**AKHIL NAIR**

Syracuse, NY 13210

[akhilNair.github.io](https://akhilNair.github.io/)

[anair05@syr.edu](mailto:anair05@syr.edu) [linkedin.com/in/akhilNair2](https://www.linkedin.com/in/husenbhagat)

+19204894259

# EDUCATION

**Master of Science, Applied Data Science** *Aug 2019 – Aug 2021*

Syracuse University, Syracuse, NY

# Bachelor of Engineering, Electronics & Telecommunication *May 2016*

University of Pune, Pune, India *First Class with Distinction*

*Coursework* : Data structures and Algorithms, Machine Learning, Information Theory, Computer Networks, Databases, Operating Systems

# WORK EXPERIENCE

**Software Engineer Tata Consultancy Services (TCS)** *January 2017 - June 2019*

* Researched, designed and implemented Natural Language Processing pipelines equipped with statistical models.
* Worked on extracting data from social media sites like Twitter using Python for Social Network Analysis (SNA) of drug complaints for adverse drug reactions (ADR)
* Applied machine learning algorithms and word embeddings on the real life patient data to perform sentiment analysis to understand the effectiveness of drugs
* Built web application on Django with ReactJs on frontend and MongoDB on the backend for sentiment analysis.
* Performed exploratory data analysis on medical reports from MAUDE and PUBMED repositories of FDA for building statistical models for classification of adverse events and drugs
* Worked on data cleaning, feature engineering, feature selection, regularization evaluation metrics of models and optimizing hyper-parameters.
* Worked on Time Series Analysis for real life data of hospitals to extract patterns and improve the effectiveness of hospitals conducting clinical trials

**Data Science Intern Quadlogix Technologies** *June 2016 - December 2016*

* Worked as a Data Science Intern to build an android application for capturing user sentiments based on reviews posted for Indian TV series.
* Wrote maintainable and scalable code in a team environment
* Researched, implemented and benchmarked machine learning algorithms like Random Forest, Support Vector Machines and Logistic Regression for prediction and classification task.
* Coordinated with the Client in order to gain better insights into their requirements

# TECHNICAL PROJECTS

**Social Media Analysis of Drug complaints on Twitter**

* Designed an end to end data science project pipeline by incorporating medical ontology to extract data from Twitter using twitter’s streaming application programming interface(API) and classify the reported drugs complaints for its authenticity and seriousness.
* Utilized Unsupervised machine learning techniques such as outlier detection and K-means clustering using Python for anomaly detection and classification tasks.
* Built Neural Network on the extracted data to classify Adverse Drug Reactions using Keras and integrated it with Tensorflow JS to visualize and fine tune the model performance metrics.
* Developed a tweet categorization module using word vectors to measure the tf-idf and cosine similarity score to categorize the tweet based on the intent.
* Worked on Time series analysis using ARIMA based approach on drug complaints to find hidden patterns about drug behavior and it’s demographics

**Data Mining**

* Developed a Generic Python code to extract information from email, Pdf’s of medical reports and litigation reports to handle all the structural variations.
* Utilize image processing algorithms to implement OCR extraction from medical reports in the form of images using Python.
* Applied machine learning algorithms like Support Vector Machines and Random Forest, to build statistical models and classify adverse events suffered by patients due to consumption of drugs using Scikit-Learn and Keras packages in Python

# SKILLS

**Programming:** Python, Java, JavaScript, SQL, NOSQL, XML

**Framework & Tools:** MySQL, Git, Jupyter Notebook, Django, Flask, ReactJs, MongoDB, NodeJs

**Familiar with:** Linux (Ubuntu, Fedora), Bash, HTML, RESTful OData v2