

PRADEEP SYAL

Address : Pa Verkuijllaan 71, 1171EB Badhoevedorp, The Netherlands
Phone : +31 647457192
Email : syalmannat@gmail.com
Date of Birth : 02-November-1988
Nationality : Indian

Educational Background

MSc Business Economics, University of Amsterdam, 2013 – 2015 7.33 /10
Bachelor of Technology in Biotechnology, Indian Institute of Technology Guwahati, 2007 – 2011 7.23 /10

Selected Relevant Courses

Applied Econometrics	Empirical Finance	OOP and Data Structures
Computational Biology	R Programming	Exploratory Data Analysis
Practical Machine Learning	Hadoop Fundamentals	Intro to MapReduce Programming

Technical Skills

Programming	R, Python, Stata, SQL, Apache Spark, Excel VBA, Pig, Hive
Tools	MS-Excel, Tableau
OS	Windows, Ubuntu Linux
Areas of Interest	Statistics, Machine Learning, Programming, Finance, Marketing

Projects

- Title** : Does Declining Patenting Activity in Technological Companies direct them towards Corporate Acquisitions or boost their internal R&D Spending?
Outline : *Ordinary Least Squares* and *Poisson* regression models were constructed to evaluate the effect of innovation levels on acquisitions and internal R&D spending. The robustness of results obtained was validated with *Logistic Regression Model*.
Link : <https://github.com/s-mnt/repo-beg/tree/master/Master's%20Thesis>
- Kaggle Challenge** : Which customers are dissatisfied with the banking experience of Santander Bank?
Outline : *Stochastic Gradient Boosting* Model was used to predict the probability that a customer is dissatisfied with the banking experience of Santander Bank using a dataset with a high class imbalance for the target variable. Final predictions on the test set were obtained with an AUROC value of 0.83.
Link : <https://github.com/s-mnt/repo-beg/tree/master/Santander%20Bank>
- DrivenData Challenge** : Develop a model that will predict the yield of DSH's fog nets using historical data about meteorological conditions and the fog net installations.
Outline : *Support Vector Machines* model, *Stochastic Gradient Boosting* Model and *Random Forests* were used to predict the yield of DSH's fog nets. Final predictions on test set were obtained with an RMSE of 3.7431.
Link : <https://github.com/s-mnt/repo-beg/tree/master/From%20Fog%20Nets%20to%20Neural%20Nets>
- Title** : Predict whether a credit card application will be approved or rejected
Outline : A *Bagged Adaboost* model was constructed to predict the outcome of credit card applications based on personal attributes of the applicants. Predictions on test set were obtained with 86.7 % accuracy rate.
Link : <https://github.com/s-mnt/repo-beg/tree/master/Credit%20Approval>
- Title** : Predict the age of abalone from its physical measurements

Outline : A *Linear Fit* multivariate model was constructed to predict the age of abalone based on its measurable physical characteristics. Predictions on test set were obtained with a mean of 9.98 and RMSE of 2.26. The accuracy of *Linear Fit* model was confirmed with the *Stochastic Gradient Boosting* model which shows only a slight improvement with RMSE of 2.22.

Link : <https://github.com/s-mnt/repo-beg/tree/master/Abalone>

- Title : Predict whether income exceeds \$50K/yr based on census data
Outline : A model ensembling technique was used to combine predictions from *Generalized Linear Model*, *Random Forests* and *Stochastic Gradient Boosting Model* to achieve an accuracy rate of 86.5 % on the test set scaling up the accuracy rates of these individual prediction models.
Link : <https://github.com/s-mnt/repo-beg/tree/master/Adult>
- Title : Distinguish between the presence and absence of cardiac arrhythmia in the patients
Outline : A model ensembling approach was used to combine predictions from tuned *Stochastic Gradient Boosting Model*, tuned *Neural Networks* and tuned *Support Vector Machines* to achieve an Area under ROC Curve equal to 0.7793, improving upon the individual AUROC achieved by *Stochastic Gradient Boosting Model* (0.7658), *Support Vector Machines* (0.7132) and *Neural Networks* (0.7177).
Link : <https://github.com/s-mnt/repo-beg/tree/master/Arrhythmia>
- Title : Predict the mpg (miles per gallon) of the vehicles
Outline : A bagged version of *multivariate adaptive regression splines (MARS)* model was used tuning its parameters to predict the 'miles per gallon' variable. Predictions on the test set were obtained with RMSE of 3.4 for mean value of 23 for the outcome variable.
Link : <https://github.com/s-mnt/repo-beg/tree/master/Auto%20MPG>
- Title : Modeling and Simulation of biodiesel production from microalgae
Outline : Hand collected data from chemical reactions occurring in microalgae from literature. Modeled and optimized the quantity of biodiesel produced by microalgae under different conditions.

Work Experience

IBM India (July 2011 – June 2013) : Test Specialist

Responsibilities

- Responsible for designing test cases and automated test case execution to test the softwares
- Responsible for constructing statistical prediction models to predict software reliability and defects
- Responsible for generating test data with appropriate probability distribution and constraints as per business needs
- Responsible for creating dashboards to report status of testing phase to client

Certifications

CFA (Chartered Financial Analyst) Level I passed

NSE (National Stock Exchange of India) Certified Market Professional Level 3

WIPO Worldwide Academy Certificate – General Course on Intellectual Property

Test Scores

- Bloomberg Aptitude Test (BAT), March 2015 : 94 percentile (600 points)
- Graduate Management Admission Test (GMAT), April 2012 : Quantitative – 92 percentile (50 points)
- Joint Entrance Examination, 2007 : 99 percentile
- All India Engineering Entrance Examination, 2007 : 99 percentile

Language Skills

- Proficient in English, Hindi and Punjabi
- Dutch : Beginner