David A. Pogrebitskiy

Boston, MA | (303) 810-6328 | pogrebitskiy.d@northeastern.edu
pogrebitskiy.com | github.com/pogrebitskiy | linkedin.com/in/davidpog

EDUCATION

Northeastern University, *Khoury College of Computer Sciences*, Boston, MA **GPA: 3.98/4.00**
Candidate for Bachelor of Science in Data Science, Mathematics Minor Expected Aug 2024

Relevant Coursework: Neural Networks, Machine Learning 2, Data Visualization, Large-Scale Storage & Retrieval, Mathematics of Machine Learning, Linear Algebra, Multi-variable Calculus, Probability & Statistics

TECHNICAL SKILLS

Languages:	Python, SQL, C++, JavaScript, Java, HTML/CSS
Libraries/Frameworks:	Pandas, NumPy, PyTorch, Scikit-learn, Tensorflow, Plotly, HuggingFace
Databases/Platforms:	MySQL, HDFS, Airflow, Dremio, MongoDB, Redis, Neo4j, Spark, ElasticSearch

EXPERIENCE

Automated Execution Analyst Co-op Jul 2023 – Present
TD Cowen / TD Securities *New York, NY*

- Deployed a Python-based FIX Protocol message translator, processing 30,000+ messages daily, which enhanced analytics capabilities for high-touch clients by improving data accuracy and processing speed
- Designed an interactive Plotly dashboard to improve interpretability of quantitative trading signals
- Automated weekly data fetching with Python integrating FINRA OTC and OneTick trade data into a Plotly-based UI, which enhanced the visualization and interpretability of order routing information
- Contributed over 200 commits to the Transactional Cost Analysis report, optimizing code and implementing new features to enhance report accuracy and efficiency

Data Science Research Assistant Jan 2024 – Present
Khoury College of Computer Sciences *Boston, MA*

- Evaluated the performance of instruction-tuned LLMs by benchmarking against various proprietary and open-source models, achieving completeness rates up to 82% in extracting numerical data for meta-analysis
- Curated a dataset of 699 randomized controlled trials (RCTs) by meticulously annotating research publications to extract critical numerical clinical findings for meta-analysis

Director of Technical Workshops Sep 2022 – Apr 2023
Disrupt: The FinTech Initiative at Northeastern *Boston, MA*

- Developed and presented four comprehensive Python lectures, covering topics from introductory programming to data manipulation and financial analysis, successfully engaging over 200 students

Data Engineering Intern Jun 2022 – Aug 2022
Space CAMP *Colorado Springs, CO*

- Implemented data streams, state tables, and a JSON API in Apache Kafka using KSQL, delivering precise satellite and mission status updates, and facilitating interaction between Python and Kafka for developers

PROJECTS

Adversarial Robustness of Neural Nets | *Python, PyTorch, Adversarial Robustness Toolbox* May 2024

- Implemented various adversarial attacks and defenses on models (VGG, LeNet, GoogLeNet, ResNet) by utilizing Python and PyTorch libraries, achieving 97% accuracy against attacks on the MNIST dataset
- Conducted experiments demonstrating up to a 50% improvement in robustness by incorporating BIM, PGD, and FGSM attacks in the training set, highlighting the effectiveness of adversarial training in model security

Author Attribution: NLP | *Python, PyTorch, HuggingFace, Scikit-learn* May 2024

- Utilized Doc2Vec and BERT models for document feature extraction, enhancing feature representation, and implemented classifiers (Logistic Regression, Random Forest, SVM, Neural Networks) to compare embedding techniques, leading to a more accurate author attribution system