

Atharva Joshi

+1 (312) 532-5083 | Boston, MA | atharvajoshi011@gmail.com | [in](https://www.linkedin.com/in/atharvajoshi011) [atharvajoshi011](https://www.linkedin.com/in/atharvajoshi011) | [Qldr7](https://github.com/ldr7)

EDUCATION

Northeastern University (NEU), Boston, MA

Dec 2023

MS in Electrical and Computer Engineering, Specialization in Machine Intelligence.

Courses: Deep Learning, Advanced Machine Learning, Reinforcement Learning, Parallel Processing.

Birla Institute of Technology & Science, Goa, India

May 2019

Bachelor of Engineering (Hons.) in Electronics & Instrumentation

Courses: Deep Learning, Machine Learning, Quantum Computing, Medical Devices, Cognitive Neuroscience.

EXPERIENCE

Northeastern University - SPIRAL Lab

Research Assistant

Jan 2023 – May 2023

- Replicated the MEDNET paper in PyTorch to perform weakly supervised semantic segmentation of Reflectance Confocal Microscopy mosaics and detect 4 types of melanoma lesions, enabling downstream development.
- Studied Vision Transformer & Contrastive-learning based approaches to compare against the MEDNET model.

Abiomed - Applied Research, Boston, USA

Data Scientist

July 2022 – Dec 2022

- Led the Left Ventricle Volume prediction, greenfield animal research project, achieving mean RMSE of 14% (0-250 mL) across hemodynamic states, constructing a deep LSTM network using implanted heart pump data as input.
- Delivered a large increase (10x) in training data using Fourier based resampling with physiological constraints.
- Deployed an LLM (BioBERT) NER pipeline to extract drug names, enabling use of ICU notes as data. Utilized HuggingFace API's & analyzed performances of multiple Large Language Models.

Dell Technologies - UDS, Bangalore, India

Software Development Engineer

July 2019 – July 2021

- Decreased database transaction latency by 15% and fixed intermittent scale down of transaction segments.
- Initiated availability of the stream tags feature in the product by implementing REST API's using Swagger.
- Improved verification of product guarantees by implementing & evaluating the horizontal scalability tests suite.

Morningstar - Quant Research, Mumbai, India

Machine Learning Engineer

Jan 2019 – June 2019

- Developed a BERT based sentiment analysis model with few shot learning to predict international equity prices using News API data extracted from built ETL pipelines. Presented results led to buy-in of key stakeholders.
- Reduced MSE of the price forecasting model by more than 50%, implementing an autoregressive LSTM network.
- Enhanced investor decision making for thematic investments by developing an NLP driven heuristic utilizing textual data from financial news providers to assess company-market theme alignment.

Tamil Nadu e-Governance Agency, India

Machine Learning Engineer

May 2017 - July 2017

- Introduced ML based technique to match missing person reports with unidentified dead bodies, building an XGBoost classifier with more than 90% F1 score. Utilized GIS data for measuring geographical proximity.

SELECT PROJECTS

- **Traffic Forecasting:** Improved the MAE on hour ahead forecasts by 6.5% by modifying the STGCN architecture. Compared performance against DCRNN, GMAN & A3T-GCN models implemented in a resource constrained setting
- **YouTube Playlist Recommendation:** Built a 7 class YouTube Playlist recommender system with an accuracy of 70% using data from transcripts of videos. PySpark is used to process transcripts, the T5 model summarizes them and OPT-30B generates embeddings that are used to compute similarity, which is the basis of recommendations.
- **nanoGPT:** Built an LLM from scratch to generate Shakespearean text. Carried out ablation studies of various components using relative training performance. Compared against Falcon, ChatGPT & fine-tuned GPT2.
- **EEG Signal Classification for Epilepsy Detection:** Constructed predictive classification models using EEG data from 500 patients, obtaining an accuracy of 96% averaged over 2000 splits of the data.

TECHNICAL SKILLS

- **Programming Languages:** C++, Python, Java, MATLAB, SQL, PySpark
- **Frameworks:** Pytorch, Tensorflow, Keras, Kubernetes, Streamlit, ChromaDB, LangChain, AWS