

NAVITA JAIN

navita.jain01@gmail.com | 518-369-3201 | Dublin, CA, 94568 | LinkedIn | GitHub

Skills

Languages | Python, Java, JavaScript, SQL, HTML

Machine Learning, Analytics | Clustering, Classification, Regressions, Feature Selection (Random Forests, k-Means, Support Vector Machines(SVM), Decision Trees), Market Basket Analysis, Text Mining, Bayesian Predictive Analysis

Libraries, Tools and Frameworks | Pandas, NumPy, SciPy, Matplotlib, scikit-learn, Jupyter Notebook, Git, Amazon Web Service(AWS), Tableau, Lucene, TkInter, MS Excel

Experience

Data Analyst, Girl Scouts of Northeastern New York Feb 2017- Oct 2018

- Performed exploratory data analysis(EDA) and statistical analysis on internal and external data to generate insights.
- Applied Random forest feature importance to identify the important feature/s for member retention.
- Used machine learning techniques to perform trend and pattern analysis on members, programs, donors data, etc.
- Demonstrated findings to directors and executive staff through presentations to facilitate informed strategic business decisions.
- Involved in large scale data migration from existing database to a new database.
- Developed a Python Application using TkInter as a support tool for data migration.
- Queried database using SQL and cleaned it using Python for modeling.

Graduate Research Assistant, Data Mining lab, University at Albany Apr 2015 - Jan 2017

- Developed efficient statistical and data mining techniques for community detection in social networks.
- Modeled topic trends in time stamp data using probabilistic and generative techniques.
- Implemented two state burst detection model on time series data for graph network using dynamic programming.

System Engineer, Infosys Limited Mar 2010 - Sep 2012

Oracle DBA, Sep 2011 - Sep 2012

- Responsible for database administrative activities for a large and complex Oracle11g and Oracle9i database system.

Salesforce Developer, Mar 2010 - Aug 2011

- Implemented, administered and maintained client's Salesforce business application.
- Created Custom Objects, Custom fields, Custom Tabs, Reports and other components as per clients requirements.

Projects

Synchronised Communities in large scale graph networks, Master's Thesis

- Synchronised behavior analysis using spike and ISI-distance on large social networks.
- Performed statistical significance and hypothesis testing using P-value to determine coordinated behavior between nodes.

Influence detection using Politeness and Likeability, Natural Language Processing and Machine Learning

- Developed text analysis model for sentiment and politeness detection.
- Implemented Naive Bayes and Support Vector Machine(SVM) classifier in Python.
- Prepared annotated Twitter dataset with politeness and sentiment labels.

Online Voting Application: BallotBox, Web Development and Software Engineering

- Build the voting module for a customizable online voting application using MVC model and performed full integration.
- Designed and deployed MySQL relational database and the web application on AWS..

Document Retrieval for Explicit Temporal Queries, Information Retrieval

- Proposed a variant of Term Frequency -Inverse Document Frequency (TF-IDF) for relevant documents retrieval in explicit time sensitive search.
- Performed detailed evaluation of proposed solution and existing Java based information retrieval software Lucene based on precision and recall.

Discovering Social Circles, Data Mining

- Proposed a solution for categorization of nodes into social circles using k-way partitioning and Jaccard Similarity.
- Evaluated the performance of proposed solution using balanced error rate and F1 score accuracy metrics.

Education

Master of Science, Computer Science, State University of New York, Albany Dec 2016

Online Course: The Data Scientist's Toolbox, Coursera, by Johns Hopkins University Jun 2014

Bachelor of Technology, Information Technology, Rishiraj Institute of Technology Sep 2009