

Рабочая программа дисциплины

Интеллектуальный анализ больших данных

<i>Направление подготовки</i>	Информационные системы и технологии
<i>Код</i>	09.03.02
<i>Направленность (профиль)</i>	Информационные технологии и большие данные
<i>Квалификация (степень) выпускника</i>	бакалавр

1. Перечень кодов компетенций, формируемых дисциплиной в процессе освоения образовательной программы

Группа компетенций	Категория компетенций	Код
Профессиональные	-	ПК-1
Профессиональные	-	ПК-2

2. Компетенции и индикаторы их достижения

Код компетенции	Формулировка компетенции	Индикаторы достижения компетенции
ПК-1	ПК-1 Способен использовать математический аппарат, методологию программирования и современные компьютерные технологии для решения практических задач получения, хранения, обработки и передачи информации.	ПК-1.1. Подбирает парадигму программирования под решение конкретной прикладной задачи; ПК-1.3. Комбинирует известные алгоритмы решения задач. ПК-1.5. Применяет знания теоретической информатики, фундаментальной и прикладной математики для анализа и синтеза информационных систем и процессов.
ПК-2	ПК-2. Способен использовать математический аппарат и современные компьютерные средства для выполнения научно-исследовательских работ по закрепленной тематике	ПК-2.1. Демонстрирует способность и готовность к построению и исследованию математических моделей различных физических, биологических, экономических и социальных систем, а также применению идей, принципов и методов математического моделирования при решении прикладных задач. ПК-2.2. Использует естественнонаучные и математические знания для ориентирования в современном информационном пространстве ПК-2.3. Способен использовать естественнонаучные и математические знания для ориентирования в современном информационном пространстве. ПК-2.4. Собирает и анализирует информацию по решаемой задаче, составляет ее математическое описание, обеспечивает накопление, анализ и систематизацию собранных данных с использованием современных достижений науки и информационных систем; ПК-2.5. Выявляет и формулирует актуальные научные проблемы; обосновывать актуальность,

		теоретическую и практическую значимость темы научного исследования, разрабатывать план и программу проведения научного исследования; ПК-2.6. Проводит научно-исследовательские работы в области математики и компьютерных наук.
--	--	---

3. Описание планируемых результатов обучения по дисциплине

3.1. Описание планируемых результатов обучения по дисциплине

Планируемые результаты обучения по дисциплине представлены дескрипторами (знания, умения, навыки).

Дескрипторы по дисциплине	Знать	Уметь	Владеть
Код компетенции	ПК-1		
	- математический аппарат, методологию программирования и современные компьютерные технологии для решения практических задач получения, хранения, обработки и передачи информации; - основные идеи, понятия и методы, определяющие стиль написания, отладки и сопровождения программ; - характеристики основных парадигм программирования;	- использовать математический аппарат, методологию программирования и современные компьютерные технологии для решения практических задач получения, хранения, обработки и передачи информации; - применять современные компьютерные технологии для решения практических задач; - делать обоснованный выбор инструментария для решения прикладных задач;	- навыками использования математического аппарата, методологии программирования и современных компьютерных технологий для решения практических задач получения, хранения, обработки и передачи информации; - математическим аппаратом для построения вычислительных моделей практических задач; - навыками использования стандартных алгоритмических моделей для решения задач хранения и обработки информации
Код компетенции	ПК-2		

	Основы операционных систем: Знать архитектуру и основные компоненты операционных систем (например, Windows, Linux). Знать принципы работы файловых систем и управления процессами. Сетевые протоколы: Знать основные сетевые протоколы (TCP/IP, HTTP, HTTPS, FTP) и их функции. Знать принципы работы сетевых устройств (маршрутизаторов, коммутаторов, брандмауэров). Инструменты конфигурирования: Знать инструменты и утилиты для конфигурирования операционных систем и сетевых устройств (например, командная строка, PowerShell, SSH). Безопасность: Знать основные принципы и методы обеспечения безопасности операционных систем и сетевых устройств.	Конфигурировать операционные системы: Уметь устанавливать и настраивать операционные системы для веб-разработки. Уметь управлять пользователями и правами доступа в операционных системах. Настраивать сетевые устройства: Уметь конфигурировать маршрутизаторы и коммутаторы для обеспечения сетевой связности. Уметь настраивать брандмауэры для защиты веб-приложений. Использовать инструменты командной строки: Уметь использовать командную строку для выполнения задач администрирования и конфигурирования. Уметь применять скрипты для автоматизации процессов конфигурирования. Обеспечивать безопасность систем: Уметь применять методы шифрования и аутентификации для защиты данных. Уметь проводить аудит безопасности и выявлять уязвимости в системах.	Практическими навыками конфигурирования: Владеть навыками установки и настройки операционных систем в средах веб-разработки. Владеть навыками конфигурирования сетевых устройств для оптимизации работы веб-приложений. Работой с документацией: Владеть умением читать и интерпретировать техническую документацию по операционным системам и сетевым устройствам. Решением проблем: Владеть навыками диагностики и устранения неполадок в операционных системах и сетевых устройствах. Работой в команде: Владеть навыками взаимодействия с другими специалистами (разработчиками, системными администраторами и) для
--	--	---	--

			эффективного конфигурирования и поддержки веб-приложений.
--	--	--	---

4. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Интеллектуальный анализ больших данных» относится к части, формируемой участниками образовательных отношений учебного плана ОПОП.

Данная дисциплина взаимосвязана с другими дисциплинами, такими как: «Математический анализ», «Линейная алгебра», «Алгоритмизация и методы программирования», «Математическая логика и дискретная математика», «Численные методы», «Системы искусственного интеллекта», «Базы данных», «Машинное обучение», «Технологии хранения и обработки больших данных».

В рамках освоения программы бакалавриата выпускники готовятся к решению задач профессиональной деятельности следующих типов:

Профиль (направленность) программы установлена путем ее ориентации на сферу профессиональной деятельности выпускников: Информационные технологии и большие данные

5. Объем дисциплины

<i>Виды учебной работы</i>	<i>Формы обучения</i>
	<i>Очно-заочное</i>
Общая трудоемкость: зачетные единицы/часы	6/216
Контактная работа:	
Занятия лекционного типа	12
Занятия семинарского типа	12
Промежуточная аттестация: экзамен, зачет	0,15
Самостоятельная работа (СРС)	191,85

6. Содержание дисциплины (модуля), структурированное по темам / разделам с указанием отведенного на них количества академических часов и видов учебных занятий

6.1. Распределение часов по разделам/темам и видам работы

6.1.1. Очно-заочная форма обучения

№ п/п	Раздел/тема	Виды учебной работы (в часах)						
		Контактная работа						Самосто ятельная работа
		Занятия лекционного типа		Занятия семинарского типа				
		Лекции	Иные учебные занятия	Практи ческие занятия	Семи нары	Лабора торные работы	Иные	
1.	Основы интеллектуального анализа данных	2			2			38

2.	Предобработка и разведочный анализ больших данных	2			2			38
3.	Методы классификации, регрессии и кластеризации для больших данных	2			2			38
4.	Продвинутые методы анализа: ассоциативные правила, уменьшение размерности, глубокое обучение	3			3			38
5.	Применение интеллектуального анализа в реальных задачах и этика	3			3			39,85
	Промежуточная аттестация	0,15						
	Итого	12			12			191,85

6.2 Программа дисциплины, структурированная по темам / разделам

6.2.1 Содержание лекционного курса

№ п/п	Наименование раздела дисциплины	Содержание раздела дисциплины
1.	Основы интеллектуального анализа данных	понятие интеллектуального анализа данных, его связь с Big Data и машинным обучением; 5V-концепция Big Data: Volume, Velocity, Variety, Veracity, Value; задачи интеллектуального анализа: классификация, регрессия, кластеризация, ассоциативные правила, выявление аномалий; жизненный цикл проекта Data Mining (CRISP-DM): понимание бизнеса, понимание данных, подготовка данных, моделирование, оценка, внедрение; источники данных: IoT, социальные сети, логи, транзакции, тексты, изображения; обзор инструментов: Python (pandas, scikit-learn), R, Weka, RapidMiner, Spark MLlib; кейс: анализ оттока клиентов (churn analysis) в телекоммуникациях.
2.	Предобработка и разведочный анализ больших данных	этапы предобработки: очистка, заполнение пропусков, удаление дубликатов, фильтрация выбросов; обработка категориальных признаков: one-hot encoding, label encoding, target encoding;

		масштабирование и нормализация: Min-Max, стандартизация (Z-score), Robust Scaler; методы отбора признаков: фильтрационные (корреляция, χ^2), обёрточные (RFE), встроенные (LASSO); разведочный анализ данных (EDA): визуализация распределений, корреляционных матриц, boxplots, pairplots; работа с несбалансированными данными: SMOTE, undersampling, class weighting; практическое задание: предобработка набора данных о кредитных заявках.
3.	Методы классификации, регрессии и кластеризации для больших данных	постановка задач классификации и регрессии в контексте Big Data; линейные модели: логистическая регрессия, линейная регрессия с регуляризацией (Ridge, LASSO); нелинейные методы: деревья решений, случайный лес, градиентный бустинг (XGBoost, LightGBM, CatBoost); методы опорных векторов (SVM) для классификации; алгоритмы кластеризации: k-means, иерархическая кластеризация, DBSCAN, Gaussian Mixture Models; метрики качества: accuracy, precision, recall, F1-score, ROC-AUC (классификация); MSE, RMSE, MAE, R^2 (регрессия); silhouette score, inertia (кластеризация); кросс-валидация для оценки моделей на больших выборках; практическое задание: прогнозирование вероятности дефолта заёмщика на большом наборе данных.
4.	Продвинутые методы анализа: ассоциативные правила, уменьшение размерности, глубокое обучение	ассоциативные правила: алгоритм Apriori, FP-Growth, метрики поддержки, доверия, lift; уменьшение размерности: PCA, t-SNE, UMAP, автоэнкодеры; нейронные сети для анализа больших данных: полносвязные сети, CNN, RNN/LSTM; глубокое обучение с подкреплением (Reinforcement Learning) для оптимизации решений; ансамблевые методы: стекинг, блендинг, голосование; AutoML для автоматизации выбора моделей и настройки гиперпараметров (H2O, TPOT); практическое задание: кластеризация пользователей мобильного приложения по поведению и построение ассоциативных правил покупок.

5.	Применение интеллектуального анализа в реальных задачах и этика	рекомендательные системы: коллаборативная фильтрация, матричная факторизация, контент-ориентированные подходы; анализ текстов (NLP): sentiment analysis, topic modeling (LDA), векторные представления слов (Word2Vec, GloVe); компьютерное зрение: классификация изображений, обнаружение объектов на больших наборах данных; этические и юридические аспекты: GDPR, ФЗ-152, объяснимость моделей (XAI); кейсы: прогнозирование спроса на основе исторических данных; выявление мошенничества в финансовых транзакциях; персонализация контента в онлайн-сервисах; современные тенденции: федеративное обучение, приватный анализ данных, edge AI; практическое задание: построение рекомендательной системы на основе истории покупок и анализ тональности отзывов.
----	---	---

6.2.2 Содержание практических занятий

№ п/п	Наименование раздела дисциплины	Содержание практических занятий
1.	Основы интеллектуального анализа данных	понятие интеллектуального анализа данных, его связь с Big Data и машинным обучением; 5V-концепция Big Data: Volume, Velocity, Variety, Veracity, Value; задачи интеллектуального анализа: классификация, регрессия, кластеризация, ассоциативные правила, выявление аномалий; жизненный цикл проекта Data Mining (CRISP-DM): понимание бизнеса, понимание данных, подготовка данных, моделирование, оценка, внедрение; источники данных: IoT, социальные сети, логи, транзакции, тексты, изображения; обзор инструментов: Python (pandas, scikit-learn), R, Weka, RapidMiner, Spark MLlib; кейс: анализ оттока клиентов (churn analysis) в телекоммуникациях.
2.	Предобработка и разведочный анализ больших данных	этапы предобработки: очистка, заполнение пропусков, удаление дубликатов, фильтрация выбросов; обработка категориальных признаков: one-hot encoding, label encoding, target encoding; масштабирование и нормализация: Min-Max, стандартизация (Z-score), Robust Scaler;

		методы отбора признаков: фильтрационные (корреляция, χ^2), обёрточные (RFE), встроенные (LASSO); разведочный анализ данных (EDA): визуализация распределений, корреляционных матриц, boxplots, pairplots; работа с несбалансированными данными: SMOTE, undersampling, class weighting; практическое задание: предобработка набора данных о кредитных заявках.
3.	Методы классификации, регрессии и кластеризации для больших данных	постановка задач классификации и регрессии в контексте Big Data; линейные модели: логистическая регрессия, линейная регрессия с регуляризацией (Ridge, LASSO); нелинейные методы: деревья решений, случайный лес, градиентный бустинг (XGBoost, LightGBM, CatBoost); методы опорных векторов (SVM) для классификации; алгоритмы кластеризации: k-means, иерархическая кластеризация, DBSCAN, Gaussian Mixture Models; метрики качества: accuracy, precision, recall, F1-score, ROC-AUC (классификация); MSE, RMSE, MAE, R^2 (регрессия); silhouette score, inertia (кластеризация); кросс-валидация для оценки моделей на больших выборках; практическое задание: прогнозирование вероятности дефолта заёмщика на большом наборе данных.
4.	Продвинутые методы анализа: ассоциативные правила, уменьшение размерности, глубокое обучение	ассоциативные правила: алгоритм Apriori, FP-Growth, метрики поддержки, доверия, lift; уменьшение размерности: PCA, t-SNE, UMAP, автоэнкодеры; нейронные сети для анализа больших данных: полносвязные сети, CNN, RNN/LSTM; глубокое обучение с подкреплением (Reinforcement Learning) для оптимизации решений; ансамблевые методы: стекинг, блендинг, голосование; AutoML для автоматизации выбора моделей и настройки гиперпараметров (H2O, TPOT); практическое задание: кластеризация пользователей мобильного приложения по поведению и построение ассоциативных правил покупок.

5.	Применение интеллектуального анализа в реальных задачах и этика	рекомендательные системы: коллаборативная фильтрация, матричная факторизация, контент-ориентированные подходы; анализ текстов (NLP): sentiment analysis, topic modeling (LDA), векторные представления слов (Word2Vec, GloVe); компьютерное зрение: классификация изображений, обнаружение объектов на больших наборах данных; этические и юридические аспекты: GDPR, ФЗ-152, объяснимость моделей (XAI); кейсы: прогнозирование спроса на основе исторических данных; выявление мошенничества в финансовых транзакциях; персонализация контента в онлайн-сервисах; современные тенденции: федеративное обучение, приватный анализ данных, edge AI; практическое задание: построение рекомендательной системы на основе истории покупок и анализ тональности отзывов.
----	---	---

6.2.3 Содержание самостоятельной работы

№ п/п	Наименование раздела дисциплины	Содержание самостоятельной работы
1.	Основы интеллектуального анализа данных	понятие интеллектуального анализа данных, его связь с Big Data и машинным обучением; 5V-концепция Big Data: Volume, Velocity, Variety, Veracity, Value; задачи интеллектуального анализа: классификация, регрессия, кластеризация, ассоциативные правила, выявление аномалий; жизненный цикл проекта Data Mining (CRISP-DM): понимание бизнеса, понимание данных, подготовка данных, моделирование, оценка, внедрение; источники данных: IoT, социальные сети, логи, транзакции, тексты, изображения; обзор инструментов: Python (pandas, scikit-learn), R, Weka, RapidMiner, Spark MLlib; кейс: анализ оттока клиентов (churn analysis) в телекоммуникациях.
2.	Предобработка и разведочный анализ больших данных	этапы предобработки: очистка, заполнение пропусков, удаление дубликатов, фильтрация выбросов; обработка категориальных признаков: one-hot encoding, label encoding, target encoding; масштабирование и нормализация: Min-Max, стандартизация (Z-score), Robust Scaler;

		методы отбора признаков: фильтрационные (корреляция, χ^2), обёрточные (RFE), встроенные (LASSO); разведочный анализ данных (EDA): визуализация распределений, корреляционных матриц, boxplots, pairplots; работа с несбалансированными данными: SMOTE, undersampling, class weighting; практическое задание: предобработка набора данных о кредитных заявках.
3.	Методы классификации, регрессии и кластеризации для больших данных	постановка задач классификации и регрессии в контексте Big Data; линейные модели: логистическая регрессия, линейная регрессия с регуляризацией (Ridge, LASSO); нелинейные методы: деревья решений, случайный лес, градиентный бустинг (XGBoost, LightGBM, CatBoost); методы опорных векторов (SVM) для классификации; алгоритмы кластеризации: k-means, иерархическая кластеризация, DBSCAN, Gaussian Mixture Models; метрики качества: accuracy, precision, recall, F1-score, ROC-AUC (классификация); MSE, RMSE, MAE, R^2 (регрессия); silhouette score, inertia (кластеризация); кросс-валидация для оценки моделей на больших выборках; практическое задание: прогнозирование вероятности дефолта заёмщика на большом наборе данных.
4.	Продвинутые методы анализа: ассоциативные правила, уменьшение размерности, глубокое обучение	ассоциативные правила: алгоритм Apriori, FP-Growth, метрики поддержки, доверия, lift; уменьшение размерности: PCA, t-SNE, UMAP, автоэнкодеры; нейронные сети для анализа больших данных: полносвязные сети, CNN, RNN/LSTM; глубокое обучение с подкреплением (Reinforcement Learning) для оптимизации решений; ансамблевые методы: стекинг, блендинг, голосование; AutoML для автоматизации выбора моделей и настройки гиперпараметров (H2O, TPOT); практическое задание: кластеризация пользователей мобильного приложения по поведению и построение ассоциативных правил покупок.

5.	Применение интеллектуального анализа в реальных задачах и этика	рекомендательные системы: коллаборативная фильтрация, матричная факторизация, контент-ориентированные подходы; анализ текстов (NLP): sentiment analysis, topic modeling (LDA), векторные представления слов (Word2Vec, GloVe); компьютерное зрение: классификация изображений, обнаружение объектов на больших наборах данных; этические и юридические аспекты: GDPR, ФЗ-152, объяснимость моделей (XAI); кейсы: прогнозирование спроса на основе исторических данных; выявление мошенничества в финансовых транзакциях; персонализация контента в онлайн-сервисах; современные тенденции: федеративное обучение, приватный анализ данных, edge AI; практическое задание: построение рекомендательной системы на основе истории покупок и анализ тональности отзывов.
----	---	--

7. Текущий контроль по дисциплине (модулю) в рамках учебных занятий

В рамках текущего контроля преподаватель самостоятельно может проводить следующие мероприятия:

№ п/п	Контролируемые разделы (темы)	Наименование оценочного средства
1.	Основы интеллектуального анализа данных	Опрос, информационный проект.
2.	Предобработка и разведочный анализ больших данных	Опрос, информационный проект, тестирование.
3.	Методы классификации, регрессии и кластеризации для больших данных	Опрос, информационный проект.
4.	Продвинутое методы анализа: ассоциативные правила, уменьшение размерности, глубокое обучение	Опрос, информационный проект.
5.	Применение интеллектуального анализа в реальных задачах и этика	Опрос, информационный проект, тестирование.

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

8.1. Основная литература:

1. Методы интеллектуального анализа данных. Задача классификации : учебное пособие / О. С. Чудинова, Р. М. Безбородникова, Е. Н. Корнейченко [и др.]. —

- Оренбург : Оренбургский государственный университет, ЭБС АСВ, 2024. — 167 с. — ISBN 978-5-7410-3284-8. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/153049.html>
2. Целых, А. Н. Принятие решений на основе методов машинного обучения : учебное пособие по курсам «Модели и методы инженерии знаний», «Методы анализа больших данных» / А. Н. Целых, Н. В. Драгныш, Э. М. Котов. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2022. — 113 с. — ISBN 978-5-9275-4246-8. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/131458.html>
 3. Хохлов, С. С. Основные методы анализа экспериментальных данных : учебное пособие / С. С. Хохлов, А. Н. Дмитриева. — Москва : Национальный исследовательский ядерный университет «МИФИ», 2025. — 126 с. — ISBN 978-5-7262-3149-5. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/156787.htm>

8.2. Дополнительная литература:

1. Целых, А. Н. Применение временных рядов для анализа больших данных : учебное пособие по курсу «Математические методы анализа больших данных» / А. Н. Целых, В. С. Васильев, Э. М. Котов. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2021. — 84 с. — ISBN 978-5-9275-3983-3. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/121929.html>
2. Целых, А. Н. Современные методы прикладной информатики в задачах анализа данных : учебное пособие по курсу «Методы интеллектуального анализа данных» / А. Н. Целых, А. А. Целых, Э. М. Котов. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2021. — 130 с. — ISBN 978-5-9275-3783-9. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/117165.html>

8.3. Периодические издания

1. Вестник Московского государственного технического университета имени Н.Э. Баумана. Серия Естественные науки. ISSN 1812-3368. <https://www.iprbookshop.ru/23124.html>
2. Открытые Системы. СУБД. ISSN 1028-7493. <https://www.iprbookshop.ru/76383.html>
3. Информационные технологии моделирования и управления. ISSN 1813-9744. <https://www.iprbookshop.ru/43350.html>

9. Перечень ресурсов информационно-телекоммуникационной сети "Интернет" (далее - сеть "Интернет"), необходимых для освоения дисциплины (модуля)

1. Федеральный портал «Российское образование» <http://www.edu.ru>
2. Электронно-библиотечная система «Научная электронная библиотека eLIBRARY.RU» <https://www.elibrary.ru>
3. Электронно-библиотечная система IPR BOOKS <https://www.iprbookshop.ru>
4. Электронная библиотечная система <https://biblioclub.ru>

10. Методические указания для обучающихся по освоению дисциплины (модуля)

Успешное освоение данного курса базируется на рациональном сочетании нескольких видов учебной деятельности – лекций, семинарских занятий, самостоятельной работы. При этом самостоятельную работу следует рассматривать одним из главных

звеньев полноценного высшего образования, на которую отводится значительная часть учебного времени.

Самостоятельная работа студентов складывается из следующих составляющих:

1. Работа с основной и дополнительной литературой, с материалами интернета и конспектами лекций;
2. Внеаудиторная подготовка к контрольным работам, выполнение докладов, рефератов и курсовых работ;
3. Выполнение самостоятельных практических работ;
4. Подготовка к экзаменам (зачетам) непосредственно перед ними.

Для правильной организации работы необходимо учитывать порядок изучения разделов курса, находящихся в строгой логической последовательности. Поэтому хорошее усвоение одной части дисциплины является предпосылкой для успешного перехода к следующей. Задания, проблемные вопросы, предложенные для изучения дисциплины, в том числе и для самостоятельного выполнения, носят междисциплинарный характер и базируются, прежде всего, на причинно-следственных связях между компонентами окружающего нас мира. Важным составляющим в изучении данного курса является решение ситуационных задач и работа над проблемно-аналитическими заданиями, что предполагает знание соответствующей научной терминологии и т.д.

Для лучшего запоминания материала целесообразно использовать индивидуальные особенности и разные виды памяти: зрительную, слуховую, ассоциативную. Успешному запоминанию также способствует приведение ярких свидетельств и наглядных примеров. Учебный материал должен постоянно повторяться и закрепляться.

При выполнении докладов, творческих, информационных, исследовательских проектов особое внимание следует обращать на подбор источников информации и методику работы с ними.

Для успешной сдачи экзамена (зачета) рекомендуется соблюдать следующие правила:

1. Подготовка к экзамену (зачету) должна проводиться систематически, в течение всего семестра.
2. Интенсивная подготовка должна начаться не позднее, чем за месяц до экзамена.
3. Время непосредственно перед экзаменом (зачетом) лучше использовать таким образом, чтобы оставить последний день свободным для повторения курса в целом, для систематизации материала и доработки отдельных вопросов.

На экзамене высокую оценку получают студенты, использующие данные, полученные в процессе выполнения самостоятельных работ, а также использующие собственные выводы на основе изученного материала.

Учитывая значительный объем теоретического материала, студентам рекомендуется регулярное посещение и подробное конспектирование лекций.

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

1. Microsoft Windows Server;
2. Семейство ОС Microsoft Windows;
3. Libre Office свободно распространяемый офисный пакет с открытым исходным кодом;
4. Информационно-справочная система: Система КонсультантПлюс (КонсультантПлюс);
5. Информационно-правовое обеспечение Гарант: Электронный периодический справочник «Система ГАРАНТ» (Система ГАРАНТ);

Перечень используемого программного обеспечения указан в п.12 данной рабочей

программы дисциплины.

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

12.1. Учебная аудитория для проведения учебных занятий, предусмотренных образовательной программой, оснащенная оборудованием и техническими средствами обучения.

Специализированная мебель:

Комплект учебной мебели (стол, стул) по количеству обучающихся; комплект мебели для преподавателя; доска (маркерная).

Технические средства обучения:

Компьютер в сборе для преподавателя, проектор, экран, колонки

Перечень лицензионного программного обеспечения, в том числе отечественного производства:

Windows 10, КонсультантПлюс, Kaspersky Endpoint Security.

Перечень свободно распространяемого программного обеспечения:

Яндекс Браузер, LibreOffice, МТС Линк.

Подключение к сети «Интернет» и обеспечение доступа в электронную информационно-образовательную среду ММУ.

12.2. Помещение для самостоятельной работы обучающихся.

Специализированная мебель:

Комплект учебной мебели (стол, стул) по количеству обучающихся; комплект мебели для преподавателя; доска (маркерная).

Технические средства обучения:

Компьютер в сборе для преподавателя; компьютеры в сборе для обучающихся; колонки; проектор, экран.

Перечень лицензионного программного обеспечения, в том числе отечественного производства:

Windows Server 2016, Windows 10, Microsoft Office, КонсультантПлюс, Kaspersky Endpoint Security.

Перечень свободно распространяемого программного обеспечения:

Яндекс Браузер, LibreOffice, МТС Линк, Gimp, Paint.net, AnyLogic, Inkscape.

Помещение для самостоятельной работы обучающихся оснащено компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду ММУ.

13. Образовательные технологии, используемые при освоении дисциплины

Для освоения дисциплины используются как традиционные формы занятий – лекции (типы лекций – установочная, вводная, текущая, заключительная, обзорная; виды лекций – проблемная, визуальная, лекция конференция, лекция консультация); и семинарские (практические) занятия, так и активные и интерактивные формы занятий – деловые и ролевые игры, решение ситуационных задач и разбор конкретных ситуаций.

На учебных занятиях используются технические средства обучения мультимедийной аудитории: компьютер, монитор, колонки, настенный экран, проектор, микрофон, пакет программ Microsoft Office для демонстрации презентаций и медиафайлов, видеопроектор для демонстрации слайдов, видеосюжетов и др. Тестирование обучаемых может осуществляться с использованием компьютерного оборудования университета.

13.1. В освоении учебной дисциплины используются следующие традиционные

образовательные технологии:

- чтение проблемно-информационных лекций с использованием доски и видеоматериалов;
- семинарские занятия для обсуждения, дискуссий и обмена мнениями;
- контрольные опросы;
- консультации;
- самостоятельная работа студентов с учебной литературой и первоисточниками;
- подготовка и обсуждение рефератов (проектов), презентаций (научно-исследовательская работа);
- тестирование по основным темам дисциплины.

13.2. Активные и интерактивные методы и формы обучения

Из перечня видов: («мозговой штурм», анализ НПА, анализ проблемных ситуаций, анализ конкретных ситуаций, инциденты, имитация коллективной профессиональной деятельности, разыгрывание ролей, творческая работа, связанная с освоением дисциплины, ролевая игра, круглый стол, диспут, беседа, дискуссия, мини-конференция и др.) используются следующие:

- диспут;
- анализ проблемных, творческих заданий, ситуационных задач;
- ролевая игра;
- круглый стол;
- мини-конференция;
- дискуссия;
- беседа.

13.3. Особенности обучения инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ)

При организации обучения по дисциплине учитываются особенности организации взаимодействия с инвалидами и лицами с ограниченными возможностями здоровья (далее – инвалиды и лица с ОВЗ) с целью обеспечения их прав. При обучении учитываются особенности их психофизического развития, индивидуальные возможности и при необходимости обеспечивается коррекция нарушений развития и социальная адаптация указанных лиц.

Выбор методов обучения определяется содержанием обучения, уровнем методического и материально-технического обеспечения, особенностями восприятия учебной информации студентами с инвалидностью и студентами с ограниченными возможностями здоровья и т.д. В образовательном процессе используются социально-активные и рефлексивные методы обучения, технологии социокультурной реабилитации с целью оказания помощи в установлении полноценных межличностных отношений с другими студентами, создании комфортного психологического климата в студенческой группе.

При обучении лиц с ограниченными возможностями здоровья электронное обучение и дистанционные образовательные технологии предусматривают возможность приема-передачи информации в доступных для них формах.

Обучающиеся из числа лиц с ограниченными возможностями здоровья обеспечены печатными и электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья.

Автономная некоммерческая организация высшего образования
«МОСКОВСКИЙ МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ»

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ
ПО ДИСЦИПЛИНЕ**

Интеллектуальный анализ больших данных

<i>Направление подготовки</i>	Информационные системы и технологии
<i>Код</i>	09.03.02
<i>Направленность (профиль)</i>	Информационные технологии и большие данные
<i>Квалификация (степень) выпускника</i>	бакалавр

1. Перечень кодов компетенций, формируемых дисциплиной в процессе освоения образовательной программы

Группа компетенций	Категория компетенций	Код
Профессиональные	-	ПК-1
Профессиональные	-	ПК-2

2. Компетенции и индикаторы их достижения

Код компетенции	Формулировка компетенции	Индикаторы достижения компетенции
ПК-1	ПК-1 Способен использовать математический аппарат, методологию программирования и современные компьютерные технологии для решения практических задач получения, хранения, обработки и передачи информации.	ПК-1.1. Подбирает парадигму программирования под решение конкретной прикладной задачи; ПК-1.3. Комбинирует известные алгоритмы решения задач. ПК-1.5. Применяет знания теоретической информатики, фундаментальной и прикладной математики для анализа и синтеза информационных систем и процессов.
ПК-2	ПК-2. Способен использовать математический аппарат и современные компьютерные средства для выполнения научно-исследовательских работ по закрепленной тематике	ПК-2.1. Демонстрирует способность и готовность к построению и исследованию математических моделей различных физических, биологических, экономических и социальных систем, а также применению идей, принципов и методов математического моделирования при решении прикладных задач. ПК-2.2. Использует естественнонаучные и математические знания для ориентирования в современном информационном пространстве ПК-2.3. Способен использовать естественнонаучные и математические знания для ориентирования в современном информационном пространстве. ПК-2.4. Собирает и анализирует информацию по решаемой задаче, составляет ее математическое описание, обеспечивает накопление, анализ и систематизацию собранных данных с использованием современных достижений науки и информационных систем; ПК-2.5. Выявляет и формулирует

		актуальные научные проблемы; обосновывать актуальность, теоретическую и практическую значимость темы научного исследования, разрабатывать план и программу проведения научного исследования; ПК-2.6. Проводит научно-исследовательские работы в области математики и компьютерных наук.
--	--	--

3. Описание планируемых результатов обучения по дисциплине

3.1. Описание планируемых результатов обучения по дисциплине

Планируемые результаты обучения по дисциплине представлены дескрипторами (знания, умения, навыки).

Дескрипторы по дисциплине	Знать	Уметь	Владеть
Код компетенции	ПК-1		
	- математический аппарат, методологию программирования и современные компьютерные технологии для решения практических задач получения, хранения, обработки и передачи информации; - основные идеи, понятия и методы, определяющие стиль написания, отладки и сопровождения программ; - характеристики основных парадигм программирования;	- использовать математический аппарат, методологию программирования и современные компьютерные технологии для решения практических задач получения, хранения, обработки и передачи информации; - применять современные компьютерные технологии для решения практических задач; - делать обоснованный выбор инструментария для решения прикладных задач;	- навыками использования математического аппарата, методологии программирования и современных компьютерных технологий для решения практических задач получения, хранения, обработки и передачи информации; - математическим аппаратом для построения вычислительных моделей практических задач; - навыками использования стандартных алгоритмических моделей для решения задач хранения и обработки информации
Код	ПК-2		

компетенции	
	Основы операционных систем: Знать архитектуру и основные компоненты операционных систем (например, Windows, Linux). Знать принципы работы файловых систем и управления процессами. Сетевые протоколы: Знать основные сетевые протоколы (TCP/IP, HTTP, HTTPS, FTP) и их функции. Знать принципы работы сетевых устройств (маршрутизаторов, коммутаторов, брандмауэров). Инструменты конфигурирования: Знать инструменты и утилиты для конфигурирования операционных систем и сетевых устройств (например, командная строка, PowerShell, SSH). Безопасность: Знать основные принципы и методы обеспечения безопасности операционных систем и сетевых устройств. Конфигурировать операционные системы: Уметь устанавливать и настраивать операционные системы для веб-разработки. Уметь управлять пользователями и правами доступа в операционных системах. Настраивать сетевые устройства: Уметь конфигурировать маршрутизаторы и коммутаторы для обеспечения сетевой связности. Уметь настраивать брандмауэры для защиты веб-приложений. Использовать инструменты командной строки: Уметь использовать командную строку для выполнения задач администрирования и конфигурирования. Уметь применять скрипты для автоматизации процессов конфигурирования. Обеспечивать безопасность систем: Уметь применять методы шифрования и аутентификации для защиты данных. Уметь проводить аудит безопасности и выявлять уязвимости в системах. Практическими навыками конфигурирования: Владеть навыками установки и настройки операционных систем в средах веб-разработки. Владеть навыками конфигурирования сетевых устройств для оптимизации работы веб-приложений. Работой с документацией: Владеть умением читать и интерпретировать техническую документацию по операционным системам и сетевым устройствам. Решением проблем: Владеть навыками диагностики и устранения неполадок в операционных системах и сетевых устройствах. Работой в команде: Владеть навыками взаимодействия с другими специалистами (разработчиками, системными

			администраторам и) для эффективного конфигурирования и поддержки веб-приложений.
--	--	--	--

3.2. Критерии оценки знаний студентов

Шкала оценивания	Индикаторы достижения	Показатели оценивания результатов обучения
ОТЛИЧНО/Зачтено	Знает:	- студент глубоко и всесторонне усвоил материал, уверенно, логично, последовательно и грамотно его излагает, опираясь на знания основной и дополнительной литературы, - на основе системных научных знаний делает квалифицированные выводы и обобщения, свободно оперирует категориями и понятиями.
	Умеет:	- студент умеет самостоятельно и правильно решать учебно-профессиональные задачи или задания, уверенно, логично, последовательно и аргументировано излагать свое решение, используя научные понятия, ссылаясь на нормативную базу.
	Владеет:	- студент владеет рациональными методами (с использованием рациональных методик) решения сложных профессиональных задач, представленных деловыми играми, кейсами и т.д.; При решении продемонстрировал навыки - выделения главного, - связкой теоретических положений с требованиями руководящих документов, - изложения мыслей в логической последовательности, - самостоятельного анализа факты, событий, явлений, процессов в их взаимосвязи и диалектическом развитии.
ХОРОШО/Зачтено	Знает:	- студент твердо усвоил материал, достаточно грамотно его излагает, опираясь на знания основной и дополнительной литературы, - затрудняется в формулировании квалифицированных выводов и обобщений, оперирует категориями и понятиями, но не всегда правильно их верифицирует.
	Умеет:	- студент умеет самостоятельно и в основном правильно решать учебно-профессиональные задачи или задания, уверенно, логично, последовательно и аргументировано излагать свое решение, не в полной мере используя научные понятия и ссылки на нормативную базу.
	Владеет:	- студент в целом владеет рациональными методами решения сложных профессиональных задач, представленных деловыми играми, кейсами и т.д.; При решении смог продемонстрировать достаточность, но не глубинность навыков

		- выделения главного, - изложения мыслей в логической последовательности. - связки теоретических положений с требованиями руководящих документов, - самостоятельного анализа факты, событий, явлений, процессов в их взаимосвязи и диалектическом развитии.
УДОВЛЕТВОРИТЕЛЬНО/Зачтено	Знает:	- студент ориентируется в материале, однако затрудняется в его изложении; - показывает недостаточность знаний основной и дополнительной литературы; - слабо аргументирует научные положения; - практически не способен сформулировать выводы и обобщения; - частично владеет системой понятий.
	Умеет:	- студент в основном умеет решить учебно-профессиональную задачу или задание, но допускает ошибки, слабо аргументирует свое решение, недостаточно использует научные понятия и руководящие документы.
	Владеет:	- студент владеет некоторыми рациональными методами решения сложных профессиональных задач, представленных деловыми играми, кейсами и т.д.; При решении продемонстрировал недостаточность навыков - выделения главного, - изложения мыслей в логической последовательности. - связки теоретических положений с требованиями руководящих документов, - самостоятельного анализа факты, событий, явлений, процессов в их взаимосвязи и диалектическом развитии.
НЕУДОВЛЕТВОРИТЕЛЬНО/Не зачтено	Знает:	- студент не усвоил значительной части материала; - не может аргументировать научные положения; - не формулирует квалифицированных выводов и обобщений; - не владеет системой понятий.
	Умеет:	студент не показал умение решать учебно-профессиональную задачу или задание.
	Владеет:	не выполнены требования, предъявляемые к навыкам, оцениваемым “удовлетворительно”.

При ответе на вопросы в рамках прохождения промежуточной аттестации (зачет/зачет с оценкой/экзамен) допускается вольная формулировка ответа, по смыслу раскрывающая содержание ответа, указанного в фонде оценочных средств, в качестве верного ответа.

4. Типовые контрольные задания (закрытого, открытого и иного типа) для проведения промежуточной аттестации, необходимые для оценки достижения компетенции, соотнесенной с результатами обучения по дисциплине

ТЕСТИРОВАНИЕ

9 СЕМЕСТР

ПК-1

1. Что такое Data Mining?
Ответ: процесс обнаружения полезных закономерностей или знаний из больших объёмов данных.
2. Какая характеристика НЕ входит в 5V-концепцию Big Data?
Ответ: Variability.
3. Какой этап НЕ входит в жизненный цикл CRISP-DM?
Ответ: развёртывание веб-приложения.
4. Какой инструмент НЕ используется для интеллектуального анализа данных?
Ответ: Apache Kafka (это брокер сообщений, а не инструмент анализа).
5. Что такое churn analysis?
Ответ: анализ оттока клиентов, прогнозирование вероятности ухода пользователя.
6. Какой метод используется для заполнения пропусков в числовых данных?
Ответ: медиана, среднее значение или KNN-импутация.
7. Что такое one-hot encoding?
Ответ: преобразование категориальных переменных в бинарные векторы.
8. Для чего применяется стандартизация (Z-score) данных?
Ответ: для приведения признаков к единому масштабу со средним 0 и дисперсией 1.
9. Какой метод отбора признаков основан на рекурсивном исключении признаков?
Ответ: RFE (Recursive Feature Elimination).
10. Что показывает boxplot в EDA?
Ответ: распределение данных, медиану, квартили и выбросы.
11. Какая метрика НЕ используется для оценки регрессионных моделей?
Ответ: ассигасу.
12. Что такое ROC-AUC?
Ответ: площадь под ROC-кривой, мера качества классификатора.
13. Какой алгоритм НЕ относится к методам градиентного бустинга?
Ответ: k-NN.
14. Для чего используется регуляризация в линейных моделях?
Ответ: для предотвращения переобучения путём штрафа за большие веса.
15. Какой алгоритм кластеризации НЕ требует указания числа кластеров заранее?
Ответ: DBSCAN.

9 СЕМЕСТР

ПК-2

1. Что измеряет silhouette score?
Ответ: качество кластеризации — насколько объекты похожи внутри кластера и отличаются от других кластеров.
2. Какая метрика используется для оценки качества регрессионной модели?
Ответ: RMSE (Root Mean Squared Error).
3. Для чего используется PCA?
Ответ: для уменьшения размерности данных путём выделения главных компонент.
4. Какой алгоритм используется для поиска ассоциативных правил?
Ответ: Apriori.
5. Что такое lift в ассоциативных правилах?
Ответ: отношение наблюдаемой поддержки к ожидаемой, если бы товары были независимы.
6. Какой тип нейронной сети подходит для анализа изображений?
Ответ: свёрточная нейронная сеть (CNN).
7. Что такое AutoML?
Ответ: автоматизация процессов машинного обучения: выбор модели, настройка гиперпараметров.
8. Какой фреймворк НЕ используется для глубокого обучения?
Ответ: scikit-learn (это библиотека классического ML).
9. Что такое XAI (Explainable AI)?
Ответ: методы объяснения решений моделей ML для повышения прозрачности и доверия.
10. Какой нормативный акт регулирует обработку персональных данных в РФ?
Ответ: ФЗ-152 «О персональных данных».
11. Что такое коллаборативная фильтрация?
Ответ: метод рекомендательных систем, основанный на схожести пользователей или товаров.
12. Какой алгоритм подходит для тематического моделирования текстов?
Ответ: LDA (Latent Dirichlet Allocation).
13. Что такое reinforcement learning?
Ответ: обучение с подкреплением — агент учится принимать решения через взаимодействие со средой и получение наград.
14. Какой инструмент поддерживает AutoML в Python?
Ответ: H2O.ai или TPOT.
15. Что такое матричная факторизация в рекомендательных системах?
Ответ: разложение матрицы пользователь-товар на скрытые факторы для прогнозирования оценок.