# Kartik Mohan

Boston, MA | +1 (857)-891-3987 | kartik.mohan99@gmail.com | LinkedIn | Github | Portfolio

## EDUCATION

**Northeastern University** - Boston, MA — *May 2023*
**Khoury College of Computer Sciences** — GPA – 4.0
*Master of Science in Data Science*
Coursework: Data Management and Processing, Supervised ML, Unsupervised ML and Data Mining, Deep Learning

**SRM Institute of Science and Technology** - Chennai, India — *July 2017 – June 2021*
*Bachelor of Technology in Computer Science and Engineering* — CGPA – 9.07
Coursework: Python Programming, Machine Learning, Probability and Statistics, Linear Algebra

## PROFESSIONAL EXPERIENCE

***Data Science Intern* | LNM Institute of Information Technology, India** — *May 2020 – July 2020*

- Deployed an end-to-end webapp on Heroku for real-time data analytics to gain insights into the COVID19 trends and identify the sentiment of people towards the vaccination drive across the globe
- Visualized and incorporated descriptive models including heat maps, geographic maps, bar charts, and line charts for representing different statistics of the COVID19 outbreak into the webapp using Plotly
- Implemented sentiment analysis of vaccination tweets leveraging NLP methodologies to extract key features using NLTK, TF-IDK Vectorizer and TextBlob libraries
- Built subsequent machine learning models using Naive Bayes and Bi-LSTM, achieving an accuracy of 84%

***Data Analyst Intern* | Tata Consultancy Services, India** — *June 2019 – Aug 2019*

- Designed a relational database system for inventory management for a library using MS SQL
- Deployed the database in an AWS RDS instance and integrated it with a front-end admin portal
- Visualized the inventory data through a dashboard by establishing a live connection between Tableau and AWS RDS

## PROJECTS

**Time Series Forecasting of Household Power Consumption**

- Analyzed 2+ million rows of time-series data to forecast the household electricity consumption using Python
- Performed EDA to provide insights into the household power consumption by day, week, semi-month, month, and year and understand the trend and seasonality
- Implemented SARIMA, LSTM, Random Forest, and Linear Regression models to predict energy consumption and compared their metrics wherein LSTM achieved the lowest MAPE of 18.7%

**Attention based Image Scene Description System (Ongoing)**

- Encoded around 8000+ images using transfer learning from the ResNet50 model pre-trained on ImageNet
- Extracted word embedding using GloVe and implemented LSTM for caption generation
- Achieved a BLEU score of 0.45 and METEOR score of 0.34

**Medical Healthcare Insurance Cost Prediction**

- Performed Exploratory Data Analysis on the Medical Health Insurance dataset consisting of 23,000 rows to determine the contributing factors and predict the health insurance cost
- Conducted Feature Engineering, Hyperparameter Tuning, and Feature Transformation to improve the model accuracy
- Implemented Linear Regression, Polynomial Regression, and Random Forest Regressor with an accuracy of 87%

**Music Genre Classification and Recommendation System**

- Generated a dataset by extracting audio features from multiple audio files using the Librosa library
- Carried out EDA to visualize the audio data using spectrograms, boxplots, and co-relation heat maps
- Leveraged XGboost, Random Forest, LGBM, CNN, and Deep Neural Network to achieve an accuracy of 93.33%
- Devised a recommendation system to generate top 10 songs using Nearest Neighbours and Cosine Similarity

## SKILLS

| | |
|---|---|
| **Programming Languages:** | Python, R, SQL, C++ |
| **Libraries & Framework:** | PyTorch, Tensorflow, Keras, NLTK, Scikit-learn, Pandas, Numpy, Plotly, Dash, TextBlob, StatsModels, SpaCy, Dplyr, Tidyverse, Ggplot2, R Shiny, Streamlit, Spark |
| **Databases:** | Microsoft SQL Server, Oracle 11g, MySQL, AWS RDS |
| **Machine Learning:** | Regression, Classification, Clustering, NLP, Computer Vision, Deep Neural Network |
| **Tools & Technologies:** | MS Excel, Tableau, Jupyter, Google Colab, RStudio, Power BI, Git, Google Data Studio |

## PUBLICATION

"Global Analysis of COVID-19 and its Impact", *Annals of RSCB*, Apr. 2021, Volume 25: Issue 4. (Link)