

Марк Андреев, Senior Machine learning engineer

Email: mark.andreev@gmail.com

Телефон: +79154690400

Github: <https://github.com/mrk-andreev>

Образование. Магистр, ВМК МГУ (Московского государственного университета имени М. В. Ломоносова)

Технологический стек

- **Java**. Spring: MVC, Data, AMQP, Kafka, Security, State machine, Apache Camel, Vert.x, GraalVM
- **Python**. Pandas, Scikit-learn, Matplotlib, XGBoost, LightGBM, Catboost, Tornado, Flask, FastAPI.
- PostgreSQL, Clickhouse, Kafka, MongoDB, RabbitMQ, Redis, Keycloak, Prometheus, Docker, Kubernetes, Helm, Airflow, Spark, Cassandra, Hadoop. // Git, Github, Gitlab, Gitlab CI
- **Cloud**. AWS: EC2, S3, RDS, CloudFront, SQS, SNS, Lambda, Batch, IAM; **Azure**: VM, Blob.

Опыт работы

Senior Machine Learning Engineer, Сентябрь 2022 - настоящее время

[Conundrum.AI](#), // Лондон, Великобритания

- Внедрил низкоуровневую оптимизацию для feature store поверх Kafka и Clickhouse (проекции, планировщик запросов на уровне приложения) | Ускорил один из самых популярных запросов в 3 раза
- Внедрил нагрузочные регрессионные тесты (Gatling) | Покрывающие 80% запросов ключевого сервиса, что значительно снижает регрессию производительности
- Реализовал оптимизацию подписок через прокси на Kafka (java 21 virtual threads, shared subscription) | Снизил нагрузку на CPU в 5 раз
- Реализовал требования безопасности для AI Платформы (аудит, L4 network policies, L7 network filter) | Удовлетворил требования ИБ на уровне сети
- Покрыл сервисы метриками производительности (Prometheus, Grafana, Alerts) | Снизил время расследования инцидента в 3 раза

Senior Machine Learning Engineer, Сентябрь 2019 - Сентябрь 2022

[Conundrum.AI](#), // Москва, Россия

- Мигрировал feature store на Kafka & Clickhouse (Колоночная OLAP БД) | Увеличил скорость запросов в 15 раз
- Реализовал низкоуровневые коннекторы для промышленных протоколов обмена данными (MQTT, OPC UA, Historian) | Снизил нагрузку на CPU сервера обмена данными в 3 раза
- Мигрировал model serving runtime на Kubernetes (KubeAPI, Helm)
- Адаптировал платформу для AKS (Azure Cloud Kubernetes) и K3s (инсталляция в закрытый контур без интернета)

Middle Machine learning engineer, Ноябрь 2017 - Сентябрь 2019

[Conundrum.AI](#), // Москва, Россия

- Реализовал feature store для хранения данных сенсоров в виде временных рядов (Java, Spring, PostgreSQL, TimescaleDB)
- Реализовал model serving runtime сервис (Python, Processes)
- Реализовал сервис управления инцидентами (Java, Spring, State machines)
- Реализовал ETL на основе S3, SQS, S3 SFTP

Junior Machine learning engineer, Май 2017 - Октябрь 2017

[Conundrum.AI](#), // Москва, Россия

- Кластеризация данных авиакомпании. Реализовал подход для разделения данных для офлайн АВ тестов
- Предсказания оттока для мобильного оператора. Реализовал подход для офлайн предсказания оттока на основе активности пользователей
- Детекция пола для посетителей веб ресурса. Реализовал подход для оффлайн предсказания пола на основе активности пользователей
- Анализ гео данных мобильного приложения. Реализовал отчет об активности пользователей на основе их гео информации
- Временные ряды для данных промышленности. Реализовал pipeline для предсказания поломок

Machine learning engineer, Октябрь 2016 - Май 2017

Big Data Indicators · Стажировка, // Москва, Россия

- Реализовал pipeline сбора и обработки данных
- Использовал тематическое моделирования для поиска трендов
- Использовал анализ тональности для предсказания трендов

Образование

Сентябрь 2016 - Июнь 2018

Факультет вычислительной математики и кибернетики Московского государственного университета имени М. В. Ломоносова, Магистр прикладной математики и информатики.

Тема: “NLP для задач макроэкономического мониторинга”.

Сентябрь 2012 - Июнь 2016

Московский энергетический институт (Национальный исследовательский университет).
Кафедра математического моделирования.

Тема: “Распознавание лиц”.

Вклад в Open Source.

- Apache Ignite. Реализован предпроцессор для кодирования целевой переменной
- Apache Ignite. Реализована интеграция с Yandex Catboost.
- Apache Ignite. Реализованы методы вычисления расстояний (BrayCurtis, Canberra, JensenShannon etc).
- Apache Camel. Исправлено взаимодействие Azure Blob Storage и Azure Blob Queue.
- Tornado Swagger. Swagger API Documentation builder for tornado server.

Conferences/Public speech

Май 2018. Волонтер Data Scientist в [EnduringNet](#) (основано Ser-Huang Poon, prof Manchester University)

Июль 2017. [Big Data approach to measure inflation expectations: the case of the Russian economy](#) (IFABS 2017 Oxford Conference), Goloshchapova, I., & Andreev M.

Май 2017. [Measuring inflation expectations of the Russian population with the help of machine learning](#) (Voprosy Ekonomiki), Goloshchapova, I., & Andreev M.