

**Автономная некоммерческая организация высшего образования
«МОСКОВСКИЙ МЕЖДУНАРОДНЫЙ УНИВЕРСИТЕТ»**

Рабочая программа дисциплины

Современные методы анализа данных и машинное обучение

Направление подготовки	Бизнес-информатика
<i>Код</i>	38.03.05
<i>Направление (профиль)</i>	Информационные системы и технологии в бизнесе
Квалификация выпускника	бакалавр

Москва
2025

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

Группа компетенций	Категория компетенций	Код
Общепрофессиональные	-	ОПК-1
Общепрофессиональные	-	ОПК-4

2. Компетенции и индикаторы их достижения

Код компетенции	Формулировка компетенции	Индикаторы достижения компетенции
ОПК-1	Способен проводить моделирование, анализ и совершенствование бизнес-процессов и информационно-технологической инфраструктуры предприятия в интересах достижения его стратегических целей с использованием современных методов и программного инструментария;	ОПК-1.1. Знает: цели и задачи анализа, моделирования и совершенствования бизнес-процессов предприятия в интересах достижения его стратегических целей с использованием современных методов и программного инструментария ОПК-1.2. Умеет: проводить обследование, моделирование, анализ бизнес-процессов и ИТИ предприятия в интересах достижения его стратегических целей ОПК-1.3. Владеет: методами моделирования при решении задачи обоснования проекта по совершенствованию бизнес-процессов и ИТИ; методами совершенствования бизнес-процессов и ИТИ предприятия на основе использования современного программного инструментария для достижения его стратегических целей
ОПК-4	Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений;	ОПК-4.1. Знает методы сбора, обработки и анализа информации, в том числе в глобальных сетях, включая программные средства, методы представления информации, а также принципы работы информационных технологий ОПК-4.2. Умеет использовать математические и статистические методы анализа данных, в том числе с использованием компьютерных технологий, для информационноаналитической поддержки принятия управленческих решений ОПК-4.3. Владеет методами и программными средствами поддержки принятия управленческих решений

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	ОПК-1		
	- Методологии моделирования бизнес-процессов (BPMN, IDEF, UML) и принципы их применения для анализа организационной деятельности - Современные подходы к оптимизации ИТ-инфраструктуры предприятия, включая методы цифровой трансформации и автоматизации процессов - Программные средства бизнес-моделирования (ARIS, Bizagi, BPMS) и анализа ИТ-систем (ERP, CRM)	- Проводить диагностику бизнес-процессов, выявлять узкие места и разрабатывать предложения по их совершенствованию - Разрабатывать модели "как есть" и "как должно быть" с использованием специализированного ПО для визуализации процессов - Оценивать эффективность предлагаемых изменений с точки зрения достижения стратегических целей организации	- Навыками работы с инструментами бизнес-моделирования (нотации BPMN, EPC) и CASE-средствами - Методами анализа и реинжиниринга бизнес-процессов с учетом требований цифровой экономики - Технологиями интеграции бизнес-процессов с корпоративными информационными системами
Код компетенции	ОПК-4		
	- Архитектуру современных информационных систем и принципы их функционирования, включая клиент-серверные технологии, облачные вычисления и распределенные базы данных - Методы и технологии обработки данных (ETL-процессы, OLAP-анализ, методы машинного обучения) для аналитической поддержки решений	- Применять специализированное ПО (BI-системы, статистические пакеты, системы визуализации) для сбора и анализа бизнес-данных - Формулировать аналитические выводы на основе обработки структурированных и неструктурированных данных - Разрабатывать рекомендации для управленческих решений на основе	- Навыками работы с аналитическими платформами (Power BI, Tableau, Qlik) и языками запросов (SQL, DAX) - Методами статистического анализа и прогнозирования бизнес-показателей - Технологиями визуализации данных для представления аналитических отчетов руководству

	- Ключевые показатели эффективности бизнес-процессов и методики их расчета с использованием информационных систем	результатов анализа данных	
--	---	----------------------------	--

4. Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Современные методы анализа данных и машинное обучение» относится к обязательной части учебного плана ОПОП.

Данная дисциплина взаимосвязана с другими дисциплинами, такими как «Экономика», «Компьютерные сети», «Информационные системы и базы данных», «Экономика организации (предприятия)».

Изучение дисциплины позволит обучающимся реализовывать универсальные и общекультурные компетенции в профессиональной деятельности.

В рамках освоения программы бакалавриата выпускники готовятся к решению задач профессиональной деятельности следующих типов: научно-исследовательский, производственно-технологический, организационно-управленческий, проектный.

Профиль (направленность) программы установлена путем ее ориентации на сферу профессиональной деятельности выпускников: бизнес-информатика.

5. Объем дисциплины

<i>Виды учебной работы</i>	<i>Формы обучения</i>	
	<i>Очная</i>	<i>Очно-заочная</i>
Общая трудоемкость: зачетные единицы/часы	5/180	5/180
Контактная работа:		
Занятия лекционного типа	36	12
Занятия семинарского типа	36	24
Промежуточная аттестация: экзамен	27	9
Самостоятельная работа (СРС)	81	135

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. Очная форма обучения

№ п/ п	Раздел/тема	Виды учебной работы (в часах)							
		Контактная работа						Самост оятельн ая работа	
		Занятия лекционного типа		Занятия семинарского типа					
		Лекци и	Иные учебны е заняти я	Практ ически е заняти я	Семин ары	Лабор аторн ые работ ы	Иные		
1.	Введение в анализ данных и машинное	3		3					12

	обучение							
2.	Предобработка и исследование данных (EDA)	3		3				12
3.	Статистические методы анализа данных	3		3				12
4.	Основы машинного обучения	3		3				12
5.	Классификация и кластеризация	3		3				12
6.	Ансамбли моделей и бустинг	3		3				12
7.	Работа с текстовыми данными (NLP)	3		3				12
8.	Рекомендательные системы	3		3				12
9.	Временные ряды и прогнозирование	3		3				12
10.	Глубокое обучение (Deep Learning)	3		3				11
11.	Развертывание ML-моделей (MLOps)	3		3				11
12.	Кейсы применения ML в бизнесе	3		3				11
	Промежуточная аттестация	27						
	Итого	36		36				81

6.1.1. Очно-заочная форма обучения

№ п/ п	Раздел/тема	Виды учебной работы (в часах)						
		Контактная работа						Самост оательн ая работа
		Занятия лекционного типа		Занятия семинарского типа				
		Лекци и	Иные учебны е заняти я	Практ ически е заняти я	Семин ары	Лабор аторн ые работ ы	Иные	
1.	Введение в анализ данных и машинное обучение	1		2				12
2.	Предобработка и исследование данных (EDA)	1		2				12
3.	Статистические методы анализа данных	1		2				12
4.	Основы машинного обучения	1		2				12
5.	Классификация и	1		2				12

	кластеризация							
6.	Ансамбли моделей и бустинг	1		2				12
7.	Работа с текстовыми данными (NLP)	1		2				12
8.	Рекомендательные системы	1		2				12
9.	Временные ряды и прогнозирование	1		2				12
10.	Глубокое обучение (Deep Learning)	1		2				12
11.	Развертывание ML-моделей (MLOps)	1		2				12
12	Кейсы применения ML в бизнесе	1		2				3
	Промежуточная аттестация	9						
	Итого	12		24				135

6.2 Программа дисциплины, структурированная по темам / разделам

6.2.1 Содержание лекционного курса

№ п/п	Наименование темы (раздела) дисциплины	Содержание лекционного занятия
1.	Введение в анализ данных и машинное обучение	Основные понятия: данные, информация, знания. Жизненный цикл анализа данных (CRISP-DM, KDD). Обзор современных методов и инструментов (Python, R, BI-системы).
2.	Предобработка и исследование данных (EDA)	Работа с пропусками, выбросами, дубликатами. Нормализация, стандартизация, кодирование категориальных данных. Визуализация данных (Matplotlib, Seaborn, Tableau).
3.	Статистические методы анализа данных	Описательная статистика (среднее, дисперсия, корреляция). Проверка гипотез (t-тест, ANOVA, χ^2). Регрессионный анализ (линейная, логистическая регрессия).
4.	Основы машинного обучения	Обучение с учителем vs без учителя. Переобучение, валидация, кросс-валидация. Метрики качества (Accuracy, Precision, Recall, F1, ROC-AUC).
5.	Классификация и кластеризация	Методы классификации (k-NN, SVM, Decision Trees, Random Forest). Кластеризация (k-means, DBSCAN, иерархическая кластеризация). Применение в бизнесе (сегментация клиентов, анализ оттока).
6.	Ансамбли моделей и бустинг	Bagging (Random Forest) vs Boosting (XGBoost, LightGBM, CatBoost). Стекинг и блендинг моделей. Оптимизация гиперпараметров (GridSearch, RandomSearch).

7.	Работа с текстовыми данными (NLP)	Токенизация, стемминг, лемматизация. Векторизация текста (TF-IDF, Word2Vec, BERT). Анализ тональности, классификация текстов.
8.	Рекомендательные системы	Коллаборативная фильтрация (user-based, item-based). Контентная фильтрация. Гибридные системы (на примере e-commerce).
9.	Временные ряды и прогнозирование	Стационарность, автокорреляция. Методы прогнозирования (ARIMA, SARIMA, Prophet). Применение в бизнесе (прогноз продаж, спроса).
10.	Глубокое обучение (Deep Learning)	Введение в нейронные сети (MLP, CNN, RNN). Фреймворки (TensorFlow, PyTorch). Применение в бизнес-аналитике (обработка изображений, прогнозирование).
11.	Развертывание ML-моделей (MLOps)	Сериализация моделей (Pickle, Joblib). Деплой моделей (Flask, FastAPI, Docker). Мониторинг и обслуживание моделей.
12.	Кейсы применения ML в бизнесе	Прогнозирование оттока клиентов (Churn Prediction). Оптимизация маркетинговых кампаний. Анализ рисков и мошенничества (Fraud Detection).

6.2.2 Содержание практических занятий

№ п/п	Наименование темы (раздела) дисциплины	Содержание лекционного занятия
1.	Введение в анализ данных и машинное обучение	Основные понятия: данные, информация, знания. Жизненный цикл анализа данных (CRISP-DM, KDD). Обзор современных методов и инструментов (Python, R, BI-системы).
2.	Предобработка и исследование данных (EDA)	Работа с пропусками, выбросами, дубликатами. Нормализация, стандартизация, кодирование категориальных данных. Визуализация данных (Matplotlib, Seaborn, Tableau).
3.	Статистические методы анализа данных	Описательная статистика (среднее, дисперсия, корреляция). Проверка гипотез (t-тест, ANOVA, χ^2). Регрессионный анализ (линейная, логистическая регрессия).
4.	Основы машинного обучения	Обучение с учителем vs без учителя. Переобучение, валидация, кросс-валидация. Метрики качества (Accuracy, Precision, Recall, F1, ROC-AUC).
5.	Классификация и кластеризация	Методы классификации (k-NN, SVM, Decision Trees, Random Forest). Кластеризация (k-means, DBSCAN, иерархическая кластеризация). Применение в бизнесе (сегментация клиентов, анализ оттока).
6.	Ансамбли моделей и бустинг	Bagging (Random Forest) vs Boosting (XGBoost, LightGBM, CatBoost). Стекинг и блендинг моделей. Оптимизация гиперпараметров (GridSearch,

		RandomSearch).
7.	Работа с текстовыми данными (NLP)	Токенизация, стемминг, лемматизация. Векторизация текста (TF-IDF, Word2Vec, BERT). Анализ тональности, классификация текстов.
8.	Рекомендательные системы	Коллаборативная фильтрация (user-based, item-based). Контентная фильтрация. Гибридные системы (на примере e-commerce).
9.	Временные ряды и прогнозирование	Стационарность, автокорреляция. Методы прогнозирования (ARIMA, SARIMA, Prophet). Применение в бизнесе (прогноз продаж, спроса).
10.	Глубокое обучение (Deep Learning)	Введение в нейронные сети (MLP, CNN, RNN). Фреймворки (TensorFlow, PyTorch). Применение в бизнес-аналитике (обработка изображений, прогнозирование).
11.	Развертывание ML-моделей (MLOps)	Сериализация моделей (Pickle, Joblib). Деплой моделей (Flask, FastAPI, Docker). Мониторинг и обслуживание моделей.
12.	Кейсы применения ML в бизнесе	Прогнозирование оттока клиентов (Churn Prediction). Оптимизация маркетинговых кампаний. Анализ рисков и мошенничества (Fraud Detection).

6.2.3 Содержание самостоятельной работы

№ п/п	Наименование темы (раздела) дисциплины	Содержание лекционного занятия
1.	Введение в анализ данных и машинное обучение	Основные понятия: данные, информация, знания. Жизненный цикл анализа данных (CRISP-DM, KDD). Обзор современных методов и инструментов (Python, R, BI-системы).
2.	Предобработка и исследование данных (EDA)	Работа с пропусками, выбросами, дубликатами. Нормализация, стандартизация, кодирование категориальных данных. Визуализация данных (Matplotlib, Seaborn, Tableau).
3.	Статистические методы анализа данных	Описательная статистика (среднее, дисперсия, корреляция). Проверка гипотез (t-тест, ANOVA, χ^2). Регрессионный анализ (линейная, логистическая регрессия).
4.	Основы машинного обучения	Обучение с учителем vs без учителя. Переобучение, валидация, кросс-валидация. Метрики качества (Accuracy, Precision, Recall, F1, ROC-AUC).
5.	Классификация и кластеризация	Методы классификации (k-NN, SVM, Decision Trees, Random Forest). Кластеризация (k-means, DBSCAN, иерархическая кластеризация). Применение в бизнесе (сегментация клиентов, анализ оттока).
6.	Ансамбли моделей и бустинг	Bagging (Random Forest) vs Boosting (XGBoost, LightGBM, CatBoost). Стекинг и блендинг моделей.

		Оптимизация гиперпараметров (GridSearch, RandomSearch).
7.	Работа с текстовыми данными (NLP)	Токенизация, стемминг, лемматизация. Векторизация текста (TF-IDF, Word2Vec, BERT). Анализ тональности, классификация текстов.
8.	Рекомендательные системы	Коллаборативная фильтрация (user-based, item-based). Контентная фильтрация. Гибридные системы (на примере e-commerce).
9.	Временные ряды и прогнозирование	Стационарность, автокорреляция. Методы прогнозирования (ARIMA, SARIMA, Prophet). Применение в бизнесе (прогноз продаж, спроса).
10.	Глубокое обучение (Deep Learning)	Введение в нейронные сети (MLP, CNN, RNN). Фреймворки (TensorFlow, PyTorch). Применение в бизнес-аналитике (обработка изображений, прогнозирование).
11.	Развертывание ML-моделей (MLOps)	Сериализация моделей (Pickle, Joblib). Деплой моделей (Flask, FastAPI, Docker). Мониторинг и обслуживание моделей.
12	Кейсы применения ML в бизнесе	Прогнозирование оттока клиентов (Churn Prediction). Оптимизация маркетинговых кампаний. Анализ рисков и мошенничества (Fraud Detection).

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.	Введение в анализ данных и машинное обучение	Опрос, проблемно-аналитическое задание, тестирование.
2.	Предобработка и исследование данных (EDA)	Опрос, проблемно-аналитическое задание, тестирование.
3.	Статистические методы анализа данных	Опрос, проблемно-аналитическое задание, тестирование
4.	Основы машинного обучения	Опрос, проблемно-аналитическое задание, тестирование
5.	Классификация и кластеризация	Опрос, проблемно-аналитическое задание, тестирование
6.	Ансамбли моделей и бустинг	Опрос, проблемно-аналитическое задание, тестирование
7.	Работа с текстовыми данными (NLP)	Опрос, проблемно-аналитическое задание, тестирование
8.	Рекомендательные системы	Опрос, проблемно-аналитическое задание, тестирование
9.	Временные ряды и прогнозирование	Опрос, проблемно-аналитическое задание, тестирование
10.	Глубокое обучение (Deep Learning)	Опрос, проблемно-аналитическое задание, тестирование

11.	Развертывание ML-моделей (MLOps)	Опрос, проблемно-аналитическое задание, тестирование
12.	Кейсы применения ML в бизнесе	Опрос, проблемно-аналитическое задание, тестирование

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

8.1. Основная учебная литература

1. Зуев, С. В. Методы анализа данных : учебное пособие / С. В. Зуев. — Москва : Ай Пи Ар Медиа, 2023. — 132 с. — ISBN 978-5-4497-2107-5. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/129059.html>

2. Кумратова, А. М. Методы хранения и анализа данных : учебное пособие / А. М. Кумратова. — Москва : Ай Пи Ар Медиа, 2022. — 143 с. — ISBN 978-5-4497-1579-1. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/119065.html>

3. Воронова, Л. И. Machine Learning: регрессионные методы интеллектуального анализа данных : учебное пособие / Л. И. Воронова, В. И. Воронов. — Москва : Московский технический университет связи и информатики, 2018. — 82 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/81325.html>

4. Ракитский, А. А. Методы машинного обучения : учебно-методическое пособие / А. А. Ракитский, К. И. Дементьева. — Новосибирск : Сибирский государственный университет телекоммуникаций и информатики, 2023. — 35 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/138857.html>

5. Сараев, П. В. Методы машинного обучения : методические указания и задания к лабораторным работам по курсу / П. В. Сараев. — Липецк : Липецкий государственный технический университет, ЭБС АСВ, 2017. — 48 с. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/83183.html>

8.2. Дополнительная литература

6. Шевченко, А. С. Методы машинного обучения : учебное пособие / А. С. Шевченко, О. В. Самарина. — Москва : Ай Пи Ар Медиа, 2025. — 143 с. — ISBN 978-5-4497-4374-9. — Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. — URL: <https://www.iprbookshop.ru/150766.html>

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

1. <https://elibrary.ru> - Научная электронная библиотека eLIBRARY.RU (ресурсы открытого доступа)

2. <https://www.rsl.ru> - Российская Государственная Библиотека (ресурсы открытого доступа)

3. <https://link.springer.com> - Международная реферативная база данных научных изданий Springerlink (ресурсы открытого доступа)

4. <https://zbmath.org> - Международная реферативная база данных научных изданий zbMATH (ресурсы открытого доступа)

5. <https://openedu.ru> - «Национальная платформа открытого образования» (ресурсы открытого доступа)

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

Успешное освоение данного курса базируется на рациональном сочетании нескольких видов учебной деятельности – лекционных занятий, практических 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:

- работа с основной и дополнительной литературой, с материалами интернета и конспектами лекций;
- внеаудиторная подготовка к контрольным работам, выполнение докладов, рефератов и курсовых работ;
- выполнение самостоятельных практических работ;
- 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. В течение семестра необходимо подготовить рефераты с использованием рекомендуемой основной и дополнительной литературы и сдать рефераты для проверки преподавателю. Важным составляющим в изучении данного курса является решение различных задач и работа над проблемно-аналитическими заданиями, что предполагает знание соответствующей научной терминологии.

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:

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

На экзамене (зачете) 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, Система ГАРАНТ, Kaspersky Endpoint Security.

Перечень свободно распространяемого программного обеспечения:

Adobe Acrobat Reader DC, Google Chrome, LibreOffice, Skype, Zoom.

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, Система ГАРАНТ, Kaspersky Endpoint Security.

Перечень свободно распространяемого программного обеспечения:

Adobe Acrobat Reader DC, Google Chrome, LibreOffice, Skype, Zoom, 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 – лекционные занятия (типы лекций – установочная, вводная, текущая, заключительная, обзорная; виды лекций – проблемная, визуальная, лекция конференция, лекция консультация) и практические занятия, так и активные и интерактивные формы занятий – диспуты, решение ситуационных задач, ролевые игры и разбор конкретных ситуаций.

На учебных занятиях используются технические средства обучения – проектор, ноутбук, проекционный экран, колонки 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 студентов-инвалидов и студентов 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**

Современные методы анализа данных и машинное обучение

Направление подготовки	Бизнес-информатика
<i>Код</i>	38.03.05
<i>Направление (профиль)</i>	Информационные системы и технологии в бизнесе
Квалификация выпускника	бакалавр

Москва
2025

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

Группа компетенций	Категория компетенций	Код
Общепрофессиональные	-	ОПК-1
Общепрофессиональные	-	ОПК-4

2. Компетенции и индикаторы их достижения

Код компетенции	Формулировка компетенции	Индикаторы достижения компетенции
ОПК-1	Способен проводить моделирование, анализ и совершенствование бизнес-процессов и информационно-технологической инфраструктуры предприятия в интересах достижения его стратегических целей с использованием современных методов и программного инструментария;	ОПК-1.1. Знает: цели и задачи анализа, моделирования и совершенствования бизнес-процессов предприятия в интересах достижения его стратегических целей с использованием современных методов и программного инструментария ОПК-1.2. Умеет: проводить обследование, моделирование, анализ бизнес-процессов и ИТИ предприятия в интересах достижения его стратегических целей ОПК-1.3. Владеет: методами моделирования при решении задачи обоснования проекта по совершенствованию бизнес-процессов и ИТИ; методами совершенствования бизнес-процессов и ИТИ предприятия на основе использования современного программного инструментария для достижения его стратегических целей
ОПК-4	Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений;	ОПК-4.1. Знает методы сбора, обработки и анализа информации, в том числе в глобальных сетях, включая программные средства, методы представления информации, а также принципы работы информационных технологий ОПК-4.2. Умеет использовать математические и статистические методы анализа данных, в том числе с использованием компьютерных технологий, для информационноаналитической поддержки принятия управленческих решений ОПК-4.3. Владеет методами и программными средствами поддержки принятия управленческих решений

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	ОПК-1		
	- Методологии моделирования бизнес-процессов (BPMN, IDEF, UML) и принципы их применения для анализа организационной деятельности - Современные подходы к оптимизации ИТ-инфраструктуры предприятия, включая методы цифровой трансформации и автоматизации процессов - Программные средства бизнес-моделирования (ARIS, Bizagi, BPMS) и анализа ИТ-систем (ERP, CRM)	- Проводить диагностику бизнес-процессов, выявлять узкие места и разрабатывать предложения по их совершенствованию - Разрабатывать модели "как есть" и "как должно быть" с использованием специализированного ПО для визуализации процессов - Оценивать эффективность предлагаемых изменений с точки зрения достижения стратегических целей организации	- Навыками работы с инструментами бизнес-моделирования (нотации BPMN, EPC) и CASE-средствами - Методами анализа и реинжиниринга бизнес-процессов с учетом требований цифровой экономики - Технологиями интеграции бизнес-процессов с корпоративными информационными системами
Код компетенции	ОПК-4		
	- Архитектуру современных информационных систем и принципы их функционирования, включая клиент-серверные технологии, облачные вычисления и распределенные базы данных - Методы и технологии обработки данных (ETL-процессы, OLAP-анализ, методы машинного обучения) для аналитической поддержки решений - Ключевые показатели эффективности бизнес-	- Применять специализированное ПО (BI-системы, статистические пакеты, системы визуализации) для сбора и анализа бизнес-данных - Формулировать аналитические выводы на основе обработки структурированных и неструктурированных данных - Разрабатывать рекомендации для управленческих решений на основе результатов анализа данных	- Навыками работы с аналитическими платформами (Power BI, Tableau, Qlik) и языками запросов (SQL, DAX) - Методами статистического анализа и прогнозирования бизнес-показателей - Технологиями визуализации данных для представления аналитических отчетов руководству

	процессов и методики их расчета с использованием информационных систем		
--	--	--	--

3.2. Критерии оценки результатов обучения по дисциплине

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.
Компетенция не достигнута		
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.

При подготовке ответа в рамках прохождения промежуточной аттестации (зачет/ зачет с оценкой/ экзамен) обучающимся разрешается использовать калькулятор и справочные таблицы.

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

7 СЕМЕСТР

ОПК-1

1. Какой из этапов CRISP-DM включает построение и обучение моделей?

- А) Понимание бизнес-задачи
- Б) Понимание данных
- В) Моделирование
- Г) Оценка результатов

✓ Ответ: В

2. Какой процесс описывает преобразование сырых данных в полезные знания?

- А) Data Mining
- Б) Data Cleaning
- В) Data Visualization
- Г) Data Storage

✓ Ответ: А

3. Какой язык программирования чаще всего используется в машинном обучении?

- А) Java
- Б) Python
- В) C++
- Г) SQL

✓ Ответ: Б

4. Какой метод используется для обнаружения выбросов в данных?

- А) Нормализация
- Б) Метод межквартильного размаха (IQR)
- В) One-Hot Encoding
- Г) Логарифмирование

✓ Ответ: Б

5. Как называется метод преобразования категориальных данных в числовые?

- А) Стандартизация
- Б) Нормализация
- В) Кодирование
- Г) Кластеризация

✓ Ответ: В

6. Какая библиотека Python чаще всего используется для визуализации данных?

- А) NumPy
- Б) Pandas
- В) Matplotlib
- Г) Scikit-learn

✓ Ответ: В

7. Какая метрика показывает степень линейной зависимости между двумя переменными?

- А) Дисперсия
- Б) Среднее арифметическое
- В) Корреляция
- Г) Медиана

✓ Ответ: В

8. Какой тест используется для сравнения средних значений двух групп?

- А) χ^2 (хи-квадрат)
- Б) t-тест
- В) ANOVA
- Г) Регрессионный анализ

✓ Ответ: Б

9. Какой тип регрессии используется для предсказания бинарных переменных?

- А) Линейная регрессия
- Б) Логистическая регрессия
- В) Полиномиальная регрессия
- Г) Гребневая регрессия

✓ Ответ: Б

10. Как называется проблема, когда модель слишком хорошо запоминает обучающие данные и плохо обобщает?

- А) Недообучение
- Б) Переобучение
- В) Регуляризация
- Г) Кросс-валидация

✓ Ответ: Б

11. Какая метрика оценивает долю правильно классифицированных объектов?

- А) Precision
- Б) Recall
- В) Accuracy
- Г) F1-score

✓ Ответ: В

12. Какой метод помогает избежать переобучения путем разбиения данных на несколько подмножеств?

- А) GridSearch
- Б) RandomSearch
- В) Кросс-валидация
- Г) Бустинг

✓ Ответ: В

13. Какой алгоритм классификации основан на расстоянии между точками?

- А) SVM
- Б) k-NN
- В) Decision Tree
- Г) Random Forest

✓ Ответ: Б

14. Какой метод кластеризации использует понятие "плотности" данных?

- А) k-means
- Б) DBSCAN
- В) Иерархическая кластеризация
- Г) Метод k-медиан

✓ Ответ: Б

15. Какой бизнес-задачей может заниматься кластеризация?

- А) Прогнозирование цен
- Б) Сегментация клиентов
- В) Обнаружение мошенничества
- Г) Анализ временных рядов

✓ Ответ: Б

16. Какой метод ансамблирования использует несколько параллельных моделей?

- А) Bagging
- Б) Boosting
- В) Stacking
- Г) Blending

✓ Ответ: А

17. Какой алгоритм бустинга использует градиентный спуск?

- А) Random Forest
- Б) XGBoost
- В) k-NN
- Г) SVM

✓ Ответ: Б

18. Какой метод оптимизации гиперпараметров перебирает все возможные комбинации?

- А) RandomSearch
- Б) GridSearch
- В) Bayesian Optimization
- Г) Genetic Algorithms

✓ Ответ: Б

19. Какой метод преобразует текст в числовой вектор на основе частоты слов?

- А) Word2Vec
- Б) TF-IDF
- В) BERT
- Г) LSTM

✓ Ответ: Б

20. Как называется процесс приведения слова к его базовой форме?

- А) Токенизация
- Б) Лемматизация
- В) Стемминг
- Г) Векторизация

✓ Ответ: Б

21. Какой метод рекомендаций основан на сходстве пользователей?

- А) Контентная фильтрация
- Б) Коллаборативная фильтрация (user-based)
- В) Гибридный подход
- Г) Матричная факторизация

✓ Ответ: Б

22. Как называется свойство временного ряда, при котором его статистические характеристики не меняются со временем?

- А) Автокорреляция
- Б) Стационарность
- В) Тренд
- Г) Сезонность

✓ Ответ: Б

23. Какой метод прогнозирования учитывает сезонность?

- А) ARIMA
- Б) SARIMA
- В) Линейная регрессия
- Г) Prophet

✓ Ответ: Б

24. Какая архитектура нейронных сетей используется для обработки изображений?

- А) MLP
- Б) CNN
- В) RNN
- Г) Transformer

✓ Ответ: Б

25. Какой инструмент используется для упаковки модели в контейнер?

- А) Flask
- Б) FastAPI
- В) Docker
- Г) Kubernetes

✓ Ответ: В

ОПК-4

1. Какой этап анализа данных включает проверку на пропущенные значения и выбросы?

- А) Визуализация данных
- Б) Предобработка данных ✓
- В) Построение моделей
- Г) Интерпретация результатов

Ответ: Б

2. Какой график НЕ используется для EDA?

- А) Гистограмма
- Б) Box plot
- В) Scatter plot
- Г) ROC-кривая ✓

Ответ: Г

3. Какая мера центральной тенденции наиболее устойчива к выбросам?

- А) Среднее арифметическое
- Б) Медиана ✓
- В) Мода
- Г) Дисперсия

Ответ: Б

4. Какой алгоритм относится к обучению с учителем?

- А) K-means
 - Б) Дерево решений ✓
 - В) DBSCAN
 - Г) Метод главных компонент
- Ответ: Б

- 5.Какой метод кластеризации требует задания числа кластеров?
- А) Иерархическая кластеризация
 - Б) K-means ✓
 - В) DBSCAN
 - Г) OPTICS
- Ответ: Б

- 6.Какой ансамблевый метод последовательно улучшает модель?
- А) Bagging
 - Б) Random Forest
 - В) AdaBoost ✓
 - Г) Voting Classifier
- Ответ: В

- 7.Какой этап НЕ входит в обработку текстовых данных?
- А) Токенизация
 - Б) Лемматизация
 - В) Нормализация
 - Г) Кластеризация ✓
- Ответ: Г

- 8.Какой подход НЕ используется в рекомендательных системах?
- А) Коллаборативная фильтрация
 - Б) Контентная фильтрация
 - В) Гибридные системы
 - Г) Линейная регрессия ✓
- Ответ: Г

- 9.Какой компонент временного ряда отражает долгосрочные изменения?
- А) Тренд ✓
 - Б) Сезонность
 - В) Шум
 - Г) Цикличность
- Ответ: А

- 10.Какая архитектура НЕ относится к глубокому обучению?
- А) Сверточные сети (CNN)
 - Б) Рекуррентные сети (RNN)
 - В) Random Forest ✓
 - Г) Трансформеры
- Ответ: В

- 11.Какой инструмент НЕ относится к MLOps?
- А) Docker
 - Б) Kubernetes

В) TensorFlow ✓

Г) Airflow

Ответ: В

12. Какой бизнес-кейс НЕ относится к применению ML?

А) Прогнозирование спроса

Б) Обнаружение мошенничества

В) Управление персоналом ✓

Г) Персонализация рекомендаций

Ответ: В

13. Какой метод НЕ используется для обработки пропущенных данных?

А) Удаление строк с пропусками

Б) Замена средним значением

В) Интерполяция

Г) Нормализация ✓

Ответ: Г

14. Какой коэффициент показывает линейную зависимость между переменными?

А) Коэффициент корреляции Пирсона ✓

Б) Коэффициент детерминации

В) Коэффициент вариации

Г) Коэффициент асимметрии

Ответ: А

15. Какой алгоритм НЕ является методом классификации?

А) Логистическая регрессия

Б) Метод k-ближайших соседей

В) Метод опорных векторов

Г) Метод главных компонент ✓

Ответ: Г

16. Какой метод бустинга использует градиентный спуск?

А) AdaBoost

Б) Gradient Boosting ✓

В) Random Forest

Г) Bagging

Ответ: Б

17. Какой метод НЕ используется в NLP?

А) TF-IDF

Б) Word2Vec

В) One-Hot Encoding

Г) K-means ✓

Ответ: Г

18. Какой тип рекомендательной системы использует историю взаимодействий пользователей?

А) Коллаборативная фильтрация ✓

Б) Контентная фильтрация

В) Демографическая фильтрация

Г) Гибридная система

Ответ: А

19.Какой метод НЕ используется для прогнозирования временных рядов?

А) ARIMA

Б) Экспоненциальное сглаживание

В) Логистическая регрессия ✓

Г) Prophet

Ответ: В

20.Какая функция активации НЕ используется в нейронных сетях?

А) ReLU

Б) Sigmoid

В) Tanh

Г) Median ✓

Ответ: Г

21.Какой инструмент используется для мониторинга ML-моделей?

А) MLflow ✓

Б) Scikit-learn

В) Pandas

Г) Matplotlib

Ответ: А

22.Какой бизнес-метрикой НЕ оценивают ML-модели?

А) Accuracy

Б) ROC-AUC

В) F1-score

Г) EBITDA ✓

Ответ: Г

23.Какой метод НЕ относится к feature engineering?

А) Создание полиномиальных признаков

Б) Нормализация данных

В) Удаление выбросов

Г) Обучение модели ✓

Ответ: Г

24.Какой алгоритм НЕ требует нормализации данных?

А) Метод k-ближайших соседей

Б) Дерево решений ✓

В) Метод опорных векторов

Г) Нейронные сети

Ответ: Б

25.Какой подход НЕ используется для борьбы с переобучением?

А) Регуляризация

Б) Увеличение данных

В) Ранняя остановка

Г) Уменьшение количества признаков ✓

Ответ: Г