OLEKSII TSEPA

+16478794301 - Toronto, ON

alextsepa001@gmail.com - imgremlin.github.io - linkedin.com/in/oleksii-tsepa

Machine Learning Engineer with 3 years of experience in various domains, including Retail, Drug Discovery, Robotics, etc. Specialize in NLP, Time-Series, and GraphML.

WORK EXPERIENCE

Competera Toronto, Canada
Machine Learning Engineer Oct 2020 - Current

- Developed the company's technical differentiator by creating **Product Embeddings** with **Graph Neural Networks** for better price elasticity estimation
- Developed a **Product Matching** system with a precision of 95% and recall of 70% to reduce the cost of manual matching by more than 40%
- Improved existing **Time-Series Model** accuracy by 7% for sales forecasting of more than 50k retail products and **deployed** in 130 retail outlets across over 25 cities in Ukraine
- Carried out **Recruitment Process**: screened 70+ resumes, evaluated 15+ test tasks, conducted 10+ tech interview

Vector Institute Toronto, Canada
Machine Learning Engineer May 2022 - Dec 2023

- Improved state-of-the-art bionic leg control accuracy (RMSE) by 20% while achieving 20x faster inference by designing novel **multimodal** architecture. Further, the model was **deployed** on Jetson Nano
- Introduced a novel Conditional Graph Fusion Layer (CongFu) designed for drug synergy predictions. CongFu outperformed state-of-the-art methods on 11 out of 12 benchmark datasets
- Improved the relative distance similarity between original and latent (2D) spaces by 10% through developing a Topology-preserving Graph Autoencoder
- Developed a model for de-novo molecule **generation** conditioned on mass spectra using **transformers** with adapted **RLHF** based on molecular similarity

RESEARCH RECORD

1. **Tsepa O**, Young A, Wang Bo. Conditional Molecule Generation based on the Mass Spectrum.
2. Wang C, **Tsepa O**, Wang Bo. Topology-preserving Graph Autoencoders.
3. **Tsepa O***, Naida B*, Goldenberg A, Wang Bo. CongFu: Conditional Graph Fusion for Drug Synergy Prediction. *Accepted to NeurIPS AI4D3 2023* [paper, code]
4. **Tsepa O**, Yang Xu. Semantic Change Detection with Graph Neural Networks. [paper, code]
5. **Tsepa O***, Burakov R*, Laschowski B, Mihailidis A. Continuous prediction of leg kinematics during walking using inertial sensors, smart glasses, and embedded computing. *Accepted to ICRA 2023.* [paper, code]
6. Kuzmenko D, **Tsepa O**, Kurbis A G, Laschowski B, Mihailidis A. Vision-Based Automated Stair Recognition for Wearable Robotics using Semi-Supervised Learning. *Accepted to IROS 2023.* [paper]
7. **Tsepa O**. Product Embeddings in Retail Industry with Graph Neural Networks. *Bachelor's thesis.* [poster]

EDUCATION

University of Toronto Sep 2022 - Dec 2023
MSc in Computer Science GPA: 4.0

Relevant Courses: Computational Models of Semantic Change, Neural Networks and Deep Learning, Topics in Computational Social Science, Imitation Learning for Robotics.

Igor Sikorsky Kyiv Polytechnic Institute Sep 2018 - June 2022
Bachelor of System Analysis GPA: 87/100

Relevant Courses: Statistics, Neural Networks, Probability Theory, Linear Algebra, Calculus, Natural Language Processing, Data Structures & Algorithms.

SKILLS

Languages	Python, SQL, C++
Machine Learning	Tabular, Natural Language Processing, Computer Vision, Graph Machine Learning
Frameworks	PyTorch, TensorFlow, PyG, transformers, NumPy, Pandas, Sklearn, LightGBM, Matplotlib
Cloud	GCP, SLURM
Tools	Git, Databricks, WandB, Linux, Bash, LaTex

COMPETITIONS & LEADERSHIP

Teaching Assistant at the University of Toronto Sep 2022 - May 2023

Supported the lectures on "Intro to Computer Programming" for 100+ students by answering ongoing questions. Conducted 1-1 **office hours** by clarifying misunderstandings in lecture materials and home assignments.

Admission Board Member at the UofT Summer Program for Ukrainians Dec 2023 - May 2023

Screened and **interviewed** candidates for the University of Toronto Summer Program for Students from Ukraine. During the process, 60+ resumes and motivation letters were reviewed and were conducted 5 interviews to asses candidates' experience, technical strength, and research potential. I also provided candidates with information about the program and answered any ongoing questions. After admission, I **supported** successful candidates in **relocating** and adapting to the new environment. [program]

Toronto Health Datathon 2023 Feb 2023

The solution aimed to **simplify diagnostic workflow** by **pruning the number of assessments** performed in the general internal medicine ward. We use past known lab test results to predict whether the patient deteriorated or was discharged. As a result, we could decrease the amount of additional tests by an average of **3%** while retaining the same outcome. [datathon, presentation]

INT20H 2022 Hackathon Feb 2022

We had to build a model to predict the segment of drivers who will stop using the service, known as churn drivers. Finally, my team won **1st place** in the competition.

Invited Lecturer at Igor Sikorsky Kyiv Polytechnic Institute Dec 2021

Gave a couple of **lectures** about **Graph Neural Networks** for a class of 60+ students as a part of the subject "Neural Networks". The material covered motivation for using GNNs, the math behind them, state-of-the-art architectures, and personal experience of using GNNs in production.

INT20H 2021 Hackathon Feb 2021

The task was to predict the book's rating on the GoodReads service. Our **multimodal** pipeline consisted of scrapping external **text** data about authors, incorporating trainable **images**, and feature generation. The LightGBM model was evaluated based on 5-fold stratified cross-validation. The model was **deployed** to the web service with Streamlit and Heroku. Finally, my team won **1st place** in the competition.

Sibur AI Challenge Jan 2021

Determined if a pair of names belong to the same company. This is necessary to speed up the processing of information about potential customers. Used transformers, over/under-sampling, and text cleaning techniques [solution]

National Bank of Ukraine IT Challenge Mar 2020

Led a team of 3 Data Scientists to create a USD/UAH exchange rate **forecasting platform**. We applied a time-series autoregressive approach consisting of feature generation (rolling statistics with different window sizes), backward feature selection, handling with nonstationarity by differentiating the target, and pseudo-labeling for a long-term forecast. Finally, my team won **2nd place** in the competition. [solution]