Soumit Bose

Data Science and Analytics

Email: bosmit90@gmail.com
Phone: +91 9163703490
LinkedIn : www.linkedin.com/in/sbose95/
GitHub : github.com/smitbose

Experience

1. HSBC Electronic Data Processing India Pvt Ltd   June 2017 –Present

Senior Analyst

Worked with multiple banking operations spread across the globe, in requirement gathering, data pipelining, solution ideation and development of cognitive solutions for optimizing banking operations.

- **ChatBot Assistant:** A ChatBot meant for banking operations to understand regularly, process information, performance and bottlenecks
  - Developed a natural language processing layer for entity and intent detection, using different solutions like Naïve Bayes classifier and Neural Network based libraries like Spacy and Dialogflow API on Google Cloud Platform for interpreting data query needs expressed in human language
  - Designed a RESTful architecture to integrate it with a messaging client, and implemented service endpoints with Python Flask
  - Created a Data model to suit chatbot efficiency and a query translation engine that converts intents and entities to SQL queries and fetches data from the database
  - Implemented end to end a Knowledge Graph that disintegrates and retains information from different documents, for quick access to process related information
- **Call Fraud Detection:** Implemented an AI real-time call fraud detector, which compares the voice of the calling customer with the previous instances of calls from the same number
  - Created a pipeline for sourcing in audio, decrypting it with mu-law encryption and splitting the channel to identify the customer track
  - Used Fourier Transform to extract audio features and did extensive feature engineering to identify key features
  - Collaborated on development of several models like Decision Trees, SVMs and Autoencoders for identifying audios that differ significantly from previous calls from the same number
- **Data Warehouse Strategic solution:** A Data Warehouse that accumulates application data from heterogenous sources into a standardized structure for different analytical dashboards
  - Gathered requirements from Business, Operations, Technology and Analytics teams and unified all the perspective to create a structure best suitable for organization's needs
  - Designed data model, for different data marts and amalgamated them into a warehouse
  - Reviewed and tested ETL pipelines for consistency and accuracy of sourced data

2. Complex Networks Research Group   May 2016 – July 2016

Research Intern

- Worked on scaling graph-analytics algorithms with distributed processing on Hadoop cluster using Apache Giraph

3. Information Technology Research Academy (MHRD) & Digital India Corporation   May 2015 – July 2015

Research Intern

- Developed a routing model for disruption tolerant communication devices using graph theory, to minimize energy consumption and maximize message delivery

Academics

1. Indian Institute of Engineering Science and Technology, Shibpur   August 2013 – May 2017
Bachelors in Engineering – Computer Science and Technology   CGPA: 8.76

2. Vivekananda Mission School   2013
Indian School Certificate (ISC) Examination   95.5%

3. Vivekananda Mission School   2011
Indian Certificate of Secondary Education (ICSE) Examination   93.2%

Achievements

1. *Spark - Global Analytics Center, HSBC* – **Sep 2019**
   Global Analytics Center award for best Innovation project.
2. *Leading Light - Global Analytics Center, HSBC* – **Q2 2018**
   GAC quarterly award recognizing excellent service delivery in terms of executing regular projects along with leading POCs
3. *Winner: Sheldon, Vision Guided Robotics Contest IIT Kharagpur* – **Jan 2016**
   Competed against top colleges in India in solving a robotics challenge of identifying numerical images and operators using Image Processing and making the bot traverse them in an order that either minimizes or maximizes the equation composing them, according to the instruction
4. *Runner-up: Aughit, Vision Guided Robotics Contest IIT Kharagpur* – **Jan 2015**
   Solved a problem of playing the famous Aughit game with a bot, being fed with live video feed of the game simulation
5. *ACM-ICPC Kharagpur Regional honorable mention* – **Dec 2014**
   Qualified to the onsite regionals of ACM-ICPC, the biggest competitive coding event, as the top team from university

Publications

1. **Social-based energy-aware multicasting in delay tolerant networks**
   In Journal of Network and Computer Applications, Volume 87, 2017, Pages 169-184, ISSN 1084-8045
   - Created a Network Routing model based on Contact Graphs that can send data packets at the same time to multiple nodes
   - Wrote an algorithm to increase probability of message delivery and reduce energy wastage, using a dynamic programming based mathematical formulation, that concluded betweenness centrality to be the best measure
   - Modified a Delay Tolerant Network Simulator written in Java, to support multicast routing, which helped in gathering corroborating data
   - Wrote an algorithm that builds the betweenness centrality score and communities in Graphs through continued interaction data among nodes

Skills

Python	Machine Learning	Statistics	SQL	Java
Google Cloud	SAS	MATLAB	Unix	Git

Personal Projects

1. **Regional Political Influence Using Twitter –** Implemented as part of Final Year Undergraduate Project, which tries to gauge user's political inclination on the basis of tweets and political interactions using links in a Social Network Graph
2. **Open-Source Contribution: Converse.js -** Contributed to an online messaging client library written using JavaScript

Hobbies

Reading books – turning few pages from a book every day, irrespective of genre, listening to music and occasionally trying out myself on a keyboard or violin and watching football and gathering any statistic that hypothesizes a favorable outcome for my favorite team week in week out.