ANASTASIYA BYELYAKOVA

nasbelyakova@gmail.com

+380684892203

linkedin.com/in/anastasiya-byelyakova

https://github.com/AnastasiyaByelyakova?tab=repositories

EXPERIENCE

Data Analyst
Kyros AI
Aug 2021 - June 2025

- ► Scraping websites to get the data, details and answers for questions (including universities, summer programs, volunteering program, etc.) extensively using NLP techniques (with spaCy and transformers for texts cleaning, NER and summarization) and AI (Gemini and ChatGPT with LangChain) including using FAISS

- ► Perform image processing for websites including background optimization and cropping

- ► Build predictive models for university ROI and student acceptance probability using machine learning

- ► Build predictive models for predicting selectivity of summer program with tensorflow, scikit-learn and XGBoost

- ► Build classification models for classifying essays examples texts into pre-defined categories with scikit-learn and XGBoost

- ► Develop advanced chatbots to answer student inquiries using AI technologies

- ► Analyze acceptance rates using AI models and statistical methods

- ► Develop Python scripts for automating data processing workflows

- ► Conduct research on educational materials contributing to new learning resource development

- ► Processing pdf files to get answers to questions i.e. numbers of admitted students by gender, how many applicants there were, universities to which high school graduates were admitted etc. Using AI and tensorflow tools.

Data Analyst
DataObrii
May 2017 - Aug 2021

- ► Created spaCy models for POS tagging and named entity recognition in text processing

- ► Developed NLTK models for text chunking and natural language processing

- ► Built web scraping solutions to extract job opening information from company platforms

- ► Implemented CNN classifiers for image recognition using transfer learning on Inception v3 (TensorFlow) For following cases:
  - ► Detecting COVID on lungs x-rays
  - ► Classifying dogs photoes based on dog breed
  - ► Classifying cosmetics photos based on cosmetic brand
  - ► Realtime cars counting on video from camera

- ► Developed LSTM models for sales volume prediction based on historical data (Keras)

- ► Created Flask APIs for MongoDB database access and user authorization systems

- ► Built LSTM and Random Forest models for financial price direction prediction

- ► Performed statistical analysis of Indian company price datasets using pandas, matplotlib, and seaborn

PhD Student & Researcher
Università degli Studi di Siena
Feb 2014 - Feb 2017

- ► Analyzed olive oil samples for acidity compliance with European Regulations No 2568/91 and No 2015/1830

- ► Performed chemical characterization of olive oil, fruits, leaves, and pomace samples

- ► Conducted antioxidant activity analysis using TEAC assay and spectrometric methods

- ► Quantified polyphenols content using Folin-Ciocalteu method

- ► Performed ion content analysis (K, Na, Ca, Mg, Mn, Cu, Zn) using atomic absorption spectroscopy

- ► Optimized RH-HPLC-MS methods for polyphenolic compound quantification

- ► Investigated olive oil production by-products for quality improvement applications

Teaching Assistant
University of Geneva

Feb 2013 - Oct 2013

- ► Contributed to research in photoswitchable Na/pH optical and potentiometric sensors
- ► Developed optical sensor nanoparticles for UV-Vis absorption and fluorescence applications
- ► Worked on fluorescence-based optical sensors for calcium detection

EDUCATION

Master of Science (MSc) - Analytical Chemistry

Taras Shevchenko National University of Kyiv

2010 - 2012

Thesis: *"Determination of 1,4-Dioxane by headspace GC with preconcentration by SPE"*

Coursework: Gas Chromatography, HPLC, Environmental Analytical Chemistry, Physical Methods of Analysis, Nanochemistry, Pharmaceutical Chemistry, Instrumental Analysis, Bioanalytical Chemistry, NMR Spectroscopy

Bachelor of Science (BSc) - Chemistry

National University of Kiev-Mohyla Academy

2005 - 2009

Thesis: *"Modification of PETF membranes by graft-copolymerization of N-VIA initiated by AIBN"*

Additional Coursework: Biochemistry, Genetics (General & Molecular), Cytology, Histology, Virology, Microbiology, Developmental Biology

SKILLS

Programming & Data Science

Python, OOP, SQL, MongoDB, Statistical Analysis, FAISS, Machine Learning, Deep Learning, N8N

Cloud Technologies

AWS S3, AWS EC2, AWS Lambda, AWS Redshift, Google Compute Engine

AI

Gemini, ChatGPT, LangChain

Technical Specializations

Natural Language Processing, Web Scraping, Data processing and visualisations, AI, Machine learning, Deep Learning

Tools & Technologies

TensorFlow, Keras, scikit-learn, pandas, numpy, selenium, playwright, matplotlib, seaborn, spaCy, transformers, NLTK, Flask, Docker, PySpark, Dask

CERTIFICATIONS

Programming for Everybody (Python) - Coursera

The Data Scientist's Toolbox - Coursera

Python for Genomic Data Science - Coursera

Algorithms for DNA Sequencing - Coursera

Bioinformatics Methods I & II - Coursera

Introduction to Genomic Technologies - Coursera

Finding Hidden Messages in DNA - Coursera

Nanotechnology and Nanosensors - Coursera

Drugs and the Brain - Coursera

Systems Biology Methods - Coursera

Python Programming Projects - Udemy

SQL Fundamentals - Sololearn

Diploma in C Programming - Alison

SQL Fundamentals - Sololearn

Diploma in C Programming - Alison