VIB YALAMANCHILI

Austin, TX

Phone: 703-869-9719 | Email: yvibhu@gmail.com

LinkedIn | GitHub | Portfolio

PROFESSIONAL SUMMARY

Multidisciplinary, business-minded data scientist with 13+ years of industry experience across multiple functions like Data Analysis, Data Mining, Data Engineering, and Data Modeling. Demonstrated ability to deliver valuable insights via data analytics and advanced data-driven methods. Experienced at regression analysis, predictive data modeling and analyzing data mining algorithms to deliver insights and solutions to complex business problems. Prepared to leverage quantitative analysis and machine learning knowledge with visualization and attention to detail.

Technical Skills:

Languages: Python, R, SQL, NoSQL, C++, JavaScript, VBA, HTML, CSS, COBOL, and JCL

Databases: MySQL, SQLite, PostgreSQL, DB2, Oracle DB, MongoDB

Data Mining, Manipulation & Visualization: pandas, SciPy, NumPy, matplotlib, seaborn, Plotly, Tableau, D3.js, Leaflet.js | *Big Data*: Hadoop, MapReduce, Hive, Spark (PySpark), Pig

Machine Learning: scikit-learn (sklearn), TensorFlow, Neural Networks, Deep Learning, Keras, Predictive Modeling, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Natural Language Processing (NLP)

Other: Git, Flask, command line (Unix/Linux), BeautifulSoup, Splinter, Web scraping, SVG, ES6 (ECMAScript 6), Bootstrap, RESTful APIs

PROJECTS

Loan Default Prediction – 2018 Source Repo

- Performed data pre-processing, Feature selection and Feature Engineering. Predicted the probability of an approved customer defaulting on their loan payment
- Train/test/determine the classification accuracy along with other model metrics -> Decision Trees, Random Forest, Bagging, Gradient Boosting, Gaussian, Passive Aggressive, Ridge, MLP, Logistic Regression, KNN, SVM

Predicting Recession based on current Economic KPIs, 2018 Source Repo

- Performed data wrangling, mining and normalization on the datasets of the economic KPIs and then used ML Algorithms to analyze and tune the results to perform a Time Series Prediction.
- Linear Discriminant Analysis, Classification, Regression, Decision Trees, Naive Bayes (NB), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Bagging, Random Forest, Boosting, AdaBoost. Implemented LSTM network.

Real Estate Trend Prediction and Demographics Dashboarding 2018 Real Estate, Demographics - Repo

- Performed data mining/ transformation on the Austin housing data to forecast future trend. Leveraged Machine Learning by training the model on various KPIs.
- Designed an interactive dashboard that shows various Austin demographics like Budget Spending, City Crime by zip code, Avg. Income/Median home prices by zip code and real-time City Traffic incidents

Technologies/Skills Used: Python, Jupyter, Anaconda, matplotlib, PySpark, MapReduce, Hive, MySQL, PostgreSQL, MongoDB, pandas, NumPy, APIs, JavaScript, D3.js, Leaflet.js, sklearn, Machine Learning, Keras, Deep Learning, TensorFlow, Neural Networks, NLP, Tableau, Flask, HTML, CSS.

EXPERIENCE

Data Scientist/Engineer, May 2017 – Current
Northstar Group Inc – Gaithersburg, MD

- Architected and implemented multiple ML algorithms using Python, sklearn, R, SQL, NoSQL, flask, Keras - Deep Learning, TensorFlow. Experience in wrangling large datasets using MapReduce, Spark (PySpark), and Hive.
- Incorporated intelligence from over 1500 features from 6 different data sources involving structured/unstructured data sets and built the data transformations to feed the ML models.
- Conceptualized, architected and implemented a predictive deep learning model to accurately predict by 87% a customer defaulting on their payment resulting in reduction of the number of accounts put in collections.
- Extensively worked in modeling solutions using techniques - logistic regression, SVM, KNN, decision trees, neural networks and mastered predictive modeling, data processing, and dashboard development (Tableau, D3.js).

Business Intelligence Analyst, Sep 2011 – May 2017
S2Tech – Chesterfield, MO
Data Scientist/Analyst

- Quickly promoted from System analyst to Business Intelligence Analyst due to demonstrated ability in Exploratory Data analytics, Quantitative Analysis, mining and statistical insights.
- Involved in design, development and deployment of multiple big data machine learning initiatives and advanced ML models based on decision trees, random forest, logistic regression, and other categorical data.
- Perform hypotheses testing on relevance and value of the Datasets, parameter selection and calibration of inputs while building solution models to business problems.
- Built APIs to return/retrieve data in JSON, XML formats from multiple data sources. Developed pipelines to analyze large datasets combining custom Python/Shell scripts with established modeling tools.
- Spearheaded an Information Extraction project and developed web scrapers to collect raw data from the web and build executive dashboards and scorecards to show the trends in the data using Tableau, PowerBI and Excel.

Sr. Data Analyst/Engineer, Nov 2010 – Sep 2011
iGate Technologies – Bridgewater, NJ

- Developed forecasting models to predict, ordinal logistic regression models to compare characteristics and attributes. Developed custom methods for controls on model effectiveness. Involved in explaining and laying out model decisions for senior management and business teams.
- Built automated pipelines to reduce repetitive transformational and extract processes for datasets, in turn, monitoring the effectiveness and performance and determine root causes for significant changes, if any.
- Developed ETLs for data sources used in production reporting for marketing and operations teams

Business Systems Analyst, Nov 2009 – Nov 2010
Alberg Software Inc – Naperville, IL

- POC between Business teams and Technical Teams. Map business requirements to technical specifications.
- Implemented demand forecasting models which improved upon forecast accuracy and reduced inefficiencies.
- Architected and implemented analytics and visualization components for the analysis platform using company-wide KPIs and provide a dashboard for senior management and other key stakeholders.

Software Engineer/Business Analyst, Aug 2005 – Nov 2009
Satyam Computer Services Ltd – Zurich, Switzerland - 2008-2009, Hyderabad, India – 2005-2008

- Responsible for implementing data mining/engineering solutions to business problems in banking, insurance, retail sales lead scoring, demand forecasting, and targeted marketing
- Applied data mining, data preparation, and outlier detection techniques to gain insights into unstructured datasets and map them to actionable items.
- Ensure relevant data, tools, and infrastructure needed are available for test and development purposes by performing ETL processes and setting up automated data stream pipelines.
- Extensively used Python, SQL and other preprocessing methods for data munging and cleanup.

EDUCATION

Data Analytics and Visualization Certification
University of Texas – Austin, TX

Bachelor of Technology - Mechanical Engineering
Acharya Nagarjuna University – Guntur, India