Lakshmimanaswitha Chimakurthi

2B, Smith Street , Boston, MA 02120
chimakurthi.l@husky.neu.edu | (617)-513-2593
Available for Full-time positions starting May 2019

Portfolio: manaswitha1001.github.io
Linkedin: manaswithachimakurthi
Github: manaswitha1001

PROFESSIONAL EXPERIENCE

Data Science Co-op
***Brigham and Women's Hospital**, Boston, MA* *May 2018 - Dec 2018*

- Collaborated with the physicians and bioinformaticians and built a pipeline to cluster the Lung-Tissue expression and methylation profiles into COPD cases/controls using Matlab and identify the clinical associations for each cluster and visualized the results using ggplot in R.
- Built a docker image for cheweb (A tool for visualizing Channing's GWAS results).
- Implemented an autoencoder neural network classifier to classify COPD case/controls on dosage values and improved the AUROC to 0.78 using the stacked approach.
- Extracted 1M Genotype data records from multiple Oracle relational databases into a simplified json structure using SQL.

EDUCATION

***Northeastern University,** Boston, MA* *Jan 2017- Present*
Candidate for a Master of Science in **Data Science** *Expected Graduation - May 2019*
Relevant Courses: Machine Learning, Algorithms, Natural Language Processing, Data Management & Processing
Information Retrieval, Database Management Systems, Information Visualization

***VR Siddhartha Engineering College,** Vijayawada, India* *June 2012 - Apr 2016*
Bachelor of Technology in **Information Technology**
Relevant Courses: Database Management Systems, Data Warehousing, Data Mining, Business Intelligence

TECHNICAL SKILLS

Key Strengths:	**Predictive Modelling, Data Mining, Text Analytics, Web-Scraping, Recommendation System Statistical Modelling, Classification, Data Visualization, Clustering , Deep Learning**
Programming Languages:	Python, R, SQL, Scala, C++, Java, Matlab
Databases:	Oracle, MySQL, MongoDB
Machine Learning:	Linear/Logistic Regression, SVM, Tree Based, Neural Networks, Clustering
ML Tools:	Scikit Learn, Pandas, Numpy, PySpark, Tensorflow, Keras, ARIMA
Data Visualization:	Tableau, Excel, ggplot, R Shiny, Plotly, Matplotlib, Seaborn, d3.js
Big data Technologies:	Hadoop, Spark
Cloud Technologies:	AWS, Elasticsearch

PROJECTS

Price Prediction of Used Cars *Mar - May 2018*

- Scraped the car listings on attivo.com using Beautiful Soup in python
- Implemented Linear Regression, Decision Trees, KNN, Boosting to predict the prices of car using the car's attributes.
- Achieved the best RMSE of 0.76 with Gradient - Boost Regressor on test data.
- Deployed the prediction model as a Flask API and hosted the application in Heroku.

Sentiment Analysis on Customer Tweets *Jan - Mar 2018*

- Processed the Customer tweets on top 6 US Airline Carriers and encoded the text data into word vectors.
- Implemented a multilayer neural network classifier on processed data using Keras in Python.
- Classified the customer tweets into positive , negative , neutral.
- Achieved an AUROC of 0.68

Movie Recommender System *Aug - Dec 2017*

- Developed a movie recommender system using collaborative filtering approach on IMDB movie ratings.
- Suggests movies based on similar users past ratings for other movies.
- Implemented using K-Means, KNN, SVM , neural network and achieved the best Precision of 0.85 with SVD.

Understand Local Business Dynamics, Neighborhood characteristics with Yelp Data *Oct - Dec 2017*

- Extracted the neighborhood level features of business dynamics from Yelp and Census data.
- Employed K-means, GMM at Zillow Neighborhood and Census Tract level to identify clusters based on population characteristics and socio-economic metrics.
- Investigated the relationship between local business dynamics and neighborhood characteristics.

Prospect of Data Related Jobs in US States *Jan - Apr 2017*

- Scraped the Glassdoor salaries for data related jobs across US states and stored the scraped data in MongoDb.
- Analyzed the data and created interactive visualizations of the top paying states, top hiring states using Plotly in R.

Activities - Winning team member for INFORMS Data Visualization Hackathon - Presented a poster on Boston Crime Data Analysis.