



Программа магистратуры

Анализ данных в науке и предпринимательстве 01.04.02.10

Институт космических и информационных
технологий

Кафедра прикладной математики и
анализа данных, ИКИТ

Кафедра высшего ресторанного
менеджмента, Институт гастрономии

Партнеры программы



Институт вычислительного моделирования
Сибирского отделения Российской академии наук



Компания «Ар Ди Сайенс»



Группа компаний «КРИС»



Торговый холдинг «Командор»

Профессии в мире Big Data: Инженер, Аналитик или Сайентист?

Наш
выпускник

**Дата-сайентист:
Универсальный эксперт**

Сочетает навыки аналитика с умением разрабатывать модели машинного обучения для предсказания будущего.

**Дата-аналитик:
Исследователь трендов**

Работает с данными с прошлым и настоящим, выделяя закономерности и создавая визуализации в Tableau.

**Дата-инженер:
Мастер инфраструктуры**

Занимается ETL-процессами: извлечением, преобразованием и загрузкой данных в хранилища с помощью SQL и Hadoop.



Python, Matlab, Нейросети, ML-модели, C++



- сильная математическая база,
- продвинутые навыки программирования
- гибкие навыки

От бизнес-кейсов до науки

Создает интеллектуальные системы рекомендаций, прогнозирует риски и моделирует сложные физические процессы.

Разные цели —
разные инструменты



Сравнение ключевого инструментария ролей

Роль	Основной стек технологий
Дата-инженер	Hadoop, Spark, Redis, SQL, MapReduce, NoSQL
Дата-аналитик	SQL, R, Python, Tableau, Datawrapper, Теория вероятностей, Статистика
Дата-сайентист	Python, ML-модели, Нейросети, C++, Matlab

Трек «Анализ данных и машинное обучение»

1. Постановка задачи и сбор данных

Формирование требований и извлечение данных из источников (ETL-процессы).

Системы хранения, обработки и управления данными

4. Внедрение моделей в программные продукты

Развертывание готовых решений в ИТ-инфраструктуру и создание работающих приложений.

Разработка приложений машинного обучения

2. Построение и отладка моделей ML

Выбор алгоритмов машинного обучения, проектирование архитектуры и обучение моделей.

Основы машинного обучения

Продвинутые методы машинного обучения

3. Оценка качества полученных моделей

Тестирование точности и эффективности алгоритмов на соответствие бизнес-целям.

Нейронные сети

Прикладные задачи анализа данных

Предпринимательский трек

Чему вы научитесь на треке

Выявление потребностей рынка

Поиск рыночных ниш и анализ запросов потребителей

Коммерциализация продуктов

Превращение разработок в области ИИ в прибыльные рыночные решения

Стартап-проекты и ИТ решения

Предпринимательство в сфере технологий искусственного интеллекта

Стартап как диплом

Возможность защитить выпускную работу в формате реального бизнес-проекта или ИТ-решения.

Маркетинг инноваций

Монетизация больших данных

Исследовательский трек

- решение прикладных задач анализа данных:
 - предложенные партнерами программы бизнес-задачи
 - проекты в рамках исследовательских направлений кафедры ПМиАД (прикладной искусственный интеллект для анализа образовательных данных, стохастические модели в финансах и медицине, нечеткая и многоагентная логика)
 - предложенные самими магистрантами задачи
- участие в проведении фундаментальных исследований ИВМ СО РАН
- семинары по анализу данных: изучение и обсуждение современных идей и направлений развития в области машинного обучения

Проектно-технологические практики

Научно-исследовательская работа

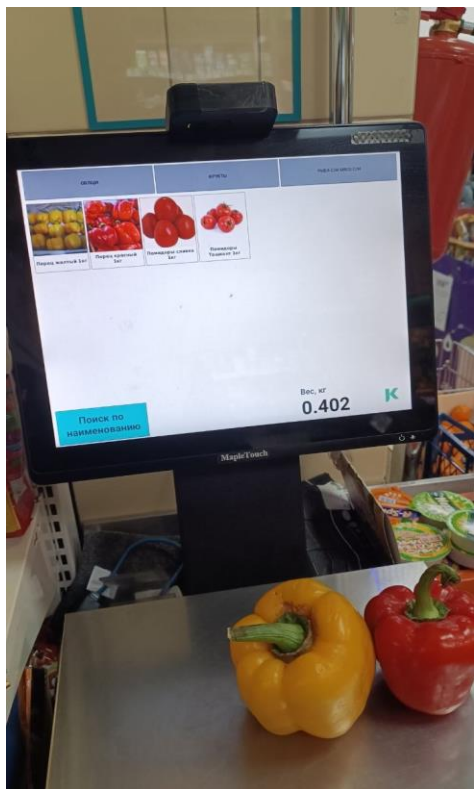
Научно-исследовательский семинар



Семинар ИКИТ по анализу данных

Пример магистерского исследования по решению бизнес-задачи

Никитин Вадим, магистрант 1 курса. Тема «Система компьютерного зрения для автоматизации процессов в розничной торговле» (задача от ООО «Командор»)

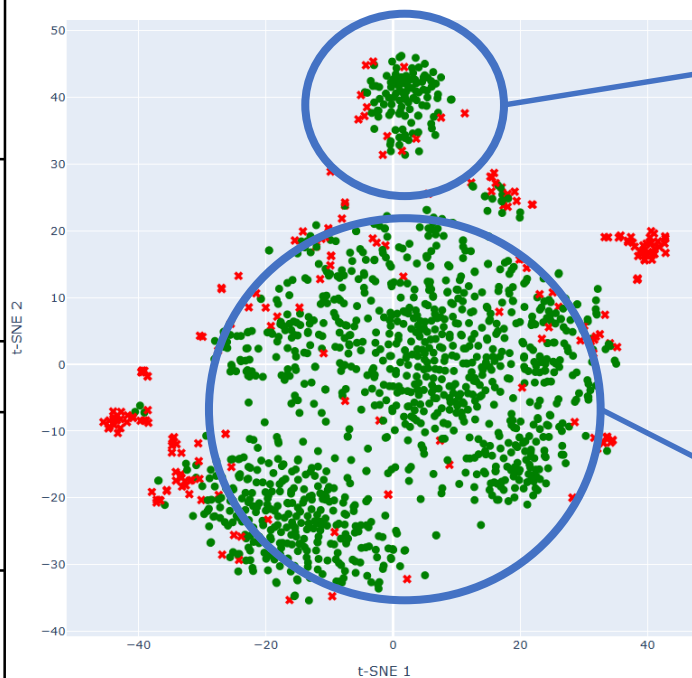


Получение изображения с камеры «Умных весов»

товар	фото	Эмбединг
Перец красный		\vec{e}_1
Перец красный		\vec{e}_2
• • •	• • •	• • •
Конфеты «Темная ночь»		\vec{e}_{n-1}
Огурец		\vec{e}_n

Векторная база данных

Очистка: осталось 1191, удалено 210



Кластеризация эмбедингов



Интерпретация принадлежности эмбединга кластеру

Пример магистерского исследования по одному из научных направлений кафедры

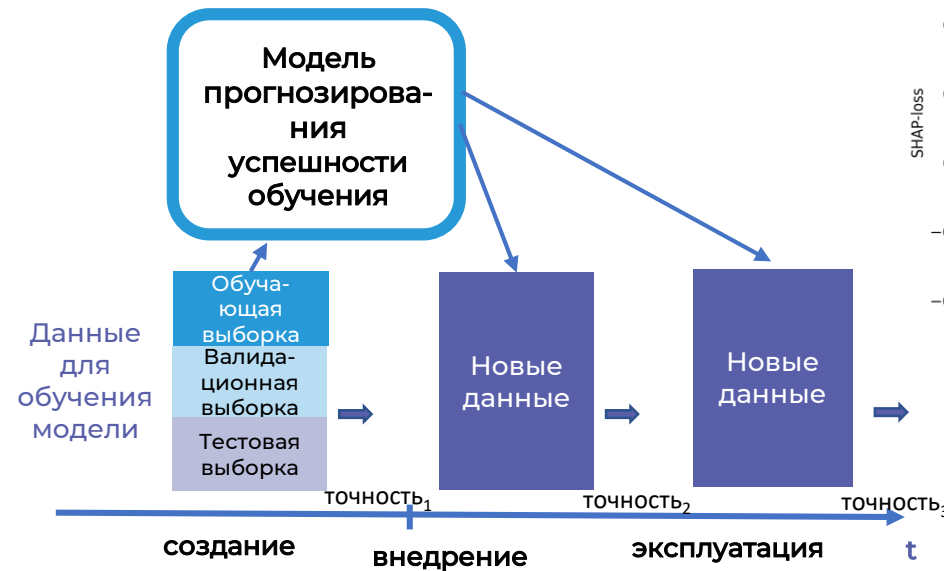
Боков Адам, магистрант 2 курса. Тема «Робастность моделей прогнозирования успешности обучения» (исследовательское направление «Прикладной искусственный интеллект для анализа образовательных данных»)



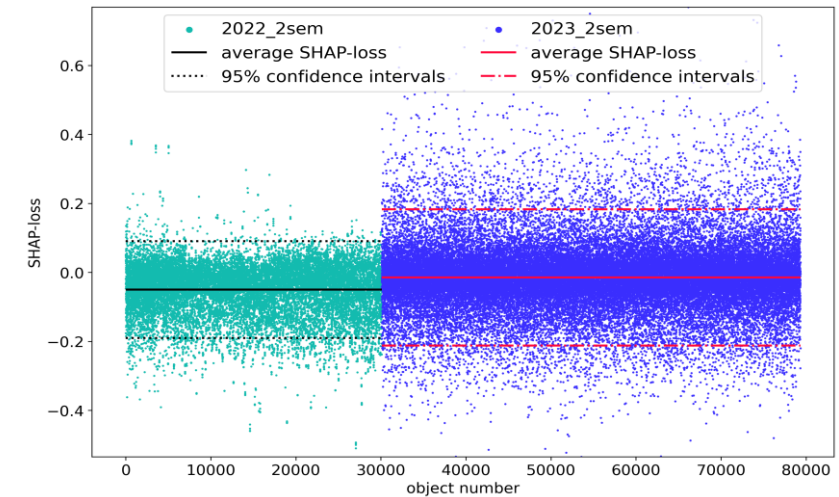
Цифровой след с электронной обучающей платформы СФУ



Сбор и обработка данных



Обучение модели прогнозирования, мониторинг качества ее работы



$$\varphi(j) = \sum_{S \subseteq \{1, \dots, p\} \setminus \{j\}} \frac{|S|!(p-|S|-1)!}{p!} (val(S \cup \{j\}) - val(S)).$$

Оценка устойчивости ее работы методами интерпретируемого ИИ и стресс-тестирования

Поступление на программу

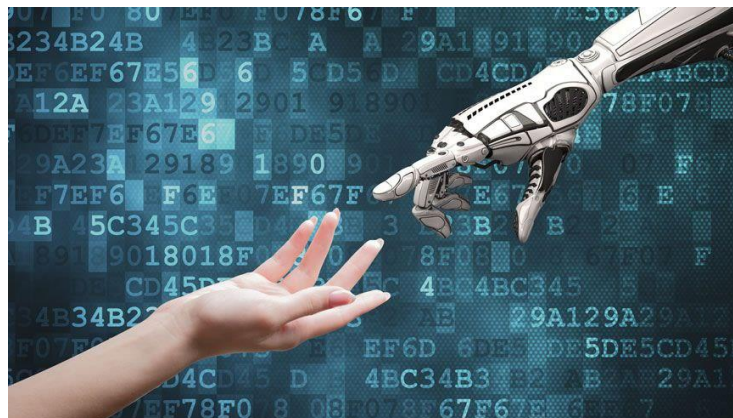
Анализ данных в науке и предпринимательстве

- Бюджетных мест 12, платных – 2
- Учебная нагрузка 15-17 академических часов в неделю
- Часть занятий проходит в электронной среде
- Вступительное испытание - устное собеседование



Сайт программы:

- учебный план и описание программы
- программа подготовки к собеседованию
- информация о преподавателях



VK программы:

- узнать больше о программе
- задать вопросы
- найти полезные материалы по машинному обучению