

Sai Raghuram Kothapalli

Boston, MA – 02119 | 857-544-5554 | k.sraghuram20@gmail.com | [LINKEDIN](#) | [GITHUB](#)

Education

Northeastern University, Boston, MA

Aug '16 – Dec '18

Master of Science, Industrial Engineering (Specialization – Data Science/Analytics)

Relevant Coursework - Machine learning, Data Mining, Deep Learning, Probability and Statistics, Algorithms

CGPIT, India

Jun '12 – May '16

Bachelor of Technology, Mechanical Engineering

Expertise

- **Tools:** Python (Numpy, scikit-learn, genism, plotly, tensorflow, keras, Seaborn), R, MySQL, Hadoop, Hive, docker, Linux
- **Analytical:** ML (sklearn, caret, mlr, mllib), language modelling, NER, NLP, Statistics (t-tests, confidence intervals)

Experience

Fidelity Investments – Data Science Intern

Jan '18 – Jun '18

Sentiment Analysis model

- Developed 2 deployable sentiment analysis models (LSTM and TCN) aimed at deriving customer sentiment
- Results: LSTM achieved 82.5% and TCN achieved 81.8% of testing accuracy, with False Positive Rate of 11%.

Customer Insights reporting

- Scraped 10K webpages and extracted keywords using RAKE for content summarization and topic extraction
- Results: with keywords and sentiments, automated reports in interactive Tableau dashboards to assess the context of chats and search bar from webpages, e.g. Bitcoin (December surge), Cambridge Analytica (Facebook scam in April)
- Tools: Python (plotly, RAKE), Tableau | Techniques: Deep learning – LSTM, GRU, CNN & TCN, web scraping, NLP

Bharat Heavy Electrical Ltd. – Business Analysis Intern

May '15 – Jul '15

- Designed a database in MySQL and created business intelligence reporting system in Tableau using their sample data
- Results: Increased the sales by 4% due to reduction in fulfillment costs and understanding of requirements of products
- Tools: MySQL, Tableau, R | Techniques: Reporting, database management – stored procedures, triggers, views

Reliance Industries Ltd. – Data Analysis Intern

Jun '14 – Jul '14

- Executed SQL scripts to extract crude oil prices data from databases and transformed the unstructured data
- Results: Visualized the data in Tableau, used time series and forecasting to predict future crude oil prices
- Tools: MS Excel, Oracle database, R, Tableau | Techniques: Data extraction and cleaning, time series forecasting

Northeastern University – Graduate Teaching Assistant

Sep '18 – Dec '18

- Facilitated cognitive and effective learning for students in Data Science and Big data courses (Python, H2O & R)

Academic Projects

Northeastern University

Software sales forecasting, Kaggle

Nov '18

- Analyzed the trend and seasonality in 2-year data and built ARIMA, SARIMA as well as Facebook Prophet forecasting models to predict sales. Implemented grid search for hyperparameter tuning of order models.
- Tools: R (astsa) | Techniques: Autocorrelation function, Dickey-fuller test, Ljung-Box test | Results: AIC – 474.40

Salary Prediction of Employees

Nov '17

- Analyzed a large (1.9 million observations) imbalanced classification dataset and converted it into balanced using SMOTE sampling and conducted exploratory data analysis using plots with feature engineering
- Tools: R (mlr) | Techniques: SVM, XGBoost, feature engineering, semi-structured data | Results: 70% testing accuracy

Anomaly detection of credit card transactions

Sep '17

- Built an L1-regularized Autoencoder in H2O platform aimed at detecting anomalies in the 2-day transactions and used the reconstruction-error to evaluate the fraudulent transactions.
- Tools: Python (H2O, keras) | Techniques: deep learning, regularization, risk analysis | Results: 85% frauds detected

Image Recognition

Dec '16

- Trained many CNN models on the MNIST dataset for getting an optimum model capable of recognizing digits using feature engineering and data augmentation techniques like rotation, zooming and shifting.
- Performed error/deviation analysis like “3” being predicted as “8” and “4” as “9”. Most misclassified was number “4”
- Tools: Python (keras) | Techniques: DNN, CNN, data augmentation, feature engineering | Results: 98.7% test accuracy