

# SkillGrow

Система рекомендаций недостающих навыков в резюме DS  
специалиста

*Курсовой проект DS-17*

Игорь Земенков

Ноябрь 2025

# Краткое описание проекта, постановка задачи



**Соискатели сталкиваются с пробелами в компетенциях**

сложно определить, чему учиться в первую очередь



**SkillGrow автоматически собирает вакансии,**  
анализирует требования и рекомендует недостающие навыки

1



## Актуальность

Рынок Data Science стремительно растёт и повышает требования к кандидатам, при этом ATS системы фильтруют резюме по признаку сходства с описанием вакансии

2



## Цель проекта

Выявить недостающие навыки в резюме и предложить путь развития / доработки резюме

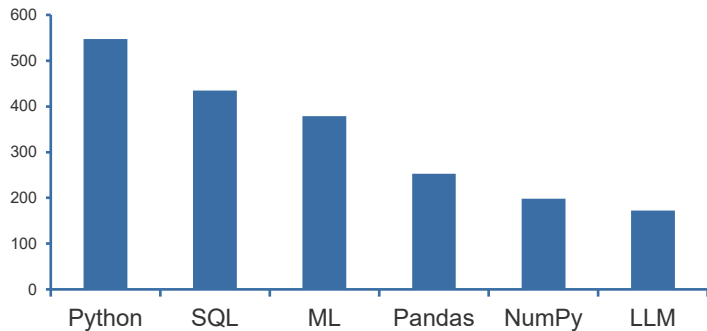
3



## Задачи

Собрать данные о вакансиях, извлечь навыки, построить модель и создать приложение.

# Портрет рынка

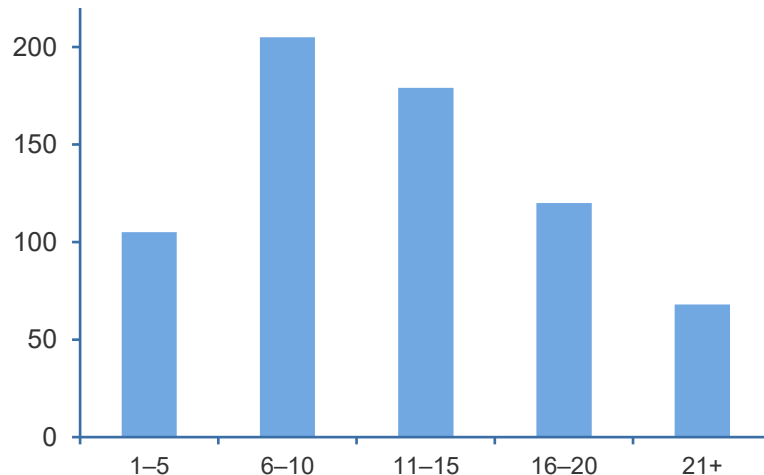


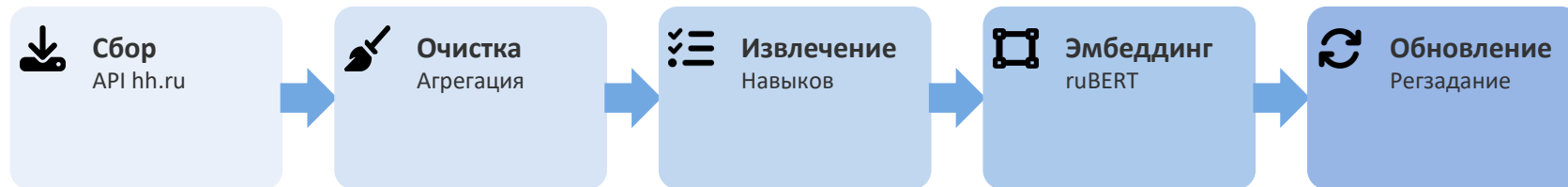
**Большинство вакансий требуют 6–15 навыков,** что подтверждает мультидисциплинарный характер работы.

Только 15% позиций ожидают менее 6 навыков, а 10% – более 20, часто для руководящих ролей.

## Инсайты:

- Python и SQL встречаются более чем в половине вакансий.
- Машинное обучение (ML) и библиотеки Pandas/NumPy занимают прочные позиции.
- LLM модели уже упоминаются в десятках вакансий.





Показатель	Значение*
Вакансий	677
Уникальных навыков	279
Среднее навыков/вакансию	12.3

Система регулярно собирает вакансии через API hh.ru, очищает и агрегирует данные в мастер-датасет, извлекает ключевые навыки и строит эмбеддинги для сравнения.

Обновления выполняются автоматически через регзадания (github workflows)

\*По состоянию на 24.11.2025

# Алгоритм рекомендации



## Загрузка резюме

Парсинг DOCX и очистка текста



## Извлечение навыков

Сопоставление словарю из вакансий



## Эмбединг

Векторизация резюме и вакансий (ruBERT)



## Поиск похожих вакансий

Cosine similarity для N ближайших



## Вывод рекомендаций

Частоты, семантическая близость и фильтрация

# Интерфейс приложения

## Настройки рекомендаций

Сколько навыков показывать на графике

15

Сколько похожих вакансий учитывать

20

## SkillGrow — рекомендации по развитию навыков для Data Science

Анализ резюме Как собираются вакансии с hh.ru Как извлекаются навыки и строится эмбединги

### Анализ резюме и рекомендации по навыкам

Загрузите ваше резюме (DOCX)

Drag and drop file here  
Limit 200MB per file • DOCX

Browse files

resumeJunior\_ds.docx 36.5KB

#### Извлечённый текст резюме

Игорь Иванов — Junior Data Scientist Контакты Email: [example@email.com](mailto:example@email.com) Телефон: +7 (900) 000-00-00 Город: Москва О себе Начинающий специалист по анализу данных. Обладаю базовыми навыками в машинном обучении, работе с данными и построении аналитических моделей. Интересуюсь разработкой рекомендательных систем и применением NLP в бизнес-задачах. Навыки: Python (pandas, numpy, matplotlib, scikit-learn) • SQL (PostgreSQL, MySQL) • Machine Learning (классические модели) • Работа с API • Data Visualization (Power BI, Tableau) • Git, Linux, Jupyter Notebook Опыт проектов Анализ продаж — исследование данных, визуализация ключевых метрик, построение прогностической модели. Классификация текстов — подготовка выборки, предобработка, обучение модели Logistic Regression. Образование Бакалавр, Прикладная математика и информатика Университет — 2020–2024

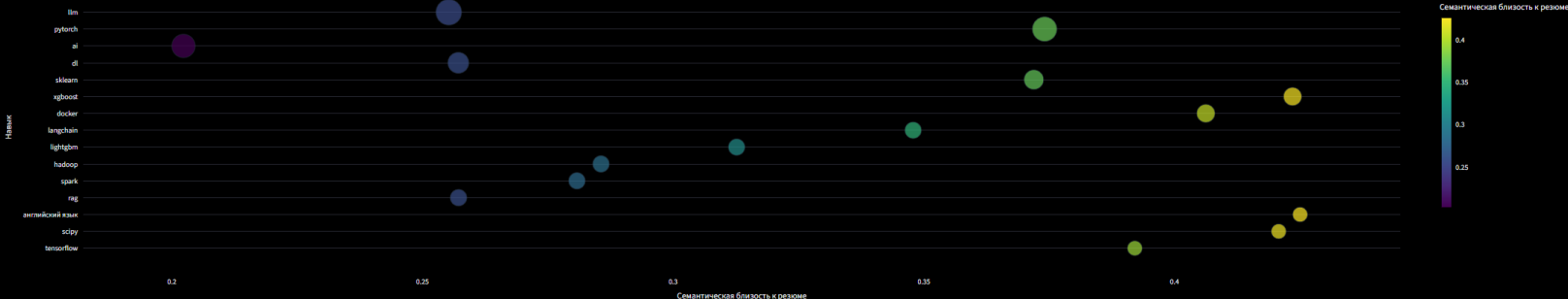
#### Найденные навыки (нормализованные)

api, business intelligence, data science, git, jupyter notebook, linux, machine learning, matplotlib, mysql, nlp, numpy, pandas, postgresql, power bi, python, scikit-learn, sql, tableau

### График приоритетных навыков

Рекомендация не означает, что у вас точно нет навыка — она показывает, какие навыки чаще всего встречаются в похожих вакансиях и не упомянуты в вашем резюме.

#### Рекомендованные навыки для развития



### Таблица рекомендаций

Навык	Частота в похожих вакансиях	Семантическая близость к резюме
llm	12	0.25

<https://skillgrow.streamlit.app/>

# Заключение и возможности развития MVP

## Достижения

- Организован сбор и обработка данных с hh.ru.
- Реализован модуль извлечения и нормализации навыков.
- Построена модель рекомендаций на основе эмбедингов.
- Создано приложение Streamlit для конечных пользователей.

## Возможности развития

- Расширение на другие области (веб-разработка, BI)
- Интеграция дополнительных источников данных
- Использование LLM для более точного извлечения навыков
- Адаптация резюме под вакансии