Akshay Srinivasan

📱+1 343-575-2491 | ✉ s.akshay0799@gmail.com | ⌥ github.com/Akshay0799 | in linkedin.com/in/akshay–srinivasan

Work Experience

Novacene AI *Ottawa, Ontario*
Data Scientist *Jun 2023 - Oct 2024*

Machine Learning
- Designed and developed an end-to-end **AutoML** pipeline for tabular classification tasks in the **Cybersecurity** domain on Azure ML studio
- Implemented components for Data Preparation, Feature Processing, Encoding, and Training to automate the management of **Alert Triages** and detect malicious activity in IT infrastructure systems
- Facilitated the deployment of this solution in real-time for use at a prominent **Big Four firm**

Natural Language Processing
- Utilized NLP algorithms such as Bi-Kmeans clustering and Clause Extraction with **Transformers, HuggingFace, and Scikit-Learn** frameworks to analyze customer feedback and reviews
- Leveraged **LLMs** through **LangChain** to perform Clustering and Theme generation, implemented **Retrieval-Augmented Generation** (RAG) for document Q&A, enhancing the accuracy and relevance of responses derived from customer data
- All efforts contributed to generating actionable marketing insights by gauging product perception and market reception

Software Development
- Implemented **Docker containerization** for backend services, seamlessly integrating Machine Learning and Natural Language Processing enrichers using **Django** framework
- Engineered an ingestion pipeline to handle **nested JSON** structures and convert them into machine readable CSVs to be used for Machine learning pipelines

National Research Council of Canada *Ottawa, Ontario*
NLP Research Intern *May 2022 - Dec 2022*
- Introduced perturbation-based approaches to create **adversarial test sets** across languages and improved the robustness of state-of-the-art Named Entity Recognition(NER) Systems using adversarial data augmentation by **5-10%**
- Utilized **Weak Supervision** to automate the annotation of datasets for the NER task by aggregating ML models, domain-specific heuristic rules and gazetteers
- Devised **few-shot learning** approaches to efficiently train transformer-based Multilingual language models on an unseen language with **40%** less training data

Education

University of Ottawa *Ottawa, Ontario*
Master's in Computer Science: Applied AI (GPA: 3.8/4) *Sept 2021 - May 2023*
- Graduate Teaching Assistant for undergraduate and graduate courses
- **Courses:** Machine Learning, Natural Language Processing, Virtual AI Assistants, Evolutionary Computing and Ethics in AI

SRM Institute of Science and Technology *Chennai, India*
Bachelor's in Computer Science Engineering (GPA: 9.2/10) *July 2017 - May 2021*

Projects

Automated Extraction of Regulatory Requirements from Legal Texts
- Built an NLP pipeline using **BiLSTM** to parse Legal texts and extract important requirements from regulations
- Trained multiple BiLSTM binary classifiers using **few-shot learning** so that different types of requirements can be identified and fetched with an accuracy of **95%**

Med Parser
- Created a web application to parse medical transcriptions and automatically generate patient reports through **Named Entity Recognition** (NER)
- Fine-tuned **BIOBERT** on biomedical NER data and set it up in a **NodeJS** server in the backend to predict and extract named entities real-time

Image Classifier API
- Built a furniture image classification pipeline using **VGG16** with **97%** accuracy via transfer learning
- Wrapped the classifier into an API using **Django** server and containerized the API through **Docker**
- Employed Github Actions to implement **CI/CD** workflows onto the codebase

Skills

Python | PyTorch | Tensorflow | HuggingFace | NLP | LLM | LangChain | Pandas | LangChain | Docker | GPU Cluster Management

Scikit-learn | Django | AI | SQL | NER | GAN | MongoDB | SpringBoot | REST | API | JAVA | ReactJS | NodeJS | Git | Linux

Publications

A Multilingual Evaluation of NER Robustness to Adversarial Inputs
Srinivasan, A., Vajjala, S. (2023). A Multilingual Evaluation of NER Robustness to Adversarial Inputs. Proceedings of the 8th Workshop on Representation Learning for NLP, ACL 2023, pp 40–53 (Paper Link)

Favourable SubPopulation Migration Strategy in Parallel Genetic Algorithm
Chandar, A., Srinivasan, A., Anand, G.,P. (2022). Favourable subpopulation migration strategy for travelling salesman problem. International Journal of Business Intelligence and Data Mining, Vol. 20, 3, 2022 (Paper Link)