APRATIM MISHRA

901 S 1st Street, Apt 1, Champaign, IL - 61820 ◦ (302)-407-2633 ◦ apratim941208@gmail.com

SUMMARY

Ph.D. candidate in Information sciences adept in machine learning and natural language processing(NLP), seeking a suitable position to leverage technical expertise. Interested in data-driven and applied products/services.

PROFESSIONAL EXPERIENCE

AstraZeneca, Maryland, Illinois May 2023 – August 2023

Data Science Intern

- Utilized protein language models (PLMs) and graph neural networks (GNNs) to derive insights, explain relevant protein features, and predict complex biological behaviors, significantly improving modeling characteristics.
- Developed quantized deep learning models (parameter efficient methods) and improved model architecture to decrease computational and memory footprint and facilitate multi-GPU scalability for faster inference times.

The Cline Centre for Advanced Social Research, Champaign, Illinois May 2021 – August 2021

Graduate NLP Programmer

- Engineered an NLP pipeline for quotation extraction and entity classification leveraging tools like spaCy and Stanford CoreNLP, and implemented testing strategies, evaluation metrics, and documentation updates.
- Fine-tuned transformer models like Bert and XLNet for precise quote classification, enhancing evaluation accuracy by focusing on 'span' matching to determine quote beginnings or endings.

Reliable Power Alternatives Corporation, Garden City, New York Nov 2018 – Aug 2019

Energy Data Analyst

- Optimized data retrieval using SQL and developed Python-based Machine Learning pipelines with PySpark and scikit-learn, achieving a 15% cost reduction for clients.
- Supported model testing and deployment using Flask, providing clients with reports on key performance parameters.

Delaware Army National Guard, Wilmington, Delaware

Research Intern September 2017 – January 2018

- Analyzed energy load trends using Python by employing models like Arima, XGBoost, LGBMBoost, and LSTMs.
- Performed one-sample and two-sample statistical tests to verify data integrity for clients and benchmark outcomes.

NYC Department of Citywide Administrative Services, New York City, New York

Energy Efficiency Intern June 2017- August 2017

- Developed an interactive dashboard in Tableau to visualize core business KPIs based on cost savings.

EDUCATION

The University of Illinois at Urbana-Champaign, IL

Ph.D. in Information Sciences August 2019 - Present

Relevant Coursework: Deep Learning, Applied Machine Learning, Network Analysis, Applied Linguistics

- Investigated 'Stance Detection' in tweets, contrasting neural and matrix-based topic modeling techniques. Designed and benchmarked tree-based models like XGBoost and Random Forest.
- Probed the dynamics of scientific 'Hype' in biomedical literature, leveraging NLP to engineer and select features like author ethnicity, gender, and topical interests; subsequently, modeled significance and categorized labels.
- Developed language models for translating natural language to the command line, introducing bash template generation methods using fine-tuned Seq-to-Seq transformers like T5 and slot-filling techniques.

The University of Delaware, Newark, DE

Master in Energy and Environmental Policy May 2018

- Presented a final paper comparing energy forecasting methodologies: univariate, single-step forecasting, and multi-step forecasting, using ARIMA, feature-based machine learning models, and deep learning sequential models.

Birla Institute of Technology and Science, India

Bachelor of Engineering in Chemical Engineering May 2016

SKILLS AND INTERESTS

- **Tools**: Python, TensorFlow, PyTorch, SQL, Linux, Tableau, Power Bi, AWS, MongoDB, Deepspeed, Accelerate
- **Packages**: Scikit-Learn, Pandas, NLTK, matplotlib, spaCy, PyTorch Lightning, PyTorch Ignite, Transformers
- **LinkedIn**: https://www.linkedin.com/in/apratim94/ **GitHub**: https://github.com/apratim-mishra