

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
СТАВРОПОЛЬСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ**

УТВЕРЖДАЮ

Директор/Декан
института экономики, финансов и
управления в АПК
Гулько Юлия Александровна

«__» _____ 20__ г.

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ (ОЦЕНОЧНЫХ МАТЕРИАЛОВ)

Б1.О.13.03 Анализ больших данных

38.03.05 Бизнес-информатика

Электронный бизнес

бакалавр

очная

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Процесс изучения дисциплины направлен на формирование следующих компетенций ОП ВО и овладение следующими результатами обучения по дисциплине:

Код и наименование компетенции	Код и наименование индикатора достижения	Перечень планируемых результатов обучения по дисциплине
ОПК-2 Способен проводить исследование и анализ рынка информационных систем и информационно-коммуникационных технологий, выбирать рациональные решения для управления бизнесом;	ОПК-2.3 Выбирает рациональные решения для управления бизнесом на основе анализа рынка информационных систем и информационно-коммуникационных технологий	знает основные категории и классы информационные системы и информационно-коммуникационные технологии, используемых в бизнесе; принципы функционирования и архитектуру современных корпоративных ИТ-решений (ERP, CRM, BI-систем); методы и критерии анализа рынка ИС и ИКТ (стоимость владения, функциональность, масштабируемость, интегрируемость); подходы к сравнительной оценке программных продуктов и ИТ-платформ; основные тенденции цифровой трансформации бизнеса и развития ИТ-рынка.
		умеет анализировать рынок информационных систем и ИКТ с использованием формализованных критериев выбора; сравнивать альтернативные программные и технологические решения с учётом потребностей бизнеса; выявлять функциональные и экономические преимущества различных ИТ-решений; обосновывать выбор информационных систем для автоматизации бизнес-процессов; использовать аналитические методы (SWOT, ТСО-анализ, матрицы сравнения) для оценки ИТ-решений; формировать аргументированные рекомендации по внедрению ИТ-продуктов в организации.
		владеет навыками навыками системного анализа и оценки рынка информационных систем и ИКТ; методиками технико-экономического обоснования выбора ИТ-решений; инструментами сравнительного анализа программных продуктов и платформ; практикой подготовки аналитических отчетов и рекомендаций для принятия управленческих решений; компетенциями выбора оптимальных ИТ-решений для цифровой трансформации бизнес-процессов.

2. Перечень оценочных средств по дисциплине

№	Наименование раздела/темы	Семестр	Код индикаторов достижения компетенций	Оценочное средство проверки результатов достижения индикаторов компетенций
1.	1 раздел. Основной раздел			
1.1.	Тема	6		Собеседование
	Промежуточная аттестация			За

3. Оценочные средства (оценочные материалы)

Примерный перечень оценочных средств для текущего контроля успеваемости и промежуточной аттестации

№ п/п	Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде (Оценочные материалы)
Текущий контроль			
Для оценки знаний			
1	Собеседование	Средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной, и рассчитанное на выяснение объема знаний обучающегося по определенному разделу, теме, проблеме и т.п.	Вопросы по темам/разделам дисциплины
Для оценки умений			
Для оценки навыков			
Промежуточная аттестация			

2	Зачет	Средство контроля усвоения учебного материала практических и семинарских занятий, успешного прохождения практик и выполнения в процессе этих практик всех учебных поручений в соответствии с утвержденной программой с выставлением оценки в виде «зачтено», «незачтено».	Перечень вопросов к зачету
---	-------	---	----------------------------

4. Примерный фонд оценочных средств для проведения текущего контроля и промежуточной аттестации обучающихся по дисциплине (модулю) "Анализ больших данных"

Примерные оценочные материалы для текущего контроля успеваемости

Текущий контроль успеваемости по дисциплине «Анализ больших данных» осуществляется с целью проверки уровня усвоения теоретического материала, сформированности практических навыков и способности обучающихся применять методы анализа данных для решения прикладных задач бизнес-информатики.

Контроль проводится в течение семестра в форме устных опросов, тестирования, выполнения лабораторных работ, мини-кейсов, аналитических заданий и самостоятельных расчетно-аналитических работ.

1. Перечень примерных тестовых заданий

Что понимается под термином Big Data?

Какие характеристики относятся к 5V-модели?

В чем основное назначение Hadoop?

Что такое Data Lake?

Чем NoSQL отличается от SQL баз данных?

Какие задачи решает предобработка данных?

Что такое ETL-процесс?

Для чего используется кластеризация данных?

В чем отличие классификации от регрессии?

Какие задачи решает машинное обучение в бизнес-аналитике?

2. Перечень вопросов для устного опроса

Объясните роль больших данных в управлении бизнесом.

Опишите основные этапы анализа данных.

В чем заключается проблема «шума» в данных?

Какие существуют источники больших данных?

Объясните принцип работы распределённых систем обработки данных.

Какие существуют виды аналитики (описательная, диагностическая, предиктивная)?

Почему важна визуализация данных при принятии решений?

Как оценивается качество модели машинного обучения?

3. Практические задания для лабораторного контроля

Загрузить набор данных и провести первичный анализ (EDA).

Выполнить очистку данных от пропусков и дубликатов.

Провести нормализацию числовых признаков.

Построить корреляционную матрицу и интерпретировать результаты.

Реализовать простую линейную регрессию.

Построить модель классификации (логистическая регрессия).

Выполнить кластеризацию методом k-means.

Подготовить SQL-запрос для анализа выборки данных.

Построить графики распределения данных.

Сформировать аналитический вывод по результатам обработки данных.

4. Мини-кейсы (ситуационные задания)

Компания внедряет BI-систему. Обоснуйте выбор архитектуры хранения данных.

Интернет-магазин фиксирует снижение конверсии. Предложите подход анализа причин.

Банк планирует внедрение системы скоринга клиентов. Определите набор данных и метод анализа.

Логистическая компания хочет оптимизировать маршруты доставки с использованием аналитики данных.

Маркетинговый отдел планирует сегментацию клиентов — предложите методологию решения задачи.

5. Критерии оценивания текущего контроля

Оценивание результатов осуществляется по следующим критериям:

правильность и полнота ответов на теоретические вопросы;

владение базовой терминологией анализа больших данных;

корректность выполнения практических заданий;

способность интерпретировать результаты анализа;

обоснованность предлагаемых решений в кейсах;

самостоятельность выполнения заданий.

6. Формы текущего контроля

тестирование (письменное и/или компьютерное);

защита лабораторных работ;

устные опросы;

разбор кейсов;

проверка самостоятельных расчетно-аналитических заданий.

Примерные оценочные материалы для проведения промежуточной аттестации (зачет, экзамен) по итогам освоения дисциплины (модуля)

I. Теоретические задания (1–20)

Раскройте понятие больших данных и перечислите их основные характеристики (5V-модель).

Опишите роль анализа больших данных в современной бизнес-информатике.

Дайте определение распределённой системы обработки данных.

Объясните принцип работы экосистемы Hadoop.

В чем отличие Data Lake от Data Warehouse?

Перечислите основные типы данных в Big Data-системах.

Опишите особенности потоковых данных и их обработку.

Какие задачи решает предварительная обработка данных?

Что такое ETL-процесс и какие его этапы?

Опишите основные виды аналитики: описательная, диагностическая, предиктивная.

В чем заключается корреляционный анализ данных?

Опишите базовые принципы регрессионного анализа.

Что такое кластеризация и где она применяется?

Объясните отличие обучения с учителем и без учителя.

Какие метрики используются для оценки моделей машинного обучения?

Опишите архитектуру Apache Spark.

Что такое NoSQL базы данных и чем они отличаются от реляционных?

Перечислите основные BI-инструменты и их назначение.

В чем заключается концепция data-driven управления бизнесом?

Опишите роль визуализации данных в аналитике.

II. Аналитические задания (21–40)

Проанализируйте преимущества использования Big Data в ритейле.

Сравните Hadoop и Spark по ключевым характеристикам.

Обоснуйте выбор Data Lake или Data Warehouse для компании.

Разработайте критерии выбора BI-системы для предприятия.

Проведите SWOT-анализ внедрения аналитики больших данных в банк.

Опишите риски при обработке неструктурированных данных.
Предложите метрики оценки эффективности аналитической системы.
Сравните SQL и NoSQL подходы к хранению данных.
Обоснуйте необходимость очистки данных перед анализом.
Проанализируйте применение машинного обучения в маркетинге.
Опишите сценарий использования кластеризации клиентов.
Сравните методы классификации и регрессии.
Объясните влияние качества данных на точность модели.
Предложите архитектуру аналитической системы для интернет-магазина.
Проанализируйте использование потоковой аналитики в логистике.
Опишите преимущества облачных платформ для Big Data.
Сравните подходы batch и stream processing.
Разработайте модель анализа поведения пользователей сайта.
Обоснуйте выбор алгоритма для задачи прогнозирования спроса.
Проанализируйте роль API в сборе данных.

III. Практико-ориентированные задания (41–60)

Выполните загрузку CSV-данных и проведите первичный анализ в Python.
Очистите набор данных от пропусков и дубликатов.
Проведите нормализацию числовых признаков.
Выполните кодирование категориальных переменных.
Постройте графики распределения данных.
Постройте корреляционную матрицу и интерпретируйте результаты.
Реализуйте линейную регрессию на учебном наборе данных.
Оцените качество модели регрессии (MAE, RMSE).
Постройте модель логистической регрессии для классификации.
Реализуйте дерево решений и интерпретируйте результаты.
Выполните кластеризацию методом k-means.
Определите оптимальное число кластеров методом локтя.
Проведите анализ временного ряда (тренд, сезонность).
Создайте SQL-запросы для анализа продаж.
Выполните JOIN нескольких таблиц данных.
Постройте ETL-процесс обработки данных.
Реализуйте базовую обработку данных в Apache Spark.
Создайте простой BI-дашборд (Power BI / Tableau).
Постройте модель прогнозирования спроса на товар.
Разработайте мини-проект анализа данных полного цикла (загрузка → обработка → анализ → выводы).

Темы письменных работ (эссе, рефераты, курсовые работы и др.)

Роль анализа больших данных в цифровой трансформации бизнеса.
Архитектуры систем обработки больших данных: Hadoop и Spark.
Сравнение Data Lake и Data Warehouse в корпоративных системах.
Проблемы качества данных и методы их решения в Big Data-средах.
Источники больших данных в современных информационных системах.
Потоковая обработка данных: технологии и бизнес-применение.
ETL и ELT-подходы: сравнительный анализ и практическое применение.
Использование NoSQL баз данных в аналитических системах.
Аналитика больших данных в электронной коммерции.
Машинное обучение как инструмент анализа больших данных.
Кластеризация клиентов: методы и бизнес-применение.
Прогнозная аналитика в управлении спросом.
Анализ пользовательского поведения на веб-платформах.
Визуализация данных как инструмент принятия решений.
BI-системы в управлении предприятием.
Использование Python в анализе больших данных.
SQL как инструмент аналитической обработки данных.
Облачные технологии в обработке больших данных.
Роль API в сборе и интеграции данных.
Применение регрессионного анализа в бизнес-аналитике.
Методы оценки качества моделей машинного обучения.
Проблемы масштабируемости систем обработки данных.
Анализ потоковых данных в реальном времени.
Big Data в банковском секторе: риски и возможности.
Аналитика данных в маркетинге и рекламе.
Предобработка данных: методы очистки и трансформации.
Интеграция разнородных данных в аналитических системах.
Использование кластерного анализа для сегментации рынка.
Цифровая экономика и роль больших данных в бизнес-решениях.
Проектирование аналитической системы для предприятия (на примере отрасли).