

Irena Zimovska

(+48) 789 702 690 | irenazimovska@gmail.com | [LinkedIn](#) | [GitHub](#)

SUMMARY

- Data Scientist with a strong academic foundation in statistical analysis and machine learning, specializing in credit risk modeling and automation.
- Delivered business intelligence solutions for credit risk institutions, streamlining scoring model development and monitoring to enhance process efficiency and ensure compliance with KNF/EBA regulations.
- Highly collaborative and communicative, with a strong commitment to ethics, teamwork, and delivering high-quality business products.
- Top 4 graduate of the Data Science & Business Analytics program (2023/2024) at WNE UW, winner of the [Semkow Award](#) for the Best Master's Thesis defended at WNE UW (2023/2024).

SKILLS

Languages

Polish (Professional Working/ C1)
English (Professional Working/ C1)
Russian (Native/ C2)
Turkish (Spoken/ B1)
German (Elementary/ A1)

Technical Skills

Python, R, SQL, SAS, Git/Gitlab, PySpark

Libraries:

Numpy, Pandas, scikit-learn, XGBoost, Matplotlib, Plotly, Dash, Streamlit, DALEX, SHAP, PyTorch, OmniXai, Captum, tidyverse/dplyr, ggplot2, Shiny, Selenium, Scrapy, BeautifulSoup, PyTest

Expertise

Econometric modelling (e.g. Linear/Logistic Regression, Tobit Regression), Machine Learning Modelling (e.g. Gradient Boosting, Random Forest, Neural Networks, PCA, t-SNE, K-means, PAM, HDBSCAN), NLP (Text data preprocessing, Sentiment Analysis, Advanced Topic modelling, Embeddings, Named Entity Recognition), Web-Scrapping, Credit Risk (PD/LGD models, KNF/EBA regulations), Statistical Analysis, Finance, Anti-Money Laundering

EXPERIENCE

FACULTY OF ECONOMIC SCIENCES, UNIVERSITY OF WARSAW

Lecturer & Course Developer

February 2025 – Pending

Developing and introducing a novel elective course for the master's students of WNE UW called: **"Applications of Explainable AI in Predictive Modelling"**.

The aim of the course is to equip students with the knowledge and skills necessary to create ML models with a focus on model interpretability, ethics, and compliance with regulations. Throughout the course the whole ML modelling pipeline is demonstrated solving Regression and Computer Vision tasks.

The course module implements Python and frameworks such as PyTorch, SHAP, OmniXAI, Captum.

DATA JUICE LAB SP. Z O. O.

Data Scientist

October 2024 – April 2025

- Development of application for LGD modelling pipeline allowing for interactive feature engineering based on PySpark.

Data Science Consultant

April 2023 – October 2024

- Development and implementation of automation software for PD model monitoring, designing backend statistical analysis modules and an interactive dashboard for end-to-end model evaluation.
- Development of a comprehensive model back-testing methodology (homogeneity analysis, PD/DR confidence intervals, etc.), ensuring compliance with KNF regulatory standards and audit requirements.
- Trained banking teams in adopting open-source tools for daily operations, delivering structured Python training sessions to enhance model development capabilities.
- Actively contributed to proof-of-concept development for client offers – research and design, preliminary project planning and backlog estimation.
- Worked with the Financial Analysis team of a global production enterprise, ensuring data verification, database construction, and monthly financial KPI reporting for decision-making support.
- Assisted in credit policy audits, analyzing factors affecting low acceptance and booking rates in loan application processes.
- Contributed to the development of AI-based recommendation system for arbitrage court decisions, leveraging machine learning techniques.
- Performed feature engineering for text-based data, including natural language processing (NLP) techniques such as lemmatization and seamless ETL pipeline integration.
- Conducted secondary verification of labeled data provided by a legal expert team, identifying and addressing annotation errors to ensure model reliability.

Stack: R, Python, SQL, numpy, pandas, Scikit-Learn, Dash, PySpark, Hadoop, Hive, H2O, Git/GitLab

FACULTY OF ECONOMIC SCIENCES, UNIVERSITY OF WARSAW

Research Investigator in Scientific Grant

December 2023 – November 2024

Investigator in Sonata 18 Research Program of the National Science Centre (NCN), Poland. Project title: Populism and economic performance: a quantitative perspective, project no: UMO-2022/47/D/HS4/01047.

- Led technical implementation of research ideas, working extensively with unstructured data to extract meaningful insights.
- Developed and maintained R-based analytical pipelines, ensuring code efficiency, reproducibility, and clarity for long-term usability.
- Designed and executed randomized experiments and regression analyses to measure the impact of data toggling, uncovering key effects.
- Owned the end-to-end coding process, becoming the principal technical support in the research team.
- Collaborated weekly in research calls, taking on key tasks and aligning technical execution with project goals.

Stack: R, Git/GitHub, Econometric Modelling

EDUCATION

FACULTY OF ECONOMIC SCIENCES, UNIVERSITY OF WARSAW

MA in Data Science and Business Analytics

September 2022 - September 2024

GPA: 4.85 / Grade: 5! (Summa Cum Laude)

Thesis title: "Exploring parliamentary activity in the Polish Sejm: analyzing parliamentary interpellations with text mining".

My research analyzed the parliamentary activity of the Polish Sejm during its IX term, focusing on how deputies used interpellations (pol. *interpelacje poselskie*) to address economic issues. I personally built an original database of interpellations, enabling an in-depth analysis using advanced NLP techniques. Leveraging **Dirichlet Multinomial Regression** for topic modeling, I uncovered key themes in these texts while incorporating metadata such as submission year and political affiliation. This allowed me to identify patterns in how different political groups prioritize and discuss economic topics. Additionally, I applied **unsupervised clustering models like HDBSCAN and PAM** to segment deputies into behavioral groups based on their activity and personal characteristics.

Awards:

- **Nagroda im. Andrzeja Semkowa na najlepszą pracę magisterską** obronioną w r. a. 2023/2024 (przyznano w lutym 2025)
- Rector scholarship award for outstanding academic performance in 2023/2024.

FACULTY OF ECONOMIC SCIENCES, UNIVERSITY OF WARSAW

BA in Finance and International Investment

September 2018 - September 2021

GPA: 4.59 / Grade: 5

Thesis title: "Personal Data Spaces: contingent valuation of user-centric control over personal data access"

The research study aimed **to estimate the value of a user-centric privacy control over personal data** on the Internet by means of contingent valuation approach. In today's data-driven economy, personal data plays a crucial role, yet individuals have limited control over its use. To address this lack of practical tools for Internet users to manage the flow and online access to personal data by third parties, the concept of Personal Data Spaces was developed. In my study, the Personal Data Spaces application served as a proxy for valuation, with the survey focusing on determining individuals' Willingness-to-Pay (WTP) for a privacy protecting app that restores user-centric control over personal data. The estimations were analyzed with econometric approach, leveraging **Tobit Regression**.

Awards:

- Rector scholarship award for outstanding academic performance in 2019/2020, 2020/2021

DECLARATION

I agree to the processing of personal data provided in this document for realising the recruitment process pursuant to the Personal Data Protection Act of 10 May 2018 (Journal of Laws 2018, item 1000) and in agreement with Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).