

# Sofia SHISHKINA



## PROFILE SUMMARY

*Data Scientist* with a solid foundation in machine learning, statistical analysis, and mathematical modeling. Experienced in data processing, visualization, and mathematical research. Strong communicator focused on turning complex data into actionable insights. Eager to apply skills in real-world data science projects.

## CONTACT DETAILS

@ LinkedIn, GitHub

+31 637247552

sepofia@gmail.com

## PERSONAL INFORMATION

Citizenship: **Amsterdam**

Languages:

- English - *Professional*
- Russian - *Native*
- Dutch - *Beginner*

## SKILLS

- **Python Libraries:** Pandas, SciPy, NumPy, Scikit-learn, PyTorch, transformers, Matplotlib, Seaborn, Dash, Psycopq2, Requests
- Docker, Git, API, Bash
- SQL, Power BI, LaTeX, Excel
- Haskell, Matlab
- Communication and team collaboration

## HOBBIES

Playing the piano, Pilates, and composing music for games

## WORK EXPERIENCE

INTERN FINANCIAL ANALYST, *Bank Saint Petersburg*. **2024.06–2024.08**

- ◇ Built a forecasting model to predict the Central Bank's Key Rate, based on a comprehensive mathematical analysis.
- ◇ Produced rigorous documentation and formal model validation for internal and external stakeholders.

HASKELL TEACHER ASSISTANT at *HSE, St. Petersburg*. **2023.09–2023.12**

- ◇ Helped with lectures, checked students' homework, and answered questions.

## PROJECTS

AAGENT5 – ANTIBODY SEQUENCE GENERATOR, *Master's Thesis*

- ◇ Collected, cleaned, and processed large datasets with biological data (antibody chain sequences and their characteristics). Conducted exploratory statistical analysis to reveal the biological patterns.
- ◇ Designed a Transformer-based model for antibody chains (letter sequence) generation with a custom tokenizer, data preprocessing, and metrics to test the model's performance.

DRAMABOT - CHAT BOT FOR PERSONALIZED DRAMA RECOMMENDATIONS

- ◇ Designed and deployed a multi-container Telegram bot system using Python, PostgreSQL, and Docker.
- ◇ Collected Asian drama series data from an open-source API, developed and deployed the database on a server, and created a bot for recommending series based on user preferences. Implemented automated database backup to Google Drive and language translation via the Google Translate API. The database is connected to the Adminer GUI for convenient and quick access.

## EDUCATION

MASTER OF COMPUTATIONAL BIOLOGY AND BIOINFORMATICS.

*Higher School of Economics, St. Petersburg, Russia.* **2022–2024**

BACHELOR OF ENGINEERING (BENG).

*Bauman Moscow State Technical University.* **2017–2021**

## CERTIFICATIONS

◇ INTRODUCTION TO DATA SCIENCE, *Skillbox.* **2024.09–2025.06**

◇ "ENGLISH FOR ANALYSTS" PROFESSIONAL TRAINING COURSE, *Yandex Practicum.* **2024.09–2024.12**