

ДМИТРИЙ

+7 (964) 969-58-38 | dimok.shalashov@gmail.com | telegram: @DmitriyShalashov | github: DmitriyShalashov

ОБРАЗОВАНИЕ

Университет

Бакалавриат, 3 курс

НИУ ВШЭ Бизнес-информатика, Москва

2023-2027

ПРОЕКТЫ

Генерация цитат | Технологии: PyTorch, RNN

2

- Генерация цитат, используя датасет цитат из книг писателей с Kaggle(на английском)
- В данном проекте я реализовал собственный слой RNN, используя базовые модули nn из библиотеки Torch и обучил нейросеть на датасете. Также сравнил качество и скорость работы собственного слоя и встроенного слоя RNN в Torch
- Ссылка: <https://github.com/DmitriyShalashov/RNN-generate-quotes>

Кредитный скоринг физических лиц | Технологии: Scikit-learn, CatBoost, Optuna

2

- Предсказание вероятности одобрения кредита
- В данном проекте используются модели линейной классификации, градиентного бустинга и случайного леса. В ходе данного проекта я оценивал качество различных моделей, используя метрики F1, recall, precision, accuracy
- Ссылка: <https://github.com/DmitriyShalashov/Credit-scoring>

Прогноз цен на недвижимость | Технологии: TensorFlow, LSTM, Self-attention

2

- Прогноз цен на недвижимость, по данным с Kaggle за 2017-2023 год для определенного региона
- В этом проекте я использовал LSTM модель, добавляя к ней механизм самовнимания(Self-attention). Проект состоит из базовых элементов Tensorflow и функции самовнимания
- Ссылка: <https://github.com/DmitriyShalashov/LSTM-House-sales>

Рекомендации для музыкальной платформы | Технологии: CF, ALS

2

- Рекомендации музыкальных треков пользователю, через сравнение векторов пользователей по метрикам
- Для данной задачи я использовал модели ALS и коллаборативную фильтрацию. Данные представляли собой таблицу users-tracks(добавление трека пользователем). В качестве метрики качества я использовал NDCG
- Ссылка: <https://github.com/DmitriyShalashov/RecSys-Music>

КУРСЫ

- Нейронные сети и обработка текста, Stepik
- MLOps. Начало, Stepik
- Основы машинного обучения, ФКН ВШЭ
- Time-series, Kaggle
- Intro to Deep Learning, Kaggle

ХАКАТОНЫ

- Reclamoton MTC Travel 2024 - лучшее Data science решение
- Кейс-чемпионат от Ахенix 2025 - составление маршрутов и ранжирование с персонализацией
- Кейс-чемпионат от Сбер 2025 - анализ и прогнозирование страховых выплат

ТЕХНИЧЕСКИЕ НАВЫКИ

Языки: Python

Инструменты разработчика: Jupiter, SQL, Docker, HuggingFace, ClearML, PySpark

Библиотеки: PyTorch, Tensorflow, CatBoost, Scikit-learn, Optuna, Pandas, NumPy

Навыки: ML, DL, Data Science