

Аннотация рабочей программы дисциплины

Б1.В.04 Анализ и визуализация данных.

Код и наименование программы подготовки, профиля: 38.03.01 «Экономика», Финансы и кредит».

Квалификация (степень) выпускника: бакалавр.

Форма обучения: очная

Цель освоения дисциплины: сформировать компетенцию в области применения информационных технологий для решения прикладных финансово-экономических задач.

Содержание дисциплины

Тема 1. Основы анализа данных и машинного обучения

Тема осваивается с применением дистанционных образовательных технологий (размещается на платформе СДО РАНХиГС)

Раздел 1. Введение в теорию анализа данных.

Основы обработки данных. Основы алгоритмизации. Существующие наборы данных, визуализация модели классификации. Объекты и признаки. Типы шкал. Показатели вариации. Линейные и нелинейные модели регрессии.

Раздел 2. Практика анализа данных и машинного обучения

Понижение размерности. Модель алгоритмов машинного обучения.

Классификаторы: виды и принципы построения. Методы оценки точности решений. Визуализация.

Тема 2. Подготовка данных к анализу

Очистка и предобработка данных.

Оценка качества данных. Фильтрация данных. Обработка дубликатов и противоречий. Выявление аномальных значений. Восстановление пропущенных значений. Сокращение размерности. Сэмплинг.

Трансформация данных.

Группировка данных. Слияние данных. Квантование. Нормализация и кодирование.

### Тема 3. Визуализация данных

Введение в визуализацию. Визуализаторы общего назначения. OLAP-анализ. Визуализаторы для оценки качества моделей. Визуализаторы, применяемые для интерпретации результатов анализа.

Визуализация данных средствами MS Power BI.

Семейство MS Power BI. Конструктор кубов и аналитических представлений.

Совместная работа с представлениями – серверные сервисы Power BI.

Решение задач объединения данных из разных источников, вычисляемые поля и Меры, элементы языка DAX, построение различных визуальных представлений по данным из учетных систем и CRM-систем.

### Тема 4. Классификация и регрессия (Статистические методы)

Введение в классификацию и регрессию. Множественная линейная регрессия. Регрессия с категориальными входными переменными. Методы отбора переменных в регрессионные модели. Ограничения применимости регрессионных моделей. Нелинейные регрессионные модели. Основы логистической регрессии. Множественная логистическая регрессия. Оценка эффективности и сравнение моделей. ROC-анализ. Оценка ошибки модели. Издержки ошибочной классификации. LIFT- и PROFIT-кривые.

### Тема 5. Классификация и регрессия (Машинное обучение)

Введение в деревья решений. Алгоритмы построения деревьев решений. Упрощение деревьев решений. Введение в нейронные сети. Принципы построения нейронных сетей. Процесс обучения нейронной сети. Обучение в условиях несбалансированности классов.

### Тема 6. Кластеризация

Введение в кластеризацию. Меры близости, используемые в алгоритмах кластеризации. Методы кластерного анализа. Иерархическая кластеризация. Алгоритм кластеризации k-means. Сети Кохонена. Карты Кохонена.

### Тема 7. Анализ и прогнозирование временных рядов

Введение в прогнозирование. Временной ряд и его компоненты. Модели прогнозирования. Исследование наборов данных и выбор метода

прогнозирования. Ошибки прогноза. Оценка адекватности выбранного метода прогнозирования. Методы сглаживания и скользящие средние. Декомпозиция временного ряда. Регрессионный анализ временных рядов.

### **Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине**

Формы и методы текущего контроля успеваемости и промежуточной аттестации. В ходе реализации дисциплины «Дифференциальные уравнения» используются следующие методы текущего контроля успеваемости обучающихся:

При проведении занятий лекционного типа: опрос.

При проведении занятий семинарского типа: решение задач, контрольная работа, тест.

**Промежуточная аттестация проводится в форме зачета с оценкой.**

## **2. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине**

### 2.1. Основная литература.

1. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2023. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/511020>
2. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2023. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/511121>

### 2.2. Дополнительная литература.

1. Чубукова, И. А. Data Mining: учебное пособие / И. А. Чубукова. — 3-е изд. — Москва, Саратов: Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 469 с. — ISBN 978-5-4497-0289-0. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <http://www.iprbookshop.ru/89404.html>

2. Миркин, Б. Г. Введение в анализ данных: учебник и практикум / Б. Г. Миркин. — Москв : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450262>
3. Подкорытова, О. А. Анализ временных рядов: учебное пособие для вузов / О. А. Подкорытова, М. В. Соколов. — 2-е изд., перераб. и доп. — Москва: Издательство Юрайт, 2020. — 267 с. — (Высшее образование). — ISBN 978-5-534-02556-9. — Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/450587>