

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
ПО ДИСЦИПЛИНЕ**
Прикладной анализ данных

Код модуля
1161465(1)

Модуль
Прикладной анализ данных

Екатеринбург

Оценочные материалы составлены автором(ами):

№ п/п	Фамилия, имя, отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Комоцкий Евгений Игоревич	без ученой степени, без ученого звания	Ассистент	аналитики больших данных и методов видеоанализа
2	Федотов Олег Васильевич	без ученой степени, без ученого звания	Старший преподаватель	интеллектуальных информационных технологий

Согласовано:

Управление образовательных программ

Е.А. Смирнова

Авторы:

- **Комоцкий Евгений Игоревич, Ассистент, аналитики больших данных и методов видеоанализа**
- **Федотов Олег Васильевич, Старший преподаватель, интеллектуальных информационных технологий**

1. СТРУКТУРА И ОБЪЕМ ДИСЦИПЛИНЫ Прикладной анализ данных

1.	Объем дисциплины в зачетных единицах	3	
2.	Виды аудиторных занятий	Лекции Практические/семинарские занятия Лабораторные занятия	
3.	Промежуточная аттестация	Экзамен	
4.	Текущая аттестация	Контрольная работа	1
		Домашняя работа	1

2. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ (ИНДИКАТОРЫ) ПО ДИСЦИПЛИНЕ МОДУЛЯ Прикладной анализ данных

Индикатор – это признак / сигнал/ маркер, который показывает, на каком уровне обучающийся должен освоить результаты обучения и их предъявление должно подтвердить факт освоения предметного содержания данной дисциплины, указанного в табл. 1.3 РПМ-РПД.

Таблица 1

Код и наименование компетенции	Планируемые результаты обучения (индикаторы)	Контрольно-оценочные средства для оценивания достижения результата обучения по дисциплине
1	2	3
ПК-6 -Способен использовать цифровые решения для работы с потоками ESG-данных и составления отчетности организации	3-1 - Описать современные информационно-коммуникативные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач 3-2 - Описывать современные техники и методики сбора данных для решения управленческих и исследовательских задач 3-4 - Сделать обзор основных методов моделирования и математического анализа,	Домашняя работа Контрольная работа Лабораторные занятия Лекции Практические/семинарские занятия Экзамен

	<p>применимых для формализации и решения задач в области техносферной безопасности</p> <p>З-5 - Описать технологии хранения и обработки больших данных в организации: базы данных, хранилища данных, распределенную и параллельную обработку данных, вычисления в оперативной памяти</p> <p>П-2 - Создавать отчеты, рефераты, статьи, оформленные в соответствии с предъявляемыми требованиями с использованием интеллектуальных информационно-аналитических систем</p> <p>У-2 - Выбирать оптимальные методы обработки и анализа данных, в том числе используя интеллектуальные информационно-аналитические системы при решении управленческих и исследовательских задач</p> <p>У-6 - Выбирать оптимальные современные методы и инструментальные средства анализа больших данных, методы интерпретации и визуализации больших данных</p>	
--	--	--

3. ПРОЦЕДУРЫ КОНТРОЛЯ И ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ В РАМКАХ ТЕКУЩЕЙ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ МОДУЛЯ В БАЛЬНО-РЕЙТИНГОВОЙ СИСТЕМЕ (ТЕХНОЛОГИЧЕСКАЯ КАРТА БРС)

3.1. Процедуры текущей и промежуточной аттестации по дисциплине

1. Лекции: коэффициент значимости совокупных результатов лекционных занятий – 0.4		
Текущая аттестация на лекциях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>контрольная работа</i>	6	60
<i>активность на занятиях</i>	8	40
Весовой коэффициент значимости результатов текущей аттестации по лекциям – 0.5		

Промежуточная аттестация по лекциям – экзамен Весовой коэффициент значимости результатов промежуточной аттестации по лекциям – 0.5		
2. Практические/семинарские занятия: коэффициент значимости совокупных результатов практических/семинарских занятий – 0.3		
Текущая аттестация на практических/семинарских занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>домашняя работа</i>	14	60
<i>активность на занятиях</i>	16	40
Весовой коэффициент значимости результатов текущей аттестации по практическим/семинарским занятиям– 1		
Промежуточная аттестация по практическим/семинарским занятиям– нет Весовой коэффициент значимости результатов промежуточной аттестации по практическим/семинарским занятиям– не предусмотрено		
3. Лабораторные занятия: коэффициент значимости совокупных результатов лабораторных занятий –0.3		
Текущая аттестация на лабораторных занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
<i>отчет по лабораторным работам</i>	16	100
Весовой коэффициент значимости результатов текущей аттестации по лабораторным занятиям - 1		
Промежуточная аттестация по лабораторным занятиям – нет Весовой коэффициент значимости результатов промежуточной аттестации по лабораторным занятиям – не предусмотрено		
4. Онлайн-занятия: коэффициент значимости совокупных результатов онлайн-занятий –не предусмотрено		
Текущая аттестация на онлайн-занятиях	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
Весовой коэффициент значимости результатов текущей аттестации по онлайн-занятиям - не предусмотрено		
Промежуточная аттестация по онлайн-занятиям – нет Весовой коэффициент значимости результатов промежуточной аттестации по онлайн-занятиям – не предусмотрено		

3.2. Процедуры текущей и промежуточной аттестации курсовой работы/проекта

Текущая аттестация выполнения курсовой работы/проекта	Сроки – семестр, учебная неделя	Максимальная оценка в баллах
Весовой коэффициент текущей аттестации выполнения курсовой работы/проекта– не предусмотрено		
Весовой коэффициент промежуточной аттестации выполнения курсовой работы/проекта– защиты – не предусмотрено		

4. КРИТЕРИИ И УРОВНИ ОЦЕНИВАНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ МОДУЛЯ

4.1. В рамках БРС применяются утвержденные на кафедре/институте критерии (признаки) оценивания достижений студентов по дисциплине модуля (табл. 4) в рамках контрольно-оценочных мероприятий на соответствие указанным в табл.1 результатам обучения (индикаторам).

Таблица 4

Критерии оценивания учебных достижений обучающихся

Результаты обучения	Критерии оценивания учебных достижений, обучающихся на соответствие результатам обучения/индикаторам
Знания	Студент демонстрирует знания и понимание в области изучения на уровне указанных индикаторов и необходимые для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.
Умения	Студент может применять свои знания и понимание в контекстах, представленных в оценочных заданиях, демонстрирует освоение умений на уровне указанных индикаторов и необходимых для продолжения обучения и/или выполнения трудовых функций и действий, связанных с профессиональной деятельностью.
Опыт /владение	Студент демонстрирует опыт в области изучения на уровне указанных индикаторов.
Другие результаты	Студент демонстрирует ответственность в освоении результатов обучения на уровне запланированных индикаторов. Студент способен выносить суждения, делать оценки и формулировать выводы в области изучения. Студент может сообщать преподавателю и коллегам своего уровня собственное понимание и умения в области изучения.

4.2 Для оценивания уровня выполнения критериев (уровня достижений обучающихся при проведении контрольно-оценочных мероприятий по дисциплине модуля) используется универсальная шкала (табл. 5).

Таблица 5

Шкала оценивания достижения результатов обучения (индикаторов) по уровням

Характеристика уровней достижения результатов обучения (индикаторов)				
№ п/п	Содержание уровня выполнения критерия оценивания результатов обучения (выполненное оценочное задание)	Шкала оценивания		
		Традиционная характеристика уровня		Качественная характеристика уровня
1.	Результаты обучения (индикаторы) достигнуты в полном объеме, замечаний нет	Отлично (80-100 баллов)	Зачтено	Высокий (В)

2.	Результаты обучения (индикаторы) в целом достигнуты, имеются замечания, которые не требуют обязательного устранения	Хорошо (60-79 баллов)		Средний (С)
3.	Результаты обучения (индикаторы) достигнуты не в полной мере, есть замечания	Удовлетворительно (40-59 баллов)		Пороговый (П)
4.	Освоение результатов обучения не соответствует индикаторам, имеются существенные ошибки и замечания, требуется доработка	Неудовлетворительно (менее 40 баллов)	Не зачтено	Недостаточный (Н)
5.	Результат обучения не достигнут, задание не выполнено	Недостаточно свидетельств для оценивания		Нет результата

5. СОДЕРЖАНИЕ КОНТРОЛЬНО-ОЦЕНОЧНЫХ МЕРОПРИЯТИЙ ПО ДИСЦИПЛИНЕ МОДУЛЯ

5.1. Описание аудиторных контрольно-оценочных мероприятий по дисциплине модуля

5.1.1. Лекции

Самостоятельное изучение теоретического материала по темам/разделам лекций в соответствии с содержанием дисциплины (п. 1.2. РПД)

5.1.2. Практические/семинарские занятия

Примерный перечень тем

1. Применение методов анализа данных с использованием программ Vortex, SPSS, Statistica и языков программирования Python, R для решения практических задач.

2. Применение методов обработки и анализа данных для решения различных практических задач.

3. Библиотека Pandas: работа с данными и визуализация.

4. Jupyter Notebook и Google Colab: интерактивная работа с данными.

5. Применения машинного обучения в бизнесе и науке.

6. Интерфейс платформы Kaggle и основные функции.

7. Кросс-валидация в оценке качества алгоритмов.

8. Уменьшение разброса с помощью усреднения: методы и примеры применения.

Примерные задания

Реализация алгоритма машинного обучения с помощью Google Colab.

Описание: Реализовать алгоритм машинного обучения (например, линейную регрессию или классификацию) для решения конкретной задачи с использованием интерактивных инструментов Google Colab. Сравнить результаты работы алгоритма с различными методами предварительной обработки данных.

Загрузить набор данных для обучения и тестирования модели из выбранного источника (например, UCI Machine Learning Repository или Kaggle) в Google Colab с использованием библиотеки pandas.

Разбить набор данных на обучающую и тестовую выборки с использованием функций

train_test_split библиотеки sklearn.

Реализовать и настроить алгоритм машинного обучения для решения конкретной задачи (например, логистической регрессии или классификации).

Провести кросс-валидацию модели на обучающей выборке с использованием K-кратной перекрестной проверки.

Выполнить предварительную обработку данных: удалить пропущенные значения, преобразовать категориальные переменные, нормализовать числовые переменные.

Оценить качество работы модели с помощью метрик точности (accuracy), полноты (recall), F1-меры и ROC-кривой.

Сравнить результаты работы модели с различными методами предварительной обработки данных и интерпретировать полученные результаты.

Результат: Модель машинного обучения, обученная на выбранном наборе данных и отчет о сравнении результатов работы модели с разными методами предварительной обработки данных, включающий выводы о наилучшем методе предварительной обработки.

LMS-платформа – не предусмотрена

5.1.3. Лабораторные занятия

Примерный перечень тем

1. Анализ данных с использованием программы Vortex.
2. Программирование на Python для анализа данных.
3. Интерактивная работа с данными в Google Colab.
4. Применение машинного обучения в бизнес-анализе.
5. Кросс-валидация качества алгоритмов машинного обучения.

LMS-платформа – не предусмотрена

5.2. Описание внеаудиторных контрольно-оценочных мероприятий и средств текущего контроля по дисциплине модуля

Разноуровневое (дифференцированное) обучение.

Базовый

5.2.1. Контрольная работа

Примерный перечень тем

1. Применение интерактивных инструментов Google Colab для анализа данных и визуализации результатов.

Примерные задания

Анализ и визуализация данных с помощью Google Colab.

Описание: Проанализировать и визуализировать набор данных с помощью интерактивных инструментов в Google Colab. Разработать и применить различные методы обработки и визуализации данных для лучшего понимания и интерпретации данных.

Загрузить набор данных из открытого источника (например, Kaggle или UCI Machine Learning Repository) в Google Colab и обработать его для анализа.

Загрузить набор данных из выбранного источника в Google Colab с использованием библиотеки pandas.

Предварительно обработать данные: удалить пропущенные значения, преобразовать категориальные переменные, нормализовать числовые переменные.

Осуществить визуализацию данных: построить гистограммы, точечные диаграммы, коробки с усами, диаграммы размаха для каждой числовой переменной в наборе данных.

Построить корреляционную матрицу для всех числовых переменных в наборе данных, чтобы выявить наличие линейных связей между переменными.

Построить деревья решений для каждой категориальной переменной в наборе данных для выявления основных паттернов и зависимостей в данных.

Интерпретировать полученные результаты и представить выводы о структуре и свойствах анализируемого набора данных.

Результат: Отчет о проведенном анализе данных, содержащий гистограммы, графики корреляции, деревья решений и выводы об основных характеристиках и зависимостях в анализируемом наборе данных.

LMS-платформа – не предусмотрена

5.2.2. Домашняя работа

Примерный перечень тем

1. Сравнительный анализ различных методов предварительной обработки данных для задач машинного обучения.

2. Исследование эффективности использования различных алгоритмов машинного обучения для решения конкретных практических задач.

3. Реализация и сравнение различных методов оценки качества моделей машинного обучения.

Примерные задания

Задание 1. Сравнение алгоритмов предварительной обработки данных.

Описание: Провести сравнительный анализ эффективности различных методов предварительной обработки для задач машинного обучения и выбрать наиболее подходящий метод для конкретного набора данных.

Загрузить набор данных, который будет использоваться для анализа (например, из UCI Machine Learning Repository).

Реализовать и настроить несколько алгоритмов машинного обучения (классификацию и регрессию) на данном наборе данных.

Провести предварительную обработку данных с использованием различных методов: удаление пропущенных значений, нормализация, стандартизация, масштабирование и т.д.

Оценить качество работы каждого алгоритма машинного обучения после предварительной обработки с использованием метрик качества (accuracy, precision, recall, F1-мера и т.п.).

Сравнить полученные результаты и выбрать метод предварительной обработки, который обеспечивает наилучшие результаты для данного набора данных.

Результат: Отчет, содержащий результаты сравнительного анализа различных методов предварительной обработки и выводы о наиболее эффективном методе для конкретного набора данных.

Задание 2. Оценка качества моделей машинного обучения.

Описание: Реализовать различные методы оценки качества моделей машинного обучения и сравнить их эффективность.

Загрузить набор данных для оценки качества модели.

Реализовать и настроить модели машинного обучения на данном наборе данных (классификацию, регрессию).

Оценить качество работы каждой модели с использованием различных метрик оценки качества (accuracy, F1-мера, ROC-AUC и т.д.).

Сравнить полученные результаты и сделать выводы о наиболее эффективной метрике оценки качества для данного набора данных и конкретной задачи машинного обучения.

Результат. Отчет, содержащий результаты оценки качества моделей и выводы о наиболее подходящей метрике качества для конкретной задачи машинного обучения.

LMS-платформа – не предусмотрена

5.3. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине модуля

5.3.1. Экзамен

Список примерных вопросов

1. Какие программы используются для анализа данных?
2. Какие языки программирования применяются для анализа данных?
3. В чем разница между Jupyter Notebook и Google Collab?
4. В каких сферах применяется машинное обучение?
5. Как работает платформа Kaggle?
6. Что такое кросс-валидация и зачем она нужна?
7. Как уменьшить разброс с помощью усреднения?
8. Какие методы анализа данных вы знаете?
9. Какие инструменты для обработки данных вы используете?
10. В чем заключается работа с данными с помощью Python и R?
11. Что такое визуализация данных и как ее можно провести?
12. Какие алгоритмы машинного обучения вы знаете?
13. Зачем нужно применять машинное обучение в бизнесе?
14. Какие функции есть у платформы Kaggle?
15. Какие метрики используются для оценки качества алгоритмов?
16. В чем заключаются преимущества использования Google Colab?
17. Какие задачи можно решить с помощью анализа данных?
18. Какие примеры применения машинного обучения вы можете привести?
19. Как можно улучшить качество алгоритмов машинного обучения?

LMS-платформа – не предусмотрена

5.4 Содержание контрольно-оценочных мероприятий по направлениям воспитательной деятельности

Направления воспитательной деятельности сопрягаются со всеми результатами обучения компетенций по образовательной программе, их освоение обеспечивается содержанием всех дисциплин модулей.