Suvro Banerjee

Address : 801 Pine St Apt 17F Seattle, WA 98101
Phone: (716) 430-3246 ; Email : suvrobaner@gmail.com

EDUCATION

University at Buffalo, State University of New York
Master of Science, Economics, GPA : 3.8 / 4.0

West Bengal University of Technology, India
Bachelor of Engineering, (ECE)

SKILLS

- Data Science, Statistical Learning, Econometrics, Predictive Modeling, Data Mining & Machine Learning.
- R, Python, SAS , SQL, AWS, Siebel CRM, Microsoft Dynamics CRM, ITIL, Analytics, Excel, Scripting & QC.
- R Packages :lattice, gmodels, MASS, leaps, glmnet, pls, tree, randomForest, e1071, support.CEs, ts, wordcloud, tm, arules, lubridate, ggplot2, queueing, lpSolve, neuralnet, gam, quantmode, forecast, ChoiceModelR.
- Python Packages: pandas, numpy, scipy, statsmodels, matplotlib, patsy.contrasts, re, sklearn, nltk.
- Certification : SAS Certified Base Programmer for SAS 9
- LinkedIn : https://www.linkedin.com/pub/suvro-banerjee/2a/544/934

PROJECTS (using R and Python)

- **Discrete Choice using Market Simulation on a Computer Choice Study** : Predict individual consumer choice and estimate Brand Loyalty, Price Sensitivity, Feature Focus, Brand Switching and Optimal Pricing. Hierarchical Bayes (HB) Multinomial Logit Model using Markov Chain Monte Carlo estimation (MCMC). Visualization : Mosaic plot, Triplot /Ternary plot, Comparative Density plot and Parallel Coordinates plot.
- **Conjoint Analysis for Mobile Communication services :** How product attributes affect purchase decisions? Choice Experiments, Multiple Linear Regression using contrasts coding to compute part-worth's and attributes relative importance. Visualization : Spine Chart.
- **Unsupervised Text Analytics of movie taglines (last 100 years) from IMDb using Bag-of-words approach:** Multidimensional Scaling, K-Means Clustering, Partitioning Around Medoids (PAM) Clustering, Hierarchical Clustering, Principal Component Analysis with the help of Biplot, Dendrograms and Silhouette plots.
- **Market Basket Analysis of Grocery Store Data :** To find Association Rules for the Market Baskets. Apriori Algorithm with the help of Item Frequency plot, Association Rules scatterplot and Matrix Bubble Chart.
- **Filtering Mobile phone Spam :** Develop a Classification model that could filter SMS spam using naive Bayes.
- **Gas mileage for automobiles :** To predict whether a given car gets high or low gas mileage (Classification Problem).Logistic Regression, Linear Discriminant Analysis, Quadratic Discriminant Analysis and Support Vector Machines.
- **Optimal Workforce estimation and scheduling for a call center of a bank :** with reference to a target wait time and minimize the overall cost of the call center operations using Queueing Model (Erlang-C) and Integer Linear Programming (to solve constrained optimization problem) with the help of Ribbon and ggplot features.

- **Predictive Model for Los Angeles Dodgers Promotion & Attendance :** Do bobblehead promotions increase attendance, controlling for the date of the game? Using MLS data for 2012 season using Multiple Linear Regression with the help of Box plot, Lattice plot and Trellis graphs.

- **Optical Character Recognition :** OCR using Support Vector Machines (SVM) using linear and radial kernels.

- **National Cancer Institute Data (High Dimensional Data) :** Unsupervised / Supervised learning of cancer cell with gene expression levels using PCA, Hierarchical and K-Means Clustering and SVM (supervised).

- **Diagnosing Breast Cancer :** To predict 'Malignant' & 'Benign' cell nuclei from the measurements of biopsied cells using k-Nearest Neighbors algorithm and analyzing the result using a confusion matrix.

- **Identifying risky bank loans :** Develop a credit approval model using Divide & Conquer (C5.0 decision tree).

- **Modeling the strength of concrete :** Predict the strength of concrete using Artificial Neural Network (ANN).

- **Sales of Child Car Seats :** Predict the sales (Regression) and Classify whether the sales for the car seats are above average at each location (Classification) using Tree with pruning, Bagging and Random Forests.

- **Time Series Analysis of major Economic Indicators :** Analysis and forecasting (24 months) of 'National Civilian Employment Rate', 'Manufacturer's New Orders-Durable Goods', 'University of Michigan Index of Consumer Sentiment' and 'New Homes Sold in the US' using the ARIMA model and Horizon and Chart Series plots.

- **Housing values in the suburbs of Boston :** Predict whether a given suburb has a crime rate above or below the median (Classification Problem) using Logistic Regression & LDA, predict the per capita crime rate (Regression Problem) using Best Subset Selection, Validation Set Approach, K-fold cross validation, Ridge & Lasso Regression, PCR and PLS and predict the median value of the owner occupied homes (Regression Problem) using Regression Trees with and without pruning, Bagging/Bootstrapping, Random Forest and Boosting.

PROFESSIONAL EXPERIENCE

Rosen Capital Advisors
Summer Intern Analyst, Los Angeles, (June 2014 - August 2014)

- Data mining/analysis of S&P 500 and implementing in-house trading models using Black Scholes.
- Implemented Microsoft Dynamics CRM to manage daily operations and report using charts and analytics.

Accenture
Senior Software Engineer, Bangalore, 2.5 years (January 2011 - July 2013)

- System Design and Application Development using Siebel Customer Relationship Management for telecom.
- Software implementation of Opportunity, Quote, Order, Asset, Service and Activity management modules.
- Migrated big data using Enterprise Integration Manager (EIM) tools from flat files & legacy to Siebel database.
- Worked extensively on functionalities using E-Scripting, Workflows, Business Services, Virtual / Business Components, Applets, Run time events, User properties, DB links, SQLs and Product configuration.
- Managed the development team, iteration wise (Agile Project Management) and worked with Analytics.

Tech Mahindra Ltd
Software Engineer, India & UK, 3.5 years (July 2007 - December 2010)

- Transitioned legacy systems to Siebel CRM from vendors, implemented functionalities using Siebel Configuration, Workflows, Scripting and Oracle SQL for a telecom firm.
- Upgraded Siebel servers and Microsoft SQL server to higher versions.
- Implemented various Interfaces using EAI integration objects, Inbound / Outbound Web services.