Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Пермский национальный исследовательский политехнический университет»

Общество с ограниченной ответственностью
«Пермский центр поддержки принятия решений»

УДК 004.67+519.816

На правах рукописи

КОЖЕМЯКИН ЛЕОНИД ВАЛЕРЬЕВИЧ

МЕТОДЫ И АЛГОРИТМЫ ПРЕДОБРАБОТКИ ОГРАНИЧЕННЫХ ДАННЫХ И ОЦЕНКИ КАЧЕСТВА НЕПОЛНОСВЯЗНЫХ НЕЙРОННЫХ СЕТЕЙ ПРИ ИХ ПОСТРОЕНИИ НА ОСНОВЕ КОРНЕЙ ПРИНЯТИЯ РЕШЕНИЙ

2.3.1. Системный анализ, управление и обработка информации, статистика

ДИССЕРТАЦИЯ

на соискание ученой степени

кандидата технических наук

Научный руководитель:

кандидат экономических наук,

доцент Алексеев А.О.

Пермь 2024

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ 5

ГЛАВА 1. ОБЗОР МЕТОДОВ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ИНФОРМАЦИИ В УСЛОВИЯХ ОГРАНИЧЕННЫХ НАБОРОВ ДАННЫХ 12

1.1. Бутстреп-методы 13

1.2. Методы генерации синтетических данных 16

1.3. Методы фильтрации данных 19

1.4. Методы системно-когнитивного анализа 23

1.5. Методы построения деревьев решений 27

1.6. Методы агрегирования и механизмы комплексного оценивания 31

1.7. Выводы по главе 1 37

ГЛАВА 2. ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ 39

2.1. Базовое представление искусственных нейронных сетей 40

2.2. Искусственные нейронные сети с каскадной и свёрточной структурой 47

2.3. Искусственные нейронные сети, основанные на корнях принятия решений 52

2.4. Выводы по главе 2 57

ГЛАВА 3. ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ ОБРАБОТКИ ДАННЫХ 59

3.1. Программный инструментарий автоматизированного системно-когнитивного анализа – интеллектуальная система «Эйдос-Xpro» 60

3.2. Программная реализация деревьев решений для классификации 64

3.3. Программный инструментарий построения нейронных сетей – «Нейросимулятор 5.0» 67

3.4. Программный модуль проектирования и обучения нейронных сетей, основанных на корнях принятия решений 70

3.5. Выводы по главе 3 76

ГЛАВА 4. ОЦЕНКА КАЧЕСТВА НЕЙРОННЫХ СЕТЕЙ,
ОСНОВАННЫХ НА КОРНЯХ ПРИНЯТИЯ РЕШЕНИЙ, НА ПРИМЕРЕ
ДИАГНОСТИКИ БОЛЕЗНИ АЛЬЦГЕЙМЕРА 77
4.1. Предварительная обработка данных с помощью АСК-анализа 77
4.2. Предварительная обработка данных с помощью деревьев решений 85
4.3. Нейросетевое моделирование в нейропакете
«Нейросимулятор 5.0» 91
4.3.1. Результаты обучения нейронных сетей
без предварительной обработки исходных данных 92
4.3.2. Результаты обучения нейронных сетей
с предварительной обработкой исходных данных 98
4.4. Нейросетевое моделирование с помощью
библиотек Keras, TensorFlow 103
4.4.1. Результаты обучения нейронных сетей
без предварительной обработки исходных данных 103
4.4.2. Результаты обучения нейронных сетей
с предварительной обработкой исходных данных 104
4.5. Тестирование программного модуля проектирования
и обучения неполносвязных нейронных сетей
класса Decision’s-Root-based Neural Network 105
4.6. Выводы по главе 4 113
ЗАКЛЮЧЕНИЕ 114
СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ 116
СПИСОК ЛИТЕРАТУРЫ 118
ПРИЛОЖЕНИЕ 1 141
ПРИЛОЖЕНИЕ 2 143
ПРИЛОЖЕНИЕ 3 148
ПРИЛОЖЕНИЕ 4 151

ПРИЛОЖЕНИЕ 5 154

ПРИЛОЖЕНИЕ 6 156

ПРИЛОЖЕНИЕ 7 157

ПРИЛОЖЕНИЕ 8 158

ПРИЛОЖЕНИЕ 9 159

ПРИЛОЖЕНИЕ 10 160

ВВЕДЕНИЕ

Актуальность исследования. Известно, что искусственные нейронные сети являются универсальным аппроксиматором [1; 2]. Однако качество аппроксимации напрямую связано как с объёмом обучающих данных, так и с выбором необходимой архитектуры нейронной сети и количеством нейронов на слоях этой сети [3–6].

На развитие искусственных нейронных сетей существенное значение оказало доказательство А.Н. Колмогоровым [7; 8] и В.И. Арнольдом [9; 10] возможности представления непрерывных функций многих переменных в виде суперпозиции функций меньшего числа переменных. Представление дискретных многозначных функций многих переменных в виде суперпозиции функций двух переменных получило применение в проектировании систем и механизмов комплексного оценивания (МКО). В 2020 г. было показано, что любой МКО может быть представлен в форме неполносвязной искусственной нейронной сети с каскадной структурой [11; 12]. В том же году независимо были получены несколько методов решения задачи синтеза МКО на основе обучающих примеров. Эти обстоятельства позволяют комплексировать технологии обучения нейронных сетей и МКО. В 2021 г. было предложено новое название для МКО – корни принятия решений [13]. Поэтому искусственные нейронные сети, построенные на основе МКО, в данной работе называются «нейронные сети, основанные на корнях принятия решений».

До настоящего исследования неизвестным являлся ответ на вопрос, будет ли достигнуто допустимое с прикладной точки зрения качество нейронных сетей, основанных на корнях принятия решений, если область определения исходных данных разделить на интервалы, каждому из которых установить дискретное значение, и по полученному набору дискретных данных идентифицировать корень принятия решений, последнее представить в виде нейронной сети и затем обучить её на исходных данных в непрерывном виде? При этом существенным ограничением для идентификации корней принятия решений является число анализируемых переменных, поскольку их количество экспоненциально

определяет вариативность структур корней принятия решений. Таким образом, сохраняет актуальность выявление наиболее значимых переменных на этапе предварительной обработки данных.

Степень разработанности темы исследования. Распространённым подходом при работе с ограниченными данными является расширение имеющейся выборки. Развитием бутстреп-методов занимались M.H. Quenouille [14], J.W. Tukey [15], B. Efron [16], А.И. Орлов [17; 18], A.R. Henderson [19], С.А. Анатольев [20], А.В. Антонов [21; 22], С.В. Соколов, В.А. Чепурко [21], К.Н. Маловик, И.А. Чумаков [22], В.К. Шитиков [23; 24], Г.С. Розенберг [24], E.S. Banjanovic, J.W. Osborne [25], А.А. Белых, А.В. Аталян [26] и др. Альтернативным способом расширения имеющейся выборки является генерация синтетических данных, которой занимались D.B. Rubin [27; 28], H. Kvarnström [29], L. Moniz [30], C.M. Bowen [31], A. Sistrunk [32], J. Awan, Z. Cai [33], V. Firoiu [34], С.И. Николенко [35], J. Jordon, L. Szpruch, F. Houssiau, M. Bottarelli, G. Cherubin, C. Maple, S.N. Cohen, A. Weller [36], А.Н. Рабчевский [37; 38] и др.

Другим подходом к анализу ограниченных данных является исключение переменных, которые могут быть признаны источниками шума. Для оценки и борьбы с шумом разработаны различные методы фильтрации, которыми занимались R.E. Kalman, R.S. Bucy [39], G.E.P. Box, G.M. Jenkins [40], D. Wilson [41], R.J. Hickey [42], A. Felinger [43], C.E. Brodley, M.A. Friedl [44], D. Gamberger, N. Lavrac, C. Groselj [45], H. Xiong [46], T.M. Khoshgoftaar, P. Rebours [47], X. Wu, X. Zhu [48; 50], A. Malossini [49], N.E. Huang [51], Ю.А. Каламбет, Ю.П. Козьмин [52; 53], С.А. Мальцев [52], А.С. Самохин [53], B. Frenay, M. Verleysen [54], J.A. Saez [55], М.В. Куликова, Г.Ю. Куликов [56] и др.

Для обработки данных без какой-либо фильтрации Е.В. Луценко предложена методология автоматизированного системно-когнитивного анализа (АСК-анализа) [57–61]. Эффективность методов АСК-анализа была многократно показана в работах В.С. Симанкова, В.Н. Лаптева [62], В.И. Лойко [63; 64], О.А. Макаревича [63], Т.П. Барановской [64], В.Е. Коржакова [65], Н.А. Чередниченко, А.П. Трунева [66], в том числе при работе с ограниченными данными.

Алгоритмы поиска иерархического представления набора решающих правил в виде деревьев решений, в том числе градиентный бустинг, относятся к методам машинного обучения, используемым для решения задач классификации и регрессии, которыми занимались C.I. Hovland [67], E.B. Hunt, J.Marin, P.J. Stone [68], L. Breiman, J. Friedman, R.A. Olshen, C.J. Stone [69], J.R. Quinlan [70; 71], S.P. Curram, J. Mingers [72], L.Rokach, O.Z. Maimon [73], А.Н. Сулейманова [74], В.Н. Гридин [75], И.Л. Кафтанников, А.В. Парасич [76], Г.С. Степович-Цветкова [77] и др.

Разработкой и применением нейронных сетей занимались А.И. Галушкин [78; 79], G.V. Cybenko [1], А.Н. Горбань [2; 80], Л.Н. Ясницкий [81–84], Ф.М. Черепанов [84–86], В.Е. Лялин [87], В.А. Лоренц [6], Ю.И. Еременко [88], А.И. Глущенко [88; 89], В.А. Головко, В.В. Краснопрошин [90], В.Ю. Столбов [91], А.В. Клюев [91; 92], В.И. Горбаченко [93; 94], А.В. Макаренко [95] и др. Каскадным нейронным сетям, ставшим основой для развития нейронных сетей глубокого обучения, посвящены работы Y. LeCun [96–98], Y. Bengio [97; 98], Э.Д. Аведьян, Г.В. Баркан, И.К. Левин [99] и др. Свёрточными нейронными сетями занимались D. Ciresan, U. Meier, L. Gambardella, J. Schmidhuber [100], A. Krizhevsky [101] под руководством G. Hinton [102], S. Ioffe [103], H.A.M. Williams, M.H. Jones, M. Nejati, M.J. Seabright, B.A. MacDonald [104], T. Wiatowski, H.A. Bolcskei [105] и др. Гибридные и нейро-нечёткие модели исследовали C.T. Lin [106], С.А. Филист [107], В.А. Алексеев [108; 109], П.В. Сараев [108–110] и др.

Проектированием систем и механизмов комплексного оценивания (корней принятия решений) занимались А.М. Черкашин [111; 112], В.А. Глотов [113–115], В.В. Павельев [112–117], В.Б. Гусев [118–120], В.Н. Бурков [121–126], Н.А. Коргин [13; 125–128; 136; 137], В.А. Сергеев [13; 126; 127], Д.А. Новиков [124; 128–130], С.А. Баркалов [131], В.А. Харитонов [132–134], Р.Ф. Шайдулин [133], А.А. Белых [132–135], А.О. Алексеев [11; 12; 136–146; 150–155] и др.

Несмотря на многообразие методов анализа данных, а также способов их применения в случае ограниченных данных, проверке подлежит качество обучения искусственных нейронных сетей, основанных на корнях принятия решений. Это определило объект, предмет, цель и задачи настоящего исследования.

Объектом исследования являются интеллектуальные системы обработки информации.

Предметом исследования являются методы и алгоритмы предварительной обработки информации в условиях ограниченного набора данных.

Цель диссертационного исследования – оценка качества неполносвязных нейронных сетей, основанных на корнях принятия решений, с помощью методов и алгоритмов предобработки данных.

Основные задачи исследования:

1. Выполнить анализ математических и инструментальных методов обработки информации и обосновать выбор методов, применимых для предварительной обработки в условиях ограниченных данных.

2. Разработать алгоритмическое обеспечение информационной системы проектирования и обучения искусственных нейронных сетей, основанных на корнях принятия решений.

3. Провести многократные вычислительные эксперименты по обучению и тестированию нейронных сетей и оценить полученное качество в условиях ограниченных данных.

Положения, выносимые на защиту, обладающие **научной новизной:**

1. Предложены правила и методы обработки ограниченных данных, отличающиеся от известных тем, что области определения исходных данных разбиваются на интервалы, образуя набор дискретных данных, на основе которых идентифицируются корни принятия решений и строятся неполносвязные нейронные сети, в конце обучаемые на исходном наборе данных. Такая предобработка данных позволяет сократить поиск нужной структуры неполносвязной нейронной сети *(п. 4 паспорта специальности 2.3.1 – «Разработка методов и алгоритмов решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта»).*

2. Разработано алгоритмическое обеспечение информационной системы проектирования и обучения искусственных нейронных сетей, основанных на

корнях принятия решений, отличающееся от известных тем, что неполносвязность нейронных сетей обеспечивается во всех узлах ее структуры. Это позволяет осуществлять многократные вычислительные эксперименты по анализу устойчивости результатов обучения неполносвязных нейронных сетей, включая кросс-валидацию в условиях ограниченных данных *(п. 5 паспорта специальности 2.3.1 – «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта»).*

3. Предложен метод оценки качества обучения нейронной сети в условиях ограниченных данных, отличающийся от известных тем, что результаты многократного обучения нейронной сети без предобработки данных сравниваются с результатами, полученными на основе предобработки, путём поиска неполносвязной нейронной сети. Результаты многократных экспериментов при разных структурах нейронных сетей показали, что качество неполносвязных нейронных сетей в среднем превосходит качество полносвязных нейронных сетей, сопоставимых по числу искомых в процессе обучения параметров *(п. 3 паспорта специальности 2.3.1 – «Разработка критериев и моделей описания и оценки эффективности решения задач системного анализа, оптимизации, управления, принятия решений, обработки информации и искусственного интеллекта»).*

Теоретическая значимость. Впервые выполнена оценка качества искусственных нейронных сетей, основанных на корнях принятия решений в условиях ограниченного набора исходных данных.

Разработанные методы использовались на лекционных и практических занятиях по дисциплинам «Дискретная математика и математическая логика», «Системы искусственного интеллекта», «Искусственный интеллект и машинное обучение».

Практическая значимость. Создано программное обеспечение, позволяющее проектировать и обучать неполносвязные нейронные сети на основе корней принятия решений. Программное обеспечение используется как составной компонент виртуальной среды интеллектуального анализа данных «Data to

Decisions» (D2D.Platform), разработанное ООО «Пермский центр поддержки принятия решений» в рамках НИОКР «Разработка и тестирование прототипа многопользовательской виртуальной среды интеллектуального анализа данных» (регистрационный номер ЕГИСУ НИОКТР 122122600007-1) [147]. Также программное обеспечение использовалось при выполнении фундаментальных исследований в рамках соглашений в 2022 [148] и 2023 [149] гг. с АНО «Научно-образовательный центр мирового уровня «Рациональное недропользование» по проекту «Новые материалы и технологии для медицины». Созданы два результата интеллектуальной деятельности, которые отнесены к категории «секрет производства (ноу-хау)» и заключены лицензионные договоры на их использование с ООО «Системы диагностик».

Методология и методы исследования. Теоретическую и методологическую основу исследования составляют общая теория систем и системный анализ, теория информации, теория принятия решений. В исследовании применялись методы автоматизированного системно-когнитивного анализа, методы и алгоритмы синтеза механизмов комплексного оценивания (корней принятия решений), методы обучения искусственных нейронных сетей. Обработка данных выполнялась с помощью различных информационных систем, таких как автоматизированная когнитивная аналитическая система «Эйдос-Xpro», нейропакет «Нейросимулятор 5.0». Помимо этого, потребовалась разработка собственных программ для ЭВМ для осуществления многократных вычислительных экспериментов по обучению деревьев решений (использовалась библиотека Sklearn.tree для Python), проектированию и обучению нейронных сетей, в том числе основанных на корнях принятия решений (использовалась платформа TensorFlow и библиотека Keras).

Степень достоверности и апробация результатов. Достоверность результатов и выводов подтверждается многократными вычислительными экспериментами, выполненными с применением различных методов обработки данных, используя независимые информационные системы. Апробация полученных результатов была проведена на следующих международных и

всероссийских конференциях: XVIII Всероссийской школе-конференции молодых ученых «Управление большими системами» (г. Челябинск, 2022) [150]; 4th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA) (г. Липецк, 2022) [151]; XXI Всероссийской научной конференции «Нейрокомпьютеры и их применение» (г. Москва, 2023) [152]; XIX Всероссийской школе-конференции молодых ученых «Управление большими системами» (г. Воронеж, 2023) [153]; XXXVI Международной научной конференции «Математические методы в технике и технологиях – ММТТ-36» (г. Нижний Новгород, 2023) [154].

Публикации. По теме диссертационной работы было опубликовано 12 научных работ [150–160], в их числе 3 статьи в рецензируемых научных изданиях [155–157], в которых должны быть опубликованы основные результаты диссертаций на соискание учёной степени кандидата наук, из них 1 статья в издании, индексируемом в международных реферативных базах и системах цитирования Web of Sciences Core Collection и Scopus [155], 1 патент на изобретение [158] (Приложение 6), 3 свидетельства о государственной регистрации программы для ЭВМ [159–161] (Приложение 7–9), а также 1 статья в сборнике конференции, индексируемом в Scopus [151].

Структура и объем работы. Содержание и логика диссертационного исследования определили его структуру и последовательность изложения материала. Диссертация состоит из введения, четырех глав, заключения, списка литературы и приложений. Работа изложена на 160 страницах машинописного текста и содержит 46 рисунков, 15 таблиц и 10 приложений. В библиографическом списке представлено 187 наименований литературных источников.

ГЛАВА 1. ОБЗОР МЕТОДОВ ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ИНФОРМАЦИИ В УСЛОВИЯХ ОГРАНИЧЕННЫХ НАБОРОВ ДАННЫХ

В современном информационном обществе каждый день генерируется и накапливается большое количество данных, которые стали ценным ресурсом, используемым во многих отраслях, таких как медицина, промышленность, финансы, маркетинг, наука и т.д. Однако, чтобы извлечь пользу из этого потока информации, необходимы эффективные подходы к анализу данных, одними из которых являются методы, основанные на нейросетевых технологиях.

Текущий уровень развития нейросетевых технологий, в основе которых лежат искусственные нейронные сети (ИНС), позволяет решать сложные и слабо формализуемые задачи. Однако для достижение высокого уровня точности получаемых ИНС, несмотря на различные варианты их структур, от простого однослойного персептрона до сложной глубокой рекуррентной сети, необходим достаточно большой набор исходных данных для обучения, валидации и тестирования нейросетевых моделей. На практике не всегда удается сформировать большой объём исходных данных. Известны различные подходы, начиная от бутстрепа – многократной генерации данных на базе имеющейся выборки, до генерации абсолютно искусственных данных на основе экспертных представлений о предметной области [152].

Так, например, в работе [37] было показано, что синтетические данные могут быть эффективно использованы для обучения искусственных нейронных сетей. Другими эффективными методами обработки ограниченных данных являются методы, основанные на фильтрации выборки [39–56], методы АСК-анализа [57–66], алгоритмы поиска иерархического представления набора решающих правил в виде деревьев решений [67–77], механизмы комплексного оценивания [111–146].

Однако не во всех исследуемых областях, например, синтетические данные могут быть сгенерированы или применены, вследствие чего появляется необходимость в предварительной обработке исходных наборов данных, позволяющая эффективно применять аппарат искусственных нейронных сетей, получая нейросетевые модели, обладающие высоким уровнем точности.

Все вышесказанное определило следующую структуру обзора методов предварительной обработки информации в условиях ограниченных данных.

1.1. Бутстреп-методы

Впервые идея «размножения выборок» была предложена M.H. Quenouille в 1949 г. [14], а ее последующее дополнение показано J.W. Tukey в 1958 г. [15], в основе которой лежит правило исключения одного наблюдения, используемое в дальнейшем B. Efron в [16] в виде «нового взгляда» на «размножение выборок», в результате чего термин «бутстреп» получил свою наибольшую известность в 1979 г. Бутстреп-метод, или бутстреппинг (bootstrap method), – это статистический метод, используемый для оценки характеристик выборки на основе повторной выборки. Он часто применяется для оценки дисперсии, доверительных интервалов, проверки гипотез и построения моделей в ситуациях, когда традиционные методы могут быть неприменимы или менее эффективны.

Применение бутстреп-методов направлено на выявление степени неопределённости разброса параметров, не имея представления о характере распределения этих параметров в исходном наборе данных. Поскольку ключевая идея бутстреп-методов состоит в дополнении или замене теоретических исследований на эксперименты, то анализ исходных данных с применением параметрического распределения заменяется большим количеством искусственно созданных «схожими» выборками, что позволяет сразу перейти к решению задач с помощью вычислительных методов, тем самым, не прибегая к оцениванию параметров и характеристик и проверки гипотез. Далее происходит вычисление необходимых статистик в рамках поставленной задачи для каждой из созданной «схожей» выборки для анализа ее распределения. Так, поиск теоретического распределения статистики заменяем на искусственно созданные «схожие» выборки, определяем их необходимые статистики и проводим анализ эмпирического распределения или выборочной функции распределения (например, для поиска доверительных интервалов будем пользоваться квантилями этого распределения).

Одной из целей бутстреп-метода является оценка неопределённости статистических оценок и улучшение качества их интерпретации. Для этого на каждой выборке вычисляется и сохраняется статистическая оценка интересующего нас параметра. Затем на основе полученного распределения оценок строится доверительный интервал, который показывает вероятный диапазон значений параметра [162].

Алгоритм работы метода начинается с получения выборки данных из исходной генеральной совокупности, например из выборки размера N. Далее существуют два способа использования бутстреп-метода: с возвратом и без возврата.

Первый способ позволяет оценить стабильность и доверительные интервалы для оценок параметров путём многократной генерации случайной выборки из исходных данных с возвращением. В данном способе случайным образом выбирается N элементов из исходного набора данных, которые добавляются в новую выборку с возвратом (то есть один и тот же элемент может быть выбран несколько раз). Второй же используется для анализа данных путём создания нескольких независимых подвыборок из исходного набора данных без повторения элементов. В этом способе выборка формируется аналогично первому способу, но без возврата, то есть каждый элемент может быть выбран только один раз. Затем для новой выборки вычисляется статистика (например, среднее или медиана), и эти действия повторяются от сотни до десятка тысяч раз для получения множества новых выборок и вычисления статистики для каждой из них. После чего, используя полученное множество статистик, создаётся доверительный интервал, и полученные результаты интерпретируются, основываясь на полученных доверительных интервалах [20; 26]. Общий алгоритм бутстрепа представлен на Рисунке 1.1 [25].

Преимущества и недостатки применения бутстрепа как статистического метода генерации данных обсуждаются в работе [17]. Преимущества бутстрепа заключаются в его простоте, универсальности (бутстреп применим к различным типам данных и распределений), нетребовательности к конкретным предположениям о распределении данных, способности обеспечивать более

надежные оценки параметров на основе имеющихся данных и возможности использования в различных статистических методах и моделях. Стоит отметить, что бутстреп-оценки параметров являются более устойчивыми к выбросам асимметрии, чем другие методы. Также бутстреп позволяет учесть неопределенность, связанную с оценками, и оценить достоверность полученных результатов. Ключевыми недостатками бутстреп-методов является высокая зависимость получаемых результатов от качества и репрезентативности исходных данных, а также вычисленная затратность при больших объёмах выборки или при использовании сложных моделей, таких как временные ряды.

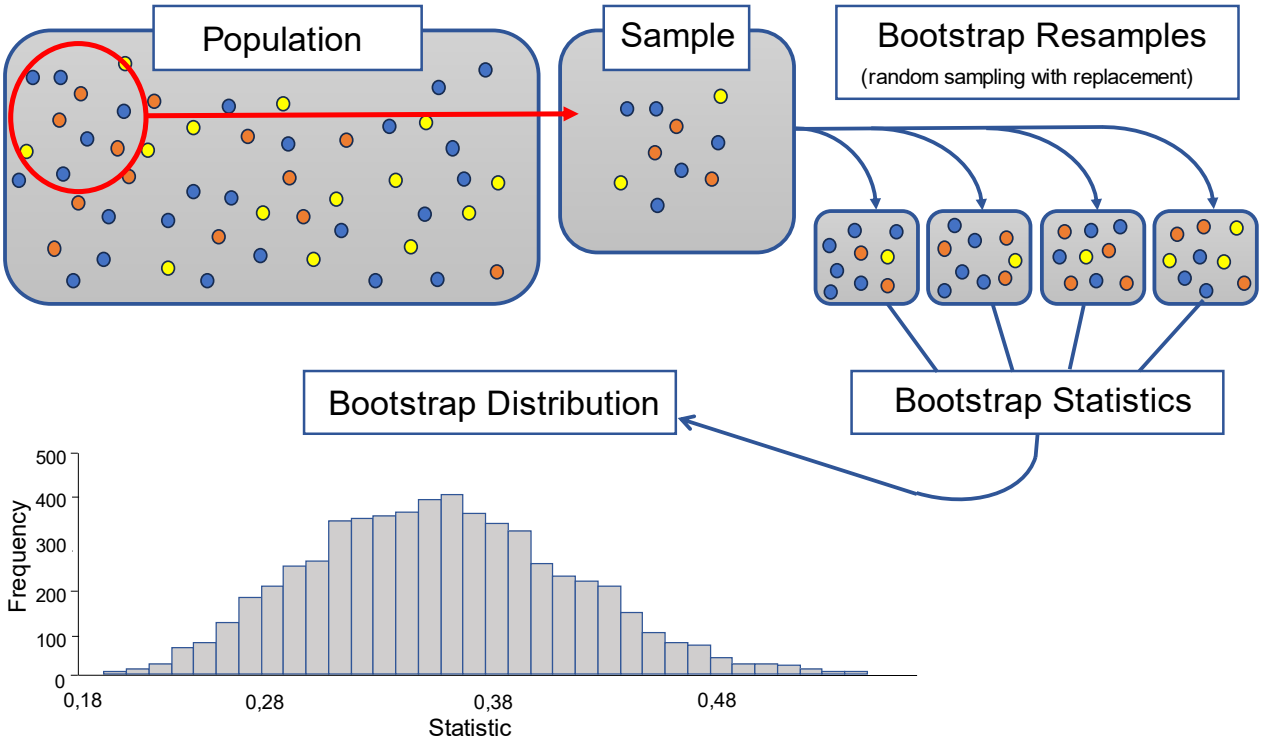


Рисунок 1.1 – Схематичное описание процесса бутстреп-метода [25, с. 2]

Бутстреп-методы широко применяются: для оценки среднего значения выборки и соответствующего доверительного интервала; для определения стандартной ошибки, что позволяет получить более точную оценку, особенно когда размер выборки мал; для оценки параметров модели и получения соответствующего доверительного интервала; для оценки выборочных корреляций между различными переменными, что позволяет определить, насколько сильна связь между двумя переменными, и выявить, насколько можно на них полагаться в анализе данных; для

оценки необходимого размера выборки для получения точных статистических выводов. На практике бутстреп-методы эффективно применяются в медицине [19], в сфере ядерной промышленности [21; 22], экологии [23; 24] и в других областях.

1.2. Методы генерации синтетических данных

С одной стороны, ограниченный объём исходных данных существенно снижает вероятность получения точных моделей в процессе анализа исследуемой области, с другой стороны, большое количество примеров из реального мира показывает, что высокоразмерные наборы данных по своей природе уязвимы при атаках на конфиденциальность и что существующие методы анонимизации не обеспечивают адекватную защиту. Это ограничивает возможности по совместному использованию таких наборов данных, что создает «узкое место» в разработке и внедрении методов машинного обучения и методов науки о данных. В таких случаях генерация синтетических данных позволяет обезличить информацию, сохраняя при этом структуру и характеристики исходных данных.

Однако в некоторых областях исследования, таких как медицина, промышленность или прогнозирование катастроф, данные могут быть ограниченными. Генерация синтетических данных позволяет создать дополнительные наборы данных для обучения и тестирования моделей машинного обучения и статистического анализа. Поэтому актуальность генерации синтетических данных в настоящее время активно возрастает.

Под генерацией синтетических данных понимается процесс создания новых данных на основе уже имеющихся без их явного копирования или повторения. Этот подход широко используется в области машинного обучения и анализа данных для создания дополнительных обучающих примеров, увеличения объёма данных, балансировки классов или защиты конфиденциальности. В первых работах [27; 28], посвященных разработке методов защиты данных, D.B. Rubin предложил идею использования оригинальных полностью синтетических данных вместо реальных. Такой подход позволял избежать риска нарушения конфиденциальности и приватности реальных людей, а также экономил время и

ресурсы, которые иногда требуются для сбора и очистки реальных данных. Необходимость извлекать полезную информацию из больших массивов данных привела к разработке сложных моделей, основанных на данных (машинное обучение). Для этих моделей роль данных в выборе модели более заметна, чем для более простых, моделей, созданных вручную. Это означает, что качество результатов модели напрямую зависит от качества данных, используемых для обучения этих моделей [36]. С тех пор синтетические данные стали широко применяться в различных областях для оценки поведения различных систем или алгоритмов, в частности, при использовании независимых, антикоррелированных, коррелированных и случайных данных.

Несмотря на популярность генерации синтетических данных, не существует общепринятого определения синтетических данных. С одной стороны, синтетические данные как искусственно созданные наборы данных имитируют структуру и распределение реальных данных для использования в обучении моделей машинного обучения. С другой стороны, синтетические данные могут включать в себя любые данные, созданные искусственно, независимо от цели их создания. Однако, в общем смысле, синтетические данные создаются искусственно с помощью специально построенных алгоритмов случайным образом или на основе заранее заданных правил. Важно учитывать контекст, в котором используются синтетические данные, чтобы определить их точное значение и применимость в конкретной области. Несмотря на отсутствие универсального определения синтетических данных, частота их использования продолжает расти в контексте исследований и разработок в области данных и искусственного интеллекта. Контролируя процесс генерации данных, конечный пользователь может регулировать объём приватной (засекреченной) информации, раскрываемой синтетическими данными, и контролировать их сходство с реальными данными.

Цель синтетических данных заключается в том, чтобы создать новый набор данных со свойствами распределения исходного набора данных, при этом сохраняя конфиденциальность участвующих в исследовании лиц. Существует два основных типа синтетических данных: частично синтетические данные,

позволяющие обнародовать определенные статистические данные или атрибуты без сохранения конфиденциальности, защищающие при этом другие аспекты данных, и полностью синтетические данные, где все статистические данные и атрибуты данных защищены [33].

Чтобы синтетические данные имели практический смысл, они должны быть похожи на оригинальные и при этом отличаться от них. Допустимое сходство или требуемое несходство зависит от задачи и представляет собой один из трех атрибутов, которые являются основополагающими при создании синтетических данных: полезность, деликатность и конфиденциальность. Полезность заключается в том, чтобы синтетические данные могли быть достаточно похожи на оригинальные, чтобы быть полезными для проведения анализа и исследования, но при этом необходимо, чтобы они не содержали конкретных личностных данных, чтобы обеспечить конфиденциальность. «Деликатность» синтетических данных выражается в ненарушении «чувствительных» аспектов оригинальных данных, например, чтобы не раскрывать коммерческие секреты или личную информацию, содержащуюся в исходном наборе данных [36].

Принцип, основные методы и алгоритмы генерации синтетических данных представлены на Рисунке 1.2.

Синтетические данные показали свою эффективность в сфере безопасности [29], в медицине при обеспечении конфиденциальности данных о пациентах [30; 31], в социальных науках [32], в математике [33]. Отдельно хочется отметить рост применения синтетических данных в сфере компьютерного зрения и распознавания образов с применением интеллектуальных нейросетевых технологий, что так же, как и более широкий обзор областей применения синтетических данных, представлено в работе А.Н. Рабчевского [38].

Ключевое преимущество синтетических данных заключается в своей способности устранять пробелы, особенно когда реальные данные недоступны или должны быть закрыты из-за конфиденциальности или рисков, связанных с конфиденциальностью или соблюдением нормативных требований. Однако качество генерируемых синтетических данных варьируется в зависимости от

качества исходных данных, что делает генерацию синтетических данных сложным и затратным процессом, особенно если требуется большое количество данных, точно отражающих реальные условия.

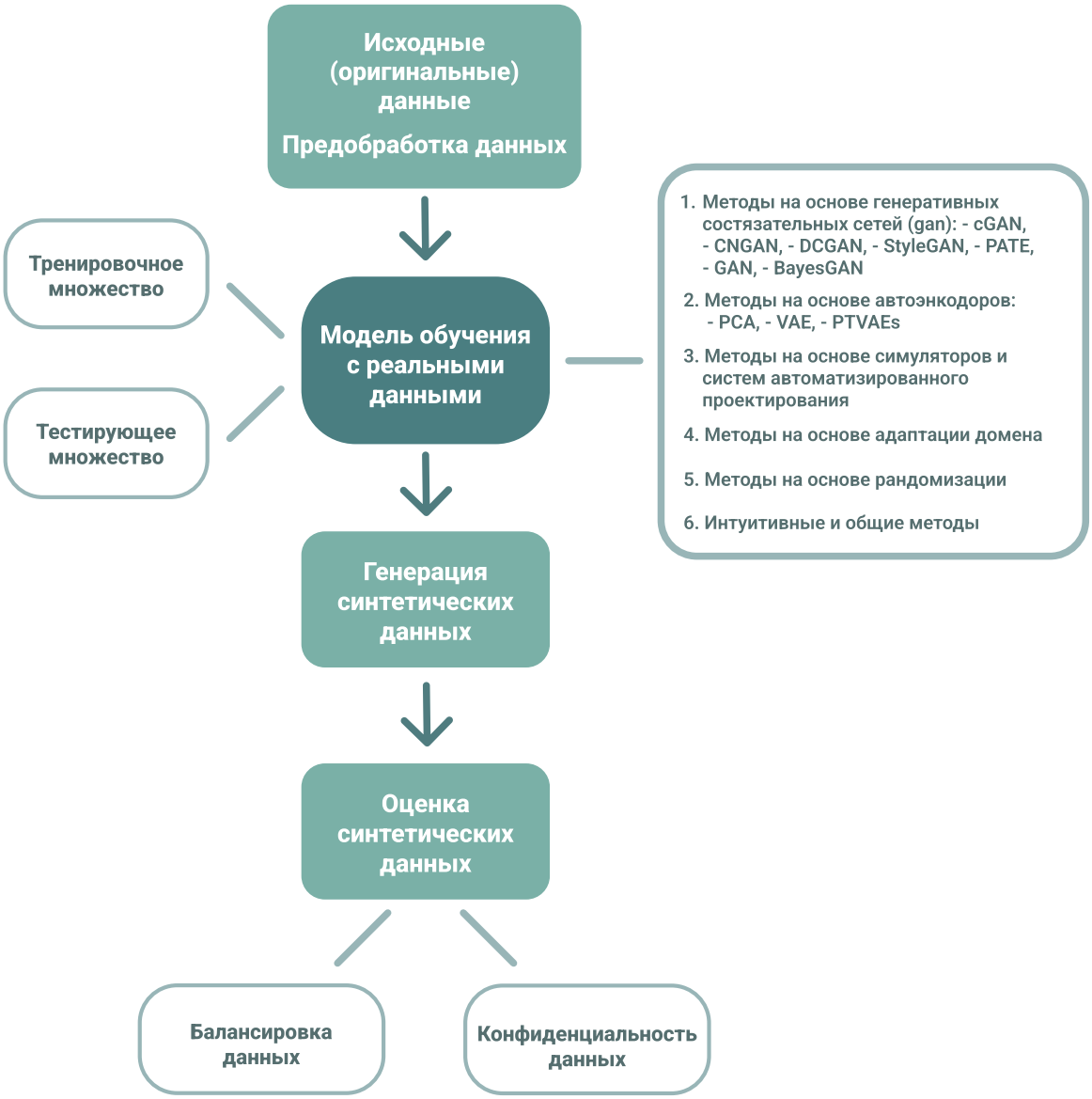


Рисунок 1.2 – Общий принцип генерации синтетических данных (составлено на основе работы [38])

1.3. Методы фильтрации данных

В условиях ограниченных данных каждый пример исходной выборки влияет на итоговую модель при анализе этих данных. Однако не все примеры в исходных данных исследуемой области релевантны. Процесс маркировки каждого обучающего примера обычно выполняется экспертами или автоматическими

системами. При неправильной маркировке появляются так называемые «шумы» в данных, которые существенно снижают качество получаемых результатов.

Выделяют два основных типа шума данных в машинном обучении и анализе данных: классовый шум (Class noise) и атрибутный шум (Attribute noise). Классовый шум происходит, когда класс объекта неправильно определен. Это может быть результатом неточных или неполных данных, ошибок в процессе формирования данных или просто случайный шум. Например, в классификации изображений классовый шум может возникнуть, если изображение кота случайно отмечено как изображение собаки. Атрибутный шум происходит, когда атрибуты объекта (то есть их характеристики или свойства) неправильно измерены или ошибочно записаны. Это может быть результатом неточных инструментов измерения, несовершенства методов сбора данных или просто случайных ошибок. Например, в анализе климата атрибутный шум может возникнуть, если температура измерена неправильно или записана с ошибкой. Шум в классе может повлиять не только на производительность системы, но и на снижение точности классификации [48].

В разных предметных областях шум в данных может появляться разным способом, например, при классификации медицинских данных классовый шум может исходить из следующих источников [42]:

1. Ошибки в процессе маркировки, которые более вероятны при работе со сложными данными, могут возникать из-за человеческого фактора, в результате усталости, рутиной работы, быстрого изучения каждого случая или нехватки времени. Субъективность также может создавать шум в классе; например, когда существует вариативность в маркировке несколькими экспертами [49].

2. Машинные ошибки. Когда машина отвечает за автоматическое нанесение меток, возникновение ошибок в конструкции или кратковременных ошибок может привести к появлению ошибочных меток.

3. Ошибки оцифровки и архивирования. При создании цифровой записи изученных дел можно по простой ошибке неправильно ввести класс. Аналогичная ситуация возникает при использовании исторических записей.

Одним из наиболее часто используемых методов уменьшения шума данных является фильтрация, которая удаляет зашумленные примеры из обучающих данных, улучшая их качество [45].

Под фильтрацией данных понимается процесс обработки и анализа информации с целью исключения из нее ненужных или нежелательных элементов, оставляя только те данные, которые соответствуют определенным критериям или условиям. Этот процесс позволяет улучшить качество данных, упростить их анализ и использование, а также повысить эффективность принятия решений на основе этих данных.

Обзор литературы показал, что в основном схемы фильтрации шума делятся на различные категории в соответствии с их характеристиками [41; 44; 45]. Наиболее подробный обзор шумовых фильтров можно найти в работе [54]. Отфильтрованные данные можно использовать в качестве входных данных для различных классификаторов [47], поэтому вычислительное время, необходимое для подготовки данных, требуется только один раз. Устранение ошибочно помеченных примеров дает некоторые преимущества в производительности классификатора, в отличие от удаления примеров с атрибутивным шумом [45; 48].

На Рисунке 1.3 показан пример преимущества обучения классификаторов в присутствии шума в классе при применении методов фильтрации.

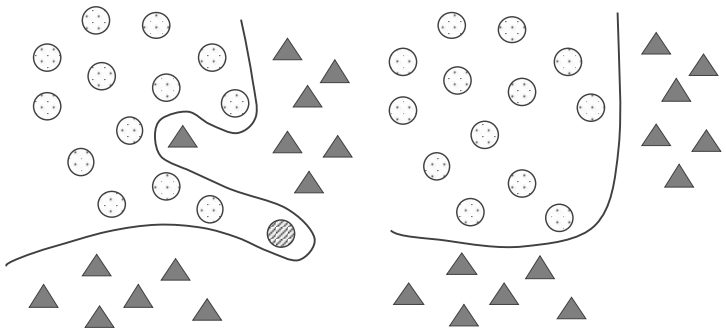


Рисунок 1.3 – Набор данных с двумя примерами с классовым шумом: *а* – примерная граница принятия решения при рассмотрении зашумленных примеров как правильных, *б* – граница принятия решения после удаления зашумленных примеров с помощью фильтрации [55]

Из Рисунка 1.3 видно, что граница принятия решения, полученная после фильтрации, является менее сложной и приводит к лучшему обобщению этих двух классов.

Основные методы фильтрации данных:

1. Сглаживание данных: этот метод используется для удаления высокочастотных шумов и колебаний в данных. Примеры методов сглаживания включают скользящее среднее, экспоненциальное сглаживание и методы регрессии. Применяется в финансовом анализе для удаления шумов на фондовом рынке и выделения основных трендов. Также используется в метеорологии для анализа временных рядов погодных данных.

2. Детектирование выбросов: данный метод позволяет выявить и удалить аномалии или выбросы в данных, которые могут искажать результаты анализа. Примеры методов детектирования выбросов включают z-оценку и методы машинного обучения. Применяется в медицинской диагностике для выявления аномалий в биомедицинских данных, таких как ЭКГ-сигналы или данные о здоровье пациентов.

3. Фильтры нижних и верхних частот: фильтры нижних частот пропускают низкочастотные компоненты сигнала, а фильтры верхних частот – высокочастотные компоненты. Эти фильтры используются для выделения определенных частотных составляющих сигнала, например, при аудиообработке или обработке речи.

Несмотря на успешное применение методов фильтрации, стоит отметить, что при фильтрации могут быть удалены примеры, содержащие ценную информацию, поэтому она не всегда обеспечивает повышение качества исходного набора данных. В исследовании [55] показана эффективность методов фильтрации в контексте классификации медицинских данных и тестируются различные шумовые фильтры, так как качество удаления шумов зависит от конкретных характеристик обрабатываемых данных.

Данные выводы также подтверждаются в работе [50], в которой авторы объясняют, что успех фильтрации данных зависит от нескольких обстоятельств,

таких как вид и природа ошибок, количество удаляемого шума или возможности классификатора справиться с потерей полезной информации, связанной с фильтрацией. Таким образом, эффективность фильтров шума зависит от помехоустойчивости и возможностей обобщения используемого классификатора.

1.4. Методы системно-когнитивного анализа

Наибольший интерес представляют системы искусственного интеллекта и машинного обучения, которые обладают универсальностью, не зависят от конкретной области исследования и могут выступать в роли системы поддержки принятия решений. Одной из таких является система, основанная на методах автоматизированного системно-когнитивного анализа (АСК-анализа), которые реализованы в отечественной автоматизированной когнитивной аналитической системе «Эйдос-Xpro» [61], разработанная профессором кафедры компьютерных технологий и систем Кубанского государственного аграрного университета (КУБГАУ) Е.В. Луценко, позволяющая последовательно обработать данные, информацию и знания любой предметной области, формализовать предметную область, провести синтез и верификацию полученных моделей.

Предварительно, система анализирует исходные данные, создавая одну или несколько классификационных и описательных шкал с градациями, количество которых зависит от выбранного метода определения размера интервала и количества наблюдений. Это может быть равномерное распределение интервалов с различным количеством наблюдений или неравномерное распределение с одинаковым количеством наблюдений в каждом интервале. Далее осуществляется синтез и верификация трех моделей статистической базы (*ABS*, *PRC1*, *PRC2*) и семь моделей системно-когнитивной базы знаний *(INF1-INF7)* (Рисунок 1.4). По умолчанию система оставляет 10 % данных от исходной выборки, удаляя наименее достоверные распознавания, но в ограниченных случаях исходного набора данных рекомендуется оставлять 90–100 % (Рисунок 1.4 – выделено красным цветом).

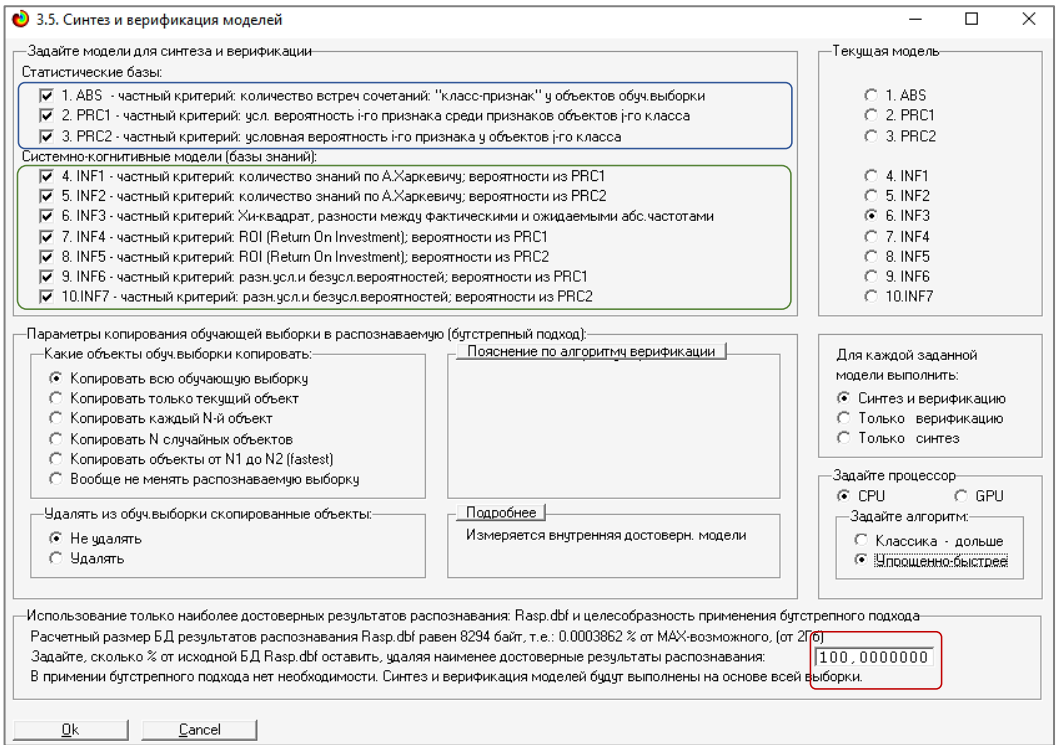


Рисунок 1.4 – Экранная форма режима «Синтез и верификация моделей» в интеллектуальной системе «Эйдос-Xpro» [61, режим 3.5]

Оценка значимости заданных статистических моделей и моделей знаний проводится путём определения количества истинных положительных (True Positive) и истинно отрицательных (True Negative), а также ложных положительных и отрицательных решений по F-мере Ван Ризбергена и L1, L2-мерам профессора Е.В. Луценко (ошибки 1-го и 2-го рода) [59, с. 25]. В Таблице 1.1. представлены модели и их частные критерии статистической базы и системно-когнитивной базы знаний, поддерживающиеся аналитической системой «Эйдос-Xpro».

Для определения способности моделей *(INF1–INF7)* к многопараметрической типизации как к свойству создавать обобщенные образы в будущих состояниях объекта управления в интеллектуальной системе «Эйдос-Xpro» используется аддитивный интегральный критерий «Сумма знаний» (I_j, далее по тексту – IC_SK) и аддитивный интегральный критерий «Семантический резонанс знаний» (I_j, далее по тексту – IC_SRK), вычисляемые по следующим формулам (1), (2) соответственно [60, с. 85–87]:

Таблица 1.1 – Аналитические формы конкретных критериев знаний в интеллектуальной аналитической системе «Эйдос-Xpro» [59, с. 10–11]

№ п/п	Модель знаний и ее частный критерий	Выражение для частного критерия знаний через *относительные* частоты	Выражение для частного критерия знаний через *абсолютные* частоты
1	*ABS,* матрица абсолютных частот	–	N_{ij}
2	*PRC1*, матрица условных и безусловных процентных распределений, в качестве $N_{\sum j}$ используется суммарное количество признаков по классу	–	$P_{ij}=\frac{N_{ij}}{N_{\sum j}}$
3	*PRC2*, матрица условных и безусловных процентных распределений, в качестве $N_{\sum j}$ используется суммарное количество объектов обучающей выборки по классу	–	$P_{ij}=\frac{N_{ij}}{N_{\sum j}}$
4	*Модель INF1*, частный критерий: количество знаний по А. Харкевичу; вероятности из PRC1, где N_j – общее количество признаков по *j*-му классу. Вероятность того, что если у объекта *j*-го класса обнаружится признак, то это *i*-й признак	$I_{ij}=\Psi\cdot\log_2\frac{P_{ij}}{P_i}$	$I_{ij}=\Psi\cdot\log_2\frac{N_{ij}N}{N_iN_j}$
5	*Модель INF2*, частный критерий: количество знаний по А. Харкевичу, 2-й вариант расчета вероятностей; вероятности из PRC2, где N_j – общее количество признаков по *j*-му классу. Вероятность того, что если у объекта *j*-го класса обнаружится признак, то это *i*-й признак	$I_{ij}=\Psi\cdot\log_2\frac{P_{ij}}{P_i}$	$I_{ij}=\Psi\cdot\log_2\frac{N_{ij}N}{N_iN_j}$
6	*Модель INF3*, частный критерий: Хи-квадрат разности между фактическими и теоретически ожидаемыми абсолютными частотами	–	$I_{ij}=N_{ij}-\frac{N_iN_j}{N}$
7	*Модель INF4*, частный критерий: ROI – Return On Investment; вероятности из PRC1, 1-й вариант расчета вероятностей: N_j – суммарное количество объектов по *j*-му классу	$I_{ij}=\frac{P_{ij}}{P_i}-1=\frac{P_{ij}-P_i}{P_i}$	$I_{ij}=\frac{N_{ij}N}{N_iN_j}-1$
8	*Модель INF5*, частный критерий: ROI – Return On Investment; вероятности из PRC2, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по *j*-му классу	$I_{ij}=\frac{P_{ij}}{P_i}-1=\frac{P_{ij}-P_i}{P_i}$	$I_{ij}=\frac{N_{ij}N}{N_iN_j}-1$
9	*Модель INF6*, частный критерий: разность условной и безусловной вероятностей; вероятности из PRC1, 1-й вариант расчета вероятностей: N_j – суммарное количество объектов по *j*-му классу	$I_{ij}=P_{ij}-P_i$	$I_{ij}=\frac{N_{ij}}{N_j}-\frac{N_i}{N}$
10	*Модель INF7*, частный критерий: разность условной и безусловной вероятностей; вероятности из PRC2, 2-й вариант расчета вероятностей: N_j – суммарное количество объектов по *j*-му классу	$I_{ij}=P_{ij}-P_i$	$I_{ij}=\frac{N_{ij}}{N_j}-\frac{N_i}{N}$

$$\mathrm{IC_SK}_j = \underbrace{\left(\overrightarrow{I_{ij}}, \overrightarrow{L_i}\right)}_{\text{скалярный вид}} = \underbrace{\sum_{i=1}^{M} I_{ij} L_i}_{\text{координатная форма}}, \qquad (1.1)$$

где M – количество градаций описательных шкал (признаков); $\overrightarrow{I_{ij}} = \left\{I_{ij}\right\}$ – вектор состояния *j*-го класса; $\overrightarrow{L_{ij}} = \left\{L_i\right\}$ – вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив-локатор), т.е.:

$$\overrightarrow{L_i} = \begin{cases} 1, \text{если } i\text{-й фактор действует} \\ n, \textit{где}: n > 0, \text{если } i\text{-й фактор действует с истинностью } n \\ 0, \text{если } i\text{-й фактор не действует} \end{cases}$$

$$\mathrm{IC_SRK}_j = \frac{1}{\sigma_j \sigma_l M} \sum_{i=1}^{M} \left(I_{ij} - \overline{I_j}\right)\left(L_i - \overline{L}\right), \qquad (1.2)$$

где M – количество градаций описательных шкал (признаков); $\overline{I_j}$ – средняя информативность по вектору класса; $\overline{L}$ – среднее по вектору объекта; σ_j – среднеквадратическое отклонение частных критериев знаний вектора класса; σ_l – среднеквадратическое отклонение по вектору распознаваемого объекта; $\overrightarrow{I_{ij}} = \left\{I_{ij}\right\}$ – вектор состояния *j*-го класса; $\overrightarrow{L_{ij}} = \left\{L_i\right\}$ – вектор состояния распознаваемого объекта, включающий все виды факторов, характеризующих сам объект, управляющие воздействия и окружающую среду (массив-локатор) [59].

Методы системно-когнитивного анализа успешно применяются при анализе изображений [58], при анализе объектов информационной безопасности [65], в работах по экологии климатологии и изучению влияния космической среды на различные глобальные процессы на Земле [66], в работах по посвященных системному анализу [62], исследованию на устойчивость системно-когнитивной модели в экономике [64] и в других областях.

1.5. Методы построения деревьев решений

Большая часть аналитических методов анализа данных представляет собой известные математические алгоритмы и методы, а их новизна заключается в различных возможностях их применения при решении конкретных проблем, обусловленных появлением новых программных и технических средств. В зависимости от необходимой информации выделяют задачи интеллектуального анализа данных, такие как классификация, кластеризация, выявление ассоциаций, выявление последовательностей и прогнозирование. Деревья решений являются одним из наиболее популярных методов машинного обучения и используются в различных областях, включая классификацию, регрессию и кластеризацию.

В классическом понимании деревья решений – это иерархическое представление набора решающих правил, состоящее из узлов и листьев, в виде графической структуры, в которой каждый узел представляет собой определённый признак или атрибут, а каждое ребро является связью между этими признаками. Деревья решений являются одним из типов контролируемого алгоритма машинного обучения, используемого для решения задач классификации и регрессии.

Основная идея деревьев решений состоит в разделении данных на более конкретные категории или классы на основе признаков или атрибутов. Далее применяются тесты на этих данных, и разбиение продолжается до тех пор, пока не будет достигнуто определённое условие остановки (например, все объекты находятся в одном классе или достигнуто максимальное количество уровней дерева). В качестве основных атрибутов разбиения узлов выбираются следующие критерии (Таблица 1.2): теоретико-информационный критерий (информационная энтропия); статистический подход (коэффициент Джини); log_loss (логарифмическая потеря).

Теоретико-информационный критерий (информационная энтропия) используется для оценки эффективности разбиения данных на подмножества и измеряет степень неопределённости или разнородности данных в каждом подмножестве. Статистический подход включает выбор и применение статистических метрик и тестов для оценки качества моделей, часто основывается на

анализе результатов, полученных из экспериментов или репрезентативных наборов данных. Log_loss (логарифмическая потеря или потери перекрестной энтропии) является метрикой качества, которая часто используется в задачах классификации и измеряет расхождение между вероятностным прогнозом модели и фактическими метками путём минимизации ошибки логарифма вероятности, что означает, что прогностическая способность модели стремится к более точным результатам.

Таблица 1.2 – Критерии разбиения узлов при построении дерева решений (составлено на основе работы [164])

Атрибут	Формула
Теоретико-информационный критерий (информационная энтропия)	$Entropy = -\sum_j p_j \cdot \log_2 p_j,$ где p_j – вероятность принадлежности классу j
Статистический подход (коэффициент Джини)	$Gini = 1 - \sum_j p_j^2,$ где p_j – вероятность принадлежности классу j
Log_loss (логарифмическая потеря или потери перекрестной энтропии)	$L_{\log}(Y,P) = -\frac{1}{N}\sum_{i=0}^{N-1}\sum_{k=0}^{K-1} y_{i,k} \log p_{i,k},$ где p_j – вероятность принадлежности классу j

Теоретико-информационный критерий (информационная энтропия) и Log_loss (логарифмическая потеря или потери перекрестной энтропии) имеют общую основу в оценке моделей с использованием информационной потери, но первый используется для выбора модели, а второй для оценки точности модели. Статистический подход же включает различные методы и процедуры в статистике [163; 164].

В процессе построения дерева решений атрибут разбиения улов используется для выбора наилучшего способа разделения данных на подмножества на каждом узле дерева. Более низкое значение атрибута указывает на более чистые и однородные подмножества данных, что является важным для построения эффективного дерева решений.

Для оценки качества деревьев решений, как и в машинном обучении, наиболее популярными являются следующие показатели (формулы (1.3)–(1.6))

[163], в основе которых лежат результаты классификации обучающих и тестирующих примеров: точность измерений, F1-мера, средняя потеря Хэмминга, корреляция Мэтьюза (MCC).

Точность измерений модели по всем классам (*Accuracy*), или доля правильных ответов полученной модели:

$$Accuracy\left(y,\hat{y}\right)=\frac{1}{n_{samples}}\sum_{i=0}^{n_{samples}-1}1\left(\hat{y}_i=y_i\right), \tag{1.3}$$

где $n_{samples}$ – количество примеров в выборке, $\hat{y}$ – прогнозируемое значение.

F1-мера (*F1-score*) – среднее гармоническое *Precision* и *Recall*:

$$F1=2\cdot\frac{Precision\cdot Recall}{Precision+Recall}, \tag{1.4}$$

где *Precision* – точность модели класса *Positive*, или доля объектов, классифицированных как положительные объекты и при этом действительно являющиеся положительными; *Recall* – полнота, или доля объектов положительного класса из всех объектов положительного класса, найденная классификатором.

Средняя потеря Хэмминга или расстояние Хэмминга:

$$L_{Ham}\left(y,\hat{y}\right)=\frac{1}{n_{samples}\cdot n_{labels}}\sum_{i=0}^{n_{samples}-1}\sum_{j=0}^{n_{labes}-1}1\left(\hat{y}_{i,j}\neq y_{i,j}\right), \tag{1.5}$$

где $n_{samples}$ – количество примеров в выборке, n_{labels} – количество классов, $\hat{y}$ – прогнозируемое значение.

Корреляция Мэтьюза (MCC) – сбалансированная мера, используемая при сильном дисбалансе классов:

$$MCC=\frac{TP\cdot TN\ -\ FP\cdot FN}{\sqrt{\left(TP\ +\ FP\right)\left(TP\ +\ FN\right)\left(TN\ +\ FP\right)\left(TN\ +\ FN\right)}}, \tag{1.6}$$

где *TP* – количество верно предсказанных положительных классов *(True Positive)*; *TN* – количество верно предсказанных отрицательных классов *(True Negative)*; *FP* – количество неверно предсказанных положительных классов *(False Positive)*; *FN* – количество неверно предсказанных отрицательных классов *(False Negative)* [163].

Деревья решений обладают рядом преимуществ, таких как интерпретируемость, возможность обработки как числовых, так и категориальных признаков, а также способность обрабатывать большие объемы данных.

Однако деревья решений также имеют некоторые недостатки, такие как склонность к переобучению, особенно на данных с большим количеством признаков, а также возможность создания сложных моделей, которые может быть трудно интерпретировать. Для решения проблем, связанных с недостатками деревьев решений, были разработаны основные алгоритмы построения деревьев решений [74]:

1. CART (Classification and Regression Tree) – алгоритм построения бинарного дерева решений, который может использоваться для выполнения задач классификации и регрессии. Он представляет собой последовательный процесс разбиения выборки на подгруппы до тех пор, пока все объекты не окажутся в одной группе. При каждом разбиении выбирается тот признак, который наилучшим образом разделяет выборку на две подгруппы так, чтобы объекты внутри каждой подгруппы были наиболее похожими по значению этого признака.

2. CHAID – используется для определения взаимодействия между переменными. Полученное дерево может быть использовано для прогнозирования значений зависимой переменной для новых данных на основе значений независимых переменных.

3. C4.5 – алгоритм оптимизирует разбиение данных на наиболее информативные атрибуты. Работает по принципу рекурсивного разбиения данных на подмножества с использованием атрибутов, чтобы найти оптимальные разделения и создать дерево принятия решений. Известен своей способностью работать с категориальными, числовыми и отсутствующими значениями данных и обрабатывать большие наборы данных. Также может автоматически устранять шум и ненужные атрибуты, улучшая качество построенной модели.

4. CN2 – часто используется для задач классификации с бинарным целевым классом и вопросами отбора признаков. Алгоритм более эффективен при работе с наборами данных, содержащих большое количество несовершенных или неоднородных данных.

Эти алгоритмы позволяют автоматически генерировать деревья решений из данных, учитывая различные критерии, такие как информационный выигрыш или прирост дисперсии.

Успешное применение деревьев решений в анализе данных и задачах классификации представлено в работах [75; 77]. В других прикладных работах, например в медицине, показано применение деревьев решений как один из эффективных инструментов интеллектуального анализа и разработки прогностических моделей. Эти модели способны предсказывать вероятность возникновения различных заболеваний или оценивать риски результатов лечения. Например, они могут определять вероятность развития сердечно-сосудистых заболеваний у пациентов на основе таких факторов, как возраст, пол, семейная история заболеваний, уровень холестерина и других соответствующих параметров [165; 166].

1.6. Методы агрегирования и механизмы комплексного оценивания

Традиционно механизмы комплексного оценивания (МКО) используются для агрегирования информации об объекте управления и описания его состояния с помощью всего одного параметра – интегральной, или комплексной оценки. Такие подходы также имеют свое применение при планировании деятельности и поддержки принятия управленческих решений [142].

В середине XX в. активно стали развиваться иерархические матричные механизмы комплексного оценивания (ММКО), для создания систем поддержки принятия решения (СППР), систем контроля, систем оценки результатов деятельности, базисным решением которых является попарная свёртка критериев. Идеи МКО основываются на трудах академика А.Н. Колмогорова [7; 8] и В.И. Арнольда [9; 10], которые предложили представлять функцию многих переменных как суперпозицию функций меньшего числа переменных.

Эти методы традиционно воспринимались как экспертные системы, идея которых лежит в определении основных элементов математического объекта, его структуры дерева критериев, описательных шкал, а также в предложении МКО путём формализации экспертного мнения. Стоит отметить, что во всех работах МКО всегда базировался на экспертных суждениях. Таким образом, иерархические ММКО исторически являются одними из экспертных методов.

Однако в 2019 г. задача синтеза МКО на основе обучающих примеров впервые была представлена на XIII Всероссийском совещании по проблемам управления (ВСПУ-2019). Результаты совещания подготовлены и представлены в сборнике трудов [167] в виде полных статей Институтом проблем управления им. В.А. Трапезникова РАН.

Продолжение и развитие идея в рамках задачи МКО, представленные годом ранее, получила в 2020 г., когда двумя командами исследователей была решена поставленная задача.

Первая команда ученых Института проблем управления им. В.А. Трапезникова (лаборатория № 57) под руководством В.Н. Буркова и Н.А. Коргина поставила задачу в 2018 г. (на публичном слушании работа [125] представлена в 2019 г.), а также в дальнейшем нашла решение этой задачи, основанное на унитарном представлении механизмов комплексного оценивания [126]. Авторы предлагают подход к идентификации МКО для заданного набора обучающих примеров, базирующийся на унитарном кодировании с использованием квадратичной формы в виде суперпозиции унитарных векторов и матриц. Задача аппроксимации и задача реализации как ее частный случай могут быть сведены к задаче максимизации некоторого полинома, получаемого для заданных бинарного дерева и набора примеров с использованием унитарного кодирования [126]. В качестве дискуссионного вопроса можно обозначить сложность выбора структуры бинарных деревьев для МКО.

Вторая команда исследователей из Перми (ПНИПУ) под руководством А.О. Алексеева рассматривала алгоритм идентификации МКО, основанный на преобразовании таблицы истинности для уменьшения размерности входных данных (в работе используется термин «табличный метод») [143; 144].

Первые идеи идентификации МКО, предложенные А.О. Алексеевым, имели ограничение по структуре. Так, в работе [11] подход механизма комплексной оценки в виде суперпозиции произведений однократно закодированных критериев и матриц применим только на МКО последовательных структур. Впоследствии алгоритм был модифицирован А.О. Алексеевым и в работе [146] представлен МКО уже на непоследовательных структурах. Используя основанные механизмы в непрерывной форме для каждого набора в обучающем примере, можно определить непрерывные комплексные значения и оценить сходимость различных механизмов.

В 2020 г. в работах [11; 12] показано, что любой МКО может быть представлен в форме неполносвязной ИНС. Идентифицированные ИНС полностью воспроизводят обучающие наборы с дискретными значениями. Соответственно возник вопрос: «Будут ли найденные ИНС обладать приемлемой для прикладных задач точностью, если их обучить на исходных данных в непрерывном виде?» [154].

Исходя из работ А.О. Алексеева, его подход к построению неполносвязной ИНС, основанной на корне принятия решения, условно можно разделить на два этапа.

На первом этапе происходит представление непрерывных данных в дискретные с помощью последовательных свёрточных преобразований. В качестве примера рассмотрим объект с четырьмя критериями $n = 4, \{X_1, X_2, X_3, X_4\}$ с последовательной свёрткой первых двух и последующих двух критериев, содержащий две матрицы свёртки (M_{12}, M_{34}) и корневую матрицу (M), представленный на Рисунке 1.5.

Пусть каждый критерий X_1, X_2, X_3, X_4 описывается с помощью трехбалльных порядковых шкал (k = 3). Тогда корень принятия решений содержит три матрицы свёртки, в которых элементы этих матриц соответствуют декартову произведению дискретных значений шкал критериев [142]. Пусть элементы матриц свёртки (M_{12}, M_{34}) и корневой матрицы (M) определены следующими значениями (Рисунок 1.6).

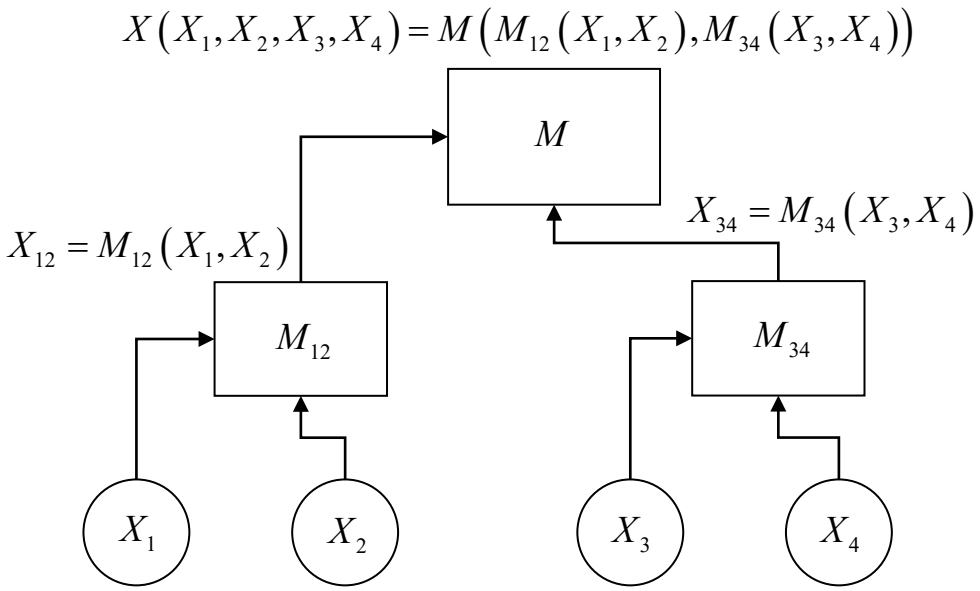


Рисунок 1.5 – Пример корня принятия решений
в общем виде для четырёх критериев

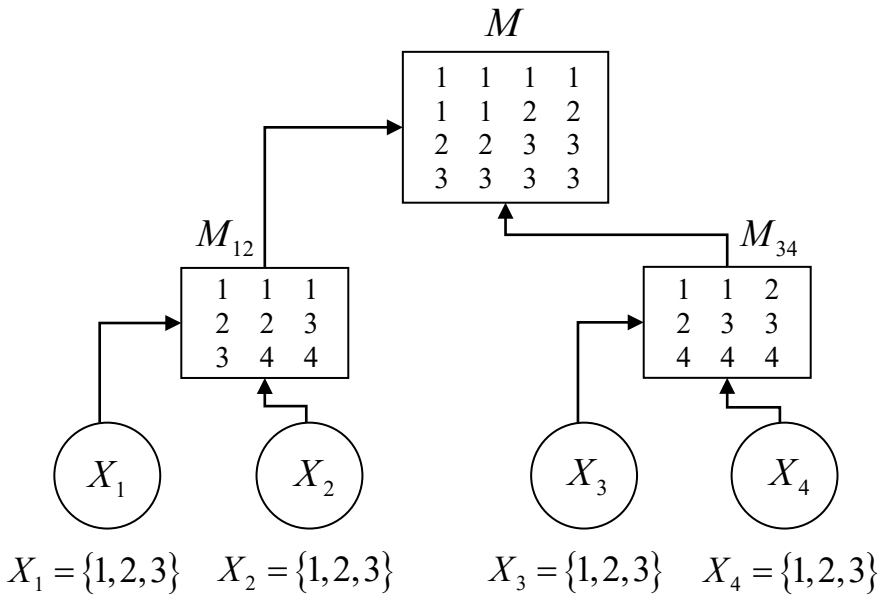


Рисунок 1.6 – Корень принятия решений для четырёх критериев,
описываемых с помощью трехбалльных порядковых шкал

На втором этапе определяется структура искомой ИНС (Рисунок 1.7) по корню принятия решений (см. Рисунок 1.6). Используя алгоритм [11, с. 402], можно заранее определить количество нейронов искомой нейронной сети на основе корня принятия решений. Так, первый скрытый слой имеет 12 нейронов, так как 4 входные переменные обозначены на множестве 3 закодированными дискретными значениями {1; 2; 3}. Для каждой из матриц свёртки из нижнего уровня корня принятия решений соответствуют 9 нейронов на втором скрытом слое, так как матрицы имеют

9 элементов. Матрица M_{12} имеет 4 значения, M_{34} имеет 4 значения, поэтому третий скрытый слой имеет 8 нейронов. Четвертый скрытый слой имеет 16 нейронов, потому что корневая матрица M имеет 16 элементов. И последний пятый скрытый слой имеет 3 нейрона, потому что решается задача определения класса (см. значения корневой матрицы M) – 1, 2 или 3. Исходя из этого, получена нейронная сеть с 5 скрытыми слоями и 69 нейронами на всех слоях, включая нейроны на 1-м входном и 1-м выходном слоях (см. Рисунок 1.7) [154].

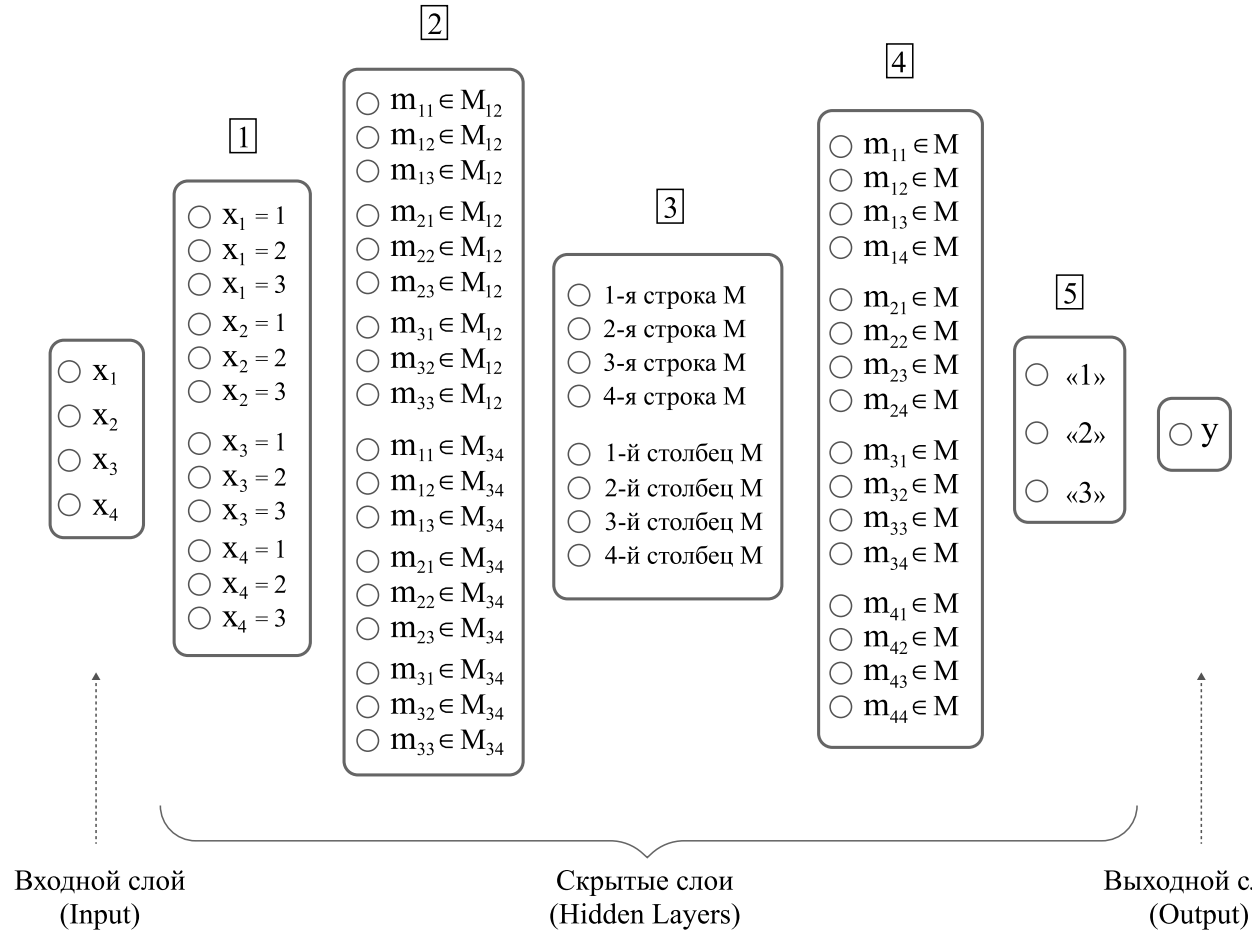


Рисунок 1.7 – Нейроны неполносвязной ИНС,
основанные на корне принятия решений (см. Рисунок 1.6.)

Таким образом, на Рисунке 1.7 представлены нейроны ИНС, основанные на корне принятия решений, где вектор нейронов на входном слое $\in R^4$, вектор нейронов на скрытом слое (1) $\in R^{12}$, вектор нейронов на скрытом слое (2) $\in R^{18}$, вектор нейронов на скрытом слое (3) $\in R^8$, вектор нейронов на скрытом слое (4) $\in R^{16}$, вектор нейронов на скрытом слое (5) $\in R^3$, вектор нейронов на выходном слое $\in R^1$.

Синаптические связи (Рисунок 1.8) определяются в полном соответствии с элементами матриц свёртки в корне решений [152]. Общее количество неизвестных параметров (синаптических коэффициентов) неполносвязной нейронной сети, предопределенной на основе корня принятия решений, равно 175. Покажем это.

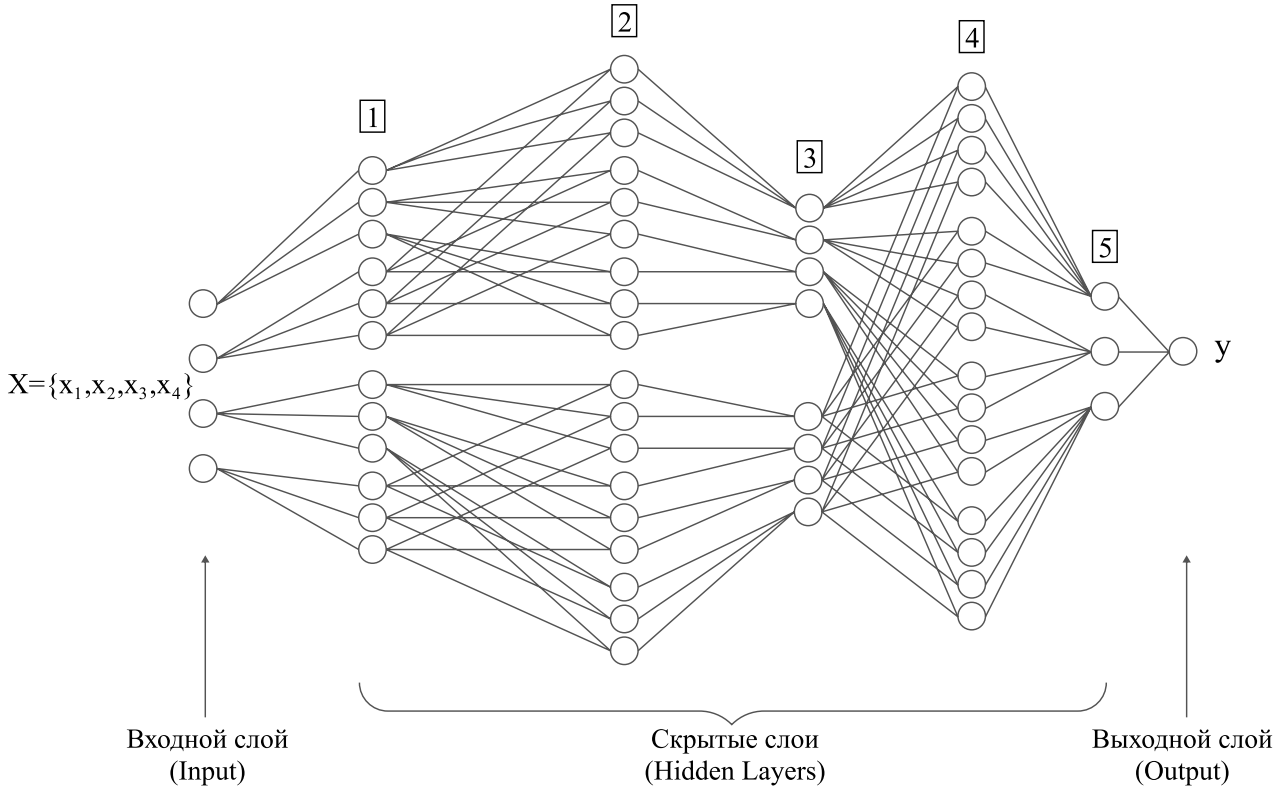


Рисунок 1.8 – Синаптические связи неполносвязной ИНС, основанной на корне принятия решений

В структуре предопределенной ИНС было 12 связей между входными сигналами и первым скрытым слоем (каждая связь генерирует два параметра математической модели, значит всего 24 параметра). Следующий слой, в котором нейроны соответствуют элементам матрицы, имеет по три параметра каждый (два входа из предыдущего слоя и одна неизвестная константа). Поскольку у нас есть две матрицы по девять элементов каждая на нижнем уровне дерева критериев, это генерирует 18 нейронов на втором слое, что означает всего 54 неизвестных параметра. Следующий слой состоит из нейронов, соответствующих матричным оценкам, а связей ровно столько, сколько нейронов было на предыдущем слое, т.е. в нашем случае 18, но к ним нужно добавить 8 параметров (определяемых числом нейронов на 3-м слое). Таким образом, всего имеем 18 и 8 параметров

(26 параметров на третьем слое). Затем имеем матрицу, содержащую 16 элементов, где каждому элементу соответствует 3 параметра (два входа и одна неизвестная константа). Это означает 48 параметров на 4-м слое. На последнем скрытом слое 3 нейрона, на которых 16 входных сигналов плюс 3 неизвестные константы (количество нейронов на текущем слое). На выходе только один сигнал. Этот сигнал связан с 3 последними нейронами. Это означает, что есть 4 параметра (3 входа плюс одна неизвестная константа). Таким образом, общее количество неизвестных параметров определяется суммой (24 + 54 + 26 + 48 + 19 + 4 = 175).

Таким образом, методы, традиционно использованные как экспертные, могут опираться на анализ данных, тем самым расширив область применения СППР, основанных на МКО.

Механизмы комплексного оценивания успешно применяются в строительном инжиниринге [168], в управлении информационными рисками [169], в управлении образовательными системами [170], в медицине в качестве системы поддержки принятия врачебных решений [150; 151; 155]. Отдельно хочется отметить, что впервые в медицинской области получен патент на изобретение «Способ диагностики болезни Альцгеймера», в основе которого лежат механизмы комплексного оценивания (корни принятия решений) [158].

В 2021 г. Н.А. Коргиным и В.А. Сергеевым было предложено новое название для МКО, традиционно применяемого для агрегирования нескольких показателей в одну комплексную оценку – корни принятия решений [13]. Далее в данном исследовании под корнями принятия решений также будет пониматься механизм комплексного оценивания.

1.7. Выводы по главе 1

В первой главе приведён обзор методов предварительной обработки информации в условиях ограниченных данных, к одному из которых относится интервальное кодирование области определения наблюдаемых параметров, что позволяет анализировать частоту наблюдений признаков в образованных диапазонах и выявлять их информационную значимость.

Такой подход реализован в методах системно-когнитивного анализа, основанных на теории информации. Эффективность этих методов и когнитивной аналитической системы «Эйдос-Xpro», автоматизирующей расчёты в данных методах, была многократно показана в работах [62–66], в том числе при работе с ограниченными данными.

Другой подход позволяет идентифицировать на основе набора дискретных данных дискретную функцию в виде суперпозиции функции двух переменных, особенностью которой является возможность её представления в графическом виде, поскольку последовательность операций над переменными можно изобразить в виде ориентированного графа – дерева критериев, а каждую дискретную функцию двух переменных можно представить в матричной форме [152]. Такое представление функций известно как МКО, или корень принятия решений. В работах [11; 12] показано, что любой МКО (корень принятия решений) может быть представлен в виде неполносвязной ИНС с каскадной структурой, что открывает широкие возможности синтеза методов идентификации МКО и нейросетевых технологий при обработке данных.

Поиск на возникший в настоящей главе вопрос о том, будут ли найденные ИНС обладать приемлемой для прикладных задач точностью, если их обучить на исходных данных в непрерывном виде, излагается в последующих главах.

ГЛАВА 2. ТЕОРЕТИЧЕСКИЕ АСПЕКТЫ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

Одним из востребованных методов анализа больших данных являются искусственные нейронные сети, часто использующиеся на практике, когда речь заходит об управлении сложными системами, в которых присутствуют в большей степени нелинейные параметры. Так, например, в задаче кластеризации, трудно формализуемая задача может быть решена путём большого человеческого ума либо с помощью адаптивных систем автоматизации. Единственный адаптивный способ нахождения зависимости, имея нелинейные параметры, – нейронная сеть. *(Математический аппарат нелинейной регрессии, реализованный с помощью искусственной нейронной сети).*

В результате искусственные нейронные сети могут решать достаточно сложные системы, которые трудны для поиска, обработки и анализа информации, при этом прогнозируя развитие ситуации, где функциональную роль занимают исследуемые объекты. Благодаря своей способности обучаться на больших объёмах, данных и адаптироваться к новым условиям, нейронные сети становятся незаменимым инструментом в таких областях, как медицина, финансы, промышленность и транспорт. Например, в медицине искусственные нейронные сети способны помочь в диагностике заболеваний, анализе медицинских изображений и персонализированном подборе лечения. В финансовой сфере они используются для прогнозирования рыночных тенденций, обнаружения мошенничества и управления рисками. В промышленности нейронные сети применяются для оптимизации производственных процессов, контроля качества и прогнозирования поломок оборудования. В транспортной отрасли они помогают в создании автономных систем управления, оптимизации маршрутов и распределении ресурсов. Нейронные сети становятся ключевым элементом в современном мире, где обработка больших данных и быстрый анализ информации являются необходимостью для успешного функционирования различных систем и отраслей.

Развитие ИНС и их практическое применение в современном мире привело к качественному росту получаемых результатов при анализе исследуемой области.

Однако это также привело к многообразию архитектур нейронных сетей и математических способов и программных продуктов для их реализации. Существующий теоретический и практический материал в литературе в достаточной степени позволяет получить представление об основах нейросетевого моделирования в различных областях исследования. Несмотря на это, в рамках логической постановки исследуемых проблем и гипотез далее в текущей работе приведен краткий обзор ключевых видов структур ИНС, основных способов их обучения и тестирования.

2.1. Базовое представление искусственных нейронных сетей

Теоретический фундамент в моделировании человеческого мозга как интеллектуальной системы заложен в трудах У. МакКаллока и У. Питтса, опубликованных в статье «Логическое исчисление идей, относящихся к нервной активности» в 1943 г. [171]. Впоследствии данная работа легла в основу нейроинформатики, тем самым запустив механизм развития и применения искусственных интеллектуальных систем в разных предметных областях.

Используя идею У. МакКаллока и У. Питтса о представлении биологического нейрона в виде модели искусственного математического нейрона, в 1958 г. в своей работе [172] американский ученый Ф. Розенблатт предложил первый персептрон с одним скрытым слоем или базовую архитектуру искусственной нейронной сети, более известную как нейронная сеть прямого распространения с одним скрытым слоем с жесткой пороговой функцией активации для всех нейронов. Параллельно с этим в период 1950–1963 гг. в СССР также развивается теория самоорганизующихся нейронных систем.

Долгое время не было возможности представить функцию многих переменных в виде суммы функций меньшего количества переменных. Получалось, что персептрон, сколько бы он нейронов ни имел, не всегда мог построить аппроксимацию любой функции нескольких переменных. Сомнения относительно возможностей персептронов развеяли советские математики – академики А.Н. Колмогоров и В.И. Арнольд, доказав, что любая функция n

переменных, определенная на единичном отрезке $E^1 = [0;1]$, всегда может быть представлена в виде композиции непрерывных функций одной переменной или бинарной операцией сложения [7–10]:

$$f(X) = f(x_1,...,x_n) = \sum_{q=1}^{q=2n+1} \left(\chi_q \left[\sum_{p=1}^{n} \psi^{pq}(x_p) \right] \right) \quad (2.1)$$

где $\chi_q(y)$ – действительные и непрерывные функции.

Из следствия теорем Арнольда – Колмогорова – Хехт-Нильсена необходимое количество нейронов в скрытых слоях персептрона можно определить по формуле:

$$\frac{N_y Q}{1+\log_2(Q)} \le N_w \le N_y \left(\frac{Q}{N_x} + 1 \right) (N_x + N_y + 1) + N_y \quad (2.2)$$

где N_y – размерность выходного сигнала; Q – число элементов множества обучающих примеров; N_w – необходимое число синаптических связей; N_x – размерность входного сигнала.

Продолжением теорем Арнольда – Колмогорова для нейронных сетей, стала работа Р. Хехт-Нильсена [173], в которой приводится доказательство универсальности нейронной сети вида двухслойного персептрона (или нейронная сеть с одним скрытым слоем), причем активационные функции его нейронов должны иметь сигмоидальную функцию.

Выбор активационной функции также находит свое подтверждение еще в одной универсальной теореме аппроксимации, доказанной Джорджем Цыбенко в 1989 г., в которой говорится, что искусственная нейронная сеть прямого распространения с одним скрытым слоем может аппроксимировать любую непрерывную функцию многих переменных с любой точностью при условии, что нейронная сеть имеет в скрытом слое достаточное количество нейронов, имеющих сигмоидальную активационную функцию [1].

Нейронные сети, нейроны которых сгруппированы в слои, образуют многослойные нейронные сети, причем если каждый нейрон предыдущего слоя связан со всеми нейронами следующего слоя, то такая нейронная сеть называется полносвязной, в противном случае нейронная сеть является неполносвязной или частично неполносвязной на определенных узлах сети.

На Рисунке 2.1 представлен общий вид многослойной искусственной нейронной сети с полносвязной сетью прямого распространения, где $\mathbf{x} = \left\{ x_1^{(1)}, x_2^{(2)}, ..., x_m^{(n)} \right\}$ – входные параметры (верхний индекс обозначает принадлежность к определенному скрытому слою); $\mathbf{y} = \left\{ y_1, ..., y_\eta \right\}$ – выходные параметры; $\mathbf{w}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\phi}$ – векторы весовых коэффициентов (синаптических связей);

$$\mathrm{W}=\begin{pmatrix} \mathbf{w}_1^T \\ \mathbf{w}_2^T \\ \dots \\ \mathbf{w}_w^T \end{pmatrix}; \quad \Theta=\begin{pmatrix} \boldsymbol{\theta}_1^T \\ \boldsymbol{\theta}_2^T \\ \dots \\ \boldsymbol{\theta}_\theta^T \end{pmatrix}; \quad \Gamma=\begin{pmatrix} \boldsymbol{\gamma}_1^T \\ \boldsymbol{\gamma}_2^T \\ \dots \\ \boldsymbol{\gamma}_\gamma^T \end{pmatrix}; \quad \Phi=\begin{pmatrix} \boldsymbol{\phi}_1^T \\ \boldsymbol{\phi}_2^T \\ \dots \\ \boldsymbol{\phi}_\phi^T \end{pmatrix}; \quad x_1^{(1)} = g^{(1)}\left(\mathbf{w}_1{}^T \cdot \mathbf{x} \right); \quad x_2^{(1)} = g^{(1)}\left(\mathbf{w}_2{}^T \cdot \mathbf{x} \right);$$

$$x_3^{(1)} = g^{(1)}\left(\mathbf{w}_3{}^T \cdot \mathbf{x} \right); \quad X^{(1)} = g^{(1)}\left(\mathrm{W} \cdot X \right); \quad X^{(2)} = g^{(2)}\left(\Theta \cdot X^{(1)} \right); \quad X^{(N)} = g^{(2)}\left(\Gamma \cdot X^{(N-1)} \right);$$

$$y = g^{(n+1)}\left(\Phi \cdot g^{(n)}\left(\Gamma \cdot g^{(n-1)}\left(\dots \cdot g^{(2)}\left(\Theta \cdot g^{(1)}\left(\mathrm{W} \cdot X \right) \right) \dots \right) \right) \right).$$

Общая задача обучения нейронной сети ставится следующим образом.

Пусть X – множество объектов, обладающее набором признаков; R – множество решений алгоритма, тогда существует неизвестная целевая функция $g' : X^n \to R$. Необходимо найти такую первообразную функции $g : X^n \to R$, которая восстанавливает оценку g'. То есть задача сводится к минимизации ошибки обучения, которая показывает разницу между требуемым (желаемым) и выходом, получающаяся в результате вычисления персептрона, где целью обучения ИНС является подбор оптимальных синаптических весов [174, с. 426].

Исходя из набора логических пар $\mathbf{d}_n{}' = \left(\mathbf{x}_n, \mathbf{r}_n \right)$, где $\mathbf{d}_n{}'$ составляет n-й прецедент, выделим три соответствующих подмножества прецедентов

$$D^{Tr} = \left\{\mathbf{d}_n{}'\right\}_{n=1}^{N_{Tr}} - \textit{Train Set},\ D^{Ts} = \left\{\mathbf{d}_n{}'\right\}_{n=1}^{N_{Ts}} - \textit{Test Set},\ D^{Tv} = \left\{\mathbf{d}_n{}'\right\}_{n=1}^{N_{Tv}} - \textit{Validation Set},$$

при условии исключения потери данных: $D^{Tr} \cap D^{Ts} \cap D^{Tv} \equiv \varnothing$ [174, с. 426].

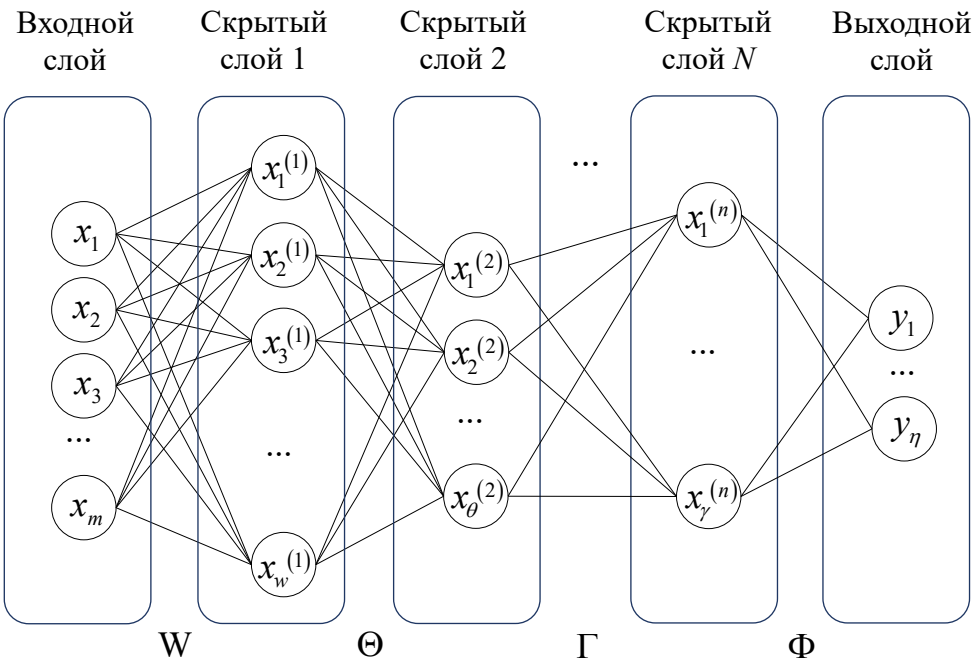


Рисунок 2.1 – Математическая модель многослойной полносвязной искусственной нейронной сети

Пусть $G_i : X_{i-1}^n \times W_i \to R$, где W – множество допустимых значений вектора параметров $\mathbf{w}$. Тогда выделяют следующие основные типы обучения нейронных сетей: параметрический – при фиксированном G_i определяется «оптимальное» значение $\tilde{\mathbf{w}}$ при минимальной ошибки функционала $L\left[G_i\left(X,\tilde{\mathbf{w}}\right),R\right] \to \min$; и структурный – поиск «оптимального» вида G_i. Исходя из подмножества прецедентов D^{Tr}, выделяют три основных режима обучения нейронных сетей: с учителем – $D^{Tr} \neq \varnothing$; без учителя – $D^{Tr} \equiv \varnothing$, и с подкреплением (взаимодействие с некоторой средой в процессе обучения) [174, с. 170, с. 426]. Таким образом, формально задача построения нейронной сети сводится к выбору алгоритма G_i в зависимости от области применения ИНС.

Согласно общепринятой технологии проектирования нейронных сетей всё множество примеров рекомендуется разбивать на следующие подмножества: обучающее (Train), тестирующее (Test) и подтверждающее (Validation) в соотношении 70 % : 20 % : 10 % соответственно [81, с. 143]. Однако в работе [81, с. 143] также отмечается, что в случаях, когда исходный набор данных ограничен (например в медицине), допускается разделять все множество только на обучающее и тестирующее в соотношении 90 : 10 %. Это также подтверждается в работе А.В. Макаренко, что, применяя инструменты машинного обучения по прецедентам, минимально выделяют два подмножества прецедентов – обучающую и тестовую выборки [95].

Суть существующих в настоящее время подходов к обучению нейронных сетей заключается в поиске минимума некоторой непрерывной функции ошибки. За последние несколько лет было предложено множество алгоритмов оптимизации, использующих различные подходы для обновления параметров модели [175].

В качестве примера эвристических методов обучения нейронных сетей, при котором идет подбор оптимальных синаптических весов нейросетевой модели, можно привести алгоритм упругого обратного распространения (*RPROP – Resilent back PROPagation*), в котором минимизация целевой функции производится методом градиентного спуска, но при корректировке весовых коэффициентов учитывается только знак соответствующей составляющей градиента [81, с. 121]:

$$\Delta w_{ij}(t) = -\eta(t)\, sign\left(\frac{\partial \varepsilon(t)}{\partial w_{ij}}\right). \tag{2.3}$$

Другой группой являются градиентные методы, в частности метод градиентного спуска ищет минимум целевой функции $\varepsilon = \varepsilon(w)$, зависящей от вектора весовых коэффициентов $w = [w_1, w_2, ..., w_n]^T$.

Усовершенствование метода градиентного спуска в простой форме привело к введению момента, когда влияние градиента на изменение весов накапливается со временем (формула (2.4)) [175, с. 126].

$$\Delta w_{ij}(t) = \mu \Delta w_{i-1\,j-1} - \alpha \nabla_w e(w), \tag{2.4}$$

где α – скорости обучения нейронной сети.

Однако в последнее время наиболее эффективными являются адаптивные методы обучения нейронных сетей, такие как Adagrad, RMSProp, Adadelta, Adam. В качестве примера приведем метод адаптивной оценки моментов (Adam) по формулам (2.5)–(2.6) [175, с. 127]:

$$m_N = \beta_1 m_{N-1} + (1-\beta_1) g_N; \tag{2.5}$$

$$\nu_N = \beta_2 \nu_{N-1} + (1-\beta_2) g^2{}_N, \tag{2.6}$$

где m_N, v_N – моменты первого и второго порядка, которые корректируются согласно формулам (2.7)–(2.8) [175, с. 127].

$$\overline{m}_N = \frac{m_N}{1-\beta_1^N}; \tag{2.7}$$

$$\overline{\nu}_N = \frac{\nu_N}{1-\beta_2^N}. \tag{2.8}$$

Пересчет полученных весов происходит по следующей формуле [175, с. 127]:

$$w^{N+1} = w^N - \frac{\eta}{\sqrt{\overline{\nu}_n + \varepsilon}} \overline{m}_N. \tag{2.9}$$

Для оценки качества нейронных сетей, используемых для большинства областей исследования, наиболее популярными являются следующие метрики (формулы (2.10)–(2.13)): средняя абсолютная ошибка (*Mean Absolute Error, MAE*), корень среднеквадратической ошибки (*Root Mean Square Error, RMSE*), коэффициент детерминации (R^2) и скорректированный коэффициент детерминации $(R^2{}_{ADJ})$ [86, с. 32–34].

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |d_i - y_i|, \tag{2.10}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (d_i - y_i)^2}, \tag{2.11}$$

где N – количество наблюдений; d_i – необходимый (желаемый) выход i-го нейрона; y_i – выходной сигнал, получаемый исходя из вычислений модели. Средняя абсолютная ошибка и корень среднеквадратической ошибки показывают среднюю ошибку модели в единицах необходимой нам переменной и лежат в интервале $[0;\infty)$.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(d_i - y_i)^2}{\sum_{i=1}^{N}(d_i - \overline{d})^2}, \tag{2.12}$$

$$R^2{}_{ADJ} = 1 - (1 - R^2)\frac{N-1}{N-n-1}, \tag{2.13}$$

где d_i – необходимый (желаемый) выход i-го нейрона; $\overline{d}$ – среднее значение величины d; N – количество наблюдений; n – количество входных параметров модели [86].

Однако в некоторых областях, например, в медицинской диагностике, дополнительно используются показатели чувствительности (*Se*, sensitivity) и специфичности (*Sp*, specificity) (формулы (2.14)–(2.15) соответственно) [86, с. 37], учитывающие различную значимость недопущения ошибок первого и второго родов, в отличие от средней абсолютной ошибки и корня среднеквадратической ошибки, в которых ошибки первого и второго рода учитываются в равной степени.

$$Se = \frac{TP}{N_p} \cdot 100\ \%, \tag{2.14}$$

$$Sp = \frac{TN}{N_N} \cdot 100\ \%, \tag{2.15}$$

где N_p – количество пациентов в выборке, у которых диагностируемое заболевание присутствует; N_N – количество пациентов в выборке, у которых диагностируемое заболевание отсутствует; *TP* – количество верно диагностированных случаев наличия заболевания; *TN* – количество верно диагностированных случаев отсутствия заболевания [86].

Как и в случае оценки качества полученных деревьев решений, для оценки качества классификаторов в машинном обучении используется метрика *Accuracy* (точность, доля правильных ответов) (формула (1.3), формула (2.16)). Другим критерием оценки является метрика – точность (*Precision*) (формула (2.17)) [176, с. 124].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (2.16)$$

$$Precision = \frac{TP}{TP + FP}. \quad (2.17)$$

В отличие от *Accuracy, Precision* менее чувствительна к дисбалансу классов и показывает, насколько хорошо алгоритм классифицирует данные в положительный класс и как этот класс соответствует реальному положительному классу.

Стоит отметить, что существуют и другие метрики, используемые для оценки качества нейросетевых моделей и их прогностической способности, которые подробно приведены в работе [86, с. 34–36].

Несмотря на успехи исследователей в области ИНС и понимание того, что ИНС обладает универсальными свойствами аппроксимации, многие задачи не удавалось решить в связи с ограниченностью скрытых слоев. Соответственно, появляются новые теории нейросетевого представления исследуемых объектов (неклассические ИНС), которые могли бы более точно описывать и прогнозировать зависимость в исследуемых переменных.

2.2. Искусственные нейронные сети с каскадной и свёрточной структурой

ИНС довольно часто используются на практике, когда речь заходит об управлении сложными экономическими системами, в которых присутствуют в большей степени нелинейные параметры, такие, например, как процесс кластеризации, когда стоит трудно формализуемая задача, решение которой можно достичь либо путём большого человеческого ума, либо с помощью адаптивных систем автоматизации. Единственный адаптивный способ нахождения зависимости,

имея нелинейные параметры, – нейронная сеть (математический аппарат нелинейной регрессии, реализованный с помощью нейронной сети).

В результате искусственные нейронные сети могут эффективно решать достаточно сложные системы, которые достаточно трудны для поиска, обработки и анализа информации, при этом прогнозируя развитие ситуации, где функциональную роль занимают исследуемые объекты.

Традиционный метод решения задачи аппроксимации, при котором составляется линейная комбинация фиксированных базисных функций, имеет среднеквадратическую ошибку аппроксимации $O\left(\left(\frac{1}{n}\right)^{\frac{2}{d}}\right)$, где n – число элементов, а d – размерность входного пространства. Так называемый феномен «проклятия размерности» делает практически невозможным применение традиционных методов в задачах, например, с большим объемом данных [99, с. 39].

Необходимость адаптации ИНС к каждой из исследуемых задач и совершенствование нейросетевых технологий привело к появлению новых видов ИНС. Если глобально подходить к определению видов нейронных сетей, то их можно классифицировать по структуре, по типу входной информации, по типу связи и способу обучения, по времени передачи сигнала и др.

Одними из неклассических ИНС являются каскадные нейронные сети (Рисунок 2.2), основанные на нелинейных рекуррентных ИНС, являясь обобщением линейной схемы двуслойной нейронной сети, позволяющие наращивать свою архитектуру в процессе обучения. Таким образом, архитектура ИНС представляет собой аналог каскада в результате последовательного добавления базовых элементов к структуре ИНС.

В работе [99] доказано, что каскадные нейронные сети на определённой области значений более эффективно описывают разрывные функции, чем классические двуслойные нейронные сети. Авторы на расчётных примерах показывают, что каскадная схема построения нейронной сети обладает теми же достоинствами, что и линейная схема (последовательная структура, точность

аппроксимации $O\left(\frac{1}{n}\right)$, независимость скорости аппроксимации от размера входных параметров).

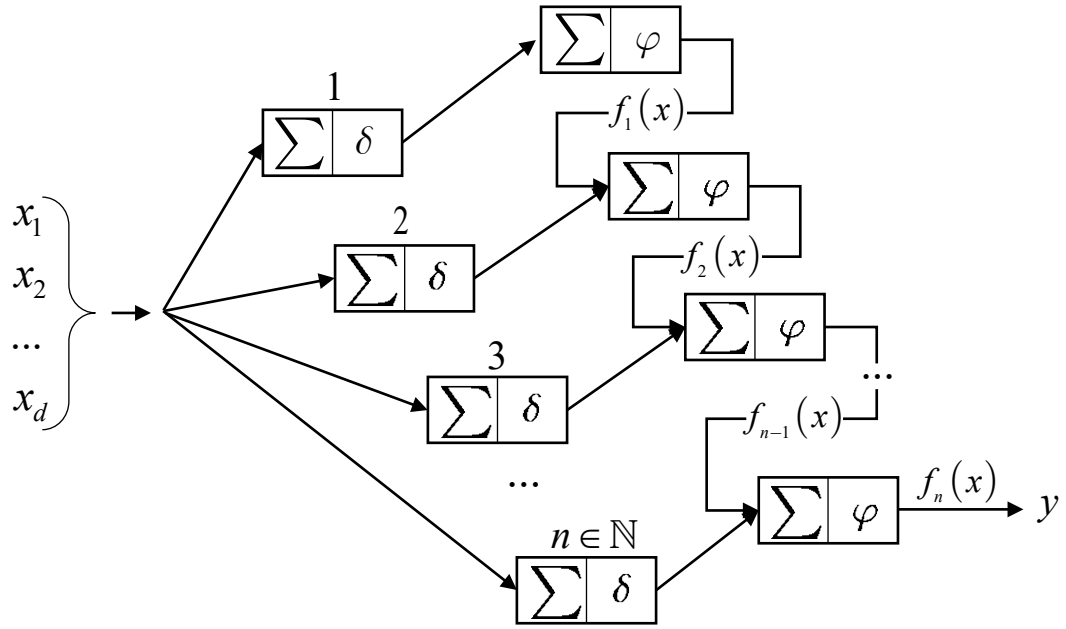


Рисунок 2.2 – Схематический рисунок каскадной нейронной сети:
$X \in \{x_1, x_2, ..., x_d\}$ – множество входных параметров; φ, δ – нелинейные функции;
$f(x)$ – отображение, являющееся многократной композицией функций базовых элементов и линейной функции $f : X \to R$ (составлено на основе [99 с. 41])

В другой работе Ян Лекун с соавторами [96] показывает применение ИНС к реальной задаче на примере распознавания рукописных цифр в почтовом индексе. Авторы предложили использовать архитектуру ИНС, основанную на принципе разделения весов. Впоследствии идеи данного исследования были дополнены и представлены в работе [97], результаты которой явились основой для свёрточных нейронных сетей (convolutional neural networks, CNN), получившие особую популярность при анализе изображений. Такие ИНС состоят из последовательной комбинации двух типов слоев: свёртка (convolution) и субдискретизация (subsampling).

Свёрточные преобразования в глубоких свёрточных нейронных сетях могут быть интерпретированы как полудискретные [105, с. 3], возвращающие функцию,

соответствующую взаимно корреляционной функции, к которым была применена свёртка, представленная в виде линейного оператора [105, с. 5]:

$$(f \cdot \mu)(x) \overset{def}{=} \int_{R^n} f(y)\mu(x-y)dy = \int_{R^n} f(x-y)\mu(y)dy, \qquad (2.18)$$

где $(f,\mu): R^n \to R$ – функции, интегрируемые относительно меры Лебега на пространстве R^n.

Общая архитектура свёрточной искусственной нейронной сети представлена на Рисунке 2.3, где C_1, C_3, C_5 – совокупность свёрточных слоев, S_2, S_4 – совокупность подвыборочных (объединяющих) слоев (субдискретизация). Свёрточная нейронная сеть представляет собой сочетание свёрточных и субдискретизационных слоёв, которые выполняют нелинейное иерархическое преобразование входного пространства образов [90].

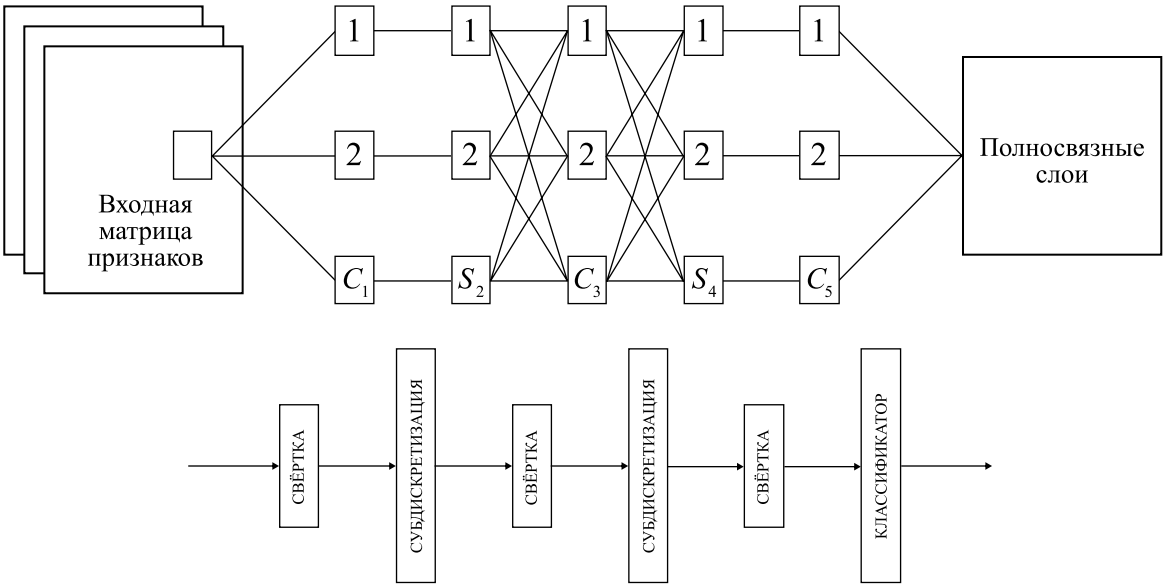


Рисунок 2.3 – Общий вид модели свёрточной искусственной нейронной сети (Рисунок оптимизирован и дополнен автором из работы [90, с. 128–130])

С 2010 г. свёрточные ИНС приобретали все более новый характер нейронных сетей, совершенствовались алгоритмы, подходы к их обучению и выбору архитектуры; появились различные модификации свёрточных ИНС, каждая из которых являлась оптимальной для конкретной области применения по какому-либо критерию.

То, что классическое представление ИНС требовало развития для решения более сложных задач с помощью нейросетевых технологий, также подтверждается в работе [106], где предлагается использовать гибридный подход обучения ИНС, сочетающий алгоритм самоорганизованного обучения и алгоритм контролируемого обучения. Предложен новый подход, который сочетает в себе преимущества нечетких систем и нейронных сетей. Это может быть полезно для решения задач управления в условиях неопределённости и нелинейности, где традиционные методы могут быть неэффективными.

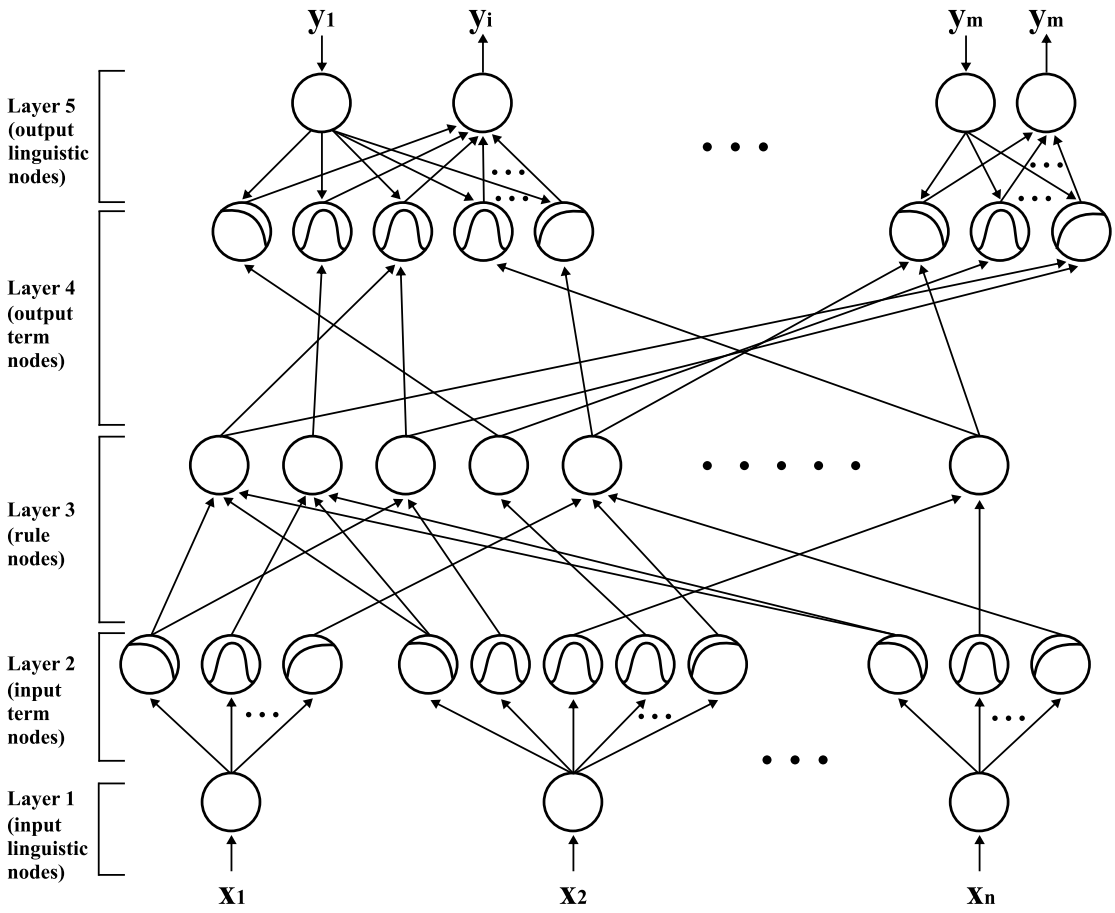


Рисунок 2.4 – Нейро-нечёткая сеть управления (NFCN) [106, с. 186]

В работе [106] приведены результаты предлагаемых гибридных систем, где для реализации нечёткой логической системы управления предлагается общая коннекционистская модель, названная нейро-нечёткой сетью управления (NFCN). Предлагаемая NFCN (Рисунок 2.4) представляет собой многослойную сеть с обратной связью, которая объединяет основные элементы и функции традиционного

нечёткого логического контроллера в коннекционистскую структуру, обладающую способностью к распределённому обучению. NFCN может быть построена на основе контролируемых обучающих примеров с помощью методов машинного обучения, а коннекционистская структура может быть обучена для разработки правил нечёткой логики и нахождения функций принадлежности.

Компьютерное моделирование управления беспилотным автомобилем и решение задачи балансировки тележки на столбах с применением различных комбинаций моделей, основанных на NFCN, и алгоритмов обучения нейро-нечёткой сети с подкреплением и без подкрепления показало эффективность и применимость предложенных гибридных подходов к нейросетевому моделированию. Нейро-нечёткая сеть управления успешно обучается и адаптируется к различным условиям работы, что подтверждается результатами моделирования [106]. Такой тип нейронной сети, предложенный в работе [106], является примером того, как современные технологии искусственного интеллекта могут быть объединены для создания более мощных и гибких систем.

2.3. Искусственные нейронные сети, основанные на корнях принятия решений

Несмотря на успехи в области построения нейросетевых моделей, проблема выбора оптимальной архитектуры ИНС является одной из ключевых задач при нейросетевом моделировании. Впервые А.О. Алексеевым был предложен алгоритм, который позволяет использовать матричные свёрточные преобразования для получения полной архитектуры нейронной сети. В работе [11, с. 402] показано, что любой МКО (корень принятия решений) может быть представлен в форме неполносвязной ИНС с каскадной структурой, что открывает широкие возможности синтеза методов идентификации МКО (корня принятия решений) и нейросетевых технологий при обработке данных [12, с. 156].

Для решения задачи МКО, состоящего из двоичного дерева, набора матриц свёртки и шкал оценки, в настоящее время применяется единый подход. Долгое время этот подход базировался на представлении механизма комплексного

оценивания в виде суперпозиции произведений однократно закодированных критериев и матриц. Идентификация древовидной структуры критериев и элементов матриц свёртки определяется с помощью логических методов и полиномов, таких как машинное обучение, а также различных алгебраических методов. Однако, согласно такому подходу, элементы обучающего набора представлены дискретными значениями.

С математической точки зрения, МКО является бесповторной переключательной дискретной *k*-арной функцией *n* переменных, представленной в виде суперпозиции *n*–1 функций двух переменных[1].

Нельзя не вспомнить знаменитые работы [7–10] академика АН СССР Андрея Николаевича Колмогорова и его ученика, впоследствии также ставшего академиком РАН, Владимира Ивановича Арнольда о представлении функции многих переменных в виде суперпозиции функций меньшего числа переменных. Собственно, идея МКО, по словам тех, кто стоял в 80-е гг. XX в. у их истоков, была продиктована именно работами [7–10]. Стоит признать, что похожие задачи встречались в середине XX в. в электротехнике и формулировались в виде представления функции многих переменных минимальным числом элементарных функций. В 2021 г. подход [11] был обобщен [146] на случай МКО с любой структурой дерева критериев, отражающего последовательность функций в суперпозиции.

Как показано в Главе 1 в разделе 1.6, по МКО (корень принятия решений) можно определить некоторые параметры неполносвязной нейронной сети с каскадной структурой на основе корня принятия решений. Такой тип искусственных нейронных сетей в [155] назван Decisions' Root-based Neural Network (DRB NN) – нейронная сеть, основанная на корне принятия решений.

Идентифицированные неполносвязные ИНС полностью воспроизводят обучающие наборы с дискретными значениями. Соответственно возник вопрос, будут ли найденные ИНС обладать приемлемой для прикладных задач точностью, если их обучить на исходных данных в непрерывном виде? [152]. Анализ

[1] Данное определение введено совместно с научным руководителем А.О. Алексеевым в работе [152].

нейросетевых технологий показал, что необходимость адаптации ИНС к каждой из исследуемых задач и совершенствование нейросетевых технологий привели к появлению новых видов ИНС и алгоритмов оптимизации, использующих различные подходы для обновления параметров модели с целью подбора оптимальных синаптических весов. В результате анализа были предложены правила и методы предварительной обработки данных и выбора архитектуры нейронной сети для её обучения (Рисунок 2.5).

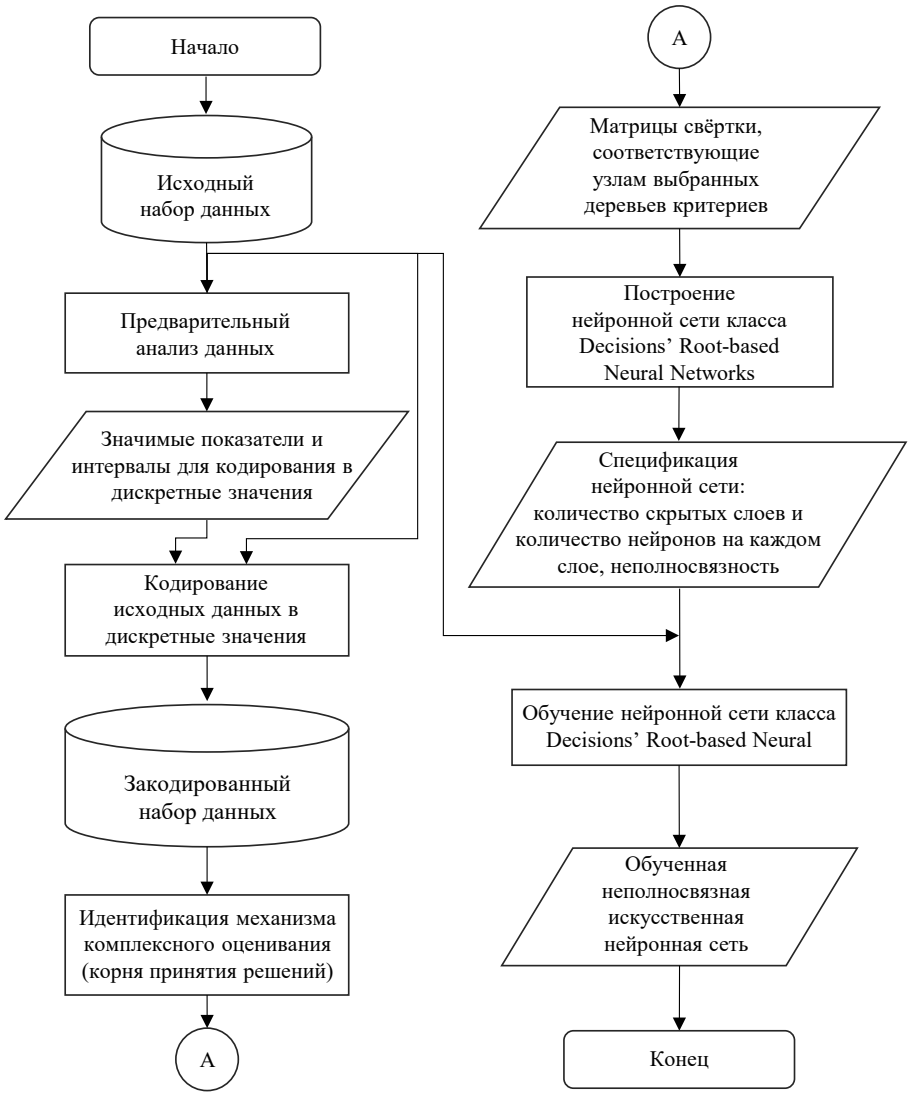


Рисунок 2.5 – Блок-схема предлагаемых правил и методов предварительной обработки данных и выбора архитектуры нейронной сети для обучения нейронной сети класса Decision's-Root-based Neural Network в ограниченных случаях исходного набора данных

Правила и методы предварительной обработки ограниченных данных включают в себя три этапа (см. Рисунок 2.5). На первом этапе с помощью методов

АСК-анализа исходные параметры, разбитые на интервалы, оцениваются по степени их информационной значимости, что позволяет выбрать для дальнейшей обработки только те параметры, которые демонстрируют эффект переключения.

На втором этапе синтезируются полные бинарные графы с поименными листьями и идентифицируется набор функций двух переменных, образуя для конкретного дерева графо-матричное представление функции многих переменных – корень принятия решений.

На третьем этапе определяется спецификация неполносвязной нейронной сети, основанной на корне принятия решений. Алгоритм обработки ограниченных данных завершается тем, что найденная нейронная сеть обучается на исходных данных в непрерывном виде.

С другой стороны, можно найти ИНС, сопоставимые по количеству неизвестных параметров, и сравнить их эффективность, зная количество неизвестных параметров, которые будут определены в процессе обучения нейронной сети.

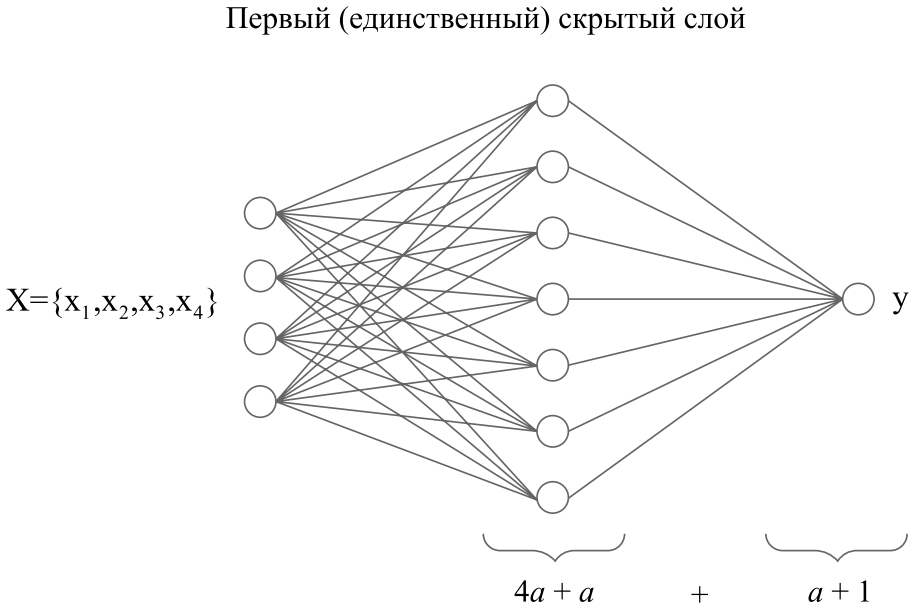


Рисунок 2.6 – Схематическое представление структуры нейронной сети с одним скрытым слоем и количеством неизвестных параметров

Итак, полносвязная нейронная сеть с одним скрытым слоем (Рисунок 2.6) с 4 входными сигналами и 1 выходным сигналом имеет следующее правило:

каждый нейрон на скрытом слое соответствует 4 входным неизвестным переменным, соответствующим входным сигналам, и 1 неизвестной константе (т.е. 5*a* неизвестных параметров); выходной сигнал получает на входе сигналы от *x* нейронов и имеет 1 неизвестную константу.

Таким образом, для вычисления количества неизвестных параметров, которые необходимо определить для полносвязной нейронной сети с одним скрытым слоем используется неравенство (2.19):

$$4a + a + 1 \geq 175, \quad a \in N. \tag{2.19}$$

Так, согласно неравенству (2.17), минимальное количество нейронов на скрытом слое для однослойной нейронной сети равняется 35, что является сопоставимым по количеству неизвестных параметров с найденной неполносвязной нейронной сетью (см. Главу 1, Рисунок 1.8).

Для нейросетевой структуры с двумя скрытыми слоями (Рисунок 2.7) общее количество синаптических связей определяется согласно выражению:

$$5a + ab + 2b + 1, \quad a,b \in N, \tag{2.20}$$

где *a* – количество нейронов на первом скрытом слое, *b* – количество нейронов на втором скрытом слое.

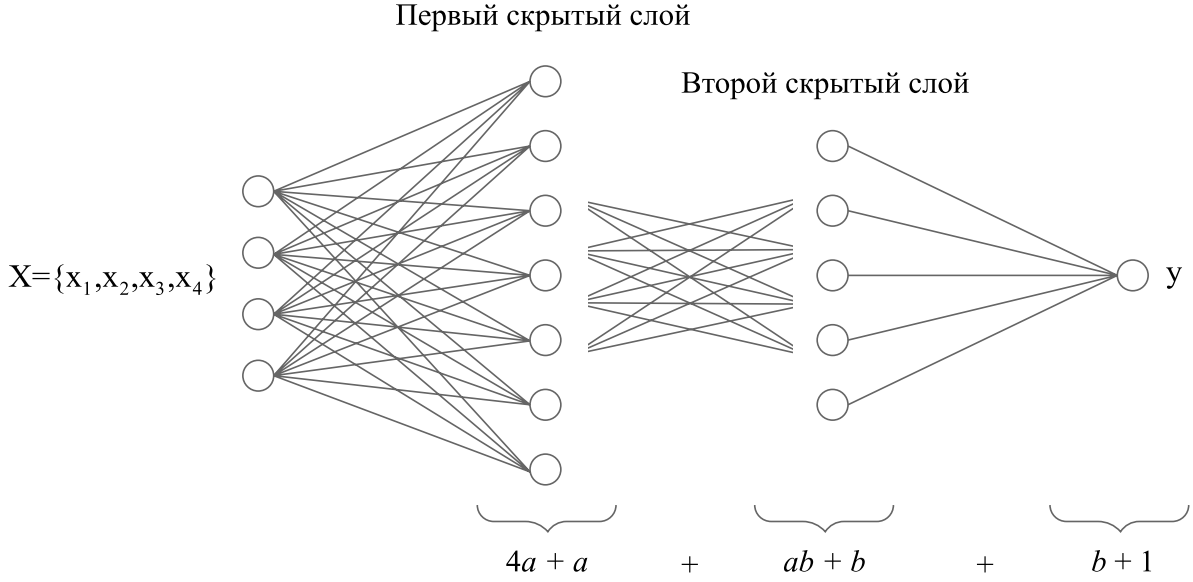


Рисунок 2.7 – Схематическое представление структуры полносвязной нейронной сети с двумя скрытыми слоями и числа неизвестных параметров

Минимальные целые значения *a* находились путём изменения количества нейронов на втором скрытом слое *b*, начиная с 1, при условии выполнения, что общее количество нейронов, определённое по выражению (4), не меньше 175. Эти значения определяют количество нейронов на первом скрытом слое (см. Рисунок 2.6). Таким образом, можно найти различные нейронные сети с двумя скрытыми слоями, сопоставимые по количеству неизвестных параметров нейронной сети, основанной на корне принятия решений.

2.4. Выводы по главе 2

В результате анализа нейросетевых технологий была предложена последовательность предварительной обработки данных (см. Рисунок 2.5), которые в совокупности с предложенными методами позволяют определить архитектуру нейронной сети для её обучения в условиях ограниченных наборах данных. Так, первый процесс может быть выполнен с помощью автоматизированной когнитивной аналитической системы «Эйдос-Xpro» [61], в которой реализованы три статистических и семь информационных моделей, основанные на теории информации. Эффективность этих методов и системы «Эйдос-Xpro» была многократно показана в работах [62–66], в том числе при работе с ограниченными данными.

Второй процесс может быть выполнен с помощью алгоритмов идентификации механизма комплексного оценивания (корней принятия решений), предложенных в работах [11; 13; 126; 127; 143; 146]. Одна из первых программ, позволяющих автоматизировать процесс идентификации механизма комплексного оценивания, появилась в 2023 г. – «Сборщик полных бинарных деревьев на именованных листьях на основе данных анализа групп эквивалентности» [177].

Третий процесс может быть выполнен с помощью методов [11; 12]. Однако до настоящего исследования программных продуктов, позволяющих строить нейронные сети по корням принятия решений в автоматическом режиме, не существовало. Одной из основных проблем в работе с неполносвязными

искусственными нейронными сетями, основанными на корнях принятия решений, является отсутствие специализированного программного обеспечения, которое позволило бы задавать структуру ИНС типа DRB NN для пользователей, не обладающих навыками профессионального программирования.

Существующая программа «Программный комплекс, реализующий работу неполносвязных нейронных сетей» [178], созданная в Пермском государственном национальном исследовательском университете Андреем Леонидовичем Гусевым и его аспирантом Александром Андреевичем Окуневым, позволяет реализовать обучение неполносвязных ИНС, однако в указанной программе есть возможность создать неполносвязную сеть только на одном слое, что не соответствует количеству синаптических связей нейронной сети, которые определяются в полном соответствии с элементами матриц свёртки в корне принятия решений.

Вышеперечисленные обстоятельства определяют необходимость в разработке специального алгоритмического и программного обеспечения информационной системы, позволяющей реализовать второй и третий этапы предлагаемых правил и методов предварительной обработки данных.

ГЛАВА 3. ИНТЕЛЛЕКТУАЛЬНЫЕ ИНФОРМАЦИОННЫЕ СИСТЕМЫ ОБРАБОТКИ ДАННЫХ

В современном мире сохраняется тенденция автоматизации технологических процессов и цифровизации бизнес-процессов в любых сферах экономики. Для этого требуются эффективные методы обработки оцифрованных сведений, результатов наблюдений, экспериментов, статистических показателей. Новые алгоритмы распознавания, обработки и структурирования таких данных находят свое применение в любой сфере человеческой деятельности. Интеллектуальные системы всё больше становятся неотъемлемой частью нашей жизни, которые имеют широкий спектр возможностей – от простой задачи автоматизации до анализа больших данных, моделирования и прогнозирования сложных процессов и систем.

Существует два обособленных типа интеллектуальных информационных систем (ИИС): экспертные системы и системы, основанные на анализе данных (на машинном обучении). В одном случае получение знаний происходит путём извлечения информации из экспертных знаний и их последующая формализация, в другом случае информация извлекается исходя из анализа данных, а уже из полученной информации формируются какие-либо данные.

Это определило область применения тех или иных информационных систем, что породило развитие и применение разных методов при решении прикладных задач. Тогда, при ограниченной области применения информационных систем, экспертные системы и системы, основанные на машинном обучении, развивались условно параллельно с периодическими пересечениями. Такой синтез методов, базирующихся на экспертных системах и машинном обучении для анализа данных, включает в себя комплексный подход к созданию и совершенствованию текущих алгоритмов, в том числе на основе нейронных сетей.

Как показано ранее в Главе 2, на основе любого корня принятия решений можно определить структуру неполносвязной искусственной нейронной сети (Decisions Root-based Neural Networks, DRB NN). Однако одной из проблем при идентификации корней принятия решений является выбор структуры графа, представляющего собой полное бинарное дерево с поимёнными листьями.

Другой проблемой при реализации нейросетевого моделирования полученной ИНС на основе корня принятия решений является то, что в настоящее время отсутствуют в открытом доступе специализированные программные продукты, позволяющие задавать структуру ИНС типа DRB NN, ориентированные на пользователей, не имеющих навыков профессионального программирования, что порождает потребность в разработке методов и алгоритмов предобработки данных при построении неполносвязных нейронных сетей, основанных на корнях принятия решений, и также программного инструментария проектирования и обучения нейронных сетей, основанных на корнях принятия решений [154].

3.1. Программный инструментарий автоматизированного системно-когнитивного анализа – интеллектуальная система «Эйдос-Xpro»

Предварительная обработка исходных данных заключается в определении наиболее значимых факторов, интервальном кодировании исходных данных и поиске МКО (корня принятия решений).

Согласно [11, с. 398] МКО определяется как кортеж (3.1):

$$<G,M,X,P>, \tag{3.1}$$

где G – граф, описывающий последовательность свёртки частных критериев, M – множество матриц свёртки, соответствующих узлам дерева критериев, X – множество шкал для оценки частных критериев, P – процедура агрегирования.

Одной из проблем при идентификации корней принятия решений является выбор структуры графа G, представляющего собой полное бинарное дерево с поимёнными листьями [179], так как их общее количество определяется по уравнению [127]:

$$\left|G\right| = \left(2l-3\right)!!, \tag{3.2}$$

где l – общее количество рассматриваемых переменных (листьев в терминах древовидного графа). Как видно из уравнения (3.1), количество структур G существенно зависит от числа переменных, так, например, для 4 переменных

количество структур равно 15, для 5 – 105, для 6 – 945, а их уже более 34 миллионов, например, для 10 переменных.

Для уменьшения числа переменных и решения задачи классификации в данном исследовании предлагается использовать методы системно-когнитивного анализа с помощью автоматизированной когнитивной аналитической системы «Эйдос-Xpro» [61], тем самым формализовав исследуемую предметную область, проведя синтез и верификацию полученных моделей.

На начальном этапе с помощью автоматизированных алгоритмов системы «Эйдос-Xpro», в режиме 2.3.2.2 (Рисунок 3.1) исходный набор исследуемой предметной области формируется в нормализованную классификационными и описательными шкалами и градациями обучающую выборку.

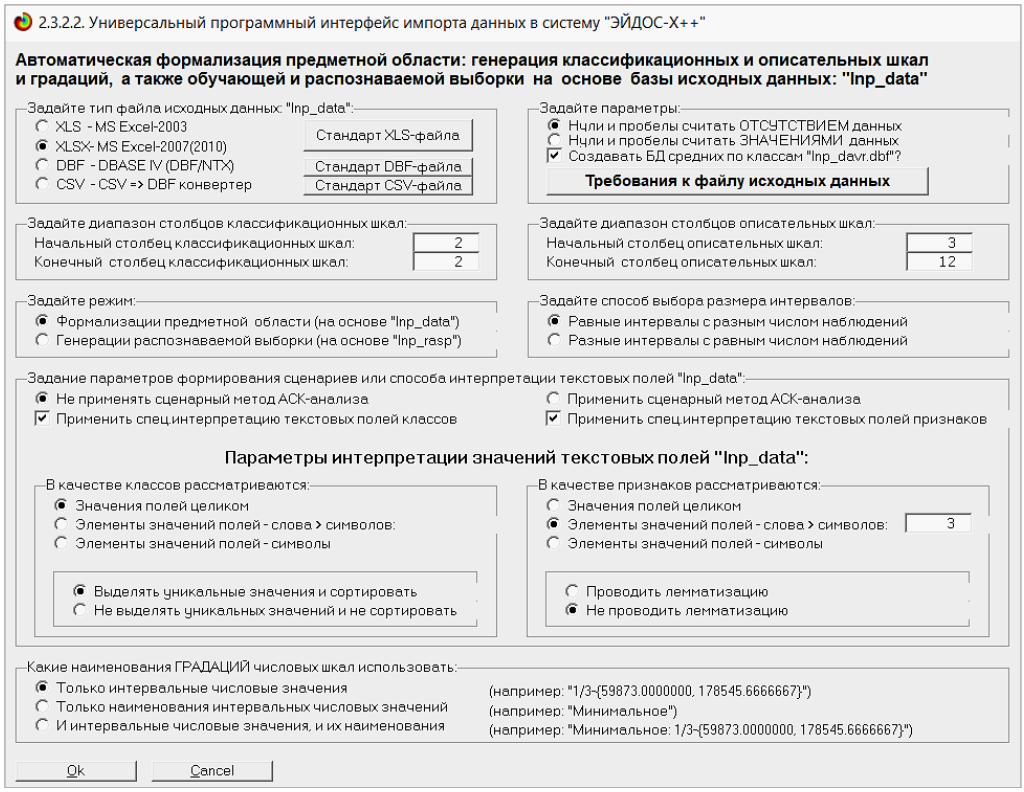


Рисунок 3.1 – Экранная форма режима «Универсальный программный интерфейс формализации предметной области и импорта данных» в интеллектуальной системе «Эйдос-Xpro» [61, режим 2.3.2.2]

Одной из задач предварительной обработки данных является определение количества интервалов, а затем их границы (ширина) для преобразования непрерывных значений в дискретные. Так, в работе [180] приведен достаточно

полный обзор существующих методов дискретизации и предложен авторский алгоритм с применением кластеризации на основе плотности и использовании инструментов RST для получения удовлетворительных результатов дискретизации.

Существует два способа разбиения области наблюдаемых значений на интервалы: разбиение области наблюдения на равные интервалы $l_1 = l_2$ с разным количеством объектов $N_1 \neq N_2$ (Рисунок 3.2, *а*) и на разные интервалы $l_1 \neq l_2$ с равным количеством объектов $N_1 = N_2$ (Рисунок 3.2, *б*).

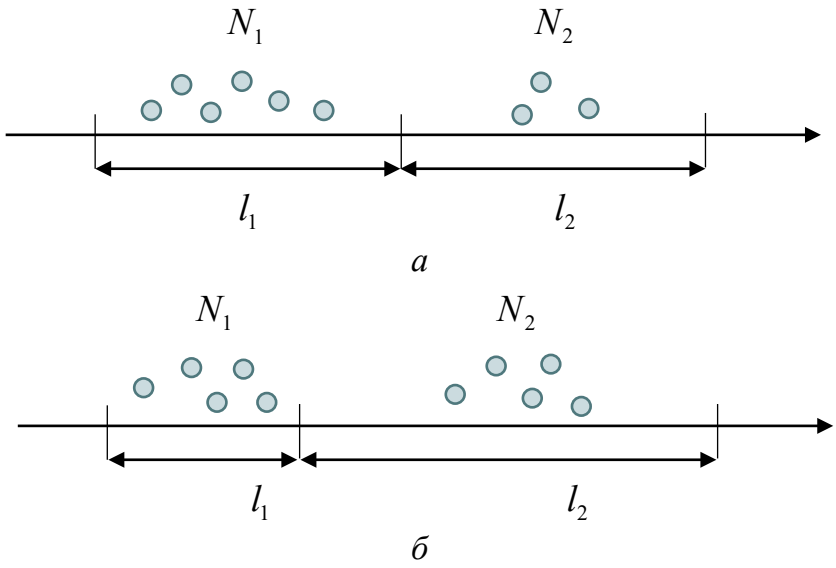


Рисунок 3.2 – Примеры (*а*, *б*) способов разбиения области наблюдаемых значений на интервалы

В зависимости от способа разбиения области наблюдаемых значений на интервалы достоверность, полнота и корректность получаемых моделей (*INF1–INF7*) исследуемой предметной области будет меняться. В данном исследовании были представлены результаты нескольких экспериментов относительно способа разбиения области наблюдаемых значений на интервалы (см. Главу 4).

Далее в результате применения АСК-анализа определяются диапазоны значений исследуемых показателей, обладающих наибольшей информационной значимостью для отнесения их к одному из классов («способствующие»), а также признаки, которые в наибольшей степени не характеризуют этот класс («препятствующие»).

Для примера покажем разбиение исследуемой области наблюдений на два равных интервала с разным числом наблюдений и определение наиболее значимых факторов (Рисунок 3.3).

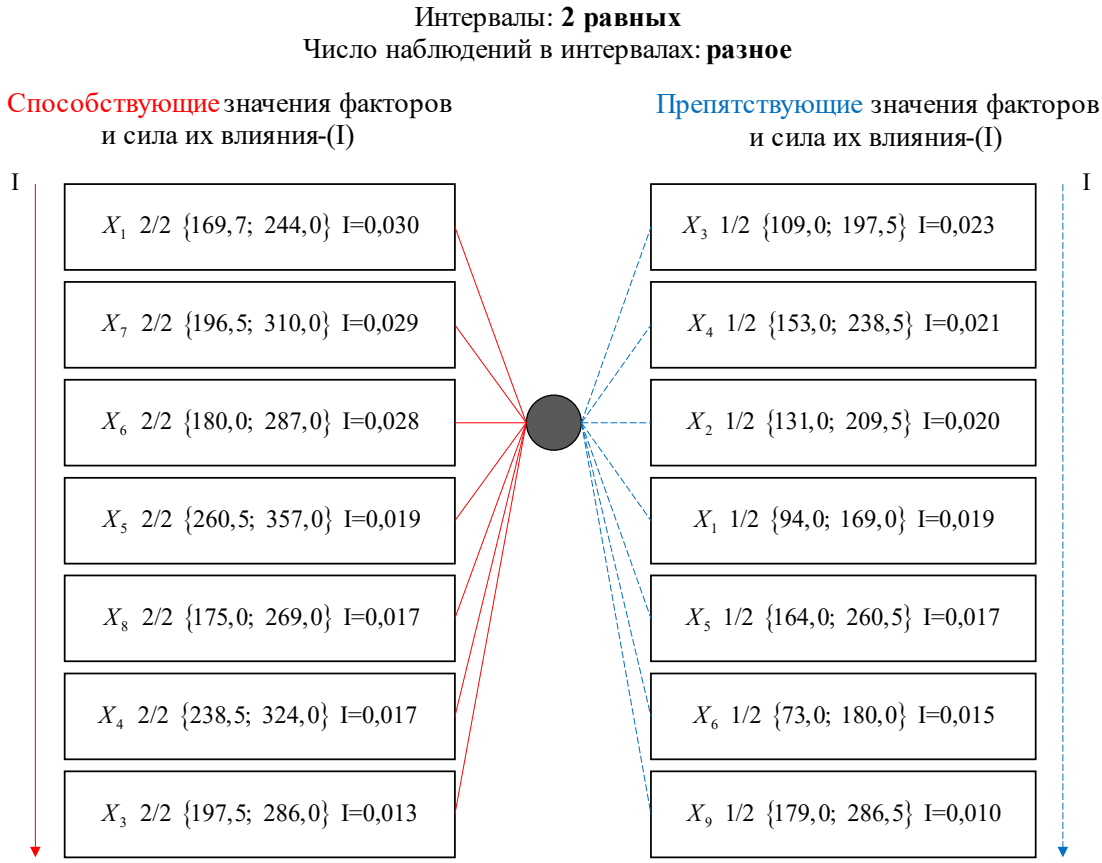


Рисунок 3.3 – Значимые факторы (переменные),
обладающие наибольшей информационной значимостью

На Рисунке 3.3 после названия переменной (например, X_1), указывается порядковый номер интервала из 2 значений, а в фигурных скобках непосредственно указаны значения интервала исследуемой переменной. В данном примере средний разброс показателя силы влияния (I) между переменными равен 0,01, поэтому нет возможности достоверно определить степень влияния какого-либо фактора конкретной модели, представленной в качестве примера на Рисунке 3.3, полученной при способе разбиения исследуемой области наблюдений на 2 равных интервала с разным числом наблюдений.

Тем не менее, когда существует много способов разделить изучаемую область наблюдений, методы системно-когнитивного анализа, используя

автоматические алгоритмы из системы «Эйдос-Xpro», показывают наиболее существенное влияние тех или иных переменных в анализируемых исходных данных, что может использоваться для выявления наиболее значимых факторов в исследуемой области и последующего уменьшения количества показателей.

3.2. Программная реализация деревьев решений для классификации

Другим инструментом, позволяющим решать задачи классификации и являющимся одним из типов контролируемого алгоритма машинного обучения, полагаются деревья решений.

Исходные данные разделяются на более конкретные категории или классы на основе признаков или атрибутов до тех пор, пока не будет выполнено условие остановки разделения или условие остановки обучения дерева решений.

В качестве основных атрибутов разбиения узлов выбираются следующие критерии (см. Таблицу 1.2): теоретико-информационный критерий (информационная энтропия); статистический подход (коэффициент Джини); log_loss (логарифмическая потеря). Для оценки качества деревьев решений наиболее часто используют следующие показатели: точность измерений (тестовые данные), F1-мера (тестовые данные), средняя потеря Хэмминга, или расстояние Хэмминга, корреляция Мэтьюза (MCC) (формулы (1.1)–(1.6)).

Основные теоретические аспекты обучения деревьев решений заключаются в следующем. Пусть все множество наблюдений D разделено на D^{Tr} – обучающее (*Train Set*) и тестирующее (*Test Set*) – D^{Ts}. Пусть D^{Tr} содержит *n* примеров, каждый из которых имеет метку класса $K_i\left(i=1,...,n\right)$, и m атрибутов $A_j\left(j=1,...,m\right)$, определяющие принадлежность объекта к какому-либо классу. В таком случае возможны следующие варианты [181]:

1. Все обучающее множество D^{Tr} принадлежит одному классу (все примеры, входящие в D^{Tr}, имеют одинаковую метку $K_i\left(i=1,...,n\right)$). В таком случае дерево решений представляет собой единственный лист класса $K_i\left(i=1,...,n\right)$, а его практическое значение не имеет смысла, так как любой новый объект данная модель будет относить к этому единственному классу $K_i\left(i=1,...,n\right)$.

2. В обучающем множестве D^{Tr} присутствуют все примеры принадлежности всех классов K_k. В таком случае D^{Tr} разбивается на подмножества $\left(D^{Tr}{}_1, D^{Tr}{}_2, ..., D^{Tr}{}_q\right)$, ассоциированные с этими классами путём выбора одного из атрибута $A_j\left(j=1,...,m\right)$ обучающего множества D^{Tr}, содержащее два и более неповторяющихся значения $\left(a_1, a_2, ...a_q,\right)$, где q – число неповторяющихся значений признака. Каждое из таких подмножеств $\left(D^{Tr}{}_1, D^{Tr}{}_2, ..., D^{Tr}{}_q\right)$ включает примеры, которые имеют такие же значения, как значения атрибута. Далее происходит перевыбор атрибута $A_j\left(j=1,...,m\right)$, и разбиение узла дерева продолжается до тех пор, пока все примеры в подмножествах $\left(D^{Tr}{}_1, D^{Tr}{}_2, ..., D^{Tr}{}_q\right)$ не будут одного класса или до выполнения дополнительных заранее определенных условий остановки обучения модели [181].

Во всех вариантах тестирующее множество D^{Ts} (*Test Set*) служит для оценки обобщающей способности модели, тем самым позволяя оценить качество получаемых моделей деревьев решений для решения задачи классификации. Стоит отметить, что, кроме задачи классификации, деревья решений являются одним из типов контролируемого алгоритма машинного обучения, используемым для решения задач регрессии.

Забегая вперед, в данной диссертационной работе представлен анализ многократных экспериментов с различными вариациями обучения и тестирования деревьев решений (см. Главу 4) на данных как с их предварительной обработкой, так и без их предобработки. Эта серия экспериментов является дополнительным исследованием, подтверждая необходимость предварительной обработки исходного набора данных при дальнейшем построении неполносвязных нейронных сетей (Decisions Root-based Neural Networks, DRB NN)[2], основанных на корне принятия решений (см. Главу 4).

Для вычислительных экспериментов был написан программный код на языке программирования Python с использованием одной из ключевых библиотек

[2] Данное определение было введено совместно с научным руководителем А.О. Алексеевым в работе [155].

для реализации деревьев решений – *sklearn.tree* [160] (получено свидетельство о государственной регистрации программы для ЭВМ (Приложение 8)). Фрагмент программного кода представлен ниже на Рисунке 3.4.

```
count = 0

data = []
while count<190:
    count = count+1

    x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.2)
    model = tree.DecisionTreeClassifier(criterion="entropy", splitter="best", min_samples_split=5)

    model.fit(x_train, y_train)
    p=model.score(x_train, y_train)
    print(p)

    y_pred = model.predict(x_test)
    a1 = accuracy_score(y_true = y_train, y_pred = model.predict(x_train))
    a2 = accuracy_score(y_true = y_test, y_pred = y_pred)
    a3 = f1_score(y_true=y_test, y_pred=y_pred, pos_label="AD")
    a4 = f1_score(y_true=y_test, y_pred=y_pred, pos_label="CON")
    a5 = hamming_loss(y_true=y_test, y_pred=y_pred)
    a6 = matthews_corrcoef(y_true=y_test, y_pred=y_pred)
```

Рисунок 3.4 – Фрагмент программного кода, реализующий серию экспериментов обучения и тестирования деревьев решений [160]

Для оценки качества деревьев решений была использована библиотека sklearn.metrics, в частности, импортированы следующие метрики: accuracy_score, f1_score, hamming_loss, matthews_corrcoef.

Разработанный программный модуль позволяет провести ряд экспериментов обучения и тестирования нескольких вариаций деревьев решений на основе исходных данных, хранящиеся в файле расширения *.xlsx. После выполнения вычислений формируется сводная таблица с метриками качества обучения и тестирования для заданного количества получаемых деревьев решений. Программный модуль сохраняет построенные деревья в графическом виде формата *.svg (Рисунок 3.5), что в дальнейшем позволяет пользователю в удобном виде использовать деревья решений как систему поддержки принятия решений.

На Рисунке 3.5 показан пример дерева решений для 10 переменных $X = \left(X_1, X_2, ..., X_{10}\right)$, который реализуется посредством выполнения программного кода (см. Рисунок 3.4, Приложение 8).

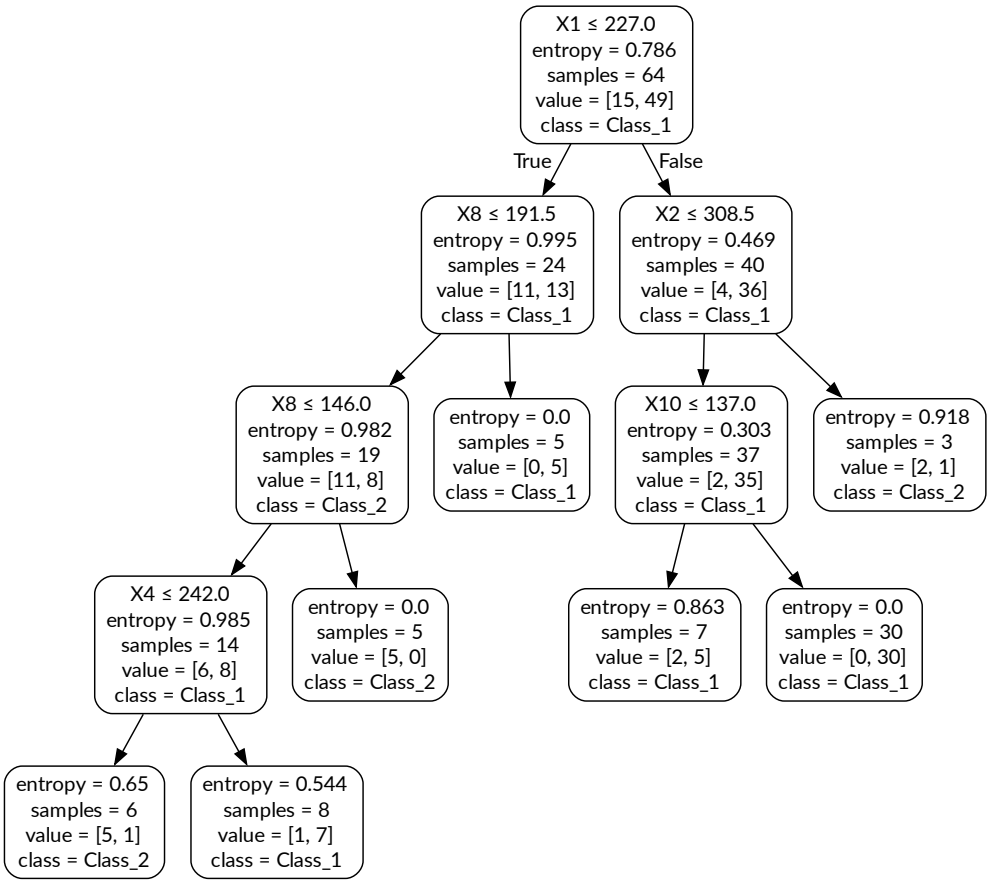


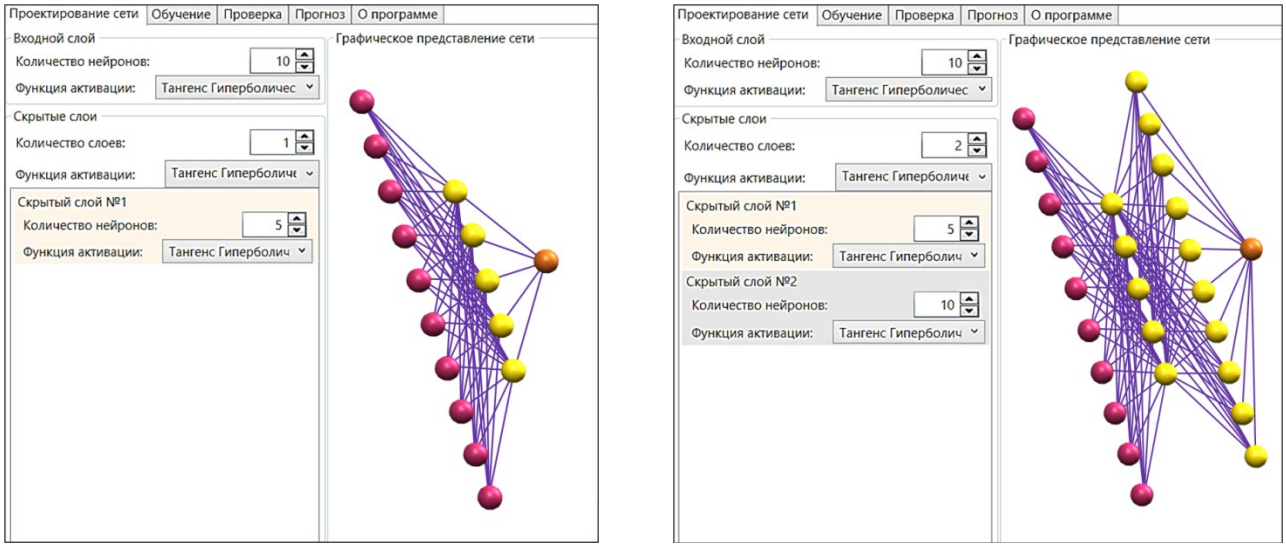
Рисунок 3.5 – Дерево решений для 10 входных переменных
с 2 классами исходных данных

В каждом узле дерева решений определено решающее правило, по которому принимается решение о том, выполняется оно (True) или нет (False), а также содержится информация об атрибуте разбиения узлов и уровне его информационной неопределённости. Также показано общее количество примеров (samples) на конкретном узле разбиения и количество примеров (value = [,]), формирующие принадлежность к тому или иному классу соответственно (Class_2 или Class_1).

3.3. Программный инструментарий построения нейронных сетей – «Нейросимулятор 5.0»

Работа с неполносвязными искусственными нейронными сетями (ИНС) каскадной структуры сопряжена с определёнными сложностями. В частности, одной из основных проблем является отсутствие специализированного программного обеспечения, которое позволило бы задавать структуру ИНС типа DRB NN для пользователей, не обладающих навыками профессионального программирования.

Например, известный нейросимулятор профессора Леонида Нахимовича Ясницкого и его ученика Фёдора Михайловича Черепанова [182] позволяет пользователю задавать число скрытых слоев ИНС и определять число нейронов на каждом слое отдельно, а также выбирать вид активационных функций для каждого нейрона отдельно. Кроме того, преимуществами нейропакета «Нейросимулятор 5.0» является простота использования пользовательского интерфейса в сочетании с обширным набором функций. Этот пакет позволяет проводить нейросетевое моделирование с выбором архитектуры нейронной сети (Рисунок 3.6) и алгоритма обучения без необходимости знаний программирования или специализированных математических пакетов для обработки информации, таких как MATLAB, Maple, STATISTICA, и др. [157].



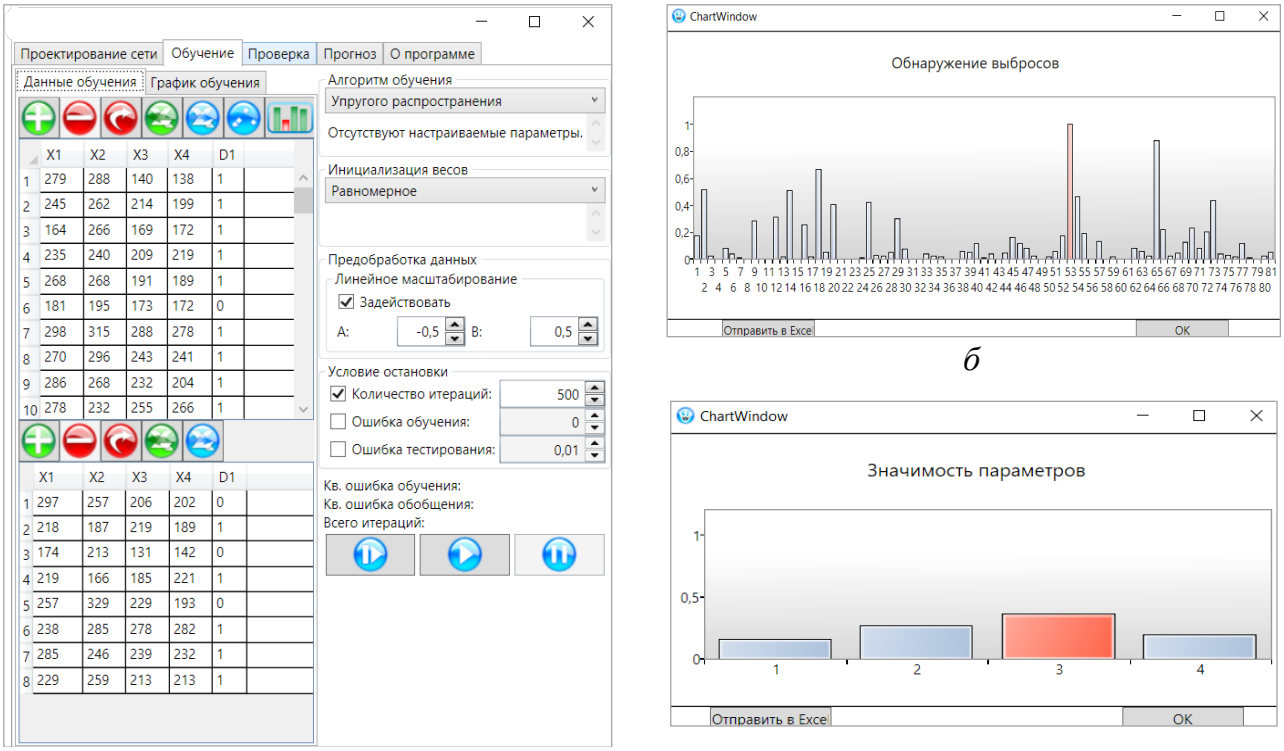
а *б*

Рисунок 3.6 – Пример графического представления выбора структуры полносвязных ИНС, реализуемых в «Нейросимулятор 5.0», с 10 входными сигналами и 1 выходным сигналом: *а* – ИНС с одним скрытым слоем и 5 нейронами на скрытом слое; *б* – ИНС с двумя скрытыми слоями и 5 нейронами на первом и 10 нейронами на втором скрытых слоях

На Рисунке 3.6. представлен пример выбора структуры полносвязных ИНС. Для ИНС с одним скрытым слоем и 5 нейронами на скрытом слое (см. Рисунок 3.6, *а*) общее число нейронов равняется 16, число связей определено на уровне 55 (+ 6 смещений), для ИНС с двумя скрытыми слоями и 5 нейронами на первом и 10 нейронами на втором скрытых слоях общее число нейронов на всех слоях

равняется 25, число связей определено на уровне 110 (+ 16 смещений) (см. Рисунок 3.6, *б*) [157].

Наряду со стандартными функциями для программ подобного рода в нейросимуляторе реализованы авторские наработки в области оптимизации и обучения нейронных сетей (Рисунок 3.7, *а*), а именно такие функции, как: обнаружение выбросов в исходных данных (Рисунок 3.7, *б*) и вычисление относительной значимости входных параметров нейросетевым энтропийным методом (Рисунок 3.7, *в*); функциональная предобработка множества и обучение с нелинейным масштабированием ошибки [182].



а *б* *в*

Рисунок 3.7 – Экранные формы оптимизации и обучения нейронных сетей, обнаружение выбросов в исходных данных и вычисление относительной значимости входных параметров в нейропакете «Нейросимулятор 5.0»

Однако «Нейросимулятор 5.0» поддерживает возможность создания только полносвязных ИНС, не обеспечивая неполносвязность во всех узлах ИНС, что является важным критерием при использовании ИНС класса Decisions' Root-based Neural Network. Решением могло бы стать принудительное присвоение нулевых

значений синаптическим коэффициентам, соответствующим весам рёбер сети, которые не попали в неполносвязную ИНС, но текущая версия нейросимулятора данной функции не поддерживает [152].

С другой стороны, как было отмечено ранее, программный комплекс [178], реализующий работу неполносвязных нейронных сетей, созданный также в Пермском государственном национальном исследовательском университете Андреем Леонидовичем Гусевым и его аспирантом Александром Андреевичем Окуневым, позволяет обучать неполносвязные ИНС. Однако в указанной программе есть возможность создать неполносвязную сеть только на одном слое. Если же неполносвязность присутствует на нескольких слоях, то данная программа не применима. Именно такой случай возникает при построении структуры неполносвязной ИНС по корню принятия решений. Число скрытых слоев в таких ИНС определено по формуле $2n–1$, где n – число переменных. Действительно, при больших значениях n получается довольно много скрытых слоев, что теоретически должно сказываться на эффективности обучения [152].

3.4. Программный модуль проектирования и обучения нейронных сетей, основанных на корнях принятия решений

В работе с неполносвязными ИНС класса Decisions’ Root-based Neural Network важным критерием является необходимость в обеспечении неполносвязности во всех узлах ИНС. Как было отмечено, нет общедоступных программных продуктов с низким уровнем требований к навыкам программирования, позволяющим пользователю задать сложную структуру ИНС.

Для решения этой проблемы требуются удобные инструменты и интерфейсы, которые могли бы позволить пользователям без специальной подготовки программирования эффективно работать с неполносвязными ИНС каскадной структуры. Такие инструменты должны быть интуитивно понятными и предоставлять возможность легко настраивать параметры и конфигурации нейронных сетей и анализировать их работу.

Таким образом, в рамках текущей диссертационной работы был разработан программный модуль [159] проектирования и обучения нейронных сетей, основанных на корнях принятия решений класса Decisions' Root-based Neural Network. Данный программный модуль является частью программного обеспечения многопользовательской виртуальной среды интеллектуального анализа данных типа SaaS, разработанной ООО «Пермский центр поддержки принятия решений» в рамках гранта Фонда содействия инновациям.

Программа предназначена для проектирования и обучения неполносвязных нейронных сетей с каскадной структурой, которая полностью определяется корнем принятия решений дискретной переключательной функцией, представленной в виде суперпозиции функций двух переменных, каждая из которых записывается в виде бинарной матрицы и располагается в определенном узле полного бинарного графа с поименными листьями (дерева критериев).



```
{
    "model": {
      "left": {
        "arguments": ["X1", "X2"],
        "conversationMatrix": [[1,1,1], [2,2,3], [3,4,4]],
        "activation": "PReLU"
      },
      "right": {
        "arguments": ["X3", "X4"],
        "conversationMatrix": [[1,1,2], [2,3,3], [4,4,4]],
        "activation": "PReLU"
      },
      "result":{
        "arguments": ["left", "right"],
        "conversationMatrix": [[1,1,1,1], [1,1,2,2], [2,2,3,3], [3,3,3,3]],
        "activation": "PReLU"
      }
    },
    "compile": {
        "useDefaultWeights": {
            "zero": 0.01,
            "one": 0.99
        },
        "metrics": ["r2", "accuracy", "sensitivity", "specificity", "precision", "AUC"],
        "optimizer": "Adam"
    },
    "outputLayerActivation": "PReLU",
    "bins": {
    },
    "kFoldMetrics": ["r2", "neg_mean_squared_error", "accuracy"]
}
```

Рисунок 3.8 – Фрагмент записи корня принятия решений
в формате JSON в модуле проектирования и обучения
нейронных сетей класса Decision's-Root-based Neural Network

В программе корень принятия решения записывается как последовательность переменных и последовательность строк, образующих матрицы (Рисунок 3.8). Программа может использоваться в открытой библиотеке Keras, включающей в себя

такие алгоритмы обучения, как SGD, RMSprop, Adagrad, Adadelta, Adam, Nadam, Adamax и др., а также поддерживает нейросетевую библиотеку TensorFlow для глубокого обучения нейронных сетей.

Функциональные возможности программы выражаются в проектировании и обучении неполносвязных нейронных сетей, основанных на корнях принятия решений (Decisions Root-based Neural Networks), настройке обучения нейронной сети, выбора метода обучения, активационных функций, метрик для оценки качества обученной нейронной сети и т.п. (Рисунок 3.9) [159].

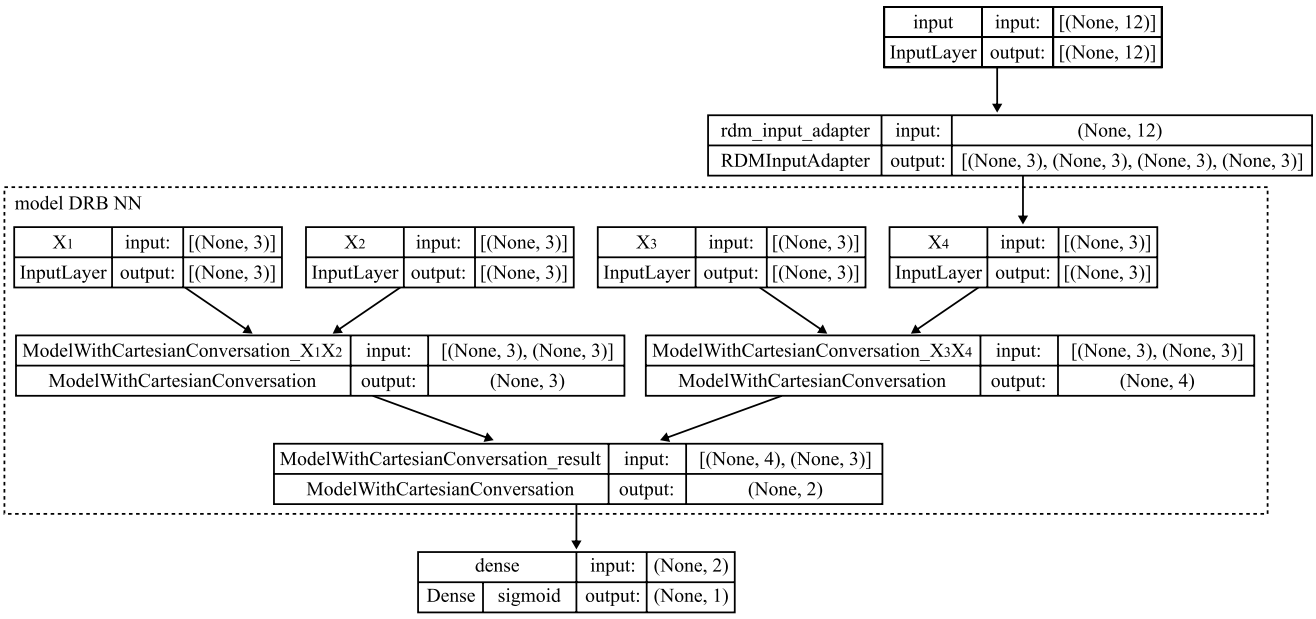


Рисунок 3.9 – Структура неполносвязной нейронной сети класса Decisions' Root-based Neural Network в модуле проектирования и обучения нейронных сетей, основанных на корнях принятия решений [159] (на примере корня принятия решений – см. Рисунок 2.6)

На одном из семинаров в ходе обсуждения полученных результатов в рамках диссертационного исследования была получена рекомендация к применению перекрёстной проверки (скользящий контроль или кросс-валидация, от англ. Cross-Validation) как одного из наиболее широко используемых методов повторной выборки данных для оценки обобщающей способности прогностических моделей.

Перекрёстная проверка похожа на метод повторной случайной выборки, но выборка осуществляется таким образом, чтобы тестовые наборы не пересекались.

При *k*-кратной перекрестной проверке обучающее множество разбивается на *k* разделяющихся подмножеств (Split) примерно одинакового размера. Здесь «Split» означает количество результирующих подмножеств. Это разбиение выполняется путём случайной выборки случаев из обучающего множества без замены. Модель обучается на *k*-1 подмножествах, которые вместе представляют собой обучающее множество. Затем модель применяется к оставшемуся подмножеству, которое обозначается как проверочное множество, и измеряется эффективность обучения. Эта процедура повторяется до тех пор, пока каждое из *k* подмножеств не станет валидационным [183].

На Рисунке 3.10 показан процесс перекрёстной проверки для $k = 5$, т.е. для 5-кратной кросс-валидации. В Split 1 первое подмножество служит валидационным набором Test set, а остальные пять подмножеств служат обучающим множеством Train set. В Split 2 второе подмножество является валидационным множеством, а остальные подмножества – обучающим множеством, и так далее.

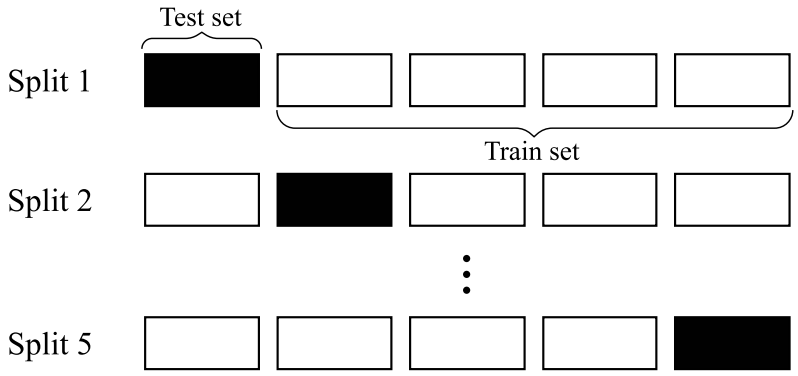


Рисунок 3.10 – Пример 5-кратной кросс-валидации данных с разбиением случайным образом на пять подмножеств (Split), где модель обучается на обучающем наборе (Tran set), а затем применяется к валидному набору (Test set)

Среди различных стратегий повторной выборки данных одной из самых простых является метод одиночного удержания, который случайным образом выбирает некоторые случаи из обучающего множества для тестового множества, а оставшиеся случаи составляют обучающее множество. Часто тестовое множество содержит от 10 до 30 % имеющихся случаев, а обучающее множество содержит от 70 до 90 % случаев [183, с. 3].

Поэтому среди основных функций для нейросетевого моделирования, в программном модуле проектирования и обучения нейронных сетей, основанных на корнях принятия решений [159], реализован алгоритм кросс-валидации.

Разработанный программный модуль проектирования и обучения неполносвязных искусственных нейронных сетей может выступать как самостоятельным программным продуктом, поскольку она реализована в библиотеке Keras с поддержкой TensorFlow, так и выступать частью программного обеспечения многопользовательский виртуальной среды интеллектуального анализа данных, алгоритм работы которой представлен на Рисунке 3.11.

Предлагаемая информационная система состоит из 5 программных модулей: обработка гетерогенных наборов данных в виде дискретных (категориальных) данных (программный модуль 1); идентификация корня принятия решений на основе полных обучающих наборов (программный модуль 2); определение неизвестных параметров в обучающих наборах для идентификации корня принятия решений (программный модуль 3); построение структуры неполносвязной искусственной нейронной сети, основанной на корне принятия решений (программный модуль 4); обучение неполносвязной искусственной нейронной сети, основанной на корне принятия решений (программный модуль 5).

Отличительной особенностью предлагаемой информационной системы от известных является то, что неполносвязность нейронных сетей обеспечивается во всех узлах её структуры, что позволяет проектировать и обучать искусственные нейронные сети, основанные на корнях принятия решений. Предлагаемая информационная система дает возможность проводить многократные вычислительные эксперименты по анализу устойчивости результатов обучения неполносвязных нейронных сетей, включая кросс-валидацию в условиях ограниченных наборов данных. Предусмотрен следующий функционал пользовательского приложения анализа данных: анализ данных, построение графиков, формирование отчета, просмотр отчета в PDF-формате, сохранение и отправка по электронной почте экспресс-отчета.

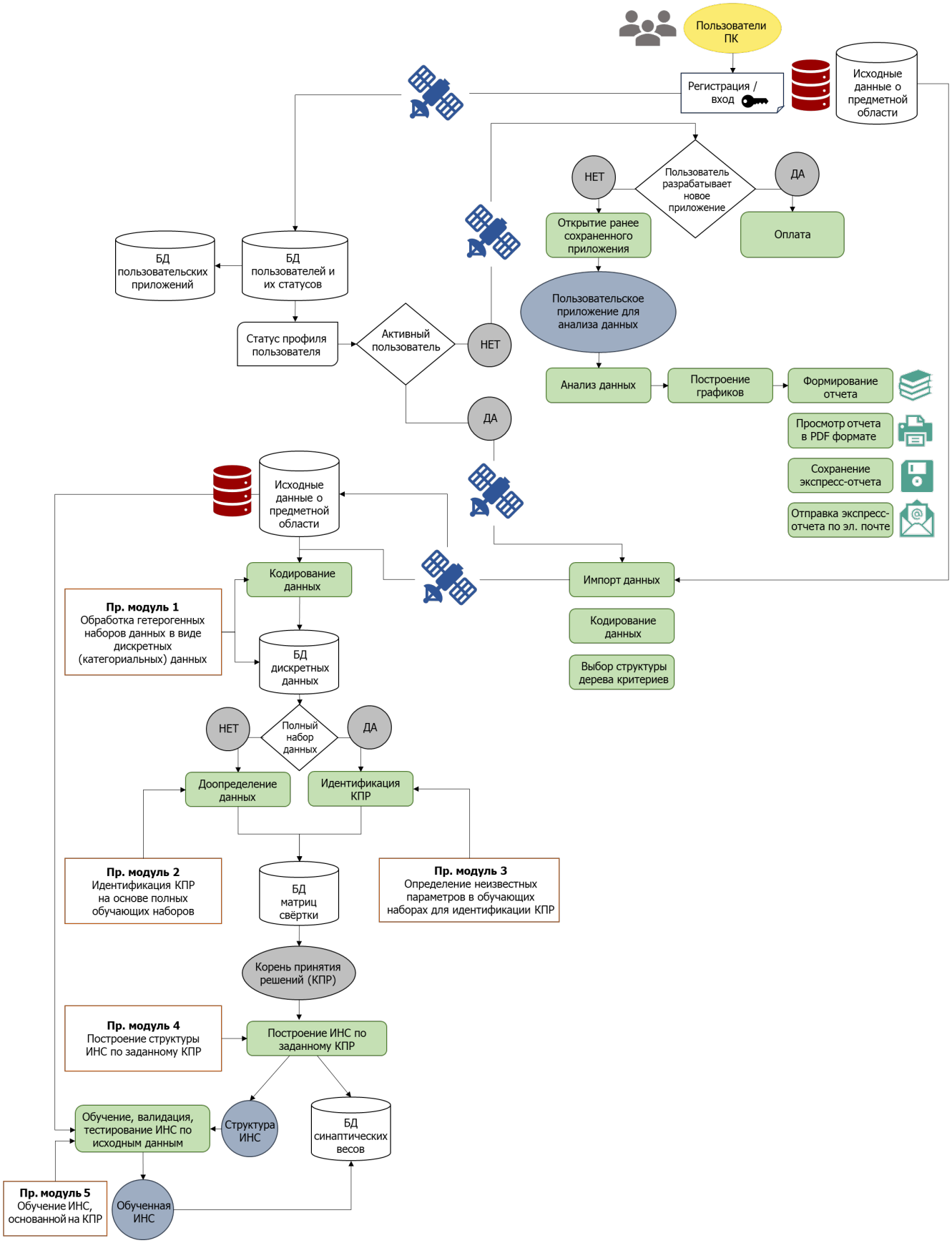


Рисунок 3.11 – Алгоритмическое обеспечение информационной системы проектирования и обучения искусственных нейронных сетей, основанных на корнях принятия решений

3.5. Выводы по главе 3

В главе рассмотрены информационные системы обработки информации, которые соответствуют предложенным правилам и методам предварительной обработки данных и выбора архитектуры нейронной сети класса DRB NN.

Описан разработанный программный модуль проектирования и обучения нейронных сетей, основанных на корнях принятия решений класса DRB NN. Среди основных функций для нейросетевого моделирования разработанный модуль имеет функциональную возможность применения перекрёстной проверки как одного из наиболее широко используемых методов повторной выборки данных для оценки обобщающей способности прогностических моделей.

Наглядно изображён алгоритм работы пяти функциональных программных модулей (см. Рисунок 3.11), вместе образующих единую информационную систему проектирования и обучения искусственных нейронных сетей, основанных на корнях принятия решений. Программные модули [159; 161] (Приложение 7, 9) информационной системы реализованы в открытой библиотеке Keras, включающей в себя такие алгоритмы обучения, как SGD, RMSprop, Adagrad, Adadelta, Adam, Nadam, Adamax и др., а также поддерживают нейросетевую библиотеку TensorFlow для глубокого обучения нейронных сетей. Информационная система позволяет осуществлять многократные вычислительные эксперименты по анализу устойчивости результатов обучения неполносвязных нейронных сетей, включая кросс-валидацию в условиях ограниченных данных.

ГЛАВА 4. ОЦЕНКА КАЧЕСТВА НЕЙРОННЫХ СЕТЕЙ, ОСНОВАННЫХ НА КОРНЯХ ПРИНЯТИЯ РЕШЕНИЙ, НА ПРИМЕРЕ ДИАГНОСТИКИ БОЛЕЗНИ АЛЬЦГЕЙМЕРА

Точность получаемых моделей при анализе данных напрямую зависит от качества и объёма исходного набора данных. И здесь не всегда работает правило «чем больше, тем лучше», так как при отсутствии репрезентативности как одного из ключевых свойств данных, используемых для построения аналитических моделей и показывающая достаточность информации в анализируемых данных, невозможно получить достоверную модель, отвечающую установленному уровню качества и точности.

Однако не во всех задачах есть возможность сформировать достаточный объём исходных данных. Существуют две предметные области, в которых кроме репрезентативности данных необходимо учитывать их ограниченность: промышленность и медицина. Требуются новые методы и алгоритмы к анализу данных в условиях их ограниченного исходного набора.

В данной главе рассмотрим практическое применение и апробацию методов и алгоритмов предварительной обработки данных с выбором структуры нейронной сети класса Decision’s-Root-based Neural Network в ограниченных случаях исходного набора данных (см. Рисунок 3.9).

4.1. Предварительная обработка данных с помощью АСК-анализа

Предварительная обработка данных методами АСК-анализа с применением универсальной когнитивной аналитической системы «Эйдос-Xpro» [61] проводилась на медицинских данных для выявления наиболее значимых показателей магнитной восприимчивости вен головного мозга (англ. Magnetic Susceptibility Value, MSV), которые могут являться маркерами болезни Альцгеймера на ранней стадии клинического проявления. Болезнь Альцгеймера – одно из наиболее часто встречающихся заболеваний головного мозга в последние

десятилетия, характеризующееся прогрессирующим ухудшением когнитивных функций, включая память, мышление и навыки, становится все более распространенным в наше время. Одним из ключевых маркеров болезни Альцгеймера являются амилоидные бляшки, которые представляют собой скопления неправильно свернутых белковых молекул в нейронах мозга. Эти бляшки могут быть обнаружены с помощью позитронно-эмиссионной томографии (ПЭТ) или магнитно-резонансной томографии (МРТ), позволяющих визуализировать структуру мозга и выявить аномальные области.

В качестве исходных данных взяты представленные в открытом доступе результаты исследования головного мозга, которые использовались в работе [184] для установления корреляционной связи между значениями магнитной восприимчивости (англ. Magnetic Susceptibility Value, MSV) вен головного мозга.

В общей сложности данные были собраны на 81 пациента, из которых у 59 человек была диагностирована болезнь Альцгеймера на основе результатов мини-теста психического состояния (MMSE), Монреальского когнитивного теста (MoCA), задачи рисования часов (CDT) и шкалы активности повседневной жизни (ADL), в то время как у 22 пожилых людей когнитивные способности оставались в норме.

Исходные данные представляли собой томографические снимки пациентов. На основе томографических снимков были количественно определены значения магнитной восприимчивости следующих вен головного мозга: базальных вен левого и правого полушария (L_BV и R_BV), внутренних мозговых вен левого и правого полушария (L_ICV, R_ICV), вен таламуса левого и правого полушария головного мозга (L_TV, R_TV), септальных вен левого и правого полушария (L_SV, R_SV), вен зубчатого ядра левого и правого полушария головного мозга (L_DNV, R_DNV) [151]. Упомянутые вены выделены на томографических снимках ниже (Рисунок 4.1).

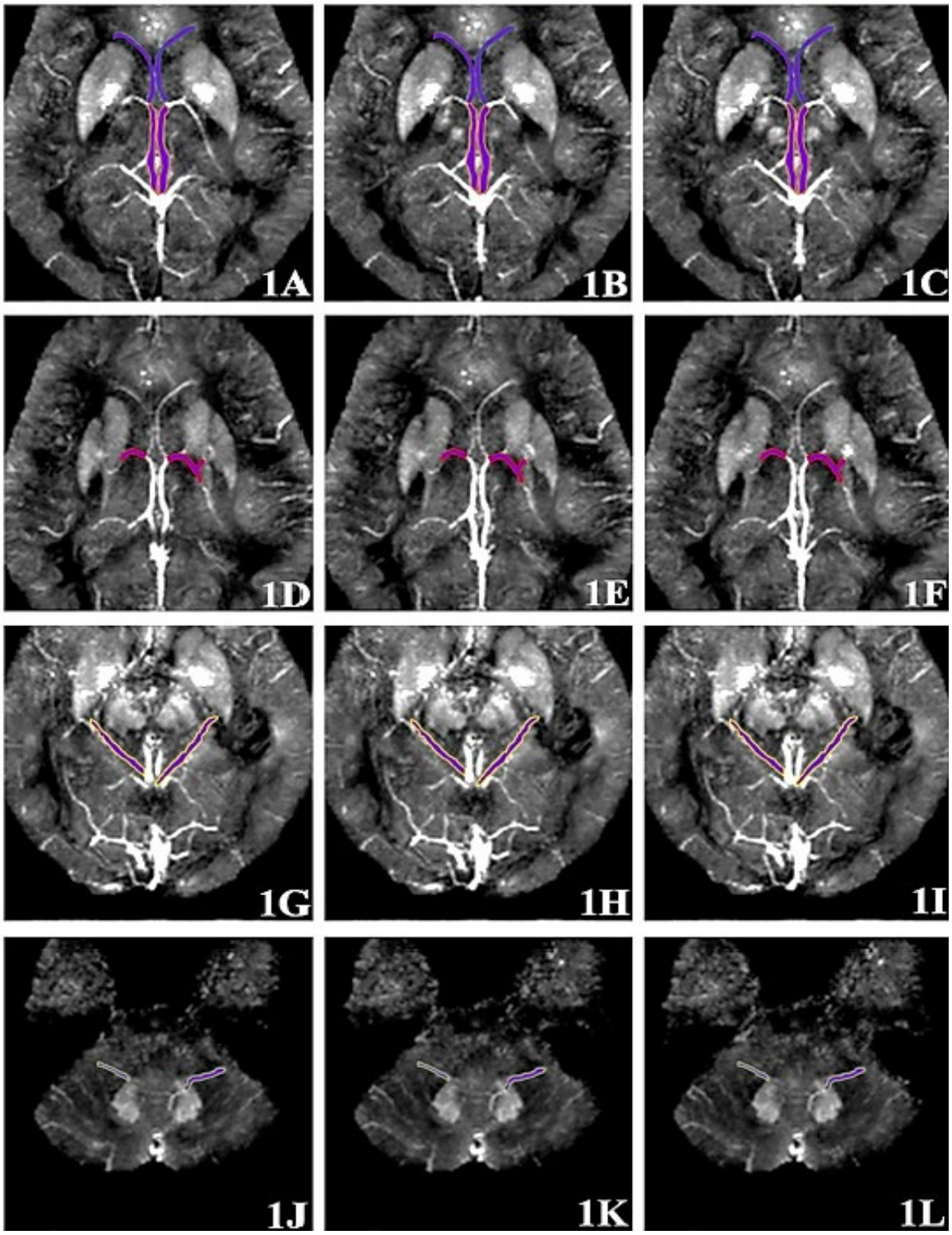


Рисунок 4.1 – Выбранные в работе [184, с. 3] области исследования вен головного мозга: 1A–1C – три непрерывных среза, на которых показаны септальные вены и внутренние мозговые вены; 1D–1F – три непрерывных среза, показывающие вены таламуса; 1G–1I – три непрерывных среза, показывающих базальные вены; 1J–1L – три непрерывных среза, показывающие вены зубчатого ядра

В настоящем исследовании использовались исключительно числовые данные, приведённые в приложении к статье [184]. Фрагмент исходных данных

представлен в Таблице 4.1. Полный набор исходных данных представлен в работе в Приложении 1.

Таблица 4.1 – Фрагмент исходных данных со значениями магнитной восприимчивости вен головного мозга

NO.	Group	L_BV	R_BV	L_ICV	R_ICV	L_TV	R_TV	L_SV	R_SV	L_DNV	R_DNV
sub001	AD	279	288	255	263	140	138	131	131	165	185
sub002	AD	274	247	223	243	239	262	190	222	204	102
sub003	AD	259	333	236	243	172	159	135	145	152	153
sub004	AD	213	252	298	301	174	154	154	135	140	127
sub005	AD	198	298	243	236	213	213	261	250	166	167
						. . .					
sub079	CON	279	288	179	189	141	153	245	214	177	149
sub080	CON	249	232	221	232	154	142	165	151	161	143
sub081	CON	300	259	295	299	216	197	163	131	162	129

В качестве признаков выступали диапазоны значений показателей магнитной восприимчивости описанных выше вен L_BV, R_BV, L_ICV, R_ICV, L_TV, R_TV, L_SV, R_SV, L_DNV, R_DNV, а в качестве классов использовались состояние пациента – больной болезнью Альцгеймера (подтвержденная группа пациентов с болезнью Альцгеймера, от англ. Alzheimer’s disease, далее AD) и не больной болезнью Альцгеймера (не подтвержденная (контрольная) группа пациентов с болезнью Альцгеймера, от англ. control, далее CON) [151].

С помощью методов АСК-анализа проведена серия экспериментов с различными способами разбиения области наблюдений по два, три и четыре равных интервала с разным числом наблюдений, а также по два, три и четыре разных интервалов с равным числом наблюдений.

В каждом из таких экспериментов в системе «Эйдос-Xpro» с помощью 2 интегральных критериев и 3 статистических и 7 информационных моделей получено по 20 значений принадлежности исследуемых субъектов группе AD и группе CON (Таблица 4.2). Наиболее достоверной моделью по F-мере Ван Ризбергена оказалась информационная модель *INF3* при разбиении области значений на три равных интервала с разным числом наблюдений (Таблица 4.3). Это также не противоречит тому, что при выборе разбиения на два интервала получается классическая задача нахождения булевой функции, при выборе четырёх

интервалов формируется 80 признаков, а исходный набор данных, который необходимо разделить на обучающее и тестовое множества, содержит всего 81 пример. Остальные результаты экспериментов статистических моделей представлены в сводных таблицах в Приложении 2. Таблицы принятых интервалов значений магнитной восприимчивости для наилучших моделей при разной спецификации разбиения исходных данных представлены в Приложении 3.

Таблица 4.2 – Сводная таблица достоверности статистических моделей с двумя интегральными критериями «Семантический резонанс знаний» (IC_SRK) и «Сумма знаний» (IC_SK) при разбиении на три равных интервала с разным числом наблюдений

№	Модель	Интеграль-ный критерий (IC)	N	TP	FN	FP	TN	Точность модели	Полнота модели	F	L1	L2
1	ABS	IC_1	81	76	5	64	17	0,543	0,938	0,688	0,752	0,668
1	ABS	IC_2	81	81	0	81	0	0,500	10	0,667	0,772	0,772
2	PRC1	IC_3	81	76	5	64	17	0,543	0,938	0,688	0,752	0,668
2	PRC1	IC_4	81	81	0	81	0	0,500	10	0,667	0,689	0,689
3	PRC2	IC_3	81	76	5	64	17	0,543	0,938	0,688	0,752	0,668
3	PRC2	IC_4	81	81	0	81	0	0,500	10	0,667	0,689	0,689
4	INF1	IC_SRK	81	54	27	27	54	0,667	0,667	0,667	0,727	0,571
4	INF1	IC_SK	81	51	30	22	59	0,699	0,630	0,662	0,700	0,540
5	INF2	IC_SRK	81	54	27	27	54	0,667	0,667	0,667	0,727	0,571
5	INF2	IC_SK	81	51	30	22	59	0,699	0,630	0,662	0,700	0,540
6	INF3	IC_SRK	81	60	21	21	60	0,741	0,741	0,741	0,791	0,569
6	INF3	IC_SK	81	60	21	21	60	0,741	0,741	0,741	0,791	0,569
7	INF4	IC_SRK	81	49	32	21	60	0,700	0,605	0,649	0,702	0,565
7	INF4	IC_SK	81	53	28	30	51	0,639	0,654	0,646	0,730	0,600
8	INF5	IC_SRK	81	49	32	21	60	0,700	0,605	0,649	0,702	0,565
8	INF5	IC_SK	81	53	28	30	51	0,639	0,654	0,646	0,730	0,600
9	INF6	IC_SRK	81	59	22	20	61	0,747	0,728	0,738	0,787	0,568
9	INF6	IC_SK	81	60	21	22	59	0,732	0,741	0,736	0,749	0,518
10	INF7	IC_SRK	81	59	22	20	61	0,747	0,728	0,738	0,787	0,568
10	INF7	IC_SK	81	60	21	22	59	0,732	0,741	0,736	0,749	0,518

В Таблице 4.2 и в таблицах Приложения 2 бирюзовым цветом обозначены значимые модели по критерию F-мера Ван Ризбергена; зеленым цветом обозначены значимые модели по критерию L1 – первая мера профессора Е.В. Луценко; желтым цветом обозначены значимые модели по критерию L2 – вторая мера профессора Е.В. Луценко; серым цветом выделены наиболее

значимые модели, которые либо совпадают с одним из трех критериев меры (двойное выделение цветами), либо значимы только статистические модели, выбранные экспертным путём, так как в текущем исследовании интерес представляют модели только системно-когнитивной базы знаний.

Таблица 4.3 – Принятые интервалы значений магнитной восприимчивости вен головного мозга для модели *INF3* при разбиении на три равных интервала с разным числом наблюдений

Вены головного мозга	Обозначение	Наблюдаемая область значений MSV	1-й интервал	2-й интервал	3-й интервал
Базальная вена левого полушария	L_BV	{153,0; 324,0}	{153,0; 210,0}	{210,0; 267,0}	{267,0; 324,0}
Внутренняя мозговая вена левого полушария	L_ICV	{179,0; 394,0}	{179,0; 250,7}	{250,7; 322,3}	{322,3; 394,0}
Вена таламуса левого полушария	L_TV	{131,0; 288,0}	{131,0; 183,3}	{183,3; 235,7}	{235,7; 288,0}
Септальная вена левого полушария	L_SV	{83,0; 310,0}	{83,0; 158,7}	{158,7; 234,3}	{234,3; 310,0}
Вена зубчатого ядра левого полушария	L_DNV	{94,0; 244,0}	{94,0; 144,0}	{144,0; 194,0}	{194,0; 244,0}
Базальная вена правого полушария	R_BV	{164,0; 357,0}	{164,0; 228,3}	{228,3; 292,7}	{292,7; 357,0}
Внутренняя мозговая вена правого полушария	R_ICV	{163,0; 411,0}	{163,0; 245,7}	{245,7; 328,3}	{328,3; 411,0}
Вена таламуса правого полушария	R_TV	{109,0; 286,0}	{109,0; 168,0}	{168,0; 227,0}	{227,0; 286,0}
Септальная вена правого полушария	R_SV	{73,0; 287,0}	{73,0; 144,3}	{144,3; 215,7}	{215,7; 287,0}
Вена зубчатого ядра правого полушария	R_DNV	{81,0; 269,0}	{81,0; 143,7}	{143,7; 206,3}	{206,3; 269,0}

В Таблице 4.2 и в таблицах Приложения 2 используются следующие обозначения: IC_1 – корреляция абсолютных частот; IC_2 – сумма абсолютных частот по признакам объекта; IC_3 – корреляция условно-относительных частот; IC_4 – сумма условно-относительных частот по признакам объекта; IC_SRK – «Семантический резонанс знаний»; IC_SK – «Сумма знаний»; N – количество всех логических объектов выборки; TP – число истинно-положительных решений; FN – число ложно-отрицательных решений, FP – число ложно-положительных решений; TN – число истинно-отрицательных решений; F – мера Ван Ризбергена; L1 – мера профессора Е.В. Луценко; L2 – мера профессора Е.В. Луценко.

В результате применения АСК-анализа были определены диапазоны значений MSV (Таблица 4.3), обладающие наибольшей информационной значимостью для отнесения пациента к классу AD (Рисунок 4.2 – «способствующие»), а также признаки, которые в наибольшей степени не характеризуют этот класс (Рисунок 4.2 – «препятствующие»). Поскольку в исследуемом наборе данных пациенты относятся всего к двум классам (AD и CON), то незначимость для класса AD, означает значимость для класса CON соответственно [151].

Таким образом, была установлена взаимосвязь между значениями MSV и принадлежностью пациента к классу AD или CON. Текущие результаты демонстрируют перспективность использования АСК-анализа и анализа MSV в целях улучшения диагностики и прогнозирования развития заболеваний класса AD.

Результаты экспериментов, в ходе которых были определены диапазоны значений MSV для своей наилучшей модели, обладающие наибольшей информационной значимостью для отнесения пациента к классу AD или классу CON, представлены в Приложении 4.

Из Рисунка 4.2 видно, что высокие значения MSV (3-и диапазоны) в венах таламуса L_TV, R_TV относятся к 3 наиболее способствующим факторам, а низкие значения MSV (1-е диапазоны) в тех же венах относятся уже к препятствующим факторам. Другими словами, L_TV и R_TV выступают в качестве индикаторов, или сигналов, указывающих на предрасположенность или наличие болезни Альцгеймера.

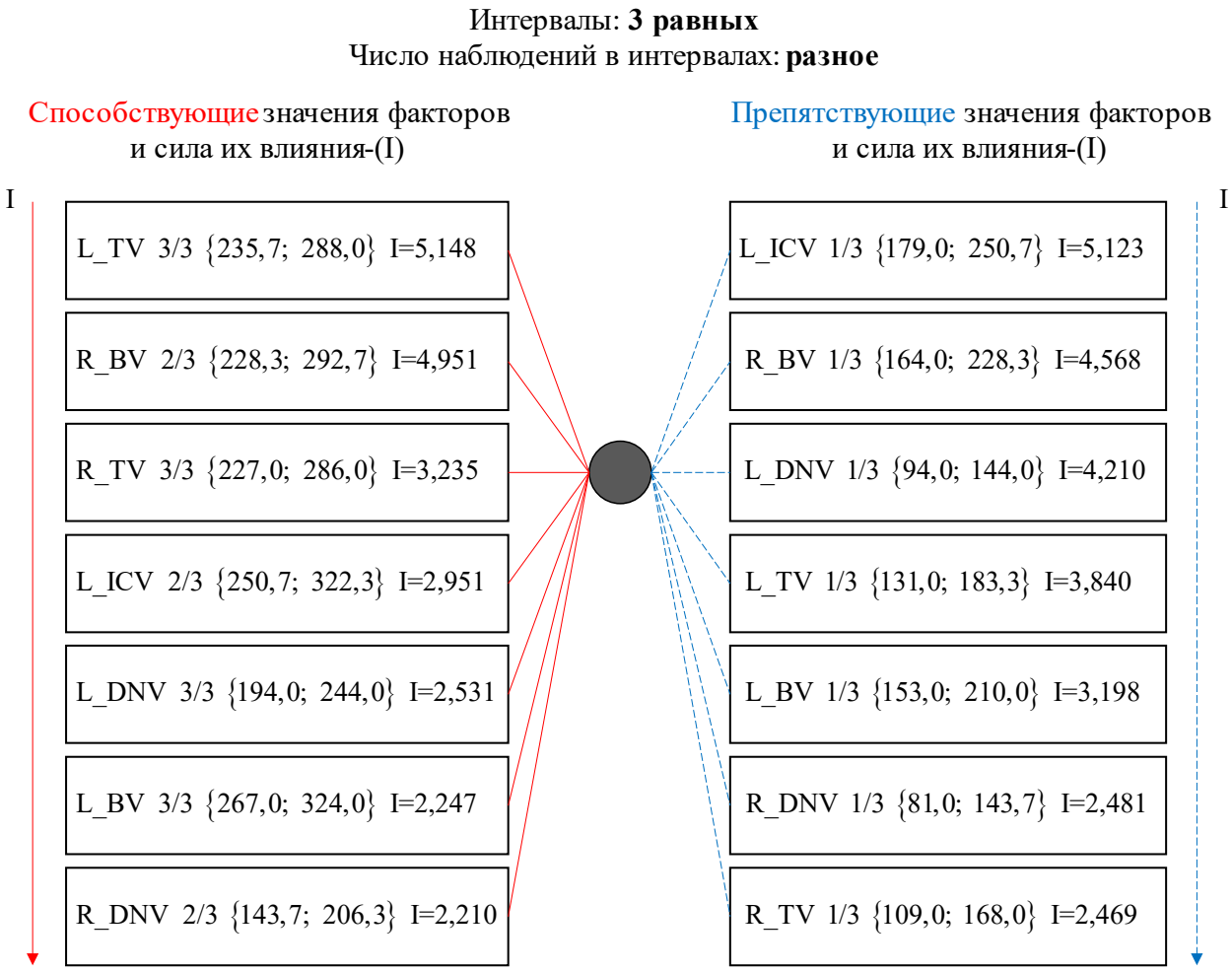


Рисунок 4.2 – Наиболее значимые факторы модели *INF3*, влияющие на возможную патологию головного мозга в виде болезни Альцгеймера (после названия переменной, соответствующей определенной вене, указывается порядковый номер интервала из трех значений, а в фигурных скобках непосредственно указаны значения интервала MSV) [151]

Результаты анализа с использованием метода АСК согласуются с данными [184], демонстрирующими, что значения магнитной восприимчивости вен таламуса левого и правого полушария головного мозга (L_TV, R_TV) и клинико-лабораторные параметры/когнитивные оценки у пациентов с болезнью Альцгеймера коррелируют с показателями левой и правой вен таламуса, а также с тем фактом, что для MSV левой и правой вен таламуса при сопоставлении группы больных болезнью Альцгеймера и контрольной группы $p < 0{,}05$.

В дополнение к этому в настоящем исследовании значимыми факторами были выбраны базальные вены R_BV и L_BV, поскольку высокие и средние значения MSV для обеих (3-й диапазон значений MSV для левой базальной вены

и 2-й диапазон значений MSV для правой соответственно) вошли в семь способствующих факторов для отнесения пациента к классу AD (см. Рисунок 4.2), в то время как низкие значения MSV (1-й диапазоны значений MSV) как для левой, так и для правой базальных вен относятся к факторам, препятствующим такому отнесению (см. Рисунок 4.2), что может быть интерпретировано как переключательный триггер.

Значимость базальных вен, в свою очередь, также имеет медико-биологическое объяснение – изменение кровотока по базальным венам связано с внутричерепной гипертензией, влияющей на перфузию, т.е. снабжение тканей мозга необходимыми веществами [185; 186].

Таким образом, в результате применения методов АСК-анализа из числа всех имеющихся показателей (L_BV, R_BV, L_ICV, R_ICV, L_TV, R_TV, L_SV, R_SV, L_DNV, R_DNV) следует считать наиболее значимыми четыре показателя (L_BV, R_BV, L_TV, R_TV), что не противоречит сведениям, приведенным в работах [184; 185]. Уменьшение числа переменных позволило решить одну из проблем при идентификации МКО (корней принятия решений) – выбор структуры графа G, представляющего собой полное бинарное дерево с поимёнными листьями, тем самым сократив количество возможных структур до 15.

4.2. Предварительная обработка данных с помощью деревьев решений

В результате обучения и тестирования деревьев решений на основе исходных данных о магнитной восприимчивости 10 вен головного мозга (см. Приложение 1) 81 пациента были получены 450 деревьев решений, где на каждый атрибут разбиения узлов дерева (информационная энтропия, коэффициент Джини, логарифмическая потеря) приходилось 150 экспериментальных вариантов.

На Рисунке 4.3 показана точность измерений на тестовых данных 150 деревьев, полученных в результате их обучения и тестирования с использованием теоретико-информационного критерия (информационная энтропия), так как применение информационной энтропии показывает лучшие результаты, чем другие атрибуты разбиения узлов дерева решений.

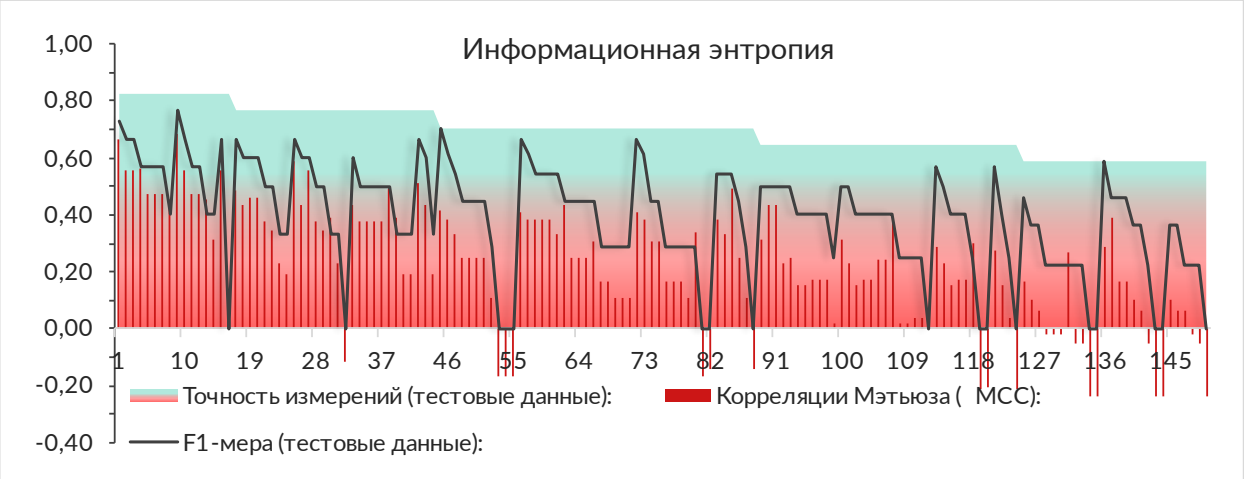


Рисунок 4.3 – Графическое представление качества обученных 150 деревьев решений на данных без их предобработки (10 вен головного мозга), где для разбиения узлов деревьев решений выбрана информационная энтропия

Для оценки качества деревьев решений были выбраны следующие метрики: точность измерений, F1-мера, средняя потеря Хэмминга, корреляция Мэтьюза (MCC). Показатели метрик оценки качества первых 10 наилучших деревьев решений среди 150 экспериментов на основе исходных данных о магнитной восприимчивости 10 вен головного мозга с применением теоретико-информационного критерия (информационной энтропии) в качестве критерия разбиения узлов представлены в Таблице 4.4.

Таблица 4.4 – Таблица результатов обучения и тестирования деревьев решений (первые 10 лучших вариантов)

№ п/п	Точность измерений (обучающие данные)	Точность измерений (тестовые данные)	F1-мера (обучающие данные)	F1-мера (тестовые данные)	Средняя потеря Хэмминга	Корреляции Мэтьюза (MCC)
1	0,984	0,824	0,870	0,727	0,176	0,663
2	0,984	0,824	0,880	0,667	0,176	0,555
3	0,984	0,824	0,880	0,667	0,176	0,555
4	0,984	0,824	0,889	0,571	0,176	0,566
5	0,984	0,824	0,889	0,571	0,176	0,471
6	0,984	0,824	0,889	0,571	0,176	0,471
7	0,984	0,824	0,889	0,571	0,176	0,471
8	0,984	0,824	0,897	0,400	0,176	0,451
9	0,969	0,824	0,857	0,769	0,176	0,685
10	0,969	0,824	0,880	0,667	0,176	0,555

Применяя вышеперечисленные метрики оценки качества, было определено наилучшее дерево решений (Рисунок 4.4). Полученное дерево может использоваться как система поддержки принятия врачебных решений или как «второе мнение» специалиста, когда врач, сравнивая полученные показатели MSV вен головного мозга пациента с диапазонами в узлах, начиная сверху вниз, в зависимости от выполненного условия (True, False) следует по дереву до конечного узла и сопоставляет с предварительным диагнозом по результатам клинического исследования пациента при диагностике болезни Альцгеймера.

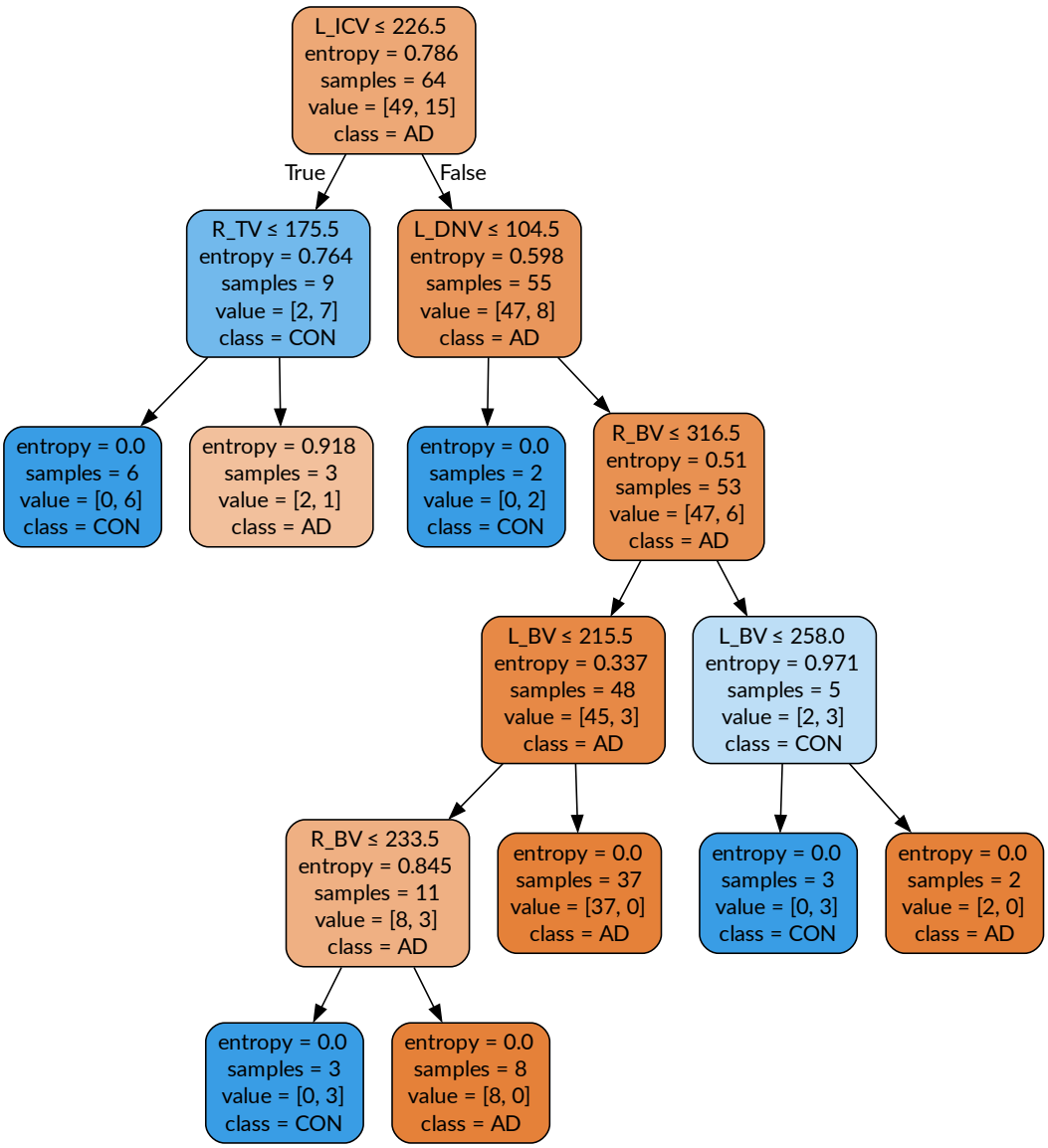


Рисунок 4.4 – Дерево решений, построенное на данных без их предобработки (10 вен головного мозга), критерий разбиения – информационная энтропия, где AD – группа пациентов с болезнью Альцгеймера, CON – контрольная группа пациентов, у которых нет проблем с когнитивными способностями

Стоит отметить, что базальные вены наиболее часто встречаются в структуре полученного дерева. Этот факт может послужить для дальнейшего исследования взаимосвязи магнитной восприимчивости вен головного мозга и болезни Альцгеймера в том, что базальные вены головного мозга могут являться ключевыми триггерами для принятия решения о полноценном обследовании пациента на наличие нейродегенеративных заболеваний. Эти результаты о значимости базальных вен коррелируют с данными, полученными с помощью использования методов АСК-анализа.

На правой ветке дерева решений (см. Рисунок 4.4) видно, что для диагностики болезни Альцгеймера при принятии врачебных решений в четырёх из пяти полученных правил специалист будет обращать внимание на базальные вены правого и левого полушария. В левой ветке дерева решений находится единственное правило, по которому диагноз пациента с когнитивными нарушениями ставится по результатам анализа вены таламуса правого полушария. Такие результаты, полученные с помощью дерева решения, являются дополнительным подтверждением выбора наиболее значимых вен для диагностики болезни Альцгеймера.

Далее был проведён ряд экспериментов обучения и тестирования деревьев решений на основе данных о магнитной восприимчивости четырех вен головного мозга (L_BV, R_BV, L_TV, R_TV), которые были выбраны как наиболее значимые в результате применения методов АСК-анализа. Были получены 450 деревьев решений, где на каждый атрибут разбиения узлов дерева (информационная энтропия, коэффициент Джини, логарифмическая потеря) приходилось 150 экспериментальных вариантов, как и в серии обучения и тестирования деревьев на 10 венах.

На Рисунке 4.5 показана точность измерений на тестовых данных 150 деревьев, полученных в результате их обучения и тестирования с использованием теоретико-информационного критерия (информационная энтропия), так как применение информационной энтропии показывает лучшие результаты, чем другие атрибуты разбиения узлов дерева решений.

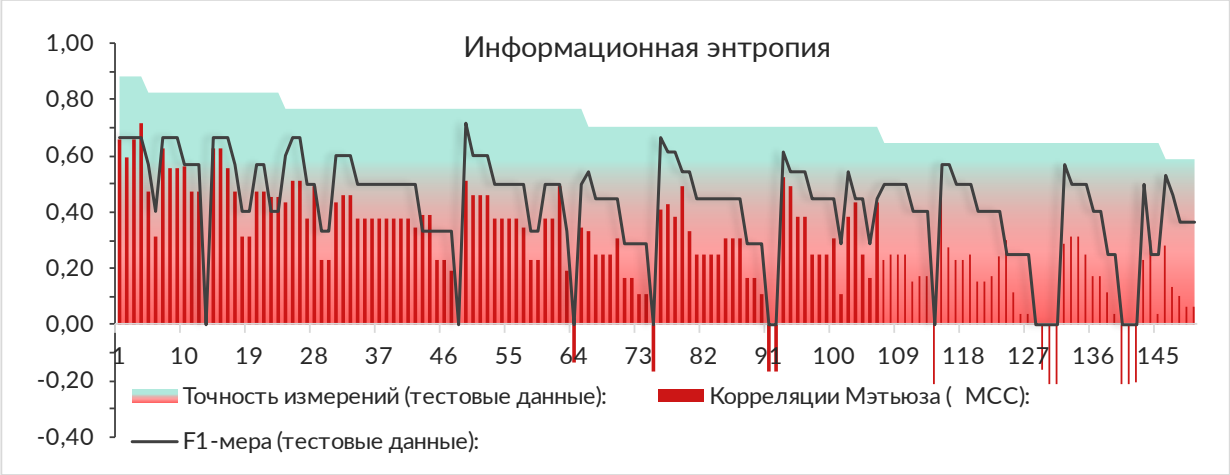


Рисунок 4.5 – Графическое представление качества обученных 150 деревьев решений на предобработанных данных (четыре вены головного мозга), где для разбиения узлов деревьев решений выбрана информационная энтропия

Показатели метрик оценки качества первых 10 наилучших деревьев решений на основе исходных данных о магнитной восприимчивости четырех вен головного мозга при использовании в качестве атрибута разбиения узлов деревьев теоретико-информационного критерия (информационная энтропия) представлены в Таблице 4.5.

Таблица 4.5 – Таблица результатов обучения и тестирования деревьев решений (первые 10 лучших вариантов)

№ п/п	Точность измерений (обучающие данные)	Точность измерений (тестовые данные)	F1-мера (обучающие данные)	F1-мера (тестовые данные)	Средняя потеря Хэмминга	Корреляции Мэтьюза (MCC)
1	0,969	0,882	0,929	0,667	0,118	0,658
2	0,953	0,882	0,929	0,667	0,118	0,595
3	0,938	0,882	0,929	0,667	0,118	0,658
4	0,922	0,882	0,917	0,667	0,118	0,717
5	1,000	0,824	0,889	0,571	0,176	0,471
6	1,000	0,824	0,897	0,400	0,176	0,310
7	0,984	0,824	0,880	0,667	0,176	0,627
8	0,984	0,824	0,880	0,667	0,176	0,555
9	0,984	0,824	0,880	0,667	0,176	0,555
10	0,984	0,824	0,889	0,571	0,176	0,566

Используя вышеперечисленные метрики оценки качества, было определено наилучшее дерево решений (Рисунок 4.6). В отличие от дерева решений, полученного на основе исходных данных о магнитной восприимчивости 10 вен

головного мозга (см. Рисунок 4.4), дерево решений, полученное на основе исходных данных о магнитной восприимчивости четырех вен головного мозга (Рисунок 4.6), имеет точность выше.

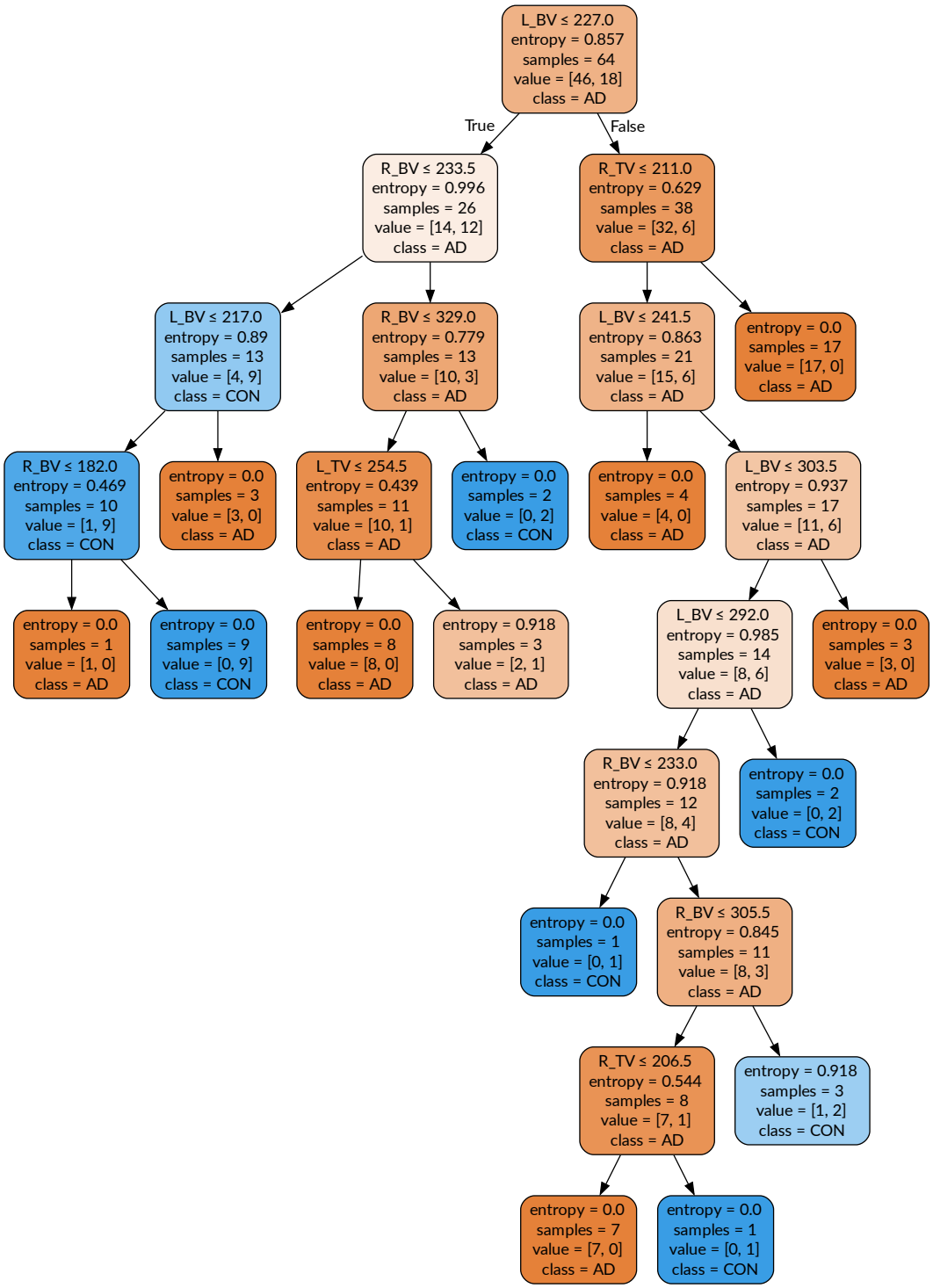


Рисунок 4.6 – Дерево решений, построенное на предобработанных данных (четыре вены головного мозга), критерий разбиения – информационная энтропия, где AD – группа пациентов с болезнью Альцгеймера, CON – контрольная группа пациентов, у которых нет проблем с когнитивными способностями

В ходе экспериментов при построении 900 деревьев решений с разными критериями разбиения узлов и разным количеством переменных было обнаружено, что точность и частота деревья решений в случае, когда они обучались и тестировались на основе исходных данных о магнитной восприимчивости четырёх вен головного мозга (L_BV, R_BV, L_TV, R_TV), были выше, чем в случае обучения и тестирования на 10 венах головного мозга (L_BV, R_BV, L_ICV, R_ICV, L_TV, R_TV, L_SV, R_SV, L_DNV, R_DNV). Это подтверждает значимость и необходимость предварительной обработки ограниченных данных для обнаружения количественной связи между исследуемыми показателями.

Результаты всех серий экспериментов обучения и тестирования деревьев решений с применением атрибутов разбиения узлов коэффициента Джини и логарифмической потери, на основе исходных данных о магнитной восприимчивости как в случае 10 вен головного мозга, так и в случае, когда исходные данные были предобработаны и выбраны четыре наиболее значимые вены, представлены в графическом и табличном виде в Приложении 5.

4.3. Нейросетевое моделирование в нейропакете «Нейросимулятор 5.0»

В исследовании [155] был изучен вопрос о влиянии предварительной обработки данных на эффективность обучения искусственных нейронных сетей при работе с ограниченными исходными наборами данных. Предварительная обработка включала в себя выделение наиболее значимых переменных и определение корня принятия решений, на основе которого выявляются структура и спецификация неполносвязной нейронной сети.

В ходе обсуждения полученных результатов на научных семинарах и конференциях была выдвинута гипотеза о возможности установления количественной взаимосвязи между изучаемыми параметрами без предварительной обработки данных. Это обстоятельство определило один из дополнительных этапов настоящего исследования – оценить качество ИНС, обученных на основе исходных данных всех 10 вен головного мозга, и установить связь между болезнью

Альцгеймера и величиной магнитной восприимчивости (MSV) вен головного мозга (на примере 81 пациента) при сохранении конфигурации скрытых слоев ИНС [156]. Для нейросетевого моделирования был выбран «Нейросимулятор 5.0» [182].

На сегодняшний день строгой теории выбора оптимального количества скрытых слоев и нейронов в скрытых слоях не существует. Поэтому выбор структур ИНС, как в работе [155], так и в текущем исследовании, не противоречит общепринятым подходам, в соответствии с которыми чаще всего используются персептроны с одним или двумя скрытыми слоями, при этом количество нейронов в скрытых слоях обычно варьируется в интервале от $N_x / 2$ до $3N_x$ [155, с. 137].

Исходный набор данных о 81 пациенте был разбит на тренировочное (65 примеров) и тестовое (16 примеров) множества в соотношении 80 : 20 %, что является приемлемым согласно [81, с. 143]. В качестве подбора и корректировки синаптических весов $w_{ij}(t)$ нейросетевой модели (цель обучения ИНС) был выбран алгоритм упругого распространения. За активационную функцию нейронов скрытых слоев, входного слоя и выходного слоя был выбран тангенс гиперболический [156].

Также хочется отметить, что в процессе проведенного исследования для получения ответа на поставленную гипотезу о возможности обнаружения количественной связи между исследуемыми показателями без какой-либо предварительной обработки данных был получен дополнительный результат влияния количества нейронов на скрытых слоях ИНС на качество обучения и обобщающую способность этих ИНС. Похожие выводы представлены в других исследованиях [6; 187], что подтверждает адекватность полученных результатов.

4.3.1. Результаты обучения нейронных сетей без предварительной обработки исходных данных

Процесс обучения нейронных сетей включает в себя подбор синаптических весов, инициализированных случайным образом на первоначальном этапе. В связи с этим оценка качества обучения ИНС должна основываться не только на наиболее успешных результатах, но и учитывать все попытки обучения в целом.

В результате проведения 800 экспериментов по обучению ИНС без предварительной обработки данных, используя все 10 вен головного мозга при ограниченных данных (65 примеров), было установлено, что ни одна из обученных на основе 10 вен головного мозга ИНС не соответствовала критериям качества обучения при тестировании (16 примеров) [157].

В Таблице 4.6 представлены результаты обучения и тестирования ИНС с одним скрытым слоем.

Таблица 4.6 – Результатов обучения полносвязных нейронных сетей с одним скрытым слоем и десятью входными сигналами [157]

ИНС	Кол-во ИНС	Средн.кв. ошибка обучения	Средн.кв. ошибка обобщения (тестирования)	Кол-во ИНС, R^2 Train >0,7	Кол-во ИНС, R^2 Test >0,7	Кол-во ИНС, средн. кв. отн._Train (%) < 20 %	Кол-во ИНС, средн. кв. отн._Test (%) < 20 %
1-5	50	0,910	12,185	50	0	45	0
1-10	50	0,460	7,497	50	0	49	0
1-15	50	0,354	7,624	50	0	49	0
1-20	50	0,314	10,545	50	0	50	0
1-25	50	0,300	9,901	50	0	49	0
1-27	50	0,237	10,341	50	0	50	0
1-30	50	0,033	16,103	50	0	45	0

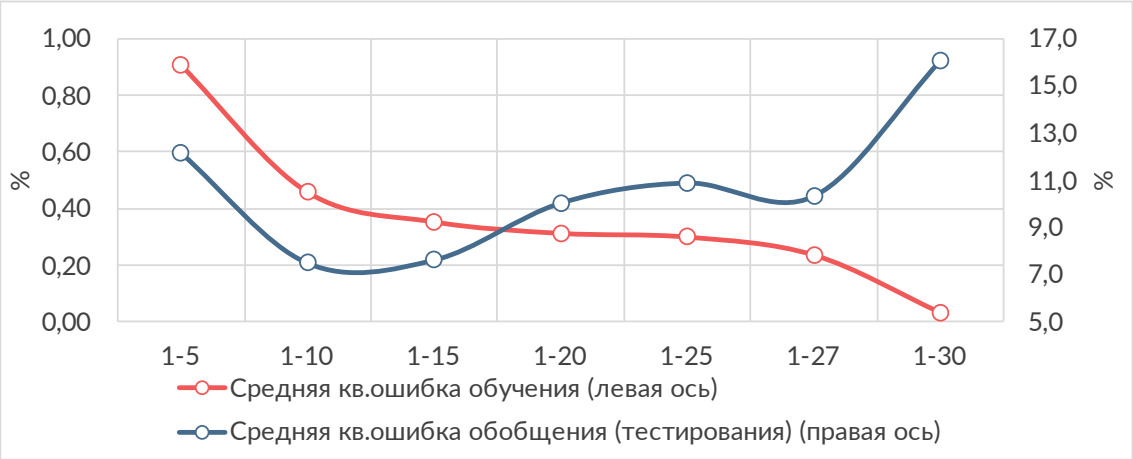


Рисунок 4.7 – Среднеквадратические ошибки обучения и обобщения (тестирования) полносвязных ИНС с одним скрытым слоем с различным числом нейронов на скрытом слое [157]

На Рисунке 4.7 видно, что среднеквадратическая ошибка обучения (левая ось) ИНС уменьшалась при увеличении количества нейронов на скрытом слое

сети, в то время как среднеквадратическая ошибки обобщения (тестирования) ведет себя немонотонно.

Для наглядного представления результатов многократных экспериментов нейросетевого моделирования предложен новый тип столбчатой диаграммы, который позволяет визуализировать качество обучения нейронных сетей.

Данная диаграмма объединяет количественные и графические элементы, что делает её удобной для анализа. По левой оси отображается показатель среднеквадратической ошибки обобщения (тестирования), а по правой оси – значения коэффициента детерминации (R^2) на тестовом множестве. Кроме того, столбцы, соответствующие ИНС, которые удовлетворяют критериям качества, выделяются зеленым цветом, а те, которые не соответствуют – красным. Это позволяет оценить степень сложности получения качественной ИНС при выбранной конфигурации [157].

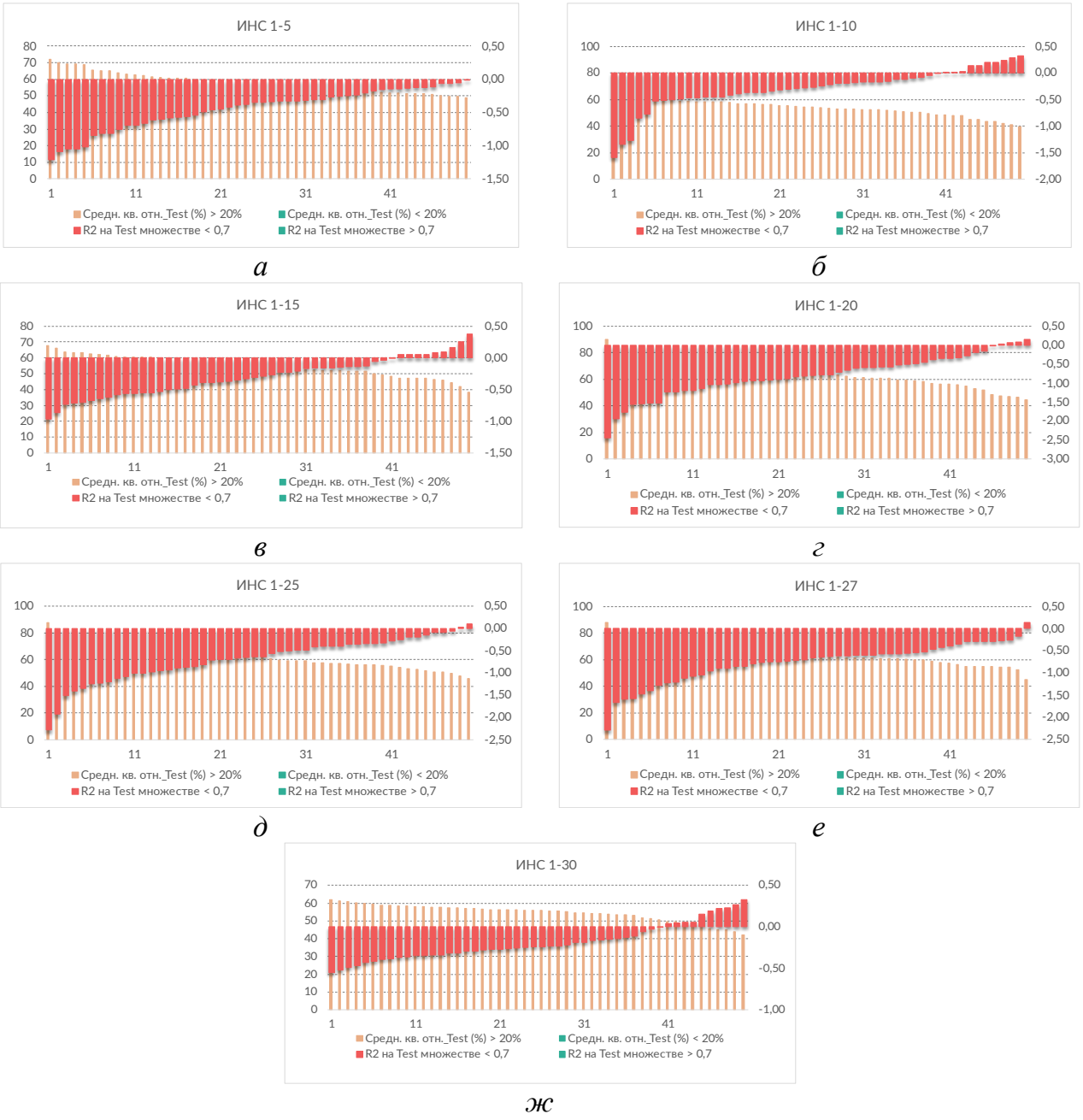
Ниже представлены предлагаемые столбчатые диаграммы для ИНС с одним скрытым слоем и с различным числом нейронов на скрытых слоях (Рисунок 4.8) [157].

Предложенный вид столбчатой диаграммы может быть адаптирован для различных конфигураций ИНС и типов данных, что делает его универсальным инструментом для анализа результатов нейросетевого моделирования.

Из Рисунка 4.8 видно, что ни одна из 350 ИНС с одним скрытым слоем не удовлетворяет заданным критериям качества. Более того, с увеличением количества нейронов на скрытом слое ошибка тестирования у ИНС 1-25, 1-27, 1-30 увеличивается, что видно из Рисунка 4.7. Здесь начинает проявляться эффект персонализации нейронов к исходным примерам, поскольку становится сопоставимо количество нейронов на скрытом слое и количество примеров для обучения [157].

Результаты многократных серий нейросетевого моделирования нейронных сетей с двумя скрытыми слоями и различным числом нейронов приведены в Таблице 4.7. Для обозначения конфигурации сети в столбце ИНС и далее по тексту используется следующая система: первое число указывает на

количество скрытых слоёв, второе число – на количество нейронов в первом скрытом слое, а третье число – на количество нейронов во втором скрытом слое. В ходе исследования было проведено девять серий экспериментов, при этом в каждой серии проводилось по 50 попыток обучения ИНС. Шаг изменения количества нейронов на скрытых слоях для разных конфигураций ИНС был установлен на уровне 5 [157].



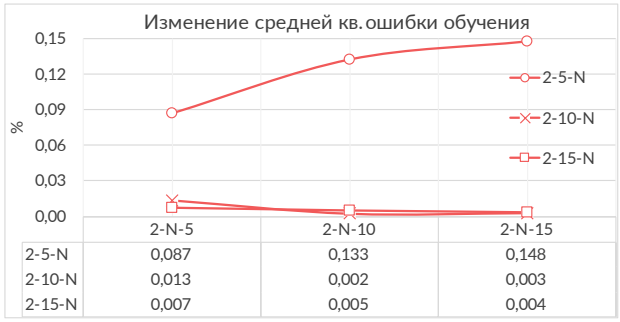
а *б* *в* *г* *д* *е* *ж*

Рисунок 4.8 – Графическое представление (*а–ж*) результатов нейросетевого моделирования полносвязных ИНС с одним скрытым слоем с различным числом нейронов на скрытом слое [157]

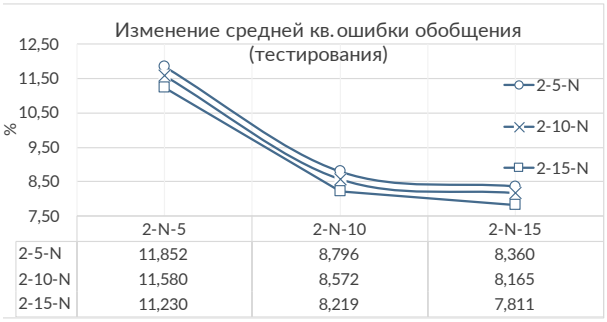
Таблица 4.7 – Результаты обучения полносвязных нейронных сетей с двумя скрытыми слоями и десятью входными сигналами [157]

ИНС	Кол-во ИНС	Средн.кв. ошибка обучения	Средн.кв. ошибка обобщения (тестирования)	Кол-во ИНС, R^2 Train >0,7	Кол-во ИНС, R^2 Test >0,7	Кол-во ИНС, средн. кв. отн._Train (%) < 20 %	Кол-во ИНС, средн. кв. отн._Test (%) < 20 %
2-5-5	50	0,087	11,852	50	0	50	0
2-5-10	50	0,133	8,796	50	0	50	0
2-5-15	50	0,148	8,360	50	0	50	0
2-10-5	50	0,013	11,580	50	0	50	0
2-10-10	50	0,002	8,572	50	0	50	0
2-10-15	50	0,003	8,165	50	0	50	0
2-15-5	50	0,007	11,230	50	0	50	0
2-15-10	50	0,005	8,219	50	0	50	0
2-15-15	50	0,004	7,811	50	0	50	0

Ниже представлены зависимости среднеквадратической ошибки обучения (Рисунок 4.9, *а*) и обобщения (тестирования) (Рисунок 4.9, *б*) ИНС с двумя скрытыми слоями с различным числом нейронов на скрытом слое. Графики представляют собой зависимости изменения средней квадратической ошибки при фиксированном числе нейронов на первом скрытом слое от числа нейронов на втором скрытом слое. Например, график 2-5-N, соответствующий значению по оси абсцисс 2-N-15, соответствует структуре ИНС 2-5-15 [157].



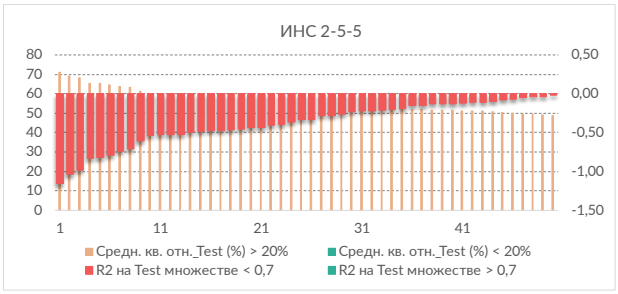
а



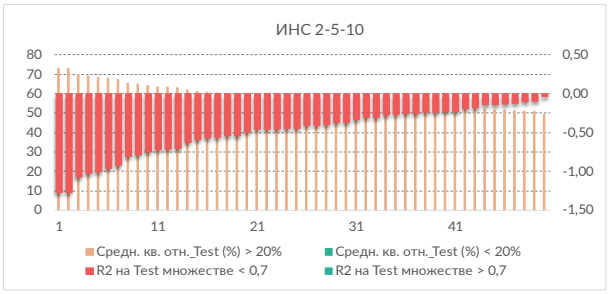
б

Рисунок 4.9 – Среднеквадратические ошибки обучения и обобщения (тестирования) полносвязных ИНС с двумя скрытыми слоями в зависимости от числа нейронов на первом скрытом слое [157]

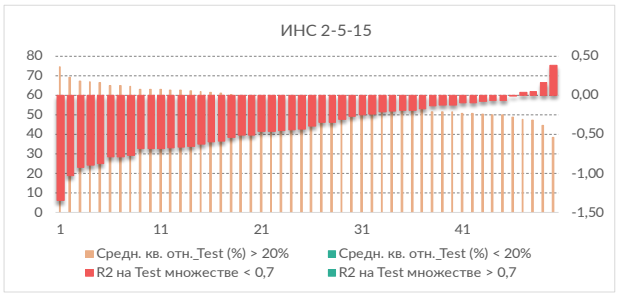
На Рисунке 4.10 представлены столбчатые диаграммы для наглядного представления результатов 450 попыток обучить ИНС с двумя скрытыми слоями.



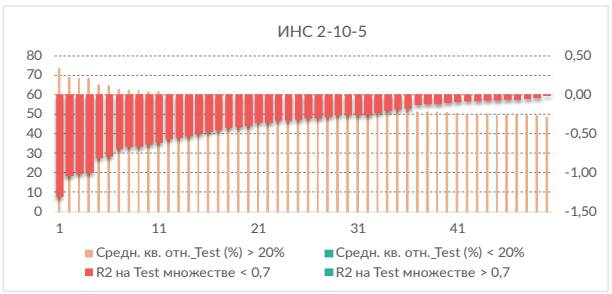
а



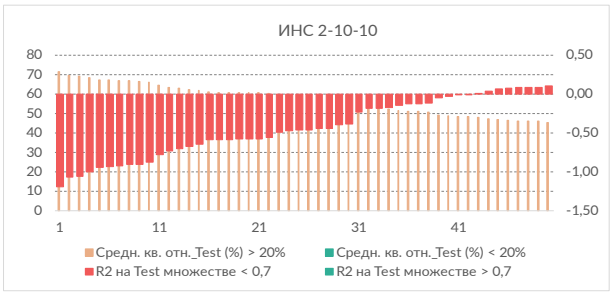
б



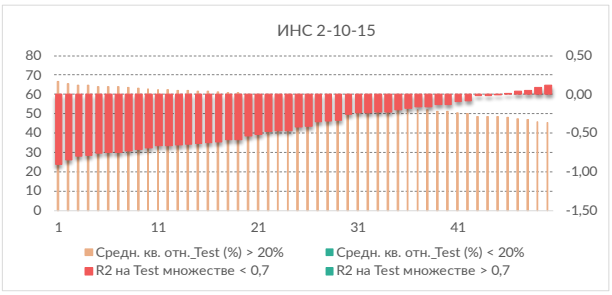
в



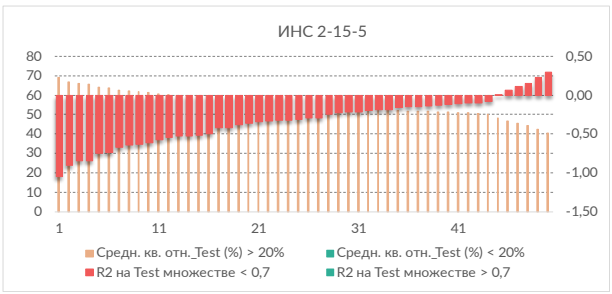
г



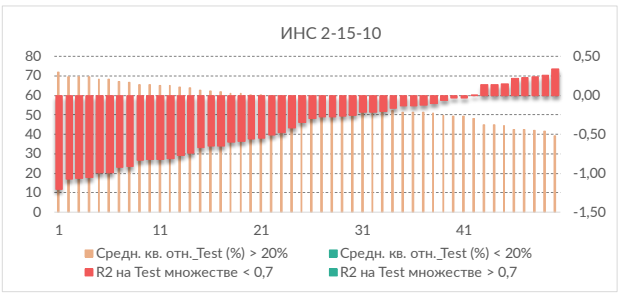
д



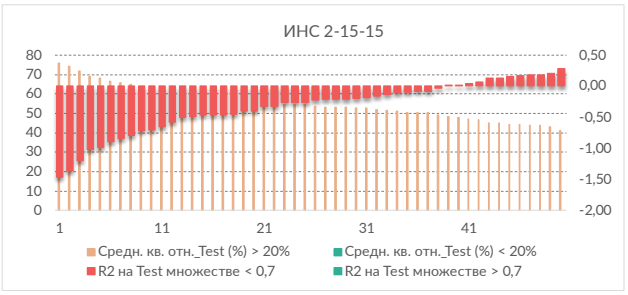
е



ж



з



к

Рисунок 4.10 – Графическое представление (*а–к*) результатов нейросетевого моделирования полносвязных ИНС с двумя скрытыми слоями с различным числом нейронов на скрытых слоях [157]

Из Рисунка 4.10 видно, что в результате многократных экспериментов в одних и тех же условиях не удалось получить ИНС с двумя слоями, которые могли бы удовлетворять заданным критериям качества, несмотря на их довольно «мягкие» значения.

В результате анализа 800 нейросетей, обученных без предварительной обработки данных, установлено, что ни одна из них не достигает требуемых показателей: среднеквадратической ошибки менее 20 % и коэффициента детерминации выше 70 %. Таким образом, гипотеза, выдвинутая на научных семинарах, о том, что качественная нейросеть может быть обучена на исходных данных без предварительной обработки, не нашла подтверждения.

С другой стороны, в работе [155] представлены результаты успешной обучаемости нейронных сетей на тех же исходных данных, которые были предварительно обработаны путём отбора наиболее значимых показателей и определения числа синаптических связей в нейронной сети, основанной на корне принятия решений, идентифицированном после интервального кодирования исходных данных. Это позволило определить конфигурацию полносвязных нейронных сетей с сопоставимым числом синаптических коэффициентов.

Однако стоит отметить, что в работе [155] представлены только лучшие результаты обучения, и остается вопрос о том, насколько легко получить такие нейронные сети? Так как в [157] предложена новая диаграмма, позволяющая наглядно показать результаты многократных обучений, следовательно, стоит привести все результаты, а не только лучшие. Это позволит сравнить результаты нейронных сетей и оценить качество получаемых моделей в зависимости от того, была ли проведена предобработка данных или нет.

4.3.2. Результаты обучения нейронных сетей с предварительной обработкой исходных данных

Ниже демонстрируются показатели качества 2400 обученных ИНС, среди которых было семь нейронных сетей с одним скрытым слоем (Таблица 4.8) и девять нейронных сетей с двумя скрытыми слоями (Таблица 5), для каждой из которых проводилось по 150 экспериментов [157].

Таблица 4.8 – Результаты обучения полносвязных нейронных сетей с одним скрытым слоем с четырьмя входными сигналами [157]

ИНС	Кол-во ИНС	Средн. кв. ошибка обучения	Средн. кв. ошибка обобщения (тестирования)	Кол-во ИНС, R^2 Train >0,7	Кол-во ИНС, R^2 Test >0,7	Кол-во ИНС, средн. кв. отн._Train (%) < 20 %	Кол-во ИНС, средн. кв. отн._Test (%) < 20 %
1-5	150	5,399	23,998	18	16	9	8
1-10	150	1,904	21,541	149	13	148	5
1-15	150	1,205	17,682	150	32	150	4
1-20	150	1,001	15,158	150	40	150	14
1-25	150	0,880	16,491	150	33	150	8
1-27	150	0,310	17,626	150	26	150	6
1-30	150	0,388	18,079	150	24	150	1

Из данных Таблицы 4.8 видим, что для каждой из рассмотренных конфигураций нейронной сети с единственным скрытым слоем существуют варианты, которые соответствуют установленным на начальном этапе исследования критериям качества. Исследование показывает, что шансы успешно обучить нейронную сеть с одним скрытым слоем повышаются при использовании 20 нейронов на этом слое (Рисунок 4.11). На этом же Рисунке 4.11 можно наблюдать, как среднеквадратическая погрешность обучения уменьшается по мере увеличения количества нейронов, в то время как среднеквадратическая погрешность при обобщении начинает возрастать, что указывает на переобученность нейронной сети [157].

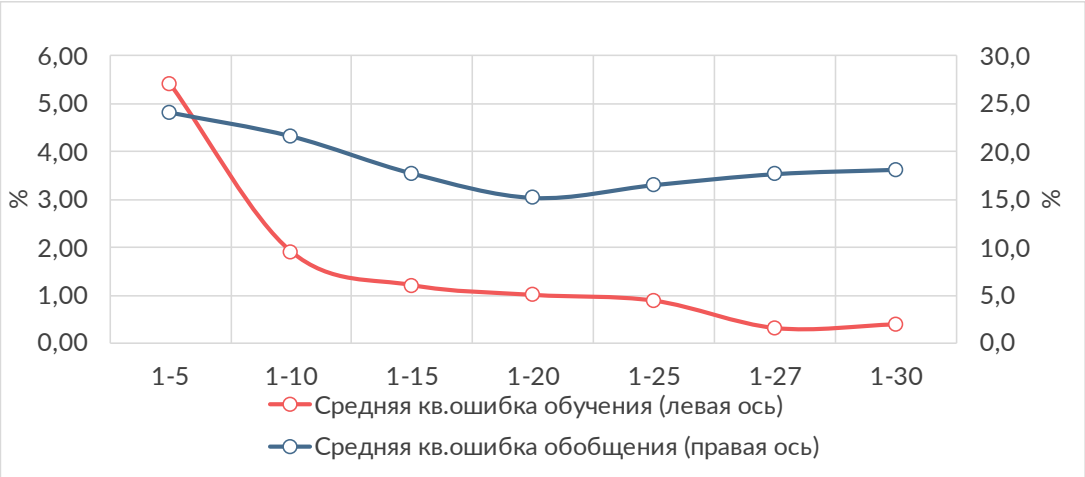


Рисунок 4.11 – Средние квадратические ошибки обучения и обобщения (тестирования) полносвязных ИНС с одним скрытым слоем и четырьмя входными сигналами в зависимости от числа нейронов на скрытом слое [157]

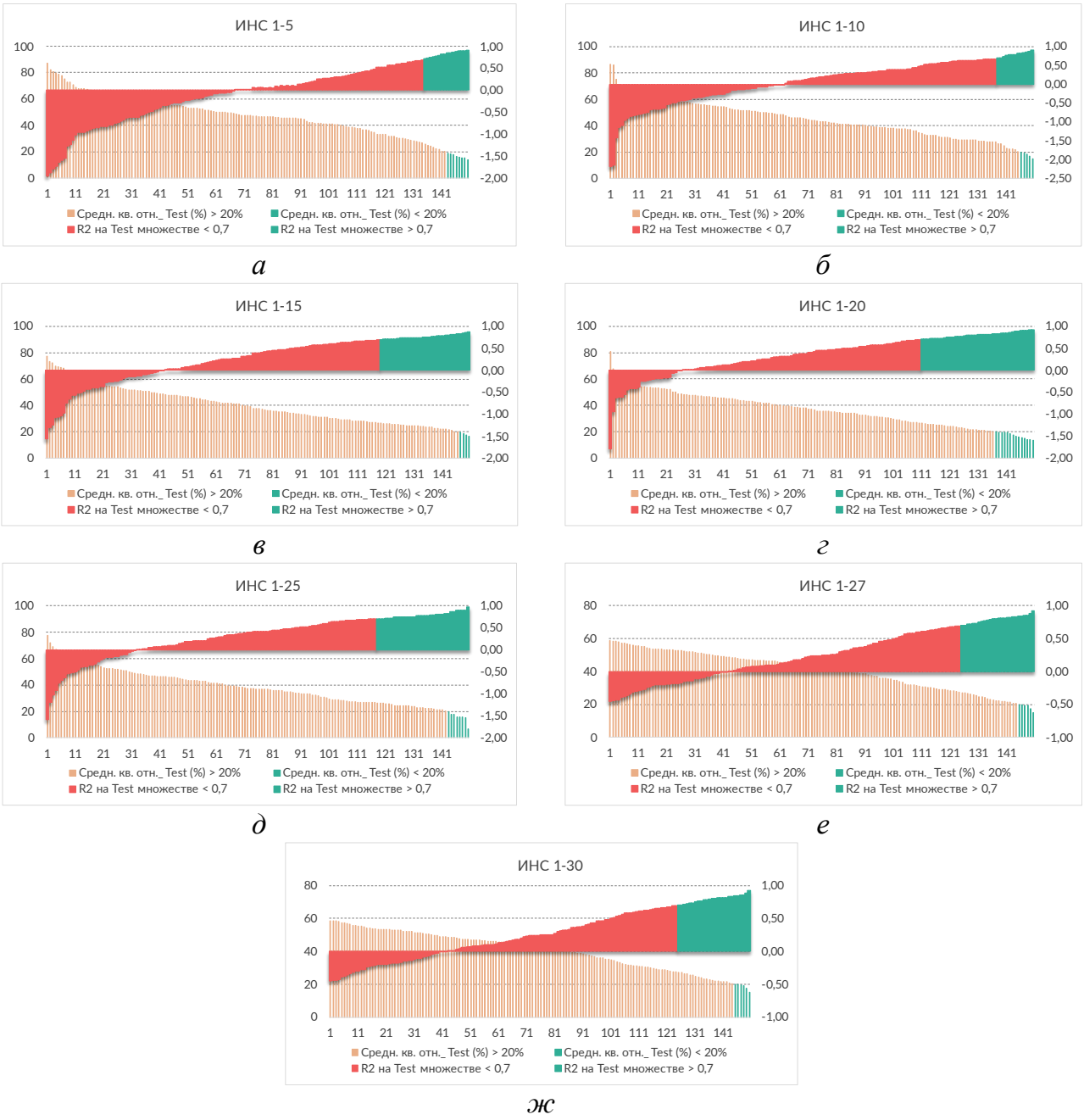
Из данных Таблицы 4.9 видим, что для каждой комбинации из двух скрытых слоев существуют нейронные сети, которые удовлетворяют установленным критериям качества. Вероятность того, что нейронная сеть будет качественно обучена с двумя скрытыми слоями оказывается выше, особенно при наличии 20 нейронов на скрытых слоях. Среди всех вариантов нейронная сеть с конфигурацией 2–10–10 демонстрирует наилучшие результаты, удовлетворяющие требованию по коэффициенту детерминации. Незначительно отстает от неё нейронная сеть с конфигурацией 2–5–15, которая тем не менее занимает ведущие позиции по количеству полученных нейронных сетей той же конфигурации, удовлетворяющих критерию по среднеквадратической ошибке при обобщении (тестировании), при этом среднее значение этой ошибки является одним из самых малых.

Таблица 4.9 – Результаты обучения полносвязных нейронных сетей с двумя скрытыми слоями с четырьмя входными сигналами [157]

ИНС	Кол-во ИНС	Средн.кв. ошибка обучения	Средн.кв. ошибка обобщения (тестирования)	Кол-во ИНС, R^2 Train >0,7	Кол-во ИНС, R^2 Test >0,7	Кол-во ИНС, средн. кв. отн._Train (%) < 20 %	Кол-во ИНС, средн. кв. отн._Test (%) < 20 %
2-5-5	150	1,355	22,081	18	18	9	9
2-5-10	150	0,716	24,733	150	14	150	7
2-5-15	150	0,509	23,325	150	17	150	5
2-10-5	150	0,543	23,421	149	21	148	11
2-10-10	150	0,132	17,886	150	26	150	8
2-10-15	150	0,060	20,492	149	21	145	7
2-15-5	150	0,173	17,049	150	24	150	11
2-15-10	150	0,057	17,221	150	16	150	6
2-15-15	150	0,020	16,649	150	19	150	5

Сравнивая данные Таблиц 4.8 и 4.9, видим, что вероятность получить качественно обученную ИНС с одним скрытым слоем выше, нежели обучать ИНС с двумя скрытыми слоями. Это может свидетельствовать о качественном отборе наиболее значимых показателей.

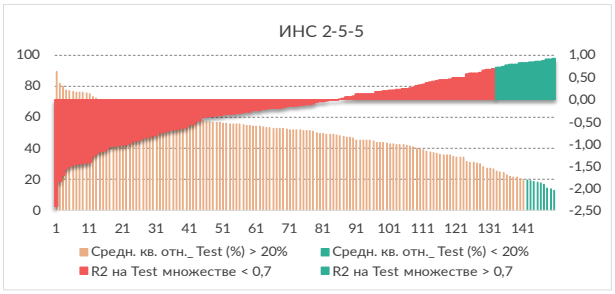
К аналогичным выводам можно прийти, глядя на диаграммы ниже, где приведены результаты обучения ИНС с одним скрытым слоем и четырьмя входными сигналами (Рисунок 4.12), а на Рисунке 4.13 приведены девять ИНС с двумя скрытыми слоями и также четырьмя входными сигналами [157].



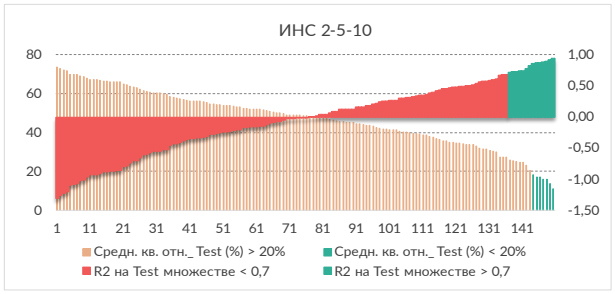
а *б* *в* *г* *д* *е* *ж*

Рисунок 4.12 – Графическое представление (*а–ж*) результатов нейросетевого моделирования полносвязных ИНС с одним скрытым слоем в зависимости от числа нейронов на скрытом слое [157]

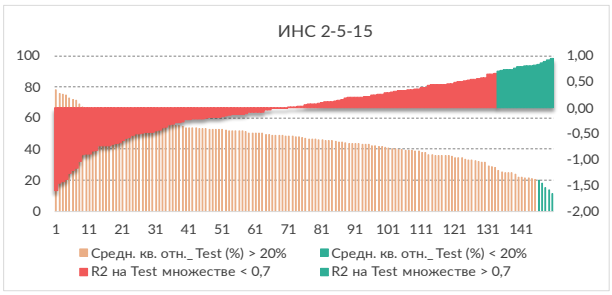
Из Рисунка 4.12 видим, что зеленые столбцы чаще встречаются на четвертой диаграмме (см. Рисунок 4.12, *г*), что согласуется с выводом выше о наибольшем качестве обучения ИНС при 20 нейронах на одном скрытом слое. Это свидетельствует о наглядности предложенной диаграммы, особенно когда они сопоставляются вместе [157].



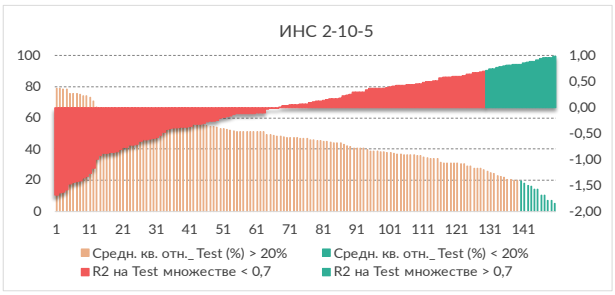
а



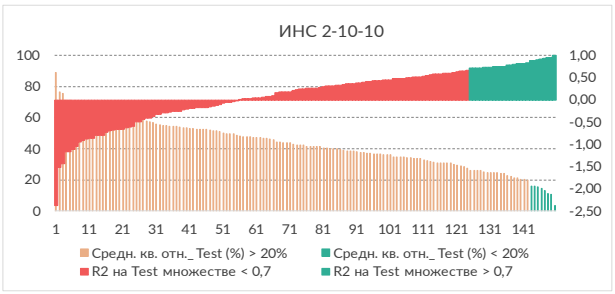
б



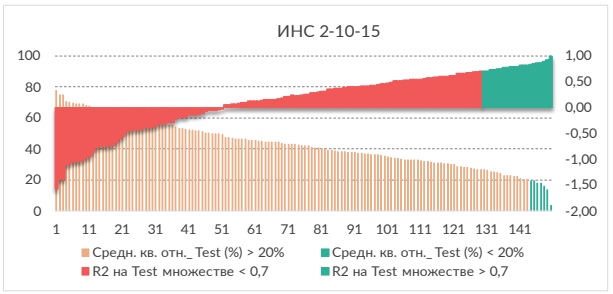
в



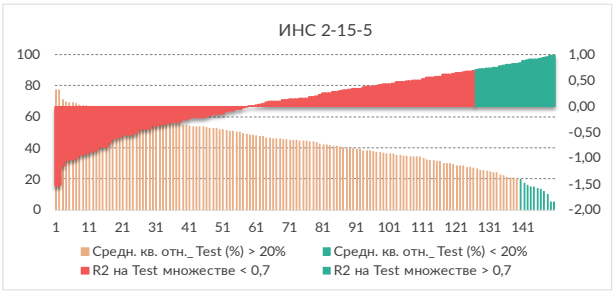
г



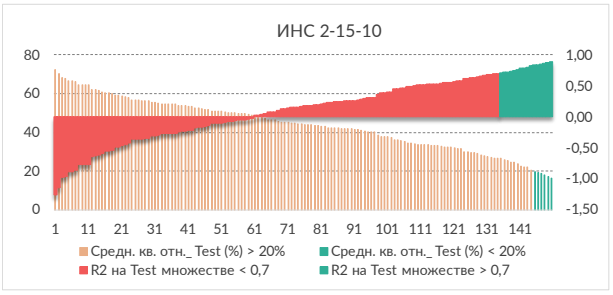
д



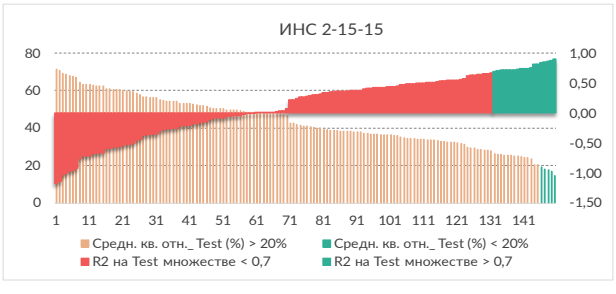
е



ж



з



к

Рисунок 4.13 – Графическое представление (*а*–*к*) результатов нейросетевого моделирования полносвязных ИНС с двумя скрытыми слоями с различным числом нейронов на скрытых слоях [157]

На Рисунке 4.13 визуально выделяются две диаграммы, соответствующие ИНС 2-10-10 (см. Рисунок 4.13, *д*) и ИНС 2-5-15 (см. Рисунок 14.13, *ж*), среди

которых преимущество ИНС 2-5-15 перед ИНС 2-10-10 еле заметно, но все же видно. Это еще раз подтверждает наглядность предложенной диаграммы [157].

4.4. Нейросетевое моделирование с помощью библиотек Keras, TensorFlow

Далее приведены результаты серии экспериментов по обучению нейронных сетей, выполненных в программном модуле (Приложение 8) для оценки качества обучения нейронных сетей [161] как на данных без их предварительной обработки (Таблица 4.10), так и на данных с предварительной обработкой (Таблица 4.11).

4.4.1. Результаты обучения нейронных сетей без предварительной обработки исходных данных

В Таблице 4.10 представлены средние результаты обобщения (тестирования) обученных ИНС, полученные в ходе многократных серий моделирования нейронных сетей как с одним скрытым слоем, так и с двумя скрытыми слоями и разным количеством нейронов на этих слоях на данных без предварительной обработки.

Таблица 4.10 – Результаты обобщения (тестирования) обученной ИНС без предварительной обработки исходных данных

ИНС	Средняя Accuracy *(Точность) TP, TN, FP, FN*	Средняя Precision *(Точность) TP, FP*	Средняя Mean Squared Error *(Среднеквадратическая ошибка)*	Средняя Sensitivity *(Чувствительность)*	Средняя Specificity *(Специфичность)*	Средняя AUC *(Area Under the ROC Curve, площадь под ROC-кривой)*	Средний R^2 *(коэф. детерм.)*
1-5	0,495	0,496	0,441	0,426	0,198	0,312	-0,837
1-10	0,503	0,525	0,425	0,436	0,188	0,312	-0,637
1-15	0,499	0,543	0,423	0,458	0,151	0,304	-0,353
1-20	0,471	0,509	0,454	0,424	0,120	0,272	-0,359
1-25	0,523	0,540	0,402	0,444	0,160	0,302	-0,569
1-30	0,471	0,525	0,450	0,423	0,165	0,294	-0,998
2-5-5	0,492	0,489	0,452	0,421	0,220	0,321	-1,015
2-5-10	0,542	0,561	0,399	0,484	0,155	0,319	-0,374
2-5-15	0,526	0,555	0,413	0,460	0,185	0,323	-0,748
2-10-5	0,497	0,563	0,428	0,473	0,131	0,302	-0,622
2-10-10	0,471	0,496	0,465	0,412	0,185	0,299	-1,009
2-10-15	0,506	0,554	0,419	0,450	0,170	0,310	-0,846
2-15-5	0,519	0,552	0,401	0,446	0,165	0,306	-0,753
2-15-10	0,476	0,544	0,450	0,440	0,150	0,295	-0,819
2-15-15	0,523	0,582	0,404	0,470	0,155	0,312	-0,681

4.4.2. Результаты обучения нейронных сетей с предварительной обработкой исходных данных

Далее приведены средние результаты обобщения (тестирования) обученных ИНС в ходе серии экспериментов по обучению нейронных сетей, выполненных в программном модуле для оценки качества обучения нейронных сетей [161] на данных с предварительной обработкой (Таблица 4.11).

Таблица 4.11 – Результаты обобщения (тестирования) обученной ИНС с предварительной обработкой исходных данных

ИНС	Средняя Accuracy *(Точность) TP, TN, FP, FN*	Средняя Precision *(Точность) TP, FP*	Средняя Mean Squared Error *(Среднеквадратическая ошибка)*	Средняя Sensitivity *(Чувствительность)*	Средняя Specificity *(Специфичность)*	Средняя AUC *(Area Under the ROC Curve, площадь под ROC-кривой)*	Средний R^2 *(коэф. детерм.)*
1-5	0,605	0,556	0,325	0,679	0,397	0,538	-0,646
1-10	0,573	0,520	0,349	0,628	0,415	0,521	-0,767
1-15	0,586	0,543	0,333	0,654	0,395	0,525	-0,684
1-20	0,597	0,557	0,318	0,674	0,367	0,521	-0,598
1-25	0,588	0,558	0,320	0,677	0,351	0,514	-0,630
1-30	0,559	0,500	0,359	0,610	0,410	0,510	-0,817
2-5-5	0,579	0,513	0,356	0,643	0,412	0,527	-0,817
2-5-10	0,569	0,504	0,360	0,622	0,417	0,520	-0,827
2-5-15	0,584	0,532	0,336	0,656	0,370	0,513	-0,692
2-10-5	0,586	0,539	0,334	0,664	0,372	0,518	-0,699
2-10-10	0,571	0,521	0,347	0,639	0,386	0,513	-0,768
2-10-15	0,545	0,490	0,375	0,594	0,433	0,513	-0,928
2-15-5	0,585	0,544	0,332	0,671	0,364	0,517	-0,698
2-15-10	0,592	0,543	0,330	0,666	0,383	0,524	-0,670
2-15-15	0,590	0,543	0,330	0,659	0,389	0,524	-0,667

Сравнивая полученные результаты, можно сделать вывод, что вероятность получить более качественно обученную ИНС выше, если проводить нейросетевое моделирование на предварительно обработанных данных. Из данных Таблиц 4.10 и 4.11 видим, что средняя Mean Squared Error (среднеквадратическая ошибка) меньше, а средняя Sensitivity (чувствительность) и средняя AUC (Area Under the ROC Curve, площадь под ROC-кривой) выше в случае обучения ИНС на предварительно обработанных данных.

4.5. Тестирование программного модуля проектирования и обучения неполносвязных нейронных сетей класса Decision's-Root-based Neural Network

Как было сказано в Главе 2, на основе корня принятия решений можно определить неполносвязную ИНС, обладающую приемлемым уровнем точности. Согласно алгоритму (см. Рисунок 3.9), на примере количественных исходных данных [184] после определения наиболее значимых факторов с помощью АСК-анализа идентифицируется корень принятия решений, описывающий свёртку категориальных значений этих факторов.

В предыдущем исследовании [151] был использован алгоритм идентификации, который применим к любой структуре критериальных деревьев без ограничений на алфавит, используемый в матрице свёртки. Сняв ограничение на алфавит, используемый для кодирования элементов матриц свёртки, можно аппроксимировать исходные данные несколькими функциями, так как при склеивании групп эквивалентности возникает вариабельность этих склеек [151].

Как уже было сказано в Главе 3, для четырех факторов возможно 15 полных бинарных деревьев с именованными листьями (формула (3.1)). Из всех возможных древовидных структур предложено воспроизводить свёртки показателей, характеризующие левые вены друг с другом, а также правые друг с другом, в связи с особенностью физиологической работы кровообращения двух полушарий. Для каждой древовидной структуры критериев можно выделить несколько корней принятия решений, два примера из возможных на основе обучающей выборки после исключения конфликтных примеров представлены на Рисунках 4.14–4.15.

Представленный на Рисунке 4.14 корень принятия решений в корневой матрице имеет два не до конца идентифицированных элемента (см. Рисунок 4.14, вторая и третья строки третьего столбца корневой матрицы), поскольку при принятии решений об исключении противоречивых примеров из исходных данных дополнительно учитывалась информация о степени сходства с классами CON или AD, что сделало корень принятия решений более достоверным.

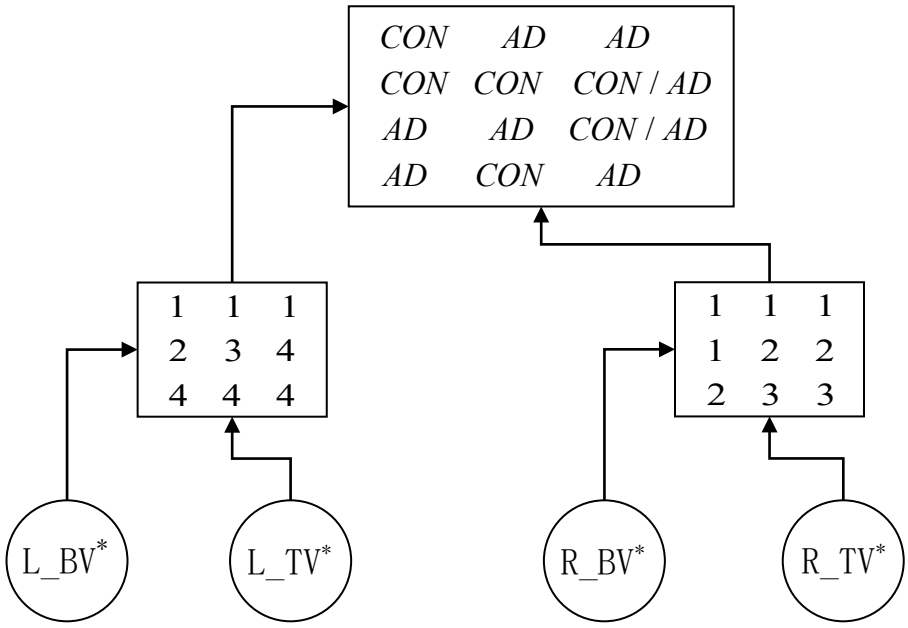


Рисунок 4.14 – Корень принятия решений, описывающий свёртку категориальных значений MSV вен по расположению в головном мозге: вена левого полушария (L_BV) с веной левого полушария (L_TV), вена правого полушария (R_BV) с веной правого полушария (R_TV) [155]

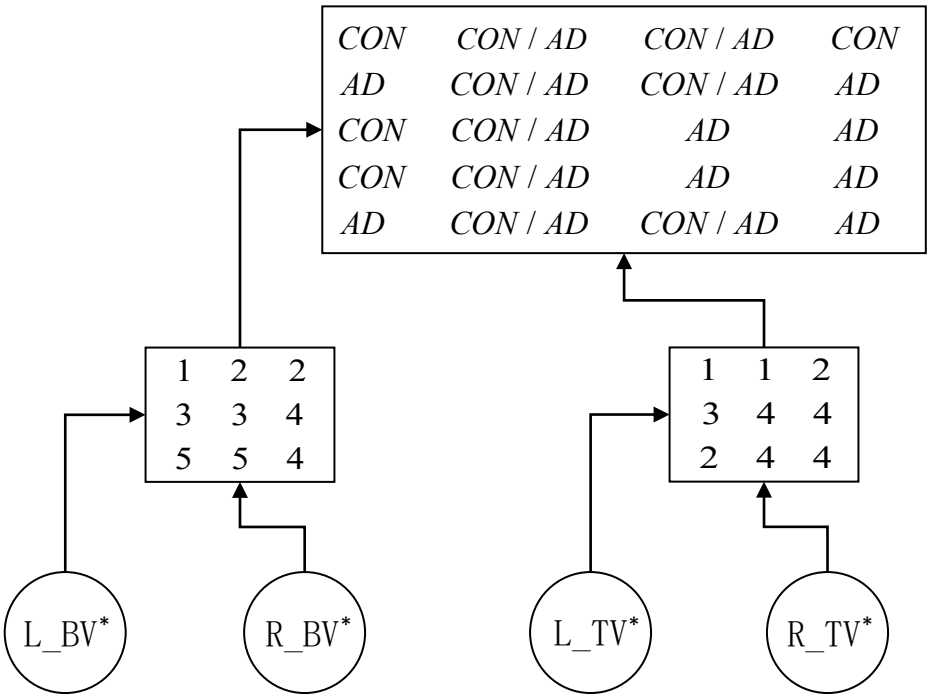


Рисунок 4.15 – Корень принятия решений, описывающий свёртку категориальных значений MSV по типу вен в головном мозге: базальная вена левого полушария (L_BV) с базальной веной правого полушария (R_BV), вена таламуса левого полушария (L_TV) с веной таламуса правого полушария (R_TV)

В действительности этим элементам соответствуют данные как здоровых пациентов, так и пациентов с клинически диагностированным заболеванием Альцгеймера. В отношении этих пациентов результаты АСК-анализа были неинформативными, поэтому сделать выбор в пользу какой-либо группы было невозможно. Точность полученного корня принятия решений составляет 78 %, в 95 % случаев будет поставлен диагноз.

Представленный на Рисунке 4.15 корень принятия решений в корневой матрице имеет восемь не до конца идентифицированных элементов (см. Рисунок 4.15: весь второй столбец, второй и третий элементы третьего столбца). Точность полученного корня принятия решений составляет 78 %, и только в 71 % случаев будет поставлен диагноз. Это подтверждает то, что с помощью экспертного мнения, которое было получено от специалистов в области неврологии, выбор способа структуры корня принятия решений, когда свёртка категориальных значений MSV происходит по расположению вен в головном мозге, является наиболее верным.

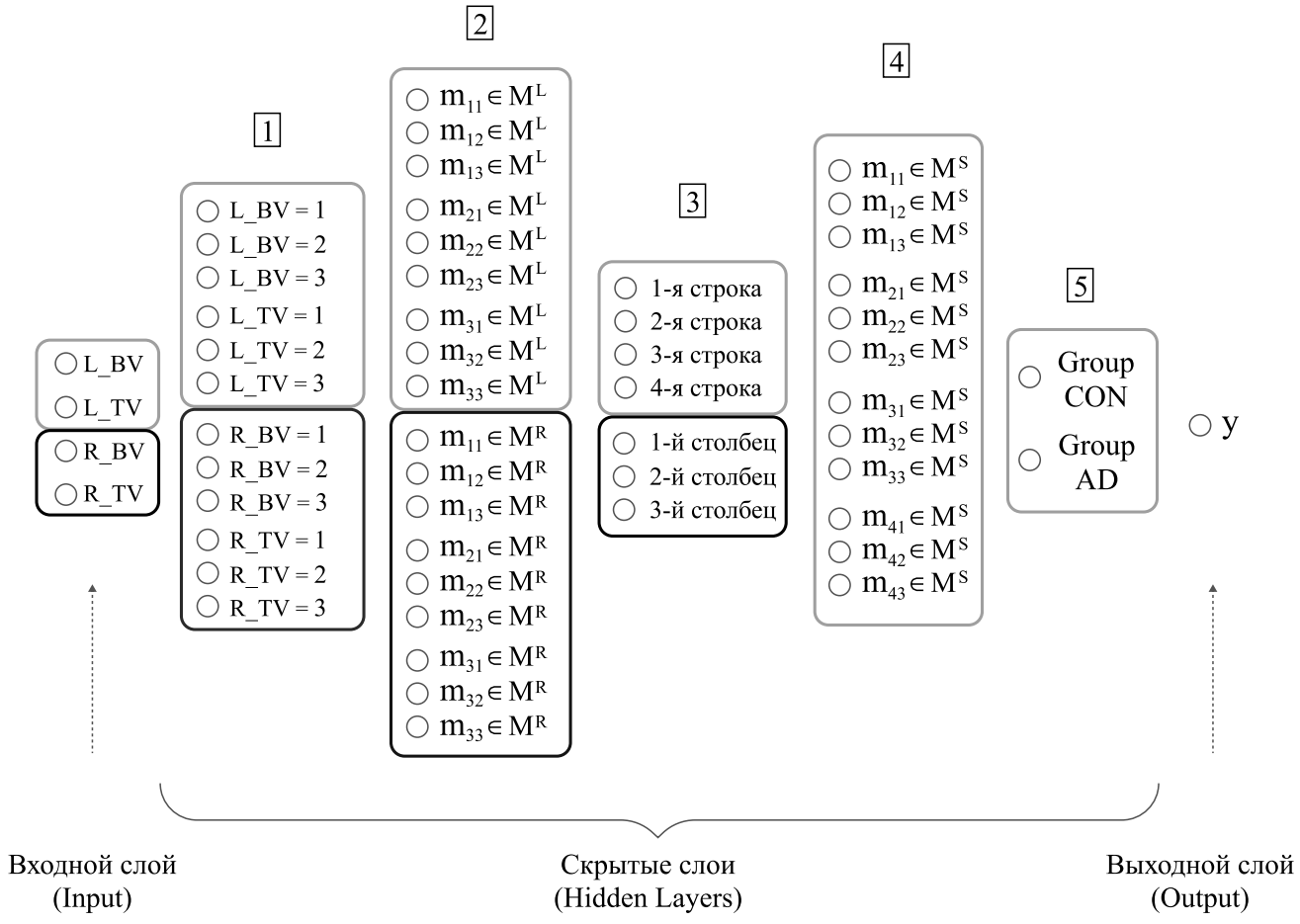


Рисунок 4.16 – Нейроны ИНС, основанной на корне принятия решений [155]

На основе корня принятия решений (см. Рисунок 4.14), описывающего свёртку категориальных значений MSV по типу вен в головном мозге, была определена структура нейронов неполносвязной ИНС (Рисунок 4.16) с синаптическими связями в соответствии с элементами матриц свёртки корня принятия решений (Рисунок 4.17).

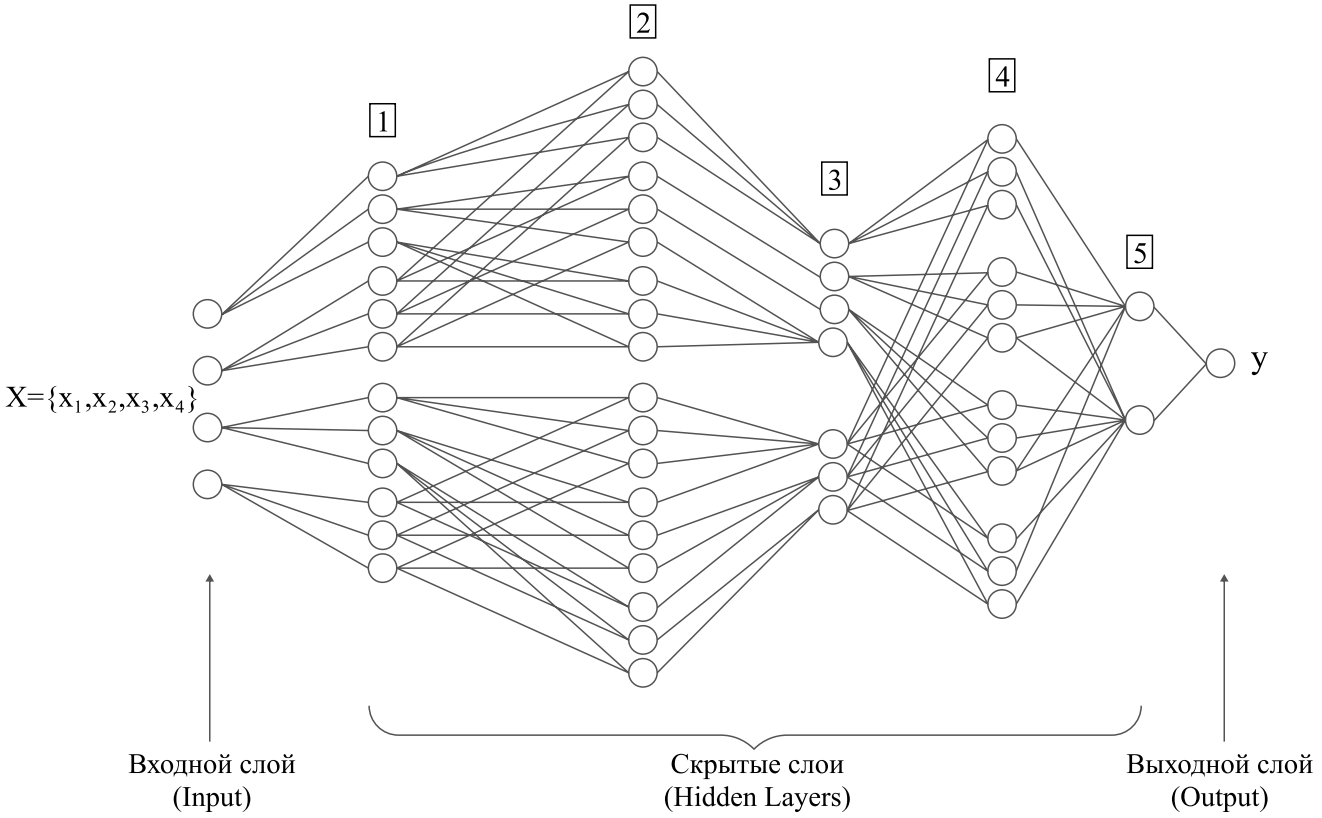


Рисунок 4.17 – Синаптические связи неполносвязной ИНС, основанной на корне принятия решений [155]

В программном модуле проектирования и обучения неполносвязных ИНС класса Decisions' Root-based Neural Network [159] неполносвязная ИНС, структура которой определена по корню принятия решений (см. Рисунок 4.14), будет иметь вид, приведенный на Рисунке 4.18.

Обучение неполносвязной ИНС класса Decision's-Root-based Neural Network проходило на множестве [184, см. прил. 1] предварительно разделенном на обучающее и валидационное в соотношении 85 : 15 % в разработанном на языке программирования Python программном модуле (Приложение 6) с использованием библиотек Keras, TensorFlow.

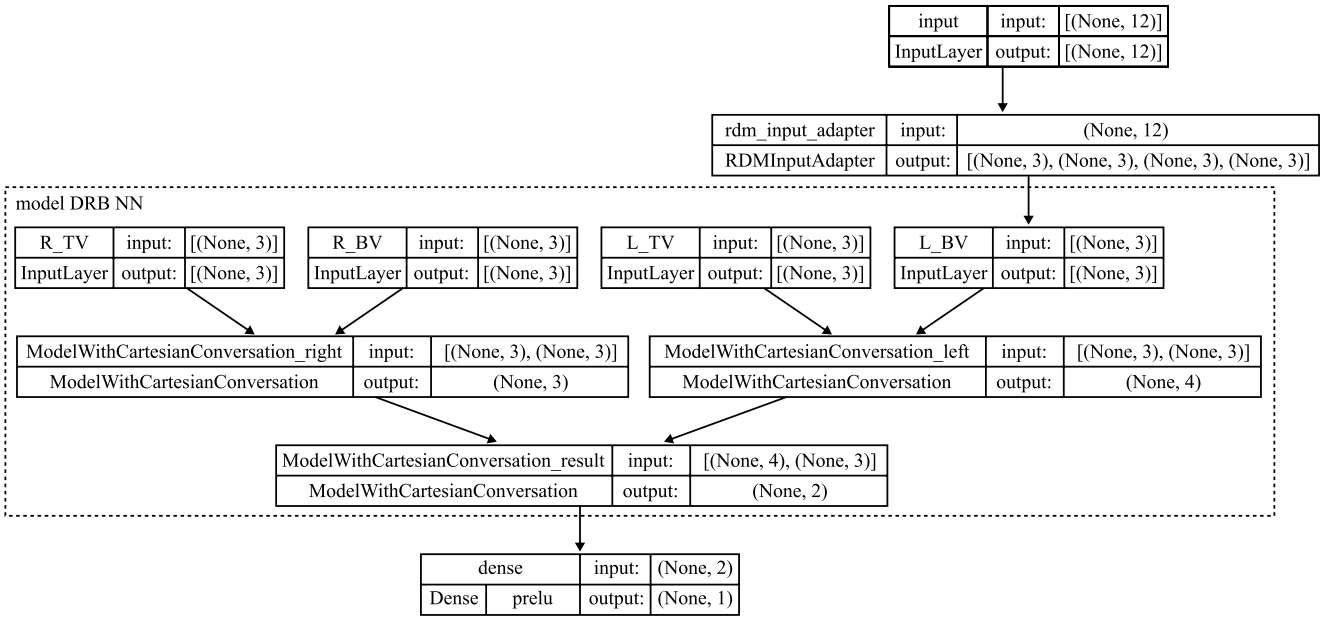


Рисунок 4.18 – Структура неполносвязной нейронной сети
класса Decision’s-Root-based Neural Network

Процесс выбора настройки гиперпараметров модели, специфичности и устойчивости обусловливается тем, что в полном соответствии с элементами матриц свёртки корня принятия решений определяется спецификация нейронной сети с нужными синаптическими связями, количеством скрытых слоев и нейронов на этих слоях, что является необходимым элементом в процессе обучения нейронной сети, так как это позволяет настроить веса и смещения связей между нейронами, чтобы сеть могла адекватно реагировать на входные данные и выполнять требуемые задачи.

В качестве метода обучения был выбран оптимизатор Adam (см. Главу 2, формулы (2.5)–(2.9)). На скрытых слоях сети была использована функция активации $\mathrm{PReLU}(x) = \max(0, x) + \gamma \min(0, x)$ – параметрический линейный или полулинейный выпрямитель (от англ. Parametric Rectified Linear Unit, ReLU), где γ – параметр, определяющий наклон отрицательной части активации. При подборе γ определено, что в нашем случае, обобщающая способность нейронной сети становилась лучше в случае $\gamma = 0$, что делает активационную функцию простой непараметрической $\mathrm{ReLU}(x) = \max(0, x)$.

Оценка качества неполносвязной нейронной сети класса DRB NN, полученной по корню принятия решения (см. Рисунок 4.18), проводилась с помощью перекрёстной проверки.

Ниже показаны изменения выбранных метрик качества обучения неполносвязной ИНС на различных эпохах в ходе 4-кратной перекрёстной проверки (Рисунок 4.19).

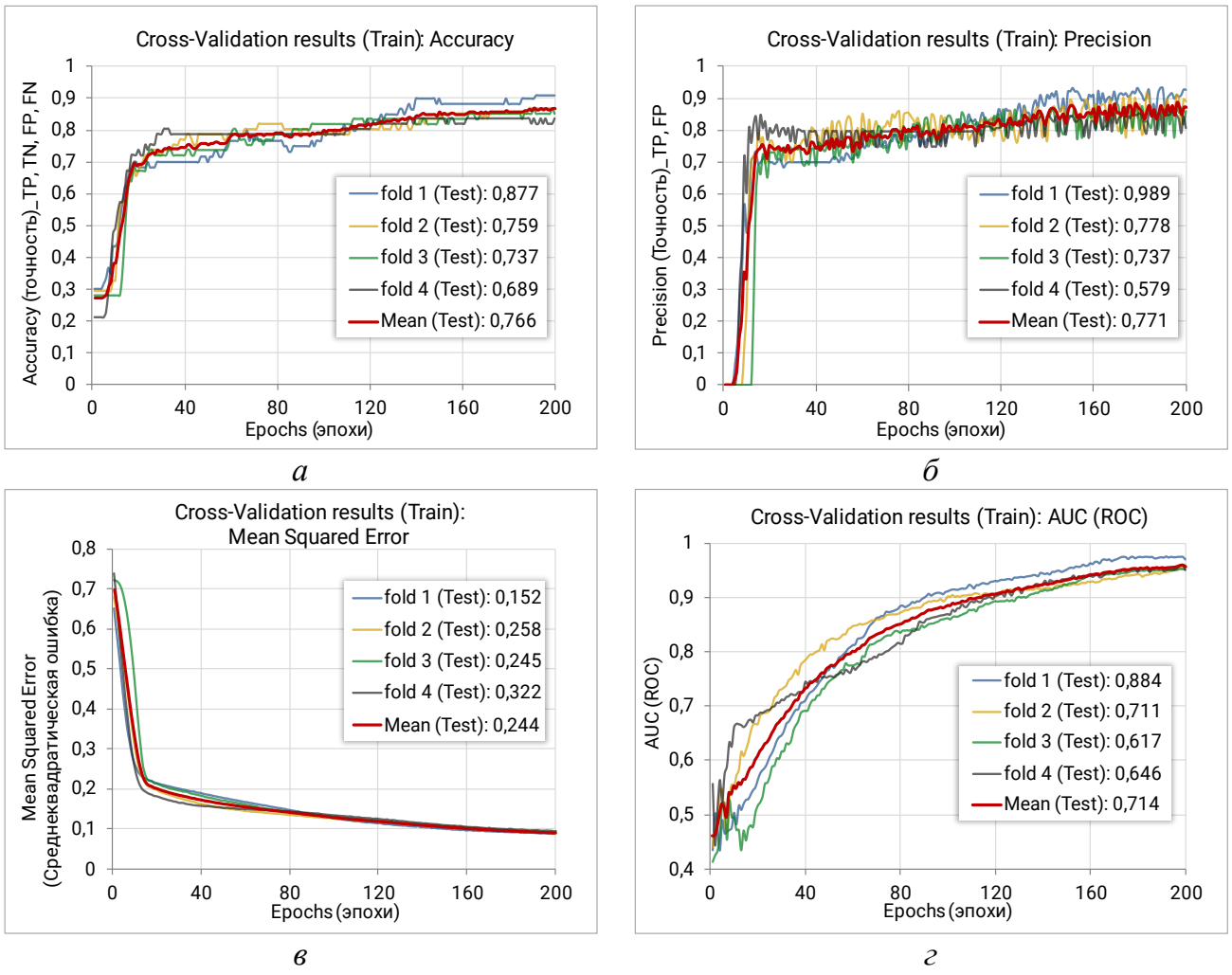


Рисунок 4.19 – Графическое представление результатов перекрёстной проверки неполносвязной ИНС при разных эпохах обучения: *а* – Accuracy *(Точность, доля правильных ответов) TP, TN, FP, FN*; *б* – Precision *(Точность) TP*, *FP*; *в* – среднеквадратическая ошибка; *г* – значения AUC *(площадь под ROC-кривой)*

Результаты кросс-валидации показывают приемлемое качество модели ИНС, полученной по корню принятия решения (см. Рисунок 4.18). Итоговые значения

полученных в ходе перекрёстной проверки метрик качества обученной неполносвязной нейронной сети представлены в Таблице 4.12.

Таблица 4.12 – Результаты перекрёстной проверки неполносвязной ИНС

Множество	Split	Accuracy *(Точность) TP, TN, FP, FN*	Precision *(Точность) TP, FP*	Mean Squared Error *(Среднеквадратическая ошибка)*	Sensitivity *(Чувствительность)*	Specificity *(Специфичность)*	AUC *(Area Under the ROC Curve, площадь под ROC-кривой)*
Train	fold1	0,910	0,928	0,090	0,944	0,767	0,928
	fold2	0,868	0,891	0,089	0,927	0,736	0,956
	fold3	0,852	0,859	0,091	0,953	0,631	0,952
	fold4	0,836	0,807	0,094	0,966	0,492	0,950
	Mean	0,867	0,871	0,091	0,933	0,660	0,946
Test	fold1	0,877	0,989	0,152	0,825	0,751	0,884
	fold2	0,759	0,778	0,258	0,875	0,615	0,711
	fold3	0,737	0,737	0,245	0,933	0,525	0,617
	fold4	0,689	0,579	0,322	0,989	0,478	0,646
	Mean	0,766	0,771	0,244	0,905	0,593	0,714

Обучение неполносвязной нейронной сети класса DRB NN проводилось в соответствии с общепринятой технологией нейросетевого моделирования. Подбор оптимальных синаптических весов осуществлялся с применением адаптивного метода обучения нейронных сетей Adam. Результаты обучения и валидации выбранной модели нейронной сети класса DRB NN представлены в Таблице 4.13, а изменения выбранных метрик качества обучения неполносвязной ИНС на различных эпохах показаны на Рисунке 4.20.

Таблица 4.13 – Показатели качества обученной неполносвязной нейронной сети класса DRB NN

DRB NN	Accuracy *(Точность, доля правильных ответов) TP, TN, FP, FN*	Precision *(Точность) TP, FP*	Mean Squared Error *(Среднеквадратическая ошибка)*	Sensitivity *(Чувствительность)*	Specificity *(Специфичность)*	AUC *(Area Under the ROC Curve, площадь под ROC-кривой)*
Train	0,941	0,952	0,064	0,978	0,880	0,983
Validation	0,923	0,900	0,170	0,925	0,875	0,925

Результаты нейросетевого моделирования неполносвязной нейронной сети, основанной на корне принятия решений, имеют следующие значения:

среднеквадратическая ошибка на обучающем множестве составила 6,4 %, на валидирующем множестве – 17 %; точность, или доля правильных ответов (*Accuracy, TP,TN,FP,FN*) на обучающем множестве – 94 %, на валидирующем множестве – 92 %, точность (*Precision, TP, FP*) полученной модели на обучающем множестве – 95 %, на валидирующем множестве – 90 %.

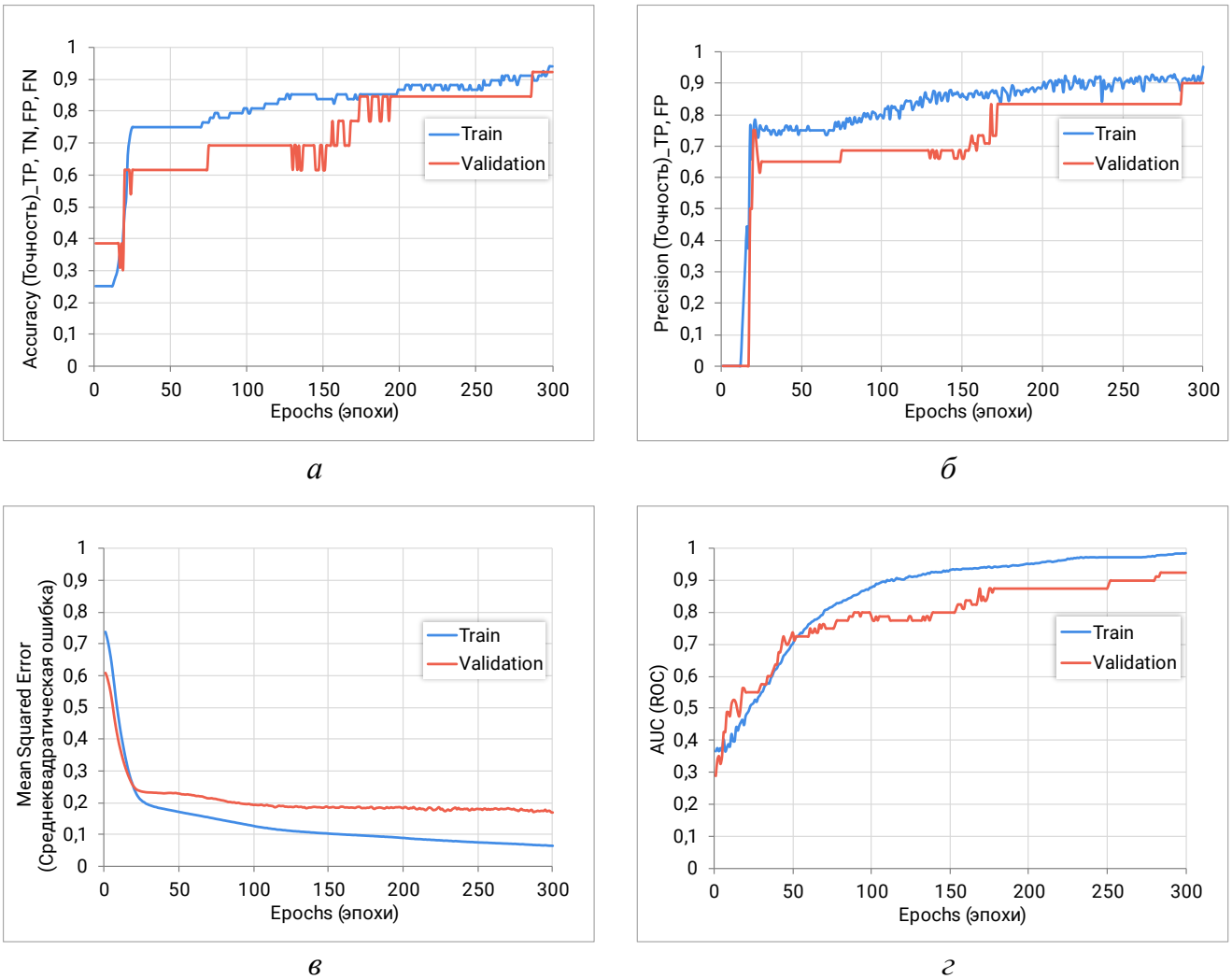


Рисунок 4.20 – Результаты обучения и валидации неполносвязной ИНС класса DRB NN: *а* – Accuracy *(Точность, доля правильных ответов) TP, TN, FP, FN*; *б* – Precision *(Точность) TP, FP; в* – среднеквадратическая ошибка; *г* – значения AUC *(площадь под ROC-кривой)*

Таким образом, полученная неполносвязная ИНС класса DRB NN обладает приемлемым качеством. В ходе многочисленных вычислительных экспериментов обучения и тестирования нейронных сетей данного класса удавалось получить

нейронные сети, обладающие схожими показателями качества. Предложенный метод оценки качества обучения нейронной сети в условиях ограниченных данных позволил сравнить результаты многократных экспериментов при разных структурах нейронных сетей, в результате чего оказалось, что качество неполносвязных нейронных сетей в среднем превосходит качество полносвязных нейронных сетей, сопоставимых по числу искомых в процессе обучения параметров.

4.6. Выводы по главе 4

В четвёртой главе показано практическое применение и апробация методов и алгоритмов предварительной обработки ограниченных наборов данных на примере анализа медицинских сведений. В ходе экспериментов, с помощью нейропакета «Нейросимулятор 5.0» получено 3200 ИНС, среди которых: 800 ИНС, обученных на данных без их предварительной обработки, из них 7 нейронных сетей с одним скрытым слоем (50 экспериментов), 9 нейронных сетей с двумя скрытыми слоями (50 экспериментов); 2400 ИНС, обученных на данных с предварительной обработкой с такой же архитектурой, из них 7 нейронных сетей с одним скрытым слоем (150 экспериментов) и 9 нейронных сетей с двумя скрытыми слоями (150 экспериментов).

Дополнительно для сравнения в разработанном программном модуле с помощью библиотек Keras, TensorFlow обучены и протестированы 3000 ИНС, среди которых: 1500 ИНС, обученных на данных без их предварительной обработки, из них 6 нейронных сетей с одним скрытым слоем (100 экспериментов) и 9 нейронных сетей с двумя скрытыми слоями (100 экспериментов); 1500 ИНС, обученных на данных с предварительной обработкой с такой же архитектурой, из них 6 нейронных сетей с одним скрытым слоем (100 экспериментов) и 9 нейронных сетей с двумя скрытыми слоями (100 экспериментов).

Результаты, полученные в ходе экспериментов, не подтверждают гипотезу, выдвинутую на научных семинарах, о том, что качественная искусственная нейросеть может быть обучена на исходных данных без их предварительной обработки.

ЗАКЛЮЧЕНИЕ

На основе выполненного исследования можно сделать следующие выводы:

1. Предложенные правила и методы обработки ограниченных данных позволили обучить неполносвязную нейронную сеть, основанную на корне принятия решений, обладающую приемлемыми значениями точности, специфичности и чувствительности. Так, в рассмотренной задаче на первом этапе 10 показателей, используемых в качестве объясняющих переменных, были с помощью методов АСК-анализа сокращены до четырех наиболее значимых. Это позволило существенно сократить выбор структуры полного бинарного графа с поименными листьями до 15 возможных вариантов, в то время как при 10 показателях существует более 34 миллионов возможных структур. На втором этапе для выбранной структуры был идентифицирован корень принятия решений, затем на третьем этапе по его параметрам была однозначно предопределена спецификация неполносвязной нейронной сети.

2. Разработанное алгоритмическое обеспечение информационной системы проектирования и обучения искусственных нейронных сетей позволяет определять архитектуру нейронной сети в соответствии с матрицами свёртки корня принятия решений. Алгоритмическое обеспечение программно реализовано в открытой библиотеке Keras, поддерживаемой в платформе TensorFlow для глубокого обучения нейронных сетей. Это позволяет проводить многократные вычислительные эксперименты по анализу устойчивости результатов обучения неполносвязных нейронных сетей, применяя любые реализованные в Keras активационные функции, методы обучения и тестирования нейронных сетей, а также метрики качества нейронных сетей, включая кросс-валидацию.

3. Сравнение результатов экспериментов, полученных как при использовании неполносвязных, так и полносвязных нейронных сетей, показало, что обобщающая способность нейронной сети относительно полносвязных нейронных сетей с одним скрытым слоем увеличилась на 2 %, полносвязных нейронных сетей с двумя скрытыми слоями на 6 %. Предложенный метод оценки качества обучения нейронной сети в условиях ограниченных данных позволил

сравнить результаты многократных экспериментов при разных структурах нейронных сетей, в результате чего оказалось, что качество неполносвязных нейронных сетей в среднем превосходит качество полносвязных нейронных сетей, сопоставимых по числу искомых в процессе обучения параметров.

Разработанные методы и алгоритмы предобработки ограниченных данных рекомендуется использовать при создании систем поддержки принятия врачебных решений, поскольку зависимости состояния здоровья от медицинских показателей в виде корня принятия решений служат основой для построения наглядных диагностических карт, которые могут быть проинтерпретированы врачами.

Найденные с помощью предложенных методов и алгоритмов корни принятия решений могут быть использованы для описания отдельной операции в последовательности действий при диагностике заболеваний. Так, найденный в ходе настоящего исследования корень принятия решений для диагностики болезни Альцгеймера стал отличительной частью запатентованного способа диагностики болезни Альцгеймера.

Перспективным направлением исследований и разработок по данной тематике является применение предложенных методов и алгоритмов в промышленности, в особенности для мелкосерийных и опытных производств, где невозможно собрать большие объемы обучающих примеров, достаточные для применения традиционных методов математической статистики или методов обучения нейронных сетей.

СПИСОК СОКРАЩЕНИЙ И УСЛОВНЫХ ОБОЗНАЧЕНИЙ

ИИС – интеллектуальные информационные системы

ИНС – искусственные нейронные сети

МСС – корреляция Мэтьюза

МКО – механизм комплексного оценивания

ММКО – матричные механизмы комплексного оценивания

СППР – система поддержки принятия решений

АСК-анализ – автоматизированный системно-когнитивный анализ

AD – подтверждённая группа пациентов с болезнью Альцгеймера

AUC *(Area Under the ROC Curve)* – площадь под ROC-кривой

CNN – свёрточные нейронные сети

CON – контрольная группа пациентов, у которых нет болезни Альцгеймера

DRB NN – *Decisions-Root-based Neural Network*, нейронная сеть, основанная на корне принятия решений

FN – количество неверно предсказанных отрицательных классов

FP – количество неверно предсказанных положительных классов

IC_1 – корреляция абсолютных частот

IC_2 – сумма абсолютных частот по признакам объекта

IC_3 – корреляция условно-относительных частот

IC_4 – сумма условно-относительных частот по признакам объекта IC_SRK – «семантический резонанс знаний»

IC_SK – «сумма знаний»

L1 – первая мера профессора Е.В. Луценко

L2 – вторая мера профессора Е.В. Луценко

L_BV – базальные вены левого полушария головного мозга

L_DNV – вены зубчатого ядра левого полушария головного мозга

L_ICV – внутренние мозговые вены левого полушария головного мозга

L_SV – септальные вены левого полушария головного мозга

L_TV – вены таламуса левого полушария головного мозга

MAE – средняя абсолютная ошибка

MSE *(Mean Squared Error)* – среднеквадратическая ошибка

MSV – магнитная восприимчивость вен головного мозга

NFCN – нейро-нечёткая сеть управления

RMSE – корень среднеквадратической ошибки

R_BV – базальные вены правого полушария головного мозга

R_DNV – вены зубчатого ядра правого полушария головного мозга

R_ICV – внутренние мозговые вены правого полушария головного мозга

R_SV – септальные вены правого полушария головного мозга

R_TV – вены таламуса правого полушария головного мозга

PReLU – параметрический линейный или полулинейный выпрямитель *(Parametric Rectified Linear Unit)*

ReLU – линейный или полулинейный выпрямитель (*Rectified Linear Unit)*

TN – количество верно предсказанных отрицательных классов

TP – количество верно предсказанных положительных классов

СПИСОК ЛИТЕРАТУРЫ

1. Cybenko, G.V. Approximation by Superpositions of a Sigmoidal function / G.V. Cybenko // Mathematics of Control Signals and Systems. – 1989. – Vol. 2, № 4. – P. 303–314.

2. Горбань, А.Н. Обобщённая аппроксимационная теорема и вычислительные возможности нейронных сетей / А.Н. Горбань // Сибирский журнал вычислительной математики. – 1998. – Т. 1, № 1. – С. 12–24.

3. Brigato, L.A. Close Look at Deep Learning with Small Data / L.A. Brigato, Luca Iocchi // 25th International Conference on Pattern Recognition (ICPR). – 2020. – P. 2490–2497.

4. Handling a Small Dataset Problem in Prediction Model by employ Artificial Data Generation Approach: A Review / M. Lateh, A. Muda, Z. Yusof, N. Muda, M. Azmi // Journal of Physics: Conference Series. – 2017. – Vol. 892. – Art. 012016. – DOI:10.1088/1742-6596/892/1/012016.

5. Olson, M. Modern Neural Networks Generalize on Small Data Sets / M. Olson, A.J. Wyner, R.A. Berk // 32nd Conference on Neural Information Processing Systems. – Montreal, Canada, 2018.

6. Лоренц, В.А. Эффект влияния количества нейронов сети на параметры ее обучаемости / В.А. Лоренц // Вестник КГПУ им. В.П. Астафьева. – 2011. – № 4.

7. Колмогоров, А.Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных / А.Н. Колмогоров // Доклады АН СССР. – 1956. – Т. 108, № 2. – С. 179–182.

8. Колмогоров, А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного и сложения / А.Н. Колмогоров // Доклады АН СССР. – 1957. – Т. 114, № 5. – С. 953–956.

9. Арнольд, В.И. О функции трех переменных / В.И. Арнольд // Доклады АН СССР. – 1957. – Т. 114, № 4. – С. 679–681.

10. Арнольд, В.И. О представлении непрерывных функций трех переменных суперпозициями непрерывных функций двух переменных / В.И. Арнольд // Математический сборник. – 1959. – Т. 48(90), № 1. – С. 3–74.

11. Alekseev, A. Identification of Integrated Rating Mechanisms Based on Training Set / A. Alekseev // Proceedings of 2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency, SUMMA 2020, Virtual, Lipetsk, 10–13 November 2020. – IEEE, 2020. – P. 398–403. – DOI: 10.1109/SUMMA50634.2020.9280751

12. Алексеев, А.О. Представление механизмов комплексного оценивания в виде искусственных нейронных сетей / А.О. Алексеев // Математика и междисциплинарные исследования – 2020: материалы Всероссийской научно-практической конференции молодых ученых с международным участием, Пермь, 12–14 октября 2020 года. – Пермь: Пермский государственный национальный исследовательский университет, 2020. – С. 156–160.

13. Коргин, Н.А. Корни принятия решений – еще один инструмент для анализа дискретных (категориальных) наборов данных / Н.А. Коргин, В.А. Сергеев. – Текст : электронный. – URL: https://youtu.be/b4dF7znmVyo (дата обращения: 02.02.2024).

14. Quenouille, M.H. Problems in Plane Sampling / M.H. Quenouille // The Annals of Mathematical Statistics. – 1949. – Vol. 20 (3). – P. 355–375. DOI: 10.1214/aoms/1177729989

15. Tukey, J.W. Bias and confidence in not quite large samples (abstract) / J.W. Tukey // The Annals of Mathematical Statistics. – 1958. – Vol. 29 (2). – P. 614. DOI: 10.1214/aoms/1177706647

16. Efron, B. Bootstrap methods: another look at the jackknife / B. Efron // The Annals of Statistics. – Institute of Mathematical Statistics. – 1979. – Vol. 7, № 1. – P. 1–26.

17. Орлов, А.И. О реальных возможностях бутстрепа как статистического метода / А.И. Орлов // Заводская лаборатория. Диагностика материалов. – 1987. – Т. 53, № 10. – С. 82–85.

18. Орлов, А.И. Пути развития статистических методов: непараметрика, робастность, бутстреп и реалистическая статистика / А.И. Орлов // Надежность и контроль качества. – 1991. – № 8. – С. 3–8.

19. Henderson, A.R. The bootstrap: a technique for data-driven statistics. Using computer-intensive analyses to explore experimental data / A.R. Henderson // Clinica chimica acta. – 2005. – Т. 359, № 1–2. – P. 1–26.

20. Anatolyev, S.A. Robustness of residual-base bootstrap to composition of serially correlated errors / S.A. Anatolyev // Journ. Statistical Computation and Simulation. – 2009. – Vol. 79, № 3. – P. 315–320.

21. Антонов, А.В. К вопросу оценки остаточной наработки восстанавливаемого оборудования ядерных энергетических установок / А.В. Антонов, С.В. Соколов, В.А. Чепурко // Надежность. – 2011. – № 4. – С. 2–13.

22. Антонов, А.В. Интервальная оценка характеристик надежности уникального оборудования / А.В. Антонов, К.Н. Маловик, И.А. Чумаков // Фундаментальные исследования. – 2011. – № 12–1. – С. 71–76.

23. Шитиков, В.К. Использование рандомизации и бутстрепа при обработке результатов экологических наблюдений / В.К. Шитиков // Принципы экологии. – 2012. – № 1 (1). – С. 4–24.

24. Шитиков, В.К. Рандомизация и бутстреп: статистический анализ в биологии и экологии с использованием R / В.К. Шитиков, Г.С. Розенберг. – Тольятти: Кассандра, 2013. – 314 с.

25. Banjanovic, E.S. Confidence Intervals for Effect Sizes: Applying Bootstrap Resampling / E.S. Banjanovic, J.W. Osborne. – Текст : электронный // Practical Assessment, Research & Evaluation. – 2016. – Vol. 21(5). – URL: http://pareonline.net/getvn.asp?v=21&n=5 (дата обращения: 02.02.2023).

26. Белых, А.А. Методы построения доверительных интервалов с применением бутстреп-моделирования / А.А. Белых, А.В. Аталян // Прикладные проблемы дискретного анализа: сборник научных трудов / под ред. О.В. Кузьмина. – Иркутск: Иркутский государственный университет, 2021. – С. 27–32. – EDN LBMYLF.

27. Rubin, D.B. Multiple imputation for survey nonresponse / D.B. Rubin. – New York: Wiley, 1987.

28. Rubin, D.B. Discussion statistical disclosure limitation / D.B. Rubin // Journal official Statistics. – 1993. – Vol. 9(2). – P. 461.

29. Kvarnström, H. On the implementation and protection offraud detection systems: Thesis for the degree of doctor of philosophy / H. Kvarnström. – Göteborg, Sweden, 2004.

30. Construction and validation of synthetic electronic medical records / L. Moniz [et. al.] // Online J. Public Health Inform. – 2009. – Vol. 1, № 1.

31. Bowen, C.M. Comparative study of differentially private synthetic data algorithms from the NIST PSCR differential privacy synthetic data challenge / C.M. Bowen, J. Snoke // arXiv. – 2019. – 1911.12704.

32. Sistrunk, A. On synthetic data generation for anomaly detection in complex social networks / A. Sistrunk, V. Cedeno, S. Biswas // CoRR. – 2020. – Vol. abs/2010.13026

33. Awan, J. One Step to Efficient Synthetic Data / J. Awan, Z. Cai // arXiv. – 2020. – bs/2006.02397. – DOI: 10.48550/ARXIV.2006.02397

34. Traning a first-order theorem prover from synthetic data / V. Firoiu [et al.] // CoRR. – 2021. – Vol. abs/2103.03798.

35. Nikolenko, S.I. Synthetic Data for Deep Learning / S.I. Nikolenko. – Springer International Publishing, Cham, 2021. – DOI: 10.1007/978-3-030-75178-4

36. Synthetic Data - what, why and how? / J. Jordon, L. Szpruch, F. Houssiau, M. Bottarelli, G. Cherubin, C. Maple, S.N. Cohen, A. Weller. – 2022. – Текст : электронный. – URL: https://arxiv.org/abs/2205.03257 (дата обращения: 09.10.2023).

37. Rabchevsky, A.N. The Role of Synthetic Data in Improving Neural Network Algorithms / A.N. Rabchevsky, L.N. Yasnitsky // Proceeding of the 2022 4th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), Lipetsk, Russian Federation, 09-11 November 2022. – IEEE, 2022. – P. 316–320. – DOI: 10.1109/SUMMA57301.2022.9973838

38. Рабчевский, А.Н. Синтетические данные и развитие нейросетевых технологий: учебное пособие / А.Н. Рабчевский. – 1-е изд. – М.: Изд-во «Юрайт», 2023. – 187 с. – (Высшее образование).

39. Kalman, R.E. New results in linear filtering and prediction theory / R.E. Kalman, R.S. Bucy // J. of Basic Engineering. – 1961. – Vol. 83, no. 1. – P. 95–108.

40. Box, G.EP. Time Series Analysis Forecasting and Control / G.EP. Box, G.M. Jenkins. – Holden-Day, San Francisco, 1970.

41. Wilson, D. Asymptotic properties of nearest neighbor rules using edited data / D. Wilson // IEEE Transactions on Systems and Man and Cybernetics. – 1972. – Vol. 2. – P. 408–421.

42. Hickey, R.J. Noise modelling and evaluating learning from examples / R.J. Hickey // Artificial Intelligence 1996. – Vol. 82. – P. 157–179.

43. Felinger, A. Data analysis and signal processing in chromatography / A. Felinger // Data Handling in Science and Technology. – Elsevier, 1998. – Vol. 21.

44. Brodley, C.E. Identifying mislabeled training data / C.E. Brodley, M.A. Friedl // Journal of Artificial Intelligence Research. – 1999. – Vol. 11. – P. 131–167.

45. Gamberger, D. Experiments with noise filtering in a medical domain / D. Gamberger, N. Lavrac, C. Groselj // ICML. – 1999. – Vol. 99. – P. 143–151.

46. Enhancing data analysis with noise removal / H. Xiong [et al.] // IEEE transactions on knowledge and data engineering. – 2006. – Vol. 18, № 3. – P. 304–319.

47. Khoshgoftaar, T.M. Improving software quality prediction by noise filtering techniques / T.M. Khoshgoftaar, P. Rebours // Journal of Computer Science and Technology. – 2007. – Vol. 22. – P. 387–396.

48. Zhu, X. Class noise vs. attribute noise: A quantitative study / X. Zhu, X. Wu // Artificial Intelligence Review. – 2004. – Vol. 22. – P. 177–210.

49. Malossini, A. Detecting potential labeling errors in microarrays by data perturbation / A. Malossini, E. Blanzieri, R.T. Ng // Bioinformatics. – 2006. – Vol. 22. – P. 2114–2121.

50. Wu, X. Mining with noise knowledge: error-aware data mining / X. Wu, X. Zhu // IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans. – 2008. – Vol. 38, № 4. – P. 917–932.

51. Wu, Z. Ensemble empirical mode decomposition: a noise-assisted data analysis method / Z. Wu, N.E. Huang // Advances in adaptive data analysis. – 2009. – Vol. 1, № 01. – P. 1–41.

52. Каламбет, Ю.А. Фильтрация шумов: окончательное решение проблемы / Ю.А. Каламбет, С.А. Мальцев, Ю.П. Козьмин. – Текст : электронный // Анатилика. – 2011. – Вып. № 1. – URL: http://www.j-analytics.ru/journal/article/3067 (дата обращения: 06.08.2020).

53. Каламбет, Ю.А. Фильтрация шумов: Сравнительный анализ методов / Ю.А. Каламбет, Ю.П. Козьмин, А.С. Самохин. – Текст : электронный //

Аналитика. – 2017. – No 5(36). – URL: http://multichrom.ru/Docs/17-5-analitica-sravn-anal.pdf (дата обращения: 06.08.2020).

54. Frénay, B. Classification in the presence of label noise: A survey / B. Frénay, M. Verleysen // IEEE Transactions on Neural Networks and Learning Systems. – 2014. – Vol. 25. – P. 845–869.

55. Sáez, José A. On the Influence of Class Noise in Medical Data Classification: Treatment Using Noise Filtering Methods / José A. Sáez, Bartosz Krawczyk, Michał Woźniak // Applied Artificial Intelligence. – 2016. – Vol. 30(6). – P. 590–609. – DOI: 10.1080/08839514.2016.1193719

56. Куликова, М.В. Численные методы нелинейной фильтрации для обработки сигналов и измерений / М.В. Куликова, Г.Ю. Куликов // Вычислительные технологии. – 2016. – Т. 21, № 4. – С. 64–98.

57. Луценко, Е.В. Автоматизированный системно-когнитивный анализ в управлении активными объектами (системная теория информации и ее применение в исследовании экономических, социально-психологических, технологических и организационно-технических систем): монография / Е.В. Луценко. – Краснодар: КубГАУ, 2002. – 605 с.

58. Луценко, Е.В. Автоматизированный системно-когнитивный спектральный анализ конкретных и обобщенных изображений в системе «Эйдос» (применение теории информации и когнитивных технологий в спектральном анализе) / Е.В. Луценко. – Текст : электронный // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ). – Краснодар: КубГАУ, 2017. – № 04(128). – С. 1–64. – IDA [article ID]: 1281704001. – URL: http://ej.kubagro.ru/2017/04/pdf/01.pdf (дата обращения: 15.07.2022).

59. Луценко, Е.В. Синтез семантических ядер научных специальностей ВАК РФ и автоматическая классификации статей по научным специальностям с применением АСК-анализа и интеллектуальной системы «Эйдос» (на примере Научного журнала КубГАУ и его научных специальностей: механизации, агрономии и ветеринарии) / Е.В. Луценко, Н.В. Андрафанова, Н.В. Потапова //

Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2019. – № 145. – DOI: 10.21515/1990-4665-145-033

60. Луценко, Е.В. Сценарный и спектральный автоматизированный системно-когнитивный анализ: научная монография / Е.В. Луценко. – Краснодар: КубГАУ, 2021. – 288 с.

61. Свидетельство о государственной регистрации программы для ЭВМ № 2022615135, Российская Федерация. Персональная интеллектуальная онлайн среда разработки «Эйдос-X Professional (Система «ЭЙДОС-Xpro») / Е.В. Луценко // Программа для ЭВМ. Базы данных. Топологии интегральных микросхем: официальный бюллетень. – 2022. – Бюл. № 4. – 1 с.

62. Симанков, В.С. Системный анализ в адаптивном управлении / В.С. Симанков, Е.В. Луценко, В.Н. Лаптев. – Краснодар: ООО «Просвещение – Юг», 2001. – 261 с. – EDN SHXMLD.

63. Луценко, Е.В. Системно-когнитивный подход к построению многоуровневой семантической информационной модели управления агропромышленным холдингом / Е.В. Луценко, В.И. Лойко, О.А. Макаревич // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2008. – № 41. – С. 193–214. – EDN JWXYBZ.

64. Барановская, Т.П. Синтез, верификация и исследование на устойчивость системно-когнитивной модели перерабатывающего комплекса региона / Т.П. Барановская, Е.В. Луценко, В.И. Лойко. – Текст : электронный // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ). – Краснодар: КубГАУ, 2014. – № 07(101). – С. 305–333. – URl: http://ej.kubagro.ru/ (дата обращения: 10.11.2022).

65. Луценко, Е.В. Системно-когнитивный анализ объектов информационной безопасности / Е.В. Луценко, В.Е. Коржаков // Вестник Адыгейского государственного университета. Серия 4: Естественно-математические и технические науки. – 2010. – № 3.

66. Чередниченко, Н.А. Моделирование смещения полюса Земли и алгоритм прогнозирования его динамики с применением АСК-анализа / Н.А. Чередниченко, Е.В. Луценко, А.П. Трунев. – Текст : электронный // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ). – Краснодар: КубГАУ, 2014. – №05(099). – С. 149–188. – IDA [article ID]: 0991405010. – URl: http://ej.kubagro.ru/ 2014/05/ pdf/10.pdf (дата обращения: 10.11.2022).

67. Hovland, C.I. Computer simulation of thinking / C.I. Hovland // American Psychologist. – 1960. – Vol. 15(11). – P. 687–693.

68. Hunt, Earl B. Experiments in Induction / Earl B. Hunt, Janet Marin, Philip J. Stone. – New York: Academic Press, 1966.

69. Classification and Regression Trees (1st ed.) / L. Breiman, J. Friedman, R.A. Olshen, C.J. Stone. – Chapman and Hall/CRC, 1984. – DOI: 10.1201/9781315139470

70. Quinlan, J.R. Induction of decision trees / J.R. Quinlan // Machine Learning. – 1986. – Vol. 1(1). – P. 81–106.

71. Quinlan, J. Ross. C4.5: Programs for Machine learning / Quinlan J. Ross. – Morgan Kaufmann Publishers, 1993.

72. Curram, S.P. Neural Networks, Decision Tree Induction and Discriminant Analysis: An Empirical Comparison / S.P. Curram, J. Mingers // The Journal of the Operational Research Society. – 1994. – Vol. 45(4). – P. 440–450. DOI: 10.2307/2584215

73. Rokach, L. Data Mining with Decision Trees: Theory and Applications / L. Rokach, O. Maimon. – World Scientific Pub Co Inc., Singapore, 2008.

74. Сулейманова, А.Н. Обзор развития алгоритмов деревьев решений / А.Н. Сулейманова // Социология: методология, методы, математическое моделирование. – 2020. – № 50–51. – С. 64–97. – EDN ZJQQKL.

75. Построение деревьев решений и извлечение правил из обученных нейронных сетей / В.Н. Гридин [и др.] // Искусственный интеллект и принятие решений. – 2013. – № 4. – С. 26–33.

76. Кафтанников, И.Л. Особенности применения деревьев решений в задачах классификации / И.Л. Кафтанников, А.В. Парасич // Вестник Южно-

Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2015. – Т. 15, № 3. – С. 26–32.

77. Степович-Цветкова, Г.С. Метод деревьев решений для задач анализа данных / Г.С. Степович-Цветкова // Актуальные вопросы научных исследований. – 2016. – С. 48–50.

78. Галушкин, А.И. Синтез многослойных систем распознавания образов: монография / А.И. Галушкин. – М.: Энергия, 1974. – 368 с.

79. Галушкин, А.И. Нейронные сети: основы теории / А.И. Галушкин. – М.: Горячая линия-Телеком, 2010. – 496 с.

80. Горбань, А.Н. Нейронные сети на персональном компьютере / А.Н. Горбань, Д.А. Россиев. – Новосибирск: Наука: Сибир. изд. фирма РАН, 1996. – 276 с.

81. Ясницкий, Л.Н. Искусственный интеллект. Элективный курс: учеб. пособие / Л.Н. Ясницкий. – М.: БИНОМ. Лаборатория знаний, 2011. – 197 с.

82. Yasnitsky, L.N. Whether Be New “Winter” of Artificial Intelligence? / L.N. Yasnitsky // Lecture Notes in Networks and Systems. – 2020. – Vol. 78. – P. 13–17.

83. Yasnitsky, L.N. Dynamic Artificial Neural Networks as Basis for Medicine Revolution / L.N. Yasnitsky, A.A. Dumler, F.M. Cherepanov // T. Antipova, A. Rocha (eds); Digital Science. DSIC18 2018. Advances in Intelligent Systems and Computing. – 2019. – Vol. 850. – P. 351–358. – DOI: 10.1007/978-3-030-02351-5_40

84. Черепанов, Ф.М. Нейросетевой фильтр для исключения выбросов в статистической информации / Ф.М. Черепанов, Л.Н. Ясницкий // Вестник Пермского университета. Серия: Математика. Механика. Информатика. – 2008. – № 4. – С. 151–155.

85. Черепанов, Ф.М. Методы повышения эффективности нейросетевых систем в условиях ограниченных объемов выборок со сложными корреляционными связями / Ф.М. Черепанов // Прикладная математика и вопросы управления. – 2019. – № 2. – С. 40–61. – DOI: 10.15593/2499-9873/2019.2.03. – EDN JDZWEQ.

86. Черепанов, Ф.М. Методы повышения эффективности нейросетевых рекомендательных систем в условиях ограниченных объемов выборок со сложными корреляционными связями (на примере диагностики и прогнозирования сердечно-

сосудистых заболеваний человека): дис. ... канд. техн. наук: 05.13.01 / Черепанов Федор Михайлович. – Пермь: Перм. нац. иссл. гос. университет, 2019. – 136 с.

87. Лялин, В.Е. Определение параметров режимов фильтрации на графике производной давления с помощью нейронной сети / В.Е. Лялин, К.А. Сидельников // Интеллектуальные системы: материалы VIII Междунар. науч.-техн. конф. – 2010. – С. 34–40.

88. Еременко, Ю.И. О разработке метода выбора структуры нейронной сети для решения задачи адаптации параметров линейных регуляторов / Ю.И. Еременко, А.И. Глущенко // УБС. – 2016. – № 62.

89. Идентификация маркировки стальных заготовок в СПЦ-1 АО «Оскольский электрометаллургический комбинат им. А.А. Угарова» на основе нейросетевого подхода / А.В. Фомин, А.И. Глущенко, Д.А. Полещенко, И.С. Зорин // Управление большими системами: сборник трудов. – 2022. – № 95. – С. 62–78. – DOI: 10.25728/ubs.2022.95.4. – EDN QFEPJR.

90. Головко, В.А. Нейросетевые технологии обработки данных: учеб. пособие / В.А. Головко, В.В. Краснопрошин. – Минск: БГУ, 2017. – 263 с. – (Классическое университетское издание).

91. Клестов, Р.А. Алгоритмическая устойчивость нейронных сетей глубокого обучения при распознавании микроструктуры материалов / Р.А. Клестов, А.В. Клюев, В.Ю. Столбов // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. – 2021. – Т. 21, № 1. – С. 159–166.

92. Аристов, Г.В. Распознавание и классификация микроструктуры металлов и сплавов с использованием глубоких нейронных сетей / Г.В. Аристов, А.В. Клюев // GraphiCon 2017: тр. 27-й Междунар. конф. по компьютерной графике и машинному зрению, г. Пермь, 2428 сент. 2017 г. / АНО науч. о-во «Графикон», Перм. гос. нац. исслед. ун-т. – Пермь: Издат. центр ПГНИУ, 2017. – С. 180–183.

93. Горбаченко, В.И. Совершенствование градиентных алгоритмов первого порядка обучения сетей радиальных базисных функций / В.И. Горбаченко // Нейрокомпьютеры и их применение: материалы XVI Всерос. науч. конф. – 2018. – С. 245–247.

94. Горбаченко, В.И. Моделирование объектов с распределенными параметрами на нейронных сетях / В.И. Горбаченко, М.М. Алкезуини // Модели, системы, сети в экономике, технике, природе и обществе. – 2019. – № 4(32). – С. 50–64. – EDN OVCOXY.

95. Макаренко, А.В. Глубокие нейронные сети: зарождение, становление, современное состояние / А.В. Макаренко // Проблемы управления. – 2020. – № 2.

96. Backpropagation Applied to Handwritten Zip Code Recognition / Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, L. Jackel // Neural Computation. – 1989. – Vol. 1, no. 4. – P. 541–551.

97. Gradient-Based Learning Applied to Document Recognition / Y. LeCun, L. Bottou, Y. Bengio, P. Haffner // IEEE Intelligent Signal Processing. – 1998. – P. 306–351.

98. LeCun, Y. Deep Learning / Y. LeCun, Y. Bengio, G. Hinton // Nature. – 2015. – Vol. 521. – P. 436–444. – DOI: 10.1038/nature14539

99. Аведьян, Э.Д. Каскадные нейронные сети / Э.Д. Аведьян, Г.В. Баркан, И.К. Левин // Автоматика и телемеханика. – 1999. – № 3. – С. 38–54. – EDN OJYREB.

100. Deep Big Simple Neural Nets Excel on Handwritten Digit Recognition / D. Ciresan, U. Meier, L. Gambardella, J. Schmidhuber // ArXiv . – 2010. –1003.0358.

101. Krizhevsky, A. ImageNet Classification with Deep Convolutional Neural Networks / A. Krizhevsky, I. Sutskever, G. Hinton // Proc. of the 25th Int. Conference NIPS. – 2012. – Vol. 1. – P. 1097–1105.

102. Improving Neural Networks by Preventing Co-Adaptation of Feature Detectors / G. Hinton [et al.] // ArXiv. – 2012. – 1207.0580.

103. Ioffe, S. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift / S. Ioffe, C. Szegedy // ArXiv. – 2015. –1502.03167.

104. Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms / H.A.M. Williams, M.H. Jones, M. Nejati, M.J. Seabright, B.A. MacDonald // Biosystems Engineering. – 2019. – Vol. 181. – P. 140–156.

105. Wiatowski, T. Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction / T. Wiatowski, H.A. Bolcskei. – IT & EE, ETH Zurich,

Switzerland, 2016. – Текст : электронный. – URL: https://arxiv.org/abs/1512.06293 (дата обращения: 05.08.2023).

106. Lin, C.T. A Neural Fuzzy Control System with Structure and Parameter Learning / C.T. Lin // Fuzzy Sets and Systems. – 1995. – Vol. 70, iss. 2–3. – P. 183–212. – DOI: 10.1016/0165-0114(94)00216-T

107. Филист, С.А. Гибридная нейронная сеть с макрослоями для медицинских приложений / С.А. Филист, О.В. Шаталова, М.А. Ефремов // Нейрокомпьютеры: разработка, применение. – 2014. – № 6. – С. 35–69.

108. Организация хранения нейро-нечётких моделей в реляционных базах данных / В.А. Алексеев, П.В. Сараев, П.А. Домашнев, О.А. Назаркин // Управление большими системами: сборник трудов. – 2016. – № 61. – С. 226–255. – EDN WKHFJL.

109. Сараев, П.В. О структуре автоматизированных систем нейро-нечеткого моделирования ANFIS на основе больших объемов данных / П.В. Сараев, М.Г. Журавлева, В.А. Алексеев // Modern informatization problems in economics and safety: Proceedings of the XXII-th International Open Science Conference, Yelm, WA, USA, 04–11 января 2017 года / editor in Chief O.Ja. Kravets. – Yelm, WA, USA: Science Book Publishing House LLC, 2017. – С. 66–71. – EDN XRXWEZ.

110. Сараев, П.В. Перспективы применения дуально-параметрических нейронных сетей в задачах обработки и анализа данных / П.В. Сараев, Ю.Е. Полозова // Гибридные и синергетические интеллектуальные системы: материалы IV Всероссийской Поспеловской конференции с международным участием, Калининград, 14–19 мая 2018 года / под ред. А.В. Колесникова. – Калининград: Издательство Балтийского федерального университета им. Иммануила Канта, 2018. – С. 363–368. – EDN XNGBPN.

111. Бурков, В.Н. Методические основы комплексной оценки результатов деятельности предприятий с учётом их прогрессивности в ВПО «Союзэлектроприбор» / В.Н. Бурков, Н.П. Гореликов, А.М. Черкашин // Приборы и системы управления. – 1982. – № 11. – С. 21.

112. Комплексный подход к управлению научно-техническим прогрессом в отрасли / В.А. Трапезников [и др.] // Вестник АН СССР. – 1983. – № 3. – С. 33–43.

113. Глотов, В.А. Метод определения коэффициентов относительной важности / В.А. Глотов, В.В. Павельев, В.М. Гречко // Приборы и системы управления. – 1976. – № 8. – С. 73–S6.

114. Глотов, В.А. Комплексное оценивание многомерных объектов: препринт / В.А. Глотов, В.В. Павельев. / Ин-т проблем управления (автоматики и телемеханики). – М., 1984. – 27 с.

115. Глотов, В.А. Методы многокритериальной стратификации и их применение / В.А. Глотов, В.В. Павельев // II Всесоюзная конференция по статистическому и дискретному анализу нечисловой информации и экспертным оценкам: тез. докладов / Всесоюзн. науч.-исслед. ин-т системных исследований. – М.: ВНИИСИ, 1984. – С. 38–45.

116. Павельев, В.В. Вариантность структуры показателей комплексной оценки в модели целенаправленного выбора / В.В. Павельев // Теория активных систем – 2009: международная научно-практическая мультиконференция «Управление большими системами – 2009»: тр. междунар. науч.-практ. конф. (Москва, 17–19 ноября 2009 г.) / общ. ред. В.Н. Бурков. Д.А. Новиков; Ин-т проблем управления им. В.А. Трапезникова РАН. – М.: ІГПУ РАН, 2009. – Т. 1. – С. 287–290.

117. Павельев, В.В. Структурная идентификация целевой функции в задачах выбора многопараметрических объектов / В.В. Павельев // Идентификация систем и задач управления SICPRO-12: тр. IX Междунар. конф., (Москва. 30 января – 2 февраля 2012 г.) / Ин-т проблем управления им. В.А. Трапезникова РАН. – М.: ИПУ РАН, 2012. – С. 783–791.

118. Гусев, В.Б. Использование непрерывных шкал при оценивании и принятии решений в сложных проблемных ситуациях / В.Б. Гусев, В.В. Павельев; Ин-т проблем управления им. В.А. Трапезникова РАН. – М.: ИПУ РАН, 2013. – 118 с.

119. Анохин, А.М. Комплексное оценивание и оптимизация на моделях многомерных объектов / А.М. Анохин, В.Б. Гусев, В.В. Павельев; Ин-т проблем управления им. В.А. Трапезникова РАН. – М.: ИПУ РАН, 2003. – 79 с.

120. Гусев, В.Б. Согласование критериев принятия решений при целевом планировании / В.Б. Гусев // Сибирский журнал индустриальной математики. – 2005. – Т. VIII, № 2 (22). – С. 32–45.

121. Бурков, В.Н. Механизмы функционирования организационных систем / В.Н. Бурков, В.В. Кондратьев. – М.: Наука, 1981. – 384 с.

122. Большие системы: моделирование организационных механизмов / В.Н. Бурков [и др.]. – М.: Наука, 1989. – 245 с.

123. Бурков, В.Н. Механизмы повышения безопасности дорожного движения / В.Н. Бурков, В.Д. Кондратьев, А.В. Щепкин. – М.: ЛИБРОКОМ, 2012. – 208 с.

124. Бурков, В.Н. Как управлять проектами: научно-практическое издание / В.Н. Бурков, Д.А. Новиков. – М.: СПНТЕГ-ГЕО, 1997. – 188 с.

125. Бурков, В.Н. Проблемы синтеза механизма комплексного оценивания на основе обучающего набора данных / В.Н. Бурков, Н.А. Коргин, О.Л. Марин // XIII Всероссийское совещание по проблемам управления (ВСПУ–2019): сб. тр. (Москва, 17–20 июня 2019 г.) / Ин-т проблем управления им. В.А. Трапезникова РАН. – М.: ИПУ РАН, 2019. – С. 2280–2284.

126. Бурков, В.Н. Идентификация механизмов комплексного оценивания на основе унитарного кода / В.Н. Бурков, Н.А. Коргин, В.А. Сергеев // Управление большими системами: сборник трудов. – 2020. – Т. 87. – С. 67–85.

127. Korgin, N. Identification of integrated rating mechanisms on complete data sets / N. Korgin, V. Sergeev // Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems (APMS 2021) / Eds. A. Dolgui, A. Bernard, D. Lemoine, G. von Cieminski, D. Romero. – Springer: Cham, Germany, 2021. – P. 610–616. (Series book: IFIP Advances in Information and Communication Technology, vol. 630). – DOI: 10.1007/978-3-030-85874-2_66

128. Иващенко, А.А. Неманипулируемые механизмы экспертизы при неограниченных множествах возможных сообщений экспертов / А.А. Иващенко, Н.А. Коргин, Д.А. Новиков // Известия Тульского государственного университета. – 2005. – Вып. 8, ч. 2. – С. 159–165.

129. Новиков, Д.А. Нечёткие сетевые системы комплексного оценивания / Д.А. Новиков, А.Л. Суханов // Проблемы информационной экономики. Вып. 6: Моделирование инновационных процессов и экономической динамики. – М.: Ленанд, 2006. – С. 279–292.

130. Андронникова, Н.Г. Процедуры нечёткого комплексного оценивания / Н.Г. Андронникова, С.В. Леонтьев, Д.А. Новиков // Современные сложные системы управления: тр. междунар. науч.-практ. конф. – Липецк, 2002. – С. 7–8.

131. Баркалов, С.А. Об определении количества канонических матриц свёртки / С.А. Баркалов, Е.А. Власова // Вестник Воронежского государственного технического университета. – 2010. – Т. 6. – С. 151–153.

132. Харитонов, В.А. Функциональные возможности механизмов комплексного оценивания с топологической интерпретацией матриц свёртки / В.А. Харитонов, И.Р. Винокур, А.А. Белых // Управление большими системами: сборник трудов. – 2007. – № 18. – С. 129–140.

133. Интеллектуальные технологии обоснования инновационных решений: монография / В.А. Харитонов [и др.]; под науч. ред. В.А. Харитонова. – Пермь: Изд-во Перм. гос. техн. ун-та, 2010. – 342 с.

134. Белых, А.А. Обоснование направления развития механизма комплексного оценивания / А.А. Белых, В.А. Харитонов // Инновационный потенциал аграрной науки – основа развития АПК: материалы всерос. науч.-практ. конф. (Пермь, 21 ноября 2008 г.) / Перм. гос. сельскохоз. академия. – Пермь, 2008. – Ч. И. – С. 232–237.

135. Системы конструирования матриц свёртки в экспертных задачах комплексного оценивания / А.А. Белых [и др.] // Вестник «УГТУ – УПИ». Строительство и образование. – 2006. – № 12 (83). – С. 24–26.

136. Алексеев, А.О. О применении обобщенных медианных схем для матричной активной экспертизы / А.О. Алексеев, Н.А. Коргин // Прикладная математика, механика и процессы управления: материалы всерос. науч.-техн. интернет-конф. студентов и молодых ученых, г. Пермь, 30 нояб. – 5 дек. 2015 г. – Пермь: Изд-во Перм. нац. исслед. политехн. ун-та, 2016. – С. 170–177.

137. Алексеев, А.О. Матричный анонимный обобщенный медианный механизм с правом делегирования сообщений / А.О. Алексеев, Н.А. Коргин // Прикладная математика и вопросы управления. – 2016. – № 4. – С. 137–156.

138. Алексеев, А.О. Классификация механизмов комплексного оценивания сложных объектов / А.О. Алексеев // Информационные и математические технологии в науке и управлении. – 2018. – № 2 (10). – С. 106–120.

139. Алексеев, А.О. Комплексное оценивание сложных объектов в условиях неопределённости / А.О. Алексеев // Прикладная математика и вопросы управления. – 2019. – № 2. – С. 103–131.

140. Алексеев, А.О. Матричные механизмы комплексного оценивания, элементы матриц свёртки которых определены в нечетком виде / А.О. Алексеев // Управление большими системами (УБС-2017): материалы XIV Всерос. шк.-конф. молодых ученых, г. Пермь, 4–8 сент. 2017 г. / Ин-т проблем упр. им. В.А. Трапезникова РАН, Перм. нац. исслед. политехн. ун-т. – М.: [б. и.], 2017. – С. 219–233.

141. Алексеев, А.О. Аддитивно-мультипликативный матричный механизм нечёткого комплексного оценивания и эквивалентный ему непрерывный механизм / А.О. Алексеев // Современные сложные системы управления: материалы XII междунар. науч.-практ. конф. ([г. Липецк]. 25–26 октября 2017 г.) / Липец. гос. техн. ун-т. ин-т проблем управления им. В.А. Трапезникова РАН. – Липецк, 2017. – Т. 1. – С. 15–20.

142. Алексеев, А.О. Математические и инструментальные методы комплексного оценивания сложных объектов в условиях неопределённости: учеб. пособие / А.О. Алексеев. – Пермь: Изд-во Перм. нац. исслед. политехн. ун-та, 2019. – 100 с.

143. Алексеев, А.О. Об одном подходе к идентификации механизма комплексного оценивания на основе обучающего множества / А.О. Алексеев // Информационные системы и технологии: достижения и перспективы: материалы II международной конференции / Министерство образования Азербайджанской республики; Сумгаитский государственный университет, 09–10 июля 2020 г., г. Сумгаит, Азербайджанская Республика. – Сумгаит: Сумгаитский государственный университет, 2020. – С. 117–120.

144. Алексеев, А.О. Об одном подходе к идентификации механизма комплексного оценивания на основе обучающего множества: доклад /

А.О. Алексеев. – Текст : электронный. – URL: https://youtu.be/MJK1a9dycfA (дата обращения: 07.12.2022).

145. Алексеев, А.О. Представление механизмов комплексного оценивания в виде искусственных нейронных сетей / А.О. Алексеев // Математика и междисциплинарные исследования – 2020: мат. Всерос. науч.-практ. конф. мол. учен. с междунар. уч., Пермь, 12–14 октября 2020 года. – Пермь: Перм. гос. нац. исслед. ун-т, 2020. – С. 156–160. – EDN TTSEOC.

146. Alekseev, A. Identification of Integrated Rating Mechanisms with Non-serial Structures of Criteria Tree / A. Alekseev. – Текст : электронный // 2021 3rd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA): proceedings: Lipetsk State Techn. Univ., Lipetsk, Russia, Nov., 10–12 2021 / Lipetsk State Techn. Univ, IEEE, IEEE Industrial Elec-tronics Society, IEEE Industry Applications Society, IEEE Russia section. – New York: IEEE, 2021. – P. 599–603. – URL: https://ieeexplore.ieee.org/document/9632221 (дата обращения: 13.12.2022).

147. Разработка и тестирование прототипа многопользовательской виртуальной среды интеллектуального анализа данных: отчет о НИОКР / А.О. Алексеев, Л.В. Кожемякин, А.В. Маккавеев, В.Н. Никитин; Пермский центр поддержки принятия решений; Фонд содействия развитию малых форм предприятий в научно-технической сфере. Рег. № НИОКТР 122122600007-1. – Пермь : Пермский центр поддержки принятия решений, 2023. – 64 с.

148. Отчет о НИР в рамках соглашения № 075-10-2022-110/6 ПНИПУ от 08.11.2022 г. (ИГК № 000000S407522RRX0002) о направлении гранта на реализацию проектов в рамках Пермского НОЦ Мирового уровня «Рациональное недропользование» Технологический проект «Инновационные химические, медицинские и фармацевтические технологии». Подпроект «Новые материалы и технологии для медицины» : отчет о НИР / В.Ю. Столбов и др. [А.О. Алексеев, Л.В. Кожемякин и др.]. – Пермь : Пермский национальный исследовательский политехнический университет, 2022. – 387 с.

149. Отчет о НИР в рамках соглашения № 075-10-2023-113/1 ПНИПУ от 12.09.2023 г. (ИГК № 0000000007523RDZ0002) о направлении гранта на

реализацию проектов в рамках Пермского НОЦ Мирового уровня «Рациональное недропользование» подпроект «Новые материалы и технологии для медицины» : отчет о НИР / В.Ю. Столбов и др. [А.О. Алексеев, Л.В. Кожемякин и др.]. – Пермь : Пермский национальный исследовательский политехнический университет, 2023. – 310 с.

150. Кожемякин, Л.В. Применение корней принятия решений для определения взаимосвязи между магнитной восприимчивостью вен головного мозга и наличием болезни Альцгеймера / Л.В. Кожемякин, В.Н. Никитин, А.О. Алексеев // Управление большими системами: труды XVIII Всероссийской школы-конференции молодых ученых, Челябинск, 05–08 сентября 2022 года. – Челябинск: Издательский центр ЮУрГУ, 2022. – С. 318–324. – EDN CAJUES

151. Kozhemyakin, L.V. Application of Decisions' Roots for Data Analysis on Example of Dataset with Magnetic Susceptibility Values of the Brain Veins and the Alzheimer's Disease / L.V. Kozhemyakin, A.O. Alekseev, V.N. Nikitin // Proceedings of 2022 4th International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency, SUMMA 2022, Virtual, Lipetsk, 09–11 November 2022. – IEEE, 2022. – P. 425–427. – DOI: 10.1109/SUMMA57301.2022.9973977.

152. Кожемякин, Л.В. Применение искусственных нейронных сетей на ограниченном исходном наборе данных / Л.В. Кожемякин, А.О. Алексеев // Нейрокомпьютеры и их применение: сборник тезисов XXI Всероссийской научной конференции, Москва, 28 марта 2023 года. – М.: МГПУ, 2023. – С. 158–160.

153. Алексеев, А.О. Концепция многопользовательской виртуальной среды интеллектуального анализа данных / А.О. Алексеев, Л.В. Кожемякин // Управление большими системами: сборник научных трудов XIX Всероссийской школы-конференции молодых ученых, Воронеж, 05–08 сентября 2023 года. – Воронеж: Воронежский государственный технический университет, 2023. – С. 442–448. – EDN TBGVHW.

154. Алексеев, А.О. Проектирование нейронных сетей на основе дискретных механизмов комплексного оценивания / А.О. Алексеев, Л.В. Кожемякин // Математические методы в технологиях и технике. – 2023. – № 9. – С. 55–58. – DOI: 10.52348/2712-8873_MMTT_2023_9_55

155. Data Preprocessing and Neural Network Architecture Selection Algorithms in Cases of Limited Training Sets – On an Example of Diagnosing Alzheimer's Disease / A.A. Alekseev, L.V. Kozhemyakin, V.N. Nikitin, J.A. Bolshakova // Algorithms. – 2023. – Vol. 16. – Art. 219. DOI: 10.3390/a16050219

156. Кожемякин, Л.В. О качестве обучения искусственных нейронных сетей без предобработки данных в условиях их ограниченного исходного набора / Л.В. Кожемякин, А.О. Алексеев // Прикладная математика и вопросы управления / Applied Mathematics and Control Sciences. – 2023. – № 3. – С. 67–83. – DOI: 10.15593/2499-9873/2023.3.05

157. Кожемякин, Л.В. О качестве обучения основанных на корнях принятия решений неполносвязных нейронных сетей в условиях ограниченных данных / Л.В. Кожемякин, А.О. Алексеев. – Текст : электронный // Инженерный вестник Дона – 2024. – № 3. – URL: http://www.ivdon.ru/ru/magazine/archive/n3y2024/9089 (дата обращения: 05.04.2024).

158. Патент на изобретение № 2819148, Российская Федерация. Способ диагностики болезни Альцгеймера / Алексеев А.О., Кожемякин Л.В., Никитин В.Н., Большакова Ю.А. – № 2023126743: заявл. 19.10.2023: опубл. 14.05.2024.

159. Свидетельство о государственной регистрации программы для ЭВМ № 2023685360, Российская Федерация. Программный модуль проектирования и обучения нейронных сетей, основанных на корнях принятия решений (Software Module for Designing and Training Decisions Root-based Neural Networks) / Алексеев А.О., Кожемякин Л.В., Маккавеев А.В. – № 2023684000: заявл. 13.11.2023: опубл. 27.11.2023.

160. Свидетельство о государственной регистрации программы для ЭВМ № 2023688880, Российская Федерация. Программный модуль диагностики болезни Альцгеймера с помощью деревьев решений / Кожемякин Л.В. – № 2023688748: заявл. 25.12.2023; опубл. 25.12.2023.

161. Свидетельство о государственной регистрации программы для ЭВМ № 2023688880, Российская Федерация. Программный модуль для оценки качества обучения нейронных сетей / Алексеев А.О., Кожемякин Л.В., Маккавеев А.В. – № 2023688748; заявл. 25.12.2023; опубл. 20.05.2024.

162. Орлов, А.И. Прикладная статистика – состояние и перспективы / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. – 2016. – № 119. – С. 44–74. – EDN WAFKKR.

163. Metrics and scoring: quantifying the quality of predictions. – Текст : электронный. – URL: https://scikit-learn.org/stable/modules/model_evaluation.html (дата обращения: 10.11.2023).

164. Michie, D. Machine Learning, Neural and Statistical Classification / Michie D., Spiegelhalter D., Taylor Charles // Technometrics. – 1999. – Vol. 37. – 10.2307/1269742.

165. Использование методов машинного обучения при коррекции поведенческих факторов риска в профилактике сердечно-сосудистых заболеваний / Г.Г. Рапаков, Г.Т. Банщиков, В.А. Горбунов, А.В. Ударатин // Вестник Череповецкого государственного университета. – 2020. – Т. 4 (97). – С. 54–68.

166. Медников, Д.А. Интеллектуальные системы поддержки принятия врачебных решений в кардиологической практике / Д.А. Медников, З.У. Протасова, Е.В. Крикунова // Юность и Знания – гарантия успеха – 2021: сборник научных трудов 8-й Международной молодежной научной конференции (16–17 сентября 2021 г.): в 3 т. – 2021. – Т. 2. – С. 18–20.

167. XIII Всероссийское совещание по проблемам управления ВСПУ-2019: труды. 17–20 июня 2019 г., Москва / под общ. ред. Д.А. Новикова. – Электрон. текстовые дан. (617 файлов: 219 МБ). – М.: ИПУ РАН, 2019. – 3290 с. – Текст : электронный. – URL: http://elibrary.udsu.ru/xmlui/bitstream/handle/123456789/18574/519.pdf?sequence=1 (дата обращения: 05.12.2023).

168. Чугайнова, А.М. Идентификация финансовых паттернов на примере строительных организаций / А.М. Чугайнова // Инновационное развитие экономики: тенденции и перспективы. – 2021. – Т. 2. – С. 254–265. – EDN MMDQHV.

169. Калашников, А.О. Управление информационными рисками организационных систем: механизмы комплексного оценивания / А.О. Калашников // Информация и безопасность. – 2016. – Т. 19, № 3. – С. 315–322. – EDN WWYFID.

170. Суровцев, И.С. Механизмы комплексного оценивания в задачах управления образовательными комплексами / И.С. Суровцев, Н.П. Глотова // Научный вестник Воронежского государственного архитектурно-строительного университета. Серия: Управление строительством. – 2005. – № 1. – С. 17–20. – EDN VNMHQK.

171. McCulloch, W.S. A Logical Calculus of Ideas Immanent in Nervous Activity / W.S. McCulloch, W. Pitts // Bull. Mathematical Biophysics. – 1943. – Vol. 5.

172. Rosenblatt, F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain / F. Rosenblatt // Cornell Aeronautical Laboratory: Psychological Review. – 1958. – Vol. 65, no. 6. – P. 386–408.

173. Hecht-Nielsen, R. Kolmogorov's mapping neural network existence theorem / R. Hecht-Nielsen // IEEE First Annual Int. Conf. on Neural Networks, San Diego. – San Diego, 1987. – Vol. 3. – P. 11–13.

174. Теория управления (дополнительные главы): учебное пособие / под ред. Д.А. Новикова. – М.: ЛЕНАНД, 2019. – 552 с.

175. Каширина, И.Л. Исследование и сравнительный анализ методов оптимизации, используемых при обучении нейронных сетей / И.Л. Каширина, М.В. Демченко // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2018. – № 4. – С. 123–132. – EDN YTRTOP.

176. Модель выявления аномальных транзакций биткоинов на основе машинного обучения / Е.В. Фельдман, А.Н. Ручай, В.К. Матвеева, В.Д. Самсонова // Челяб. физ.-матем. журн. – 2021. – Т. 6, вып. 1. – С. 119–132.

177. Свидетельство о государственной регистрации программы для ЭВМ № 2023618561, Российская Федерация. Сборщик полных бинарных деревьев на именованных листьях на основе данных анализа групп эквивалентности / Сергеев В.А., Коргин Н.А. Зарег. 26.04.2023.

178. Свидетельство о государственной регистрации программы для ЭВМ №2021665634, Российская Федерация. Программный комплекс, реализующий работу неполносвязных нейронных сетей / Гусев А.Л., Окунев А.А. // Программа

для ЭВМ. Базы данных. Топологии интегральных микросхем: официальный бюллетень. – 2021. – Бюл. № 10. – 1 с.

179. Darwin Meets Graph Theory on a Strange Planet: Counting Full n-ary Trees with Labeled Leafs / J. Barnett, H. Carreia, P. Johnson, M. Laughlin, K. Willsom // Alabama J. Math. – 2010. – Vol. 35. – P. 16–23.

180. Верескун, В.Д. Интеллектуальный анализ данных: дискретизация значений атрибутов с использованием теории приближенных множеств и кластеризации / В.Д. Верескун, А.Н. Гуда, М.А. Бутакова // Вестник Ростовского государственного университета путей сообщения. – 2018. – № 3(71). – С. 76–84. – EDN XZOBJJ.

181. Деревья решений: общие принципы. – Текст : электронный. – URL: https://loginom.ru/blog/decision-tree-p1 (дата обращения: 11.02.2024).

182. Черепанов, Ф.М. Свидетельство о государственной регистрации программы для ЭВМ № 2014618208, Российская Федерация. Нейросимулятор 5.0 / Ф.М. Черепанов, Л.Н. Ясницкий // Программа для ЭВМ. Базы данных. Топологии интегральных микросхем: официальный бюллетень. – 2014. – Бюл. № 9. – 1 с.

183. Berrar, D. Life Sciences Help Module / D. Berrar. – Текст : электронный // Cross-Validation. – 2018. – URL: 10.1016/B978-0-12-809633-8.20349-X (дата обращения: 15.01.2024).

184. Correlation Between Cerebral Venous Oxygen Level and Cognitive Status in Patients with Alzheimer’s Disease Using Quantitative Susceptibility Mapping / Y. Liu, J. Dong, Q. Song, N. Zhang, W. Wang, B. Gao, S. Tian, C. Dong, Z. Liang, L. Xie, Y. Miao // Frontiers in Neuroscience. – 2021. – Vol. 14. – Art. 570848. – DOI: 10.3389/fnins.2020.570848. PMID: 33536866; PMCID: PMC7848136.

185. Состояние венозного церебрального кровообращения при дисциркуляторной энцефалопатии на разных стадиях заболевания / Л.В. Шагал, М.А. Барабанова, Г.Г. Музлаев, И.С. Блуменау, Е.В. Ухина // Кубанский научный медицинский вестник. – 2007. – No. 1–2. – P. 187–190.

186. Transcranial ultrafast ultrasound localization microscopy of brain vasculature in patients / C. Demené, J. Robin, A. Dizeux, B. Heiles, M. Pernot, M. Tanter, F. Perren //

Nat Biomed Eng. – 2021. – Vol. 5(3). – P. 219–228. DOI: 10.1038/s41551-021-00697-x. Epub 2021 Mar 15. PMID: 33723412; PMCID: PMC7610356.

187. Марахимов, А.Р. Повышение обобщающей способности нейронных сетей и селекция обучающих выборок / А.Р. Марахимов, К.К. Худайбергенов, Н.А. Игнатьев // Проблемы вычислительной и прикладной математики. – 2020. – № 2 (26). – С. 99–107. – EDN MBLCIX.

ПРИЛОЖЕНИЕ 1

NO	Group	L_BV	R_BV	L_ICV	R_ICV	L_TV	R_TV	L_SV	R_SV	L_DNV	R_DNV
sub001	AD	279	288	255	263	140	138	131	131	165	185
sub002	AD	274	247	223	243	239	262	190	222	204	102
sub003	AD	259	333	236	243	172	159	135	145	152	153
sub004	AD	213	252	298	301	174	154	154	135	140	127
sub005	AD	198	298	243	236	213	213	261	250	166	167
sub006	AD	249	234	256	236	142	109	139	131	158	160
sub007	AD	185	169	195	213	161	179	123	145	108	155
sub008	AD	202	312	290	292	192	232	191	200	150	153
sub009	AD	164	266	259	244	169	172	127	85	158	134
sub010	AD	211	235	234	253	206	250	103	101	198	133
sub011	AD	229	259	295	244	213	213	180	187	114	171
sub012	AD	324	250	314	293	248	193	141	122	160	153
sub013	AD	255	235	254	284	217	194	165	137	178	181
sub014	AD	287	292	343	343	227	191	152	156	128	195
sub015	AD	219	166	235	287	185	221	148	142	145	122
sub016	AD	312	232	313	307	227	208	154	182	191	170
sub017	AD	233	298	364	345	205	161	298	254	169	147
sub018	AD	280	243	305	266	273	245	197	180	180	194
sub019	AD	210	282	309	314	269	236	203	201	125	168
sub020	AD	285	246	257	263	239	232	124	73	145	147
sub021	AD	245	262	272	266	214	199	146	141	134	148
sub022	AD	292	164	394	337	239	243	199	200	137	200
sub023	AD	237	221	251	265	207	184	114	127	159	123
sub024	AD	238	285	348	338	278	282	152	150	216	199
sub025	AD	278	232	253	261	255	266	168	136	188	167
sub026	AD	270	296	244	253	243	241	200	212	150	143
sub027	AD	222	199	326	346	246	242	176	137	112	190
sub028	AD	245	260	298	303	247	233	181	168	185	149
sub029	AD	230	240	242	244	265	284	243	253	145	148
sub030	AD	218	187	230	228	219	189	310	287	122	126
sub031	AD	258	293	201	163	193	204	128	101	193	128
sub032	AD	286	268	314	330	232	204	182	133	173	194
sub033	AD	268	268	229	247	191	189	198	161	145	168
sub034	AD	278	289	246	268	286	286	146	150	176	198
sub035	AD	223	243	242	261	255	252	178	143	158	154
sub036	AD	254	287	303	299	193	197	139	151	156	145
sub037	AD	260	290	292	287	252	275	110	133	239	159
sub038	AD	245	274	299	321	266	258	165	156	180	185
sub039	AD	265	197	289	304	263	266	230	201	161	159
sub040	AD	298	315	288	291	288	278	166	159	189	179
sub041	AD	224	219	284	266	218	208	145	142	129	154
sub042	AD	236	247	298	299	177	175	167	201	197	119
sub043	AD	307	220	244	174	148	197	179	185	235	142
sub044	AD	263	265	304	307	239	249	120	182	244	269
sub045	AD	264	298	266	262	236	238	178	168	174	166
sub046	AD	238	272	317	243	194	143	163	172	131	187
sub047	AD	153	312	251	237	254	233	263	248	199	163
sub048	AD	310	231	312	325	246	265	186	201	178	160
sub049	AD	222	255	374	290	248	251	260	243	128	187
sub050	AD	274	299	283	296	235	243	120	155	226	185

Окончание Таблицы

NO	Group	L_BV	R_BV	L_ICV	R_ICV	L_TV	R_TV	L_SV	R_SV	L_DNV	R_DNV
sub051	AD	262	353	357	411	244	284	215	198	184	200
sub052	AD	252	280	257	230	269	224	237	218	217	178
sub053	AD	186	257	331	289	221	223	108	120	205	141
sub054	AD	235	240	236	267	209	219	133	148	149	146
sub055	AD	301	269	266	234	241	241	213	206	202	89
sub056	AD	265	208	286	299	268	276	183	196	152	199
sub057	AD	216	301	301	312	147	140	129	144	128	132
sub058	AD	287	278	203	232	207	180	155	149	156	168
sub059	AD	240	231	256	263	150	186	156	167	134	142
sub060	CON	220	259	271	272	255	233	144	145	141	151
sub061	CON	181	195	208	216	173	172	148	144	134	163
sub062	CON	216	223	247	222	145	151	126	122	95	146
sub063	CON	257	329	293	316	229	193	155	175	154	144
sub064	CON	194	208	224	282	211	169	135	147	112	81
sub065	CON	209	223	238	263	188	174	156	124	123	88
sub066	CON	204	199	320	291	224	239	184	167	94	135
sub067	CON	174	213	214	208	131	142	83	99	165	154
sub068	CON	225	357	310	343	262	251	160	183	133	176
sub069	CON	288	298	183	186	143	152	205	210	123	121
sub070	CON	259	235	218	242	151	209	147	149	121	132
sub071	CON	297	257	286	281	206	202	165	123	160	121
sub072	CON	212	232	250	231	165	197	130	134	165	133
sub073	CON	243	318	322	321	218	203	157	179	133	191
sub074	CON	187	218	233	290	226	205	145	148	123	151
sub075	CON	212	224	240	253	198	231	159	139	145	148
sub076	CON	213	200	311	299	234	240	175	169	199	195
sub077	CON	179	222	210	212	134	134	110	100	173	154
sub078	CON	223	346	313	341	267	259	165	178	167	177
sub079	CON	279	288	179	189	141	153	245	214	177	149
sub080	CON	249	232	221	232	154	142	165	151	161	143
sub081	CON	300	259	295	299	216	197	163	131	162	129

ПРИЛОЖЕНИЕ 2

Сводная Таблица достоверности статистических моделей
с двумя интегральными критериями IC_SK и IC_SRK при разбиении
на два равных интервала с разным числом наблюдений

№	Модель	Интегральный критерий (IC)	N	TP	FN	FP	TN	Точность модели	Полнота модели	F	L1	L2
1	ABS	IC_1	81	69	12	59	22	0,539	0,852	0,660	0,691	0,593
1	ABS	IC_2	81	81	0	81	0	0,500	10	0,667	0,754	0,754
2	PRC1	IC_3	81	69	12	59	22	0,539	0,852	0,660	0,691	0,593
2	PRC1	IC_4	81	81	0	81	0	0,500	10	0,667	0,673	0,673
3	PRC2	IC_3	81	69	12	59	22	0,539	0,852	0,660	0,691	0,593
3	PRC2	IC_4	81	81	0	81	0	0,500	10	0,667	0,673	0,673
4	INF1	IC_SRK	81	43	38	38	43	0,531	0,531	0,531	0,589	0,559
4	INF1	IC_SK	81	47	34	25	56	0,653	0,580	0,614	0,628	0,508
5	INF2	IC_SRK	81	43	38	38	43	0,531	0,531	0,531	0,589	0,559
5	INF2	IC_SK	81	47	34	25	56	0,653	0,580	0,614	0,628	0,508
6	INF3	IC_SRK	81	44	37	37	44	0,543	0,543	0,543	0,599	0,557
6	INF3	IC_SK	81	44	37	37	44	0,543	0,543	0,543	0,599	0,557
7	INF4	IC_SRK	81	43	38	38	43	0,531	0,531	0,531	0,585	0,554
7	INF4	IC_SK	81	49	32	32	49	0,605	0,605	0,605	0,657	0,556
8	INF5	IC_SRK	81	43	38	38	43	0,531	0,531	0,531	0,585	0,554
8	INF5	IC_SK	81	49	32	32	49	0,605	0,605	0,605	0,657	0,556
9	INF6	IC_SRK	81	44	37	37	44	0,543	0,543	0,543	0,599	0,557
9	INF6	IC_SK	81	44	37	37	44	0,543	0,543	0,543	0,601	0,558
10	INF7	IC_SRK	81	44	37	37	44	0,543	0,543	0,543	0,599	0,557
10	INF7	IC_SK	81	44	37	37	44	0,543	0,543	0,543	0,601	0,558

Сводная Таблица достоверности статистических моделей
с двумя интегральными критериями IC_SK и IC_SRK при разбиении
на четыре равных интервала с разным числом наблюдений

№	Модель	Интегральный критерий (IC)	N	TP	FN	FP	TN	Точность модели	Полнота модели	F	L1	L2
1	ABS	IC_1	81	76	5	66	15	0,535	0,938	0,682	0,733	0,666
1	ABS	IC_2	81	81	0	81	0	0,500	10	0,667	0,768	0,768
2	PRC1	IC_3	81	76	5	66	15	0,535	0,938	0,682	0,733	0,666
2	PRC1	IC_4	81	81	0	81	0	0,500	10	0,667	0,687	0,687
3	PRC2	IC_3	81	76	5	66	15	0,535	0,938	0,682	0,733	0,666
3	PRC2	IC_4	81	81	0	81	0	0,500	10	0,667	0,687	0,687
4	INF1	IC_SRK	81	57	24	19	62	0,750	0,704	0,726	0,828	0,647
4	INF1	IC_SK	81	52	29	19	62	0,732	0,642	0,684	0,779	0,614
5	INF2	IC_SRK	81	57	24	19	62	0,750	0,704	0,726	0,828	0,647
5	INF2	IC_SK	81	52	29	19	62	0,732	0,642	0,684	0,779	0,614
6	INF3	IC_SRK	81	57	24	24	57	0,704	0,704	0,704	0,821	0,659
6	INF3	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,821	0,659
7	INF4	IC_SRK	81	45	36	10	71	0,818	0,556	0,662	0,776	0,636
7	INF4	IC_SK	81	57	24	26	55	0,687	0,704	0,695	0,797	0,637
8	INF5	IC_SRK	81	45	36	10	71	0,818	0,556	0,662	0,776	0,636
8	INF5	IC_SK	81	57	24	26	55	0,687	0,704	0,695	0,797	0,637
9	INF6	IC_SRK	81	55	26	18	63	0,753	0,679	0,714	0,829	0,659
9	INF6	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,799	0,626
10	INF7	IC_SRK	81	55	26	18	63	0,753	0,679	0,714	0,829	0,659
10	INF7	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,799	0,626

Сводная Таблица достоверности статистических моделей
с двумя интегральными критериями IC_SK и IC_SRK при разбиении
на два разных интервала с равным числом наблюдений

№	Модель	Интегральный критерий (IC)	N	TP	FN	FP	TN	Точность модели	Полнота модели	F	L1	L2
1	ABS	IC_1	81	48	33	32	49	0,600	0,593	0,596	0,669	0,578
1	ABS	IC_2	81	81	0	81	0	0,500	10	0,667	0,762	0,762
2	PRC1	IC_3	81	48	33	32	49	0,600	0,593	0,596	0,669	0,578
2	PRC1	IC_4	81	81	0	81	0	0,500	10	0,667	0,679	0,679
3	PRC2	IC_3	81	49	32	32	49	0,605	0,605	0,605	0,669	0,569
3	PRC2	IC_4	81	81	0	81	0	0,500	10	0,667	0,679	0,679
4	INF1	IC_SRK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
4	INF1	IC_SK	81	43	38	25	56	0,632	0,531	0,577	0,603	0,514
5	INF2	IC_SRK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
5	INF2	IC_SK	81	43	38	25	56	0,632	0,531	0,577	0,603	0,514
6	INF3	IC_SRK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
6	INF3	IC_SK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
7	INF4	IC_SRK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
7	INF4	IC_SK	81	49	32	32	49	0,605	0,605	0,605	0,627	0,523
8	INF5	IC_SRK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
8	INF5	IC_SK	81	49	32	32	49	0,605	0,605	0,605	0,627	0,523
9	INF6	IC_SRK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
9	INF6	IC_SK	81	49	32	32	49	0,605	0,605	0,605	0,628	0,524
10	INF7	IC_SRK	81	49	32	32	49	0,605	0,605	0,605	0,668	0,568
10	INF7	IC_SK	81	49	32	32	49	0,605	0,605	0,605	0,628	0,524

Сводная Таблица достоверности статистических моделей
с двумя интегральными критериями IC_SK и IC_SRK при разбиении
на три разных интервала с равным числом наблюдений

№	Модель	Интегральный критерий (IC)	N	TP	FN	FP	TN	Точность модели	Полнота модели	F	L1	L2
1	ABS	IC_1	81	54	27	29	52	0,651	0,667	0,659	0,750	0,608
1	ABS	IC_2	81	81	0	81	0	0,500	10	0,667	0,771	0,771
2	PRC1	IC_3	81	54	27	29	52	0,651	0,667	0,659	0,750	0,608
2	PRC1	IC_4	81	81	0	81	0	0,500	10	0,667	0,693	0,693
3	PRC2	IC_3	81	54	27	29	52	0,651	0,667	0,659	0,750	0,608
3	PRC2	IC_4	81	81	0	81	0	0,500	10	0,667	0,693	0,693
4	INF1	IC_SRK	81	53	28	30	51	0,639	0,654	0,646	0,744	0,613
4	INF1	IC_SK	81	48	33	18	63	0,727	0,593	0,653	0,700	0,542
5	INF2	IC_SRK	81	53	28	30	51	0,639	0,654	0,646	0,744	0,613
5	INF2	IC_SK	81	48	33	18	63	0,727	0,593	0,653	0,700	0,542
6	INF3	IC_SRK	81	53	28	28	53	0,654	0,654	0,654	0,749	0,612
6	INF3	IC_SK	81	53	28	28	53	0,654	0,654	0,654	0,749	0,612
7	INF4	IC_SRK	81	53	28	28	53	0,654	0,654	0,654	0,748	0,611
7	INF4	IC_SK	81	54	27	27	54	0,667	0,667	0,667	0,736	0,582
8	INF5	IC_SRK	81	53	28	28	53	0,654	0,654	0,654	0,748	0,611
8	INF5	IC_SK	81	54	27	27	54	0,667	0,667	0,667	0,736	0,582
9	INF6	IC_SRK	81	53	28	28	53	0,654	0,654	0,654	0,749	0,612
9	INF6	IC_SK	81	53	28	28	53	0,654	0,654	0,654	0,729	0,587
10	INF7	IC_SRK	81	53	28	28	53	0,654	0,654	0,654	0,749	0,612
10	INF7	IC_SK	81	53	28	28	53	0,654	0,654	0,654	0,729	0,587

Сводная Таблица достоверности статистических моделей
с двумя интегральными критериями IC_SK и IC_SRK при разбиении
на четыре разных интервала с равным числом наблюдений

№	Модель	Интегральный критерий (IC)	N	TP	FN	FP	TN	Точность модели	Полнота модели	F	L1	L2
1	ABS	IC_1	81	59	22	22	59	0,728	0,728	0,728	0,810	0,613
1	ABS	IC_2	81	81	0	81	0	0,500	10	0,667	0,778	0,778
2	PRC1	IC_3	81	59	22	23	58	0,720	0,728	0,724	0,810	0,618
2	PRC1	IC_4	81	81	0	81	0	0,500	10	0,667	0,703	0,703
3	PRC2	IC_3	81	59	22	22	59	0,728	0,728	0,728	0,810	0,613
3	PRC2	IC_4	81	81	0	81	0	0,500	10	0,667	0,703	0,703
4	INF1	IC_SRK	81	58	23	29	52	0,667	0,716	0,690	0,802	0,643
4	INF1	IC_SK	81	49	32	14	67	0,778	0,605	0,681	0,766	0,594
5	INF2	IC_SRK	81	58	23	29	52	0,667	0,716	0,690	0,802	0,643
5	INF2	IC_SK	81	49	32	14	67	0,778	0,605	0,681	0,766	0,594
6	INF3	IC_SRK	81	57	24	24	57	0,704	0,704	0,704	0,805	0,634
6	INF3	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,805	0,634
7	INF4	IC_SRK	81	57	24	24	57	0,704	0,704	0,704	0,806	0,637
7	INF4	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,792	0,616
8	INF5	IC_SRK	81	57	24	24	57	0,704	0,704	0,704	0,806	0,637
8	INF5	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,792	0,616
9	INF6	IC_SRK	81	57	24	24	57	0,704	0,704	0,704	0,805	0,634
9	INF6	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,790	0,613
10	INF7	IC_SRK	81	57	24	24	57	0,704	0,704	0,704	0,805	0,634
10	INF7	IC_SK	81	57	24	24	57	0,704	0,704	0,704	0,790	0,613

ПРИЛОЖЕНИЕ 3

Принятые интервалы значений
магнитной восприимчивости вен головного мозга для модели *INF1*
при разбиении на два равных интервала с разным числом наблюдений

Обозначение	Наблюдаемая область значений MSV	1-й интервал	2-й интервал
L_BV	{153,0; 324,0}	{153,0; 238,5}	{238,5; 324,0}
L_ICV	{179,0; 394,0}	{179,0; 286,5}	{286,5; 394,0}
L_TV	{131,0; 288,0}	{131,0; 209,5}	{209,5; 288,0}
L_SV	{83,0; 310,0}	{83,0; 196,5}	{196,5; 310,0}
L_DNV	{94,0; 244,0}	{94,0; 169,0}	{169,0; 244,0}
R_BV	{164,0; 357,0}	{164,0; 260,5}	{260,5; 357,0}
R_ICV	{163,0; 411,0}	{163,0; 287,0}	{287,0; 411,0}
R_TV	{109,0; 286,0}	{109,0; 197,5}	{197,5; 286,0}
R_SV	{73,0; 287,0}	{73,0; 180,0}	{180,0; 287,0}
R_DNV	{81,0; 269,0}	{81,0; 175,0}	{175,0; 269,0}

Принятые интервалы значений
магнитной восприимчивости вен головного мозга для модели *INF1*
при разбиении на четыре равных интервала с разным числом наблюдений

Обозначение	Наблюдаемая область значений MSV	1-й интервал	2-й интервал	3-й интервал	4-й интервал
L_BV	{153,0; 324,0}	{153,0; 195,75}	{195,75; 238,5}	{238,5; 281,25}	{281,25; 324,0}
L_ICV	{179,0; 394,0}	{179,0; 232,75}	{235,75; 286,5}	{286,5; 340,25}	{340,25; 394,0}
L_TV	{131,0; 288,0}	{131,0; 170,25}	{170,25; 209,5}	{209,5; 248,75}	{248,75; 288,0}
L_SV	{83,0; 310,0}	{83,0; 139,75}	{139,75; 196,5}	{196,5; 253,25}	{253,25; 310,0}
L_DNV	{94,0; 244,0}	{94,0; 131,5}	{131,5; 169,0}	{169,0; 206,5}	{206,5; 244,0}
R_BV	{164,0; 357,0}	{164,0; 212,25}	{212,25; 260,5}	{260,5; 308,75}	{308,75; 357,0}
R_ICV	{163,0; 411,0}	{163,0; 225,0}	{225,0; 287,0}	{287,0; 349,0}	{349,0; 411,0}
R_TV	{109,0; 286,0}	{109,0; 153,25}	{153,25; 197,5}	{197,5; 241,75}	{241,75; 286,0}
R_SV	{73,0; 287,0}	{73,0; 126,5}	{126,5; 180,0}	{180,0; 233,5}	{233,5; 287,0}
R_DNV	{81,0; 269,0}	{81,0; 128,0}	{128,0; 175,0}	{175,0; 222,0}	{222,0; 269,0}

Принятые интервалы значений
магнитной восприимчивости вен головного мозга для модели *INF3*
при разбиении на два разных интервала с равным числом наблюдений

Обозначение	Наблюдаемая область значений MSV	1-й интервал	2-й интервал
L_BV	{153,0; 324,0}	{153,0; 243,0}	{243,0; 324,0}
L_ICV	{179,0; 394,0}	{179,0; 266,0}	{266,0; 394,0}
L_TV	{131,0; 288,0}	{131,0; 218,0}	{218,0; 288,0}
L_SV	{83,0; 310,0}	{83,0; 159,0}	{159,0; 310,0}
L_DNV	{94,0; 244,0}	{94,0; 158,0}	{158,0; 244,0}
R_BV	{164,0; 357,0}	{164,0; 255,0}	{255,0; 357,0}
R_ICV	{163,0; 411,0}	{163,0; 266,0}	{266,0; 411,0}
R_TV	{109,0; 286,0}	{109,0; 208,0}	{208,0; 286,0}
R_SV	{73,0; 287,0}	{73,0; 151,0}	{151,0; 287,0}
R_DNV	{81,0; 269,0}	{81,0; 153,0}	{153,0; 269,0}

Принятые интервалы значений
магнитной восприимчивости вен головного мозга для модели *INF4*
при разбиении на три разных интервала с равным числом наблюдений

Обозначение	Наблюдаемая область значений MSV	1-й интервал	2-й интервал	3-й интервал
L_BV	{153,0; 324,0}	{153,0; 223,0}	{223,0; 262,0}	{262,0; 324,0}
L_ICV	{179,0; 394,0}	{179,0; 246,0}	{246,0; 295,0}	{295,0; 394,0}
L_TV	{131,0; 288,0}	{131,0; 198,0}	{198,0; 239,0}	{239,0; 288,0}
L_SV	{83,0; 310,0}	{83,0; 146,0}	{146,0; 178,0}	{178,0; 310,0}
L_DNV	{94,0; 244,0}	{94,0; 145,0}	{145,0; 173,0}	{173,0; 244,0}
R_BV	{164,0; 357,0}	{164,0; 234,0}	{234,0; 274,0}	{274,0; 357,0}
R_ICV	{163,0; 411,0}	{163,0; 253,0}	{253,0; 292,0}	{292,0; 411,0}
R_TV	{109,0; 286,0}	{109,0; 193,0}	{193,0; 236,0}	{236,0; 286,0}
R_SV	{73,0; 287,0}	{73,0; 143,0}	{143,0; 178,0}	{178,0; 287,0}
R_DNV	{81,0; 269,0}	{81,0; 143,7}	{143,7; 206,3}	{206,3; 269,0}

Принятые интервалы значений

магнитной восприимчивости вен головного мозга для модели *INF4*

при разбиении на четыре разных интервала с равным числом наблюдений

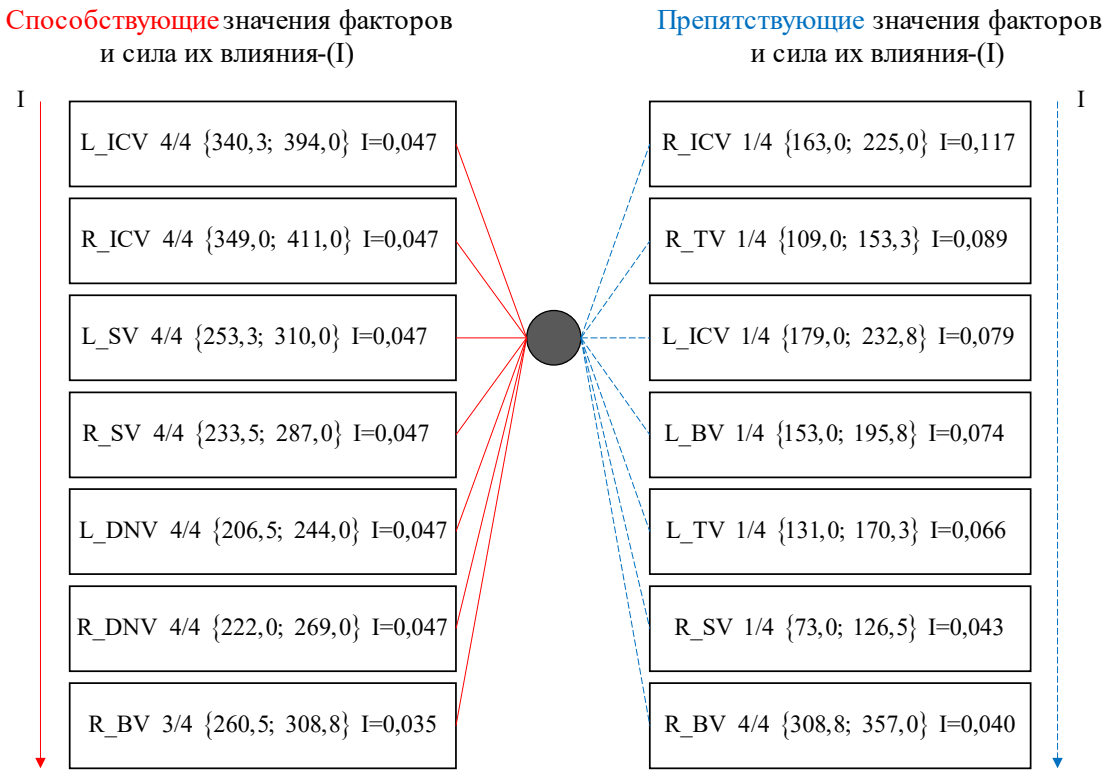
Обозначение	Наблюдаемая область значений MSV	1-й интервал	2-й интервал	3-й интервал	4-й интервал
L_BV	{153,0; 324,0}	{153,0; 216,0}	{216,0; 243,0}	{243,0; 270,0}	{270,0; 324,0}
L_ICV	{179,0; 394,0}	{179,0; 238,0}	{238,0; 266,0}	{266,0; 303,0}	{303,0; 394,0}
L_TV	{131,0; 288,0}	{131,0; 185,0}	{185,0; 218,0}	{218,0; 246,0}	{246,0; 288,0}
L_SV	{83,0; 310,0}	{83,0; 139,0}	{139,0; 159,0}	{159,0; 183,0}	{183,0; 310,0}
L_DNV	{94,0; 244,0}	{94,0; 134,0}	{134,0; 158,0}	{158,0; 178,0}	{178,0; 244,0}
R_BV	{164,0; 357,0}	{164,0; 224,0}	{224,0; 255,0}	{255,0; 288,0}	{288,0; 357,0}
R_ICV	{163,0; 411,0}	{163,0; 243,0}	{243,0; 266,0}	{266,0; 299,0}	{299,0; 411,0}
R_TV	{109,0; 286,0}	{109,0; 180,0}	{180,0; 208,0}	{208,0; 242,0}	{242,0; 286,0}
R_SV	{73,0; 287,0}	{73,0; 136,0}	{136,0; 151,0}	{151,0; 185,0}	{185,0; 287,0}
R_DNV	{81,0; 269,0}	{81,0; 141,0}	{141,0; 153,0}	{153,0; 176,0}	{176,0; 269,0}

ПРИЛОЖЕНИЕ 4

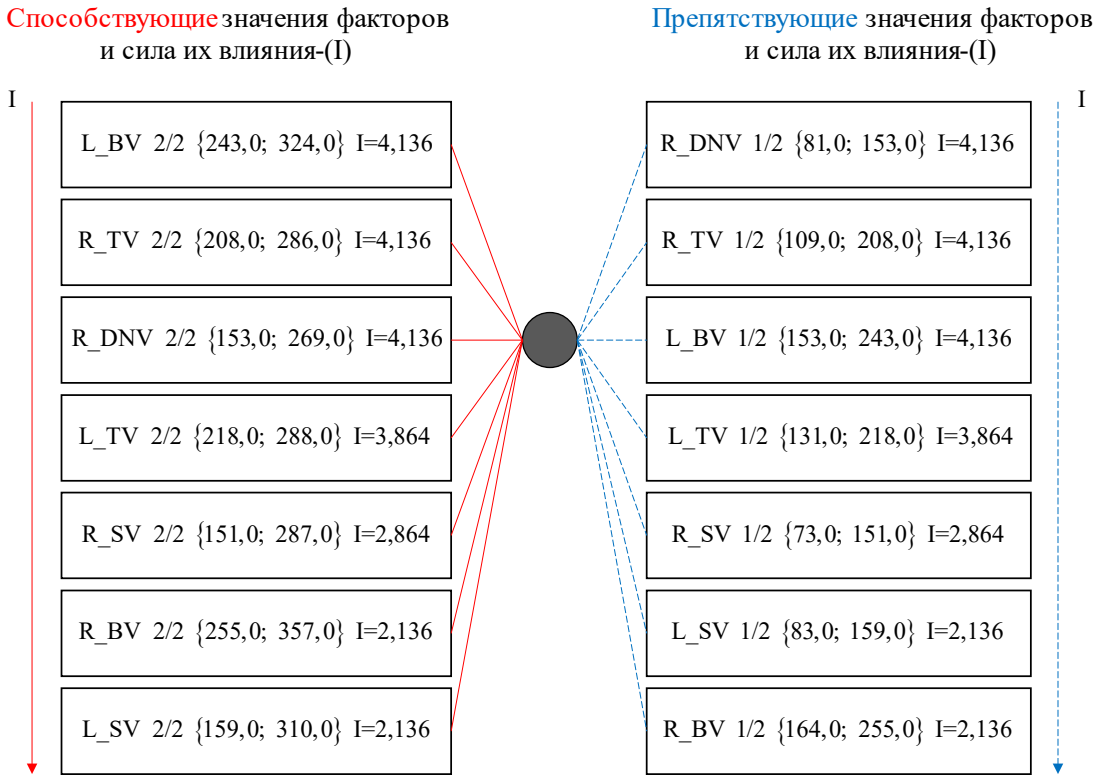
Интервалы: **2 равных**
Число наблюдений в интервалах: **разное**



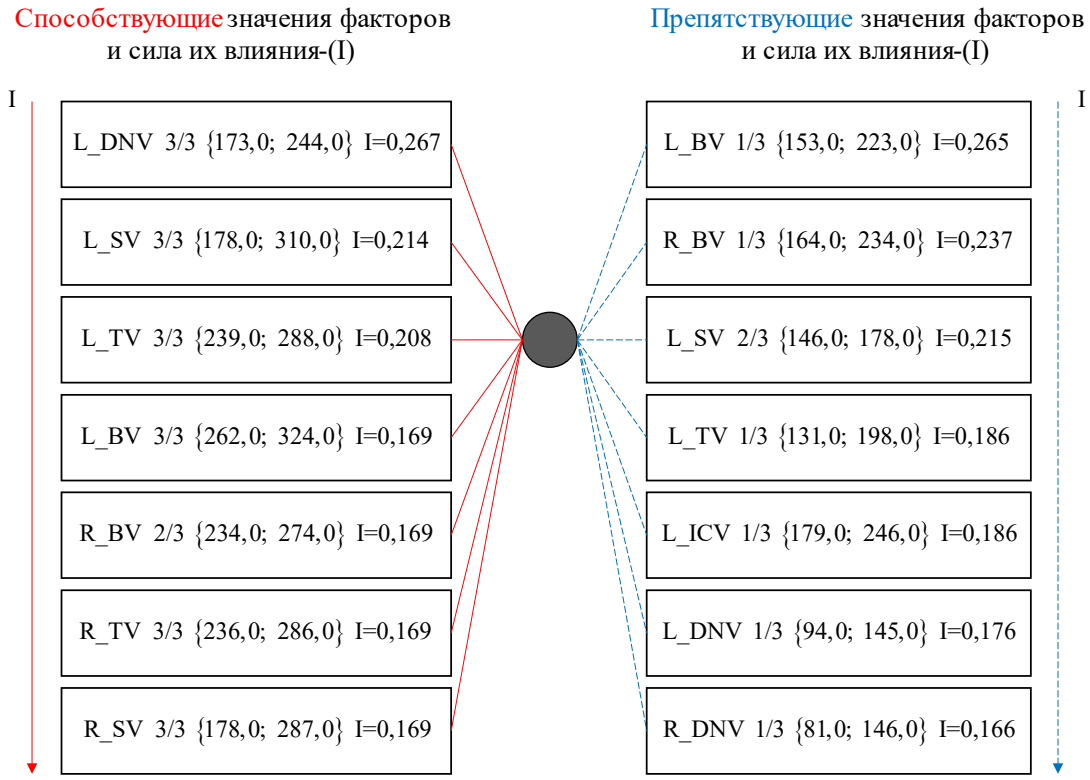
Интервалы: **4 равных**
Число наблюдений в интервалах: **разное**



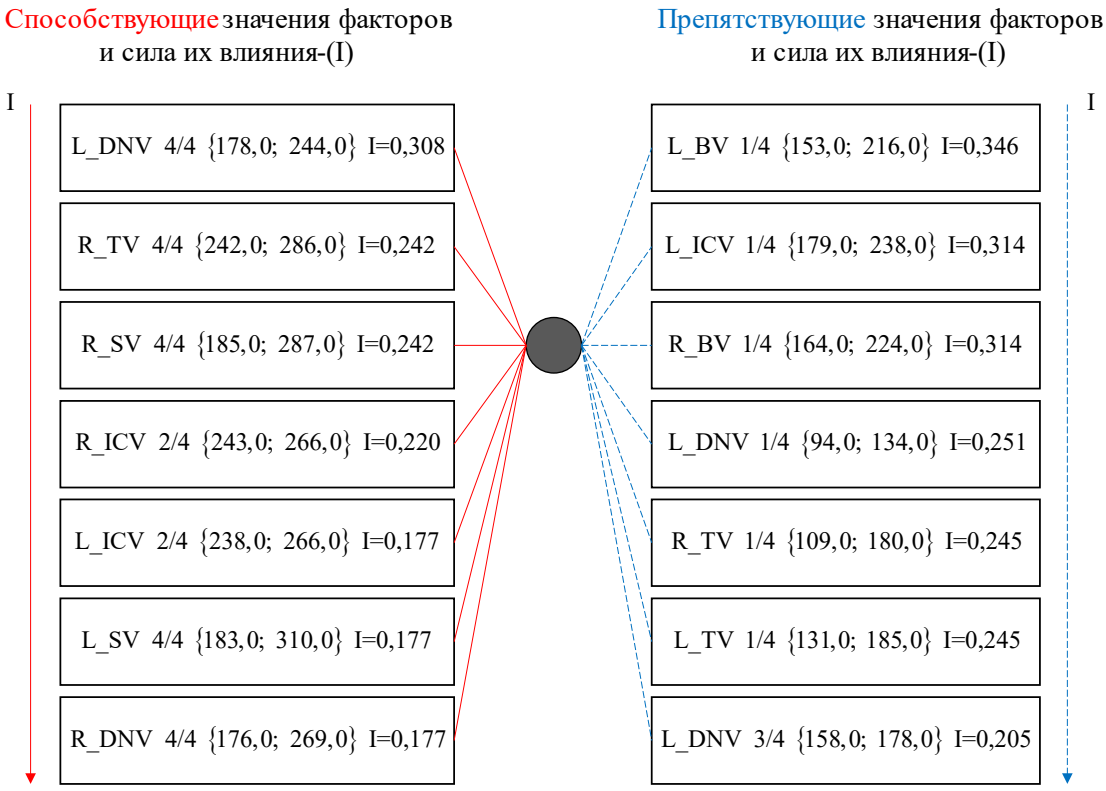
Интервалы: **2 разных**
Число наблюдений в интервалах: **равное**



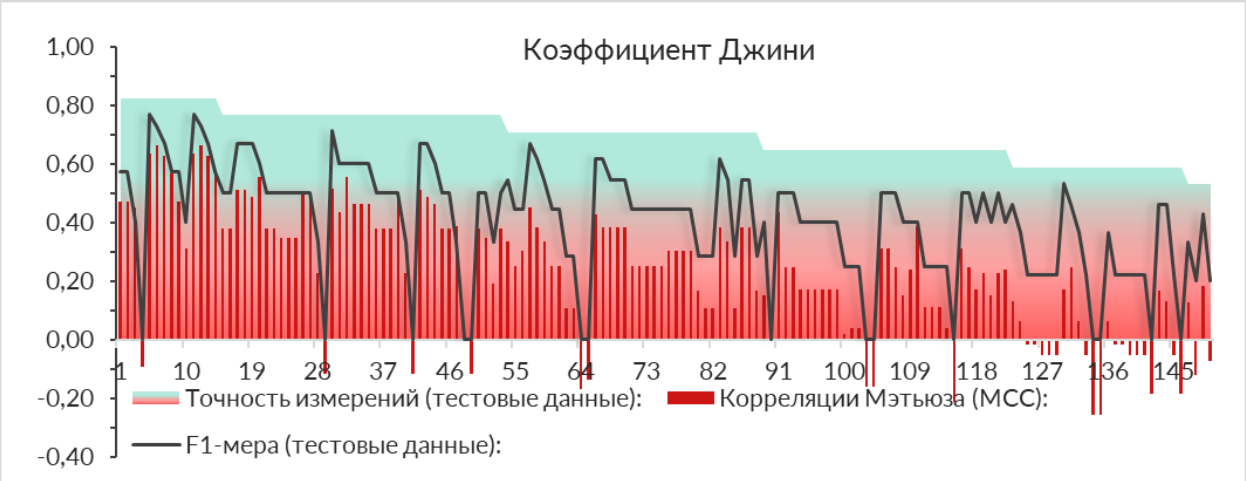
Интервалы: **3 разных**
Число наблюдений в интервалах: **равное**



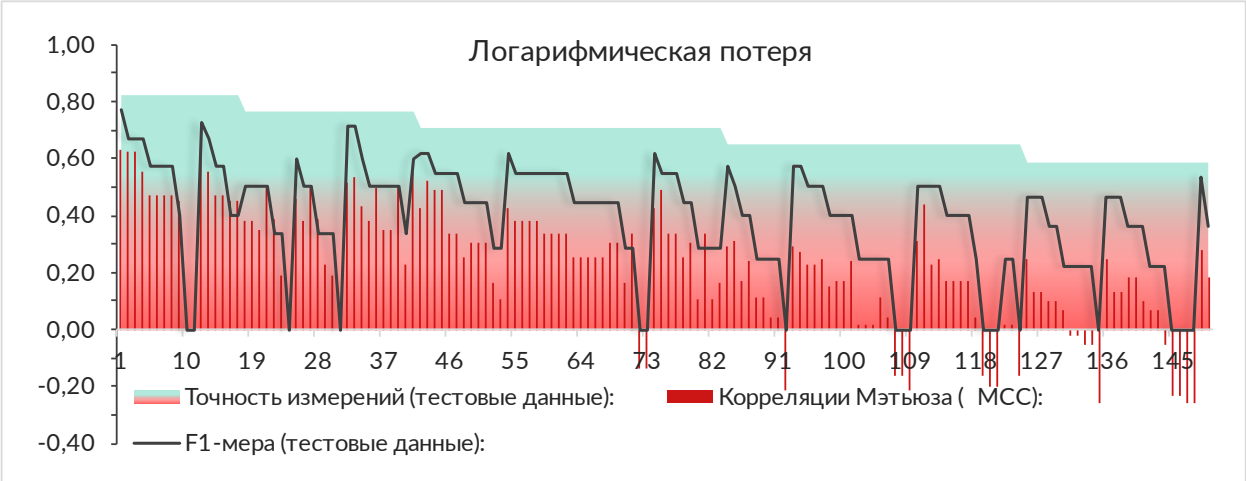
Интервалы: **4 разных**
Число наблюдений в интервалах: **равное**



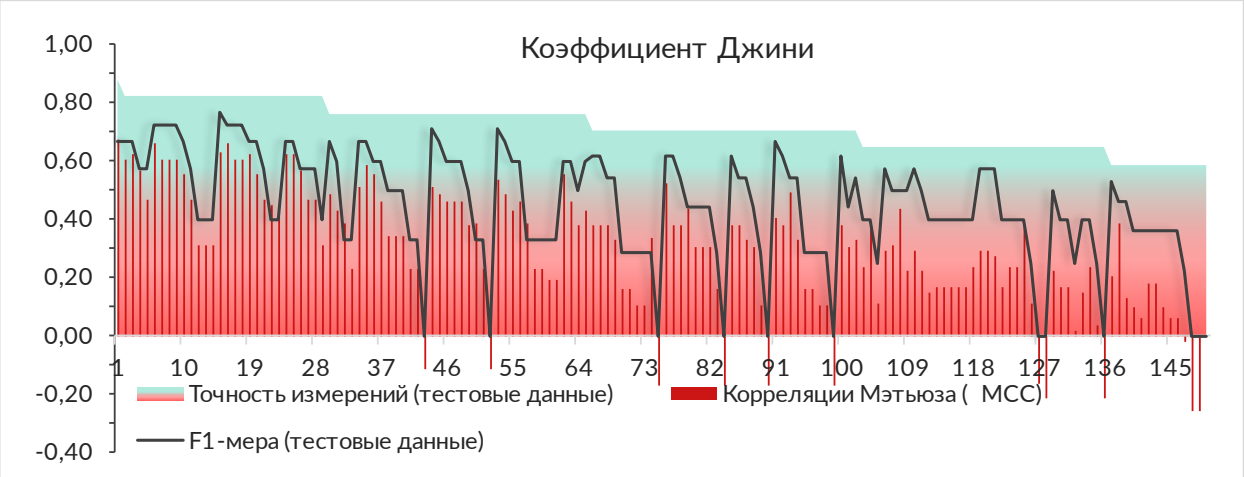
ПРИЛОЖЕНИЕ 5



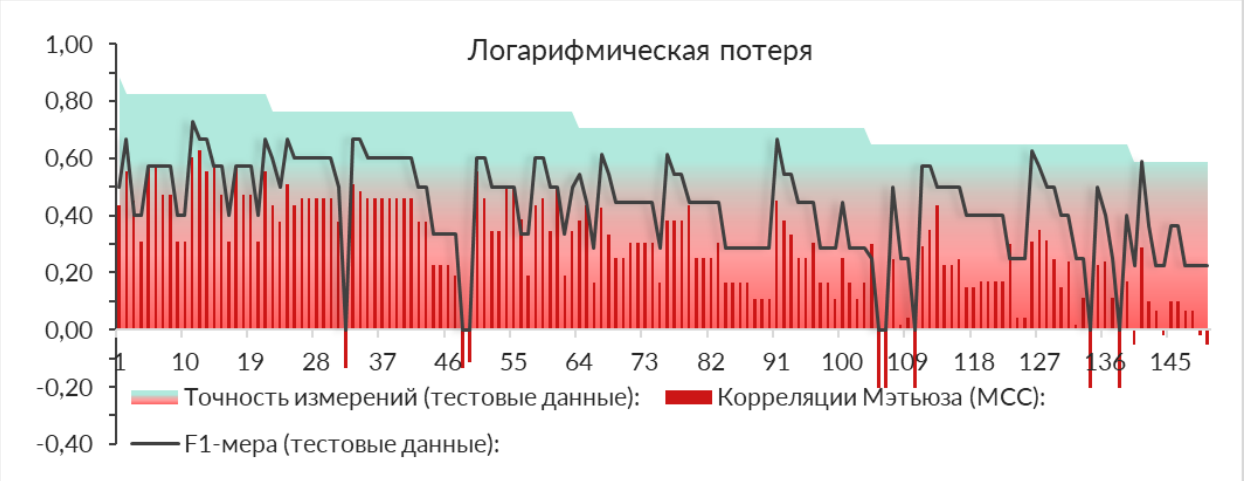
Графическое представление качества обученных 150 деревьев решений на данных без их предобработки, где для разбиения узлов деревьев решений выбран коэффициент Джини



Графическое представление качества обученных 150 деревьев решений на данных без их предобработки, где для разбиения узлов деревьев решений выбрана логарифмическая потеря



Графическое представление качества обученных 150 деревьев решений на предобработанных данных, где для разбиения узлов деревьев решений выбран коэффициент Джини



Графическое представление качества обученных 150 деревьев решений на предобработанных данных, где для разбиения узлов деревьев решений выбрана логарифмическая потеря

ПРИЛОЖЕНИЕ 6

РОССИЙСКАЯ ФЕДЕРАЦИЯ

ПАТЕНТ

НА ИЗОБРЕТЕНИЕ

№ 2819148

СПОСОБ ДИАГНОСТИКИ БОЛЕЗНИ АЛЬЦГЕЙМЕРА

Патентообладатель: ***федеральное государственное автономное образовательное учреждение высшего образования "Пермский национальный исследовательский политехнический университет" (RU)***

Авторы: ***Алексеев Александр Олегович (RU), Кожемякин Леонид Валерьевич (RU), Никитин Владислав Николаевич (RU), Большакова Юлия Алексеевна (RU)***

Заявка № **2023126743**
Приоритет изобретения **19 октября 2023 г.**
Дата государственной регистрации
в Государственном реестре изобретений
Российской Федерации **14 мая 2024 г.**
Срок действия исключительного права
на изобретение истекает **19 октября 2043 г.**

Руководитель Федеральной службы по интеллектуальной собственности

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат 429b6a0fe3853164baf96f83b73b4aa7
Владелец Зубов Юрий Сергеевич
Действителен с 10.05.2023 по 02.08.2024

Ю.С. Зубов

ПРИЛОЖЕНИЕ 7

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2023685360

Программный модуль проектирования и обучения нейронных сетей, основанных на корнях принятия решений (Software Module for Designing and Training Decisions Root-based Neural Networks)

Правообладатель: ***ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ "ПЕРМСКИЙ ЦЕНТР ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ" (RU)***

Авторы: ***Алексеев Александр Олегович (RU), Кожемякин Леонид Валерьевич (RU), Маккавеев Александр Валерьевич (RU)***



Заявка № **2023684000**

Дата поступления **13 ноября 2023 г.**

Дата государственной регистрации в Реестре программ для ЭВМ ***27 ноября 2023 г.***

Руководитель Федеральной службы по интеллектуальной собственности

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат 429b6a0fe3853164baf96f83b73b4aa7
Владелец **Зубов Юрий Сергеевич**
Действителен с 10.05.2023 по 02.08.2024

Ю.С. Зубов

ПРИЛОЖЕНИЕ 8

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2023688880

«Программный модуль диагностики болезни Альцгеймера с помощью деревьев решений»

Правообладатель: ***Федеральное государственное автономное образовательное учреждение высшего образования «Пермский национальный исследовательский политехнический университет» (ПНИПУ) (RU)***

Автор(ы): ***Кожемякин Леонид Валерьевич (RU)***



Заявка № **2023688748**
Дата поступления **25 декабря 2023 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ ***25 декабря 2023 г.***

Руководитель Федеральной службы по интеллектуальной собственности

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат 429b6a0fe3853164baf96f83b73b4aa7
Владелец **Зубов Юрий Сергеевич**
Действителен с 10.05.2023 по 02.08.2024

Ю.С. Зубов

ПРИЛОЖЕНИЕ 9

РОССИЙСКАЯ ФЕДЕРАЦИЯ

СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2024662382

Программный модуль для оценки качества обучения нейронных сетей

Правообладатель: ***Федеральное государственное автономное образовательное учреждение высшего образования «Пермский национальный исследовательский политехнический университет» (ПНИПУ) (RU)***

Авторы: ***Кожемякин Леонид Валерьевич (RU), Алексеев Александр Олегович (RU), Маккавеев Александр Олегович (RU)***

Заявка № **2024661802**
Дата поступления **28 мая 2024 г.**
Дата государственной регистрации
в Реестре программ для ЭВМ ***28 мая 2024 г.***

Руководитель Федеральной службы по интеллектуальной собственности

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
Сертификат 429b6a0fe3853164baf96f83b73b4aa7
Владелец Зубов Юрий Сергеевич
Действителен с 10.05.2023 по 02.08.2024

Ю.С. Зубов

ПРИЛОЖЕНИЕ 10

Министерство промышленности и торговли Пермского края

Автономная некоммерческая организация
«Пермский научно-образовательный центр мирового уровня
«Рациональное недропользование»
(АНО Пермский НОЦ)
614013, г. Пермь, ул. Куйбышева, 14, телефон/факс (342) 239-10-40,
ОГРН 1195958032753, ИНН 5902055649

Справка об использовании результатов
диссертации Кожемякина Леонида Валерьевича
при выполнении фундаментальных научных исследований
по Технологическому проекту «Инновационные химические,
медицинские и фармацевтические технологии»,
по подпроекту «Новые материалы и технологии для медицины»

Настоящей справкой подтверждаем, что Кожемякин Леонид Валерьевич участвовал в 2022 году в подготовке двух разделов отчета о научно-исследовательской работе в рамках соглашения № 075-10-2022-110/6 ПНИПУ от 08.11.2022 г. (ИГК № 000000S407522RRX0002), в частности:

– выполнил литературный обзор исследований и публикаций в области прогнозирования болезни Альцгеймера;

– разработал математические модели для диагностирования болезни Альцгеймера в виде искусственных нейронных сетей с одним и двумя скрытыми слоями.

В 2023 году Кожемякин Леонид Валерьевич участвовал в подготовке двух разделов отчета о научно-исследовательской работе в рамках соглашения № 075-10-2023-113/1 ПНИПУ от 12.09.2023 г. (ИГК № 0000000007523RDZ0002), в частности:

– разработана программа для ЭВМ для диагностики болезни Альцгеймера с помощью деревьев решений. В результате обучения и тестирования деревьев решений на основе данных о магнитной восприимчивости вен головного мозга 81 пациента, были получены 150 вариантов деревьев решений;

– проведена серия вычислительных экспериментов по обучению и тестированию полносвязных искусственных нейронных сетей с предобработкой данных и без предобработки данных;

– получены результаты обучения и тестированию искусственной нейронной сети класса Decisions Root-based Neural Networks для диагностики болезни Альцгеймера. Точность полученной нейронной сети составила 89%, среднеквадратичная ошибка 11%.

Директор АНО «Пермский НОЦ»
«Рациональное недропользование» П. Ю. Илюшин