Jeevanshi Sharma

LinkedIn, GitHub, Medium

Email : jeevansh@ualberta.ca
Mobile : +1-587-937-4711

Education

- **University of Alberta** — Edmonton, AB, Canada
  Master's in Computing Science - Multimedia Specialization (GPA 3.5/4) — *Sep'23 – Nov'24*
- **Zakir Husain College of Engineering and Technology, AMU** — Aligarh, UP, India
  Bachelor of Technology in Electrical Engineering (GPA 8.57/10) — *Aug'18 – Sep'22*

Experience

- **Okaki Health Intelligence Inc** — Remote / Calgary, AB, Canada
  Machine Learning Engineering Intern — *Nov'23 - Nov'24*
  - **NLP Pipeline Engineering and Deployment**: Engineered NLP pipelines for automating cognitive assessments like MoCA, optimizing AI-driven analytics, and integrating scalable architecture using Azure OpenAI.
  - **Advancing Audio-Speech NLP**: Scaled speech-to-text integration using Whisper, WhisperX, and Azure Speech models for precise transcription and diarization with unified dataclass output. Enhanced the transcription library's structure by implementing object-oriented principles, which improved code maintainability, modularity, and ease of scaling.
  - **Realtime-AI Assistant**: Built and optimized a real-time medical assistant for clinical assessments with vision capabilities using LiveKit's Pipeline.
- **Ersilia Open Source Initiative** — Remote / Barcelona, Spain
  Outreachy Intern — *Dec'22 - Mar'23*
  - **Deep Learning and NLP Integration**: Developed deep learning models (e.g., Biomed-RoBERTa) for drug taxonomy via textual bioassay data embeddings.
  - **Software Development and Optimization**: Built the Auto-TabNet package, leveraging Optuna for hyperparameter tuning. Contributed to debugging and documentation.
- **Emplay Inc** — Remote / Dublin, CA, USA
  Data Science Intern — *Oct'21 - Apr'22*
  - **API Development**: Developed and deployed the Topic-Suggestion API for generating relevant keyphrases and tags.
  - **Keyphrase Extraction**: Implemented supervised keyphrase extraction methods with precision scores of 0.37.
  - **NLP Content Moderation**: Built a content moderation pipeline to detect toxicity using pre-trained language models.
- **Dalhousie University** — Remote / Halifax, NS, Canada
  MITACS Globalink Research Intern — *Jun'21 - Sep'21*
  - **Enhanced Natural Language Understanding**: Optimized pre-processing of datasets for compositional tasks, enabling efficient pre-training of language models such as ALBERT-xxlarge-v2, BERT-large-uncased, and RoBERTa-large.
  - **Advanced Semantic Evaluation**: Improved semantic evaluation techniques and accuracy by 15-20% for albert-xxlarge-v2 on benchmark datasets (ex: COPA, Winogender, aNLI, PDP), advancing zero-shot common-sense reasoning tasks.

Projects

- **Budget AI**: Built REST APIs for a client application using MicroServices architecture. Developed an LSTM model for expense forecasting integrating AI educational content. Winner of CalHacks 8.0 ***Microsoft Cloud Challenge: Building for Social Good*** and ShellHacks ***Code What Matters w/ JP Morgan Chase***.
- **Locust Location and Apprehension Module (LLAM)**: Developed a web app to track locusts using satellite imagery and machine learning, predicting swarm attacks with 76% accuracy. Winner of ***Microsoft Azure AI Hackathon '21***.
- **Dashboard for Indian Ministry of Statistics and PI**: Built a data visualization dashboard for the National Account Statistics, with an admin panel to upload, modify, and display statistical data as graphs. First Runner-Up at ***Smart India Hackathon - 2019 edition***.

Publications

- **Machine Learning For Classification Of Antithetical Emotional States**: 2022 *IEEE Xplore*
  J. Sharma, R. Maheshwari, Y. U. Khan
- **Evaluating Performance of Different Machine Learning Algorithms for the Acute EMG Hand Gesture Datasets**: 2022 *Journal of Electronics and Informatics* J. Sharma, R. Maheshwari, S. Khan, A. A. Khan
- **Evaluating CNN with Oscillatory Activation Function**: 2022 *arXiv Preprint* J. Sharma

Skills

- **Programming Languages**: Python, SQL
- **Programming Frameworks**: PyTorch, Flask, Huggingface, Langchain, Fast API
- **Tools**: Docker, Azure AI, AWS Bedrock, Linux Command Line, Git, Pandas, Numpy
- **Technologies**: Deep Learning, NLP, Vector Databases, Transformer Architectures, Prompt Engineering, RAG