

CHAITRA VISHWANATHA HEGDE

cvh255@nyu.edu | <https://www.linkedin.com/in/hegdechaitra/> | <https://github.com/HegdeChaitra>

+1(201) 238-8359 | 212 Beacon Avenue, Jersey City, New Jersey - 07306

EDUCATION

MASTER OF SCIENCE, DATA SCIENCE - GPA: 3.783/4.0

2019 (Expected)

New York University, Center for Data Science

New York, USA

- Relevant Course Work: Deep Learning for Medicine, Machine Learning, Big Data, Probability and Statistics for Data Science, Introduction to Data Science
- Received Moore-Sloan research grant to conduct research on image segmentation task using Deep Learning

BACHELOR OF ENGINEERING, COMPUTER SCIENCE - GPA: 3.8/4.0

2013 - 2017

Visvesvaraya Technological University

Karnataka, India

- Relevant Course Work: Algorithms and Data Structure, Object Oriented Programming in C++, Graph Theory, Operating Systems

EXPERIENCE

AI Labs, NYU Langone's Department of Radiology, NYC, USA

June '18 - Aug '18

- Working as AI research intern on brain tissue segmentation task using novel Deep Learning methods

NYU Center for Data Science, Junior Data Scientist, NYC, USA

Jan '18 - Present

- Working with a post-doc Moore Sloan Data Science fellow at NYU CDS on modeling human mobility in urban areas
- Predicted geo-temporal variables by studying massive amount of highly sparse and unorganized datasets from NYC Open Data
- Discovered and Unearthed un-common hypothesis in context of human mobility using Foursquare data, NYC Taxi data, Ambulance data and Phone call data

Grinions.com, Developer Intern, Bangalore, India

Jan '16 - Mar '16

- Worked on a user interactive Web Application for collecting contributions for social events from attendees
- Wrote Well organized codes using Python's Flask framework and MySQL to handle huge traffic on the website
- Built an attractive user interface for the website using Google's Material Design Lite, HTML-5, CSS and JavaScript

TECHNICAL SKILLS

Programming Languages :

Python, C/C++, SQL, HTML, CSS

Machine Learning Libraries :

PyTorch, PySpark, Sci-kit learn, Pandas, Matplotlib, Numpy, nltk, Seaborn, GeoPandas

Software and Facilities :

Jupyter Notebook, AWS, HPC, UNIX, GitHub

RELEVANT DATA SCIENCE PROJECTS

Knee Cartilage Segmentation Using Deep Learning Methods

Feb - May '18

- Performed highly imbalanced multi-class segmentation of knee cartilages in MRIs using novel deep learning models like U-Net, V-Net and ensemble of similar such models. Final model achieved state-of-the-art results and surpassed a human radiologist level
- Performed Perturbation analysis to understand important features of MRI. Built confidence maps that showed the model's confidence of the voxel-level predictions.

Dataset Search Engine

Feb - May '18

- Built a sophisticated search engine to search for datasets in NYC Open Data portal that contains approximately 1500 datasets (~650 GB)
- Developed 11 advanced search functionalities, currently are not supported by NYC Open Data website, to help the user find relevant datasets
- Used PySpark to parallelize, hence speed up the execution of code. Created data summaries to quickly speed up the search process

NYC Taxi Demand Prediction

Feb - Apr '18

- Built a Regression Tree based taxi demand prediction model to predict the number of taxis at a given space and time in NYC with 0.9 R-square performance
- Enhanced model's performance by adding novel taxi-demand relevant features from multiple open source datasets. Identified demand drivers using feature importance maps

Study on Factors influencing NYC Housing Prices

Apr - Apr '18

- Runner-up of "Enigma Datathon", conducted at NYU Center for Data Science
- Extracted and Processed 5 NYC open datasets to characterize a location in NYC in term of approachability, safety, standard of living, etc.
- Combined these location relevant features with the housing price dataset to enhance house price prediction model's performance and understand factors driving the house prices in NYC

Bank Marketing Campaign Analysis

Sep - Dec '17

- Analyzed the prior marketing campaigns of a Portuguese Bank using various ML techniques like Logistic Regression, Random Forests, Decision Trees, Gradient Boosting and AdaBoost and predicted if the user will buy the Bank's term deposit or not
- Recommended, the marketing team, ways to better target customers using feature importance maps and business intuition

Yelp Recommendation Engine

Sep - Dec '17

- Winner of "Best First Year Project" award at NYU CDS Academy Awards
- Built a recommendation engine for recommending restaurants to Yelp users using traditional models like Cosine similarity based model, SVD and Alternating Least Square model; Rating Matrix was very sparse with sparsity of 99.4%
- Developed advanced models like Stochastic Gradient Descent - based model, Neural Network based model, Random Forest Regressor based model and an ensemble model to achieve higher performance on the sparse rating matrix completion task

Virtual HR System

Feb - May '17

- Built an user friendly, interactive Android application from scratch to help recruitment wing of the undergrad college.
- Tackled the cold-start problem of prediction by representing each of the student and the company's requirement as a vector of values based on the information provided during signing into the service. Cosine similarity and Jaccard similarity were used to compute the probability that a student would get placed in each of the companies

EXTRA-CURRICULAR ACTIVITIES

- CDS Leadership Circle, Marketing and Student relation representative, New York | Art of Living, Volunteer, Bangalore, India
- BIT Cultural Club, Bangalore, India | Technoholix, Bangalore, India | Freelance Photographer