

## ОТЗЫВ

официального оппонента, доктора технических наук, профессора, заведующего кафедрой «Компьютерные технологии» федерального государственного бюджетного образовательного учреждения высшего образования «Пензенский государственный университет»

Горбаченко Владимира Ивановича

на диссертационную работу Черепанова Федора Михайловича «Методы повышения эффективности нейросетевых рекомендательных систем в условиях ограниченных объемов выборок со сложными корреляционными связями (на примере диагностики и прогнозирования сердечно-сосудистых заболеваний человека)» по специальности 05.13.01 – Системный анализ, управление и обработка информации (технические и информационные системы)

### **1. Актуальность темы диссертационного исследования**

Диссертационное исследование посвящено разработке методов повышения эффективности инструмента нейросетевого моделирования в случае, когда невозможно собрать достаточно большой массив данных, необходимых для построения качественной нейронной сети, а также в случае, если моделируемый процесс имеет сильные скрытые взаимосвязи входных переменных модели. Предложены методы для выявления аномальных примеров, определения значимых параметров нейросетевой модели, настройки чувствительности получаемых нейросетевых моделей к ошибкам первого или второго рода, а также несколько методов повышения точности прогнозирования развития во времени процессов с использованием нейросетевых моделей.

Указанные выше проблемы возникают практически в каждой задаче моделирования на основе данных, в том числе с применением нейронных сетей, в связи с чем публикуется множество научных работ, посвященных попыткам их решения. Однако в данной диссертационной работе рассматриваются круг задач с ограничениями в виде малого объема исходных данных и сложных внутренних связей между входными параметрами моделируемых объектов или процессов. Одной из таких задач является диагностика и прогнозирование заболеваний сердечно-сосудистой системы человека, и на её примере выполняется тестирование эффективности предложенных методов и разработанной информационной системы, реализованной на основе аппарата нейросетевого моделирования с применением этих методов (далее – нейросетевая система).

Таким образом, диссертационная работа Черепанова Ф.М., посвященная повышению эффективности нейросетевых систем в условиях ограниченных

объемов выборок со сложными корреляционными связями, является актуальной и носит практический характер.

## **2. Обоснованность и достоверность полученных результатов**

Основные результаты, полученные Черепановым Ф.М. базируются на научно обоснованных методах исследования, и подтверждаются теоретическими исследованиями, сравнительными тестами. Для каждого разработанного метода предложена адекватная методика тестирования, включающая сравнение с результатами существующих методов на тех же наборах данных. Показано преимущество предложенных методов по сравнению с существующими в рассматриваемых условиях.

Метод выявления аномальных наблюдений с использованием нейросети, базируется на том факте, что присутствие в обучающей выборке сильного отклонения влияет на сходимость процесса обучения нейронной сети. Похожий подход используется в группе методов анализа удаленных остатков, которые применяются в регрессионном анализе для решения той же задачи, что отмечено в работе. Также в работе приводится обзор других методов выявления аномальных наблюдений с указанием их достоинств и недостатков. Эффективность предложенного метода сравнивалась с такими методами выявления выбросов как: расстояний Махаланобиса, DFFITS для множественной линейной и нелинейной регрессионных моделей. Сравнение показало эффективность предложенного в диссертационной работе метода по сравнению с остальными, заключающуюся в получении нейросетевой модели с более низким уровнем ошибки при удалении из обучающего множества выявленных аномальных наблюдений каждым из методов.

Метод вычисления информативности входных параметров посредством анализа нейросетевой модели основан на сравнении ошибки нейросетевой модели при поочередном «исключении» входных параметров. Под исключением понимается удаление данного параметра из модели, либо замена исходных значений данного параметра на случайные или специально подобранные, некоррелированные с изначальными, значения. Данный подход основан на положениях теории информации и используется в нескольких методах вычисления информативности входных параметров нейросетевой модели, обзор которых приведен в тексте диссертации. Группа данных методов отличается между собой процедурой исключения входных параметров, и предложенный в диссертационной работе метод отличается тем что для данной процедуры используется случайное перемешивание значений входного параметра, что не нарушает структуру сети, и статистическое распределение данных, исключая влияние этих характеристик на процедуру оценки информативности. Это

позволяет получать более точные оценки, что подтверждается результатами сравнения данного метода с другими методами вычисления информативности на искусственно созданных и реальных наборах данных. Сравнение проводилось как с классическим коэффициентом корреляции Пирсона, так и с другими наиболее точными (на основании данных обзора публикаций) методами оценки информативности нейросетевых моделей. Процедуры сравнения эффективности обоснованы и грамотно построены.

Метод настройки чувствительности алгоритмов обучения нейросети имеет математическое обоснование. Эффективность данного метода оценивается путем сравнения чувствительности нейросетевой модели при выявлении пациентов с риском заболевания «Инфаркт миокарда», используя предложенный метод, с методом сдвига порога классификации «болен\здоров», при фиксированном уровне специфичности. Также в обзоре упоминается метод получения нейросетевых моделей с заданными операционными характеристиками, путем обучения на двух специально подготовленных выборках, однако сравнение с ним не производится, вероятно, по причине отсутствия подходящего набора данных.

Метод экспертной коррекции прогноза нейросетевой модели дополняет результаты оценки развития заболевания, полученные посредством изменения отдельных входных параметров, подаваемых на вход нейронной сети, экспертными знаниями о том как может развиваться данный процесс во времени. В диссертационной работе рассматривается случай нейросетевых моделей диагностирования заболеваний сердечно-сосудистой системы человека, а в качестве экспертных знаний используется Европейская шкала SCORE, показывающая риск наступления смерти от сердечно-сосудистого заболевания в зависимости от возраста, пола и пристрастия к курению. Тестирование метода проводилось на тестовом множестве, дополненном фактическими данными о развитии заболеваний у пациентов, прошедших повторное обследование. В ходе тестирования проводилось сравнение изначальной погрешности нейросетевой модели, погрешности нейросетевой модели с увеличением параметра возраст, и погрешности при увеличении параметра возраст с последующим применением метода экспертной коррекции. Результаты тестирования показали повышение точности прогнозирования при применении метода экспертной коррекции для нескольких видов заболеваний сердечно-сосудистой системы. Кроме того, в публикациях, посвященных этому методу приводятся результаты применения данного метода и их обсуждение с профессиональными врачами-кардиологами, и делается вывод об их адекватности. Данный метод разработан в соавторстве, что указано в тексте диссертации.

Метод скользящего окна для повышения точности прогнозирования с использованием нейросетевой модели также служит для повышения точности прогнозирования развития процессов сложных объектов во времени с использованием нейросетевых моделей. Данный метод предлагает применение методологии скользящего окна для получения множества обучающих примеров, которое используется при обучении нейросетевой модели способной непосредственно выдавать прогноз на будущие периоды времени значений моделируемого процесса. В диссертационной работе приводится описание исследования, где данный подход успешно применяется для прогнозирования скорости ветра и делается вывод о возможности его применения для прогнозирования развития сложных нелинейных процессов. В диссертационной работе Черепанова Ф.М. приводится процедура формирования множества обучающих примеров для многомерного входного вектора на примере задачи диагностики заболеваний, делается развитие данного метода для получения нейросетевых моделей предназначенных для выдачи рекомендаций, и приводится описание процедуры их применения. Далее приводятся результаты тестирования данного метода для прогнозирования развития нескольких форм сердечно-сосудистых заболеваний, при этом результаты сравниваются с диагнозом, выставленным врачом по тем же пациентам, но на основании уточненных данных. Также погрешность прогнозирования моделей, полученных с применением метода скользящего окна, сравнивается с результатами тестирования предыдущего метода повышения точности прогнозирования. Результаты сравнения показывают меньшую погрешность метода скользящего окна, что объясняется способностью нейросетевой модели учитывать индивидуальные особенности каждого конкретного пациента.

Предложенные методы реализованы в виде программных модулей и включены в состав нейросетевой системы диагностики и прогнозирования рисков заболеваний сердечно-сосудистой системы человека. Создание системы осуществлялось с применением методологии объектно-ориентированного проектирования. В диссертационной работе описаны этапы проектирования и обоснование полученных на каждом из этапов артефактов. Отдельная глава посвящена обоснованию выбора средств разработки. Разработанная нейросетевая система была протестирована на тестовом наборе данных, показав свою адекватность, а также протестирована в ходе внедрения в медицинскую практику (имеется акт о внедрении с результатами тестирования).

На основании результатов тестирования отдельных методов и системы в целом делаются выводы о более высокой эффективности данных методов в условиях малого объема выборки обучающих данных и сложных внутренних связях между входными параметрами процесса, по сравнению с другими

аналогичными методами. Таким образом, полученные Черепановым Ф.М. результаты научно обоснованы, а достоверность подтверждена результатами тестирования и опытной эксплуатации разработанной нейросетевой системы.

### **3. Новизна и значимость результатов диссертационного исследования**

Новизну и значимость результатов диссертационного исследования для науки представляют:

1. Метод выявления аномальных наблюдений с использованием нейросети, отличающийся применением нейросетевой модели в сочетании с анализом удаленных остатков, позволяющий повысить точность обнаружения примеров, препятствующих успешному обучению нейросетевой модели (соответствует п. 12 «Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации» паспорта специальности 05.13.01).
2. Метод вычисления информативности входных параметров посредством анализа нейросетевой модели, отличающийся методом замены значений входных параметров, позволяющий более точно оценить степень влияния входных параметров на моделируемый процесс (соответствует п. 12 «Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации» паспорта специальности 05.13.01).
3. Метод настройки чувствительности алгоритмов обучения нейросети к ошибкам первого и второго рода, отличающийся применением модифицированной функции потерь, и позволяющий получать нейросетевые модели с требуемыми характеристиками чувствительности к соответствующим ошибкам (соответствует п. 5 «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации» паспорта специальности 05.13.01).
4. Метод экспертной коррекции и метод скользящего окна, позволяющие повысить точность прогнозирования развития процессов со сложными внутренними связями с применением нейронных сетей (соответствует п. 12 «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации» и п.13 «Методы получения, анализа и обработки экспертной информации» паспорта специальности 05.13.01).

Практическая значимость диссертационного исследования заключается в реализации предложенных методов в виде программного кода, а также в разработке действующей нейросетевой системы диагностики и прогнозирования

заболеваний сердечно-сосудистой системы человека, что подтверждено свидетельством о государственной регистрации программы для ЭВМ и актом о внедрении (соответствует п. 5 «Разработка специального математического и алгоритмического обеспечения систем анализа, оптимизации, управления, принятия решений и обработки информации» паспорта специальности 05.13.01). Кроме того, полученные научные результаты могут применяться при создании иных информационных систем, в которых используется методология нейросетевого моделирования. Всё вышесказанное позволяет сделать вывод о том, что данное диссертационное исследование обладает научной новизной, а также научной и практической значимостью.

#### **4. Соответствие диссертационного исследования установленным критериям**

Диссертационное исследование Черепанова Ф.М. является завершённой научно-квалификационной работой и содержит новые научные результаты и положения. Судя по оригинальности работы, на основе проверки текста диссертации с помощью сервиса «Антиплагиат», можно с уверенностью считать диссертационную работу Черепанова Ф.М. самостоятельно выполненной. Текст работы хорошо структурирован и написан научным языком, работа грамотно оформлена.

Результаты, полученные Черепановым Ф.М. в ходе выполнения диссертационной работы «Методы повышения эффективности нейросетевых рекомендательных систем в условиях ограниченных объемов выборок со сложными корреляционными связями (на примере диагностики и прогнозирования сердечно-сосудистых заболеваний человека)» обоснованы, и имеют практическую значимость, которая заключается в создании действующей нейросетевой системы диагностики и прогнозирования заболеваний сердечно-сосудистой системы человека, что подтверждено свидетельством о государственной регистрации программы для ЭВМ и актом о внедрении. Помимо этого, полученные результаты используются в учебном процессе, что также подтверждено соответствующими актами.

Положения, выносимые на защиту аргументированы, предложенные методы протестированы на различных наборах данных и сопоставлены с иными актуальными методами. Результаты тестирования разработанной нейросетевой системы отображены в акте внедрения и сопоставлены с заключениями врачей-экспертов, сделанными на основании данных дополнительных лабораторных и инструментальных исследований.

Автор ссылается на отечественные и зарубежные публикации, заимствования корректны, и имеют соответствующие ссылки на источник. Результаты, полученные в соавторстве, имеют соответствующую отметку.

Основные результаты исследований опубликованы в шести статьях, опубликованных в изданиях, входящих в Перечень рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, из них четыре в изданиях, индексируемых в Scopus.

### **5. Достоинства и недостатки диссертационной работы**

К достоинствам диссертационной работы можно отнести широту исследования, включающего обзор и решение нескольких проблем, присущих методу нейросетевого моделирования и попытку выработки универсального подхода при их решении. Необходимо также отметить применение предложенных методов при разработке успешно действующей нейросетевой системы диагностики и прогнозирования сердечно-сосудистых заболеваний, что носит практический и социально-значимый характер.

К недостаткам диссертационной работы следует отнести следующие замечания:

1. Термин "рекомендательные системы" имеет устоявшееся значение в искусственном интеллекте и машинном обучении, отличное от того, которое подразумевается в диссертации.
2. Рис.2 – это фактически отбор признаков методом обертки (wrapper method). Новизна не подчеркнута.
3. Когда говорят об ошибках первого и второго рода, чувствительности и специфичности, то обычно, рассматривают выбор точки отсечения, F-метрику, ROC-анализ. В автореферате не показаны преимущества предлагаемого подхода.

Стоит отметить, что указанные недостатки не снижают ценности научных и практических результатов, полученных в ходе выполнения диссертационного исследования.

### **Заключение по работе**

Диссертационная работа Черепанова Федора Михайловича «Методы повышения эффективности нейросетевых рекомендательных систем в условиях ограниченных объемов выборок со сложными корреляционными связями (на примере диагностики и прогнозирования сердечно-сосудистых заболеваний человека)» является законченной научно-квалификационной работой, в которой содержится решение сложной и актуальной задачи, связанной с повышением

эффективности нейросетевых систем в условиях ограниченных объемов выборок со сложными корреляционными связями.

Диссертация соответствует паспорту специальности 05.13.01 «Системный анализ, управление и обработка информации (технические и информационные системы)», содержит новые научные и практические результаты и отвечает требованиям к диссертациям на соискание ученой степени кандидата наук, установленным п. 9 действующего «Положения о порядке присуждения ученых степеней ВАК РФ», утвержденным Постановлением Правительства РФ от 24 сентября 2013 г. № 842. Автор диссертационной работы Черепанов Федор Михайлович заслуживает присуждения ему ученой степени кандидата технических наук по специальности 05.13.01 «Системный анализ, управление и обработка информации (технические и информационные системы)».

### Официальный оппонент

Доктор технических наук по специальности 05.13.15 «Вычислительные машины, комплексы и компьютерные сети», профессор,  
заведующий кафедрой «Компьютерные технологии»  
Пензенского государственного университета

  
Горбаченко Владимир Иванович

« 27 » 08 2019 г.

Телефон: +7 (987) 526-80-68  
e-mail: [gorvi@mail.ru](mailto:gorvi@mail.ru)

Подпись Горбаченко В.И. заверяю:  
Ученый секретарь Ученого совета  
ФГБОУ ВО ПГУ, к.т.н., доцент

  
О.С. Дорофеева

« 28 » 08 2019 г.

Адрес организации:  
440026, г. Пенза, ул. Красная 40, ФГБОУ ВО «Пензенский государственный университет»  
Телефон: +7 (8412) 56-35-11, e-mail: [rector@pnzgu.ru](mailto:rector@pnzgu.ru)