



ДАНИЛ ШАЛАГИН

ML-разработчик

✉ shalagin.danil@mail.ru

📧 @Dani1 Sh

🐙 Sh41lD4n1n

Пройденные курсы

Обработка данных: Big Data, Data Bases

Машинное обучение: Trees, KNN, logistic regression

Глубокое обучение: модели для анализа текста (BERT, ELMo, Transformer), GAN

Компьютерное зрение: YOLO, RCNN, image preprocessing

HardSkills

🐍 Python: numpy, scipy, pandas, sklearn, pytorch, seaborn

✂ Linux Docker Git Apache Airflow Spark Hadoop

✂ Databases: PostgreSQL

✂ Other: C++

Soft Skills

Отвественный,
Командный работник,
целеустремленный

Дополнительно

language, English (Upper Intermediate)

О себе

Инженер в области машинного обучения, имеющий опыт работы над проектами CV, ML, NLP, и проектами связанных с анализом данных. Во время работы над групповыми проектами и стажировками, я научился выстраивать коммуникации в команде, стремился доводить проекты до значимых результатов, управлял проектами и внимательно относился к правкам от руководителей.

Во время обучения я получил навыки обработки изображений (python, openCV), работал с большими данными (с Spark, Hadoop, SQL), изучал основы ML и Deep learning, занимался обработкой текста: сбор и обработка (beautiful soap, nltk). Также во время обучения создавал проекты в областях Машинного зрения и Машинного обучения.

Образование

2019 - 2023 Информатика и вычислительная техника Innopolis University
Диплом: Локальные методы с предобуславливанием для задач федеративного обучения

Опыт

2021 - 2022 ML-разработчик, стажер Innopolis University
– Разработал ML-модель для подсчета количества животных на изображениях.
– Создал симуляции (модель очереди, модель перемещения клиентов и модель мнений). Применял модели машинного обучения для прогнозирования параметров моделирования
– Создание структуры базы данных. Применил Apache Airflow для преобразования данных.

Проекты

Применение федеративного обучения нейронных сетей 2023 🏆
– Создана модель федеративного обучения (на PyTorch) на основе моделей ResNet, VGG, Logistic regression, которые классифицируют изображения из набора данных CIFAR-50 и пользовательского набора данных и используют прекодинтеры (Adam).
– Реализованная модель удовлетворяет теоретическому определению и условиям Federated learning.

• YOLO based image search 2022 🏆
– Применили и внедрили модель глубокого обучения YOLO для поиска животных на картинке. После настройки и обучения наша модель имела точность 0,7, запоминание 0,8 и предположение 0,9.

• Обработка текста 2022
– Собирали и производили поиск по тексту с beautiful soap
– Проводил обработку текста (токенизация, стемминг, лемматизация, word embeddings)

• CNN for iris center selection 2021 🏆
– Предварительно обработанный набор данных лиц таким образом, чтобы он содержал только изображения глаз в серой шкале.
– Построили CNN-модель для предсказания центра радужной оболочки глаза.