Dipankar Srirag

Student Visa (Subclass 500) - Limited Working Rights
dipankar.srirag@gmail.com | Website | Google Scholar | LinkedIn | Sydney NSW

SUMMARY

Research assistant, **Casual Academic**, and **Master of Information Technology** graduate at the University of New South Wales with strong background in Natural Language Processing (**NLP**), Large Language Models (**LLM**s), Machine Learning and Data Science. My Master's thesis involves adapting LLMs to be dialect-robust, supervised by **Dr. Aditya Joshi**.

EDUCATION

Master of Information Technology, WAM: 81.3 — Sep 2022 - Sep 2024
University of New South Wales — Sydney NSW

- *Courses*: Principles of Programming; Data Services and Engineering; Data Analytics for Graphs; Big Data Management; Computer Vision; Neural Networks and Deep Learning.

Bachelor of Information Technology, GPA: 8.38 — Jul 2018 - Jul 2022
Manipal Academy of Higher Education — Manipal INDIA

- *Courses*: Machine Learning for Data Analytics, Information Retrieval, Engineering Mathematics, Cyber Security.

EXPERIENCE

Research Assistant — May 2024 - Present
UNSW — Sydney NSW

- Collect Google Places reviews and Reddit comments across three locales (Australia, India, and the UK) to build a benchmark dataset for sentiment and sarcasm classification, funded by **Google Research Scholar** grant.
- Develop better data sampling strategies to make the existing datasets more challenging for the LLMs.
- Designing and executing experiments involving all classes of LLMs (encoder-only, decoder-only, and encoder-decoder), which involve in-context learning and parameter-efficient fine-tuning.

Casual Academic — May 2024 - Present
UNSW — Sydney NSW

- Leads deep learning tutorials for 40 students across two groups, covering topics such as neural networks, ensemble learning, unsupervised learning, and reinforcement learning.
- Mentors student groups through their end-to-end artificial intelligence projects.
- Grades assignments with detailed, constructive feedback to enhance understanding.

RPA Analyst Intern — Jan 2022 - Jun 2022
Deloitte Touche Tohmatsu India LLP — Bangalore INDIA

- Trained to execute end-to-end process automation solutions using Automation Anywhere, Blue Prism, and UiPath, to enhance operational efficiency and cut down manual errors.
- Built a bot leveraging functionalities of BluePrism, to achieve a perfect score of 100% on online RPA Challenge for Invoice Processing.

PREPRINTS AND SUBMISSIONS

Srirag, D., Joshi, A. and Eisenstein, J. (2024) 'Predicting the Target Word of Game-playing Conversations using a Low-Rank Dialect Adapter for Decoder Models', arXiv.
Submitted to ACL Rolling Review.

Srirag, D., Sahoo, N. R. and Joshi, A. (2024) 'Evaluating Dialect Robustness of Language Models via Conversation Understanding', arXiv.
Submitted to Social Impact Track at AAAI 2025.

PROJECTS

Improving Dialect Robustness of Language Models using Game-playing Conversations

- Proposes a novel methodology to evaluate language models using a pre-existing dataset of dialogues.
- Evaluate open-source (Llama, Mistral, and Gemma) and closed-source (GPT-3.5 and 4 Turbo) models on novel task of masked target word prediction on conversations from dialogue games.
- Proposes a novel adapter-based architecture to make pre-trained decoder models (Mistral and Gemma) robust to other dialects of English.
- Train dialect adapters using a pseudo-parallel corpus of naturally occurring conversation pairs.
- Bridge the performance gap between Standard American English and Indian English by 12% on word similarity and 25% on accuracy on the masked target word prediction task.

Comparative Analysis of Abstractive Summarisation Techniques

- Conducted a comprehensive study comparing the effectiveness of static versus dynamic, context-aware embeddings in dialogue summarisation, using models like Seq2Seq, BART-base, and FLAN-T5.
- Achieved marked improvements in summarisation models, with fine-tuned BART-base showing a ROUGE-1 score increase from 0.2866 to 0.4215 outperforming a custom Seq2Seq model.
- Identified key strengths and weaknesses in NLP models, emphasising the balance between computational resource needs and the precision of context-aware embeddings in producing coherent and accurate summaries.

SKILLS

Programming Languages: Python, R, SQL, C.
Libraries: PyTorch, TensorFlow, Keras, scikit-learn, LangChain.
Data Visualisation tools: Matplotlib, Seaborn, Plotly, Tableau.
Version Control and Collaboration: Git, GitHub, Jupyter Notebooks.
Others: Report writing, Research methods, LaTeX.

COMMUNITY

Treasurer - UNSW Motorsport Society