# Kartik Ullal

Boston, MA 02215 • (857) 334-3460
kartikullal99@gmail.com • www.linkedin.com/in/kartikullal • Portfolio Website

## EDUCATION

**Master's of Science, Data Science** Expected May 2023
**Northeastern University**, Khoury College of Computer Sciences, Boston CGPA: 3.83/4

- Related Courses: Linear Algebra and probability for Data Science, Introduction to Data Management and Processing, Machine Learning, Natural Language Processing

**Bachelor's in Engineering, Information Technology** June 2021
**Mumbai University**, Thadomal Shahani Engineering College, Mumbai CGPA: 3.71/4 (9.18/10)

- Related Courses: Database Management Systems, Data Mining, Business Intelligence, Artificial Intelligence, Big Data Analytics, Object Oriented Programming

## EXPERIENCE

**Data Engineer Co-op** July 2022 - Present
Wayfair LLC, Boston, Massachusetts

- Developed a self-healing dedupe functionality for their data pipeline in Google BigQuery and integrated it with slack and Email using Python, which increased Service Level Objective(SLO) meetup from 86% to 92%
- Designed an anomaly detection algorithm by creating a time series forecasting model using Prophet, and send alerts when the true rates are 2 standard deviations away from forecasted rates to quickly resolve the anomalies
- Built a variety of ETL procedures, data models, and Looker dashboards for FinTech and Loyalty in the Strategic Programs Team and resolved the data needs for a group of 100+ business stakeholders across various channels

**Data Science Intern** June 2019 - December 2019
Navlakhi Education, Mumbai, Maharashtra

- Designed database and preprocessed data using MySql and Python for JEE testing software and the Mumbai Fashion Academy website
- Designed a simple recommendation system for Mumbai Fashion Academy using item-item collaborative filtering
- Managed a team of six interns and reviewed requirements for the management team from Mumbai Fashion Academy

## PROJECTS

**Resume Parser and Job Similarity Scoring, Northeastern University** January 2022 - May 2022
**GitHub Link:** Resume and Job Similarity Scorer

- Collaborated with two engineers to parse a resume from pdf to JSON format using PyMuPDF and build a Named Entity Recognizer using Spacy to recognize skills from the resume and provided job description.
- Utilized the Genism and Huggingface libraries from Python to convert the list of skills into word embeddings using Doc2Vec and SentenceBERT, and calculated similarity using Cosine Similarity.
- Evaluated the NER model with an accuracy of 85%, and received a 92% similarity score for my resume and a data science job description.

**Predicting Depression, Stress, and Anxiety scores, Northeastern University** September 2021 - December 2021
**GitHub Link:** DASS Prediction

- Coordinated with two engineers to build a machine learning model to predict the scores for depression, stress, and anxiety from the DASS survey, using only personality traits and demographic data
- Preprocessed and cleaned the data to make it fit for prediction, and visualized different parameters to find a high correlation with depression, stress, and anxiety scores
- Implemented a StackCVRegressor Algorithm, that used predictions made by Lasso regression, Ridge regression, ElasticNet regression, Gradient Boosting regression, LightGBM, and XGBoost regression as features to predict the scores, with an RMSE of 10.

**Detection of Covid-19 using Chest X-rays, Hack India Crisis Hackathon** March 2020 - April 2020
**GitHub Link:** Covid-19 Detection

- Cooperated with a team of four engineers to detect the presence of Covid-19 among patients with the classification abilities of Dense Convolutional Neural Networks (DenCOvseNets) on Chest X-rays
- Integrated, preprocessed, and resized images of five different datasets to form a complete dataset of 17,194 samples of chest X-rays
- Trained the classification model with different train-validation split ratios and achieved an accuracy of 93.496% on the testing data

## TECHNICAL KNOWLEDGE

**Programming Languages:** R(ggplot2, dplyr), Python(Scikit-Learn, Numpy, Pandas, PyTorch, Tableau), SQL
**Skills:** Machine Learning, Deep Learning, Data Analytics, Natural Language Processing, Data Visualization (Tableau, Looker, PowerBI), Kronos, Google Cloud Platform, Excel, PyTorch, BigQuery