

Онлайн-магистратура «Науки о данных и искусственный интеллект»

24 месяца

продолжительность программы

4 семестра

дисциплины вуза и Практикума



1 семестр

Базовая часть
(обязательная для всех треков)

Программирование на Python
Предподготовка и анализ данных
Машинное обучение с учителем
Базы данных и язык SQL
BI и методы визуализации данных
Теория вероятностей
Математическая статистика
A/B тестирование

2 семестр

Трек Data Engineering

Алгоритмы и структуры данных
Анализ бизнес-метрик и продвинутый SQL
Практическая бизнес-аналитика
Введение в системный анализ
Операционные системы семейства Unix
Практический хакатон
Soft Skills

Трек Data Science

Алгоритмы и структуры данных
Машинное обучение без учителя
Принципы разработки на Python
Операционные системы семейства Unix
Практический хакатон
Прогнозирование временных рядов
Soft Skills

ТАЙМЛАЙН ПРОГРАММЫ [1 / 2]

3 семестр



Трек Data Engineering

Работа с данными в хранилище
Автоматизация подготовки данных
Принципы проверки качества входных данных
Аналитические базы данных и хранилища
Потоковая обработка данных и их хранение в озере
Облачные технологии и проектный практикум
Практический хакатон
Soft Skills

Трек Data Science

СПЕЦИАЛИЗАЦИЯ ML

Разработка пайплайнов подготовки данных и обучения модели
Принципы и практики для улучшения базовой модели
Вывод модели машинного обучения в производственную среду
Практический MLOps
Рекомендательные системы
Uplift-моделирование
Практический хакатон
Soft Skills

Трек Data Science

СПЕЦИАЛИЗАЦИЯ CV

Работа с изображениями и текстами
Конволюционные нейронные сети и решаемые ими задачи
Задача детекции компьютерного зрения
Задача сегментации компьютерного зрения
Работы с видеопотоком
Генеративные модели машинного обучения
Практический хакатон
Soft Skills

Трек Data Science

СПЕЦИАЛИЗАЦИЯ NLP

Работа с изображениями и текстами
Классические и глубокие модели по работе с текстом
Задача преобразования последовательностей
Модели на базе архитектуры трансформер
Большие языковые модели
Генеративные модели машинного обучения
Практический хакатон
Soft Skills

4 семестр

Подготовка
выпускного проекта

Преддипломная практика
Выполнение и защита выпускной квалификационной работы

ТАЙМЛАЙН ПРОГРАММЫ [2 / 2]



Семестр 1

4 месяца

БАЗОВАЯ
ЧАСТЬ

Программирование на Python

Яндекс Практикум

Научитесь программировать на Python, изучите ключевые структуры данных, принципы ООП и обработку исключений. Поймёте, как писать чистый и эффективный код, работать с ключевыми библиотеками для анализа данных. Разберётесь в основах тестирования и инструментах контроля версий.

Инструменты

Git, GitHub, VS Code, Flake8, Pytest, bash, venv.

Примеры проектов

Создадите программу для регистрации и отслеживания продуктов в холодильнике, а также реализуете классическую игру «Змейка» на Python.

Предподготовка и анализ данных

Яндекс Практикум

Научитесь очищать данные от выбросов, пропусков и дубликатов, а также преобразовывать разные форматы данных.

Инструменты

Python, Pandas.

Примеры проектов

Проанализируете данные о клиентах банка и определите, какие показатели влияют на их кредитоспособность.

Машинное обучение с учителем

МИФИ

Познакомитесь с библиотекой scikit-learn и циклом создания модели машинного обучения. Узнаете, как подготовить данные для машинного обучения и проверить работу модели. Освоите линейные и нелинейные модели. Научитесь решать задачи регрессии и классификации. Освоите новые метрики качества модели и подходы для их улучшения. Изучите основы линейной алгебры.

Инструменты

Scikit-learn, SVM, kNN, NumPy.

Примеры проектов

Разработаете модель машинного обучения, которая поможет владельцам молочной фермы принимать эффективные решения.

Разберётесь, что такое база данных, и научитесь писать сложные запросы к ней. Поймёте, как считать ключевые бизнес-метрики через SQL-запросы и подключаться к базам данных из Python.

Инструменты

SQL, PostgreSQL.

Примеры проектов

Напишите ряд запросов разной сложности к базе данных, в которой хранится информация о венчурных инвесторах, стартапах и инвестициях в них.

С помощью Python и SQL подключитесь к базе данных, посчитаете и визуализируете ключевые метрики сервис-системы вопросов и ответов о программировании.

BI и методы визуализации данных

Познакомьтесь с принципами бизнес-аналитики и научитесь превращать данные в ценные бизнес-решения. Освойте методы сбора, обработки и визуализации данных. Изучите основные метрики, показатели эффективности и способы их интерпретации. Будете работать с инструментами BI, строить дашборды и автоматизировать отчётность.

Инструменты

DataLens, Superset.

Примеры проектов

Создайте интерактивный дашборд для анализа продаж и клиентского поведения.

Соберёте данные из базы, подготовите их для анализа и визуализируете ключевые метрики. Настроите автоматическое обновление отчёта и создадите удобный интерфейс с дашбордом в BI-системе. Проанализируете тренды, выявите аномалии и сформируете рекомендации для бизнеса.

Теория вероятностей

Познакомьтесь с основами теории вероятностей и их применением в анализе данных. Изучите законы распределения случайных величин, методы оценки вероятностей и основные теоремы. Научитесь работать с дискретными и непрерывными распределениями, вычислять математическое ожидание, дисперсию и ковариацию. Разберётесь в методах байесовского анализа, законах больших чисел и центральной предельной теореме.

Инструменты

Python, NumPy, SciPy, Statsmodels, Matplotlib, Seaborn.

Примеры проектов

Смоделируете систему предсказания спроса с учётом неопределённости.

Разработаете вероятностную модель, оценивающую спрос на товары с учётом сезонности и случайных факторов. Проведёте анализ исторических данных, построите распределения вероятностей, примените методы Монте-Карло для прогнозирования и оцените доверительные интервалы. Визуализируете результаты и сделаете выводы о стабильности модели.

Познакомьтесь с основами математической статистики и методами проверки гипотез. Научитесь работать с основными распределениями, рассчитывать доверительные интервалы и анализировать дисперсию. Освойте методы оценки параметров и проверки статистических зависимостей. Разберётесь в ключевых подходах к анализу данных, научитесь применять их в задачах машинного обучения и инженерии данных.

Инструменты

Python, NumPy, SciPy, Statsmodels, Matplotlib, Seaborn.

Примеры проектов

Проведёте статистический анализ факторов, влияющих на успеваемость студентов. Соберёте и обработаете данные о студентах, их занятиях, оценках и внешних факторах. Исследуете основные статистические закономерности, примените ANOVA, регрессионный анализ и тесты гипотез для выявления значимых факторов. Рассчитаете доверительные интервалы и проверите гипотезу о влиянии учебных привычек на итоговые результаты. Оформите выводы в виде аналитического отчёта с визуализациями.

A/B тестирование

В модуле вы закрепите полученные знания и освоите навыки подготовки и проведения A/B-тестирования. Разберётесь, как найти баланс между получением полезных результатов и затраченными ресурсами. Узнаете, как корректно использовать A/B-тестирование при принятии решений.

Инструменты

Python, NumPy, SciPy, Statsmodels, Matplotlib, Seaborn.

Примеры проектов

Проведёте A/B-тестирование нового продукта в аудиосервисе. Возьмёте на себя роль аналитика приложения для прослушивания аудиокниг. Определите целевую аудиторию для продвижения новой подписки на детские аудиокниги и спроектируете эксперимент. Рассчитаете ключевые метрики, выберете оптимальный объём выборки и оцените статистическую значимость результатов. Проведёте анализ данных, сформируете портрет пользователя и обоснуете бизнес-решение на основе A/B-теста.

4 месяца



Алгоритмы и структуры данных

Яндекс Практикум

Познакомьтесь с принципами применения алгоритмов для решения различных задач. Научитесь находить и реализовывать альтернативные способы решения, оценивать их эффективность. Разберётесь в ключевых структурах данных и освоите их применение для оптимального решения конкретных задач.

Инструменты

Python

Примеры проектов

Разработаете поисковую систему для текстовых документов.

Создадите поисковый алгоритм, который индексирует текстовые файлы и позволяет быстро находить нужные документы по ключевым словам. Реализуете основные структуры данных (хэш-таблицы, деревья, графы) для хранения индексов. Оптимизируете поиск с помощью префиксных деревьев (Trie) или обратного индекса. Оцените производительность алгоритма и сравните различные подходы. Оформите код в виде модульного проекта и протестируете его.

Анализ бизнес-метрик и продвинутый SQL

Яндекс Практикум

Познакомьтесь с ключевыми бизнес-метриками и научитесь анализировать их с помощью SQL. Освоите методы расчёта конверсий, когортного и ретеншн-анализа, юнит-экономики. Разберётесь в сложных SQL-запросах, включая оконные функции, CTE, подзапросы и агрегированные вычисления.

Инструменты

Метрики (CTR, Retention и др.)
SQL.

Примеры проектов

Создадите набор SQL-запросов для расчёта бизнес-метрик: конверсии, среднего чека, LTV и ретеншн-анализа.

Практическая бизнес-аналитика

Яндекс Практикум

Познакомьтесь с основами бизнес-анализа и подходами к решению бизнес-задач. Научитесь выявлять потребности заинтересованных сторон, формулировать требования и анализировать бизнес-процессы. Освоите методы сбора и обработки данных, построения моделей процессов и оценки эффективности решений. Разберётесь в инструментах бизнес-аналитики и научитесь применять их для оптимизации бизнеса и принятия стратегических решений.

Инструменты

Draw.io, UML, BPMN, IDEF0.

Примеры проектов

Поработаете с интервью с заказчиком по вопросам внутренней логистики. Проанализируете ASIS, разработаете TOBE. Проанализируете и разработаете автоматизацию процессов.



Познакомьтесь с принципами системного подхода и научитесь анализировать сложные системы с точки зрения их структуры, функций и взаимодействия компонентов. Освойте методы моделирования систем, выявления их слабых мест и оптимизации процессов. Узнаете, как проектировать архитектуру систем, учитывая технические и функциональные требования. Разберётесь в подходах к интеграции систем, управлению изменениями и обеспечению их устойчивости. Научитесь применять системное мышление для решения бизнес-задач.

Инструменты

Draw.io, UML, BPMN, IDEF0.

Примеры проектов

Разработаете архитектуру системы обработки данных для логистической компании. Проанализируете текущие бизнес-процессы компании и выявите узкие места. Смоделируете бизнес-процесс с помощью BPMN и сформируете требования к системе обработки данных. Создадите UML-диаграммы для описания структуры системы. Опишете варианты интеграции с внешними сервисами и разработаете концептуальную архитектуру будущей системы.

Операционные системы семейства Unix

МИФИ

Познакомьтесь с основами операционных систем семейства Unix и их архитектурой. Научитесь работать с файловой системой, процессами и управлением ресурсами. Освойте базовые команды Shell, автоматизацию задач с помощью Bash-скриптов, а также разберётесь в механизмах многозадачности, сетевого взаимодействия и безопасности. Научитесь настраивать окружение, управлять пользователями и анализировать работу системы.

Инструменты

Bash, Shell scripting, SSH.

Примеры проектов

Настроите и автоматизируете работу серверного окружения. Развернёте виртуальную машину с Linux, настроите пользователей, права доступа и сетевое взаимодействие. Создадите Bash-скрипт для автоматизации установки программ, мониторинга системы и логирования ошибок. Настроите SSH-доступ и разработаете план резервного копирования.

Практический хакатон

МИФИ

Программа дисциплины поможет углубить и применить на практике полученные теоретические знания.

Вы получите опыт работы с реальными данными от заказчиков. Сможете развить навыки организации процесса и самостоятельного принятия решений. Соберёте кейсы для портфолио и защитите свой проект.

Познакомьтесь с гибкими методологиями разработки, научитесь организовывать процессы в команде и адаптироваться к изменениям. Освойте бизнес-коммуникации, научитесь понимать заказчика и формулировать задачи. Разовьёте критическое мышление, научитесь анализировать информацию, аргументировать решения и правильно эскалировать проблемы. Освойте навыки презентации, научитесь понятно объяснять сложные идеи и работать с визуализацией данных. Будете эффективно взаимодействовать с командой и бизнесом, разберётесь в основах управления людьми и проектами.

Инструменты

Agile, Scrum, Kanban, управление процессами, выявление требований, переговоры, стресс-менеджмент, критическое мышление, презентация данных, сторителлинг, управление командой, лидерство.

Примеры проектов

Проанализируете существующие рабочие процессы команды, выявите узкие места и предложите способы оптимизации. Настроите систему управления задачами с учётом гибких методологий. Проведёте интервью с «заказчиком» для уточнения требований. Разработаете план улучшения коммуникации в команде и представите его в формате визуализированной презентации. Научитесь аргументировать решения перед бизнесом и анализировать их эффективность.

Семестр 2

DATA — TPEK
SCIENCE —

4 месяца

+

Алгоритмы и структуры данных

Яндекс Практикум

Познакомьтесь с принципами применения алгоритмов для решения различных задач. Научитесь находить и реализовывать альтернативные способы решения, оценивать их эффективность. Разберётесь в ключевых структурах данных и освоите их применение для оптимального решения конкретных задач.

Инструменты

Python

Примеры проектов

Разработаете поисковую систему для текстовых документов.

Создадите поисковый алгоритм, который индексирует текстовые файлы и позволяет быстро находить нужные документы по ключевым словам. Реализуете основные структуры данных (хэш-таблицы, деревья, графы) для хранения индексов. Оптимизируете поиск с помощью префиксных деревьев (Trie) или обратного индекса. Оцените производительность алгоритма и сравните различные подходы. Оформите код в виде модульного проекта и протестируете его.

+

+

Научитесь применять машинное обучение для наборов данных, не имеющих явного признака для предсказания. Выучите практики, позволяющие кластеризовать данные, уменьшать их размерность и находить аномалии.

Инструменты

Scikit-learn

Примеры проектов

Разделите покупателей магазина на сегменты и поймёте, чем эти сегменты между собой отличаются.

Принципы разработки на Python

Познакомьтесь с принципами разработки промышленного программного обеспечения. Научитесь проектировать будущие программы, работать с системами контроля версий и управлять зависимостями. Освойте практики тестирования, отладки и развёртывания.

Инструменты

Python, VS Code, flake8, Pylint.

Примеры проектов

Напишите программу, позволяющую автоматизировать рутинные операции на рабочем месте.

Операционные системы семейства Unix

Познакомьтесь с основами операционных систем семейства Unix и их архитектурой. Научитесь работать с файловой системой, процессами и управлением ресурсами. Освойте базовые команды Shell, автоматизацию задач с помощью Bash-скриптов, а также разберётесь в механизмах многозадачности, сетевого взаимодействия и безопасности. Научитесь настраивать окружение, управлять пользователями и анализировать работу системы.

Инструменты

Bash, Shell scripting, SSH.

Примеры проектов

Настройте и автоматизируете работу серверного окружения. Развернёте виртуальную машину с Linux, настройте пользователей, права доступа и сетевое взаимодействие. Создадите Bash-скрипт для автоматизации установки программ, мониторинга системы и логирования ошибок. Настройте SSH-доступ и разработаете план резервного копирования.

Программа дисциплины поможет углубить и применить на практике полученные теоретические знания.

Вы получите опыт работы с реальными данными от заказчиков. Сможете развить навыки организации процесса и самостоятельного принятия решений. Соберёте кейсы для портфолио и защитите свой проект.

Прогнозирование временных рядов

Временные ряды описывают, как параметры, например, объём потребления электроэнергии или количество заказов такси, меняются с течением времени. Вы научитесь анализировать ряды, искать тренды и выявлять сезонность. Узнаете, как создавать табличные данные и задачу регрессии из временных рядов.

Инструменты

Анализ временных рядов, задача регрессии.

Примеры проектов

Построите модель и спрогнозируете пиковые нагрузки в такси.

4 месяца



Работа с данными в хранилище

Яндекс Практикум

Познакомьтесь с основными технологиями, необходимыми для строительства хранилищ данных. Изучите различные подходы к их проектированию и научитесь работать с требованиями заказчика, выбирая оптимальные решения для конкретных задач. Разберётесь в концепции витрин данных, освоите их построение и обновление. Овладеете методами инкрементальной загрузки и транзакционной обработки. Узнаете, как оптимизировать запросы для повышения эффективности работы с данными.

Инструменты

SQL, PostgreSQL.

Примеры проектов

Построите витрину с инкрементальной загрузкой для аналитики аудитории интернет-магазина. Приведёте в порядок модель данных и осуществите миграцию в текущем хранилище.

Автоматизация подготовки данных

Яндекс Практикум

Автоматизируете пайплайн работы с данными: настройте автоматическую выгрузку данных из источников и освоите методы их регулярной и инкрементальной загрузки в базу данных.

Инструменты

Python, Airflow, PostgreSQL.

Примеры проектов

Построите для e-commerce проекта пайплайн автоматизированного получения, обработки и загрузки данных (ETL) от источников до витрины.

Принципы проверки качества входных данных

Яндекс Практикум

Разберёте принципы оценки качества данных и настройте шаги, необходимые для проверки.

Инструменты

Python, Airflow, PostgreSQL.

Примеры проектов

Оцените качество входных данных в своём проекте и настройте проверку.

Построите DWH с нуля на реляционной СУБД и познакомитесь с MongoDB как источником данных. Изучите организацию хранилища в Vertica, освоите базовые операции с данными и создадите простое хранилище данных в Vertica.

Инструменты

PostgreSQL,
MongoDB, Vertica,
Airflow.

Примеры проектов

Спроектируете и реализуете DWH для стартапа. Построите DWH для высоконагруженной системы данных с использованием Vertica.

Потоковая обработка данных и их хранение в озере

Познакомитесь с архитектурой Data Lake и принципами его наполнения данными из различных источников. Освоите обработку данных в MPP-системах и научитесь использовать PySpark и Airflow для управления обработкой данных. Разберётесь в особенностях потоковой обработки и построите свою стриминговую систему. В завершение курса создадите витрину данных с использованием real-time потоков.

Инструменты

Hadoop, MapReduce,
HDFS, Apache Spark
(PySpark), Kafka,
Spark Streaming.

Примеры проектов

Построите Data Lake для сбора данных о просмотрах видео. Автоматизируете загрузку и обработку данных в нём. Реализуете real-time обработку данных с помощью Kafka и Spark Streaming, чтобы анализировать активность пользователей, выявлять тренды и аномалии.

Облачные технологии и проектный практикум

Познакомитесь с принципами облачных технологий и разберётесь в инфраструктуре Yandex Cloud. Научитесь развёртывать DWH, работать с Redis и Kafka для хранения и потоковой обработки данных. Освоите контейнеризацию в Docker, управление микросервисами через Kubernetes и настройку сервисов с Helm. Напишите свой первый облачный сервис с PostgreSQL, подготовите окружение и логику работы. Разберётесь в визуализации данных с DataLens и научитесь создавать интерактивные дашборды. В завершение курса подготовите инфраструктуру для хранения и обработки данных в облаке.

Инструменты

Яндекс Облако,
Kubernetes,
Kubectl, Redis,
PostgreSQL.

Примеры проектов

Разработаете инфраструктуры хранения и обработки данных в облаке.

Программа дисциплины поможет углубить и применить на практике полученные теоретические знания.

Вы получите опыт работы с реальными данными от заказчиков. Сможете развить навыки организации процесса и самостоятельного принятия решений. Соберёте кейсы для портфолио и защитите свой проект.

Soft Skills

Познакомьтесь с гибкими методологиями разработки, научитесь организовывать процессы в команде и адаптироваться к изменениям. Освойте бизнес-коммуникации, научитесь понимать заказчика и формулировать задачи. Разовьёте критическое мышление, научитесь анализировать информацию, аргументировать решения и правильно эскалировать проблемы. Освойте навыки презентации, научитесь понятно объяснять сложные идеи и работать с визуализацией данных. Будете эффективно взаимодействовать с командой и бизнесом, разберётесь в основах управления людьми и проектами.

Инструменты

Agile, Scrum, Kanban, управление процессами, выявление требований, переговоры, стресс-менеджмент, критическое мышление, презентация данных, сторителлинг, управление командой, лидерство.

Примеры проектов

Проанализируете существующие рабочие процессы команды, выявите узкие места и предложите способы оптимизации. Настроите систему управления задачами с учётом гибких методологий. Проведёте интервью с «заказчиком» для уточнения требований. Разработаете план улучшения коммуникации в команде и представите его в формате визуализированной презентации. Научитесь аргументировать решения перед бизнесом и анализировать их эффективность.

4 месяца

СПЕЦИАЛИЗАЦИЯ ML

Разработка пайплайнов подготовки данных и обучения модели

Яндекс Практикум

Научитесь подготавливать и трансформировать данные. Сможете найти и устранить ошибки в данных.

Создадите модель и настроите DVC-пайплайн её обучения.

Инструменты

DVC, AirFlow, MLFlow.

Примеры проектов

Обучите модель, предсказывающую события на данных Яндекс Недвижимости.

Принципы и практики для улучшения базовой модели

Яндекс Практикум

Добавите новые признаки в данные. Будете версионировать запуски и считать метрики. Определите, где модель работает неправильно. Подберёте параметры модели, используя современные инструменты.

Инструменты

MLFlow, Optuna.

Примеры проектов

Отойдёте от baseline-модели, полученной ранее. Проведёте эксперименты, применяя продвинутый Feature Engineering и тюнинг гиперпараметров.

Вывод модели машинного обучения в производственную среду

Яндекс Практикум

Поймёте, как устроен процесс релиза модели в продакшен. Создадите микросервис с помощью FAST API для релиза модели в продакшен. Научитесь контейнеризировать микросервис с моделью с помощью Docker.

Инструменты

Docker, FAST API, Yandex Облако.

Примеры проектов

Создадите сервис на базе ML-модели и выкатите его в облачную инфраструктуру.

Познакомьтесь с принципами MLOps и основными этапами жизненного цикла моделей машинного обучения. Научитесь автоматизировать развёртывание, мониторинг и обновление моделей. Освойте инструменты CI/CD, контейнеризацию и оркестрацию. Разберётесь в методах управления данными, версионирования и масштабирования ML-систем.

Инструменты

Docker, Kubernetes, CI/CD, Prometheus, Grafana, Apache Airflow.

Примеры проектов

Разработаете и внедрите ML-модель для предсказания оттока клиентов в онлайн-сервисе. Оформите модель в виде API-сервиса, контейнизируйте её с помощью Docker, настройте CI/CD-пайплайн для автоматического развёртывания, а также реализуете мониторинг качества предсказаний с помощью инструментов Prometheus и Grafana. Выкатите сервис в продакшен.

Рекомендательные системы

Узнаете принципы коллаборативной фильтрации и контентных рекомендаций. Научитесь применять базовые модели для получения рекомендаций и проведёте валидацию и оценку их качеств. Будете использовать бустинг для улучшения рекомендаций и разберётесь в типовой архитектуре рекомендательной системы.

Инструменты

U2U, I2I, ALS, implicit, Catboost, Uvicorn.

Примеры проектов

Разработаете рекомендательную систему для улучшения предсказаний в сервисе Яндекс Музыка. Оформите сервис и выкатите его в продакшен.

Uplift-моделирование

Сможете объяснить внутренние принципы работы uplift-модели и самостоятельно её построить. Научитесь улучшать перфоманс модели через эксперименты. Подготовите uplift-модель к релизу в продакшен.

Инструменты

S/T-learner, scikit-uplift.

Примеры проектов

Решите задачу uplift-моделирования и решите, на какую аудиторию следует ориентироваться бизнесу и какой финансовый эффект это даст.

Практический хакатон

Программа дисциплины поможет углубить и применить на практике полученные теоретические знания.

Вы получите опыт работы с реальными данными от заказчиков. Сможете развить навыки организации процесса и самостоятельного принятия решений. Соберёте кейсы для портфолио и защитите свой проект.

Познакомьтесь с гибкими методологиями разработки, научитесь организовывать процессы в команде и адаптироваться к изменениям. Освойте бизнес-коммуникации, научитесь понимать заказчика и формулировать задачи. Разовьёте критическое мышление, научитесь анализировать информацию, аргументировать решения и правильно эскалировать проблемы. Освойте навыки презентации, научитесь понятно объяснять сложные идеи и работать с визуализацией данных. Будете эффективно взаимодействовать с командой и бизнесом, разберётесь в основах управления людьми и проектами.

Инструменты

Agile, Scrum, Kanban, управление процессами, выявление требований, переговоры, стресс-менеджмент, критическое мышление, презентация данных, сторителлинг, управление командой, лидерство.

Примеры проектов

Проанализируете существующие рабочие процессы команды, выявите узкие места и предложите способы оптимизации. Настроите систему управления задачами с учётом гибких методологий. Проведёте интервью с «заказчиком» для уточнения требований. Разработаете план улучшения коммуникации в команде и представите его в формате визуализированной презентации. Научитесь аргументировать решения перед бизнесом и анализировать их эффективность.

Семестр 3

DATA — ТРЕК
SCIENCE —

4 месяца

СПЕЦИАЛИЗАЦИЯ CV

Работа с изображениями и текстами

МИФИ

Поймёте принципы работы с текстовыми и медиаданными (изображения, видео). Разбёрете, чем их предобработка отличается от табличных данных. Познакомитесь с задачами, которые решаются в областях NLP и CV, и изучите базовые подходы к их решению.

Инструменты

OpenCV, Pillow, PyTorch, NLTK, spaCy.

Примеры проектов

Научитесь классифицировать изображения и тексты для разных бизнес-задач.

Конволюционные нейронные сети и решаемые ими задачи

МИФИ

Познакомьтесь с основами конволюционных нейронных сетей (CNN) и их архитектурами. Научитесь применять CNN для разного рода задач. Освойте методы увеличения данных (data augmentation). Проведёте дообучение и валидацию современных свёрточных архитектур, таких как EfficientNet.

Инструменты

OpenCV, PyTorch, Albumentations.

Примеры проектов

Обучите конволюционную нейронную сеть (CNN) на датасете изображений дорожных знаков, используя предобученные архитектуры.

Задача детекции компьютерного зрения

МИФИ

Познакомьтесь с основами задачи детекции в компьютерном зрении и её ключевыми методами. Научитесь использовать свёрточные нейросети и современные архитектуры, такие как YOLO и Faster R-CNN. Проведёте обучение моделей на размеченных данных, освоите техники аугментации и методы повышения точности детекции. Разберётесь в процессах валидации, оценки качества и оптимизации производительности детекторов.

Инструменты

OpenCV, PyTorch, Albumentations, YOLO, Faster R-CNN.

Примеры проектов

Разметите датасет с примерами объектов, таких как оставленные предметы, транспортные средства, люди. Обучите и протестируете модель на основе современных архитектур, например, YOLO или Faster R-CNN.

Задача сегментации компьютерного зрения

МИФИ

Познакомьтесь с основными методами сегментации изображений, включая классические алгоритмы и современные нейросетевые подходы. Научитесь применять свёрточные нейросети и трансформеры для сегментации, а также проводить валидацию и оценку качества предсказаний. Разберётесь в архитектурах популярных моделей, таких как U-Net, DeepLab и Mask R-CNN, и научитесь адаптировать их к различным задачам компьютерного зрения.

Инструменты

Torchvision / KerasCV, Hugging Face, Transformers.

Примеры проектов

Разработаете модель сегментации для автоматического выделения дорожной разметки на спутниковых снимках.

Познакомьтесь с основами обработки видеопотока и ключевыми алгоритмами компьютерного зрения для работы с видео. Научитесь извлекать и анализировать кадры, детектировать и отслеживать объекты, а также применять методы оптического потока. Освойте свёрточные нейросети и трансформеры для анализа видео, разберётесь в принципах кодирования и потоковой обработки. Реализуете свой первый проект по анализу видеопотока в реальном времени.

Инструменты

FFmpeg, PyTorch, YOLO.

Примеры проектов

Разработаете систему автоматического детектирования и отслеживания объектов на видеопотоке с уличных камер.

Генеративные модели машинного обучения

Познакомьтесь с ключевыми архитектурами генеративных моделей, таких как GAN. Поймёте принципы создания и обучения моделей для генерации реалистичных изображений, стилизации и улучшения качества картинок. Освойте методы оценки качества сгенерированных изображений и разберётесь в применении генеративных моделей в различных задачах, от художественного рендеринга до повышения разрешения.

Инструменты

StyleGAN, Stable Diffusion.

Примеры проектов

Разработаете генеративную модель для улучшения качества изображений в старых фотографиях.

Практический хакатон

Программа дисциплины поможет углубить и применить на практике полученные теоретические знания.

Вы получите опыт работы с реальными данными от заказчиков. Сможете развить навыки организации процесса и самостоятельного принятия решений. Соберёте кейсы для портфолио и защитите свой проект.

Познакомьтесь с гибкими методологиями разработки, научитесь организовывать процессы в команде и адаптироваться к изменениям. Освойте бизнес-коммуникации, научитесь понимать заказчика и формулировать задачи. Разовьёте критическое мышление, научитесь анализировать информацию, аргументировать решения и правильно эскалировать проблемы. Освойте навыки презентации, научитесь понятно объяснять сложные идеи и работать с визуализацией данных. Будете эффективно взаимодействовать с командой и бизнесом, разберётесь в основах управления людьми и проектами.

Инструменты

Agile, Scrum, Kanban, управление процессами, выявление требований, переговоры, стресс-менеджмент, критическое мышление, презентация данных, сторителлинг, управление командой, лидерство.

Примеры проектов

Проанализируете существующие рабочие процессы команды, выявите узкие места и предложите способы оптимизации. Настроите систему управления задачами с учётом гибких методологий. Проведёте интервью с «заказчиком» для уточнения требований. Разработаете план улучшения коммуникации в команде и представите его в формате визуализированной презентации. Научитесь аргументировать решения перед бизнесом и анализировать их эффективность.

Семестр 3

DATA — ТРЕК
SCIENCE —

4 месяца

СПЕЦИАЛИЗАЦИЯ NLP

Работа с изображениями и текстами

МИФИ

Поймёте принципы работы с текстовыми и медиаданными (изображения, видео). Разбёрете, чем их предобработка отличается от табличных данных. Познакомитесь с задачами, которые решаются в областях NLP и CV, и изучите базовые подходы к их решению.

Инструменты

OpenCV, Pillow, PyTorch, NLTK, spaCy.

Примеры проектов

Научитесь классифицировать изображения и тексты для разных бизнес-задач.

Классические и глубинные модели по работе с текстом

МИФИ

Познакомьтесь с классическими и глубинными методами обработки текста. Научитесь работать с моделями на основе статистических подходов, векторных представлений слов и нейросетевых архитектур. Освойте различные рекуррентные сети. Проведёте анализ текстов и их классификацию. Разберётесь в принципах работы современных NLP-моделей и их применении в реальных задачах.

Инструменты

sраCy,
TF-IDF,
RNN,
LSTM,
BeautifulSoup.

Примеры проектов

Разработаете систему автоматической классификации отзывов для онлайн-магазина. Разработаете и обучите NLP-модель для автоматической классификации отзывов клиентов.

Задача преобразования последовательностей

МИФИ

Познакомьтесь с основными подходами к моделированию последовательностей в NLP. Научитесь работать архитектурами моделей для машинного перевода, текстовой суммаризации и генерации. Разберётесь в механизме внимания и обучите модели для обработки последовательностей.

Инструменты

PyTorch,
Attention,
seq2seq.

Примеры проектов

Разработаете модель для автоматической суммаризации новостных статей.

Модели на базе архитектуры трансформер

МИФИ

Познакомьтесь с ключевыми принципами архитектуры трансформеров и их применением в задачах обработки естественного языка. Научитесь работать с предобученными моделями, такими как BERT, адаптировать их под конкретные задачи с помощью дообучения.

Инструменты

Hugging Face,
Transformers,
PyTorch.

Примеры проектов

Используете предобученную модель BERT, адаптируете её под ваш корпус данных.

Познакомьтесь с принципами работы больших языковых моделей (LLM) и их архитектурами, такими как GPT. Научитесь дообучать и адаптировать предобученные модели под специфические задачи.

Инструменты

LLM,
Fine-Tuning.

Примеры проектов

Обучите и доработаете предобученную LLM-модель на юридических текстах, реализуете механизм обработки длинных контекстов.

Генеративные модели машинного обучения

Познакомьтесь с принципами работы генеративных моделей в обработке естественного языка. Научитесь обучать и дообучать трансформеры, создавать текстовые модели для генерации, суммаризации и дополнения текста. Освойте методы оценки качества генерации.

Инструменты

Hugging Face,
PyTorch,
BLEU.

Примеры проектов

Разработаете текстовую модель для автоматического написания кратких аннотаций к новостям.

Практический хакатон

Программа дисциплины поможет углубить и применить на практике полученные теоретические знания.

Вы получите опыт работы с реальными данными от заказчиков. Сможете развить навыки организации процесса и самостоятельного принятия решений. Соберёте кейсы для портфолио и защитите свой проект.

Познакомьтесь с гибкими методологиями разработки, научитесь организовывать процессы в команде и адаптироваться к изменениям. Освойте бизнес-коммуникации, научитесь понимать заказчика и формулировать задачи. Разовьёте критическое мышление, научитесь анализировать информацию, аргументировать решения и правильно эскалировать проблемы. Освойте навыки презентации, научитесь понятно объяснять сложные идеи и работать с визуализацией данных. Будете эффективно взаимодействовать с командой и бизнесом, разберётесь в основах управления людьми и проектами.

Инструменты

Agile, Scrum, Kanban, управление процессами, выявление требований, переговоры, стресс-менеджмент, критическое мышление, презентация данных, сторителлинг, управление командой, лидерство.

Примеры проектов

Проанализируете существующие рабочие процессы команды, выявите узкие места и предложите способы оптимизации. Настроите систему управления задачами с учётом гибких методологий. Проведёте интервью с «заказчиком» для уточнения требований. Разработаете план улучшения коммуникации в команде и представите его в формате визуализированной презентации. Научитесь аргументировать решения перед бизнесом и анализировать их эффективность.

Семестр 4

4 месяца

Преддипломная практика

МИФИ

Выполнение и защита выпускной квалификационной работы

МИФИ