Supriya Dara

343-549-4606 darasupriya99@gmail.com LinkedIn Portfolio Github

Education

University of Ottawa Sept 2022 - April 2024

Master of Computer Science concentration in Applied Artificial Intelligence (GPA: 9.30 / 10)

- **Relevant Coursework:** ML, NLP, AI-For Cybersecurity, Distributed Databases and Transaction Processing, Ethics for AI.

Experience

Machine Learning Intern Nov 2021 – July 2022

Aarth Software

- Led **NLP project** to extract and process recipes from 100+ websites, achieving **80% accuracy** using **BERT** for ingredient parsing, text classification, and entity recognition.
- Developed an ensemble ML model using **regression algorithms**, achieving **93% accuracy** in predicting hospital stay lengths for diabetic patients using **SVM**.
- Built a Python web scraper with **BeautifulSoup** for automating metadata extraction from URLs, demonstrating data engineering, achieving a **98% success rate** in scraping product schema from the top 5 Google search results.
- Actively participated in and contributed to **team meetings**, delivering comprehensive progress updates.

Data Science Intern May 2021 – Aug 2021

Exposys Data Labs

- Engineered a Diabetes Prediction model in Python, applying **KNN** algorithm and achieving **88%** accuracy.
- Authored a detailed project report, documenting methodology, analysis, and results, highlighting practical applications of ML techniques.
- Collaborated with **cross-functional teams** to translate project requirements into technical solutions.

Machine Learning Intern May 2020 – Aug 2020

National Instruments

- Implemented and fine-tuned a sentiment analysis model in Python, achieving **95%** accuracy through ML algorithms and optimization techniques.
- Leveraged **Tableau** and **Power BI** for data visualization, presenting sentiment trends to support data-driven decision-making.

Projects

Log-based Anomaly Detection | *Python, PyTorch, Scikit-learn, Transformers* Link

- Designed a novel methodology that leverages RoBERTa for transforming unstructured log data, achieving an **F1-score of 0.99** in identifying anomalous logs, surpassing traditional TF-IDF methods.

Phishing URL Detection | *Python, Scikit-learn, PyTorch, Transformers, SHAP* Link

- Led RoBERTa-based project detecting phishing URLs, analyzing lexical, syntactic, and semantic features, and integrated SHAP for model interpretability with **98.34% accuracy**.

Automated Code Review with NLP | *Python, Scikit-learn, NLTK, PyTorch, Transformers* Link

- Created a RoBERTa model for optimizing code review in software development, achieving **97.7% accuracy** in analyzing GitHub comments for both accuracy and efficiency.

Automated Essay Scoring System using LSTM and NLP | *Python, TensorFlow, Keras, NLTK, LSTM* Link

- Constructed a 2-layer LSTM network and used NLP techniques to evaluate and rate essays, achieving a **QWK score of 0.92**.

Technical Skills

- **Languages:** C, Python, Java, SQL, R (learning).
- **Technologies/Frameworks:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, PyTorch, TensorFlow, Keras, NLTK, SHAP, OpenCV.
- **Concepts:** : Statistical Modeling, Deep Learning, Large Language Models (LLMs)

Soft Skills

- Excellent written and oral communication skills honed through roles such as **Head Teaching Assistant.**
- Demonstrated adaptability and a commitment to **continuous learning** through internships, workshops, and mentoring, constantly enhancing my expertise in emerging technologies.
- Remarkable **problem-solving** and **analytical skills** developed through various academic projects.
- Fostered **collaboration** and strengthened community engagement by coordinating 10+ events as part of **Voice4Girls**, promoting teamwork and collective impact.