KRISHNA GOLLAPUDI

http://www.krishnagollapudi.com/ | https://www.linkedin.com/in/gkrishna19/ | https://github.com/KrishnaGolllapudi | vgollap1@uncc.edu | +(1)980-598-0789

Summary:

- Data Science Graduate with 2 years of professional and research experience in Data Science and Analytics
- Proficient in Python, R, Tableau, SQL with the ability to translate data into meaningful insights solving business problems
- Expertise in Data Mining, Statistical Modeling, Information Visualization, Machine Learning, Text Analytics, Natural Language Processing

EDUCATION:

Master of Science: **Information Technology, GPA: 3.9** **Jan 2017 - May 2018**
University of North Carolina at Charlotte – Charlotte, North Carolina

Bachelor of Engineering: **Electronics and Communications, GPA: 3.5** **Aug 2011 - May 2015**
Osmania University – Hyderabad, India

WORK EXPERIENCE:

PROLYTICS LLC, Charlotte, North Carolina, USA. **– Data Scientist Intern** **Jan 2018 – Apr 2018**

- Performed advanced statistical analysis and predictive modeling on MLB, NBA Draft data (3 years of NBA data) and NCAA college stats
- Adapted Machine Learning algorithms in predicting player's match up analysis based on their game position, historical NBA data
- Data wrangling of MLB data and predictive analysis on the performance of each player in future matches
- Developed an LSTM RNN to project player's expected performance in the draft over 2-3 years

ACCENTURE, Hyderabad, India. **– Associate Software Engineer** **Mar 2016 - Dec 2016**

- Contributed as a Database Developer with Global Resource Management project for client: Microsoft with Agile methodology
- Modified the web test scripts according to the API changes and created pipelines in Azure Data Factory(ADF), SQL Jobs
- Executed constant/load tests in azure which included performance monitoring, performance test analysis, performance tuning

HCL CDC, Hyderabad, India**. – Software Engineer Intern** **Dec 2015 – Feb 2016**

- Worked on an Internal project of HCL for Customer Query Tracking System
- Responsibilities include system design, creating E-R diagram, tables and stored procedures

Infosys Limited, Mysore, India. **– Systems Engineer Trainee** **June 2015 – Nov 2015**

- Trained on PYTHON, JAVA, HTML, CSS3, JavaScript, Oracle SQL
- Designed and developed an SQL Database system for an internal Business Enterprise Application

TECHNICAL SKILLS:

Programming Languages/Tools:	Python, R, SQL, SAS, Java, Pig, Hive, C++, Tableau, MS Excel, Google Analytics
Statistics/Machine Learning:	Regression, Classification, Rule Mining, ANOVA, NLP, Text Analysis, Bayesian statistics, TF-IDF, Information Retrieval, Forecasting, Survival Analysis, Time-Series Analysis, Deep Learning
Libraries (Python and R):	pandas, scikit-learn, scipy, numpy, matplotlib, cluster, xgboost, tidytext, Tensor flow, Convolution2D, H2O
Database:	MySQL, MongoDB, Oracle SQL, SQL/PL, Teradata, MS SQL Server
Cloud Services:	Amazon Web Services, Microsoft Azure, Google Cloud, Digital Ocean

PROJECTS:

Predict Housing Prices – Kaggle competition (Supervised Machine Learning**) | Tools Used:** *Python, Tableau*

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

Surprising Discoveries for Online Health Information (Unsupervised Machine Learning, NLP**) | Tools Used:** *R*

- Developed a computational approach using R programming to identify "surprising" pattern from **diabetes** related news corpus
- Applied the unsupervised machine learning techniques to achieve the surprising discovery from given text corpus of 10000 documents
  Techniques Used: Clustering, Cosine Similarity, PAM & Word Cloud, SK-means

Hire Heroes USA Client Management (Data Analysis and Visualization**) | Tools Used:** *SAS, R, Excel and Tableau*

- EDA for Data Insights and Logistic Regression, Decision Tree were performed at each stage to improve the process of hiring
- Text mining was used to generate features and predictive modelling techniques were used to model the quantities of interest
  Techniques Used: Exploratory Data Analysis, Predictive Analysis, Regression Analysis, Text Mining, Decision Trees, Tableau Visualization

Movie Recommendation Search Engine (Recommender System**) | Tools Used:** *R, Shiny*

- Prepared a collaborative filtering recommender (CFR) system for recommending movies to users based on genre
- The Similarity Calculation Method was based on Cosine Similarity and the Nearest Neighbors was set to 30
  Techniques Used: Association Rule Mining, Cosine Similarity, KNN

Lending Club - Loan Status Prediction (Supervised Machine Learning**) |Tools Used:** *Python, Jupyter Notebook, Scikit-learn, Pandas, Numpy*

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

Spatial and Time-Series Analysis of SFO Crimes (Time Series Analysis**) | Tools Used:** *Python, Numpy, Seaborn, Matplotlib, Pandas, ARIMA*

- Performed spatial distribution over time and time series analysis for 15-year dataset of reported incidents from SFPD
- Trained and fine-tuned an ARIMA model to forecast the number of theft incidents per month
  Techniques Used: Exploratory Data Analysis, Time-Series Analysis, Spatial Analysis, Matplotlib Visualization