

ДАНИИЛ МИРОНОВ

JUNIOR DATA SCIENTIST

КОНТАКТЫ

Телефон: +79999718671
Telegram: @mdaniel42
Почта: daniiomir@yandex.ru
Github: github.com/daniiomir
Kaggle: kaggle.com/danielmironov

О СЕБЕ

24 года. Хорошие soft skills.
Технический английский на уровне чтения документации и статей.

НАВЫКИ

C Python SQL
Numpy Pandas SciPy
PyTorch
OpenCV
Unix Git Docker
Catboost, XGBoost, LightGBM
Classic algorithms (regressions, decision trees, SVM, k-means)

ПЕТ-ПРОЕКТЫ

Мини библиотека глубокого обучения в стиле pytorch с нуля
https://github.com/daniiomir/mutilayer_perceptron

Присутствуют: слои (Dense, Dropout), активации (ReLU, Tanh, SoftMax, Sigmoid), оптимизаторы (Momentum, RMSProp, Adam), функции потерь (CrossEntropy, MSE)

ОПЫТ

Стажер data scientist в ПАО "СБЕРБАНК" в команде DSHR с 08.07.2020 по 15.02.2021

Работал над двумя задачами.

1) Модель OCR

Перебирал и обучал комбинации архитектур для улучшения качества SOTA OCR модели. Добавлял новые фичи в pipeline модели (определение градуса наклона документа и быстрой классификации документа, test-time augmentation). Проверял качество разметки датасетов.

Рефакторил архитектуру проекта.

2) Модель оттока сотрудников

Дано: 50 ГБ захэшированных табличных данных о почтовой переписке сотрудников.

Необходимо было проверить возможность построения модели с ROC-AUC 0.7-0.8 и выяснить причину высокой разделимости данных. Провел EDA, почистил данные, выделил частотные признаки. Построил модель с качеством ROC-AUC ~0.75. Выявил важные признаки и проверил гипотезу о присутствии данных из будущего.

ОБРАЗОВАНИЕ

Студент

School 21 от Сбербанка (L'ecole 42) С 20.04.2019 по настоящее время

Разработка на C/Python

Принципы написания чистого и эффективного кода

Работа над собственными и командными проектами

Бакалавр

РАНХиГС (юриспруденция)

С 2015 по 2019