

Sai Raghuram Kothapalli

[LinkedIn](#) | [GitHub](#) | [Gmail](#) | 8575445554

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

CGPIT, India

Jun '12 – May '16

Bachelor of Technology, Mechanical Engineering

Expertise

- **Tools:** Python (numpy, sklearn, plotly), R, MySQL, Oracle, SQLServer, Toad modeler, Tableau, PowerBI, QlikSense
- **Analytical:** Machine learning, Text mining, A/B testing, NLP, Statistics (p-value, z-test, t-test, ANOVA), Google analytics

Experience

Aiera Software Inc. – Data Scientist

Apr '19 – Jul '19

- Designed a pipeline to extract data and performed time series analysis for efficient stock prediction, using historicals
- Solved bug tracking issue by designing an efficient pipeline for logging errors and reporting to Bugsnag
- Tools: PyCharm (Python), R | Techniques: Logging, Dickey-fuller test, Ljung-Box test | Results: 0.56 validation RMSE

Fidelity Investments – Data Science Intern

Jan '18 – Jun '18

Sentiment Analysis model

- Developed 2 sentiment analysis models aimed at quantifying customer sentiment for customer and rep. conversations
- 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
- Analyzed the competitors and the locations mentioned by customers for Named Entity Recognition using NERTagger
- Results: Automated keywords and sentiment score reports in interactive Tableau dashboards to assess 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 – Aug '15

- Designed a database and automated business intelligence reporting system in Tableau using R scripts for daily scraping
- Performed market basket analysis, to recommend 'cross selling' and 'product placement' to the clients
- Results: Reduced fulfillment costs based on client buying behavior analysis, thus boosting sales by 4%
- Tools: MySQL, Tableau, R | Techniques: Reporting, database management, recommendation, market basket analysis

Reliance Industries Ltd. – Data Analysis Intern

May '14 – Aug '14

- Transformed and structured shipment production data by executing SQL scripts on internal databases
- Built and published reports on operation efficiency, cost reduction and capacity planning for the stakeholders
- Results: Analyzed and published dashboards in Tableau supporting recommended key metrics for operational logistics
- Tools: MS Excel, Oracle database, R, Tableau | Techniques: Data extraction and cleaning, time series forecasting

Academic Projects

A/B testing

Mar '19

- Solved business problem of student enrolment by conducting A/B test in Python on an education website (Udacity)
- Stated hypothesis using click-through-probability as the metric and quantified power of test to detect the effect
- Using p-value of the Z-test, concluded that the experiment group is statistically significant and ready to be launched

Market Basket Analysis

Feb '19

- Analyzed market baskets of a retail store to quantify customer buying behavior and provide better recommendations
- Created associations between products using Apriori algorithm, ex. how is buying of milk and coffee beans associated
- Created a recommendation engine (1.7 RMSE) for personalized suggestions using user's profile and buying patterns

Anomaly detection of credit card transactions

Oct '18

- Built an L1-regularized Autoencoder in H2O platform aimed at detecting anomalies in the 2-day credit card transactions and used the reconstruction-error as the metric to evaluate the fraudulent transactions.

Software sales forecasting, Kaggle

Nov '18

- Analyzed the trend and seasonality in 2-year data and built SARIMA models to predict sales in R with AIC of 474.4
- Performed Dickey-fuller test to test null hypothesis for presence of unit root and Ljung-Box test for lack of fit

Data Visualization Competition

Sep '18

- Won 1st prize by creating a metric to gauge the efficiency of response of EMS and reporting on Tableau dashboard