

Аннотация дисциплины Б.1.1.25 Дисциплина. Машинное обучение и анализ данных

Дисциплина "Машинное обучение и анализ данных" изучается обучающимися по основной профессиональной образовательной программе "Экономическая кибернетика" направления подготовки "09.03.02 Информационные системы и технологии".

Дисциплина изучается в 6, 7 семестре. Общая трудоемкость дисциплины составляет 192/7 часов/з.ед. Самостоятельная работа заключается в выполнении работ, указанных в разделе 4.

В ходе изучения дисциплины осуществляется текущий контроль в форме технологии рейтингового контроля в соответствии с технологической карты дисциплины, размещенной на электронном курсе, а также промежуточный контроль в форме зачет, экзамен.

Целью изучения дисциплины является формирование следующих компетенций:

1. ОПК-1 Способен применять естественнонаучные и общетехнические знания, методы математического анализа и моделирования, теоретического и экспериментального исследования в профессиональной деятельности
2. ОПК-6 Способен разрабатывать алгоритмы и программ, пригодные для практического применения в области информационных систем и технологий

В ходе изучения дисциплины последовательно рассматриваются темы:

1. Основные концепции машинного обучения:
 1. Моделирование.
 2. Обучение с учителем.
 3. Обучение без учителя.
 4. Обучение с подкреплением.
 5. Кластеризация.
 6. Классификация.
 7. Регрессия.
 8. Оценка и выбор моделей.
2. Математические основы машинного обучения:
 1. Линейная алгебра.
 2. Теория вероятностей и статистика.
 3. Оптимизация и методы оптимизации.
 4. Теория информации.
 5. Методы регуляризации.
3. Введение в анализ данных:
 1. Основы статистики и вероятности.
 2. Методы визуализации данных.
 3. Работа с данными с использованием Python и библиотеки pandas.
4. Статистический анализ данных:
 - 1) Сбор данных
 - 2) Статистический анализ (информационные критерии, проверка нормальности, оценка дисперсии, корреляционный анализ)
 - 3) Предварительная обработка (удаление выбросов, заполнение пропущенных значений)
 - 4) Преобразование данных (например, нормализация, кодирование категориальных переменных)
 - 5) Разбиение данных на обучающую и тестовую выборки
5. Визуализация данных:
 1. Основы визуализации данных.
 2. Инструменты для визуализации данных (например, matplotlib, seaborn, Plotly).
 3. Визуализация данных в трехмерном пространстве.

4. Визуализация временных рядов.
5. Визуализация геопространственных данных.
6. Понимание и интерпретация визуализаций данных.
6. Обучение с учителем, без учителя, с подкреплением на языке программирования Python для анализа экономических показателей.
Обучение с учителем:
 1. Введение в язык программирования Python для экономического анализа.
 2. Построение регрессионных моделей для анализа экономических данных.
 3. Анализ временных рядов и прогнозирование экономических показателей.
Обучение без учителя:
 1. Анализ факторной структуры и кластерный анализ данных.
 2. Понижение размерности данных с использованием методов PCA и t-SNE.
Обучение с подкреплением:
 1. Построение и отладка моделей машинного обучения для анализа экономических показателей.
 2. Оптимизация гиперпараметров моделей.
 3. Использование алгоритмов усиления моделей для улучшения точности предсказаний.
7. Интеграция машинного обучения для решения задач экономики:
 1. Анализ финансовых рынков с помощью алгоритмов машинного обучения для определения трендов, паттернов и прогнозирования цен.
 2. Автоматизация финансового анализа для принятия решений, например, для определения оптимального портфеля инвестиций или для управления рисками.
 3. Оптимизация процессов ведения бизнеса с использованием методов машинного обучения, например, в сферах логистики, управления запасами и ценообразования.
 4. Применение методов машинного обучения для анализа и прогнозирования потребительского спроса и поведения потребителей.
8. Предобработка данных для машинного обучения:
 1. Масштабирование признаков: Нормализация или стандартизация признаков для обеспечения одинаковой важности каждого из них.
 2. Создание новых признаков: Извлечение новых признаков из существующих данных для улучшения производительности модели.
9. Машинное обучение и анализ данных с целью анализа деятельности организаций:
 1. Прогнозирование продаж и потребительского спроса.
 2. Анализ клиентского поведения и предсказание оттока.
 3. Оптимизация производственных процессов.
 4. Анализ финансовых показателей и прогнозирование финансовых результатов.
 5. Рекомендательные системы для персонализированного маркетинга и продаж.

Основными стратегическими образовательными технологиями являются: лекционные занятия, практические и лабораторные занятия, процедуры самообучения.

В рамках указанных технологий применяются тактические образовательные технологии: задания, классическая лекция, лекция с элементами мозгового штурма.