VAIBHAV SHARMA

New York City | vaibhav47sharma@gmail.com | (434) 249-1469
Website | LinkedIn | GitHub

EDUCATION

University of Virginia, Charlottesville, VA Jul 2018 - May 2019

- Master of Science in Data Science, GPA: 3.85/4.00
- Relevant Coursework: Machine Learning, Data Mining, Data Visualization, Statistical Computing, Linear Models

Birla Institute of Technology & Science, Pilani, India Aug 2010 - Jul 2015

- Bachelor of Technology in Electrical & Electronics Engineering
- Master of Science in Mathematics

SKILLS

Python (pandas, Matplotlib, seaborn, scikit-learn, TensorFlow, NLTK, Gensim, Scrapy, Beautiful Soup), PySpark, SQL, SQLite, R (ggplot2, tidyr, dplyr, plotly, tm, tidytext), Java, Git, Tableau, UNIX, AWS, SAS, D3.js

PROFESSIONAL EXPERIENCE

Goibibo, Gurgaon, India

(Launched in 2009, Goibibo is one of the largest online Flights and Hotels aggregation and booking engines in India)

I developed and executed the functionality of the flight plus hotel vertical of the Goibibo Android mobile app. Won several awards and was rewarded a double promotion in March 2018 after which I spearheaded revamping the analytics engine of the app

Software Engineer Jun 2015 - Mar 2018

Senior Software Engineer II Apr 2018 - Jun 2018

Product & System Design and Android Development

- Wait-less flight booking solution
  - Conducted A/B tests to identify bottlenecks in the booking process
  - Implemented a wait-free solution on the mobile app resulting in a faster booking process
  - Led to an increase of 1% in conversion rates, lower app dropouts and improved user reviews
- Flight grouping using sorting and hashing: Project to analyze lower conversion rates in the Flights vertical
  - Analyzed conversion rates in the Flights vertical using search result page (SRP) product impressions data
  - Optimized SRP results on price, baggage options and booking history to increase end-user flexibility
  - Led to a 0.5% increase in click rates from SRP and a 0.3% increase in conversion rates

Analytics

- Mobile and Product Analytics
  - Developed 10 dashboards for six business verticals for consistent data reporting
  - Increased cross-functional readability across all verticals for the marketing team
  - Utilized to record and monitor performance metrics by the backend technology team
- Customer Journey Analytics
  - Ideated and implemented a user tracking plan to identify funnel drop-off
  - Led to an overall 4.1 rating in the Google Play Store for the first time in over 1.5 years
  - Expanded scope of the project to other business verticals; heavy implementation for user complaint resolution

PROJECTS & PUBLICATIONS

Android Application Development

Developed and released two android applications (Free Open Wi-Fi Connect and Wi-Fi Hotspot Tethering) on Google Play Store

URL: https://play.google.com/store/apps/details?id=com.easyway.freewifi

- The app efficiently searched and connected to free open Wi-Fi networks
- 500,000+ total downloads and a rating of 4.3/5 on Google Play Store
- Incorporated user feedback, and monitored crash and click rate to release new improved versions of the app

Designed and Developed a VR Chatbot for Army Training

A research project sponsored by the Army Research Laboratory and the Data Science Institute, UVA

- Tools/Software Used: Python, SQL, Amazon Mechanical Turk, Django Framework, R, Unity, JIRA
- Developed a speech based chatbot within a data-driven framework to train Army personnel to be culturally sensitive
- The system analyzed the cultural competence of a user through speech input via dialogue and decision trees
- Full-fledged product equipped with a VR front end and a backend written in Python Django framework where the speech input is converted into text, classification models are run, feedback selected, and the feedback and score sent via an API
- Leveraged NLP/text mining for feature generation - TF-IDF, LDA, word2vec, and GloVe
- Classification Algorithms used – Logistic, Random Forest, Decision Trees, Gradient Boosting, and Neural Networks. The best model (Gradient Boosting) gave an F1-score of 0.92
- 80% of the users declared that the system accurately judged what they were saying and provided apt feedback
- Paper on Data Generation and Acquisition - *https://datascience.virginia.edu/projects/siri-translate-laojia*
- Research paper approved for publication in the IEEE SIEDS Journal (June 2019)
- Presented research at the IEEE SIEDS 2019 Symposium, Charlottesville, VA

Natural Language Processing for Sentiment Analysis

- Built ETL pipeline to analyze and conduct sentiment analysis on 0.5 million Amazon Fine Food reviews
- Text mining techniques (TF-IDF, Word2Vec, n-grams model, LDA) for feature generation and classification models for prediction
- Logistic Regression on bi-gram (Word2Vec) achieved a 93% accuracy with 0.85 AUC

Deep Learning for Detection of Neural Granger Causality in the Financial Market

- Leveraged MLP and RNN to build models that could detect correlation in stock prices
- Used stock market data for 5 stocks spanning a period of 5 years
- Models were successfully able to detect Granger causality between 3 out of the 5 stocks

Developing an Image Processing Pipeline for Image Recognition

- Developed an image processing pipeline which automatically searches for images, standardizes its dimensions and applies 10 different image augmentations to expand the image corpus for the purpose of training image recognition models