Arjun Sridharkumar

(+91) 9321782011 arjunsridhar9720@gmail.com

Experience

Uoft(University of Toronto)

Fall 2020 - Present

Research Intern

- Interning under Dr Shurui Zhou[Uoft].
- Working to generate on demand documentation using blogs as a major source.
- Blogs (especially medium) has become more prevalent over the years and often helps in installing/getting started while learning new technologies. On the other hand, the entry point for a API/technology is convoluted and hence other developers experiences can aid first time users.

Concepts: Software Engineering, Social computing, Mining S/W repositories
Tech: Python

Tata Research Design and Development Centre

Data Scientist Researcher *June 2018 - Present*

Part of the Software Research group focused towards enriching our AI product Ignio [https://digitate.com].

- Designed and developed statistical and machine learning based algorithms to predict outages/failures in the IT Enterprise.
- Implemented a machine learning based prediction model for TCS New York marathon with a accuracy of over 90%.
- Derived drug misuse insights from past data to aid one of the largest pharmaceutical retailers in Canada.
- Designed a regression based queuing based solution to aid PostNord (Scandinavian Courier company) to optimize their overall operations.
- Introduced an analytics based What-if scenario modelling at one of the largest banks in UK (NBS) thereby saving their planning time.
- Performed entity extraction from structured/unstructured text data using statistical and contextual analysis for Woolsworth(Australian retail chain) and Meijer (American retail chain).
- Designed a Question/Answering system that enables users to literally talk to data and ask questions. (Currently working to patent this)
- Predicted and visualized network congestion and failure of batch jobs for Meijer (American retail chain) during Black Friday sale.
- Designed algorithms to detect and visualize network fault localization and root cause analysis.
- Designed a NLP based Q&A system for questions in English to aid our clients in interacting with Ignio. [Talk it Out].
- Winner of the Digitate Ideathon 2020 (Ranked 1 among 400 other employees)

Concepts used: Time-Series Forecasting, Classification, Clustering, Data Visualization, Data Engineering, Data Analytics, Network Science, NLP, Health informatics, AI in ITOPS
Tech Stack: Python, R, d3.js, Tableau, Excel

Britomatics Engineering

Deep learning Intern *June 2017 - June 2018*

- Incorporated a KNN classification algorithm to sort emails and auto-forward them to respective departments.
- Implemented a extractive contextual summarization technique to summarize mails based on previous conversations to aid in saving employee effort and time. An additional aim to summarize was to generate replies (similar to Gmail smart compose).

Concepts used: Semantic Similarity, Word2Vec, NLP, Sentence Similarity, Classification, Optimization.

Softalk Consultancy Services

Data Visualization Intern *Dec 2016 - Jan 2017*

- Designed a social media website for one of their clients using PHP framework YII.
- Mentored and taught visualization and web design techniques to junior interns and students.

Concepts Used: Edward Tufte's Principles, Data Visualization, Change mapping.

Aditya Birla Group

Data Security Intern *May 2016 - June 2016*

- Formalized a decision tree based classification mechanism to detect leakage of confidential data through emails.
- Designed anomaly detection and time series profiling techniques to classify internet usage behaviour of employees.
- Consolidated and summarized all the vulnerabilities and prevention steps for the last quarter of 2015.

Concepts used: Data Security, Classification, Malware detection.

Education

Fr. Conceicao Rodrigues Institute of Technology
(University of Mumbai)

Bachelors' of Engineering (B.E.) in Information Technology — CGPA: **8.5** — *2014 - 2018*

South Indian Association of High School and Junior College

12th Grade — Percentage: **80.31%** — *2012- 2014*

10th Grade — Percentage: **88.55%.** — *2000 – 2012*

Projects

Predicting Flight delay *March 2017*
- Aims to predict delays in flight schedules at any airport.
- Implemented using pattern detection, temporal correlation analysis and trends
- Demonstrated an accuracy of over 93%.

Concepts used: Temporal Pattern Analysis, Correlation Analysis, Trend Detection, Time series forecasting.

Document retrieval using KNN *June 2017*
- Aims to find similar documents/articles based on term frequency- inverse document frequency (tf-IDF) values
- Implemented on Wikipedia dataset to yield related articles/pages given any article input.
- Demonstrated an accuracy of over 95%.

Concepts used: Classification, Natural Language Processing (NLP).

Analysis of medical history of US citizens *June 2017*
- The Behavioural Risk Factor Surveillance Systems (BRFFS) dataset was explored to understand the distribution of diseases across the country.
- Devised models to determine the relationships between affordability of doctors and general health in certain areas.
- Determined the impact of smoking, food habits and diet on the general health of the population.

Concepts Used: Sampling, Data Visualization, Correlation, Clustering.

Sentiment Analysis of Amazon product reviews *June 2017*
- Designed a classification-based algorithm for sentiment analysis of product reviews.
- Determined the most positive and negative reviews for a kid's toy.

Concepts Used: Natural language processing, Classification

Certifications

Issued by Datacamp:

Deep Learning in Python	*Nov 2018*
Intermediate Python for Data Science	*May 2018*
Intermediate R	*May 2018*
Sentiment Analysis in R: the tidy way	*Mar 2018*

Issued by Coursera:

Conducted by deeplearning.ai

Neural Networks and Deep Learning	*Jan 2018*

Conducted by Duke University

Introduction to Probability and Data	*Nov 2017*

Conducted by John Hopkins University

Introduction to R Programming	*June 2017*

Teaching experience

Computer Society of India - Chapter 2016
- Presided over the Computer Society of India (CSI) Chapter 2016 and led a team of 20 passionate students.
- Organized 7 Workshops that enlightened around 200 fellow students and peers with niche topics such as ethical hacking, virtual reality, etc.
- Conducted a 4-hour long Workshop on Python for data science for 50 students that illustrated the whole journey from data cleaning to predictions.

Honors and Awards

- Qualified among the top 1000 students out of 100,000 students from all over the country for Smart India Hackathon 2018
- Qualified among the top 1000 students from over 200,000 students in "CodeVita" a National Coding Competition conducted by Tata Consultancy Services (TCS)

Other Skill-sets

- Video Editing (Adobe premiere pro),
- Basic Photoshop