

Федеральное государственное образовательное бюджетное учреждение
высшего образования

**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ
РОССИЙСКОЙ ФЕДЕРАЦИИ»
(Финансовый университет)**

Краснодарский филиал Финуниверситета

Кафедра «Математика и информатика»

СОГЛАСОВАНО

ООО «Портал-Юг»
Генеральный директор



Е.В. Мостовой

«20» февраля 2024 г.

УТВЕРЖДАЮ

Краснодарский филиал
Финуниверситета

Директор



Э.В. Соболев

«20» февраля 2024 г.

Хроль Е.В.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ТЕХНОЛОГИИ ПРОДВИНУТОЙ АНАЛИТИКИ**
студентов, обучающихся по направлению подготовки
38.03.05 Бизнес-информатика
в соответствии с образовательными стандартами Краснодарского филиала
Финансового университета
(программа подготовки бакалавров)

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

*Одобрено кафедрой «Математика и информатика»
(протокол № 13 от 27.02.2024)*

Краснодар 2024

УДК 004(073)
ББК 65.290с51
Д30, Х94

Рецензенты: доктор физико-математических наук, профессор кафедры «Математика и информатика» Калайдин Е.Н.. Кандидат физико-математических наук, доцент кафедры «Математика и информатика» Кирий В.А.

Хроль Е.В. «Технологии продвинутой аналитики». Рабочая программа дисциплины для студентов, обучающихся по направлению подготовки 38.03.05 «Бизнес-информатика» – Краснодар: Краснодарский филиал Финуниверситета, кафедра «Математика и информатика», 2024 г.

Дисциплина Технологии продвинутой аналитики относится к модулю профиля по направлению подготовки 38.03.05 Бизнес-информатика.

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

Рабочая программа дисциплины

ТЕХНОЛОГИИ ПРОДВИНУТОЙ АНАЛИТИКИ

Учебное издание

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

Усл. п.л. 2,0. Изд. № _от.

Тираж 100 экз.

Заказ № .

Отпечатано в Краснодарском филиале Финуниверситета

© Хроль Е.В.
© Краснодарский филиал Финуниверситета, 2024

СОДЕРЖАНИЕ

1. Наименование дисциплины	5
2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения, соотнесенных с планируемыми результатами обучения по дисциплине	5
3. Место дисциплины в структуре образовательной программы	5
4. Объем дисциплины в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	6
5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	6
5.1. Содержание дисциплины	6
5.2. Учебно-тематический план	9
5.3. Содержание семинаров, практических занятий	9
6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	10
6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	10
6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю	11
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	12
7.1. Описание показателей и критериев оценивания компетенций	12
7.2. Вопросы для оценки знаний и умений, характеризующих формирование компетенций	14
7.3. Тесты	17
8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	19
9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	20
10. Методические материалы, определяющие процедуры оценивания знаний и умений, характеризующих степень сформированности компетенций	20
11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	22
11.1. Комплект лицензионного программного обеспечения:	22
11.2. Современные профессиональные базы данных и информационные справочные системы:	22
11.3. Сертифицированные программные и аппаратные средства защиты информации:	22
12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	22

1. Наименование дисциплины

Б1.В.03.ДВ.02.03 «Технологии продвинутой аналитики».

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

Дисциплина «Технологии продвинутой аналитики» обеспечивает формирование следующих компетенций: ПКП-3.

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (владения, умения и знания), соотнесенные с компетенциями/индикаторами достижения компетенции
ПКП-3	Способность предлагать различные варианты инфраструктурных решений для поддержки ИТ/ИС	1. Анализирует текущий уровень инфраструктурных решений предприятия/организации	Знать: текущий уровень инфраструктурных решений предприятия/организации Уметь: анализировать текущий уровень инфраструктурных решений предприятия/организации
		2. Формирует и обосновывает варианты технологического слоя архитектуры предприятия/организации	Знать: варианты технологического слоя архитектуры предприятия/организации Уметь: формировать и обосновывать варианты технологического слоя архитектуры предприятия/организации

3. Место дисциплины в структуре образовательной программы

Дисциплина «Технологии продвинутой аналитики» относится к Модулю «Информационно-аналитические технологии», формируемого участниками образовательных отношений по направлению подготовки 38.03.05 - Бизнес-информатика ОП «Цифровая трансформация управления бизнесом» профиль «Бизнес-аналитика».

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

Очно-заочная форма обучения.

Вид учебной работы по дисциплине	Всего (в з/ед. и часах)	Семестр 8 (в часах)
Общая трудоемкость дисциплины	3/108	108
Контактная работа - Аудиторные занятия	38	38
Лекции	12	12
Семинары, практические занятия	26	26
Самостоятельная работа	34	34
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	Экзамен	Экзамен
Контроль	36	36

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

5.1.Содержание дисциплины

Тема 1. Deepnote

Бесплатный онлайн-блокнот Deepnote для специалистов по данным, фокусирующийся в основном на совместном использовании в реальном времени и абстрагировании всех понятий, усложняющих работу аналитика, таких как настройка окружения и инфраструктуры. Используемые языки: Jupyter-совместимый, языки программирования Jupyter (например, Python, R), SQL. Коннекторы: подключения к библиотекам Jupyter (например, SQLAlchemy, psycopg2), к хранилищам данных и озерам данных (AWS, GCP и т.д.), базам данных (PostgreSQL, MongoDB и т.д.), предоставляемое файловое хранилище. Визуализация данных Deepnote: Jupyter (например, библиотеки Matplotlib, Altair, Plotly), пользовательский интерфейс для построения диаграмм.

Машинное обучение - поддерживает интеграции: встроенная интеграции с MindsDB, интеграции с Weights & Biases, Comet.ml, Neptune.ai.

MindsDB привносит машинное обучение в базы данных, используя концепцию AI Tables — это модели машинного обучения, хранящиеся в виде виртуальных таблиц внутри базы данных. Позволяют выполнять прогнозы временных рядов, регрессии и классификации в своей базе данных и получать результаты, запрашивая таблицу AI с помощью простых операторов SQL. Weights & Biases помогает отслеживать проекты машинного обучения. Используйте его для регистрации гиперпараметров и вывода метрик из ваших прогонов или для

визуализации и сравнения результатов. Comet позволяет специалистам по данным и командам отслеживать, сравнивать, объяснять и оптимизировать эксперименты и модели на протяжении всего жизненного цикла модели. Neptune - это легкий инструмент управления экспериментами, который помогает отслеживать эксперименты с машинным обучением.

Тема 2. Jetbrains Datalore

Jetbrains Datalore - это мощная онлайн-среда для Jupyter-ноутбуков, разработанная в JetBrains: использование интеллектуальной помощи по написанию кода для Python в онлайн-блокнотах Jupyter, запуск кода на мощных процессорах и графических процессорах, сотрудничество в режиме реального времени и легкость в возможностях делиться результатами. Анализ данных с помощью умного кодирования и автоматизации без кода. Создание интерактивных приложений для работы с данными, созданными из записных книжек всего за несколько кликов.

Функции: Jupyter-совместимый; языки программирования Jupyter (например, Python, R); подключение к библиотекам Jupyter (например, SQLAlchemy, psycpg2), базам данных (MySQL, PostgreSQL), хранилищам данных и озерам данных (AWS, GCP и т.д.), предоставляемое файловое хранилище; визуализация данных Jupyter (например, Matplotlib, Altair, Plotly) и пользовательский интерфейс для построения диаграмм; полная реактивность в реальном времени; встроен планировщик записной книжки.

Тема 3. Платформа Mode

Mode - платформа, объединяющая команды по обработке данных и бизнес-группы вокруг данных для достижения бизнес-результатов. В Mode SQL, R, Python и визуализация данных связаны друг с другом, чтобы помочь вам получать ценную информацию: анализ данных с помощью SQL, используя блокноты R или Python (нативные блокноты R и Python). В отличие от многих других инструментов бизнес-аналитики, в Mode есть все функции и возможности, необходимые для решения вопросов расширенной аналитики, - и все это на одной платформе. Переключение с SQL на полностью оборудованную среду для ноутбуков Python и R одним кликом мыши.

Mode не совместима с Jupyter, языки программирования: SQL, используя блокноты R или Python. Подключение к базам данных (MariaDB, PostgreSQL и др.), хранилищам данных и озерам данных (Amazon Redshift, Google BigQuery и др.). Помимо библиотек Python использует пользовательский интерфейс для построения диаграмм. Интерактивные информационные панели без кода, которые может создать каждый: создание группой аналитиков данных дашбордов в рекордно короткие сроки, а бизнес-группы могут использовать их в качестве отправной точки для самообслуживания.

Тема 4. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.

Библиотека Polars использует эффективные алгоритмы распараллеливания и кэширования для ускорения выполнения аналитических задач: уменьшение количества избыточных копий; эффективное кэширование памяти; сведение конфликтов при параллелизме к минимуму; реализована на Rust, а не на Python (Rust намного лучше реализует параллелизм, нежели Python).

Dask: планировщик низкого уровня и частичная замена Pandas высокого уровня, ориентированный на выполнение кода на вычислительных кластерах. Ray: низкоуровневый фреймворк для распараллеливания кода Python между процессорами или кластерами. Modin: замена Pandas на базе Dask или Ray. Vaex: частичная замена Pandas, использующая ленивую оценку и сопоставление памяти, чтобы позволить разработчикам работать с большими наборами данных на стандартных машинах.

Modin с Ray в качестве бэкенда. Установив их, можно добиться значительное преимущество, изменив всего одну строку («import pandas as pd» на «import modin.pandas as pd»). В отличие от других инструментов, Modin стремится достичь полной совместимости с Pandas. Dask, более крупный и, следовательно, более сложный проект. Dask также предоставляет Dask.dataframe , библиотеку более высокого уровня, подобную Pandas, которая помогает справиться с наборами данных вне ядра. Vaex, предназначенный для работы с большими объемами данных на стандартном ноутбуке. Его замена Pandas охватывает некоторые API Pandas, но больше ориентирована на исследование и визуализацию.

5.2. Учебно-тематический план

№ п/п	Наименование темы (раздела) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа- Аудиторная работа			Самостоя тельная работа	
			Общая	Лекции	Практ. и семинарск ие занятия		
1	Тема 1. Deepnote	18	9	3	6	9	Выполнение индивидуальных заданий
2	Тема 2. JetBrains Datalore	18	10	3	7	8	Выполнение индивидуальных заданий
3	Тема 3. Платформа Mode	18	9	3	6	9	Выполнение индивидуальных заданий
4	Тема 4. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.	18	10	3	7	8	Выполнение индивидуальных заданий
Контроль		36					
В целом по дисциплине		108	38	12	26	34	Контрольная работа

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
Тема 1. Deepnote	1. Разработка сценариев ABC, XYZ, ABC- XYZ анализов в Deepnote 2. Визуализация результатов аналитики: в библиотеки Altair, Plotly. Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий
Тема 2. JetBrains Datalore	1. Разработка сценариев RFM-анализа в Datalore 2. Создание интерактивных приложений на основе анализа данных. Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий

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
Тема3. Платформа Mode	1. Разработка сценариев когортного анализа в Datalore 2. Создание интерактивных информационных панелей на основе анализа данных. Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий
Тема4.Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.	1. Базовая аналитика и описательные статистики на больших объемах структурированных данных с использованием библиотек Polar, Vaex, Dask, Modin, Ray. 2. Визуальная аналитика больших объемах структурированных данных Основная литература: 1,2 Дополнительная литература: 1,2,3	Выполнение и защита практических заданий

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

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

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Тема 1. Deepnote	Машинное обучение в Deepnote - интеграции: интеграции с Weights & Biases, Comet.ml, Neptune.ai.	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет - источников.
Тема 2. JetBrains Datalore	Создание интерактивных приложений на основе ABC-XYZ анализов в JetBrains Datalore	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет - источников.
Тема 3. Платформа Mode	Создание интерактивных информационных панелей на основе совместного ABC-XYZ, RFF и когортного анализа	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет - источников.

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Тема 4. Библиотеки python для анализа больших объемов структурированных данных: Polar, Vaex, Dask, Modin, Ray.	Машинное обучение в Dask	Изучение методических материалов по теме в электронном виде и рекомендуемых разделов основной и дополнительной литературы, интернет - источников.

6.2.Перечень вопросов, заданий, тем для подготовки к текущему контролю

Примерные темы для контрольной работы:

Студенту в начале семестра предлагается выбрать кейс, на основе которого разрабатываться аналитический проект. Для этого можно использовать датасеты с Kaggle <https://www.kaggle.com/datasets?search=customer>, содержащие клиентские данные (всего 2466 различных Datasets).

Например:

1. Customer Personality Analysis.
2. Telco Customer Churn.
3. Shop Customer Data.
4. Customer Segmentation.
5. Mall Customer Segmentation Data.
6. Credit Card customers.
7. Starbucks Customer Survey.
8. Airlines Customer satisfaction.
9. Loan Prediction Based on Customer Behavior.
10. Brazilian E-Commerce Public Dataset by Olist.

Задание:

1. Импортировать данные в BigQuery. Обогащать датасет, проводя клиентскую аналитику: выполнить ABC-XYZ, RFM, когортный -анализ, используя библиотеки Python.
2. Создать, обучить, оценить по метрикам качества, объяснить модели машинного обучения библиотеки Python; создать интерактивный отчет (дашборд) в Datalore (или Mode, Deepnote).

Критерии балльной оценки различных форм текущего контроля успеваемости

Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях кафедры «Математика и информатика» Краснодарского филиала Финансового университета.

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

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «Технологии продвинутой аналитики».

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

7.1.Описание показателей и критериев оценивания компетенций

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетворительно»	«удовлетворительно»	«хорошо»	«отлично»	
ПКП-3 Способность предлагать различные варианты инфраструктур-ных решений для поддержки ИТ/ИС					
Анализирует текущий уровень инфраструктурных решений предприятия/организации					
Знать: Текущий уровень инфраструктурных решений предприятия/организации	Фрагментарное представление о текущем уровне инфраструктурных решений предприятия/организации	Неполные представления о текущем уровне инфраструктурных решений предприятия/организации	Сформированные, но содержащие отдельные пробелы представления о текущем уровне инфраструктурных решений предприятия/организации	Сформированные систематические представления о текущем уровне инфраструктурных решений предприятия/организации	Вопросы для оценки знаний и умений, тестовые задания.
Уметь: Анализовать текущий уровень инфраструктурных решений предприятия/организации	Фрагментарное умение анализировать текущий уровень инфраструктурных решений предприятия/организации	Несистематическое умение анализировать текущий уровень инфраструктурных решений предприятия/организации	В целом успешное, но содержащее отдельные пробелы умение анализировать текущий уровень инфраструктурных решений предприятия/организации	Сформированное умение анализировать текущий уровень инфраструктурных решений предприятия/организации	Вопросы для оценки знаний и умений, тестовые задания.

Планируемые результаты освоения компетенции (индикатора достижения компетенции)	Уровень освоения				Оценочное средство
	«неудовлетвори тельно»	«удовлетвори тельно»	«хорошо»	«отлично»	
рных решений предприятия/ организации	инфраструкту рных решений предприятия/ организации	инфраструкту рных решений предприятия/ организации	умение анализироват ь текущий уровень инфраструкту рных решений предприятия/ организации	инфраструкту рных решений предприятия/ организации	
Формирует и обосновывает варианты технологического слоя архитектуры предприятия/организации					
Знать: Варианты технологичес кого слоя архитектуры предприятия/ организации	Фрагментарн ое представлени е о вариантах технологичес кого слоя архитектуры предприятия/ организации	Неполные представлени я о вариантах технологичес кого слоя архитектуры предприятия/ организации	Сформирован ные, но содержащие отдельные пробелы представлени я о вариантах технологичес кого слоя архитектуры предприятия/ организации	Сформирован ные систематичес кие представлени я о вариантах технологичес кого слоя архитектуры предприятия/ организации	Вопросы для оценки знаний и умений, тестовые задания.
Уметь: Формировать и обосновывать варианты технологичес кого слоя архитектуры предприятия/ организации	Фрагментарн ое умение формировать и обосновывать варианты технологичес кого слоя архитектуры предприятия/ организации	Несистематич еское умение формировать и обосновывать варианты технологичес кого слоя архитектуры предприятия/ организации	В целом успешное, но содержащее отдельные пробелы умение формировать и обосновывать варианты технологичес кого слоя архитектуры предприятия/ организации	Сформирован ное умение формировать и обосновывать варианты технологичес кого слоя архитектуры предприятия/ организации	Вопросы для оценки знаний и умений, тестовые задания.

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

Шифр компетенции	Вопросы	Правильный ответ
ПКП-3	1. Какие существуют методы анализа данных?	Описательная статистика; частотный анализ, корреляционный, регрессионный, дисперсионный, кластерный, факторный, дискриминантный; деревья решений; байесовский анализ.
	2. Дайте определение машинному обучению?	Разработка алгоритмов для автоматической обработки данных и обучение на основе этих данных.
	3. Какие основные алгоритмы машинного обучения вы знаете?	Кластеризация; классификация; регрессионный; обнаружение аномалий; снижение размерности; обучение с подкреплением и без учителя; нейросети.
	4. В чем заключается задача классификации?	Научить компьютерную программу определять, к какому классу относится объект на основе его характеристик.
	5. Дайте определение кластеризации?	Разделения данных на группы, так чтобы объекты в одной группе были более похожи друг на друга, чем на объекты из других групп.
	6. Дайте определение регрессионному анализу?	Используется для предсказания значения зависимой переменной.
	7. Дайте определение анализу временных рядов?	Изучение закономерностей изменения данных с течением времени.
	8. Дайте определение прогнозированию?	Предсказывания будущих значений какого-либо показателя
	9. Дайте определение нейронной сети?	Математическая модель, имитирующая работу нервной системы живых организмов.
	10. Дайте определение глубокому обучению?	Подраздел машинного обучения, использующего нейронные сети с большим количеством слоев.
	11. Дайте определение обработке естественного языка?	Анализ и обработка текстов на естественном языке.
	12. Дайте определение рекомендательной системе?	На основе предпочтений пользователей предлагает им товары или услуги, которые могут их заинтересовать.
	13. Дайте определение коллаборативной фильтрации?	Основывается на предположении, похожие наборы элементов, могут иметь схожие предпочтения.
	14. Дайте определение системе поддержки принятия решений?	Компьютерная система, помогающая принимающим решениям, собирать, анализировать и интерпретировать данные для принятия решений.

Шифр компетенции	Вопросы	Правильный ответ
	15. Дайте определение text mining?	Автоматическое извлечение информации из больших объемов текстовых данных.
	16. Дайте определение информационному поиску?	Процесс поиска информации в больших коллекциях документов с использованием различных методов.
	17. В чем суть задачи оптимизации?	Нахождение решения, которое бы максимизировало или минимизировало целевую функцию при заданных ограничениях.
	18. Какие методы оптимизации вы знаете?	Метод линейного программирования, нелинейного, динамическое программирование, стохастическая и выпуклая оптимизации, генетические алгоритмы.
	19. Дайте определение data mining?	Обнаружение полезных знаний из больших объемов данных.
	20. Дайте определение анализу данных?	Исследование и интерпретация данных с целью извлечения полезной информации.
	21. Дайте определение статистическому анализу?	Изучает методы обработки и анализа данных.
	22. Какие задачи можно решать с помощью машинного обучения?	Классификация, регрессия, кластеризация, обнаружение аномалий, обучение с подкреплением, обработка естественного языка, рекомендательные системы, распознавание образов, прогнозирование временных рядов.
	23. Какие алгоритмы кластеризации вы знаете?	K-means, иерархическая кластеризация, DBSCAN.
	24. Какие есть методы оценки качества кластеризации?	Внутриклассовая дисперсия, расстояние до центроидов.
	25. Какие существуют подходы к решению задачи классификации?	Машинное обучение, правила ассоциации.
	26. Какие метрики оценки качества классификатора вы знаете?	Точность, полнота, F1-мера.
	27. Как выбрать оптимальный гиперпараметр для классификатора?	Метод перебора, случайного поиска, стохастической аппроксимации.
	28. Как устроены рекуррентные нейронные сети?	Обработка последовательностей данных.
	29. Какие архитектуры сверточных нейронных сетей вы знаете?	AlexNet, VGGNet, GoogleNet, ResNet и DenseNet.

Шифр компетенции	Вопросы	Правильный ответ
	30.Как работает автоэнкодер?	Сжатие и восстановление данных.
	31.Как работают генеративно-состязательные сети?	Генерация новых данных.
	32.Как работают рекомендательные системы на основе коллаборативной фильтрации?	Используют сходство между пользователями или предметами для рекомендации.
	33.В чем особенность систем поддержки принятия решений основанных на анализе больших данных?	Используют алгоритмы машинного обучения и статистического анализа для выявления закономерностей и тенденций в данных.
	34.Как обрабатывать текстовые данные с использованием text mining?	Анализ текста и классификация, извлечение информации.
	35.Как оптимизировать распределение ресурсов с использованием алгоритмов машинного обучения?	Оптимизация распределения ресурсов путем обучения модели на исторических данных и использования ее для прогнозирования оптимального распределения в будущем.
	36.Как применять методы data mining для анализа данных в бизнесе?	Прогнозирование продаж или спроса, выявление тенденций и закономерностей в данных, сегментация клиентов и определение их потребностей.
	37.Какие примеры использования машинного обучения в бизнесе вы можете привести?	Прогнозирование продаж, сегментация клиентов, обнаружение мошенничества, улучшение рекомендаций.
	38.Какие существуют способы визуализации результатов анализа данных с использованием машинного обучения?	Тепловые карты; столбчатые, круговые, точечные, пузырьковые диаграммы; гистограммы; линейные графики; карты плотности; леса Ишикавы; древовидные карты.

7.3. Тесты

Шифр компетенции	Тестовые задания	Правильный ответ
ПКП-3	1. Выберите из следующих технологий продвинутой аналитики ту, которая использует алгоритмы машинного обучения для анализа больших объемов данных: А) Text mining; Б) Анализ социальных сетей; В) Обработка естественного языка; Г) Все из вышеперечисленного.	Г
	2. Выберите из следующих методов Text mining тот, который используется для разделения текстов на разные категории: А) Анализ текста; Б) Классификация текста; В) Извлечение информации; Г) Нет правильного ответа.	Б
	3. Укажите пример использования машинного обучения в бизнес-аналитике: А) Прогнозирование спроса; Б) Сегментация клиентов; В) Обнаружение мошеннических транзакций; Г) Все из перечисленного.	Г
	4. Укажите пример системы поддержки принятия решений, основанной на анализе больших данных: А) Системы прогнозирования продаж; Б) Системы управления рисками; В) Системы обнаружения мошенничества; Г) Все из перечисленных.	Г
	5. Определение генеративно-сопоставительная сеть (GAN): А) Сеть, использующая машинное обучение для прогнозирования будущего; Б) Сеть, состоящая из двух частей - генератора и дискриминатора; В) Сеть, предназначенная для обнаружения аномалий в данных; Г) Сеть, которая извлекает информацию из больших объемов данных.	Б
	6. Выберите из нижеперечисленного пример применения data mining: А) Определение наиболее эффективных маркетинговых кампаний; Б) Прогнозирование объема продаж; В) Сегментация клиентской базы; Г) Все перечисленное.	Г
	7. Выберите из нижеперечисленного определение «автоэнкодер»: А) Это алгоритм машинного обучения, используемый для сжатия и распаковки данных;	

Шифр компетенции	Тестовые задания	Правильный ответ
	Б) Это метод data mining, предназначенный для классификации текстов; В) Это алгоритм, применяемый для распознавания образов; Г) Это инструмент для определения наиболее эффективных маркетинговых кампаний.	
	8. Выберите систему, которая может быть отнесена к системам поддержки принятия решений: А) Система прогнозирования объема продаж; Б) Система управления рисками; В) Система обнаружения мошенничества; Г) Система управления рисками и обнаружения мошенничества	А
	9. Укажите особенность систем поддержки принятия решений на базе анализа больших данных: А) Они позволяют анализировать огромные объемы данных для принятия решений; Б) Они обеспечивают высокую точность прогнозов; В) Они позволяют создавать модели, предсказывающие будущее; Г) Они используются только для обнаружения мошеннических транзакций.	А
	10. Выберите то, что относится к методам визуализации результатов анализа данных: А) Тепловые карты; Б) Столбчатые диаграммы; В) Круговые диаграммы; Г) Все вышеперечисленные.	Г
	11. Укажите, что относится к способам оптимизации распределения ресурсов: А) Алгоритмы машинного обучения; Б) Методы data mining; В) Технологии продвинутой аналитики; Г) Все вышеперечисленное.	Г
	12. Выберите определение «рекуррентные нейронные сети»: А) Нейронные сети, способные обрабатывать временные ряды; Б) Алгоритмы, используемые для прогнозирования временных рядов; В) Алгоритмы, применяемые для распознавания речи; Г) Инструменты для анализа социальных сетей.	А

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

Основная литература:

1. Кутергин, В. А. Бизнес-инжиниринг. Модельная интерпретация управления изменениями : учебное пособие для вузов / В. А. Кутергин. — Санкт-Петербург : Лань, 2022. — 396 с. — ISBN 978-5-8114-8531-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/193355> . — Режим доступа: для авториз. пользователей.
2. Прокопенко, Н. Ю. Аналитические информационные системы поддержки принятия решений : учебное пособие / Н. Ю. Прокопенко. — Нижний Новгород : ННГАСУ, 2020. — 142 с. — ISBN 978-5-528-00395-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/164866> . — Режим доступа: для авториз. пользователей.
3. Сиганьков, А. А. Управленческий контроллинг бизнес-процессов : учебное пособие / А. А. Сиганьков. — Москва : РТУ МИРЭА, 2020. — 72 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/163923> . — Режим доступа: для авториз. пользователей.

Дополнительная литература:

4. Алетдинова, А. А. Интеллектуальный анализ больших данных : учебное пособие / А. А. Алетдинова, М. Ш. Муртазина. — Новосибирск : НГТУ, 2023. — 66 с. — ISBN 978-5-7782-4899-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/404567> — Режим доступа: для авториз. пользователей.
5. Телишев, А. М. Организация коллаборативного производства : монография / А. М. Телишев. — Казань : КНИТУ-КАИ, 2021. — 152 с. — ISBN 978-5-7579-2541-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/248942> . — Режим доступа: для авториз. пользователей.

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

1. <https://cloud.google.com/bigquery/docs/sandbox> - страница BigQuery sandbox.
2. <https://www.anaconda.com/> - страница загрузки Anaconda.
3. <https://cloud.yandex.ru/services/datalens> - сервис визуализации и анализа данных Яндекс.
4. Портал открытых данных Российской Федерации <https://data.gov.ru>
5. <https://cloud.yandex.ru/training/corpplatform> - практический курс «Построение корпоративной аналитической платформы»
6. <https://practicum.yandex.ru/ycloud/> - бесплатный курс «Инженер облачных сервисов»
7. [https://rise.articulate.com/share/BtQjK0gEy1lktRKR6q2hPZ5KnRDJhB8k#/-](https://rise.articulate.com/share/BtQjK0gEy1lktRKR6q2hPZ5KnRDJhB8k#/) Ростелеком бизнес. Платформа управления данными

10.Методические материалы, определяющие процедуры оценивания знаний и умений, характеризующих степень сформированности компетенций

Рекомендации по подготовке к лекционным занятиям

Изучение дисциплины требует систематического и последовательного накопления знаний и практических навыков, следовательно, пропуски отдельных лекций необходимо сразу наверстывать посредством самостоятельного изучения пропущенной темы и консультаций с преподавателем, ведущим занятия.

Рекомендации по подготовке к практическим (семинарским) занятиям

Студентам следует на каждое практическое занятие приходить с результатами выполненной домашней работы предыдущего семинара. Такое требование связано с тем, что сложные программы обсуждаются и выполняются несколько семинаров подряд, и для работы по теме текущего семинара используется результаты работы на предыдущем семинаре и соответствующей домашней работы.

Самостоятельная работа студентов включает в себя выполнение различного рода заданий, которые ориентированы на более глубокое усвоение материала изучаемой дисциплины и приобретение практических навыков по дисциплине Технологии продвинутой аналитики.

К выполнению заданий для самостоятельной работы предъявляются следующие требования: задания должны выполняться самостоятельно. Результатом выполнения задания является контрольная работа. Задание может быть выполнено как на компьютере студента (домашнем или в компьютерном классе), так и на компьютере преподавателя (домашнем или установленным в компьютерном классе).

Студентам следует:

—руководствоваться графиком самостоятельной работы, определенным РПД

–выполнять все плановые задания, выдаваемые преподавателем для самостоятельного выполнения

–разбирать на семинарах и консультациях ошибки в программах и прочие непонятные вопросы.

Форма промежуточной аттестации по дисциплине – *экзамен*.

Критерии оценивания знаний и умений, характеризующих степень сформированности компетенций:

- оценкой **«отлично»** оценивается полное освоение компетенций по данной дисциплине. Оценка выставляется при получении обучающимся 86 и более баллов. При этом он:

знает: текущий уровень инфраструктурных решений предприятия/организации; варианты технологического слоя архитектуры предприятия/организации.

умеет: анализировать текущий уровень инфраструктурных решений предприятия/организации; формировать и обосновывать варианты технологического слоя архитектуры предприятия/организации.

- оценкой **«хорошо»** оценивается освоение компетенций по данной дисциплине, однако в ответах допускаются неточности и незначительные ошибки. Оценка выставляется при получении обучающимся от 70 до 85 баллов. При этом он:

хорошо знает: текущий уровень инфраструктурных решений предприятия/организации; варианты технологического слоя архитектуры предприятия/организации.

хорошо умеет: анализировать текущий уровень инфраструктурных решений предприятия/организации; формировать и обосновывать варианты технологического слоя архитектуры предприятия/организации.

- оценкой **«удовлетворительно»** оценивается освоение компетенций по данной дисциплине, однако в ответах допускаются отдельные ошибки. Оценка выставляется при получении обучающимся от 50 до 69 баллов. При этом он:

плохо знает: текущий уровень инфраструктурных решений предприятия/организации; варианты технологического слоя архитектуры предприятия/организации.

плохо умеет: анализировать текущий уровень инфраструктурных решений предприятия/организации; формировать и обосновывать варианты технологического слоя архитектуры предприятия/организации.

- оценка **«неудовлетворительно»** выставляется в том случае, если компетенции не освоены, ответы содержат существенные ошибки и обучающимся получено менее 50 баллов. При этом он:

не знает: текущий уровень инфраструктурных решений предприятия/организации; варианты технологического слоя архитектуры предприятия/организации.

не умеет: анализировать текущий уровень инфраструктурных решений предприятия/организации; формировать и обосновывать варианты технологического слоя архитектуры предприятия/организации.

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

11.1.Комплект лицензионного программного обеспечения:

Пакет офисных программ;

Антивирус Kaspersky.

11.2.Современные профессиональные базы данных и информационные справочные системы:

Информационно-правовая система «Консультант Плюс»;

Информационно-правовая система «Гарант»;

Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>

11.3.Сертифицированные программные и аппаратные средства защиты информации:

Не предусмотрены.

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

Помещения для проведения лекций, семинарских занятий, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы.