

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования



**Пермский национальный исследовательский
политехнический университет**

УТВЕРЖДАЮ

Проректор по образовательной
деятельности

 А.Б. Петроченков

« 06 » марта 20 23 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Дисциплина: Интеллектуальный анализ данных
(наименование)

Форма обучения: очная
(очная/очно-заочная/заочная)

Уровень высшего образования: магистратура
(бакалавриат/специалитет/магистратура)

Общая трудоёмкость: 180 (5)
(часы (ЗЕ))

Направление подготовки: 09.04.04 Программная инженерия
(код и наименование направления)

Направленность: Разработка программно-информационных систем
(наименование образовательной программы)

1. Общие положения

1.1. Цели и задачи дисциплины

Раскрывается понятие интеллектуального анализа данных как части рынка информационных технологий, рассматриваются основные методологические вопросы применения инструментов Data Mining. Подробно излагаются задачи и методы интеллектуального анализа, а также различные аспекты практического применения Data Mining.

1.2. Изучаемые объекты дисциплины

Понятие интеллектуального анализа данных. Задачи Data Mining
Задача кластеризации.

1.3. Входные требования

Не предусмотрены

2. Планируемые результаты обучения по дисциплине

Компетенция	Индекс индикатора	Планируемые результаты обучения по дисциплине (знать, уметь, владеть)	Индикатор достижения компетенции, с которым соотнесены планируемые результаты обучения	Средства оценки
ОПК-3	ИД-1ОПК-3	Знает методы получения новых знаний и умений, в том числе в новых областях знаний, связанных с профессиональной деятельностью; порядок поиска, систематизации и оценки достоверности научно-технической информации из различных источников, в т.ч. с использованием информационных технологий	Знает методы получения новых знаний и умений, в том числе в новых областях знаний, связанных с профессиональной деятельностью; порядок поиска, систематизации и оценки достоверности научно-технической информации из различных источников, в т.ч. с использованием информационных технологий	Экзамен
ОПК-3	ИД-2ОПК-3	Умеет, в том числе с помощью информационных технологий приобретать новые знания, расширять свое мировоззрение	Умеет, в том числе с помощью информационных технологий приобретать новые знания, расширять свое мировоззрение	Защита лабораторной работы
ОПК-3	ИД-3ОПК-3	Владеет информационно-коммуникационными технологиями в сфере профессиональной деятельности	Владеет информационно-коммуникационными технологиями в сфере профессиональной деятельности	Защита лабораторной работы

3. Объем и виды учебной работы

Вид учебной работы	Всего часов	Распределение по семестрам в часах	
		Номер семестра	
		3	
1. Проведение учебных занятий (включая проведение текущего контроля успеваемости) в форме:	72	72	
1.1. Контактная аудиторная работа, из них:			
- лекции (Л)	18	18	
- лабораторные работы (ЛР)	24	24	
- практические занятия, семинары и (или) другие виды занятий семинарского типа (ПЗ)	26	26	
- контроль самостоятельной работы (КСР)	4	4	
- контрольная работа			
1.2. Самостоятельная работа студентов (СРС)	72	72	
2. Промежуточная аттестация			
Экзамен	36	36	
Дифференцированный зачет			
Зачет			
Курсовой проект (КП)			
Курсовая работа (КР)			
Общая трудоемкость дисциплины	180	180	

4. Содержание дисциплины

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
3-й семестр				
Основы интеллектуального анализа данных.	6	8	8	24
Понятие интеллектуального анализа данных Data Mining как часть рынка информационных технологий Набор данных и их атрибутов Задачи Data Mining Основы анализа данных				
Методы интеллектуального анализа данных.	6	8	8	24
Методы интеллектуального анализа данных Задачи классификации и прогнозирования Деревья решений Задача регрессии Задача кластеризации Алгоритм k-средних (k-means) Метод косинусов Поиск ассоциативных правил Задача визуализации многомерных данных Метод иерархий целей				

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
Процесс применения интеллектуальных технологий	6	8	10	24
Основные этапы интеллектуального анализа Инструментальные средства анализа данных Применение Data Mining для решения бизнес-задач				
ИТОГО по 3-му семестру	18	24	26	72
ИТОГО по дисциплине	18	24	26	72

Тематика примерных практических занятий

№ п.п.	Наименование темы практического (семинарского) занятия
1	Задачи классификации, кластеризации и регрессии для прогнозирования
2	Метод иерархии целей
3	Алгоритм k-средних (k-means)

Тематика примерных лабораторных работ

№ п.п.	Наименование темы лабораторной работы
1	Очистка баз данных. Сжатие баз данных
2	Визуализация многомерных данных
3	Задача кластеризации

5. Организационно-педагогические условия

5.1. Образовательные технологии, используемые для формирования компетенций

Проведение лекционных занятий по дисциплине основывается на активном методе обучения, при котором учащиеся не пассивные слушатели, а активные участники занятия, отвечающие на вопросы преподавателя. Вопросы преподавателя нацелены на активизацию процессов усвоения материала, а также на развитие логического мышления. Преподаватель заранее намечает список вопросов, стимулирующих ассоциативное мышление и установление связей с ранее освоенным материалом.

Практические занятия проводятся на основе реализации метода обучения действием: определяются проблемные области, формируются группы. При проведении практических занятий преследуются следующие цели: применение знаний отдельных дисциплин и креативных методов для решения проблем и принятия решений; отработка у обучающихся навыков командной работы, межличностных коммуникаций и развитие лидерских качеств; закрепление основ теоретических знаний.

Проведение лабораторных занятий основывается на интерактивном методе обучения, при котором обучающиеся взаимодействуют не только с преподавателем, но и друг с другом. При этом доминирует активность учащихся в процессе обучения. Место преподавателя в интерактивных занятиях сводится к направлению деятельности обучающихся на достижение целей занятия.

При проведении учебных занятий используются интерактивные лекции, групповые дискуссии, ролевые игры, тренинги и анализ ситуаций и имитационных моделей.

5.2. Методические указания для обучающихся по изучению дисциплины

При изучении дисциплины обучающимся целесообразно выполнять следующие рекомендации:

1. Изучение учебной дисциплины должно вестись систематически.
2. После изучения какого-либо раздела по учебнику или конспектным материалам рекомендуется по памяти воспроизвести основные термины, определения, понятия раздела.
3. Особое внимание следует уделить выполнению отчетов по практическим занятиям, лабораторным работам и индивидуальным комплексным заданиям на самостоятельную работу.
4. Вся тематика вопросов, изучаемых самостоятельно, задается на лекциях преподавателем. Им же даются источники (в первую очередь вновь изданные в периодической научной литературе) для более детального понимания вопросов, озвученных на лекции.

6. Перечень учебно-методического и информационного обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Печатная учебно-методическая литература

№ п/п	Библиографическое описание (автор, заглавие, вид издания, место, издательство, год издания, количество страниц)	Количество экземпляров в библиотеке
1. Основная литература		
1	Люгер Д. Ф. Искусственный интеллект: Стратегии и методы решения сложных проблем : пер. с англ. 4-е изд. Москва [и др.] : Вильямс, 2003. 863 с.	33

2	Соловьев И. А., Червяков А. В., Репин А. Ю. Вычислительная математика на смартфонах, коммуникаторах и ноутбуках с использованием программных сред Python : учебное пособие. Санкт-Петербург [и др.] : Лань, 2011. 265 с. 14,28 усл. печ. л.	5
3	Чубукова И. А. Data Mining : учебное пособие. 2-е изд., испр. Москва : ИНТУИТ : БИНОМ. Лаб. знаний, 2008. 382 с.	4
2. Дополнительная литература		
2.1. Учебные и научные издания		
1	Методы и модели анализа данных: OLAP и Data Mining : учебное пособие / Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Санкт-Петербург : БХВ-Петербург, 2004. 336 с.	10
2.2. Периодические издания		
	Не используется	
2.3. Нормативно-технические издания		
	Не используется	
3. Методические указания для студентов по освоению дисциплины		
	Не используется	
4. Учебно-методическое обеспечение самостоятельной работы студента		
	Не используется	

6.2. Электронная учебно-методическая литература

Вид литературы	Наименование разработки	Ссылка на информационный ресурс	Доступность (сеть Интернет / локальная сеть; авторизованный / свободный доступ)
Дополнительная литература	Методы интеллектуального анализа данных	https://www.ibm.com/developerworks/ru/library/ba-data-mining-techniques/	сеть Интернет; свободный доступ

6.3. Лицензионное и свободно распространяемое программное обеспечение, используемое при осуществлении образовательного процесса по дисциплине

Вид ПО	Наименование ПО
Операционные системы	Debian (GNU GPL)
Офисные приложения.	LibreOffice 6.2.4. OpenSource, бесплатен.
Среды разработки, тестирования и отладки	NetBeans (SUN PUBLIC LICENSE)
Среды разработки, тестирования и отладки	PIP (The Python Package Installer) Free
Среды разработки, тестирования и отладки	PostgreSQL (PostgreSQL License)

6.4. Современные профессиональные базы данных и информационные справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Наименование	Ссылка на информационный ресурс
База данных Scopus	https://www.scopus.com/
База данных научной электронной библиотеки (eLIBRARY.RU)	https://elibrary.ru/
Научная библиотека Пермского национального исследовательского политехнического университета	http://lib.pstu.ru/
Электронно-библиотечная система Лань	https://e.lanbook.com/
Электронно-библиотечная система IPRbooks	http://www.iprbookshop.ru/
Информационные ресурсы Сети КонсультантПлюс	http://www.consultant.ru/

7. Материально-техническое обеспечение образовательного процесса по дисциплине

Вид занятий	Наименование необходимого основного оборудования и технических средств обучения	Количество единиц
Лабораторная работа	Компьютер	15
Лабораторная работа	Ноутбук, проектор	1
Лекция	Ноутбук, проектор	1
Практическое занятие	Компьютер	15
Практическое занятие	Ноутбук, проектор	1

8. Фонд оценочных средств дисциплины

Описан в отдельном документе

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

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
для проведения промежуточной аттестации обучающихся по дисциплине
«Интеллектуальный анализ данных»
Приложение к рабочей программе дисциплины

Направление подготовки:	09.04.04 Программная инженерия
Направленность (профиль) образовательной программы:	Разработка программно-информационных систем
Квалификация выпускника:	«Магистр»
Выпускающая кафедра:	Информационных технологий и автоматизированных систем
Форма обучения:	Очная

Курс: 2

Семестр: 3

Трудоёмкость:

Кредитов по рабочему учебному плану:	5	ЗЕ
Часов по рабочему учебному плану:	180	ч.

Форма промежуточной аттестации:

Экзамен: 3 семестр

Пермь 2023 г.

Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине является частью (приложением) к рабочей программе дисциплины. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине разработан в соответствии с общей частью фонда оценочных средств для проведения промежуточной аттестации основной образовательной программы, которая устанавливает систему оценивания результатов промежуточной аттестации и критерии выставления оценок. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине устанавливает формы и процедуры текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине.

1. Перечень контролируемых результатов обучения по дисциплине, объекты оценивания и виды контроля

Согласно РПД освоение учебного материала дисциплины запланировано в течение одного семестра (3-го семестра учебного плана). Предусмотрены аудиторские лекционные и лабораторные занятия, а также самостоятельная работа студентов. В рамках освоения учебного материала дисциплины формируется компоненты компетенций *знать, уметь, владеть*, указанные в РПД, которые выступают в качестве контролируемых результатов обучения по дисциплине (таблица 1.1).

Контроль уровня усвоенных знаний, освоенных умений и приобретенных владений осуществляется в рамках текущего, рубежного и промежуточного контроля при изучении теоретического материала, сдаче отчетов по лабораторным работам и экзамена. Виды контроля сведены в таблицу 1.1.

Таблица 1.1. Перечень контролируемых результатов обучения по дисциплине

Контролируемые результаты обучения по дисциплине (ЗУВы)	Вид контроля					
	Текущий		Промежуточный /рубежный		Итоговый	
	С	ТО	ОЛР	Т/КР		Экзамен
Усвоенные знания						
З.1 Знать методы получения новых знаний и умений, в том числе в новых областях знаний, связанных с профессиональной деятельностью; порядок поиска, систематизации и оценки достоверности научно-технической информации из различных источников, в т. ч. с использованием информационных технологий		ТО1		КР1		ТВ
Освоенные умения						
У.1 уметь, в том числе в с помощью информационных технологий, приобретать новые знания, расширять свое мировоззрение			ОЛР1,2,3,4	КР2		ПЗ
Приобретенные владения						
В.1 владеть информационно-коммуникационными технологиями в сфере профессиональной деятельности			ОЛР5,6,7			КЗ

С – собеседование по теме; *ТО* – коллоквиум (теоретический опрос); *КЗ* – кейс-задача (индивидуальное задание); *ОЛР* – отчет по лабораторной работе; *Т/КР* – рубежное тестирование (контрольная работа); *ТВ* – теоретический вопрос; *ПЗ* – практическое задание; *КЗ* – комплексное задание экзамена.

Итоговой оценкой достижения результатов обучения по дисциплине является промежуточная аттестация в виде экзамена, проводимая с учетом результатов текущего и рубежного контроля.

2. Виды контроля, типовые контрольные задания и шкалы оценивания результатов обучения

Текущий контроль успеваемости имеет целью обеспечение максимальной эффективности учебного процесса, управление процессом формирования заданных компетенций обучаемых, повышение мотивации к учебе и предусматривает оценивание хода освоения дисциплины. В соответствии с Положением о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся по образовательным программам высшего образования – программам бакалавриата, специалитета и магистратуры в ПНИПУ предусмотрены следующие виды и периодичность текущего контроля успеваемости обучающихся:

- входной контроль, проверка исходного уровня подготовленности обучаемого и его соответствия предъявляемым требованиям для изучения данной дисциплины;

- текущий контроль усвоения материала (уровня освоения компонента «знать» заданных компетенций) на каждом групповом занятии и контроль посещаемости лекционных занятий;

- промежуточный и рубежный контроль освоения обучаемыми отдельных компонентов «знать», «уметь» заданных компетенций путем компьютерного или бланочного тестирования, контрольных опросов, контрольных работ (индивидуальных домашних заданий), защиты отчетов по лабораторным работам, рефератов, эссе и т.д.

Рубежный контроль по дисциплине проводится на следующей неделе после прохождения темы дисциплины, а промежуточный – во время каждого контрольного мероприятия дисциплины;

- межсессионная аттестация, единовременное подведение итогов текущей успеваемости не менее одного раза в семестр по всем дисциплинам для каждого направления подготовки (специальности), курса, группы;

- контроль остаточных знаний.

2.1. Текущий контроль усвоения материала

Текущий контроль усвоения материала в форме собеседования или выборочного теоретического опроса студентов проводится по каждой теме. Результаты по 4-балльной шкале оценивания заносятся в книжку преподавателя и учитываются в виде интегральной оценки при проведении промежуточной аттестации.

2.2. Рубежный (промежуточный) контроль

Рубежный (промежуточный) контроль для комплексного оценивания усвоенных знаний, усвоенных умений и приобретенных владений (таблица 1.1) проводится в форме защиты лабораторных работ и рубежных контрольных работ (после изучения каждого модуля учебной дисциплины).

2.2.1. Защита лабораторных работ

Всего запланировано 7 лабораторных работ. Типовые темы лабораторных работ приведены в РПД.

Защита лабораторной работы проводится индивидуально каждым студентом. Типовые шкала и критерии оценки приведены в общей части ФОС образовательной программы.

2.2.2. Рубежная контрольная работа

Согласно РПД запланировано 2 рубежные контрольные работы (КР) после освоения студентами учебных тем дисциплины. Первая КР по теме «Выявление и упорядочение целей проекта», вторая КР – по теме «Выявление закономерностей в наборах данных и сжатие баз данных».

Типовые задания первой КР:

1. Расчет, классификация целей по методу иерархий.
2. Составление регрессионной модели.

Типовые задания второй КР:

1. Рациональное кодирование данных. Группировка данных методами кластерного анализа.
- 2.. Построение дерева решений, вывод модели и алгоритма принятия решения, прогноз.

Типовые шкала и критерии оценки результатов рубежной контрольной работы приведены в общей части ФОС образовательной программы.

2.3. Промежуточная аттестация (итоговый контроль)

Допуск к промежуточной аттестации осуществляется по результатам текущего и рубежного контроля. Условиями допуска являются успешная сдача всех лабораторных работ и положительная интегральная оценка по результатам текущего и рубежного контроля.

Промежуточная аттестация, согласно РПД, проводится в виде экзамена по дисциплине устно по билетам. Билет содержит теоретические вопросы (ТВ) для проверки усвоенных знаний, практические задания (ПЗ) для проверки освоенных умений и комплексные задания (КЗ) для контроля уровня приобретенных владений всех заявленных компетенций.

Билет формируется таким образом, чтобы в него попали вопросы и практические задания, контролирующие уровень сформированности всех заявленных компетенций. Форма билета представлена в общей части ФОС образовательной программы.

2.3.1. Типовые вопросы и задания для экзамена по дисциплине

Типовые вопросы для контроля усвоенных знаний:

1. Технологическая схема анализа данных
2. Методы построения моделей знаний, постановки и решения задач интеллектуального анализа данных
3. Гносеологическая связь методов кластеризации, регрессии, классификации и ассоциации

4. Оптимизация базиса предметной области

Типовые вопросы и практические задания для контроля освоенных умений:

1. Представление базы данных в виде базы знаний
2. Построение индуктивного дерева решений по базе данных и представление ее в виде модели
3. Очистка базы данных и оптимизация множества концептов
4. Прогноз данных по построенной модели

Типовые комплексные задания для контроля приобретенных владений:

1. Составить алгоритм построения регрессионной модели N-го порядка
2. Составить алгоритм построения дерева классов
3. Составить алгоритм поиска решения на модели предметной области в виде графа

2.3.2. Шкалы оценивания результатов обучения на экзамене

Оценка результатов обучения по дисциплине в форме уровня сформированности компонентов *знать, уметь, владеть* заявленных компетенций проводится по 4-х балльной шкале оценивания путем выборочного контроля во время экзамена.

Типовые шкала и критерии оценки результатов обучения при сдаче экзамена для компонентов *знать, уметь и владеть* приведены в общей части ФОС образовательной программы.

3. Критерии оценивания уровня сформированности компонентов и компетенций

3.1. Оценка уровня сформированности компонентов компетенций

При оценке уровня сформированности компетенций в рамках выборочного контроля при экзамене считается, что *полученная оценка за компонент проверяемой в билете компетенции обобщается на соответствующий компонент всех компетенций, формируемых в рамках данной учебной дисциплины.*

Типовые критерии и шкалы оценивания уровня сформированности компонентов компетенций приведены в общей части ФОС образовательной программы.

3.2. Оценка уровня сформированности компетенций

Общая оценка уровня сформированности всех компетенций проводится путем агрегирования оценок, полученных студентом за каждый компонент формируемых компетенций, с учетом результатов текущего и рубежного контроля в виде интегральной оценки по 4-х балльной шкале. Все результаты контроля заносятся в оценочный лист и заполняются преподавателем по итогам промежуточной аттестации.

Форма оценочного листа и требования к его заполнению приведены в общей части ФОС образовательной программы.

При формировании итоговой оценки промежуточной аттестации в виде

экзамена используются типовые критерии, приведенные в общей части ФОС образовательной программы.

Перечень типовых ситуационных заданий и кейсов для проверки умений и владений представлен в приложении 1.

Полный перечень теоретических вопросов и практических заданий в форме утвержденного комплекта экзаменационных билетов хранится на выпускающей кафедре.

Приложение 1.

Типовые ситуационные задания и кейсы для проверки умений и владений

Задание № __. (анализ кейс-стадии)

Проверяемые результаты обучения: y1; v1

Задание. Внимательно прочитайте текст предложенного кейса, найдите ответ и ответьте на вопросы задания.

Критерии оценки ситуационных заданий

Оценка «пять» ставится, если обучающийся осознанно излагает и оценивает суть данной ситуации, с аргументацией своей точки зрения, умеет анализировать, обобщать и предлагает верные пути решения складывающейся ситуации.

Оценка «четыре» ставится, если обучающийся понимает суть ситуации, логично строит свой ответ, но допускает незначительные неточности при определении путей решения.

Оценка «три» ставится, если обучающийся ориентируется в сущности складывающейся ситуации, но нуждается в наводящих вопросах, не умеет анализировать и не совсем верно намечает пути решения ситуации.

Оценка «два» ставится, если обучающийся не ориентируется и не понимает суть данной ситуации, не может предложить путей ее решения, либо допускает грубые ошибки.

Ситуация 1. При наблюдении деятельности кредитного отдела банка были описаны 14 случаев получения (отказа) заемщику в кредите.

Во время собеседования в отделе с клиентом анализировались кредитная история, долги, поручители и доходы потенциального клиента.

Определить важность каждого из документов при решении вопроса о выдаче кредита.

Определить правило выдачи кредита в условиях системного кризиса.

Осуществить для клиента 15 с заданными параметрами кредитной истории, текущего долга, поручителей, дохода расчетный риск возврата клиентом кредита банка.

Оценить сложность полученного правила и размерность несжатого дерева.

Ситуация 2. В магазине имеется набор из 20 товаров различного веса, размера, стоимости, уровня дневного спроса с указанием их изготовителей. Описание товаров находится в базе данных.

Классифицируйте товары для оптимального их расположения на полках супермаркета.

Приведите дерево классов.

Укажите в дереве классов уровень доверия при принятии каждого из решений.

Ситуация 3. Для министерства транспорта составлено двухъярусное дерево целей и подцелей, полученное экспертным путем с численными оценками экспертов.

Развитие городского транспорта электрическими машинами, дизельными машинами, подземным транспортом предполагается по критериям: удобство проезда, окупаемость транспорта, доступность районов города для населения.

Оцените численно важность каждой из целей для системы и очередность реализации целей.

Составьте уравнение принятия решений методом иерархий.

Ситуация 4. Имеется группа людей с экспертной численной оценкой их некоторой способности. Разбейте множество людей на группы наилучшим способом для дальнейшего их обучения.

Определите расчетом оптимальное количество групп.

Определите расчетом состав каждой группы.

Приложение 2. Форма билета для экзамена

09.04.01 «Информатика и вычислительная техника»

09.04.01 Разработка программно-информационных систем

Кафедра «Информационные технологии и автоматизированные системы»

МИНОБРНАУКИ РОССИИ
ФГБОУ ВПО «Пермский национальный исследовательский
политехнический университет» (ПНИПУ)

Дисциплина «Интеллектуальный анализ данных»

БИЛЕТ № 1

1. Представление обобщенных знаний по заданным данным, метод, инструментарий (*контроль знаний*)
2. Построение дерева решений по базе данных (*контроль умений*)
3. Составить алгоритм построения регрессионной модели N-порядка (*контроль умений и владений*)