

PRANAV SURESH MAGADI

(312) 721-2956 · psures6@uic.edu · <https://linkedin.com/in/pranav-suresh-magadi> · <https://psureshmagadi17.github.io>

- Analytical data scientist with 3+ years of relevant academic and practical experience within the emerging machine learning space
- Proficient in Python, SQL, R, Tableau, Statistics, Machine Learning, Data Mining, Project Management, Reporting & Consulting
- Proactive learner and a strong team player adept at delivering innovative, reliable, cost-saving solutions that drive key business decisions
- Motivated, team oriented, reflecting strong leadership abilities, problem-solving, time management, presentation & communication skills

EDUCATION

Master of Science (M.S.) - Management Information Systems, University of Illinois at Chicago (GPA: 3.89) 08/2018 - 12/2019

Bachelor of Engineering (B.E.) - Computer Science, BNM Institute of Technology, India (GPA: 3.3) 08/2012 – 06/2016

CORE COMPETENCIES

- Data Visualization
- Research
- Data Science
- Statistical Analysis
- Healthcare Analytics
- Data Analysis
- Predictive Analytics
- Business Intelligence

TECHNICAL PROFICIENCIES

Programming/Scripting: Python, R, MS-SQL, JavaScript, Java, HTML, Unix
Productivity Tools: MS Office, MS Excel, Oracle Middleware, Git
Statistical Techniques: Hypotheses testing, Anova, Chi-sq, Correlation

Analytical/Data Viz. Tools: Tableau, PowerBI, SPSS
Database/Big Data: Spark, SSMS, SQL Server
Packages: numpy, pandas, ggplot2, sklearn, seaborn

EXPERIENCE HIGHLIGHTS

Data Scientist Intern - Blue Rock Healthcare It, Inc., Chicago, IL | R, Python, Tableau, Claims data, Client Facing 6/2019 - Present

- Implemented data pipeline for de-identification, ETL, report generation and analysis solutions for 5+ clients
- Published dashboards showcasing 4 cost-saving opportunities costing clients \$100K+ while also providing analysis on individual physician expenditures for 17 different segments such as drugs, ER visits, hospitalization, etc. using healthcare data
- Simulated target prices & cost of treatments using monte-carlo technique improving data-driven risk-sharing decision making for practices

Data Science Intern - UChicago-Center for Translational Data Science | R, Python, Image Processing, Agile 7/2019 - Present

- Built random forest model able to predict fire with 80% precision based on 16 variables including vegetation, water-vapor, etc.
- Utilized geospatial data such as NASA's GOES-R satellite NETCDF files and US Forest Department fire event shapefiles
- Mitigated class imbalance by resampling, generating synthetic samples, & identifying classification thresholds using ROC curves

Data Analyst Intern - GSG Consultants, Chicago, IL | R, Excel, Text Analytics, REST API 9/2018 - 12/2018

- Performed sentiment analysis of 20,000+ competitor's tweets to identify words/phrases used to generate higher engagement
- Researched competitive landscape on social media platforms evaluating post frequency, time of posts, likes, retweets, etc.
- Delivered strategic recommendations to client on how to enhance online presence increasing customer engagement by 30%

Middleware Application Developer - Accenture Solutions, India | Oracle Middleware, SOAP, SQL, Data Analysis 11/2016 - 6/2018

- Developed SOA web services responsible for receiving 800+ orders, invoices, receipts, and bank details
- Produced technical design document and error reports analysis by data extraction using SQL helping speed up bug fixes by 20%

ACADEMIC PROJECTS

08/2018 - 12/2019

Multi-Label Text Classification Model | Unstructured Data, R, Tableau, Python, Google AutoML, Natural Language Processing (NLP)

- Devised text classification model with 4 labels capable of identifying consumers, sellers, education/news associated with opioids
- Trained classifier using transfer learning and extracted data driven insights including location and tweet frequency of users and sellers
- Analyzed 160K+ tweets resulting in ability to identify users, sellers, news, education/awareness tweets with 91% precision

Text Mining and Sentiment Analysis | Excel, R, NLP

- Analyzed sentiment dictionaries and evaluated classification models helping predict sentiment polarity of 50K Yelp reviews
- Achieved 84% accuracy with support vector machine after tokenization, stop-words removal, lemmatization, & sentiment analysis

Regression Modelling | Excel, Python

- Produced screening tool identifying probability of developing chronic kidney disease with an 86% precision rate by cleaning data, performing exploratory data analysis (EDA), normalization, data manipulation & applying logistic regression on 5K patient records
- Model can be implemented in hospitals to validate prescription of CKD tests on patients and reduce healthcare costs

Data Visualization – Interactive Dashboards/Stories | Tableau, Business Statistics, Excel

- Analyzed Bureau of Transportation Statistics airline data set to identify metrics that report On-time Performance of United Airlines
- Implemented stories, dashboards, LOD expression to extract actionable insights and perform competitive analysis based on key metrics
- Recommended routes that will generate high profit and possible growth opportunities for United airlines based on network maps

Data Mining / Classification & Predictive Modeling | Excel, R, Structured Data

- Improved cost effectiveness of direct marketing campaign for national veteran's organization with 60K donors
- Built ML model using random forest & SVM that predicts donor/non-donor & expected donation amount with 90% accuracy
- Introduced feature selection and dimensionality reduction techniques on large datasets of 60K records and 150 attributes