

Manjit Ullal

Boston, MA 02118 | +1 (857)-869-4157 | ullal.m@northeastern.edu | [LinkedIn](#) | [GitHub](#)

EDUCATION

Northeastern University, Khoury College of Computer Sciences, Boston, MA

Expected - Dec 2021

Candidate for a Master's of Science in Data Science

Related Courses: Probability, Statistics, Machine Learning, Deep Learning, Data Mining, Big Data, Natural Language Processing.

Visvesvaraya Technological University, Belgaum, India

Aug 2006 - May 2010

Bachelor's Degree in Computer Science

Related Courses: Algorithms, Data Structures, Database Management, Calculus, Linear Algebra.

SKILLS

Languages	Python, R, Java, Javascript, MATLAB
Framework	TensorFlow, Pytorch, Keras, Scikit-learn, SciPy, NumPy, Pandas, Shap, Nltk, Spacy, Flask
ML Algorithms	Regression, Classification, Clustering, SVM, Random Forest, SVD, PCA, Kmeans, Deep Learning, Artificial Neural networks, CNN, RNN, Artificial Intelligence, Naive Bayes, XGBoost, Statistical Modelling, Autoencoders Recommendation Systems, NLP, Reinforcement Learning
Cloud	AWS, EC2, Sagemaker, GCP, Lambda, Docker, Kubernetes, Elasticsearch, Kafka
Visualization	Bokeh, Seaborn, Matplotlib, Tableau, Plotly, Tableau, Big Query
Database	SQL, MongoDB, Hadoop, Bigtable, Postgre, Spark, Pyspark, Mapreduce
Tools	Git, GitHub, Jupyter Notebook, REST

PROFESSIONAL EXPERIENCE

Northeastern University (Boston, Massachusetts) – Data Science Teaching Assistant

Jan 2020 - Dec 2020

- Tutor data science through case studies. Illustrate techniques for data wrangling, transforming and modeling.

Rose Spatio Temporal Lab Northeastern University (Boston, Massachusetts) – Research Assistant

May 2020 - Oct 2020

- Collaborated with Adobe Research on a deep neural network model to predict unevenly sampled spatiotemporal events as neural point processes for events forecasting. Modeled the data as a continuous sequence of point-process.
- Designed GRU/LSTM/Transformer neural network using Pytorch on Foursquare/ NYC taxi/ Earthquake datasets to learn nonlinear dependency over space and time. Demonstrated superior performances on benchmarks.
- Derived inference by approximate the temporal intensity function with an GRU recurrent neural network and estimate the spatial intensity with Kernel density estimation.

Infosys Limited (Singapore, Singapore) – Data Analyst

May 2018 - Jul 2019

- Designed and implemented machine learning pipeline by ingesting employee data to generate insightful features. Built well-tuned predictive time-series models to predict budget and access risk, with optimal test results.
- Performed exploratory analysis to identify patterns in large data. Engineered features thereby reducing root mean square error of the model. The model Improved decision-making in resource alignment and reduced resource allocation time by 20%.
- Interacted with stakeholders and tested hypothesis by setting up A/B and statistical test. communicated the insights through visualization. Created dashboard to allow analytics for stakeholders. Evaluated tools to satisfy business requirements.
- Rated incident severity through NLP model based on sentiment analysis. Discussed metrics & validated results with users.
- Discovered key metrics impacting KPI, wrangle & transformed dataset. Improved customer KPI by 32% through automation.

PROJECTS

Computer Vision

[Technologies used: Python, Pytorch, OpenCV, CNN, Transfer Learning, ImageNet]

- Coded a program that can detect hand from webcam and segment the hand and count the number fingers using OpenCV.
- Segment the hand using absolute difference between background and frame and used convex hull method to count fingers.
- Programmed VGG19 deep CNN network to transfer the style of one image onto other.
- Noisy student training of efficientnet to classify images using semi-supervised approach, achieving accuracy of 88.4%

Question and Answering NLP

[Technologies used: Python, Pytorch, nltk, Bert, LSTM, RNN, Sklearn, Shap, Albert]

- Answer questions relating to a passage using series of NLP models on Stanford QA dataset. Provide endpoint API for inference.
- Experimented with Bert, Electra, Alberta to solve the problem of unanswerable questions, achieved F1 score of 92.1 on test.

Study on Recommendation System

[Technologies used: Python, Pytorch, Tensorflow, Sklearn, surprise, deep learning]

- Studied the pros and cons of content, collaborative filtering algorithms to deep learning-based models on Movielens dataset.
- Applied neural collaborative filtering to remedy the shortcoming of traditional models.
- Employed t-sne dimensionality reduction to visualize/cluster relationship in high dimensional big data.
- For large dataset generated missing user ratings by implementing the alternating least squares (ALS) algorithm in spark.mllib

Towards Artificial Intelligence Author/Contributor [link](#)

- I write articles covering Data Science, Machine Learning and Artificial Intelligence on popular Medium Publication.