

# ИСЛАМОВ АЙРАТ

+7-952-409-55-00 ◇ Москва

[ajratislamow@gmail.com](mailto:ajratislamow@gmail.com) ◇ [tg: @ajratik](https://t.me/ajratik) ◇ [Github](https://github.com)

## ОБРАЗОВАНИЕ

---

2024-2026: НИУ ВШЭ, Москва, Факультет компьютерных наук,  
Прикладная математика и информатика, "Master of Data Science". Магистр.

2021-2023: НИУ ВШЭ, Санкт-Петербург,  
МАЙНОР Обработка и анализ данных

2020-2024: НИУ ВШЭ, Москва, Факультет социальных наук,  
Психология, Когнитивная психология и психофизиология. Бакалавр.  
GPA 8.04

## НАВЫКИ

---

Языки программирования	Python, SQL (Postgres, Hive), R
Библиотеки Python	scikit-learn, pyspark, matplotlib, seaborn, pandas, numpy, FuzzyWuzzy
Библиотеки R	ggplot2, Plotly, caret, jsonlite, tidymodels, lme4, randomForest
Владение	HUE, IBM SPSS, MS Excel, Jamovi

## ОПЫТ РАБОТЫ

---

Управление аналитики по расходам и Process mining,  
ПАО Сбербанк

Стажер-аналитик

Ноябрь 2024 — май 2025

- Создано более 10 "цифровых двойников" на аналитической платформе Process Mining (процессы кредитования, вкладов, дебетовых и кредитных карт). Аналитика логов мобильного приложения СБОЛ и процессов обслуживания клиентов банка.
  - подготовка разметки процесса для последующего обогащения данных, парсинг разметки клиентов
  - интеллектуальный поиск текстовых совпадений событий в базе данных ClickStream (FuzzyWuzzy)
  - подготовка логики фильтрации выгрузки и написание скрипта pyspark для выгрузки Hadoop-data
  - обогащение данных, добавление новых признаков (scikit-learn)
  - работа с неатомарными, нестандартными данными и структурами (JSON и другие)
  - загрузка данных на платформу Process Mining, генерация дэшборда и добавление необходимых для заказчика элементов аналитической визуализации
  - написание итогового скрипта HiveSQL для вывода дэшборда в прод, оптимизация запроса
- Систематизировал методы pyspark и HiveSQL для выгрузки данных ClickStream из БД в рамках создания обучающего курса для внутренних клиентов.  
Создал скрипт с разбором часто используемых функций для обработки Hadoop-данных.

Департамент психологии НИУ ВШЭ

Учебный ассистент по дисциплине «Статистика для анализа данных»

Сентябрь 2022 — декабрь 2022

- проведение консультационных занятий для студентов, осваивающих дисциплину СДАД.
- подготовка учебных материалов (код и презентации) для занятий по освоению базовых инструментов анализа данных в Python: работа с pandas, numpy, matplotlib, seaborn.

- обсуждение и формирование учебного плана дисциплины для улучшения качества освоения ее студентами.

## Лаборатория нейробиологических основ когнитивного развития,

### Департамент психологии НИУ ВШЭ

Стажер-аналитик

Октябрь 2021 — июль 2022

- сбор данных поведенческого эксперимента для проверки экспериментальных гипотез проекта «Math and mental attention in children and adults» под руководством Marie Arsalidou.
- предобработка данных (удаление выбросов, объединение в единый датасет, взвешивание, агрегирование исследуемых метрик по испытуемым)
- проведение статистических тестов для проверки гипотез исследования (включая проверку выполнения условий для тестов, а также использование преобразований данных для их соблюдения)
- визуализация полученных данных в виде диаграмм, форматированных для публикации в научных отчетах и статьях

## ПРОЕКТЫ

---

### Курсовой проект по разработке моделей для предсказания цен на дома Python, sklearn [GitHub](#)

- Предварительно проведен EDA и Feature Engineering
- Построена модель линейной регрессии с регуляризацией (Ridge LR).
- VIF-based корректировка модели.

### Разработка AI - ассистента для суммаризации лекций Python, YCloudML, ollama [GitHub](#)

- Разработана внутренняя логика по обработке транскрипций лекции в умный конспект с таймкодами ключевых моментов.
- Функционал суммаризатора встроен в tg-бота.

### EDA и обогащение датасета по Netflix Python, scipy, FuzzyWuzzy [GitHub](#)

- Датасет обогащен по непрямому соответствию строк с помощью расстояния Левинштейна.
- Подобраны и проведены статистические тесты для проверки гипотез о критериях успешного фильма для платформы Netflix.

### Разработка приложения-компаньона для игры Dota 2) R, CatBoost [GitHub](#)

- Обучена и затюнингована модель CatBoost (на более чем 40000 игровых онлайн сессий), полученная модель в зависимости от союзного и вражеского набора персонажей рекомендует пользователю оптимальный выбор героя для увеличения шансов на победу.

### ВКР по исследованию восприятия пользовательских интерфейсов R, Python, PsychoPy, JavaScript

- Разработан 11 минутный эксперимент с имитацией работы с интерфейсом в среде PsychoPy + JavaScript (для онлайн версии эксперимента). [GitHub](#)
- Обработаны экспериментальные данные (генерация единого датасета, агрегация данных по испытуемым, исключение выбросов, проверка данных на нормальность), построена Mixed ANOVA с визуализацией в ggplot2, изучены и обобщены методы робустного трехфакторного анализа. [GitHub](#)

## ЯЗЫКИ

---

- Русский (родной)
- Английский (B2)