

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Ливанов Дмитрий Викторович
Должность: Ректор
Дата подписания: 27.10.2023 10:11:56
Уникальный программный ключ:
c6d909c49c1d2034fa3a0156c4eaa51e7232a3a2

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Анализ естественного языка

Цель дисциплины:

- приобретение навыков решения практических заданий с помощью библиотек и ресурсов для обработки естественных языков.

Задачи дисциплины:

- сформировать навыки обработки неструктурированного текста на естественном языке;
- сформировать навыки работы с современными математическими методами и компьютерными алгоритмами для решения этих задач.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- методы анализа проблемных ситуаций как систем;
- методы обмена деловой информацией в устной и письменной формах на государственном языке Российской Федерации и не менее чем на одном иностранном языке;
- о современном состоянии исследований в рамках тематической области своей профессиональной деятельности;
- методы анализа задач, а также методы планирования их решения;
- методы и средства разработки программного обеспечения, методы управления проектами разработки программного обеспечения, способы организации проектных данных, нормативно-технические документы (стандарты и регламенты) по разработке программных средств и проектов.

уметь:

- осуществлять поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации;
- представлять результаты академической и профессиональной деятельности на различных научных мероприятиях, включая международные;

- оценивать актуальность исследований в области информатики и вычислительной техники и их практическую значимость;
- анализировать и вычислять решения задач, понимать и учитывать на практике границы применимости получаемых решений;
- выбирать средства разработки, оценивать сложность проектов, планировать ресурсы, контролировать сроки выполнения и оценивать качество полученного результата.

владеть:

- навыками разработки стратегий достижения поставленной цели как последовательности шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности;
- навыками использования современных средств информационно-коммуникационных технологий для академического и профессионального взаимодействия;
- профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации;
- навыками самостоятельного приобретения, развития и применения математических, естественнонаучных, социально-экономических и профессиональных знаний для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте;
- методами разработки технического задания, составления планов, распределения задач, тестирования и оценки качества программных средств.

Темы и разделы курса:

1. Основные NLP задачи и их использование в бизнесе

Классификация текстов. Классификация отдельных слов в тексте. Машинный перевод. Поиск ответа на вопрос. Задачи генерации текста. Этические проблемы использования NLP моделей.

2. Векторные представления текстов

Статистические методы. Word2Vec.

3. Рекуррентные нейронные сети

Рекуррентные нейронные сети. RNN для задачи текстовой классификации.

4. Transfer learning in NLP

Архитектура Transformer. Transfer learning.

5. Задача Token Classification

Задача NER.

6. Задача Question answering

Решение задачи Question answering

7. Transformers в продакшене

Введение. Дистилляция. Квантизация.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Английский для профессиональных коммуникаций

Цель дисциплины:

- целью данного курса является развитие практических навыков использования английского языка для профессионального общения в сфере бизнеса и информационных технологий.

Задачи дисциплины:

- обеспечить усвоение основ бизнес и IT терминологии на английском языке;
- формировать навыки использования наиболее употребительных грамматических конструкций в типичных ситуациях профессионального общения;
- формировать умение поддерживать разговор профессиональной направленности на английском языке;
- формировать умение составлять деловую документацию на английском языке;
- формировать умение презентовать собственный продукт /проект на английском языке.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основную профессиональную терминологию на английском языке по выбранному направлению;
- наиболее употребительные грамматические конструкции для использования в типичных ситуациях профессионального общения;
- устойчивые выражения для поддержания разговора профессиональной направленности на английском языке;
- структурные и стилистические особенности делового письма;
- структурные и стилистические особенности презентации собственного продукта/проекта на английском языке.

уметь:

- употреблять основную профессиональную терминологию в типичных ситуациях профессионального общения на английском языке;
- употреблять в речи соответствующие ситуации общения грамматические конструкции;
- поддерживать разговор профессиональной направленности на английском языке;
- читать аутентичные тексты по профильной тематике с пониманием общей идеи, с извлечением информации и с детальным пониманием прочитанного;
- описать производственный процесс, разрабатываемый продукт, написать CV, техническое задание, бизнес-план, инструкцию, повестку дня деловой встречи, протокол заседания с - учетом норм речевого этикета стран изучаемого языка;
- создавать презентацию собственного продукта и/или проекта на английском языке.

владеть:

- межкультурной профессионально ориентированной коммуникативной компетенцией в разных видах речевой деятельности;
- различными коммуникативными стратегиями:
- учебными стратегиями для организации своей учебной деятельности;
- стратегиями рефлексии и самооценки в целях самосовершенствования личных качеств и достижений;
- стратегиями восприятия, анализа, создания устных и письменных текстов разных типов;
- интернет-технологиями для выбора оптимального режима получения информации;
- компенсаторными умениями, помогающими преодолеть «сбои» в коммуникации, вызванные объективными и субъективными, социокультурными причинами.

Темы и разделы курса:

1. Работа в международной компании

Написание CV. Поиск вакансии онлайн. Подача заявления о приеме на работу. Подготовка к собеседованию. Написание мотивационного письма. «Неудобные» вопросы в ходе стрессового собеседования. Работа в команде, распределение ролей в команде. Преимущества и недостатки работы в команде. Лидерство. Взаимоотношения в команде.

2. Деловые контакты в профессиональной области

Деловые контакты в определенной профессиональной деятельности. Коммуникация с коллегами, партнерами и клиентами компании. Особенности межкультурной коммуникации в международной компании. Основы эффективной коммуникации с клиентами компании. Удержание клиентов и расширение клиентской базы. Переговоры с партнерами. Совещания внутри компании.

3. Разработка продукта и его продвижение

Описывание этапов разработки программного продукта. Объяснение понятия риска и неопределенности, анализировать проектные риски, предлагать меры, направленные на смягчение рисков. Описывание программного продукта; чтение и обсуждение концепции PPPP: product, price, place, promotion. Участие в дискуссии об эффективных способах продвижения нового продукта, включая размещение рекламной информации о продукте в социальных сетях. Сообщение о SWOT- анализе как инструменте маркетинга.

4. Презентация продукта

Презентация нового продукта. Основные характеристики эффективной презентации. Язык презентации. Структура презентации: введение, основная часть, выводы. Создание эффективных слайдов. Управление презентацией. Взаимодействие с аудиторией. Вопросы и ответы. Оценка эффективности презентации. Pitch-презентация продукта для потенциального инвестора.

5. Компания

Создание компании. Организационно-правовые формы (виды) компаний. Разработка бизнес плана. Миссия компании. Анализ рынка. Анализ конкурентной среды. Операционный, маркетинговый и финансовый планы. Start up бизнес в IT сфере

6. Финансы

Финансы компании. Анализ статистических данных деятельности компании. Источники финансирования деятельности компании: собственные и заемные. Финансовое планирование деятельности компании. Фондовый рынок. Основные участники фондового рынка. Виды ценных бумаг: акция, облигация, фьючерс, опцион. Стартовый капитал. Первичные (IPO) и вторичные торги. Инвестиции. Стратегии инвестирования. Венчурный бизнес.

7. Производственная деятельность

Производственная деятельность компании. Описание производственного процесса. Инновационные разработки в производстве. Концепция бережливого производства. Закупки материалов и оборудования. Описание оборудования.

8. Информационные технологии

Компьютеры в современном мире. Основные компоненты компьютера. Устройства ввода и вывода данных. Запоминающие устройства. Виды программного обеспечения. Программирование. Базы данных. Интернет. Компьютеры будущего. Новые технологии: искусственный интеллект, нейросети, интернет вещей, блокчейн. Кибербезопасность

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Глубокое обучение в науках о данных

Цель дисциплины:

- сформировать у студентов навыки решения прикладных задач при помощи глубоких нейронных сетей.

Задачи дисциплины:

- изучить модель искусственного нейрона и искусственной нейронной сети;
- изучить алгоритмы обучения нейронных сетей;
- изучить популярные в настоящее время архитектуры глубоких нейронных сетей;
- изучить способы применения глубоких нейронных сетей для задач компьютерного зрения и анализа текстов;
- изучить программные системы обучения глубоких нейронных сетей.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- процедуры критического анализа, методики анализа результатов исследования и разработки стратегий проведения исследований, организации процесса принятия решения;
- современные коммуникативные технологии на государственном и иностранном языках;
- современное программное и аппаратное обеспечение информационных и автоматизированных систем;
- принципы, методы и средства анализа и структурирования профессиональной информации;
- методы и средства разработки программного обеспечения, методы управления проектами разработки программного обеспечения, способы организации проектных данных, нормативно-технические документы (стандарты и регламенты) по разработке программных средств и проектов;
- современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач.

уметь:

- принимать конкретные решения для повышения эффективности процедур анализа проблем, принятия решений и разработки стратегий;
- применять на практике коммуникативные технологии, методы и способы делового общения;
- разрабатывать программное и аппаратное обеспечение информационных и автоматизированных систем для решения профессиональных задач;
- анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров;
- выбирать средства разработки, оценивать сложность проектов, планировать ресурсы, контролировать сроки выполнения и оценивать качество полученного результата;
- обосновывать выбор современных информационно-коммуникационных и интеллектуальных технологий, разрабатывать оригинальные программные средства для решения профессиональных задач.

владеть:

- методами установления причинно-следственных связей и определения наиболее значимых среди них; методиками постановки цели и определения способов ее достижения; методиками разработки стратегий действий при проблемных ситуациях;
- методикой межличностного делового общения на государственном и иностранном языках, с применением профессиональных языковых форм и средств;
- методами модернизации программного и аппаратного обеспечения информационных и автоматизированных систем для решения профессиональных задач;
- методами подготовки научных докладов, публикаций и аналитических обзоров с обоснованными выводами и рекомендациями;
- методами разработки технического задания, составления планов, распределения задач, тестирования и оценки качества программных средств;
- методами разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач.

Темы и разделы курса:**1. Введение в Deep Learning**

Введение. Машинное обучение и типы данных. Нейронные сети. Обучение нейронной сети. Задача компьютерного зрения. Ядро свёртки. Deep Learning. Типы слоёв. Обработка последовательностей. От распознавания к синтезу. Состязательные сети. Причины успеха.

Теоретическое резюме. Практика. Открываем ноутбук в Colab. Практика. Линейная модель. Практика. Бустинг. Практика. Нейронная сеть.

2. Основные фреймворки

Уровни абстракции. Вспоминаем линейную классификацию. Граф вычислений. MLP. TensorFlow. Пример обучения. Детали градиентного спуска: цепное правило. Граф вычисления производных. Пример одного нейрона. Цепное правило и граф производных. Практика. TF в Google Colab. Практика. Знакомство с PyTorch. Подробнее о тензорах. Обучение нейросети.

3. Сверточные нейронные сети

Введение в свёртки. Как применять нейросеть. Операция свёртки. Простой свёрточный слой. Практика. Простой свёрточный слой. Усложняем свёрточный слой. Практика. Усложняем свёрточный слой. Пулинг слой. Практика. Пулинг слой. Первая свёрточная сеть. Практика. Первая свёрточная сеть. Современные архитектуры. Inception V3.

4. Задача оптимизации

Введение. Функции активации. Инициализация весов. Влияние learning rate и масштаба признаков на сходимость. Batch-нормализация. Dropout-регуляризация. Производная функции. Стохастический градиентный спуск (SGD). Adam — Adaptive Moment Estimation. Матричные операции.

5. Finetuning & Transfer Learning

Введение. Как получить такие картинки. Перенос обучения. Fine-tuning. Автокодировщики. Un-pooling. Поиск похожих изображений. Быстрый KNN. Пространство представлений

6. Обучение с подкреплением (Reinforcement Learning)

Введение. Классы обучения. Обучение с подкреплением. Примеры применения обучения с подкреплением. FROZEN LAKE: Состояния и действия. FROZEN LAKE: Действия и награды. FROZEN LAKE: s , a , r , s' . FROZEN LAKE: Эпизоды и Цель обучения. Оптимальная политика. Оценка состояний и действий. Оптимальная Q-функция. Q-функция: таблица. Q-LEARNING. Q-LEARNING: Аппроксимация и Обучение. Трюк 1: EXPLORATION vs EXPLOITATION. Q-LEARNING: Обучение. Трюк 2: EXPERIENCE REPLAY. DEEP Q-LEARNING. Другие методы RL

7. What's next? Продвинутое нейронные сети

Введение. Продвинутое задачи. Автоэнкодеры. Часть 1. Автоэнкодеры. Часть 2. Сиамские нейронные сети. ONE-SHOT LEARNING. Обучение ONE-SHOT LEARNING. Генеративные состязательные сети. Генератор и дискриминатор. Синтез изображений. Практика. Aautoencoder. Практика. VAE. Практика. GAN

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Задачи генерации в NLP

Цель дисциплины:

• познакомить магистрантов с компьютерной лингвистикой (КЛ) как научно-практическим направлением, его краткой историей, задачах, связанных с обработкой основных лингвистических и текстовых параметров языковых произведений, а также с методами и технологиями, используемыми в рамках компьютерной лингвистики.

Задачи дисциплины:

- создать представление о компьютерной лингвистике как новейшей научно-практической области исследований, ее возникновении в контексте смежных наук и ее современной организации;
- познакомить магистрантов с основными лингвистическими технологиями, реализующими анализ предложения (текста) по уровням лингвистической разметки и основными приемами автоматической генерации текстов;
- познакомить магистрантов с основными типами ресурсов, создающимися и используемыми компьютерными программами для решения конкретных задач в исследовательских целях, при разработке лингвистических технологий и в приложениях;
- соединить интуитивные и традиционные представления о свойствах естественно-языковых текстов со способами их формализации и моделирования в работах по компьютерной лингвистике;
- выработать у магистрантов элементарные практические навыки по применению компьютерно-лингвистических методов к языковому материалу и использованию лингвистических технологий.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- специфические особенности задач генерации текстов;
- формулировки постановок задач генерации текстов и метрики оценки качества;
- основные модели для решения “этих” задач.

уметь:

- решать задачи средствами huggingface transformers.

владеть:

- state of the art инструментами NLP для решения задач генерации текстов.

Темы и разделы курса:**1. Особенности использования трансформеров для генерации текстов**

Расширяем базовое представление об архитектуре трансформера до вида энкодер и декодер. Определяем отличия специфики задач генерации от задач извлечения информации.

2. Задача машинного перевода

Изучаем постановку задачи машинного перевода и основные модели и метрики оценки качества.

3. Задача генерации текстов

Изучаем постановку задачи машинного перевода и основные модели и метрики оценки качества

Изучаем постановку задач генерации текстов, знакомимся со основными моделями и корпусами для их предобучения.

4. Задача абстрактивной саммаризации

Изучаем постановку задачи абстрактивной саммаризации.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Инжиниринг данных

Цель дисциплины:

- сформировать у студентов навыки инжиниринга данных.

Задачи дисциплины:

- сформировать понимание ключевых способов интеграции, обработки, хранения больших данных;
- сформировать навыки разработки и взаимодействия с компонентами экосистемы Hadoop, распределенными хранилищами и платформами.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- процедуры критического анализа, методики анализа результатов исследования и разработки стратегий проведения исследований, организации процесса принятия решения;
- основные принципы профессионального и личностного развития, исходя из этапов карьерного роста и требований рынка труда; способы совершенствования своей деятельности на основе самооценки;
- математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности;
- методы и средства разработки программного обеспечения, методы управления проектами разработки программного обеспечения, способы организации проектных данных, нормативно-технические документы (стандарты и регламенты) по разработке программных средств и проектов;
- современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач.

уметь:

- принимать конкретные решения для повышения эффективности процедур анализа проблем, принятия решений и разработки стратегий;

- решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний;
- выбирать средства разработки, оценивать сложность проектов, планировать ресурсы, контролировать сроки выполнения и оценивать качество полученного результата;
- обосновывать выбор современных информационно-коммуникационных и интеллектуальных технологий, разрабатывать оригинальные программные средства для решения профессиональных задач.

владеть:

- методами установления причинно-следственных связей и определения наиболее значимых среди них;
- методиками постановки цели и определения способов ее достижения; методиками разработки стратегий действий при проблемных ситуациях;
- способами управления своей познавательной деятельностью и ее совершенствования на основе самооценки и принципов образования в течение всей жизни;
- методами теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте;
- методами разработки технического задания, составления планов, распределения задач, тестирования и оценки качества программных средств;
- методами разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач.

Темы и разделы курса:

1. Вводный модуль

Профессии в Data Science. Архитектура ЭВМ: Цифровой-логический уровень. Архитектура ЭВМ: Аппаратный уровень. Операционные системы. Сетевые технологии.

2. Тестирование

Введение. Чтение сообщений об ошибках. Исключения. Стэк вызова. Что такое исключение. Обработка исключений. Осознанная обработка исключений. Иерархия исключений. Детали try-except. Отладка: введение. Отладка исключений. Отладка через print. Отлов багов. Использование pdb. Смысл тестирования. Автоматическое тестирование. Типы тестов.

3. Подготовка модели в Production

Модели в вакууме никому не нужны. Жизненный цикл моделей машинного обучения. Сохранение и загрузка моделей. Сериализация. Бинарные файлы: Pickle. Сохранение

пайплайна. Практика: Pickle. Форматы PMML и ONNX-ML. Что такое воспроизводимость? Офлайн- и онлайн-модели.

4. Введение в SQL

Введение в базы данных. Системы управления базами данных (СУБД). Metabase. Введение в SQL. SELECT. Таблицы. Типы данных в SQL. Кортеж. ORDER BY. LIMIT и OFFSET.

5. Агрегатные функции

Основные агрегатные функции. DISTINCT и GROUP BY. HAVING.

6. Связи таблиц

Схема данных. Советы по написанию запросов. Знакомство с датасетом. Таблица city. Знакомство с датасетом. Таблица customer. Знакомство с датасетом. Таблица driver. Знакомство с датасетом. Таблица truck. Знакомство с датасетом. Таблица shipment. Организация данных в БД. Организация в Postgres. information_schema.tables. information_schema.columns. information_schema.table_constraints.

7. Соединения таблиц

JOIN. Синтаксис. Комментирование. Работа с INNER JOIN. OUTER JOIN. Виды. OUTER JOIN. LEFT JOIN. FULL OUTER JOIN. CROSS JOIN.

8. Сложные объединения

UNION. UNION и ограничения типов данных. UNION ALL и промежуточные итоги. Другие способы применения UNION. EXCEPT и INTERSECT.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Инструменты Big Data

Цель дисциплины:

- овладение студентами базовыми алгоритмами и технологиями по обработке и анализу данных, изучение инструментария для обработки, в том числе “больших данных” (Big Data), для их применения в реальных проектах.

Задачи дисциплины:

- приобретение студентами навыков по обработке и анализу данных, способность выбирать необходимые инструменты и алгоритмы анализа данных в зависимости от характера данных, структуры и т.п., а также потребностей организации по их анализу.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- подходы к организации хранилищ данных в современной компании, а также тенденциях их развития и способы взаимодействия с ними;
- традиционные и «нетрадиционные» источники данных для бизнес-анализа;
- решаемые в процессе анализа данных задачи;
- основные способы извлечения данных;
- основные подходы и методы анализа данных.

уметь:

- планировать работы по выполнению проектов, связанных с анализом, в том числе больших, данных;
- использовать инструментарий для извлечения данных из различных источников (БД, публичные web-сервисы и т.п.);
- использовать инструментарий для анализа данных (статистические пакеты и т.п.), в том числе в рамках современных парадигм обработки данных больших объемов данных (map-reduce и т.п.).

Владеть:

- навыками постановки задачи анализа данных в интересах компании, способами предобработки и предварительной визуализации данных;
- навыками построения аналитических моделей и методов их оценки;
- навыками донесения результатов аналитических исследования до бизнес-спонсоров и коллег.

Темы и разделы курса:**1. Введение в проблематику обработки больших объемов данных**

Ценность данных для современной компании. Понятие BigData. Понятие Data Science. Методология, хранения, обработки и анализа данных.

Задачи специалиста по обработке данных. Жизненный цикл проекта анализа данных. Представление результатов анализа.

2. Основы анализа данных

Статистический анализ данных и машинное обучение. Пакеты для статистического анализа данных. Визуализации данных.

Кластеризация данных. Правила ассоциативности.

Регрессионный анализ: линейная и логистическая регрессия.

Вероятностные графовые модели. Сети Байеса. Наивный Байес.

Деревья принятия решений. Анализ временных рядов.

3. Специальные методы анализа больших данных

Рекомендательные системы (recommender systems): основные понятия, история возникновения и базовые подходы к построению, неперсонализированные и персонализированные системы, современные рекомендательные системы: совместная фильтрация (пользователь-пользователь, продукт-продукт), фильтрация содержимого.

Особенности анализа потоковых данных. Поточные алгоритмы и соответствующий инструментарий. Понятие кэширования данных, основные подходы и их применение. Понятие хэширования, основные подходы, специальные виды хэширования (LSH).

Вероятностные графовые модели. Основные понятия и классификация. Сети Байеса. Способы задания и использования. Основные структуры данных и типы запросов к сети. Анализ потоков влияния. Инструменты для формирования моделей. Вероятностное программирование.

4. Технологии анализа данных

Подходы к анализу неструктурированного текста. Технология MapReduce. Анализ данных при помощи средств БД (in-database analysis)

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Компьютерное зрение

Цель дисциплины:

- усвоение студентами основных задач, моделей, методов и алгоритмов в области компьютерного зрения.

Задачи дисциплины:

- познакомить студентов с основами и современными методами компьютерного зрения и обработки изображения, включая извлечение семантической и метрической информации из изображений;
- сформировать у студентов практические навыки работы с изображениями и решения прикладных задач анализа изображений.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- методы анализа проблемных ситуаций как систем;
- информационно-коммуникационные технологии для поиска и анализа профессиональной информации;
- о современном состоянии исследований в рамках тематической области своей профессиональной деятельности;
- методы анализа задач и пути их решения;
- основы проектирования и элементы архитектурных решений информационных систем;
- современные языки программирования и методы параллельной обработки данных.

уметь:

- осуществлять поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации;
- применять знание информационно-коммуникационных технологий для решения поставленной задачи, формулирования выводов и оценки полученных результатов;

- оценивать актуальность исследований в области информатики и вычислительной техники и их практическую значимость;
- анализировать и вычислять решения, понимает и учитывает на практике границы применимости получаемых решений;
- применять в практической деятельности профессиональные стандарты в области информационно-коммуникационных технологий;
- реализовывать и применять численные методы решения прикладных задач в профессиональной сфере деятельности, используя пакеты программного обеспечения, операционные системы, электронные библиотеки, сетевые технологии.

владеть:

- навыками разработки стратегий достижения поставленной цели как последовательности шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности;
- навыками адаптации зарубежных комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий;
- профессиональной терминологией, используемой в современной научно-технической литературе, обладает навыками устного и письменного изложения результатов научной деятельности в рамках профессиональной коммуникации;
- навыками самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте;
- опытом оценки качества, надежности и эффективности информационной системы в конкретной профессиональной сфере.

Темы и разделы курса:

1. Работа с изображениями

Что такое изображение? Скринкаст: openCV. Image Processing. Скринкаст: skimage. Подготовка данных в основных задачах CV. Скринкаст: разметка. Скринкаст: класс Dataset. Аугментации. Скринкаст: аугментации.

2. Hard & Software. Часть I

Сервер и виртуальный сервер. Вычисления на GPU и CPU. Создание сервера и подключение к нему. Настройка окружения сервера.

3. Hard & Software. Часть II. TensorFlow и Keras для задач CV

Принципы работы библиотек для глубокого обучения. Построение моделей в Keras. Слои и их параметры в Keras. Компиляция моделей в Keras. Обучение моделей в Keras.

Визуализация обучения в Keras. Кастомизация моделей и обучения в Keras. Работа с данными в Keras. Часть I. Работа с данными в Keras. Часть II.

4. Задача классификации изображений. Часть 1

Задача классификации изображений. Сверточные нейросети. Архитектура LeNet-5 и датасет MNIST. Примеры классических подходов к классификации. AlexNet: победа над классическими подходами. Визуализация свёрточных сетей на примере ZFNet. Архитектуры VGG и GoogleNet

5. Задача классификации изображений. Часть 2

Семейство Inception и Inception-v3. Семейство архитектур ResNet. Франкенштейн Inception-ResNet-V2. Групповые свёртки и архитектура ResNeXt. Dense-блок и архитектура DenseNet. SE-блок и архитектура SE-ResNet. NASNet: генерация архитектур. EfficientNet: эффективное масштабирование архитектур.

6. Перенос обучения в CV

Перенос обучения и его стратегии. Предобученные сети как экстракторы признаков. Дообучение предобученных сетей. Преимущества и подводные камни.

7. Распознавание объектов, часть 1

Задачи распознавания объектов. Современные методы детектирования объектов. Семейство R-CNN. Семейство YOLO.

8. Распознавание объектов. Часть II

Метрики качества моделей распознавания. Использование переноса обучения для распознавания объектов. Сферы применения алгоритмов распознавания.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Математика и алгоритмы машинного обучения

Цель дисциплины:

- ознакомление студентов с методологией и основными алгоритмами машинного обучения.

Задачи дисциплины:

- знакомство с различными постановками задач машинного обучения, знакомство с важнейшими моделями и метриками классификации и регрессии, овладение методологией машинного обучения, получение навыков выбора, обучения и оценки моделей.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные алгоритмы для решения задач классификации и регрессии.

уметь:

- преобразовывать данные к виду и формату, удобному для анализа;
- очищать данные;
- осуществлять первичный анализ и визуализацию данных;
- выбирать модель для решения конкретной задачи и обосновывать свой выбор;
- обучать выбранную модель, добиваясь необходимого качества;
- выбирать метрики для оценки модели.

владеть:

- средствами из специализированных библиотек языка Python для обучения и оценки моделей машинного обучения.

Темы и разделы курса:

1. Основные понятия математической статистики на Python

Студент повторяет основные понятия ТВиМС и учится применять эти знания для решения практических задач по статистике с использованием средств языка Python.

2. Основные понятия линейной алгебры и математического анализа

Студент повторяет основные понятия линейной алгебры и математического анализа, необходимые для дальнейшего усвоения материала. Основное внимание уделяется матричным операциям и оптимизации методом градиентного спуска.

3. Теория машинного обучения

Студенты знакомятся с определением машинного обучения, типами алгоритмов и основными задачами МО. Рассматриваются вопросы определения бизнес-требований, сбора и подготовки данных, обучения, тестирования и внедрения моделей.

4. Задача регрессии

Студенты знакомятся с постановкой задачи регрессии, изучают её решение с использованием линейной регрессии, оценивают качество решения задачи регрессии.

5. Задача классификации

Студенты знакомятся с постановкой задачи классификации, изучают логистическую регрессию. Знакомятся с метриками качества классификации.

6. Решающие деревья и случайный лес

Студенты знакомятся с решением задач классификации и регрессии с использованием решающих деревьев. Разбирают алгоритмы построения решающих деревьев. разбивают концепцию бэггинга и учатся строить случайный лес решающих деревьев для решения задач классификации и регрессии.

7. Основные алгоритмы построения ансамблей. Бустинг. Стекинг

Студенты знакомятся с концепцией ансамблей, разбирают концепцию градиентного бустинга. Разбирают решение задач с использованием основных библиотек градиентного бустинга. Знакомятся с концепцией стэкинга. Решают практические задачи с использованием стекинга.

8. Обучение без учителя. Кластеризация

Студенты знакомятся с постановкой задачи кластеризации, основными алгоритмами и метриками оценки качества. Решают практические задачи.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Менеджмент для наук о данных

Цель дисциплины:

- сформировать навыки поиска скрытых закономерностей, прогнозирования развития событий и оптимизации ключевых бизнес-процессов.

Задачи дисциплины:

- подготовить и обучить специалистов новейшим навыками управления проектами;
- сформировать у студентов понимание основных концепций и инструментов для управления, планирования и реализации IT-проектов.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные понятия и методы фундаментальных дисциплин;
- методы управления проектами; этапы жизненного цикла проекта;
- математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности;
- общие принципы исследований, методы проведения исследований;
- методы и средства разработки программного обеспечения, методы управления проектами разработки программного обеспечения, способы организации проектных данных, нормативно-технические документы (стандарты и регламенты) по разработке программных средств и проектов.

уметь:

- применять знания, полученные в области фундаментальных наук, и использовать их в профессиональной деятельности;
- разрабатывать и анализировать альтернативные варианты проектов для достижения намеченных результатов; разрабатывать проекты, определять целевые этапы и основные направления работ;

- решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний;
- формулировать принципы исследований, находить, сравнивать, оценивать методы исследований;
- выбирать средства разработки, оценивать сложность проектов, планировать ресурсы, контролировать сроки выполнения и оценивать качество полученного результата.

владеть:

- методикой выбора методов решения задач профессиональной деятельности на основе теоретических знаний;
- методами установления причинно-следственных связей и определения наиболее значимых среди них; методиками постановки цели и определения способов ее достижения; методиками разработки стратегий действий при проблемных ситуациях;
- методами теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте;
- методами проведения исследований для решения практических задач профессиональной деятельности;
- методами разработки технического задания, составления планов, распределения задач, тестирования и оценки качества программных средств.

Темы и разделы курса:

1. IT в современном мире

Buzzwords в IT. Frontend. Backend. Сетевая инфраструктура. Бизнес-аналитика (Business Intelligence).

2. Оценка навыков и ресурсов

Первый проект. Выбор задачи. Как писать ТЗ. Конкретизация решения. Влияние трендов. Карта ресурсов. Осознанный выбор технологий.

3. Методологии управления проектами

Waterfall и Agile. Типы проектов. Разница между водопадным и гибким подходом. Kanban.

4. Начало проекта

Понятие проекта в IT. Проектный треугольник. Чек лист на первую неделю проекта.

5. Планирование проекта Homework

Этапы проекта. План и дорожная карта проекта. Брифование заказчика. Критерии успеха. Планирование в Waterfall-проекте. Планирование в Agile-проекте.

6. Сбор требований

Работа с требованиями на проекте. Требования в Waterfall. Требования в Agile. Бэклог продукта.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Методология научных исследований

Цель дисциплины:

Сформировать систему ключевых представлений о научном познании и научных методах исследования, которые могут понадобиться студентам при работе над академическими и научно-прикладными проектами в рамках магистратуры.

Задачи дисциплины:

- ознакомление с общенаучными методами исследования и познания;
- формирование умения разбираться в методологических подходах к научным исследованиям;
- формирование методологической базы для организации научно-исследовательской деятельности;
- развитие критического мышления и навыков ведения аргументированной дискуссии;
- изучение технологий, инструментов поиска научной информации и способов проверки качества инструментов отбора информации и критериев ее отбора.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные методологические принципы науки;
- базовую схему «тема - проблема-гипотеза -объект-предмет-методы исследования-сбор данных -анализ данных- выводы»;
- особенности научного метода постижения мира;
- различия между "проблемой" и "задачей" научного исследования;
- различение субъекта и объекта исследования;
- виды научного исследования, различия между фундаментальным и прикладным исследованием;
- теоретические и эмпирические методы исследования;
- краткую историю возникновения и развития научных традиций;

- ключевые события и особенности научной революции;
- прикладные функции науки;
- междисциплинарные исследовательские методы;
- общие методы сбора и анализа данных для исследования;
- цели и задачи проведения научного исследования;
- этапы проведения научного исследования;
- общие методы формулирования и проверки гипотез;
- логику и методы сбора и анализа научных источников;
- принципы цитирования в научных работах;
- принципы представления принципов, методов и результатов исследования;
- общепринятые в науке методы обработки и интерпретации количественных данных.

уметь:

- объяснить предпосылки и развитие научного метода познания мира;
- объяснить особенности исследовательского подхода к исследованию мира;
- определять критерии, по которым деятельность определяется как исследование;
- понимать, чем отличаются информация и знание;
- ставить исследовательский вопрос и определять цель исследования;
- использовать фундаментальные и прикладные методы исследования в своей научно-исследовательской деятельности;
- выбирать подходящий метод исследования под свою исследовательскую задачу;
- анализировать академические источники и валидировать их (оценивать их качество, избегать лженаучных публикаций),
- использовать основные векторы поиска научной информации и способы его оптимизации;
- опознавать основные признаки нарушения научно-исследовательской этики;
- обосновывать выбранные методы к исследованию проблемы в формате краткой презентации,
- определять критерии наукоемкости прикладного проекта.

владеть:

- навыками аргументации собственного мнения;
- навыками сбора и анализа источников для исследования;

- навыками оценки выборок для количественных методов исследования и анализа этих выборок;
- базовыми междисциплинарными методами качественных и количественных исследований;
- навыками написания научного аргументированного текста;
- методами научного исследования, принятыми в профессиональной области.

Темы и разделы курса:

1. Общенаучные методы исследования и познания

Специфические особенности научной картины мира. Ключевые методологические принципы научных исследований. Онтологический и эпистемологический уровни познавательной деятельности, проблема их взаимосвязи. Различение знания и информации. Научное знание, его специфические особенности, сопоставление с вненаучными формами знания (миф, искусство, религия). Основные методологические принципы науки: причинности, объяснения, соответствия. Институциональные формы науки. Различение "проблемы" и "задачи" в процессе научного исследования. Рассмотрение базовой схемы : тема - проблема-гипотеза -объект-предмет-методы исследования-сбор данных -анализ данных- выводы. Проблема выделения структурного элемента в процессе исследования. Проблема выбора критерияльной базы для анализа данных.

Краткая история развития науки. Возникновение исследовательской традиции в культурах Древнего Китая, Древней Индии и Древней Греции. Научная методология Аристотеля. Возникновение первых научных традиций в структуре философских систем Древнего Китая, Древней Индии и Древней Греции. Наука Древнего Китая: опора на нумерологию, отсутствие атомизма и формальной логики. Наука Древней Индии: атомизм школы вайшешика и логика ньяя. Научные программы древнегреческой философии: пифагорейство, ионийская традиция, атомизм, школа элеатов. Категория идеального. Платон. Систематизация научной методологии в трудах Аристотеля. Основные категории науки Аристотеля. Теория силлогизмов. Физика Аристотеля как пример качественной науки. Роль наблюдений и наблюдателя. Телеологизм и учение о сущности и энтелехии. Отношение исследователя и окружающего мира в науке Аристотеля.

Научная революция Нового Времени и формирование современного типа научного знания. Начало эпохи европейского Нового Времени. Базовые социокультурные процессы: Реформация, Великие географические открытия, Научная революция. Вклад великих астрономов в трансформацию научной методологии. Н.Коперник, Т.Браге, И.Кеплер, Г.Галилей. Математизация естествознания. Новое понимание экспериментальной практики. Отказ от телеологизма и изучения сущностей. Изменение статуса исследователя. Р.Декарт и его методологические правила.

Становление европейского естествознания в XVIII-XIX веках, развитие методов экспериментального и теоретического исследования природы. "Методологические правила" И.Ньютона. "Программа П.С.Лапласа" и ее реализация французскими физиками: механика, электродинамика, термодинамика, оптика. Новые подходы к использованию математики: дифференциальные уравнения в частных производных и теоретико-групповой подход. Представление о методах классической физики как "методологическом стандарте"

решения проблем других наук. Опора на редукционизм и структурализм. Роль механицизма в развитии биологии, социально-гуманитарных наук и экономики. Кризис механицизма и редукционизма. Формирование историко-генетического подхода к пониманию сложных систем.

Революция в естествознании в первой половине XX века. Современные проблемы методологии науки. Проблемы соотношения теоретического и эмпирического уровней в познании. Проблема редукции. Проблема социального контекста. Закат механицизма и детерминизма в физической теории первой половины XX века. Роль вероятностной интерпретации поведения физических объектов и концепции нелокальности взаимодействий. Становление "неклассического" типа рациональности, антиредукционизма и холизма. Влияние идей теоретической физики на развитие биологии и социально-гуманитарных наук. Роль методологических проблем в развитии современной

науки. Проблема соотношения эмпирического и теоретического уровней познания. Проблема взаимосвязи нормативности методологических правил, интуиции и творчества. Проблема абсолютизма и релятивизма в понимании истины. Проблема социального контекста науки, влияния вненаучных факторов на процесс научного познания.

2. Теории и методы прикладного исследования

Основы работы с данными. Логика работы с данными (изучаем генеральную совокупность, иметь дело вынуждены с выборками). Ошибки, связанные с некорректной подготовкой выборок (избыточный поиск, ошибка выжившего, игнорирование сезонности, ложное знание будущего, отсутствие статистичности). Описание выборки. Отчёт о подтверждении

качества экономико-статистической модели. Ошибки работы с временными рядами. Стационарность, автокорреляция, мультиколлинеарность. Причинность. Примеры этих ошибок.

Сумма экономики. Время: процентные инструменты как «машина времени», направленная в будущее, долгосрочный и краткосрочный периоды, темп, прошлое, настоящее, будущее.

Средства производства и технологии: смена ведущих факторов производства, принципы экономического развития, разделение труда. Теория ценности. Проблема прибыли. Цели хозяйствования: прибавочный продукт, преобразование ландшафта против одомашнивания человека, цеховая организация производства. Состав экономической теории. Детерминизм общественных отношений в результате природных условий и наилучших доступных технологий. Гармония финансовой системы.

3. Организация научно-исследовательской деятельности

Критерии отбора и проверки источников: хищники и охотники; судьи и гибриды; инструменты.

Хищническая периодика и издательства, токсичные авторы. Нарушения академической этики: плагиат (по уровням – копипаст, плагиат идей, переводной плагиат и т.д.), научный подлог, множественные публикации, публикации с загадочным авторством.

Ложное (приписное) авторство + административный ресурс.

Некорректное цитирование. Само- и перекрестное цитирование. Искажение фактологии/авторства, битые ссылки и т.п.

Jeffrey Beall's list, <https://www.antiplagiat.ru>, www.plagiarismcheck.org, www.duplichecker.com, <https://vroniplag.de/>, Диссернет и Диссеропедия (журналы, вузы).

Квартили и процентыль. Белые и черные списки библиотек (примеры: НЭБ, а также РГБ – в перспективе) и университетов (МФТИ, НИУ ВШЭ, МГУ им. М.В. Ломоносова и др.).

Редакционные практики в России и за рубежом. Ретракция статей. Добровольная (пример Фрэнсис Арнольд), принудительная (кейс Ильи Медведева). Внедрение в работу Clarivate и Elsevier элементов пакетной ретракции (термин Григория Консона).

РАН, ВАК Мин. науки и высш. обр. РФ, университеты.

Сайты академической и quasi-академической периодики.

WoS vs Scopus (предметные рейтинги и рэнкинг), хедлайнеры академической периодики (периодичность публикаций и их вес).

Интеграция ссылок на фактологию и идеи, представленные на классических исследовательских ресурсах, порталах научно-просветительского профиля, а также неакадемических сайтах.

Качество и глубина ссылок в контексте проблематизации основного объекта исследования.

Выявление/формулирование исследовательских трендов.

3.2. Поиск информации для академических исследований. Сайты и плагины

NathTrust, Unpaywall, Open Access Button, Master Journal List, группы журналов университетов Расселовской группы (Oxford Journals Online, Cambridge Journals Online и др.) и Лиги Плюща (Harvard Journals / Harvard Library и т.п.).

Ведущая международная академическая периодика: Nature, Science и др.

PLOS One, SAGE, JSTOR, TEI.

Агрегаторы и сайты: Academia.edu, Google Scholar, Google Books, Frontiers, arXiv.org, ERIC, Psycjournals, диссертационные разделы сайтов университетов России и зарубежья.

Библиотеки, библиотечные системы и цифровые визитки исследователей: Библиотека Конгресса США, РГБ, eLibrary, КиберЛенинка, Истина (МГУ и др.), Gale, ResearcherID, ORCID.

Сайты издательских групп – Springer Link, Wiley Online Library, Taylor & Francis и др.

Издательства университетов – топ-20 QS, THE, ARWU.

Сайты государственных и частных архивов, специализированные зарубежные агрегаторы: Factiva и др., отечественные: EastView, Public.ru и др.

Поисковики со спорным правовым статусом (Sci-Hub); профильные порталы / группы (Arzamas, Постнаука, Всенаука, просветительские материалы РГБ и т.д.).

Блогеры/влогеры/ведущие тематических программ (Артем Оганов и др.) на телеканалах и в социальных сетях.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Методы обучения с подкреплением

Цель дисциплины:

- познакомить студента с задачами обучения с подкреплением и современными методами их решений.

Задачи дисциплины:

- дать практические навыки в использовании методов для решения задач с обучением с подкреплением.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- метрики оценки качества задач;
- модели для решения задач.

уметь:

- формулировать постановку задачи в виде задач обучения с подкреплением;
- формулировать функцию потерь;
- оценивать качество решения задач.

владеть:

- методами решения задач обучения с подкреплением как с помощью классических методов ml, так и deep learning.

Темы и разделы курса:

1. Введение в reinforcement learning

Постановка задачи обучения с подкреплением, основные применения на практике таких подходов.

2. Deep Q-learning

Формулировка обучения с подкреплением через Q-learning, знакомство с архитектурами нейронных сетей, используемых в данной задаче.

3. Методы основанные на политике

Изучение подхода к решению задач reinforcement learning без оптимизации Q-функционала

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Модели ML в продакшн

Цель дисциплины:

- сформировать у студентов навыки веб-программирования и инструментов для настройки продукта с машинным обучением.

Задачи дисциплины:

- изучить процесс создания веб-сервиса с машинным обучением;
- изучить систему мониторинга качества модели;
- изучить прототипы веб-сервисов и веб-приложений.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- стратегию достижения поставленной цели как последовательность шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности;
- фундаментальные научные знания в области физико-математических наук;
- междисциплинарные связи в области математики и физики и способен их применять при решении задач профессиональной деятельности;
- нормативную документацию для стандартизации принятых решений и унификации разработанных изделий.

уметь:

- анализировать проблемную ситуацию как систему, выявляя ее составляющие и связи между ними;
- осуществлять поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации;
- организовывать и координировать работу участников проекта, способствует конструктивному преодолению возникающих разногласий и конфликтов;

- учитывать в своей социальной и профессиональной деятельности интересы, особенности поведения и мнения (включая критические) людей, с которыми работает/взаимодействует, в том числе посредством корректировки своих действий;
- предвидеть результаты (последствия) как личных, так и коллективных действий;
- обобщать и критически оценивать опыт и результаты научных исследований в области профессиональной деятельности;
- применять знания и навыки по использованию информационно-коммуникационных технологий для поиска и изучения научной литературы, применения прикладных программных продуктов;
- применять знания в области физико-математических наук для решения поставленной задачи, формулирования выводов и оценки полученных результатов;
- аргументированно выбрать способ проведения научного исследования;
- находить, анализировать и обобщать информацию об актуальных результатах исследований в рамках тематической области своей профессиональной деятельности;
- выдвигать гипотезы, строить математические модели для описания изучаемых явлений и процессов, оценить качество разработанной модели;
- применять теоретические и (или) экспериментальные методы исследований к конкретной научной задаче и интерпретировать полученные результаты;
- самостоятельно определять особенности и качество разрабатываемого проекта.

владеть:

- теоретическими и (или) экспериментальными методами исследований к конкретной научной задаче и интерпретировать полученные результаты.

Темы и разделы курса:

1. Работа с моделями машинного обучения

Управление моделями машинного обучения. Определение моделей машинного обучения. Использование моделей машинного обучения. Понятие контейнера. Процесс обертывания модели машинного обучения в Docker.

2. Написание веб-сервиса

Развертывание модели машинного обучения. Примеры запросов на выборку данных для модели машинного обучения. Обзор сервисов для развертывания моделей машинного обучения. Принципы верстки. Протоколы и спецификации.

3. Создание визуального прототипа

Методы разработки естественного языка. Способы создания прототипов. Визуальные интерфейсы для создания прототипов. Быстрое прототипирование. MVP (minimum viable product) прототип. Прототипы интерфейса. Прототип функционального web-продукта.

4. Работа с удаленными серверами

Понятие удаленного сервера. Программное обеспечение для работы с удаленными серверами. Облачные сервисы.

5. Системы мониторинга

Создание тестовых данных для контроля и мониторинга алгоритмов. Инструменты и программное обеспечение для мониторинга систем.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Программирование на Python

Цель дисциплины:

- освоение программирования на Python для применения в сфере анализа данных.

Задачи дисциплины:

- формирование систематизированных знаний и навыков в области программирования, ознакомление с принципами работы современных языков программирования, с актуальными парадигмами программирования;
- формирование навыков владения основными принципами программирования на высокоуровневом языке Python и прикладными аспектами его применения;
- изучение дисциплины направлено на развитие у обучающихся алгоритмического мышления, формирование навыков реализации различных алгоритмов на высокоуровневом языке программирования, применение знаний к решению прикладных задач.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные принципы профессионального и личностного развития, исходя из этапов карьерного роста и требований рынка труда;
- способы совершенствования своей деятельности на основе самооценки;
- современное программное и аппаратное обеспечение информационных и автоматизированных систем;
- аппаратные средства и платформы инфраструктуры информационных технологий, виды, назначение, архитектуру, методы разработки и администрирования программно-аппаратных комплексов объекта профессиональной деятельности;
- современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач.

уметь:

- решать задачи собственного профессионального и личностного развития, включая задачи изменения карьерной траектории;
- расставлять приоритеты;
- разрабатывать программное и аппаратное обеспечение информационных и автоматизированных систем для решения профессиональных задач;
- анализировать техническое задание;
- разрабатывать и оптимизировать программный код для решения задач обработки информации и автоматизированного проектирования;
- обосновывать выбор современных информационно-коммуникационных и интеллектуальных технологий;
- разрабатывать оригинальные программные средства для решения профессиональных задач.

Владеть:

- способами управления своей познавательной деятельностью и ее совершенствования на основе самооценки и принципов образования в течение всей жизни;
- методами модернизации программного и аппаратного обеспечения информационных и автоматизированных систем для решения профессиональных задач;
- методами составления технической документации по использованию и настройке компонентов программно- аппаратного комплекса;
- методами разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач.

Темы и разделы курса:

1. Основы синтаксиса

Python — это... Почему Python? Как работать с документацией по Python. Типы данных. Основные математические операции в Python. Переменные. Типы переменных. Ввод данных пользователем и вывод на экран. Операторы сравнения и логические операторы. Условные операторы. Циклы. Функции.

2. Структуры данных

Списки. Индексация и срезы. Методы работы со списками. Сортировка списка. Цикл for и списки. Вложенные списки. Кортеж. Операции с кортежами. Словари. Обращение к ключам словаря. Добавление и удаление элементов. Перебор элементов словаря. Множества. Перебор значений множества. Добавление и удаление элементов у множеств. Операции над множествами.

3. Функции и сбор данных

Введение. Функции в Python. Проверка аргументов. Аргументы по умолчанию. Продвинутое передачу аргументов. Область видимости переменных. Замыкание. Рекурсия. Итераторы и генераторы. Lambda-функции. Итераторы и функции. Декораторы. Функции: закрепление знаний. Модули в Python. Форматы данных. Регулярные выражения. Источники данных и работа с файлами.

4. Продвинутое методы обработки данных

Модуль NumPy. Типы данных. Модуль NumPy. Массивы. Модуль NumPy. Действия с массивами. Модуль NumPy. Операции с векторами. Модуль NumPy. Случайные числа. Введение в Pandas. Pandas.Series. Pandas.DataFrame. Работа с различными источниками данных в Pandas. Знакомимся с данными: недвижимость. Исследование структуры DataFrame. Статистические методы. Фильтрация данных в DataFrame. Feature Engineering. Базовые операции со столбцами DataFrame. Работа с датами в DataFrame. Создание и преобразование столбцов с помощью функций. Тип данных Category. Сортировка данных в DataFrame. Группировка данных в DataFrame. Сводные таблицы. Объединение DataFrame: знакомимся с новыми данными. Объединение DataFrame: concat. Объединение DataFrame: join, merge.

5. Визуализация и работа с данными

Обзор типов визуализации. Знакомимся с новыми данными: коронавирус. Графические возможности библиотеки Pandas. Графические возможности библиотеки Matplotlib. Графические возможности библиотеки Seaborn. Графические возможности библиотеки Plotly. Искусство визуализации. Работа с изображениями: skimage. Цветные и чёрно-белые изображения. Быстрые numpy-преобразования. Фильтры. Геометрические преобразования.

6. Объектно-ориентированное программирование

Парадигмы и принципы программирования. Концепция ООП и значимость принципов. Классы. Магический метод `__init__`. Методы и функции. Наследование. Наследование класса. Полиморфизм в Python. Статические методы. Декораторы класса: `@property`, `@classmethod`. Исключения. Тонкости обработки исключений. Собственные классы исключений. Работа с импортом. Работа с файлами. Контекстные менеджеры. Ключевое слово `with`, принципы создания собственных контекстных менеджеров.

7. Алгоритмы и структуры данных

O-нотация. Основные структуры данных. Основные структуры данных: список, стек, очередь. Нелинейные структуры данных: графы и деревья. Создание собственных структур. Алгоритмы поиска. Алгоритмы сортировки.

8. Системы контроля версий

Современные системы управления кодом. Git. Введение. Базовые команды Git. Магия Git flow. Полезные возможности Git

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Продвинутая визуализация данных

Цель дисциплины:

- дать студентам практические навыки в применении инструментов для визуализации данных.

Задачи дисциплины:

- дать практические навыки в использовании библиотек для продвинутой визуализации данных.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- продвинутые библиотеки для визуализации данных;
- методы снижения размерности.

уметь:

- применять методы визуализации данных.

владеть:

- продвинутыми методами визуализации данных.

Темы и разделы курса:

1. Нелинейные методы снижения размерности

Методы снижения размерности, такими как t-sne

2. Методы кластеризации для визуализации данных

Алгоритмы кластерного анализа для визуализации данных

3. Продвинутый Pandas

Основы работы с библиотекой для визуализации данных

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Проектный практикум

Цель дисциплины:

- обучить студентов применять теоретические знания и навыки, полученные в рамках других курсов по машинному обучению, на практике. Студенты должны научиться решать реальные задачи машинного обучения, используя различные методы и алгоритмы, а также оценивать эффективность своих решений и улучшать их качество. Кроме того, целью дисциплины является развитие командной работы и умений в области проектирования и реализации проектов в области машинного обучения.

Задачи дисциплины:

1. Ознакомление с основными методами и алгоритмами машинного обучения.
2. Приобретение практических навыков решения задач машинного обучения на реальных данных.
3. Изучение процесса предобработки данных, выбора признаков и построения моделей.
4. Оценка качества моделей и выбор наиболее эффективного алгоритма для решения задачи.
5. Разработка и реализация проекта в области машинного обучения с использованием командной работы.
6. Анализ полученных результатов и определение возможных улучшений проекта.
7. Подготовка отчетности и презентации проекта.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- иметь представление о содержании научной деятельности, о современных исследованиях в области, соответствующей профилю образовательной программы;
- методы теоретического и экспериментального исследования в области, соответствующей профилю образовательной программы;
- принципы организации проведения экспериментов и испытаний;
- принципы оформления результатов проектных работ;

- основные методы и алгоритмы машинного обучения;
- процесс предобработки данных и выбора признаков;
- способы оценки качества моделей.

уметь:

- провести обзор имеющегося материала для решения поставленной задачи;
- использовать выбранный метод или сочетать различные методы в решении поставленной задачи;
- применять современные методы сбора и обработки данных при работе над проектом;
- строить деятельность на основе выполнения технологических требований и нормативов, придерживаться правовых и этических норм, принятых в профессиональной деятельности;
- оформлять и представлять результаты выполненной работы;
- применять методы машинного обучения для решения задач на реальных данных;
- выбирать наиболее эффективный алгоритм для решения задачи;
- разрабатывать и реализовывать проект в области машинного обучения с использованием командной работы.

владеть:

- навыками анализа научной и технической информации в области, соответствующей профилю образовательной программы.
- умением анализировать результаты и определять возможные улучшения проекта;
- навыками подготовки отчетности и презентации проекта.

Темы и разделы курса:

1. Подготовительный этап

Объединение в проектные команды. Встреча с заказчиками проектных задач. Выбор и получения проектных задач.

2. Обзор и анализ информации по теме проекта

Анализ выбранных задач, выбор методов машинного обучения и направления решения.

3. Проектная деятельность

Проведение исследования, генерация решений, написания кода. Ревью промежуточных решений.

4. Подготовка отчета

Оформление результатов по проделанной работе и выступление на семинаре.

5. Подготовительный этап

Объединение в проектные команды. Встреча с заказчиками проектных задач. Выбор и получения проектных задач.

6. Обзор и анализ информации по теме проекта

Анализ выбранных задач, выбор методов машинного обучения и направления решения.

7. Проектная деятельность

Проведение исследования, генерация решений, написания кода. Ревью промежуточных решений.

8. Подготовка отчета

Оформление результатов по проделанной работе и выступление на семинаре.

9. Подготовительный этап

Объединение в проектные команды. Встреча с заказчиками проектных задач. Выбор и получения проектных задач.

10. Обзор и анализ информации по теме проекта

Анализ выбранных задач, выбор методов машинного обучения и направления решения.

11. Проектная деятельность

Проведение исследования, генерация решений, написания кода. Ревью промежуточных решений.

12. Подготовка отчета

Оформление результатов по проделанной работе и выступление на семинаре.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Системы хранения и обработки данных

Цель дисциплины:

- изучить принципы построения и разработки систем хранения и обработки данных;
- получить навыки настройки систем хранения данных, проектирования и разработки процесса наполнения систем хранения данных, реализации запросов к системам хранения данных.

Задачи дисциплины:

- приобретение студентами навыков проектирования архитектур, применения специализированных инструментов и разработки программных систем для работы с данными.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные этапы подготовки данных и основы работы с базами данных;
- существующие методы работы с данными, области их применения, достоинства и недостатки;
- основные цифровые технологии и алгоритмы их работы.

уметь:

- применять технологии отбора, обработки, интерпретации и анализа информации;
- отбирать технологии работы с информацией в зависимости от класса задач в области данных.

владеть:

- методами и технологиями обработки и представления информации;
- технологиями извлечения информации из различных источников.

Темы и разделы курса:

1. Теоретические основы баз данных (БД)

Основные способы хранения БД в ЭВМ. Основы реляционной алгебры. Банки данных и знаний. Типология БД. Документальные БД. Фактографические БД. Гипертекстовые и мультимедийные БД.

2. Проектирование баз данных

Установка, конфигурирование и администрирование систем управления баз данных (СУБД) /Пр/

Структура и функции СУБД. Жизненный цикл БД. Виды установки СУБД (ручная установка, автоматизированная, обновление). Безопасность базы данных. Модель безопасности на основе ролей. Проектирование и разработка БД

3. Детальная разработка баз данных

Аналитическая обработка данных. Базы данных NOSQL. Понятие о BIG DATA. Колоночные базы данных. Основы фракталов. Фрактальная математика. Фрактальные методы в архивации. Управление складами данных. Особенности проектирования и разработки OLTP и OLAP приложений.

4. Структурированный язык запросов SQL

Язык SQL-запросов

Структура и семантика языка запросов

Типы запросов: создание таблиц, заполнение таблиц, выборка, сортировка, отображение полей, вставка, удаление

Типы соединений в запросах

Подзапросы. Агрегирование запросов

Реализация запросов инструментальными средами и средствами

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Современные методы DevOps

Цель дисциплины:

- усвоение студентами подхода к разработке программного обеспечения, который связывает человеческие ресурсы, процессы и технологии для получения значимых результатов и приобретение навыка быстрого переноса программного обеспечения через стадии, способствуя быстрому выпуску версий.

Задачи дисциплины:

- подготовить и обучить специалистов с ценнейшими навыками DevOps-инженера;
- сформировать у студентов знания о DevOps, как о подходе к решению задач и организации процесса разработки и эксплуатации.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- методы анализа проблемных ситуаций как систем;
- информационно-коммуникационные технологии для поиска и анализа профессиональной информации;
- методы планирования путей решения профессиональных задач;
- методику установки и администрирования программных систем.

уметь:

- осуществлять поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации;
- применять знание информационно-коммуникационных технологий для решения поставленной задачи, формулирования выводов и оценки полученных результатов;
- применять аналитические и вычислительные методы решения, понимать и учитывать на практике границы применимости получаемых решений;
- реализовывать техническое сопровождение информационных систем.

Владеть:

- навыками разработки стратегий достижения поставленной цели как последовательности шагов, предвидя результат каждого из них и оценивая их влияние на внешнее окружение планируемой деятельности и на взаимоотношения участников этой деятельности;
- навыками адаптации зарубежных комплексы обработки информации и автоматизированного проектирования к нуждам отечественных предприятий;
- методами самостоятельного приобретения, развития и применения математических, естественнонаучных, социально-экономических и профессиональных знаний для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте;
- практическим опытом интеграции информационных систем с использованием аппаратно- программных комплексов.

Темы и разделы курса:

1. Что такое DevOps и в чем его ценность

DevOps как философия и способ мышления. DevOps как подход к решению задач и организации процесса разработки и эксплуатации. Мифы о DevOps. Когда нужны практики DevOps. Когда DevOps не нужен. Профессия DevOps-инженер.

2. Системы контроля версий

Современные системы управления кодом. Git. Введение. Базовые команды Git. Магия Git flow. Полезные возможности Git.

3. Работа с облачными инфраструктурами

Знакомство с облачными инфраструктурами. AWS — Amazon Web Services. GCP — Google Cloud Platform. Microsoft Azure. Яндекс.Облако. MCS — Mail.ru Cloud Solutions. Другие провайдеры.

4. Управление конфигурациями

Предпосылки появления подхода IaC. Что такое Infrastructure as Code (IaC). Ansible. Ansible Playbook. Продвинутое написание плейбуков. Ansible-роли. Лучшие практики использования Ansible.Puppet.

5. Система сборки

Что такое Docker. Создание Docker-образа. Работа с сетью и дисковой подсистемой в Docker. Основные команды Docker: работа с образами и контейнерами. Где не стоит применять Docker. Экосистема Docker.

6. Локальная разработка и тестирование с Docker-Compose

Знакомство с Docker Compose. Формат конфигурационного файла и запуск Docker Compose. Особенности работы с сетью и дисковой подсистемой в Docker Compose. Лучшие практики Docker Compose. Тестирование.

7. Continuous Integration

Введение в CI. Основные понятия. Реализация CI. Обзор наиболее распространенных систем CI.

8. Основы Kubernetes

Введение в Kubernetes. Архитектура Kubernetes и основные компоненты. Установка K8S. Kubectl. Управление объектами Kubernetes. Манифесты. Работа с секретами.

9. DO-9. Helm, непрерывная доставка приложений в Kubernetes

Введение в Helm. Использование готовых Helm чартов. Написание Helm Chart. Лучшие практики. Helm в CI/CD.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Современные модели машинного обучения

Цель дисциплины:

- сформировать теоретические и практические знания в области обучения машин, современных методов восстановления зависимостей по эмпирическим данным, включая дискриминантный, кластерный и регрессионный анализ.

Задачи дисциплины:

- освоить методы корректной формулировки задач в терминах машинного обучения;
- овладеть навыками практического решения задач интеллектуального анализа данных.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- основные принципы и проблематику теории обучения машин;
- основные методы и алгоритмы решения задач обучения по прецедентам;
- основные области применения этих методов и алгоритмов;
- классификации, кластеризации и регрессии.

уметь:

- формализовать постановки прикладных задач анализа данных;
- использовать методы обучения по прецедентам для решения практических задач;
- оценивать точность и эффективность полученных решений.

владеть:

- основными понятиями теории машинного обучения;
- навыками самостоятельной работы при решении типовых задач;
- культурой постановки и моделирования практически значимых задач;

- навыками теоретического анализа реальных задач, решаемых с помощью алгоритмов обучения по прецедентам.

Темы и разделы курса:

1. Прогнозирование временных рядов

- Задача прогнозирования временных рядов. Примеры приложений.
- Экспоненциальное скользящее среднее. Модель Хольта. Модель Тейла-Вейджа. Модель Хольта-Уинтерса.
- Адаптивная авторегрессионная модель.
- Следящий контрольный сигнал. Модель Тригга-Лича.
- Адаптивная селективная модель. Адаптивная композиция моделей.
- Локальная адаптация весов с регуляризацией.

2. Поиск ассоциативных правил

- Понятие ассоциативного правила и его связь с понятием логической закономерности.
- Примеры прикладных задач: анализ рыночных корзин, выделение терминов и тематики текстов.
- Алгоритм APriori. Два этапа: поиск частых наборов и рекурсивное порождение ассоциативных правил. Недостатки и пути усовершенствования алгоритма APriori.
- Алгоритм FP-growth. Понятия FP-дерева и условного FP-дерева. Два этапа поиска частых наборов в FP-growth: построение FP-дерева и рекурсивное порождение частых наборов.
- Общее представление о динамических и иерархических методах поиска ассоциативных правил.

3. Нейронные сети глубокого обучения

- Свёрточные нейронные сети (CNN). Свёрточный нейрон. Pooling нейрон. Выборка размеченных изображений ImageNet.
- Свёрточные сети для сигналов, текстов, графов, игр
- Рекуррентные нейронные сети (RNN). Обучение рекуррентных сетей: Backpropagation Through Time (BPTT).
- Сети долгой кратковременной памяти (Long short-term memory, LSTM).
- Рекуррентная сеть Gated Recurrent Unit (GRU).
- Автокодировщики. Векторные представления дискретных данных.
- Перенос обучения (transfer learning).
- Самообучение (self-supervised learning).

- Генеративные состязательные сети (GAN, generative adversarial net).
- #### 4. Эвристические, стохастические, нелинейные композиции
- Стохастические методы: бэггинг и метод случайных подпространств
 - Простое голосование (комитет большинства). Алгоритм ComBoost. Идентификация нетипичных объектов (выбросов).
 - Преобразование простого голосования во взвешенное.
 - Обобщение на большое число классов.
 - Случайный лес.
 - Анализ смещения и вариации для простого голосования.
 - Смесь алгоритмов (квазилинейная композиция), область компетентности, примеры функций компетентности.
 - Выпуклые функции потерь. Методы построения смесей: последовательный и иерархический.
 - Построение смеси алгоритмов с помощью EM-подобного алгоритма.
- #### 5. Ранжирование
- Постановка задачи обучения ранжированию. Примеры.
 - Признаки в задаче ранжирования поисковой выдачи: текстовые, ссылочные, кликовые. TF-IDF. PageRank.
 - Критерии качества ранжирования: Precision, MAP, AUC, DCG, NDCG, pFound.
 - Ранговая классификация, OC-SVM.
 - Парный подход: RankingSVM, RankNet, LambdaRank.
- #### 6. Рекомендательные системы
- Задачи коллаборативной фильтрации, транзакционные данные и матрица субъекты - объекты.
 - Корреляционные методы user-based, item-based. Задача восстановления пропущенных значений. Меры сходства субъектов и объектов.
 - Латентные методы на основе би-кластеризации. Алгоритм Брегмана.
 - Латентные методы на основе матричных разложений. Метод главных компонент для разреженных данных (LFM, Latent Factor Model). Метод стохастического градиента.
 - Неотрицательные матричные разложения. Метод чередующихся наименьших квадратов ALS.
 - Модель с учётом неявной информации (implicit feedback).
 - Рекомендации с учётом дополнительных признаков данных. Линейная и квадратичная регрессионные модели, libFM.

- Измерение качества рекомендаций. Меры разнообразия (diversity), новизны (novelty), покрытия (coverage), догадливости (serendipity).

7. Тематическое моделирование

- Задача тематического моделирования коллекции текстовых документов.
- Вероятностный латентный семантический анализ PLSA. Метод максимума правдоподобия. EM-алгоритм. Элементарная интерпретация EM-алгоритма.
- Латентное размещение Дирихле LDA. Метод максимума апостериорной вероятности. Сглаженная частотная оценка условной вероятности.
- Небайесовская интерпретация LDA и её преимущества. Регуляризаторы разреживания, сглаживания, частичного обучения.
- Аддитивная регуляризация тематических моделей. Регуляризованный EM-алгоритм, теорема о стационарной точке (применение условий Каруша–Куна–Таккера).
- Рациональный EM-алгоритм. Онлайн-EM-алгоритм и его распараллеливание.
- Мультимодальная тематическая модель.
- Регуляризаторы классификации и регрессии.
- Регуляризаторы декоррелирования и отбора тем.
- Внутренние и внешние критерии качества тематических моделей.

8. Обучение с подкреплением

- Задача о многоруком бандите. Жадные и эpsilon-жадные стратегии. Метод UCB (upper confidence bound). Стратегия Softmax.
- Среда для экспериментов.
- Адаптивные стратегии на основе скользящих средних. Метод сравнения с подкреплением. Метод преследования.
- Постановка задачи в случае, когда агент влияет на среду. Ценность состояния среды. Ценность действия.
- Жадные стратегии максимизации ценности. Уравнения оптимальности Беллмана.
- Метод временных разностей TD. Метод Q-обучения.
- Градиентная оптимизация стратегии (policy gradient). Связь с максимизацией log-правдоподобия.
- Постановка задачи при наличии информации о среде в случае выбора действия. Контекстный многорукий бандит.
- Линейная регрессионная модель с верхней доверительной оценкой LinUCB.
- Оценивание новой стратегии по большим историческим данным.

9. Активное обучение

- Постановка задачи машинного обучения. Основные стратегии: отбор объектов из выборки и из потока, синтез объектов.
- Сэмплирование по неуверенности. Почему активное обучение быстрее пассивного.
- Сэмплирование по несогласию в комитете. Сокращение пространства решений.
- Сэмплирование по ожидаемому изменению модели.
- Сэмплирование по ожидаемому сокращению ошибки.
- Синтез объектов по критерию сокращения дисперсии.
- Взвешивание по плотности.
- Оценивание качества активного обучения.
- Введение изучающих действий в стратегию активного обучения. Алгоритмы ϵ -active и EG-active.
- Применение обучения с подкреплением для активного обучения. Активное томпсоновское сэмплирование.

Аннотации к рабочим программам дисциплин.

Направление: 01.04.02 Прикладная математика и информатика

Направленность: Науки о данных

Соревновательный DS

Цель дисциплины:

- обучить студентов особенностям применения методов машинного обучения соревновательных задачах по ds.

Задачи дисциплины:

- дать студентам нужные практические навыки в применении методов машинного обучения в соревновательных задачах по ds.

Перечень планируемых результатов обучения по дисциплине (модулю)

В результате освоения дисциплины обучающиеся должны

знать:

- специфику соревновательных задач;
- текущие state of the art методы.

уметь:

- строить самостоятельно решения для задач по ds;
- применять различные state of the art методы для увел. производительности. моделей;
- преобразовывать данные для увел. производительности. моделей;
- строить схемы валидации моделей.

владеть:

-навыками и инструментами необходимыми для успешного решения задач в хакатонах по ds.

Темы и разделы курса:

1. Продвинутая предобработка данных

Продвинутые техники предобработки данных, аугментаций данных, генераций новых признаков.

2. Оптимизация гиперпараметров

Продвинутые библиотеки для оптимизации гиперпараметров.

3. Продвинутая регуляризация моделей

Продвинутые техники регуляризации моделей.