Venkata Prudhvi Raj Indana

Bloomington 47408 (INDIANA) |812-650-8032 |indana.prudhvi.2050@gmail.com |https://github.com/prudhv16

EDUCATION

Indiana University Bloomington — May 2017
Master of Science in Data Science — **GPA: 3.56/4.00**

Gitam University, Visakhapatnam — May 2012
Bachelor of Engineering in Electronics and Communication — **GPA: 3.6/4.00**

TECHNICAL SKILLS

Programming Languages: Python, R, C#, SQL, Javascript, Java, Mainframe, CICS
Databases: - MYSQL, PostgreSQL, Oracle SQL, SQL Server, Microsoft Access, IBM DB2
Tools: - Tableau, R-Studio , Jupyter, PIG, Google Analytics, JIRA, Bugzilla, Hive, WEKA, Git, Github, Markdown, Excel, Redshift, Confluence, Google Cloud, Remedy.
Statistical Methods and Mathematics: - Hypothesis testing, Probability, Calculus, Linear Algebra
Data Science Libraries:- scikit-learn, pandas, numpy, matplotlib, seaborn, ggplot2, RMySQL, Gmodels, class, wordcloud

ACADEMIC COURSES AND PROJECTS

Courses: Introduction to Statistics, Algorithm Analysis and Design, Data Mining, Exploratory Data Analysis, Artificial Intelligence, Applied Machine Learning, Information Retrieval, Machine Learning, Multivariate Data Analysis, Machine Learning in Bioinformatics.

Credit Card Fraud Detection - Kaggle *(Python, Tableau)*

- Performed under-sampling, re-sampling to solve imbalanced dataset (skewed dataset) problem along with 10 fold cross validation
- Utilized logistic Regression algorithm to classify transactions as Fraud/ Non-Fraudulent and used ROC/ PR curve as performance metric.

House Prices: Advanced Regression Techniques – Kaggle (Python)

- Created an Ensemble learning approach with k-best features for predicting house prices along with gradient boosting, lasso and ridge.
- Generated a plot of N-dimensions vs RMSE to show effect of PCA on RMSE. Created a plot of parameters vs RMSE to show bias vs Variance Trade off. Ranked Top 10% (out of 2300+ teams) in this challenge.

Spam Filter / Topic Classification *(Python)*

- Implemented of Bayesian spam filter using bag of words and comparing it with decision tree implementation using entropy based greedy Algorithm obtaining an accuracy of 87%.
- Spam filter code is extended to classify genera of a 20 Topic in fully supervised (72%), semi supervised (47%), unsupervised (6%).

Data Analysis on mammography mass data set– UCI *(Python, R)*

- Performed Regression Analysis and determined 6 significant features that affects the prediction of breast biopsy using p-values.
- Developed linear regression model in predicting the malignance with an accuracy of 94%.
- Calculated the performance metrics and used recall score to evaluate the model performance.

Image Orientation Detection (Python)

- Implementation of KNN to train / predict orientation (multiclass problem) of input image using accuracy and Confusion matrix as metric.
- Implementation of Adaboost, Neural network models for the same problem and comparing its result with the results from KNN.

WORK EXPERIENCE

Software Developer Intern — **Bloom Insurance agency, Bloomington, IN** — **Jun 2017-present**

- Responsible for enhancing and testing of Bloom's technology product – Ascend Web Application. Which is an online quoting engine and sales management application for insurance agents.
- Responsible for design and development of web applications for real time insurance-agent management system.

Software Developer — **United Health Group** — **Dec 2014-Aug 2015**

- Requirement gathering by having daily meetings with project sponsors and stakeholders to understand business problem.
- Engaged in bringing the structure to each request and translate requirements into an analytic approach.
- Developed a classification model based on claim data to predict the probability of new claim being rejected.
- Achieved an accuracy of 86% after model parameter tuning and 10 fold cross validation.

Data Analyst Intern — **Arborgold Software, Bloomington, India** — **May 2016 – Aug 2016**

- Analyzing data Creating a new reporting tool (Owners app Dashboard) generating various reports plotted on web interface using MVC C# , Javascript, SQL Server, server and kendo UI.

Software Developer — **Computer Science Corporation, India** — **Jul 2012 – Dec 2014**

- Conduct A/B tests to measure/ improve performance of UI/UX and KPI in a data driven manner.
- Requirement gathering by having daily meetings with project sponsors and stakeholders to understand business problem.
- Implemented a Logistic regression model to access risk index of new customer to diseases using 10 fold cross validation and achieved 73% accuracy. Assisting clients in creating Move it FTP jobs to importing data from client location to local server.
- Build data pipelines that import data from local server to IBM Z/OS, validates input data and rejects / recycles / processes payments for claims in data using stored procedures written in COBOL. Generate daily reports/ Support.