

**Федеральное государственное образовательное бюджетное учреждение
высшего образования
«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ
ПРАВИТЕЛЬСТВЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

**Краснодарский филиал Финуниверситета
Кафедра «Математика и информатика»**

СОГЛАСОВАНО

ООО «Портал-Юг»
Генеральный директор



Е.В. Мостовой

«20» февраля 2025 г.

УТВЕРЖДАЮ

Краснодарский филиал
Финуниверситета
Директор



Э.В.Соболев

«20» февраля 2025 г.

Калайдин Е. Н.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
АНАЛИЗ БОЛЬШИХ ДАННЫХ**

для студентов, обучающихся по направлению подготовки
01.04.02 Прикладная математика и информатика,
ОП «Анализ больших данных и машинное обучение в экономике и финансах»
(программа магистратуры)

*Одобрено кафедрой «Математика и информатика»
(протокол № 12 от 11.02.2025)*

*Рекомендовано Ученым советом Краснодарского филиала Финуниверситета
(протокол № 24 от 18.02.2025)*

Краснодар 2025

УДК 004.4
ББК 32.973.26-018.2я73
К17

Рецензенты: Кирий В.А., доцент кафедры «Математика и информатика», канд. физ.-мат. наук,
Коренева О.В., доцент кафедры «Математика и информатика», канд. техн. наук.

Калайдин Е. Н. Рабочая программа дисциплины «Анализ больших данных» для обучающихся магистратуры по направлению 01.04.02 «Прикладная математика и информатика», профиль «Анализ больших данных и машинное обучение в экономике и финансах». – Краснодар: Краснодарский филиал Финуниверситета, кафедра «Математика и информатика» 2025.

Дисциплина Анализ больших данных относится к модулю направленности программы подготовки 01.04.02 «Прикладная математика и информатика», направленность программы магистратуры «Анализ больших данных и машинное обучение в экономике и финансах».

В рабочей программе дисциплины определены ее цель, требования к результатам освоения дисциплины, содержание программы, тематика аудиторных занятий, формы самостоятельной работы, оценочные средства для текущего контроля и промежуточной аттестации, учебно-методическое и информационное обеспечение.

Учебное издание

Калайдин Е. Н.

Анализ больших данных

Рабочая программа дисциплины

Компьютерный набор и верстка Е. Н. Калайдин
Формат 60х90/16. Гарнитура Times New Roman
Усл.п.л. 1,75. Изд. № - 2023. Тираж экз.

Заказ №

Отпечатано в Краснодарском филиале Финансового университета

© Калайдин Е. Н., 2025
© КФ Финансового университета, 2025

СОДЕРЖАНИЕ

1. Наименование дисциплины	4
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине	4
3. Место дисциплины в структуре образовательной программы	6
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	6
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	7
5.1. Содержание дисциплины	7
5.2. Учебно-тематический план	7
5.3. Содержание семинаров, практических занятий	9
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	10
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	10
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по данной дисциплине	11
Дискриминантный	22
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	26
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	26
10. Методические указания для обучающихся по освоению дисциплины	28
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	28
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	29

1.Наименование дисциплины

«Анализ больших данных».

2.Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

В совокупности с другими дисциплинами модуля учебная дисциплина «Анализ больших данных» обеспечивает формирование следующих компетенций: ПКН-2; ПКН-6; ПК-2

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
ПКН-2	Способность использовать современные информационные системы и математические методы для решения математических, естественнонаучных, социально-экономических и профессиональных задач предметной области	1. Демонстрирует знание современных информационных систем на основе и с применением математических моделей и методов.	Знать: современные информационные системы на основе и с применением математических моделей и методов Уметь: выявлять современные информационные системы на основе и с применением математических моделей и методов
		2. Применяет полученные знания при решении математических, естественнонаучных, социально-экономических и профессиональных задач в междисциплинарном контексте.	Знать: теоретические основы при решении математических, естественнонаучных, социально-экономических и профессиональных задач в междисциплинарном контексте Уметь: применять полученные знания при решении математических, естественнонаучных, социально-экономических и профессиональных задач в междисциплинарном контексте
		3. Владеет методами анализа Big Data, использует для решения профессиональных задач на макро-, мезо- и микроуровнях, в том числе на уровне финансового рынка.	Знать: методы анализа Big Data, теоретические основы для решения профессиональных задач на макро-, мезо- и микроуровнях, в том числе на уровне финансового рынка Уметь:

			владеет методами анализа Big Data, использует для решения профессиональных задач на макро-, мезо- и микроуровнях, в том числе на уровне финансового рынка
ПKN-6	Способность анализировать и оценивать эффективность применения методов прикладной математики и информатики	1. Демонстрирует знание основных методов прикладной математики и информатики, применяемых в различных предметных областях.	Знать: основные методы прикладной математики и информатики, применяемых в различных предметных областях Уметь: различать основные методы прикладной математики и информатики, применяемые в различных предметных областях
		2. Владеет методологией математического моделирования для решения профессиональных задач.	Знать: теоретические основы методологии математического моделирования для решения профессиональных задач Уметь: владеть методологией математического моделирования для решения профессиональных задач
ПК-2	Способность применять лингвистические и сетевые модели машинного обучения, предобученные на больших наборах данных.	1. Демонстрирует знание современных предобученных лингвистических моделей или инструментов создания предобучения сетевых моделей.	Знать: современные предобученные лингвистические модели или инструменты создания предобучения сетевых моделей Уметь: различать современные предобученные лингвистические модели или инструменты создания предобучения сетевых моделей
		2. Владеет методологией и инструментарием использования предобученных моделей для создания решения прикладных задач, разрабатывает на базе известных решений технологию решения прикладной задачи с помощью предобученных моделей.	Знать: методологию и инструментарий использования предобученных моделей для создания решения прикладных задач, теоретические основы для разработки на базе известных решений технологии решения прикладной задачи с помощью предобученных моделей Уметь: владеть методологией и

			инструментарием использования предобученных моделей для создания решения прикладных задач, разрабатывать на базе известных решений технологию решения прикладной задачи с помощью предобученных моделей
		3. Владеет практическим навыком применения технологии использования предобученных моделей для решения прикладных задач.	Знать: теоретические основы технологий использования предобученных моделей для решения прикладных задач Уметь: владеть практическим навыком применения технологии использования предобученных моделей для решения прикладных задач

3. Место дисциплины в структуре образовательной программы

Дисциплина «Анализ больших данных» относится к модулю направленности программы магистратуры по направлению подготовки 01.04.02 - Прикладная математика и информатика.

4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Вид учебной работы по дисциплине	Всего (в часах)	Модуль 3 (в часах)
Общая трудоемкость дисциплины	180	180
Контактная работа - Аудиторные занятия	40	40
<i>Лекции</i>	<i>10</i>	<i>10</i>
<i>Семинары, практические занятия</i>	<i>30</i>	<i>30</i>
Самостоятельная работа	140	140
В семестре	104	104
Контроль (подготовка к зачету/экзамену)	36	36
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	экзамен	экзамен

5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

5.1. Содержание дисциплины

Предмет, методы и инструменты анализа.

Данные в экономике: объекты и признаки. Разведочный анализ и подтверждающий анализ. Программные средства анализа больших данных.

Разведочный анализ.

Подготовка данных. Выборочные оценки, визуализация, описательная статистика. Формулировка гипотез.

Проверка статистических гипотез.

Критерии согласия. Параметрические и непараметрические критерии. Критерии Стьюдента и Манна-Уитни. Дисперсионный анализ и критерий Краскела- Уоллиса. Post hoc тесты.

Множественная регрессия.

Линеаризация модели. Проблема мультиколлинеарности. Отбор факторов. Оценка параметров. Оценка качества модели.

Многомерные статистические методы.

Факторный анализ и метод главных компонент. Методы многомерной классификации: классификация с обучением и без обучения. Кластерный анализ. Дискриминантный анализ.

5.2. Учебно-тематический план

п/ п	Наименование тем (разделов) дисциплины	Трудоемкость в часах				Формы текущего контроля успеваемости	
		Всего	Контактная работа - Аудиторная работа				Самостоятельная работа
			Общая, в т.ч.:	Лекции	Семинары, практические занятия		

1.	Предмет, методы и инструменты анализа.	36	8	4	4	28	Решение задач на практических занятиях. Самостоятельная работа. Опрос.
2.	Разведочный анализ.	36	8	2	6	28	Решение задач на практических занятиях. Самостоятельная работа. Опрос.
3.	Проверка статистических гипотез.	36	8	2	6	28	Решение задач на практических занятиях.
4.	Множественная регрессия.	36	8	2	6	28	Решение задач на практических занятиях. Самостоятельная работа. Опрос.
5.	Многомерные статистические методы.	36	8	2	6	28	Решение задач на практических занятиях. Самостоятельная работа. Опрос.
В целом по дисциплине		180	40	10	30	140	Согласно учебному плану: контрольная работа
Итого в %		100	22	5	17	78	

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Предмет, методы и инструменты анализа.	Статистическое наблюдение в туризме. Импорт и экспорт данных. Знакомство с возможностями анализа больших данных с помощью электронных таблиц и языков программирования. Рекомендуемые источники: (8.1, 8.3)	Опрос. Проверка самостоятельной работы. Решение задач в интерактивной форме.
Разведочный анализ.	Предварительная обработка данных: форматирование, удаление пропусков и «мусора», аккуратизация, преобразование наборов данных. Выделение поднаборов. Визуализация формы распределения, рассеяния, связи, тенденции. Рекомендуемые источники: (8.2, 8.3)	Опрос. Проверка самостоятельной работы. Решение задач в интерактивной форме
Проверка статистических гипотез.	Формулировка и проверка гипотез о форме распределения и о числовых характеристиках признаков, характеризующих туристический бизнес. Рекомендуемые источники: (8.1)	Решение задач в интерактивной форме, проверка самостоятельной работы.
Множественная регрессия.	Анализ корреляционной матрицы. Спецификации регрессионной модели. Анализ качества модели множественной регрессии. Корректировка модели. Рекомендуемые источники: (8.1.)	Решение задач в интерактивной форме, проверка самостоятельной работы.
Многомерные статистические методы.	Снижение размерности модели и группировка признаков в туризме с использованием метода главных компонент. Задача классификации туристических объектов. Классификация с учителем. Классификация без учителя. Рекомендуемые источники: (8.1)	Решение задач в интерактивной форме, проверка самостоятельной работы.

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

6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Предмет, методы и инструменты анализа.	Наборы данных в туризме. Сбор данных о туристических объектах.	Работа с учебной литературой. Выполнение домашних заданий к занятию.
Разведочный анализ.	Сравнительный анализ доступных инструментальных средств подготовки, автоматизации описательной статистики, визуализации данных	Работа с учебной литературой. Знакомство со средами разработки. Разбор вопросов по теме занятия. Выполнение домашних заданий к занятию.
Проверка статистических гипотез.	Применимость параметрических критериев при проверке гипотез о значениях характеристик признаков в области туризма. Нормальное распределение признаков в туризме.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Множественная регрессия.	Нелинейные регрессионные модели в туризме.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Многомерные статистические методы.	Деревья решений. Машинное обучение.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию

Критерии бальной оценки различных форм текущего контроля успеваемости содержится в соответствующих методических рекомендациях кафедры.

7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по данной дисциплине

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Анализ больших данных».

Фонд оценочных средств включает контрольные материалы для проведения текущего контроля и промежуточной аттестации.

7.1. Описание показателей и критериев оценивания компетенций

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
ПKN-2 Способность использовать современные информационные системы и математические методы для решения математических, естественнонаучных, социально-экономических и профессиональных задач предметной области					
Демонстрирует знание современных информационных систем на основе и с применением математических моделей и методов.					
Знать: современные информационные системы на основе и с применением математических моделей и методов	Фрагментарное представление о современных информационных системах на основе и с применением математических моделей и методов	Неполные представления о современных информационных системах на основе и с применением математических моделей и методов	Сформированные, но содержащие отдельные пробелы представления о современных информационных системах на основе и с применением математических моделей и методов	Сформированные систематические представления о современных информационных системах на основе и с применением математических моделей и методов	Вопросы для оценки знаний и умений
Уметь: выявлять современные информационные системы на основе и с применением математических моделей и методов	Фрагментарное умение выявлять современные информационные системы на основе и с применением математических моделей и методов	Несистематическое применение умений выявлять современные информационные системы на основе и с применением математических	В целом успешное, но содержащее отдельные пробелы умение выявлять современные информационные системы на основе и с применением	Сформированное умение выявлять современные информационные системы на основе и с применением математических моделей и	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетвор ительно»	«удовлетвор ительно»	«хорошо»	«отлично»	
		моделей и методов	м математиче ских моделей и методов	методов	
Применяет полученные знания при решении математических, естественнонаучных, социально-экономических и профессиональных задач в междисциплинарном контексте.					
Знать: теоретическ ие основы при решении математиче ских, естественно научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	Фрагментар ное представлен ие о теоретическ их основах при решении математиче ских, естественно научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	Неполные представлен ия о теоретическ их основах при решении математиче ских, естественно научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	Сформиров анные, но содержащие отдельные пробелы представлен ия о теоретическ их основах при решении математиче ских, естественно научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	Сформиров анные систематиче ские представлен ия о теоретическ их основах при решении математиче ских, естественно научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	Вопросы для оценки знаний и умений
Уметь: применять полученные знания при решении математиче ских, естественно научных, социально-	Фрагментар ное умение применять полученные знания при решении математиче ских, естественно научных,	Несистемат ическое умений применять полученные знания при решении математиче ских, естественно	В целом успешное, но содержащее отдельные пробелы умение применять полученные знания при	Сформиров анное умение применять полученные знания при решении математиче ских, естественно	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетвор ительно»	«удовлетвор ительно»	«хорошо»	«отлично»	
экономичес ких и профессион альных задач в междисцип линарном контексте	социально- экономичес ких и профессион альных задач в междисцип линарном контексте	научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	решении математиче ских, естественно научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	научных, социально- экономичес ких и профессион альных задач в междисцип линарном контексте	
Владеет методами анализа Big Data, использует для решения профессиональных задач на макро-, мезо- и микроуровнях, в том числе на уровне финансового рынка.					
Знать: методы анализа Big Data, теоретическ ие основы для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Фрагментар ное представлен ие о методах анализа Big Data, теоретическ их основах для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Неполные представлен ия о методах анализа Big Data, теоретическ их основах для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Сформиров анные, но содержащие отдельные пробелы представлен ия о методах анализа Big Data, теоретическ их основах для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Сформиров анные систематиче ские представлен ия о методах анализа Big Data, теоретическ их основах для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
Уметь: владеет методами анализа Big Data, использует для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Фрагментар ное умение владеть методами анализа Big Data, использоват ь для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Несистемат ическое применение умений владения методами анализа Big Data, использова ния для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	В целом успешное, но содержащее отдельные пробелы умение владеть методами анализа Big Data, использоват ь для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Сформиров анное умение владеть методами анализа Big Data, использоват ь для решения профессион альных задач на макро-, мезо- и микроуровн ях, в том числе на уровне финансовог о рынка	Вопросы для оценки знаний и умений
ПКН-6 Способность анализировать и оценивать эффективность применения методов прикладной математики и информатики					
Демонстрирует знание основных методов прикладной математики и информатики, применяемых в различных предметных областях.					
Знать: основные методы прикладной математики и информатик и, применяемы х в различных предметных областях	Фрагментар ное представлен ие об основных методах прикладной математики и информатик и, применяем	Неполные представлен ия об основных методах прикладной математики и информатик и, применяем ых в	Сформиров анные, но содержащие отдельные пробелы представлен ия об основных методах прикладной математики и	Сформиров анные систематиче ские представлен ия об основных методах прикладной математики и информатик	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетвор ительно»	«удовлетвор ительно»	«хорошо»	«отлично»	
	ых в различных предметных областях	различных предметных областях	информатик и, применяем ых в различных предметных областях	и, применяем ых в различных предметных областях	
Уметь: различать основные методы прикладной математики и информатик и, применяемы е в различных предметных областях	Фрагментар ное умение различать основные методы прикладной математики и информатик и, применяем ые в различных предметных областях	Несистемат ическое применение умений различать основные методы прикладной математики и информатик и, применяем ые в различных предметных областях	В целом успешное, но содержащее отдельные пробелы умение различать основные методы прикладной математики и информатик и, применяем ые в различных предметных областях	Сформиров анное умение различать основные методы прикладной математики и информатик и, применяем ые в различных предметных областях	Вопросы для оценки знаний и умений
Владеет методологией математического моделирования для решения профессиональных задач.					
Знать: теоретическ ие основы методологии математичес кого моделирован ия для решения профессиона льных задач	Фрагментар ное представлен ие о теоретическ их основах методологи и математиче ского моделирова ния для	Неполные представлен ия о теоретическ их основах методологи и математиче ского моделирова ния для решения	Сформиров анные, но содержащие отдельные пробелы представлен ия о теоретическ их основах методологи и математиче	Сформиров анные систематиче ские представлен ия о теоретическ их основах методологи и математиче ского	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетвор ительно»	«удовлетвор ительно»	«хорошо»	«отлично»	
	решения профессион альных задач	профессион альных задач	ского моделирова ния для решения профессион альных задач	моделирова ния для решения профессион альных задач	
Уметь: владеть методологии математичес кого моделирован ия для решения профессиона льных задач	Фрагментар ное умение владеть методологи ей математиче ского моделирова ния для решения профессион альных задач	Несистемат ическое применение умений владеть методологи ей математиче ского моделирова ния для решения профессион альных задач	В целом успешное, но содержащее отдельные пробелы умение владеть методологи ей математиче ского моделирова ния для решения профессион альных задач	Сформиров анное умение владеть методологи ей математиче ского моделирова ния для решения профессион альных задач	Вопросы для оценки знаний и умений
ПК-2 Способность применять лингвистические и сетевые модели машинного обучения, предобученные на больших наборах данных.					
Демонстрирует знание современных предобученных лингвистических моделей или инструментов создания предобучения сетевых моделей.					
Знать: современные предобученн ые лингвистиче ские модели или инструменты создания предобучени я сетевых моделей	Фрагментар ное представлен ие о современны х предобучен ных лингвистич еских моделях или	Неполные представлен ия о современны х предобучен ных лингвистич еских моделях или инструмент	Сформиров анные, но содержащие отдельные пробелы представлен ия о современны х предобучен ных лингвистич	Сформиров анные систематиче ские представлен ия о современны х предобучен ных лингвистич еских	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
	инструментах создания предобучения сетевых моделей	ах создания предобучения сетевых моделей	еских моделях или инструментах создания предобучения сетевых моделей	моделях или инструментах создания предобучения сетевых моделей	
Уметь: различать современные предобученные лингвистические модели или инструменты создания предобучения сетевых моделей	Фрагментарное умение различать современные предобученные лингвистические модели или инструменты создания предобучения сетевых моделей	Несистематическое применение умений различать современные предобученные лингвистические модели или инструменты создания предобучения сетевых моделей	В целом успешное, но содержащее отдельные пробелы умение различать современные предобученные лингвистические модели или инструменты создания предобучения сетевых моделей	Сформированное умение различать современные предобученные лингвистические модели или инструменты создания предобучения сетевых моделей	Вопросы для оценки знаний и умений
Владеет методологией и инструментарием использования предобученных моделей для создания решения прикладных задач, разрабатывает на базе известных решений технологию решения прикладной задачи с помощью предобученных моделей.					
Знать: методологию и инструментальный использован предобученных моделей для создания решения	Фрагментарное представление о методологии и инструментарии использования	Неполные представления о методологии и инструментарии использования предобучен	Сформированные, но содержащие отдельные пробелы представления о методологии и инструментарии	Сформированные систематические представления о методологии и инструментарии	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
прикладных задач, теоретические основы для разработки на базе известных решений технологии решения прикладной задачи с помощью предобученных моделей	предобученных моделей для создания решения прикладных задач, теоретических основ для разработки на базе известных решений технологии решения прикладной задачи с помощью предобученных моделей	ных моделей для создания решения прикладных задач, теоретических основ для разработки на базе известных решений технологии решения прикладной задачи с помощью предобученных моделей	арии использования предобученных моделей для создания решения прикладных задач, теоретических основ для разработки на базе известных решений технологии решения прикладной задачи с помощью предобученных моделей	использования предобученных моделей для создания решения прикладных задач, теоретических основ для разработки на базе известных решений технологии решения прикладной задачи с помощью предобученных моделей	
Уметь: владеть методологией и инструментарием использования предобученных моделей для создания решения прикладных задач, разрабатыва	Фрагментарное умение владеть методологией и инструментарием использования предобученных моделей для создания решения прикладных	Несистематическое применение умений владеть методологией и инструментарием использования предобученных моделей для создания решения	В целом успешное, но содержащее отдельные пробелы умение владеть методологией и инструментарием использования предобученных	Сформированное умение владеть методологией и инструментарием использования предобученных моделей для создания решения прикладных	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетвор ительно»	«удовлетвор ительно»	«хорошо»	«отлично»	
ть на базе известных решений технологии решения прикладной задачи с помощью предобучен ных моделей	задач, разрабатыва ть на базе известных решений технологии решения прикладной задачи с помощью предобучен ных моделей	прикладных задач, разрабатыва ть на базе известных решений технологии решения прикладной задачи с помощью предобучен ных моделей	моделей для создания решения прикладных задач, разрабатыва ть на базе известных решений технологии решения прикладной задачи с помощью предобучен ных моделей	задач, разрабатыва ть на базе известных решений технологии решения прикладной задачи с помощью предобучен ных моделей	
Владеет практическим навыком применения технологии использования предобученных моделей для решения прикладных задач.					
Знать: теоретическ ие основы технологий использован ия предобученн ых моделей для решения прикладных задач	Фрагментар ное представлен ие о теоретическ их основах технологий использова ния предобучен ных моделей для решения прикладных задач	Неполные представлен ия о теоретическ их основах технологий использова ния предобучен ных моделей для решения прикладных задач	Сформиров анные, но содержащие отдельные пробелы представлен ия о теоретическ их основах технологий использова ния предобучен ных моделей для решения прикладных задач	Сформиров анные систематиче ские представлен ия о теоретическ их основах технологий использова ния предобучен ных моделей для решения прикладных задач	Вопросы для оценки знаний и умений
Уметь: владеть практически м навыком применения	Фрагментар ное умение владеть практическ	Несистемат ическое применение умений	В целом успешное, но содержащее	Сформиров анное умение владеть	Вопросы для оценки знаний и умений

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
технологии использования предобученных моделей для решения прикладных задач	им навыком применения технологии использования предобученных моделей для решения прикладных задач	владения практическим навыком применения технологии использования предобученных моделей для решения прикладных задач	отдельные пробелы умение владеть практическим навыком применения технологии использования предобученных моделей для решения прикладных задач	практическим навыком применения технологии использования предобученных моделей для решения прикладных задач	

7.2. Вопросы для оценки знаний и умений, характеризующих формирование компетенций

Шифр компетенции	Вопросы	Правильный ответ
ПKN-2	1. Что такое большие данные?	Big Data
	2. Какие основные характеристики больших данных?	Объем, разнообразие, скорость изменения
	3. Какие методы используются для анализа больших данных?	Машинное обучение, статистический анализ
	4. Что такое Hadoop?	Open-source проект Apache
	5. Как работает MapReduce?	Параллельная обработка, фазы Map и Reduce
	6. На основе чего используется Spark?	In-мемогу вычислений
ПKN-6	7. Какие подходы к визуализации данных используются для больших данных?	Графики, диаграммы, карты
	8. Какие метрики используются для оценки эффективности методов анализа больших данных?	Точность, полнота, F-мера, ROC-кривая
	9. Какие виды машинного обучения используются для анализа больших данных?	Классификация, регрессия, кластеризация

	10. Как работает алгоритм MapReduce?	Параллельная обработка
	11. Какое преимущество имеют in-memory вычисления?	Быстрое выполнение операций
	12. Как используется Spark?	Анализ данных в режиме реального времени
	13. Для чего используются сверточные нейронные сети?	Для обработки изображений и других типов данных
ПК-2	14. Для чего используется метод главных компонент?	Для деления матрицу исходных данных X на две части
	15. Какой метод применяется для изучения взаимосвязей между значениями переменных	Метод факторного анализа
	16. Какой анализ используется для принятия решения о том, какие переменные различают (дискриминируют) две или более возникающие совокупности (группы)?	Дискриминантный
	17. Цель какого анализа главной целью определяет нахождение групп схожих объектов в выборке?	Кластерного
	18. Для чего используются деревья решений?	Для решения задач классификации
	19. Перечислите один из критериев согласия	Критерий согласия хи-квадрат
	20. Что является основным недостатком наивного Байесовского метода?	Зависимость от распределения данных

7.3. Практико-ориентированные задания

Шифр компетенции	Практико-ориентированные задания	Правильный ответ
ПКН-2	1. Вы работаете в компании, которая занимается анализом текстовых данных. Вам необходимо определить настроение (положительное или отрицательное) отзывов о продукте на основе текста. Какую модель машинного обучения вы будете использовать?	LSTM (Long Short-Term Memory)
	2. Вам предоставлен набор данных с информацией о клиентах банка, включая их демографические характеристики и историю кредитования. Вам нужно предсказать, будет ли клиент задолженность по кредиту. Какую модель машинного обучения вы будете использовать?	Random Forest
	3. Вам необходимо классифицировать изображения пациентов на основе рентгеновских снимков для определения наличия или отсутствия пневмонии. Какую модель машинного обучения будет рациональнее использовать?	Convolutional Neural Network (CNN)
	4. Вам предоставлен набор данных с информацией	Time Series Forecasting

	о продажах продуктов в разных регионах. Вам нужно предсказать будущие продажи на основе исторических данных. Какую модель машинного обучения вы будете использовать?	
	5. Вам необходимо определить, является ли электронное письмо спамом или не спамом на основе его содержания и метаданных. Какую модель машинного обучения вы будете использовать?	Naive Bayes Classifier
	6. Вам предоставлен набор данных с информацией о покупках клиентов в интернет-магазине. Вам нужно сегментировать клиентов на основе их покупательского поведения. Какую модель машинного обучения вы будете использовать?	K-means Clustering
ПКН-6	7. Укажите метод анализа данных: представление данных в виде графиков и диаграмм	Визуализация
	8. Укажите метод анализа данных: группировка похожих точек данных вместе	Кластеризация
	9. Укажите метод анализа данных: предсказание числовых значений на основе входных данных	Регрессия
	10. Укажите метод анализа данных: разделение данных на различные категории или группы	Классификация
	11. Укажите метод анализа данных: выбор наиболее значимых переменных для модели	Отбор признаков
	12. Укажите метод анализа данных: определение структуры данных без меток классов	Обучение без учителя
ПК-2	13. Вам необходимо анализировать текстовые данные из социальных сетей для определения общественного мнения о продукте или услуге. Какую информационную систему и математический метод вы будете использовать?	Natural Language Processing (NLP)
	14. Вам необходимо определить оптимальное размещение рекламных объявлений на сайте для максимизации конверсии. Какую информационную систему и математический метод вы будете использовать?	A/B Testing
	15. Вам необходимо анализировать временные ряды данных о ценах на акции для прогнозирования будущих изменений. Какую информационную систему и математический метод вы будете использовать?	ARIMA или LSTM
	16. Вам необходимо проанализировать данные о погоде в разных регионах и предсказать вероятность возникновения наводнений. Какую информационную систему вы будете использовать?	Random Forest
	17. Вам предоставлены данные о клиентах ритейловой компании. Используя информационную систему и математические методы, определите сегменты клиентов для более эффективного маркетинга. Что вы будете использовать?	Hierarchical Clustering

	18. Вам необходимо определить, какие факторы оказывают наибольшее влияние на цену недвижимости на основе данных о характеристиках домов. Какую модель машинного обучения вы будете использовать?	Gradient Boosting (например, XGBoost или LightGBM)
--	--	--

7.4. Тесты

Шифр компетенции	Тестовые задания	Правильный ответ
ПКН-2	Что такое MapReduce? A) Это метод параллельного программирования B) Это язык программирования C) Это технология хранения данных D) Это алгоритм машинного обучения	D
	В каком году был создан Hadoop? A) 2000 B) 2003 C) 2005 D) 2008	D
	Что такое MapReduce? A) Это метод параллельного программирования B) Это язык программирования C) Это технология хранения данных D) Это алгоритм машинного обучения	D
	Как называется распределенная файловая система Hadoop? A) HDFS B) HBase C) Hive D) HIVE	A
ПКН-6	Какой язык используется для написания запросов в Hive? A) SQL B) PL/SQL C) Java D) Scala	A
	В какой СУБД используется язык SQL? A) Oracle B) MySQL C) PostgreSQL D) Все перечисленные	D
	Что из перечисленного является методом кластеризации, основанным на плотности? A) Иерархическая кластеризация B) Метод k-средних C) DBSCAN D) OPTICS	C
	Что такое машинное обучение? A) Это процесс обучения компьютера на основе данных	A

	В) Это процесс анализа данных С) Это процесс классификации данных D) Это процесс прогнозирования данных	
ПК-2	Что из нижеперечисленного является метрикой для оценки качества кластеризации? А) Индекс силуэта В) Коэффициент корреляции Пирсона С) Коэффициент дивергенции Кульбака-Лейблера D) Коэффициент Джини	А
	Какой из следующих методов не является методом определения числа факторов? А) Каменицкий-Трофимова В) Бартлетта С) Велша D) Кайзера	А
	Что из нижеследующего не является этапом дискриминантного анализа? А) Оценка различий между группами В) Выбор предикторов С) Построение дискриминантной функции D) Интерпретация результатов	Д
	Какую информацию дает каноническая дискриминантная функция? А) Относительную важность предикторов В) Вероятность принадлежности к группе С) Различия между группами D) Ничего из вышеперечисленного	В

8.Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература:

1. Соловьев В. И. Анализ данных в экономике: теория вероятностей, прикладная статистика, обработка и визуализация данных в MicrosoftExcel: учебник для направления бакалавриата "Экономика и управление" / В. И. Соловьев; Финунивер- ситет. - Москва: Кнорус, 2019. - 498 с. - Текст: непосредственный. - То же. - 2021. - ЭБС BOOK.ru. - URL: <https://book.ru/book/938856>. Текст: электронный
2. Криволапов С.Я. Статистические вычисления на платформе JupyterNotebookс использованием Python: учебник / С. Я. Криволапов. — Москва: КноРус, 2022. — 431 с. — ЭБС BOOK.ru. - URL: <https://book.ru/book/943660>. — Текст: электронный.
3. Анализ больших данных : учебное пособие / И. Б. Тесленко, В. Е. Крылов, А. М. Губернаторов [и др.]. — Москва : КноРус, 2023. — 295 с. — ISBN 978-5-406-10550-4. — URL: <https://book.ru/book/950469>. — Текст : электронный.

Дополнительная литература:

1. Глебов В. И. Анализ данных в экономике. Сборник задач: учебник / В.И. Глебов, С. Я. Криволапов. — Москва: КноРус, 2022. — 578 с. — ЭБС BOOK.ru. — URL: <https://book.ru/book/943011>. — Текст: электронный.
2. Целых, А. Н. Применение временных рядов для анализа больших данных : учебное пособие / А. Н. Целых, В. С. Васильев, Э. М. Котов. — Ростов-на-Дону : ЮФУ, 2021. — 83 с. — ISBN 978-5-9275-3983-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/247151>. — Режим доступа: для авториз. пользователей.
3. Управленческий учет и анализ. С примерами из российской и зарубежной практики / НИЦ ИНФРА-М, Петрова Валентина Ивановна, Петров Алексей Юрьевич, Кобищан Ирина Валентиновна, Козельцева Елена Андреевна, 2024. - 303 с. - Текст : электронный. - URL: <https://znanium.ru/catalog/document?id=435533>. – Режим доступа: по подписке.

9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://org.fa.ru/>
2. Сайт Департамента математики Факультета информационных технологий

- и анализа больших данных. <http://www.fa.ru/org/dep/dm/Pages/Home.aspx>
3. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>
 4. Электронно-библиотечная система BOOK.RU
<http://www.book.ru> (<http://www.book.ru/>)
 5. Электронно-библиотечная система «Университетская библиотека
ОНЛАЙН» <http://biblioclub.ru/>
 6. Электронно-библиотечная система Znanium <http://www.znaniy.com>
(<http://www.znaniy.com/>)
 7. Электронно-библиотечная система издательства «ЮРАЙТ»
 8. <https://www.urait.ru/>
 9. Электронно-библиотечная система издательства «Лань»
 10. <https://e.lanbook.com/>
 11. Деловая онлайн-библиотека AlpinaDigital <http://lib.alpinadigital.ru/>
 12. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
(<http://elibrary.ru/>)
 13. Национальная электронная библиотека Б.Ир://нэб.рф/
 14. Массовый открытый онлайн-курс/специализация «Машинное обучение и
анализ данных» / МФТИ и Яндекс.
<https://www.coursera.org/specializations/machine-learning-data-analysis>
 15. Массовый открытый онлайн-курс/специализация “Recommender Systems”/
University of Minnesota - <https://www.coursera.org/specializations/recommender-systems>
 16. Массовый открытый онлайн-курс/специализация “Machine Learning”/
Stanford University - <https://www.coursera.org/learn/machine-learning/home/welcome>
 17. Профессиональный ресурс по машинному обучению -
<https://stackoverflow.com> (<https://stackoverflow.com/>)
 18. Профессиональный ресурс по машинному обучению -
 19. <https://stackoverflow.com> (<https://stackoverflow.com/>)
 20. Платформа для соревнований по машинному обучению - www.kaggle.com

10. Методические указания для обучающихся по освоению дисциплины

Самостоятельная работа студентов проходит аудиторно и внеаудиторно. Организации самостоятельной работы служит учебно-тематический план изучения дисциплины. В этом плане указана тематика лекций, практических занятий, вопросы и задания для самостоятельного изучения.

Домашние задания следует выполнять регулярно при подготовке к практическим занятиям. В большинстве своем задания являются типовыми, и образцы их решения содержатся в рекомендованных пособиях, в материале лекций и практических занятий. Если то или иное задание вызвало затруднение, необходимо обратиться к преподавателю на консультации или ближайшем практическом занятии. Регулярность в выполнении домашних заданий — важный фактор освоения дисциплины. Даже небольшие отклонения от графика могут спровоцировать серьезное отставание и в дальнейшем — риск получения неудовлетворительных оценок в ходе текущей и промежуточной аттестации. Для выполнения домашних заданий следует завести отдельную тетрадь. Контроль выполнения домашних заданий осуществляется в ходе практических занятий в процессе выборочного собеседования.

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем

11.1. Комплект лицензионного программного обеспечения

1. ОС Astra Linux
2. Libre Office
3. Антивирус Kaspersky

11.2 Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Гарант»
2. Информационно-правовая система «Консультант Плюс»
3. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>
4. Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>

11.3 Сертифицированные программные и аппаратные средства защиты информации - не используются

11.4 Электронная таблица Calc LibreOffice

11.5 Прикладной программный пакет для эконометрического моделирования Gretl

11.6 Программная среда R.

12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Практические занятия по дисциплине проходят в компьютерном классе.