

## Education

### Syracuse University: Syracuse, New York, USA

Master of Science, Applied Data Science (3.9/4.0)

August '18-May '20

### University of Pune: Pune, Maharashtra, India

Bachelor of Engineering, Computer Engineering

August '13-May '17

## Skills

**Programming Languages:** Python, Scala, R, C++, Java

**Database:** T-SQL, MSSQL, MySQL, PostgreSQL, Hive, Impala, Hadoop, MongoDB, Redis, Cassandra, Apache Kafka

**Data Engineering:** Query Optimization, Indexing, Data Warehousing, Extract Transform Load (ETL), Data Pipeline

**Tools:** Microsoft Excel, Microsoft Access, Pivot table, Power Pivot, vlookup, Power Query, Docker, Tableau, Power BI, Microsoft SQL Server Integration Services, Microsoft SQL Server, Business Intelligence

**Machine Learning:** Linear Regression, Non-Linear Regression, Logistic Regression, Random Forest, Naïve Bayes, K-Mean Clustering, Principal Component Analysis, Multilayer Perceptron, Data Manipulation, Business Analytics

**Statistical Analysis:** Hypotheses Testing, Confidence Interval, Logit, Probit, Time-Series Forecasting

## Experience

### Data Analyst Intern

August '19-Present

*Health promotion, Syracuse University: Syracuse, New York*

- Developing a standalone database software for in-office usage using python and SQL to store & efficiently query data
- Coded a data pipeline to collect, clean, munge, transform & combine data from 5 sources to evaluate campus compliance
- Researching focus group to restructure & analyze data to understand student response & impact of wellness activities
- Analyzing outreach, impact & reception of various wellness activities & campaigns conducted by the office

### Team Lead/ Developer (Project: JobIQ)

May '19-August '19

*iConsult Collaborative, School of Information Studies: Syracuse, New York*

- Lead & delegated a team of 5 in development of a software assisting students with job search and interviews
- Designed the data logic & middle tier using python & MongoDB
- Implemented the interface between middle tier & backend
- Developed a text parser that extracted keywords, archived the text in the database & keyed it using the e-mail address.

## Projects

### Merger Data Warehouse

- Designing data warehouse to contain data from 2 different businesses & focus on market research for the merger
- Performing merging, staging & ETL procedure on the data from different databases to the integrated data warehouse
- Generating insightful business intelligence to be used in future business activities

### Handwriting Recognition (<https://git.io/Jv3Vy>)

- Transformed and munged features for 690,000 images after rescaling grayscale levels that reduced computational requirements for machine learning
- Implemented principal component analysis to achieve dimensionality reduction while retaining the variance explained across the data
- Trained and tuned random forest classifier and multi-layer perceptron to identify characters
- Achieved the highest accuracy of 63% with neural networks, despite computational limitations

### Austin Weather Prediction (<https://git.io/fjsbN>)

- Cleaned, refactored & feature engineered weather data from Austin & predicted weather
- Engineered Gaussian Naïve Bayes & multiclass logistic regression models in python with highest accuracy of 81.72%

## Publication

**'Survey and Analysis of ECG Interpretation Techniques'**, International Journal of Pure and Applied Mathematics (IJPAM) ISSN: 1314-3395 (online version). (<https://acadpubl.eu/jsi/2017-114-7-ICPCIT-2017/articles/7/43.pdf>)

## Soft Skills

Communication skills, Decision making, Can-do Attitude, Teamwork, Detail oriented, Reliability, Driven, Consistency