

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

Краснодарский филиал Финуниверситета

Кафедра математики и информатики

СОГЛАСОВАНО

ООО «Портал-Юг»
Генеральный директор



Е.В. Мостовой

«20» февраля 2024 г.

УТВЕРЖДАЮ

Краснодарский филиал
Финуниверситета

Директор



Э.В.Соболев

«20» февраля 2024 г.

Калайдин Е.Н.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ВВЕДЕНИЕ В НАУКИ О ДАННЫХ**

для студентов, обучающихся по направлению подготовки
01.04.02 Прикладная математика и информатика,
Направленность программы магистратуры:
Анализ больших данных и машинное обучение в экономике и финансах

*Рекомендовано Ученым советом Краснодарского филиала Финуниверситета
(протокол № 12 от 20.02.2024)*

*Одобрено кафедрой «Математика и информатика»
(протокол № 13 от 13.02.2024)*

Краснодар 2024

УДК 303.436
ББК 60.601.3
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. Введение в науки о данных. Рабочая программа дисциплины для студентов, обучающихся по направлению подготовки 01.04.02 «Прикладная математика и информатика», направленность программы магистратуры «Анализ больших данных и машинное обучение в экономике и финансах». - Краснодар: Краснодарский филиал Финансового университета при Правительстве Российской Федерации, Кафедра математики и информатики, 2024.

Дисциплина **Введение в науки о данных** относится к Общенаучному модулю направления 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, место в структуре образовательной программы, требования к результатам освоения дисциплины, содержание программы, тематика практических занятий, формы самостоятельной работы, оценочные средства для текущего контроля и промежуточной аттестации, учебно-методическое и информационное обеспечение.

Калайдин Е.Н.

Введение в науки о данных

*Формат 60*90/16. Гарнитура Times New Roman*

Усл. п.л. 2,0. Изд. № _от.

Тираж 100 экз.

Заказ № .

Отпечатано в Краснодарском филиале Финуниверситета

© Калайдин Е.Н., 2024
© Краснодарский филиал Финансового университета, 2024

СОДЕРЖАНИЕ

1. Наименование дисциплины4
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине4
3. Место дисциплины в структуре образовательной программы6
4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся6
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий7
6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине11
7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины24
8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины25
9. Методические указания для обучающихся по освоению дисциплины265
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем276
11. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине276

1. Наименование дисциплины

Дисциплина «Введение в науки о данных».

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

Дисциплина «Введение в науки о данных» обеспечивает инструментарий формирования следующих компетенций: УК-1, УК-3.

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/индикаторами достижения компетенции
УК-1	Способность к абстрактному мышлению, критическому анализу проблемных ситуаций на основе системного подхода, выработке стратегии действий	1. Использует методы абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности.	Знать: методы абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности Уметь: использовать методы абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности.
		2. Демонстрирует способы осмысления и критического анализа проблемных ситуаций.	Знать: способы осмысления и критического анализа проблемных ситуаций Уметь: демонстрировать способы осмысления и критического анализа проблемных ситуаций

		3.Предлагает нестандартное решение проблем, новые оригинальные проекты, вырабатывает стратегию действий на основе системного подхода	Знать: теоретические основы нестандартных решений проблем, новых оригинальных проектов, выработки стратегии действий на основе системного подхода Уметь: предлагать нестандартное решение проблем, новые оригинальные проекты, вырабатывать стратегию действий на основе системного подхода
УК-3	Способность определять и реализовывать приоритеты собственной деятельности в соответствии с важностью задач, методы повышения ее эффективности.	1. Объективно оценивает свои возможности и требования различных социальных ситуаций, принимает решения в соответствии с данной оценкой и требованиями.	Знать: теоретические основы оценки своих возможностей и требований различных социальных ситуаций, принятия решения в соответствии с данной оценкой и требованиями Уметь: объективно оценивать свои возможности и требования различных социальных ситуаций, принимать решения в соответствии с данной оценкой и требованиями
		2. Актуализирует свой личностный потенциал, внутренние источники роста и развития собственной деятельности.	Знать: теоретические основы для актуализации своего личностного потенциала, внутренних источников роста и развития собственной деятельности Уметь: осуществлять актуализацию своего личностного потенциала, внутренних источников роста и развития собственной деятельности
		3. Определяет приоритеты собственной деятельности в соответствии с важностью задач.	Знать: приоритеты собственной деятельности в соответствии с важностью задач Уметь: определять приоритеты собственной деятельности в соответствии с важностью задач

		4. Определяет и демонстрирует методы повышения эффективности собственной деятельности.	Знать: методы повышения эффективности собственной деятельности Уметь: определять и демонстрировать методы повышения эффективности собственной деятельности
--	--	----------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------

3. Место дисциплины в структуре образовательной программы

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

В процессе изучения дисциплины происходит знакомство с основными понятиями, методологиями, моделями, методами, методиками и технологиями обработки больших данных и машинного обучения, знания о которых будут положены в основу формирования профессиональных компетенций выпускника направления подготовки «Прикладная математика и информатика» направленность программы магистратуры «Анализ больших данных и машинное обучение в экономике и финансах». При этом студенты приобретают опыт практического использования изучаемых технологий в практических задачах, связанных с финансовыми технологиями.

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 з/е и часах)	Модуль 2 (в часах)
Общая трудоемкость дисциплины	5/180	180
Контактная работа - Аудиторные занятия	48	48
<i>Лекции</i>	<i>16</i>	<i>16</i>
<i>Семинары, практические занятия</i>	<i>32</i>	<i>32</i>
Самостоятельная работа	132	132
В семестре	96	96
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

1. От данных к ценности: Основы наук о данных

Основные определения: интеллектуальный анализ данных, большие данные, машинное обучение.

Методы и задачи интеллектуального анализа данных, машинного обучения и обработки больших данных.

Области применения методов и технологий интеллектуального анализа данных, машинного обучения и обработки больших данных.

Примеры задач машинного обучения: поиск информации в интернете, распознавание изображений, лиц, эмоций, пола, возраста, ..., распознавание речи, языка, эмоциональной окраски текстов, прогнозирование продаж, прогнозирование оттока клиентов, кредитный скоринг, рекомендательные системы и др.

Основные характеристики больших данных и их влияние на сбор, хранение, обработку и анализ данных (4V).

Критерии аналитических задач, решение которых предпочтительно с использованием технологий Big Data.

Цикл обработки данных: поиск данных, сбор данных, очистка данных, трансформация данных, интеллектуальный анализ данных, интерпретация и практическое применение результатов.

Статистические основы обработки данных: краткий обзор выборочного метода, методов описательной статистики, построения интервальных оценок и проверки гипотез.

Обзор современных технологий машинного обучения: R, Python, Spark, Microsoft Azure ML.

2. Практическое использование моделей классификации и регрессии

Принципы разработки и оценки систем машинного обучения. Основные классы моделей машинного обучения: классификация с обучением, регрессионный анализ, кластерный анализ, поиск аномалий.

Задача прогнозирования продаж. Модель множественной линейной регрессии. Методы оценки качества моделей регрессии. Коэффициент детерминации, средняя абсолютная ошибка предсказания, средняя относительная ошибка предсказания.

Задача кредитного скоринга. Модель логистической регрессии и ее реализация в R/Python/Azure ML. Методы оценки качества моделей классификации: доля правильных ответов, точность, полнота, F1, AUC. Цены ошибок первого и второго рода.

Модели деревьев и лесов решений и их компьютерная реализация.

Усиление (бустинг) деревьев решений.

Нейронные сети и их компьютерная реализация.

Машины опорных векторов и их компьютерная реализация.

Технологии улучшения моделей машинного обучения. Подбор оптимальных параметров моделей. Важность подготовки данных. Генерация синтетических признаков. Работа с пропущенными данными. Работа с несбалансированными выборками.

3. Практическое использование моделей кластерного анализа и поиска аномалий

Кластерный анализ и его компьютерная реализация. Задача сегментирования потребителей. Иерархические агломеративные алгоритмы. Метод К-средних.

Методы оценки качества моделей кластерного анализа. Расстояния между кластерами, расстояния между объектами внутри кластеров. Характеристики центров кластеров.

Задача снижения размерности факторного пространства. Метод главных компонент и его компьютерная реализация. Геометрическое и экономическое содержание метода главных компонент. Использование методов снижения размерности для улучшения качества моделей машинного обучения. Использование методов снижения размерности для выявления латентных факторов.

Поиск аномалий. Задача поиска мошеннических транзакций.

4. Анализ текстов, изображений, эмоций и рекомендательные системы

Принципы анализа текстовой информации.

Принципы анализа графической информации.

Принципы анализа эмоциональной окраски текстов.

Принципы создания рекомендательных систем.

Интеллектуальные сервисы и чат-боты.

Перспективы развития систем обработки больших данных и машинного обучения.

Финансовые технологии, основанные на обработке данных и машинном обучении: интеллектуальные кредитные сервисы, интеллектуальные страховые сервисы, интеллектуальные сервисы интернета вещей.

Современные практические исследования Финансового университета в области обработки данных и машинного обучения: индикаторы смены состояний финансовых рынков, анализ сложных сетей в управлении городами, энергетике, на транспорте, предиктивный ремонт оборудования, системы скоринга.

5.2. Учебно-тематический план

№ п/п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Аудиторная работа			Самостоятельная работа	
			Общая, в т.ч.:	Лекции	Семинары, практические занятия		
1	От данных к ценности: Ос-новы наук о данных	40	12	2	8	30	Лаборатор-ные работы (построение компьютер-ной системы визуализации данных, анализ конкретной ситуации и обсуждение результатов)
2	Практическое использование моделей классификации и регрессии	40	12	6	8	30	
3	Практическое использование моделей кластерного ана-лиза и поиска аномалий	40	12	4	8	30	
4	Анализ тек-стов,изображений, эмоций и рекомендательные системы	60	12	4	8	42	
	В целом по дисциплине	180	48	16	32	132	Контрольная работа
	Итого в %	100	27	9	18	73	

5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарских, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
1. От данных к ценности: Основы наук о данных	Построение компьютерной системы визуализации данных, анализ конкретной ситуации и обсуждение результатов. Построение компьютерной системы визуализации данных, анализ конкретной ситуации и обсуждение результатов - 100% от трудоемкости семинарского занятия. <i>Рекомендуемые источники:</i> п.8,[1], [2],[6]; п.9, [4]-[7]	Лабораторная работа, интерактивная форма.
2. Практическое использование моделей классификации и регрессии	Построение системы прогнозирования продаж. Построение системы кредитного скоринга. Оптимизация системы кредитного скоринга. <i>Рекомендуемые источники:</i> п.8, [1]-[8]; п. 9, [4]-[9]	Лабораторная работа, интерактивная форма.
3. Практическое использование моделей кластерного анализа и поиска аномалий	Кластерный анализ и его компьютерная реализация. Иерархические агломеративные алгоритмы. Метод К-средних. Методы оценки качества моделей кластерного анализа. Расстояния между кластерами, расстояния между объектами. Поиск аномалий. Задача поиска мошеннических транзакций. Сегментирование потребителей. <i>Рекомендуемые источники:</i> п.8, [2]-[5]; п. 9, [4]-[9]	Лабораторная работа, интерактивная форма.
4. Анализ текстов, изображений, эмоций и рекомендательные системы	Финансовые технологии, основанные на обработке данных и машинном обучении: интеллектуальные кредитные сервисы, интеллектуальные страховые сервисы, интеллектуальные сервисы интернета вещей. Построение рекомендательного сервиса. Анализ эмоциональной окраски записей в социальных сетях. <i>Рекомендуемые источники:</i> : п.8, [2]-[5]; п. 9, [4]-[9]	Лабораторная работа, интерактивная форма.

6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Введение в науки о данных».

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

6.1. Описание показателей и критериев оценивания компетенций

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
УК-1 Способность к абстрактному мышлению, критическому анализу проблемных ситуаций на основе системного подхода, выработке стратегии действий					
Использует методы абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности.					
Знать: методы абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	Фрагментарное представление о методах абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	Неполные представления о методах абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	Сформированные, но содержащие отдельные пробелы представления о методах абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	Сформированные систематические представления о методах абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
Уметь: использовать методы абстрактного мышления,	Фрагментарное умение использовать методы абстрактного	Несистематическое применение умений использования	В целом успешное, но содержащее отдельные пробелы	Сформированное умение использовать методы абстрактного	Вопросы для оценки знаний и умений, практико-

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности и	мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	я методов абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	умение использовать методы абстрактного мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности	мышления, анализа информации и синтеза проблемных ситуаций, формализованных моделей процессов и явлений в профессиональной деятельности и	ориентированные задания, тесты
Демонстрирует способы осмысления и критического анализа проблемных ситуаций.					
Знать: способы осмысления и критического анализа проблемных ситуаций	Фрагментарное представление о способах осмысления и критического анализа проблемных ситуаций	Неполные представления о способах осмысления и критического анализа проблемных ситуаций	Сформированные, но содержащие отдельные пробелы представления о способах осмысления и критического анализа проблемных ситуаций	Сформированные систематические представления о способах осмысления и критического анализа проблемных ситуаций	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
Уметь: демонстрировать способы осмысления и критического анализа проблемных ситуаций	Фрагментарное умение в демонстрации способов осмысления и критического анализа проблемных ситуаций	Несистематическое применение способов осмысления и критического анализа проблемных ситуаций	В целом успешное, но содержащее отдельные пробелы в демонстрации и способов осмысления и критического анализа проблемных ситуаций	Сформированное умение в демонстрации способов осмысления и критического анализа проблемных ситуаций	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
Предлагает нестандартное решение проблем, новые оригинальные проекты, вырабатывает					

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
стратегию действий на основе системного подхода					
Знать: нестандартные решения проблем, новые оригинальные проекты, вырабатывать стратегию действий на основе системного подхода	Фрагментарное представление о нестандартных решениях проблем, новых оригинальных проектах, на основе системного подхода	Неполные представления о нестандартных решениях проблем, новых оригинальных проектах, на основе системного подхода	Сформированные, но содержащие отдельные пробелы в нестандартных решениях проблем, новых оригинальных проектах, на основе системного подхода	Сформированные систематические представления о нестандартных решениях проблем, новых оригинальных проектах, на основе системного подхода	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
Уметь: предлагать нестандартное решение проблем, новые оригинальные проекты, вырабатывать стратегию действий на основе системного подхода	Фрагментарное умение предлагать нестандартное решение проблем, новые оригинальные проекты, вырабатывать стратегию действий на основе системного подхода	Несистематическое применение умений предлагать нестандартное решение проблем, новые оригинальные проекты, вырабатывать стратегию действий на основе системного подхода	В целом успешное, но содержащее отдельные пробелы умение предлагать нестандартное решение проблем, новые оригинальные проекты, вырабатывать стратегию действий на основе системного подхода	Сформированное умение предлагать нестандартное решение проблем, новые оригинальные проекты, вырабатывать стратегию действий на основе системного подхода	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
УК-3 Способность определять и реализовывать приоритеты собственной деятельности в соответствии с важностью задач, методы повышения ее эффективности.					
Объективно оценивает свои возможности и требования различных социальных ситуаций, принимает решения в соответствии с данной оценкой и требованиями.					
Знать: возможности и требования	Фрагментарное представление о	Неполные представления о	Сформированные, но содержащие	Сформированные систематические	Вопросы для оценки знаний и

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
различных социальных ситуаций	возможностях и требованиях различных социальных ситуаций	возможностях и требованиях различных социальных ситуаций	отдельные пробелы представления о возможностях и требованиях различных социальных ситуаций	ские представления о возможностях и требованиях различных социальных ситуаций	умений, практико-ориентированные задания, тесты
Уметь Объективно оценивать свои возможности и требования различных социальных ситуаций, принимать решения в соответствии с данной оценкой и требованиями и	Фрагментарное умение объективно оценивать свои возможности и требования различных социальных ситуаций, принимать решения в соответствии с данной оценкой и требованиями	Несистематическое применение умений объективно оценивать свои возможности и требования различных социальных ситуаций, принимать решения в соответствии с данной оценкой и требованиями и	В целом успешное, но содержащее отдельные пробелы умение объективно оценивать свои возможности и требования различных социальных ситуаций, принимать решения в соответствии с данной оценкой и требованиями и	Сформированное умение объективно оценивать свои возможности и требования различных социальных ситуаций, принимать решения в соответствии с данной оценкой и требованиями и	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
Актуализирует свой личностный потенциал, внутренние источники роста и развития собственной деятельности					
Знать: теоретические основы для актуализации и своего личностного потенциала, внутренних источников роста и развития	Фрагментарное представление об актуализации своего личностного потенциала, внутренних источников роста и развития собственной	Неполные представления об актуализации и своего личностного потенциала, внутренних источников роста и развития собственной	Сформированные, но содержащие отдельные пробелы представления об актуализации и своего личностного потенциала, внутренних	Сформированные систематические представления об актуализации своего личностного потенциала, внутренних	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
собственной деятельности и	деятельности	деятельности	источников роста и развития собственной деятельности	источников роста и развития собственной деятельности и	
Уметь: осуществлять актуализацию своего личностного потенциала, внутренних источников роста и развития собственной деятельности и	Фрагментарное умение осуществлять актуализацию своего личностного потенциала, внутренних источников роста и развития собственной деятельности	Несистематическое применение умений осуществлять актуализацию своего личностного потенциала, внутренних источников роста и развития собственной деятельности	В целом успешное, но содержащее отдельные пробелы умение осуществлять актуализацию своего личностного потенциала, внутренних источников роста и развития собственной деятельности	Сформированное умение осуществлять актуализацию своего личностного потенциала, внутренних источников роста и развития собственной деятельности и	Вопросы для оценки знаний и умений, практико- ориентированные задания, тесты
Определяет приоритеты собственной деятельности в соответствии с важностью задач					
Знать: приоритеты собственной деятельности и в соответствии с важностью задач	Фрагментарное представление о приоритетах собственной деятельности в соответствии с важностью задач	Неполные представления о приоритетах собственной деятельности в соответствии с важностью задач	Сформированные, но содержащие отдельные пробелы представления о приоритетах собственной деятельности в соответствии с важностью задач	Сформированные систематические представления о приоритетах собственной деятельности и в соответствии с важностью задач	Вопросы для оценки знаний и умений, практико- ориентированные задания, тесты

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
Уметь: определять приоритеты собственной деятельности и в соответствии с важностью задач	Фрагментарное умение определять приоритеты собственной деятельности в соответствии с важностью задач	Несистематическое применение умений определять приоритеты собственной деятельности в соответствии с важностью задач	В целом успешное, но содержащее отдельные пробелы умение определять приоритеты собственной деятельности в соответствии с важностью задач	Сформированное умение определять приоритеты собственной деятельности и в соответствии с важностью задач	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
Определяет и демонстрирует методы повышения эффективности собственной деятельности					
Знать: методы повышения эффективности собственной деятельности	Фрагментарное представление о методах повышения эффективности собственной деятельности	Неполные представления о методах повышения эффективности и собственной деятельности	Сформированные, но содержащие отдельные пробелы представления о методах повышения эффективности собственной деятельности	Сформированные систематические представления о методах повышения эффективности собственной деятельности	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты
Уметь: определять и демонстрировать методы повышения эффективности собственной деятельности	Фрагментарное умение определять и демонстрировать методы повышения эффективности собственной деятельности	Несистематическое применение умений определять и демонстрировать методы повышения эффективности и собственной деятельности	В целом успешное, но содержащее отдельные пробелы умение определять и демонстрировать методы повышения эффективности собственной деятельности	Сформированное умение определять и демонстрировать методы повышения эффективности собственной деятельности	Вопросы для оценки знаний и умений, практико-ориентированные задания, тесты

6.2. Задания или иные материалы, необходимые для оценки знаний и умений, характеризующих формирование компетенций в процессе освоения ОП ВО

6.2.1 Вопросы для оценки знаний и умений, характеризующих формирование компетенций

Шифр компетенции	Вопросы	Правильный ответ
УК-1	1. Что такое системный подход?	Способ анализа сложных систем
	2. Какие методы используются для анализа данных?	Статистические методы, машинное обучение, data mining.
	3. Что означает термин “стратегия”?	План действий для достижения определенной цели
	4. Какие ключевые элементы включает в себя системный подход?	Цель, входные данные, выходные данные, ограничения и обратная связь
	5. Что такое data mining?	Процесс поиска скрытых закономерностей и взаимосвязей в больших наборах данных
	6. Как называются параметры, значения которых задается до начала обучения модели и не изменяется в процессе обучения?	Гиперпараметры модели
	7. Для чего используется линейная регрессия?	Для предсказания непрерывной целевой переменной
	8. Какие метрики используются для оценки качества модели машинного обучения?	Точность, полнота, F1-мера
	9. Какой из машинного обучения используется для решения задачи классификации?	Метод опорных векторов
	10. Какие этапы включает процесс извлечения знаний из данных?	Подготовка данных, выбор модели, обучение, оценка
УК-3	11. Астроном хочет построить модель, которая сможет разбить известные науке звезды на группы по их характеристикам, чтобы лучше изучить их особенности. К какому типу относится данная задача?	Кластеризации
	12. Чему соответствует геометрическое содержание метода главных компонент?	РСА находит направления, вдоль которых данные имеют наибольшую дисперсию.
	13. Какие виды аномалий в данных вы знаете?	Выбросы, отсутствующие значения, дубликаты
	14. Какие методы можно использовать для извлечения ключевых слов из текста?	TF-IDF, RAKE, TextRank
	15. Какие основные принципы композиции графических элементов?	Принципы баланса, акцента, направления и пропорции

	16. Какие методы используются для определения эмоциональной окраски текста?	Анализ тональности, использование эмоциональных словарей, машинное обучение
	17. Что такое рекомендательные системы и какие виды они бывают?	Коллаборативная фильтрация, контент-фильтрация
	18. Как называется процесс обработки естественного языка, который используется для анализа и понимания текста, который вводит пользователь?	Natural Language Processing (NLP)
	19. Какие перспективы развития систем обработки больших данных и машинного обучения вы видите?	Развитие облачных технологий, улучшение алгоритмов машинного обучения
	20. Какие финансовые технологии основаны на обработке данных и машинном обучении?	Алгоритмический трейдинг, кредитный скоринг

6.2.2 Практико-ориентированные задания

Шифр компетенции	Практико-ориентированные задания	Правильный ответ
УК-1	1. Опишите основные этапы жизненного цикла проекта анализа данных.	Идентификация, подготовка данных, анализ, интерпретация, деплоймент
	2. Какие типы данных могут быть использованы для анализа? Приведите примеры	Числовые (температура), категориальные (цвет), текстовые (отзывы)
	3. Вам необходимо предобработать данные перед анализом. Какие методы можно использовать для этого?	Удаление выбросов, заполнение пропущенных значений, нормализация
	4. Задача: оценить качество моделей машинного обучения. Какие метрики используются для этого?	Точность, полнота, F1-мера, AUC-ROC
	5. Какие техники можно использовать для балансировки классов в задачах классификации?	Undersampling, oversampling, использование взвешенных функций потерь
	6. Какие виды задач машинного обучения с учителем вы знаете? Приведите примеры.	Классификация (разделение объектов на классы), регрессия (предсказание числового значения)
	7. Вам необходимо улучшить производительность модели машинного обучения? Что вы будете использовать?	Сложные модели, подбор оптимальных гиперпараметров
	8. Приведите примеры типов графиков, которые используются для визуализации распределения числовых данных.	Гистограмма, ящик с усами
	9. Какие методы машинного обучения можно использовать для задачи классификации?	Логистическая регрессия, метод опорных векторов, случайный лес
УК-3	10. У компании задача: к презентации проекта необходимо визуализировать данные? Какие инструменты лучше всего	Matplotlib, Seaborn, Tableau

	использовать?	
	11. У вас есть «шум» в данных. По каким причинам он мог возникнуть?	Некорректные данные, выбросы
	12. В ходе анализа возник "перекос" (skewness) в распределении данных. Как его можно исправить?	Можно применить логарифмирование
	13. Какой метод можно использовать для оценки значимости признаков в модели машинного обучения в случайном лесу?	Важность признаков,
	14. Какой метод можно использовать для оценки значимости признаков в модели машинного обучения в линейной регрессии?	Коэффициенты
	15. Вам необходимо определить наиболее популярные товары или услуги в определенном регионе. Какой первый шаг вы предпримите?	Сбор данных о продажах товаров и услуг
	16. Вам предоставлен набор данных о продажах компании. Какие показатели можно выявить для того, чтобы сделать выводы о динамике продаж?	Средние, медиана, максимум, минимум. Продажи растут/падают
	17. Необходимо провести анализ текстовых отзывов о продукте/услуге с помощью методов обработки естественного языка. Что вы будете использовать?	Natural Language Processing
	18. Какие преимущества и недостатки имеет такой метод кластеризации как K-средних?	Прост в реализации, но чувствителен к начальным значениям центров кластеров

6.2.3 Тесты

Шифр компетенции	Тестовые задания	Правильный ответ
УК-1	1. Определите в каких случаях целесообразно использовать машинное обучение. а) Когда существует большое количество данных и требуется выявить закономерности и зависимости между ними. б) Когда требуется автоматизация процесса принятия решений. в) Когда необходимо создать систему, способную обучаться на основе опыта и адаптироваться к новым данным. г) Все варианты верны.	d)
	2. Выберите основные этапы, которые существуют в процессе анализа данных? а) Сбор данных, обработка данных, анализ данных. б) Подготовка данных, моделирование данных, интерпретация результатов.	d)

	<ul style="list-style-type: none"> с) Получение данных, очистка данных, визуализация данных. д) Определение задачи, сбор данных, обработка и анализ данных, интерпретация результатов и принятие решений. 	
	<p>3. “Большие данные” — это ...</p> <ul style="list-style-type: none"> а) Данные, которые имеют большой объем, скорость поступления и разнообразие. б) Данные, имеющие высокую ценность для организации. с) Данные, полученные из различных источников. д) Данные, хранящиеся в базах данных. 	a)
	<p>4. Выберите задачу, которая решается с помощью регрессионного анализа.</p> <ul style="list-style-type: none"> а) Классификация объектов. б) Прогнозирование значений зависимой переменной. с) Выявление взаимосвязей между переменными. д) Идентификация аномалий в данных. 	b)
	<p>5. Кластеризация — это ...</p> <ul style="list-style-type: none"> а) Разделение данных на группы (кластеры) таким образом, чтобы объекты внутри одной группы были похожи, а объекты разных групп отличались. б) Метод обработки данных, позволяющий выявить скрытые закономерности и взаимосвязи. с) Задача классификации объектов по заданным классам. д) Задача определения структуры данных и их взаимосвязей. 	a)
	<p>6. Анализ главных компонент — это ...</p> <ul style="list-style-type: none"> а) Метод уменьшения размерности данных путем определения основных компонентов, на которые приходится наибольший вариационный размах. б) Статистический метод обработки данных, который позволяет определить наиболее значимые факторы, влияющие на исследуемый процесс. с) Алгоритм машинного обучения для прогнозирования значений зависимой переменной. д) Технология обработки больших данных. 	a)
УК-3	<p>7. Определите метод, который можно использовать для оптимизации использования времени.</p> <ul style="list-style-type: none"> а) Составление списка дел и приоритетов б) Разбиение задач на подзадачи с) Установка сроков выполнения задач д) Отдых и переключение между задачами 	a)
	<p>8. Выберите один или несколько вариантов ответа,</p>	c)

	<p>которые не являются технологией, которую можно использовать для повышения эффективности работы в области анализа данных?</p> <ul style="list-style-type: none"> a) Машинное обучение b) Визуализация данных c) Социальные сети d) Базы данных 	
	<p>9. Целью анализа данных является ...</p> <ul style="list-style-type: none"> a) Поиск закономерностей и зависимостей b) Классификация и прогнозирование c) Обработка и визуализация d) Принятие решений 	a)
	<p>10. Data Science — это ...</p> <ul style="list-style-type: none"> a) Наука о методах анализа данных для получения новых и полезных знаний b) Процесс обработки и анализа данных с использованием математических и статистических методов c) Область деятельности, связанная с анализом и обработкой данных с помощью информационных технологий d) Все перечисленные ответы 	d)
	<p>11. Выберите из ниже представленных методов, который не относится к методам повышения эффективности работы</p> <ul style="list-style-type: none"> a) Улучшение организации рабочего процесса b) Оптимизация использования времени c) Использование современных технологий d) Занятия спортом 	d)
	<p>12. Определите какой из этих вариантов ответа не является приоритетом в работе.</p> <ul style="list-style-type: none"> a) Важные и срочные задачи b) Задачи, требующие творческого подхода c) Рутинная работа d) Задачи, направленные на улучшение результатов 	c)

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

8.1. Основная литература

1. Котельников, Е. В. Введение в машинное обучение и анализ данных : учебное пособие / Е. В. Котельников, А. В. Котельникова. — Киров : ВятГУ, 2024. — 68 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/390698> — Режим доступа: для авториз. пользователей.

2. Калинина, В. Н., Анализ данных. Компьютерный практикум : учебное пособие / В. Н. Калинина, В. И. Соловьев. — Москва : КноРус, 2022. — 166 с. — ISBN 978-5-406-09229-3. — URL: <https://book.ru/book/942681>. — Текст : электронный.

7.2. Дополнительная литература

1. Колдаев, В. Д. Структуры и алгоритмы обработки данных : учебное пособие / В.Д. Колдаев. — Москва : РИОР : ИНФРА-М, 2021. — 296 с. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/2833. - ISBN 978-5-369-01264-2. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1230215>. — Режим доступа: по подписке.

2. Максимова, О. В. Математическая статистика и анализ данных : учебное пособие / О. В. Максимова. — Москва : МИСИС, 2024. — 172 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/395690> — Режим доступа: для авториз. пользователей.

3. Ростовцев, В. С. Искусственные нейронные сети / В. С. Ростовцев. — 3-е изд., стер. — Санкт-Петербург : Лань, 2024. — 216 с. — ISBN 978-5-507-47362-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/364517> — Режим доступа: для авториз. пользователей.

4. Шихова, О. А. Анализ данных в экономике с использованием возможностей MS Excel : учебно-методическое пособие / О. А. Шихова. — Вологда : ВГМХА им. Н.В. Верещагина, 2022. — 101 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/313985> — Режим доступа: для авториз. пользователей.

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

1. Информационно-образовательный портал Финансового университета при Правительстве Российской Федерации <http://portal.ufrf.ru/>.
2. Сайт кафедры Департамента анализа данных и машинного обучения. <http://findata.fa.ru>
3. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/> (<http://library.fa.ru/files/elibfa.pdf>)
4. Документация по Azure ML <https://docs.microsoft.com/ru-ru/azure/machine-learning/studio/>
5. Applied Machine Learning / Microsoft. - <https://www.edx.org/course/applied-machine-learning-microsoft-dat203-3x-1>
6. Data Science Essentials / Microsoft. - <https://www.edx.org/course/data-science-essentials-microsoft-dat203-1x-3>
7. Principles of Machine Learning / Microsoft. - <https://www.edx.org/course/principles-machine-learning-microsoft-dat203-2x-3>
8. Профессиональный ресурс по машинному обучению. <https://machinelearning.ru>
9. Специализация «Машинное обучение и анализ данных» / МФТИ и Яндекс. - <https://www.coursera.org/specializations/machine-learning-data-analysis>
10. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>
11. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
12. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
13. Электронно-библиотечная система Znanium <http://www.znaniyum.com>
14. Электронно-библиотечная система издательства «ЮРАЙТ» <https://urait.ru/>
15. Электронно-библиотечная система издательства Проспект <http://ebs.prospekt.org/books>
16. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
17. Электронная библиотека Издательского дома «Гребенников» <https://grebennikon.ru/>
18. Деловая онлайн-библиотека Alpina Digital <http://lib.alpinadigital.ru/>
19. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
20. Национальная электронная библиотека <http://нэб.рф/>
21. Финансовая справочная система «Финансовый директор» <http://www.lfd.ru/>
22. Ресурсы информационно-аналитического агентства по финансовым рынкам Cbonds.ru <https://cbonds.ru/>
23. СПАРК <https://spark-interfax.ru/>
24. Academic Reference <http://ar.cnki.net/ACADREF>
25. Bank Focus <http://library.fa.ru/resource.asp?id=527>
26. Пакет баз данных компании EBSCO Publishing, крупнейшего агрегатора научных ресурсов ведущих издательств мира <http://search.ebscohost.com>
27. Электронные продукты издательства Elsevier

<http://www.sciencedirect.com>

28. Emerald: Management eJournal Portfolio <https://www.emerald.com/insight/>
29. Информационно-аналитическая база данных EMIS Global <https://www.emis.com/php/companies/overview/index>
30. Реферативная база данных по математике MathSciNET <https://mathscinet.ams.org/mathscinet/>
31. Oxford Scholarship Online <https://oxford.universitypressscholarship.com/>
32. Коллекция научных журналов Oxford University Press <https://academic.oup.com/journals/>
33. ProQuest: База данных Business Ebook Subscription на платформе Ebook Central <https://search.proquest.com/>
34. ProQuest Dissertations & Theses A&I <https://search.proquest.com/>
35. База данных RUSLANA компании Bureau van Dijk <https://ruslana.bvdep.com/>
36. Scopus <https://www.scopus.com>
37. Электронная коллекция книг издательства Springer: Springer eBooks <http://link.springer.com/>
38. Интерактивная финансовая информационная система компании Bloomberg
39. Система Thomson Reuters Eikon
40. Web of Science <http://apps.webofknowledge.com>

9. Методические указания для обучающихся по освоению дисциплины

Лекционные и практические занятия проводятся в соответствии с тематическим планом, изложенным в разделе 5. Лекции следует конспектировать. Работа над конспектом лекции предполагает последующую его доработку. Вопросы, не в полной мере раскрытые на лекции, и оставленные на самостоятельную доработку студентам, следует разрешать, привлекая рекомендованную преподавателем литературу. В процессе доработки конспекта лекции знания по дисциплине, как правило, углубляются, расширяются и закрепляются. При работе с рекомендованной литературой желательно вести записи. Информацию целесообразно конспектировать, систематизируя новые знания при помощи построения логических цепочек с причинно-следственной связью. Рекомендуется полностью прорабатывать материал лекции до проведения следующего занятия с тем, чтобы иметь возможность обсудить с преподавателем пройденный материал и задать дополнительные вопросы по теме.

На практических занятиях работа ведется с использованием методических указаний по решению задач, примерами решения, а также с кейсами, разработанными на основе фактического материала. Подготовка к практическим занятиям является обязательной и предполагает работу с учебной и методической литературой. Домашние задания по курсу практических занятий являются систематическими. Контроль осуществляется регулярно на каждом занятии.

Контрольные работы проводятся в форме письменного экспресс-опроса, открытых и закрытых тестовых заданий, решения ситуационных и практических

задач по темам учебной дисциплины.

Формой промежуточной аттестации по дисциплине является экзамен. Критерии текущего контроля успеваемости устанавливает преподаватель и сообщает студентам на первом занятии.

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

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

10.1. Современные профессиональные базы данных и информационные справочные системы

1. Информационно-правовая система «Гарант»
2. Информационно-правовая система «Консультант Плюс»
3. Система комплексного раскрытия информации «СКРИН» -

<http://www.skrin.ru/>

10.2. Сертифицированные программные и аппаратные средства защиты информации - не предусмотрены

10.3. Anaconda и Python

10.4. Azure ML

11. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

Материально-техническая база Краснодарского филиала Финансового университета соответствует действующим противопожарным правилам и нормам, обеспечивает проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Образовательный процесс обеспечивается специальными помещениями, которые представляют собой аудитории для проведения занятий лекционного типа, занятий семинарского типа, выполнения курсовых работ, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, помещения для самостоятельной работы студентов и помещения для хранения и профилактического обслуживания учебного оборудования. Специальные

помещения укомплектованы специализированной мебелью и техническими средствами обучения, обеспечивающими представление учебной информации большой аудитории, демонстрационным оборудованием.

Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой, обеспечивающей доступ к сети Интернет и электронной информационно-образовательной среде Краснодарского филиала Финансового университета.