Lakshmimanaswitha Chimakurthi

2B,Smith Street | Boston, MA 02120 | chimakurthi.l@husky.neu.edu | 617-513-2593
Linkedin: manaswithachimakurthi Portfolio: manaswitha1001.github.io Github:manaswitha1001

Available for Full-time positions starting May 2019

EDUCATION

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

Programming Languages:	Python, R, Scala, SQL, C++, Java, MATLAB, Awk
Databases:	Oracle, MySQL, MongoDB, Neo4J
Machine Learning:	Linear/Logistic Regression, SVM, Tree Based, Neural Networks, Clustering, PCA
ML Tools:	Scikit Learn, Pandas, Numpy, PySpark, Tensorflow, Keras, ARIMA, dplyr, BeautifulSoup
Data Visualization:	Tableau, Excel, ggplot, R Shiny, Plotly, Matplotlib, d3.js
Big data Technologies:	Hadoop, Spark
Cloud Technologies:	AWS, Elasticsearch
Containers:	Docker

PROFESSIONAL EXPERIENCE

Data Science Co-op *May 2018 - Dec 2018*
Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA

- Performed an entropy based clustering on various Lung-Tissue expression and methylation datasets and identified the cases/controls of COPD and clinical associations for each cluster and visualized the results with ggplot in R.
- Developed a docker image for cheweb (A tool for visualizing Channing's GWAS results).
- Developed a neural network classifier to identify COPD case/controls using stacked autoencoders on dosage values.
- Efficiently written sql queries to extract genotype data from multiple Oracle relational databases into a json structure.

INDEPENDENT PROJECTS

- **Efficient Mining of Quantitative Weighted Association Rules(Java)**
  Implemented an extension to weighted apriori algorithm in Java, which identifies the weighted association rules between the products purchased by the customers and also finds out the most frequently purchased items in a store.
- **Price Prediction Of Used Cars (Python, Flask)**
  Developed a model using Gradient Boosting which predicts the price of the used cars using the features and achieved an RMSE of .76 and deployed a working app in Heroku using Flask API.
- **Sentiment Analysis on Customer Tweets (Python, Keras)**
  Developed a sentiment analyzer using neural network which predicts the sentiment of the customer's tweets posted on Twitter.

ACADEMIC PROJECTS - *Northeastern University, Boston, MA*

- **Movie Recommender System (Python)**
  Developed a movie recommender system using collaborative filtering approach that suggests movies based on users past ratings for other movies using several ml algorithms and compared the results.
- **Understand Local Business Dynamics and Neighborhood characteristics with Yelp Data (Python)**
  Clustered the Yelp businesses data and Census data to identify how the local business dynamic patterns associate with population characteristics of the neighborhood and visualized the results using matplotlib.
- **Prospect of Data Related Jobs in US (R, MongoDB, Plotly)**
  Scraped the Glassdoor salaries for data related jobs across US states and stored the data in MongoDb.
  Queried the database and Identified the top paying states, top hiring states and visualized the results using Plotly.
- **Search Engine (Python)**
  Developed an Inverted Index and Search Engine over 80k document collection and provided a ranked list of documents using Okapi, BM25 retrieval models for a given set of user queries.

ACTIVITIES

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