STEPAN TYTARENKO

New York, NY 10023 | +1 (743) 217 3870 | stytarenko@fordham.edu | LinkedIn | Github | Website

EDUCATION

Fordham University, New York, NY — **August 2022 - December 2024**
Master of Science in Computer Science — **GPA: 4.0 / 4.0**

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine — **September 2017 - July 2021**
Bachelor of Science in Computer Science with high honors — **GPA: 97.74 / 100**

WORK EXPERIENCE

Graduate Researcher, *Fordham University*, New York, USA — **August 2023 - Present**

- Developed a PyTorch framework and authored a research proposing efficient Language Models fine-tuning process
- Conducted empirical studies showing a **15%** improvement in fine-tuning stability across diverse datasets, enhancing the robustness of Large Language Models (LLM) across multiple Natural Language Processing (NLP) tasks

Machine Learning & Data Science intern, *Volvo Financial Services***,** Greensboro, USA — **May 2023 - August 2023**

- Built DataBricks pipeline using PySpark for automatic monthly data cleaning, processing and model fine-tuning
- Applied a stack of an **XGBoost** model with **Random Forest** in Python, achieving **94.5%** accuracy on 20 years of data, and developed an ensemble model interpretation dashboard for non-technical audience
- Transitioned from classification to survival analysis on time series; used Deep Learning with Recurrent Neural Networks (RNN), achieving a c-index of **0.92** and **~96%** accuracy in a **95%** confidence interval

Software Engineer*, DataArt*, Kharkiv, Ukraine — **September 2021 - July 2022**

- Engineered a distributed computing solution that enabled the processing of datasets up to **10** times larger
- Designed Redis-based cache for the CRM system with **20 million** records, reducing average request latency by **~25%**
- Utilized AWS SQS for fault-tolerant distributed system's micro-services synchronization with worst-case delay <**60s**

Software Engineer, *Distributed Lab,* Kharkiv, Ukraine — **April 2020 - September 2021**

- Authored the system with **~$2.5M** USD live market cap, **100+** active validators, and **~1000** transactions per second
- Optimized data preprocessing pipelines, resulting in a **30%** reduction in data cleaning and transformation time
- Conducted resource utilization analysis, optimizing model inference processes resulting in **~45%** latency reduction

PROJECTS / PUBLICATIONS

First author - Transformer Context Attribution for Fake News Detection (2024) (*Accepted to ICLR*, *Github*)

- Created novel Framework and Loss function to incorporate transformer models for Natural Language Processing (NLP), promising better generalizability cross-dataset by **10-15%** in low power environment with limited GPU access

First author - Space Model Framework for Conceptual Modelling in NLP (2023) (*AAAI***,** *GitHub*)

- Implemented a state-of-the-art NLP framework that outperforms existing solutions by **2-11%** F1-score and accuracy
- Received “Best Paper Award” at AAAI ReLM 2024 for the novel Task-specific Context Attribution in PyTorch

Co-author - An Explainable Hate Speech Detection Framework (2023) (Accepted to *ICLR***,** *Github***)**

- Built a novel contextual attribution approach implementation using PyTorch, improving the overall classification F1 by **~1.5%** and F1 of Intersection over union (IOU) of rationale labels using BERT by **3.5%**

Co-author - Multimodal News Source Classification with NLP Transformers (2021) (*ICTERI***,** *GitHub*)

- Designed a winning Python pipeline for the Kaggle competition in classification of News in Ukrainian language, and collaborated on writing a research paper using transformers architectures such as RoBERTa, ELECTRA, and XLM-R

Real-time Gestures recognition application (2019) *Github*

- Computer Vision CNN-based application for real-time gestures recognition at University Startup World Cup 2019, won the “Social Media Award” among **40** teams from around the world. Finalist of the UNICEF funding for startups

SKILLS

Machine Learning & Data Science*:* Python, C++, SQL (Postgres) / NoSQL (MongoDB, Redis), PySpark, DataBricks, Tensorflow/Keras, PyTorch, MLFlow, Sklearn, AWS, Docker