

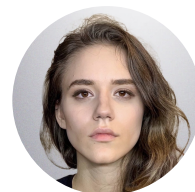
# ФЕДИНА НАДЕЖДА

@ fedina.no@phystech.edu

+7 (963) 625-17-63

Москва

<https://github.com/Vento13>



## ОПЫТ РАБОТЫ

### Data Science Intern – Сбер.IT

Решала задачу подбора персональных рекомендаций: формулирование гипотез, сбор данных, генерация таргетов, препроцессинг данных, написание ML-модели, настройка гиперпараметров, скоринг и анализ результатов.

Декабрь 2021 - Февраль 2022

## КВАЛИФИКАЦИЯ

### Математическая статистика

- Институтский курс "Теория вероятности и начала статистики"
- Курс "Основы статистики" от Bioinformatics Institute на Stepic  
<https://stepik.org/cert/1084833>

### Machine learning/Python

- "Математика и Python для анализа данных" (Яндекс & МФТИ):  
<https://www.coursera.org/account/accomplishments/certificate/MK8VZR5ZWDSW>
- "Обучение на размеченных данных" (Яндекс & МФТИ):  
<https://www.coursera.org/account/accomplishments/certificate/K6DABMEPPBFL>
- "Поиск структуры в данных" (Яндекс & МФТИ):  
<https://coursera.org/share/9b4f1191cf43737451ef9fb6bd043719>
- Курс лекций "Математические основы машинного обучения" К.В. Воронцова
- Факультативный курс "Машинное обучение" МФТИ

### SQL

- Факультетский курс по SQL
- Курс "SQL for Data Science" от Калифорнийского университета:  
<https://www.coursera.org/account/accomplishments/certificate/JA753A6R6CB2>

## ПРОЕКТЫ

### Kaggle: House Prices - Advanced Regression Techniques

- Score: 0.12015. На ноябрь 2021 - top 7%
- Предобработка данных и признаков, подбор гиперпараметров байесовскими методами оптимизации, построение метарегрессора над XGBoost, Gradient Boosting, Lasso, Ridge, Elastic regressor
- <https://github.com/Vento13/House-Prices-Advanced-Regression-Techniques>

### Recipes

Тематическое моделирование на Gensim

<https://github.com/Vento13/Recipes>

## ОБРАЗОВАНИЕ

### МФТИ

Прикладная математика и физика

2018 - настоящее время

## НАВЫКИ

Python:

Scikit-learn Numpy Pandas  
Matplotlib SciPy Seaborn  
XGBoost PyTorch Hadoop  
PySpark Hive

Machine learning/Data science:

Reinforcement learning  
Data preprocessing  
Визуализация данных  
Метрические методы классификации  
Линейные методы классификации  
Байесовские методы классификации  
Методы восстановления регрессии  
Кластеризация Ранжирование  
Тематическое моделирование  
Ансамблевые методы  
Полносвязные нейронные сети CNN  
RNN

Git Team work Agile Jira