

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

« 20 » марта 20 23 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Дисциплина: Статистические методы анализа данных и технологии DataMining

(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

Цель дисциплины в формировании знаний, умений и навыков проведения самостоятельных исследований методами Data Mining и эффективного использования результатов уже готовых статистических исследований
Задачи дисциплины: - знать методы и средства интеллектуального анализа данных; - уметь самостоятельно использовать в практической деятельности интеллектуальный анализ данных с помощью информационных технологий - уметь решать задачи статистического анализа с применением методов моделирования - владеть навыками сбора и анализа информации в соответствующей профессиональной сфере, а также экспериментального исследования объектов профессиональной деятельности.

1.2. Изучаемые объекты дисциплины

Методы и инструменты анализа данных Data Mining

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-7	ИД-1ОПК-7	Знает основные методы и средства получения, хранения и использования информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях	Знает основные методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях	Индивидуальное задание
ОПК-7	ИД-2ОПК-7	Умеет проводить поиск информации по различным показателям с использованием различных поисковых технологий и извлекать необходимую информацию из информационных систем	Умеет организовывать поиск информации по различным критериям с использованием различных поисковых технологий и извлекать необходимую информацию из информационных систем	Индивидуальное задание

Компетенция	Индекс индикатора	Планируемые результаты обучения по дисциплине (знать, уметь, владеть)	Индикатор достижения компетенции, с которым соотнесены планируемые результаты обучения	Средства оценки
ОПК-7	ИД-3ОПК-7	Владеет методами и средствами работы с современными базами данных и методами получения и обработки информации в глобальных компьютерных сетях	Владеет методами и средствами работы с современными хранилищами данных и методами получения и обработки информации в глобальных компьютерных сетях	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:	62	62	
1.1. Контактная аудиторная работа, из них:			
- лекции (Л)	16	16	
- лабораторные работы (ЛР)	16	16	
- 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)	82	82	
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				

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
Типовые задачи Data Mining и методы их решения	4	4	8	25
<p>Задачи классификации и методы решения. Постановка задач, ключевые понятия и определения. Методы, применяемые для решения задач классификации: индукция деревьев решений; байесовские сети (Bayesian Networks); нейронные сети (neural networks); статистические методы, в частности, линейная регрессия. Преимущества деревьев решений.</p> <p>Интуитивность деревьев решений. Точность. Процесс конструирования дерева решений. Метод "ближайшего соседа".</p> <p>Регрессионный анализ. Последовательность этапов регрессионного анализа. Классические методы регрессионного анализа: множественная и логистическая регрессии, выбор переменных для анализа. Установление формы зависимости. Определение функции регрессии. Оценка неизвестных значений зависимой переменной. Уравнение регрессии. Случайное отклонение. Коэффициент множественной корреляции R Сравнение качества построенных моделей.</p> <p>Задачи кластеризации: постановка задачи, ключевые понятия и определения, метод k-средних и EM-алгоритм. Задача понижения размерности, метод независимых компонент (independent component analysis). Меры сходства. Иерархический кластерный анализ в SPSS. Алгоритм k-средних (k-means). Классические методы кластеризации – Метод иерархической кластеризации (tree clustering). Классические методы понижения размерности: метод главных компонент (principal component analysis), факторный анализ (Factor Analysis).</p> <p>Прогнозирование временных рядов – основные понятия (тренд, сезонность, календарные эффекты, разложение ряда), классическая модель АРПСС (ARIMA), экспоненциальное сглаживание, анализ лагов.</p> <p>Нейронные сети (Neural networks): элементы нейронных сетей, обучение нейронных сетей, модели нейронных сетей, программное обеспечение для работы с нейронными сетями.</p> <p>Пакет Matlab</p>				
Подготовка и предварительный анализ данных, введение в Data Mining,	6	6	8	25
<p>Общая концепция методологии Data Mining и технологии реализации. Обзор задач, решаемых методами Data Mining.</p> <p>Классификация методов Data Mining по различным</p>				

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
<p>признакам.</p> <p>Этапы интеллектуального анализа данных: анализ предметной области, постановка задачи, подготовка данных. Процесс подготовки данных, понятия качества данных, грязных данных, этапы очистки данных. Этапы процесса Data Mining, связанные с построением, проверкой, оценкой, выбором и коррекцией моделей.</p> <p>Процесс Data Mining как последовательность этапов и как последовательность работ, выполняемых исполнителями ролей Data Mining. Подготовка и предварительный анализ данных</p> <p>Анализ взаимосвязей (корреляций) между переменными – коэффициенты корреляции Пирсона, Спирмена и Кендалла. Сравнение групп – критерии Стьюдента, Манна – Уитни и дисперсионный анализ (ANOVA). Инструменты очистки и редактирования данных, основные функции инструментов очистки данных, классификация ошибок в данных, которые возникают в результате использования средств очистки данных.</p> <p>Инструменты очистки данных.</p>				
Методы анализа данных и используемые приложения	6	6	10	32
<p>Задачи поиска ассоциативных правил. Методы поиска ассоциативных правил. Алгоритм Apriori. Приложения с применением ассоциативных правил.</p> <p>Задачи и методы визуализации. Способы представления информации в одно-, двух-, трехмерном измерениях и более. Принципы качественной визуализации. Основные тенденции в области визуализации.</p> <p>Визуализация инструментов Data Mining.</p> <p>Визуализация Data Mining моделей. Приложения Data Mining и используемые программные продукты СППР, их типы и компоненты. OLAP-технологии, архитектуры OLAP-серверов, интеграции Data Mining и OLAP. Технология хранилищ данных и преимущества их использования для процесса Data Mining. Процесс Data Mining и методологии CRISP и SEMMA.</p> <p>Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. Система STATISTICA Data Miner, средства анализа и схема работы интеграции Data Mining и OLAP. Технология хранилищ данных и преимущества их использования для процесса Data Mining. Процесс</p>				

Наименование разделов дисциплины с кратким содержанием	Объем аудиторных занятий по видам в часах			Объем внеаудиторных занятий по видам в часах
	Л	ЛР	ПЗ	СРС
Data Mining и методологии CRISP и SEMMA. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования. Система STATISTICA Data Miner, средства анализа и схема работы.				
ИТОГО по 3-му семестру	16	16	26	82
ИТОГО по дисциплине	16	16	26	82

Тематика примерных практических занятий

№ п.п.	Наименование темы практического (семинарского) занятия
1	Этапы интеллектуального анализа данных: анализ предметной области, постановка задачи, подготовка данных, понятия качества данных, грязных данных, этапы очистки данных .
2	Процесс Data Mining как последовательность этапов и как последовательность работ, выполняемых исполнителями ролей Data Mining.
3	Анализ взаимосвязей (корреляций) между переменными – коэффициенты корреляции Пирсона, Спирмена и Кендалла.
4	Модели дисперсионного анализа (ANOVA).
5	Выбор вида зависимости, набора значимых факторов, проверка на значимость.
6	Методы, применяемые для решения задач классификации: индукция деревьев решений; байесовские сети (Bayesian Networks); нейронные сети (neural networks) Методы, применяемые для решения задач кластеризации: метод k-средних и EM-алгоритм.
7	Задача понижения размерности. Метод независимых компонент (independent component analysis).
8	Система STATISTICA Data Miner, средства анализа и схема работы.

Тематика примерных лабораторных работ

№ п.п.	Наименование темы лабораторной работы
1	Анализ данных, представленных в виде временных рядов.
2	Анализ взаимосвязей между переменными (корреляционный анализ).
3	Регрессионный анализ: установление вида связи и проверка на качество.
4	Иерархический кластерный анализ в SPSS. Итеративная кластеризация в SPSS.

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	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. Дополнительная литература		
2.1. Учебные и научные издания		

1	Берикашвили В. Ш., Оськин С. П. Статистическая обработка данных, планирование эксперимента и случайные процессы : учебное пособие для бакалавриата и магистратуры. 2-е изд., испр. и доп. Москва : Юрайт, 2019. 164 с. 12,72 усл. печ. л.	6
2.2. Периодические издания		
	Не используется	
2.3. Нормативно-технические издания		
	Не используется	
3. Методические указания для студентов по освоению дисциплины		
1	Дубнов П. Ю. Обработка статистической информации с помощью SPSS. Москва : АСТ : NT Press, 2004. 221 с.	4
4. Учебно-методическое обеспечение самостоятельной работы студента		
1	Бродягин В. В. Основы компьютерных технологий решения геологических задач : учебное пособие. Пермь : Изд-во ПГТУ, 2008. 307 с.	28

6.2. Электронная учебно-методическая литература

Вид литературы	Наименование разработки	Ссылка на информационный ресурс	Доступность (сеть Интернет / локальная сеть; авторизованный / свободный доступ)
Методические указания для студентов по освоению дисциплины	Бродягин В. В. Основы компьютерных технологий решения геологических задач : учебное пособие / В. В. Бродягин. - Пермь: Изд-во ПГТУ, 2008.	http://elib.pstu.ru/Record/RUPSTUbooks130806	локальная сеть; свободный доступ

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

Вид ПО	Наименование ПО
Операционные системы	Windows 10 (подп. Azure Dev Tools for Teaching)
Офисные приложения.	Adobe Acrobat Reader DC. бесплатное ПО просмотра PDF
Офисные приложения.	LibreOffice 6.2.4. OpenSource, бесплатен.
Офисные приложения.	Microsoft Office Professional 2007. лиц. 42661567
Офисные приложения.	МойОфис Стандартный. , реестр отечественного ПО, необходима покупка лицензий.
Прикладное программное обеспечение общего назначения	Dr.Web Enterprise Security Suite, 3000 лиц, ПНИПУ ОЦНИТ 2017
Прикладное программное обеспечение общего назначения	MATLAB 7.9 + Simulink 7.4 Academic, ПНИПУ 2009 г.

Вид ПО	Наименование ПО
Прикладное программное обеспечение общего назначения	Oracle VM VirtualBox (GNU GPL 2)
Прикладное программное обеспечение общего назначения	WinRAR (лиц.№ 879261.1493674)
Системы управления проектами, исследованиями, разработкой, проектированием, моделированием и внедрением	Deductor Academic Deductor Academic (Free)
Системы управления проектами, исследованиями, разработкой, проектированием, моделированием и внедрением	IBM SPSS Statistic Base
Среды разработки, тестирования и отладки	Embarcadero Delphi 2007, лиц.№ 33948, 137 лиц. ПНИПУ 2008 г.
Среды разработки, тестирования и отладки	Java (JDK + JRE) Sun License (GPL) свободное ПО
Среды разработки, тестирования и отладки	Язык R

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

Наименование	Ссылка на информационный ресурс
База данных Elsevier "Freedom Collection"	https://www.elsevier.com/
База данных Scopus	https://www.scopus.com/
База данных Springer Nature e-books	http://link.springer.com/ http://jwww.springerprotocols.com/ http://materials.springer.com/ http://zbmath.org/ http://npg.com/
База данных Web of Science	http://www.webofscience.com/
База данных Wiley Journals	http://onlinelibrary.wiley.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/
Виртуальный читальный зал Российской государственной библиотеки	https://dvs.rsl.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	Компьютерный класс	10
Лекция	Лекционная аудитория: проектор и компьютер	1
Практическое занятие	Компьютерный класс	10

8. Фонд оценочных средств дисциплины

Описан в отдельном документе

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
**«Пермский национальный исследовательский политехнический
университет»**
Электротехнический факультет
Кафедра «Информационных технологий и автоматизированных систем»

УТВЕРЖДЕНО
на заседании кафедры ИТАС
протокол № ___ от __.__.2022
Заведующий кафедрой
_____ Р.А.Файзрахманов

УЧЕБНО-МЕТОДИЧЕСКИЙ КОМПЛЕКС ДИСЦИПЛИНЫ

«Статистические методы анализа данных и технологии DataMining»

Программа академической магистратуры

Направление 09.04.04 Программная инженерия

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

для проведения промежуточной аттестации обучающихся по дисциплине

Профиль программы магистратуры: 09.04.04.51 Разработка программно-информационных систем

Квалификация выпускника: магистр

Выпускающая кафедра: Информационные технологии и автоматизированные системы

Форма обучения: очная

Курс: 2Семестр(-ы): 3

Трудоёмкость:

Кредитов по рабочему учебному плану: 5 ЗЕ

Часов по рабочему учебному плану: 180 ч

Виды итогового контроля:

Зачет 3 семестр

Пермь 2022

Фонд оценочных средств дисциплины «Статистические методы анализа данных и технологии DataMining» разработан на основании:

- положения о проведении текущего контроля успеваемости и промежуточной аттестации обучающихся по образовательным программам высшего образования – программам бакалавриата, специалитета и магистратуры в ПНИПУ;
- компетентностной модели выпускника ОПОП по направлению подготовки 09.04.04 Программная инженерия (уровень магистратуры)», программе магистратуры 09.04.04.51 Разработка программно-информационных систем;
- базового учебного плана очной формы обучения по направлению подготовки 09.04.04 Программная инженерия (уровень магистратуры)», программе магистратуры 09.04.04.51 Разработка программно-информационных систем рабочей программы дисциплины «Статистические методы анализа данных и технологии DataMining».

1. Перечень формируемых частей компетенций, этапы их формирования и контролируемые результаты обучения

1.1. Формируемые части компетенций

Согласно КМВ ОПОП учебная дисциплина «Статистические методы анализа данных и технологии DataMining» участвует в формировании 1-й компетенций: **ОПК-7**. В рамках учебного плана образовательной программы в 3-м семестре на этапе освоения данной учебной дисциплины формируются следующие дисциплинарные части компетенций:

ИД-1ОПК-7 Знает основные методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях

ИД-2ОПК-7 Умеет организовывать поиск информации по различным критериям с использованием различных поисковых технологий и извлекать необходимую информацию из информационных систем

ИД-3ОПК-7 Владеет навыками выбора методов решения, установления ограничений к решениям научно-технической задачи в сфере профессиональной деятельности на основе нормативно-технической документации и знания проблем отрасли и опыта их решения

1.2. Этапы формирования дисциплинарных частей компетенций, объекты оценивания и виды контроля

Согласно РПД освоение учебного материала дисциплины запланировано в течение одного семестра (3-го семестра базового учебного плана) и разбито на 3 учебных модуля. В каждом модуле предусмотрены аудиторские лекционные и лабораторные занятия, а также самостоятельная работа студентов. В рамках освоения учебного материала дисциплины формируются компоненты дисциплинарных компетенций *знать, уметь, владеть*, указанные в РПД, и которые выступают в качестве контролируемых результатов обучения (табл. 1.1).

Контроль уровня усвоенных знаний, усвоенных умений и приобретенных владений осуществляется в рамках текущего, рубежного и промежуточного контроля при изучении теоретического материала, сдаче отчетов по лаборатор-

ным работам и дифференцированного зачета. Виды контроля сведены в таблицу 1.1.

Таблица 1.1. Контролируемые результаты обучения по дисциплине

Контролируемые результаты обучения по дисциплине (ЗУВы)	Вид контроля		
	Текущий	Промежуточный	Итоговый
	ТО	ОЛР	Зачёт
Усвоенные знания			
ИД-1 _{ОПК-7} . Знает основные методы и средства получения, хранения, переработки и трансляции информации посредством современных компьютерных технологий, в том числе в глобальных компьютерных сетях	ТО2		ТВ
Освоенные умения			
ИД-2 _{ОПК-72} . Умеет проводить поиск информации по различным показателям с использованием различных поисковых технологий и извлекать необходимую информацию из информационных систем		ОЛР2	
Приобретенные владения			
ИД-3 _{ОПК-7} . Владеет методами и средствами работы с современными базами данных и методами получения и обработки информации в глобальных компьютерных сетях		ОЛР3	

ТО – теоретический опрос;

ОЛР – отчет по лабораторной работе;

ТВ – теоретический вопрос;

Итоговой оценкой освоения дисциплинарных компетенций (результатов обучения по дисциплине) является промежуточная аттестация в виде зачета, проводимая с учетом результатов текущего и рубежного контроля.

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

2.1. Текущий контроль

Текущий контроль для оценивания знаниевого компонента дисциплинарных частей компетенций (табл. 1.1) в форме 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. Промежуточный контроль

Промежуточный контроль для комплексного оценивания усвоенных знаний, освоенных умений и приобретенных владений дисциплинарных частей компетенций (табл. 1.1) проводится согласно графику учебного процесса, приведенного в РПД, в форме защиты лабораторных работ.

2.2.1. Защита лабораторных работ

Всего запланировано 4 лабораторных работы. Типовые темы лабораторных работ приведены в РПД.

Защита лабораторной работы проводится индивидуально каждым студентом. Шкала и критерии оценки приведены в Общей части ФОС.

Результаты защиты лабораторных работ по 4-балльной шкале оценивания заносятся в книжку преподавателя и учитываются в виде интегральной оценки при проведении промежуточной аттестации.

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 промежуточного контроля.

Критерии выведения итоговой оценки промежуточной аттестации приведены в Общей части ФОС.

2.3.1. Типовые вопросы и задания для подготовки к экзамену по дисциплине

Типовые вопросы для контроля усвоенных знаний:

1. Что такое data mining, компоненты data mining.
2. Цель применения сантимерт анализа (три типа моделей, привести пример по каждой модели).
3. Что такое препроцессинг документов для целей тематического моделирования и сантимерт анализа.
4. Цели и задачи кластерного анализа.
5. Направления в кластерном анализе.
6. Меры близости в кластерном анализе.
7. Алгоритм K-means
8. Алгоритм иерархической кластеризации
9. Достоинство и недостатки кластерного подхода.
10. Особенности тематического моделирования. Цель применения тематического моделирования.
11. Достоинства и недостатки тематического моделирования.
12. Суть метода наименьших квадратов (МНК). Цель применения метода МНК.
13. Суть модели скользящего среднего (МСС). Цели применения СС.
14. Суть модели Брауна (МБ). Цели применения МБ.
15. Принцип выбора адекватной модели в прогнозировании.
16. Цель применения регрессионной модели.
17. Суть регрессионной модели (смысл коэффициентов регрессии).
18. Виды регрессионных моделей.
19. Достоинства и недостатки регрессионных моделей.
20. Особенности сетевого анализа.
21. Суть применения сетевого анализа.

22. Понятие о нейронной сети.
23. Особенности нейронного моделирования.
24. Достоинства и недостатки нейронного моделирования.

Типовые вопросы и практические задания для контроля освоенных умений:

1. Как преобразование Бокса-Кокса улучшает качество модели?
2. В чем отличие между искусственным интеллектом, машинным обучением и Data Science
3. В чем разница между линейной и логистической регрессией
4. Как выбрать число k для алгоритма кластеризации "метод k-средних"
5. Как бы вы справились с разными формами сезонности при моделировании временных рядов?
6. Кластерный анализ в сегментировании рынка, проблемы применения и их решение с помощью факторного анализа.
7. Сегментирование рынка: объекты сегментирования рынка, переменные сегментирования рынка, данные для сегментирования рынка.
8. Меры сходства и расстояние между потребителями при сегментировании рынка с применением кластерного анализа. Кластерный анализ.
9. Важность переменных сегментирования рынка .
10. Кластеры потребителей и кластерный анализ.
11. Кластерный анализ и математические характеристики кластера.
12. Методы кластерного анализа спорные объекты и перекрывающиеся кластеры .
13. Сегментирование рынка и факторный анализ.
14. Факторный анализ и кластерный анализ использование при сжатии информации.
15. Функции полезности и сегментирование рынка.
16. Сегментирование рынка и профайл потребителя из сегмента рынка.
17. Влияние переменных сегментирования рынка на результаты применения кластерного анализа.
18. Позиционирование товаров и репозиционирование брендов. Репозиционирование и сегментирование рынка.
19. Стратегический потенциал сегментов рынка.

Типовые комплексные задания для контроля приобретенных владений:

1. Классификация инструментов Data Mining
2. Программное обеспечение Data Mining для поиска ассоциативных правил
3. Программное обеспечение для решения задач кластеризации и сегментации
4. Программное обеспечение для решения задач классификации
5. Программное обеспечение Data Mining для решения задач оценивания и прогнозирования
6. Инструменты Data Mining. Система PolyAnalyst
7. Алгоритмы кластеризации
8. Алгоритмы классификации

9. Алгоритмы ассоциации
10. Модули текстового анализа
11. Визуализация.
12. Программные продукты и система STATISTICA
13. Средства анализа STATISTICA
14. Инструменты Oracle
15. Прогнозирующие модели
16. Аналитическая платформа Deductor

3. Оценка уровня сформированности компонентов дисциплинарных компетенций

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

Общая оценка уровня сформированности всех дисциплинарных 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 магистерской программы.

При формировании итоговой оценки промежуточной аттестации в виде экзамена используются следующие критерии:

$$OЗ = 0.5 * OЗЭ + 0.3 * OЗРТ + 0.2 * OЗТ,$$

где OЗ – общая оценка уровня сформированности знаний, OЗЭ – оценка знаний при ответе на билет экзамена, OЗРТ – средняя оценка знаний при рубежных тестированиях, OЗТ – оценка знаний при текущем контроле. (Все оценки по 4-х балльной шкале 2,3,4,5.)

$$OУ = 0.2 * OУЛ + 0.3 * OУИ + 0.5 * OУЭ,$$

где OУ – общая оценка уровня сформированности умений, OУЛ – оценка умений по итогам защиты лабораторных работ, OУЭ – оценка умений по итогам сдачи экзамена, OУИ – оценка умений по итогам индивидуального задания.

$$OВ = 0.2 * OВЛ + 0.3 * OВИ + 0.5 * OВЭ,$$

где OВ – общая оценка уровня сформированности владений, OВЛ – оценка владений по итогам защиты лабораторных работ, OВЭ – оценка владений по

итогах сдачи экзамена, ОВИ - оценка владений по итогам индивидуального задания.

Формула пересчета ОЗ, ОУ и ОВ в итоговую оценку:

$$\text{Итоговая оценка} = 0.4 \cdot \text{ОВ} + 0.3 \cdot \text{ОУ} + 0.3 \cdot \text{ОЗ}$$