

# Liana Lotarets

## Data Scientist

### CONTACTS

---

+380 67 12 11 662 [🔗](#)

Telegram [🔗](#)

GitHub [🔗](#)

lyanalotarets@gmail.com [🔗](#)

LinkedIn [🔗](#)

Kharkiv, Ukraine

### SUMMARY

---

Junior Data Scientist with a strong foundation in mathematics and expertise in Python. Experienced in data cleaning, exploratory data analysis (EDA) using Pandas, Matplotlib, and Seaborn, and implementing supervised learning algorithms (classification and regression), as well as working with data gathering (BeautifulSoup), unsupervised learning algorithms (clustering), deep learning algorithms (CNN and RNN), and deploying ML applications with Streamlit. I aim to further develop my skills in Data Science and Machine Learning by collaborating with seasoned professionals and contributing to the company's success.

### HARD SKILLS

---

- |                 |                                   |                 |
|-----------------|-----------------------------------|-----------------|
| • Mathematics   | • EDA                             | • Keras         |
| • Python        | • Matplotlib, Seaborn             | • CNN           |
| • NumPy         | • Scikit-learn, LightGBM, XGBoost | • OOP           |
| • Pandas        | • Supervised Learning             | • BeautifulSoup |
| • Data Cleaning | • Deep Learning                   | • SQL, NoSQL    |

### SOFT SKILLS

---

- |                     |                       |               |
|---------------------|-----------------------|---------------|
| • Lifelong Learning | • Desire to learn     | • Helpfulness |
| • Task prioritising | • Information seeking |               |

### LANGUAGES

---

English – intermediate

Ukrainian – native

Japanese – elementary

### PROJECTS

---

#### MNIST classification + OOP, [GitHub 🔗](#)

*Tools/Technologies:* Python, OOP, Keras, Matplotlib, Scikit-learn, Deep Learning, CNN

*Description:* MNIST classification using OOP three models: Random Forest, Feed-Forward Neural Network, Convolutional Neural Network.

*Achievements:* All models performed excellently. *Random Forest* (accuracy 97%, size 137.51 MB) is not the best choice, as it is less accurate and significantly larger than the other models. The *Convolutional Neural Network* (accuracy 99%, size 10.4 MB) is a much better option compared to Random Forest. However, if model size is a critical factor, the *Feed-Forward Neural Network* (accuracy 98%, size 5.39 MB) is also a good alternative.

#### Customer Outflow, [GitHub 🔗](#)

*Tools/Technologies:* Pandas, Matplotlib, Seaborn, Scikit-learn, Dash, Plotly, Bootstrap, Docker, Koyeb

*Description:* An educational team project focused on analyzing customer churn for telecommunication companies. My task was to build and train models. When selecting the best model, the size of the model was also taken into account.

*Achievements:* The best model achieved an accuracy of approximately 98% while being memory-efficient.

### **Binary Prediction of Poisonous Mushrooms, [kaggle](#)**

*Tools/Technologies:* Python, Pandas, Scikit-learn

*Description:* A Pet-project based on a Kaggle dataset from the competition *Binary Prediction of Poisonous Mushrooms*. The goal of this competition is to predict whether a mushroom is edible or poisonous based on its physical characteristics.

*Achievements:* The most challenging part of the task was data preprocessing, as the data was not pre-cleaned. The model demonstrates high accuracy: *Private Score is 0.98012, which means 98% accuracy*.

### **Quotes and Authors, [GitHub](#)**

*Tools/Technologies:* Python, BeautifulSoup, NoSQL

*Description:* The website <http://quotes.toscrape.com> is being scraped. The goal is to extract information about quotes (tags, author, text) and their authors (full name, date of birth, place of birth, and description).

*Achievements:* Two JSON files were obtained: authors.json contains information about authors, and quotes.json contains quotes and information about them.

### **Ukraine's birth rate (1950–2019), [GitHub](#)**

*Tools/Technologies:* Python, Pandas, Matplotlib

*Description:* Analysis of the table *Birth Rate in Regions of Ukraine (1950–2019)* from the website [Population of Ukraine](#).

*Achievements:* Local birth rates in the late 1990s and early 2000s were found to be the lowest during the period from 1950 to 2019. Additionally, it can be concluded that right-bank regions generally have higher birth rates than left-bank regions.

## **WORK EXPERIENCE**

---

### **Freelance, Math Tutor**

2018 – Present

- Successfully assisted an entrant in preparing for master's studies by providing foundational knowledge in numerical methods, optimization methods, probability theory and statistics.
- Successfully assisted entrants in preparing for the bachelor's degree entrance exam in mathematics.

### **National Research Foundation of Ukraine (NRFU), Grantees**

2021 – 2022

- Published a scientific article with the results of research supported by the National Research Foundation of Ukraine funded by the Ukrainian State budget in frames of project 2020.02/0096 “Operators in infinite-dimensional spaces: the interplay between geometry, algebra and topology”.

### **Akhiezer Foundation, Grantees**

2023 – 2024

- Published a scientific article with the results of research supported by the Akhiezer Foundation.

## **EDUCATION**

---

### **V. N. Karazin Kharkiv National University**

2015–2019, Bachelor's degree, Mathematics



2019–2021, Master's degree, Mathematics

### **IT School GoIT**

June 2024 – January 2025, Data Scientist

## PUBLICATIONS

---

- LOTARETS, LIANA (2022) "Geodesics of fiberwise cigar soliton deformation of the Sasaki metric," *Turkish Journal of Mathematics*: Vol. 46: No. 1, Article 10. [DOI](#) 
- LOTARETS, LIANA (2024) "Twisted Sasaki metric on the unit tangent bundle and harmonicity," *Turkish Journal of Mathematics*: Vol. 48: No. 2, Article 4. [DOI](#) 
- Lotarets, L. (2024). A characteristic property of Sasakian manifolds. *Proceedings of the International Geometry Center*, 17(3), 218-231. [DOI](#) 