EMRUL HASAN

(647)-857-2014 | emrul.phy@gmail.com | Toronto, ON
emrulphy | emrulhasan-nlp | emrulhasan-nlp.github.io

PROFILE

Machine Learning Engineer with 5 years of experience in **ML, DL, NLP, and Recommendation Systems**. Well-versed in a wide range of ML frameworks including **PyTorch, Scikit-learn, Pandas, NumPy**, etc. Proven track record of **research and publications** in top-tier peer-reviewed journals and conferences. Enthusiastic about **data-driven problem-solving and AI**.

EDUCATION

Ph.D. in Computer Science/AI, Toronto Metropolitan University, Toronto — Jan 2021 -Present
Graduate Courses: Machine Learning, Deep Learning, Advanced Natural Language Processing

RELEVANT SKILLS & KNOWLEDGE

Languages: Python, C, MATLAB, SQL
Technologies/Frameworks: PyTorch, Scikit-Learn, Pandas, NumPy, MySQL, NLTK, SpaCy, Huggingface, Git, Linux, CUDA

WORK HISTORY

Research Assistant | Toronto Metropolitan University, Toronto — Jan 2021- Present

- Applied **deep learning and NLP** techniques to develop a recommendation system by leveraging customer review.
- **Designed, trained, and fine-tuned machine learning,** and developed a recommendation system, achieving 23% and 19% performance improvement in terms of MAE and MSE compared to the state-of-the-art models.
- Employed **prompt engineering and fine-tuning** techniques to **LLMs (e.g. Llama)** to summarize the document.
- Wrote research papers for publication and effectively **communicated complex analytical** concepts and results to non-technical audiences through presentations at conferences and seminars.

Teaching Assistant, Toronto Metropolitan and Northeastern University, Toronto — Jan 2021 - Present

- Facilitated tutorials on diverse topics including **database management, statistics, Python, ML, NLP, and SQL**.
- Assessed assignments and final projects for a class of 200 students, offering constructive feedback and support.

MACHINE LEARNING PROJECTS

Document Summarization — Jan 2021-April 2021

- Developed document summarizer by using transformer decoder architecture e.g**. T5, BART, and Pegasus**.
- Conducted rigorous testing and evaluation of document **summarization models, fine-tuning parameters** to achieve optimal performance and accuracy in summarizing long text.

Fake News Classification — Jan 2021-April 2021

- Developed a Fake News classifier for early detection, leading to a decreased risk of rumor propagation.
- Experimented with various classification algorithms such as **Naïve Bayes, Logistic Regression, XGboost, SVM, LSTM, and BERT**, achieving 97% accuracy with transformer encoder (e.g. BERT).

Salary Prediction — Sep 2021-Dec 2021

- Conducted extensive exploratory data analysis, including dealing with outliers, null values, categorical features, and correlation analysis to identify the best features.
- Applied various regression models (e.g**. Linear Regression, Random Forest, Gradient Boosting, and XGBoost regression**) to obtain the best-performing model on salary prediction from the job description.
- **Fine-tuned hyperparameter** on the best performing model (**e.g. Gradient Boosting**), achieved an 11% performance improvement compared to the closest performing model (e. g. XGBoost).