

ДАНИИЛ МИРОНОВ

JUNIOR DATA SCIENTIST

КОНТАКТЫ

Телефон: +79999718671

Telegram: @mdaniel42

Почта: daniilomir@yandex.ru

Github: github.com/daniilomir

Kaggle: kaggle.com/danielmironov

НАВЫКИ

C

Python

SQL

Unix

Git

OpenCV

Docker

Numpy, Pandas, SciPy

Catboost, XGBoost, LightGBM

Classic algorithms (regressions, decision trees, SVM, k-means)

PyTorch

О СЕБЕ

24 года. Хорошие soft skills.

Технический английский на уровне чтения документации и статей.

ПРОЕКТЫ

Определение заболеваний на листьях cassava

<https://www.kaggle.com/danielmironov/cassava-leaf-disease-classification-pytorch/>

Предсказание будущих цен на товары

<https://www.kaggle.com/danielmironov/stacking-models/>

ОПЫТ

Стажер data scientist в ПАО "СБЕРБАНК" в команде DSHR

с 08.07.2020 по 15.02.2021

Работал над двумя задачами.

1) Проект распознавания фотографий документов.

Перебирал и обучал комбинации архитектур для улучшения качества SOTA OCR модели. Добавлял новые фичи в pipeline модели (модели для определения градуса документа и быстрой классификации документа, test-time augmentation). Проверял качество разметки датасета с помощью поиска большого количества пересечений полей документа.

2) Проверка качества модели оттока сотрудников

Было дано 50 ГБ захэшированных табличных данных о почтовой переписке сотрудников. Необходимо было проверить возможность построения модели с качеством ROC-AUC 0.7-0.8 и выяснить причину слишком хорошей разделимости данных. Преобразовал данные в датасет и провел EDA. 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