PRANEETH BOMMA

praneeth.bomma401@gmail.com | +1 (980)-318-6668
https://www.linkedin.com/in/pbomma | https://github.com/pbomma | http://praneethbomma.com/

OBJECTIVE

Seeking a full-time position in the field of Data Science and Analytics with 2 years of Professional, Research experience in Data Science/Analytics
Proficient in Python, R, SQL, Tableau with the ability to translate given data providing meaningful insights solving business problems

EDUCATION

Master of Science in Information Technology, GPA 4.0/4.0 **Jan 2017–May 2018**
The University of North Carolina at Charlotte, North Carolina
Coursework includes: Business Intelligence and Analytics, Machine Learning, Knowledge Discovery in Databases, Deep Learning, Text Analytics, Data Visualization, Cloud Data Storage, Modern Data Science Systems, IT Project Management

Bachelor of Technology in Electronics and Communications, GPA 3.5/4.0 **Jul 2012 – Aug 2016**
Jawaharlal Nehru Technological University – Hyderabad, India

CERTIFICATIONS: Google Academy: Google Analytics Individual Qualification, Advanced Google Analytics, Coursera - SQL for Data Science, R Programming

TECHNICAL SKILLS

Programming:	SQL, R, Python, Core Java, SAS, Hive, Pig, Scala, C
BI/Analytics Tools:	MS Excel, Tableau, Google Analytics, WEKA
Web Languages:	HTML5, CSS3, jQuery, JavaScript
Databases:	MySQL, PL/SQL, Oracle SQL, MS SQL, Postgresql, NoSQL, Hbase
Cloud Services:	Amazon Web Services, Microsoft Azure, Google Cloud, Cloudera
Machine Learning:	Regression, Classification Models, Deep Learning, ANOVA, PCA, Bayesian Statistics, NLP, Time-series
Libraries(R & Python):	Pandas, Scikit-Learn, Numpy, Plotly, NLTK, TensorFlow, Matplotlib, Pytorch, caret, forecast, glmnet, ggplot2, dply*r*

PROFESSIONAL EXPERIENCE

DATA ANALYST INTERN, *Informative Technologies Inc,* *Charlotte, NC, USA* **Jan 2018–June 2018**

ReviveNC:

- Performed exploratory data analysis, data pre-processing and cleaning using Python.
- Applied K-means clustering algorithm and did segmentation to group the unsupervised data into related clusters
- Built a classification model that classifies the new cases into a cluster using KNN so that E-waste recycling can be achieved effectively

Lookit:

- Collaborated with client and development team to set scope, specifications and identify new requirements for the application
- Created entity relationship (ER) diagrams, developed a relational database and implemented database objects
- Designed a ranking and sorting algorithm to order images based on parameters of relevance including location, timestamp and rating

DATA ANALYST INTERN, *Aakruti Digipress Pvt. Ltd,* *Hyderabad, INDIA* **May 2016–Dec 2016**

- Built a time series ARIMA models using the company's historical data.
- Generate customer-oriented report and visualization in Tableau for clients to develop marketing campaign strategy
- Predict revenue, order quantity and future demand in conjunction with technical charts, trends, demands at peak volume for company

ACADEMIC PROJECTS

Lending Club - Loan Status Prediction| Techniques: feature selection, feature extraction, classification**| Tools**: Python

- Performed feature selection, extraction, built classification and ensemble methods to predict borrowers who tend to default
- Applied k-fold cross validation to select best parameters of model and obtained 91% prediction accuracy using ensemble methods

Finding Surprising Documents on Online Health Information| Techniques: Tokenizing, Clustering, Word Cloud, SK-means**| Tools**: *R*

- Developed a computational approach using R programming to identify "surprising" news from a news corpus related to diabetes
- Applied unsupervised machine learning techniques to get the surprising elements from a text corpus of 10000 documents

Spatial and Time-Series Analysis of SFO Crimes| Techniques: EDA, Time Series Analysis, Spatial Analysis**| Tools**: *Python*

- Performed spatial distribution over time and time series analysis for a 15 year dataset of reported incidents from SFPD
- Trained and fine-tuned an ARIMA model to forecast the number of theft incidents per month

Find me a job that is a perfect fit for me | Techniques: Web Scraping, BeautifulSoup, SMTP **| Tools**: *Python*

- Built a custom tool by scraping Indeed.com for new job postings in multiple cities with a predefined recency and search criteria
- Mails consolidated list of relevant job openings to my email ID. Task can be repeated every few days

Hire Heroes USA- Client Management| Techniques: Regression, Decision Trees, Text mining **| Tools**: *SAS, R, Excel and Tableau*

- Developed a Multiple Regression Model which explains 27% of the variation in how quickly a veteran gets hired
- Used Text mining to categorize data into text topics and decision trees to explain correlation between client & employers of HHUSA

Predict Housing Prices – Kaggle competition| Techniques: Feature Scaling, k-fold, Gradient Boost, Grid search| **Tools**: *Python, Tableau*

- Performed exploratory data analysis, feature scaling, k-fold cross validation and grid search to achieve most approximate prediction
- Achieved an accuracy of 85 percent in predicting housing prices of King county housing data using Gradient Boosting

Twitter Text Analysis - Movies Success| Techniques: NLTK, sentimental analysis, tag cloud| **Tools**: *Python, NLTK, matplotlib, TweepyAPI*

- Tweets crawled using the Tweepy API in Python were pre-processed to create a corpus for analysis using NLTK module
- Applied sentiment analysis & created a tag cloud of top 50 words in the tweets to understand the audience sentiments about movie