



## Александр Семенюк

Телефон: +375 29 297 43 00

Email: asemenyuk91@gmail.com

Telegram: @asemenyuk91

Город проживания: Гродно, РБ

Дата Рождения: 30.10.1991

GitHub: <https://github.com/AlexanderSemeniuk/portfolio>

## Аналитик данных/Data Scientist

Анализирую данные

Строю модели

Разрабатываю дашборды

Обучаюсь и с удовольствием берусь за сложные и интересные проекты

## Стэк

- Python
- Pandas, Sklearn, Numpy, Pyspark
- Catboost, Lightgbm, Xgboost
- Matplotlib, Seaborn, Plotly
- YData Profiling, LightAutoML
- FastApi, Grafana, AirFlow
- Работаю с Git
- ChatGPT
- SQL(SQLite)
- Power BI
- Excel
- GA, Яндекс метрика
- Docker
- Power Query
- Сводные таблицы
- Парсинг данных по api

## HARD Skills

- Провожу Бриф заказчика:
  - анализирую потребности
  - уточняю цели
  - формирую требования;
- Загружаю данные из разных источников(файлы, БД, api);
- Провожу разведочный анализ данных (EDA) :
  - проверяю данные на корректность и полноту
  - чищу данные
  - заполняю пропуски разными способами
  - работаю с выбросами и аномалиями
  - определяю взаимосвязи, исследую основные характеристики
  - feature engineering
  - готовлю данные для ML(оhe, labelenc, standartsc, minmaxsc ...)
  - визуализирую данные;

- Генерирую и проверяю гипотезы(АБ тест, Бутстрэп);
- Строю модели машинного обучения (Modeling):
  - выбор метрик
  - обучение и тестирование модели
  - работа с фичами
  - подбор гиперпараметров модели
  - автоматизированный pipeline;
- Оцениваю готовую модель (Evaluation):
  - кросс-валидация
  - оценка метрик
  - выбор наилучшей модели;
- Упаковываю модель Deployment\*:
  - сериализация модели
  - развертывание модели
  - мониторинг работы модели;
- Строю дашборды;
- Провожу сегментацию и кластерный анализ;
- Делаю когортный анализ;
- Считаю метрики(продуктовые, маркетинговые, продажи);
- Ищу инсайты и зоны роста продукта;
- Презентую результаты исследований;

### SOFT Skills

- Умею слушать(владею техниками активного слушания)
- Умею проводить презентации(опыт 4 года)
- Умею проводить мозговой штурм с использованием различных техник(есть разработанный тренинг)
- Владею навыками деловой переписки
- Умею вести переговоры с ЛПРами
- Умею выявлять и формировать потребности
- Умею работать с возражениями
- Устойчив к неадекватному поведению и манипуляциям
- Умею обучать(успешно прошел обучение профессии Бизнес-тренер)
- Рационально отношусь к критике, умею как принимать, так и давать обратную связь.

### Проектная деятельность на платформе Sillbox

**Machine Learning Engineer**, июль 2023 - по настоящее время

Реализовал ряд проектов:

1. Модель кредитного риск-менеджмента  
[https://github.com/AlexanderSemeniuk/portfolio/tree/main/project ML\\_risk\\_management](https://github.com/AlexanderSemeniuk/portfolio/tree/main/project ML_risk_management)
2. Сводный дашборд для продуктовой команды EduGO  
[https://github.com/AlexanderSemeniuk/portfolio/tree/main/project\\_p roduct\\_dashboard](https://github.com/AlexanderSemeniuk/portfolio/tree/main/project_p roduct_dashboard)

3. АБ тест маркетинговой активности крупного магазина спортивных товаров, выгрузка данных, заполнение потерянных данных на основе ML модели

[https://github.com/AlexanderSemeniuk/portfolio/tree/main/project\\_sportswear](https://github.com/AlexanderSemeniuk/portfolio/tree/main/project_sportswear)

С остальными проектами можно ознакомиться по ссылке

<https://github.com/AlexanderSemeniuk/portfolio>

***Достижения и результаты:***

- Разработал модель которая предсказывает дефолт клиента по кредиту, качество модели по метрике  $\text{roc\_auc} > 0,7635$ . Такой результат получился благодаря правильной подготовке данных, перебору более 10 моделей и подбора их гиперпараметров.
- Разработал сводный дашборд для продуктовой команды EduGO, который позволяет ежедневно/еженедельно мониторить продуктовые метрики и их динамику. Выявил какие каналы и лендинги работают эффективнее, нашел канал по которому заявки не доходят до обработки, внес 4 предложения по оптимизации работы продукта. Проект был признан куратором одним из самых детализированных и качественных проектов за все время.
- Собрал данные из нескольких источников, заполнил пропущенные данные с помощью модели классификации и провел АБ тест маркетинговой активности крупного магазина спортивных товаров. Проверил валидность применения статистического критерия АБ теста с помощью Бутстрэпа, нашел минимальный детектируемый эффект.