




VINOTHINI PUSHPARAJA

+1-551-208-9119 

vinothini.pushparaja@gmail.com 

linkedin.com/in/vinothinip 

github.com/vpushparaja22/ 



OBJECTIVE

A go-getter data science graduate, passionate about solving real-world problems using statistical modeling, machine learning, and AI. My goal is to be the go-to person wherever data holds the key to solving a problem, however complex/challenging it may be.



EXPERIENCE

Student Volunteer – Campus Ministry | Saint Peter's University

January 2017 – May 2018

- Volunteered to sort donations and staff foods, clothes, and toiletries in the universities pantry/closet.
- Cleaning the streets of Jersey City as a part of Junkyard Dogs clean-up group, towards a better city.
- Sorted and organized food donations for the Community Foodbank of New Jersey. As a part of it, we prepared and served meals for around 30,000 people in and around New Jersey.

Data Scientist | Software Engineer | Wipro Technologies Ltd

June 2014 – June 2016

- Implemented machine learning and data mining techniques like K-NN, Naïve Bayes etc. to improve existing predictions of user's TV content search from search queries for vast amounts of data.
- Implemented recommendation systems using collaborative filtering and content based filtering methods.
- Resolved and troubleshoot issues escalated by customers and internal systems, identified, developed, implemented and deployed appropriate solutions to ensure system integrity.
- Developed a Python script to transfer regional data from the string file and collected in Excel sheets.
- Awarded as Top Player for being a quick learner and enhancing the performance of a team of 15 people.
- Acquired knowledge of using tools like Python, SQL, Java, Shell commands, Perforce, BugBase, Jenkins, & Agile Tools.



EDUCATION

MS in Data Science with Concentration in Business Analytics | Saint Peter's University

AUGUST 2016 – MAY 2018, GPA: **3.91**

COURSEWORK: Statistical Programming (SAS, R), Data Visualization (Tableau, Python, Plotly), Data Mining (AWS, PostgreSQL), Data Analysis (Statistics), Machine Learning (R, Python), Database & Data Warehousing (SQL), Big Data Analytics (Hadoop, Hive), Marketing Analytics (Python), Predictive Analytics and Financial Modeling (R).

BE in Computer Science Engineering | Anna University

JUNE 2010 – APRIL 2014, GPA: **3.25**

COURSEWORK: Data Warehousing and Data Mining, Artificial Intelligence, Software Engineering, Design and Analysis of Algorithms, Data Structures, Database Management Systems, Object Oriented Programming, Web Technology.



EXPERTISE

- Machine learning/AI: Python (Scikit-learn, Pandas, Numpy), R, NLP, Neural Network, CNN, RNN
- Data Visualization: Tableau, QlikView
- Database: MySQL, PostgreSQL
- Data Analytics: SAS
- Big Data: Hadoop, Hive, AWS (S3 | EC2)
- Other: Excel (Pivot Tables | VLOOKUP), Git, HTML, CSS, Google Analytics, RESTful



PROJECTS

Yelp Restaurant Image Classification | Saint Peter's University | January 2018 – May 2018

Multi-label classified 2,000 restaurants considering 237,000 images using machine learning models. The models were tested with 10,000 restaurants with 1.2 million images. Features for the models were extracted from images by pre-trained deep neural network called ResNet. Machine learning models were trained and the models performance was tested based on accuracy, precision and recall. **Tools:** Python, Scikit-learn, Numpy.

Bag of Words meets Bag of Popcorn | Saint Peter's University | September 2017 – December 2017

For IMDB movie reviews, performed sentiment analysis for each review using Bag of Words model. Utilized Bag of Words a Word2Vec method to understand the meaning and semantic relationships among words. Fed top 1000 words to machine learning models like random forest, logistic regression and support vector machine (SVM). **Tools:** Python, NLP.

Predicting Stock Prices using Time-Series and Neural Nets | Saint Peter's University | September 2017 - December 2017

In this project, we built a time series model based on quarterly results of the healthcare industry. Different variables like assets, market cap, P/E ratio. 20 Quarterly results have been analyzed and portfolio has been created with five buckets. A model with neural nets has resulted in better predictions. **Tools:** R.

Safe and Smart Parking | Solaria Labs, Liberty Mutual | January 2017 – May 2017

The Project focused on predicting safer parking location to the driver/user, by analyzing the data from public garage and street parking locations with Auto thefts reported in the city. Our analysis is performed with the public data from Seattle, Washington. An expected result would be a UI containing a map of both safe and unsafe parking lots as green and red spots. We utilized Logistic regression to predict whether the location is safe or not. **Tools:** R, R Shiny, Leaflet.

Visualization of Olympics Data 1896 – 2008 | Saint Peter's University | January 2017 - May 2017

Created Tableau Dashboard with interactive views, quick filters, & drill downs, to visualize Olympic medals obtained country wise from the year of 1896 – 2008. Click the link to view the dashboard: [Olympics1896-2008](#).

Breast Cancer Risk Prediction | Saint Peter's University | January 2017 – May 2017

For the breast cancer dataset consisting of 30 attributes, which was used to predict what diagnosis to be used for the patient either malignant or benign. A linear discriminant analysis and k nearest neighbors were modeled to classify malignant or benign breast cancer. K-nearest neighbors (KNN) achieved an accuracy of 98%. **Tools:** R

Prediction of Heart Disease | Saint Peter's University | September 2016 – December 2016

Modeled a support vector machine and logistic regression for heart disease dataset, which was used to predict the probability of heart disease events. Out of 76 raw attributes, only 14 of them were actually used. Utilized logistic regression algorithm which has provided with an accuracy of 84%. **Tools:** R