



SHAPING TOMORROW

By Salamat Akinova



НЕМНОГО ОБО МНЕ

Я - Саламат. По профессии я дизайнер интерьера и проработав некоторое время в этой отрасли я поняла, что это не моё - сидеть в офисе (сейчас я считаю иначе), мне хотелось путешествовать, жить за пределами своей страны, расширять кругозор и с тех пор я посещаю свой дом только раз в год или того меньше. Мне всё нравится в моей жизни, но я поняла, что мне нужно что-то новое и я решила начать с получения новых знаний или даже смены профессии. Выбрала профессию Data Scientist, потому что с детства у меня хорошо получалось с математикой и даже моя нынешняя работа связана с цифрами, а ещё эта отрасль сейчас развивается и я хочу быть 'в теме' :)

- Я совсем новичок в мире IT, но я считаю, что терпение и труд всё перетрут.
- В математике у меня базовые знания, не было необходимости пользоваться excel, только для некоторых отчётов в работе.

DATA SCIENTIST

- Мне интересен ML Engineer и изучив эту роль, пришла к пониманию, что нужно начать с Data Scientist.

РЫНОК ВАКАНСИЙ

Junior Data Scientist (Python, ML) / Дата-сайентист

Что вы будете делать:

- Помогать в сборе и подготовке данных для моделей прогнозирования;
- Участвовать **в разработке и тестировании моделей машинного обучения;**
- Проверять гипотезы, находить инсайты. **Анализировать данные и создавать отчеты;**
- Изучать и реализовывать лучшие практики командной разработки под руководством более опытных коллег.

Что мы ожидаем от вас:

- **Понимание принципов машинного обучения и статистики;**
- Знание классических алгоритмов и структур данных;
- **Уверенное знание Python и ML-стека (pandas, numpy, scikit-learn, что-либо из lightgbm/xgboost/catboost);**
- **Опыт работы с git;**
- **Опыт или желание изучать PySpark;**
- Самостоятельность и стремление к постоянному обучению в контексте практических задач.



https://hh.ru/vacancy/124449246?query=data+scientist&hhtmFrom=vacancy_search_list





Чем предстоит заниматься:

Анализировать новые схемы фрода и предлагать механизмы защиты;

Встраивать ML модели в существующие системы антифрода;

Выявлять абьюзивный контент и абьюзивное поведение на платформе;

Разрабатывать и тестировать новые признаки определения подозрительной активности пользователей.

Ты - наш идеал, если:

Имеешь опыт реализации и применения моделей машинного обучения в продакшене;

У тебя есть **понимание основных методов и алгоритмов Machine Learning**;

Уверенно владеешь Python (numpy, scipy, pytorch, sklearn);

Алгоритмическая подготовка: **знание основных алгоритмов и структур данных**;

Имеешь понимание и опыт работы с A/B-экспериментами, знание стоящей за ними статистики;

Умение работать в команде, грамотно отстаивать свою точку зрения - это все про тебя.



[https://hh.ru/vacancy/123111578?
query=data+scientist&htmlFrom=vacancy_search_list](https://hh.ru/vacancy/123111578?query=data+scientist&htmlFrom=vacancy_search_list)

ПЛАН ОБУЧЕНИЯ

Требование Математика и статистика	Мой уровень Школьная база есть, но без углубления	План закрытия Разбирать простые задачи ежедневно
Python для анализа данных	Начинающий	Изучить pandas и numpy Сделать 3–4 маленьк проекта по анализу датасета.
Визуализация и графика	Навыков пока нет	Освоить matplotlib/seaborn. Научиться строить 3–4 типа графиков на датасете.
Машинное обучение (введение)	Пока ноль	Отложить на 2-й месяц. Вернуться после освоения Python и статистики.

Приоритеты на ближайший месяц

1. Python для анализа данных → освоить pandas и numpy через практику.
2. Базовая статистика → закрепить среднее, медиану, дисперсию, вероятности.
3. Визуализация → научиться строить графики для анализа.

Примерный график (на 4 недели)

- Неделя 1: повтор Python, начать pandas (чтение датасетов, фильтрация, сортировка).
- Неделя 2: простая статистика + первые визуализации (matplotlib).
- Неделя 3: практика: анализ датасета Titanic (описание данных + графики).
- Неделя 4: мини-проект (анализ фильмов или продаж).