JIMIT DHOLAKIA

jdholakia@cs.stonybrook.edu | linkedin.com/in/jimit105 | github.com/jimit105 | jimit105.medium.com

EDUCATION

Stony Brook University **Stony Brook, NY**
Master of Science in Computer Science, GPA: 3.64/4.0 *Expected Dec 2022*
Relevant Coursework: Data Science, Big Data Analytics, Data Mining, Algorithms, Computer Networks, Teaching Assistantship

KJ Somaiya College of Engineering **Mumbai, India**
Bachelor of Technology in Computer Engineering, CGPA: 9/10 *May 2018*
Focus Areas: Software Development, Database Management, Machine Learning, Software Project Management

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, Java, C, C++, SQL
Web Technologies: Flask, FastAPI, React.js, HTML, CSS, Bootstrap
Data Science: scikit-learn, PyTorch, TensorFlow, OpenCV, spaCy, NLTK, PySpark, Hive, Pandas, NumPy
Others: MongoDB, Redis, Git, Docker, Linux, Agile Methodologies, LaTeX, Streamlit

WORK EXPERIENCE

Amazon Web Services **Seattle, WA**
Software Development Engineer Intern *May 2022 – Aug 2022*

- Prototyped a responsive user interface for an upcoming AWS product using React.js and Python, including unit testing using the Jest framework, by collaborating with cross-team members resulting in accelerated go-to-market
- Spearheaded the development of new backend functionality, using Java, to improve the productivity of the users

SUNY RF (Dept. of Biomedical Informatics, Stony Brook University) **Stony Brook, NY**
Senior Research Project Aide - NLP *Mar 2022 – May 2022*

- Conducted research using NLP technologies including spaCy, BERT, and Zero-shot text classification on the PASC data
- Assisted professors to analyze & interpret research findings, then presented recommendations based on data

Jio Platforms, a subsidiary of **Reliance Industries Ltd.** **Mumbai, India**
Data Scientist *Jul 2018 – Jul 2021*

- Pioneered a Document Validation System to automate the processing of KYC documents reducing the time from 15 minutes to 2 minutes for over 5000 documents daily using Python, OpenCV, OCR, Fuzzy String Matching, and Kafka
- Conceptualized an algorithm to find potential duplicates from Material Master Data, estimated to have 10-40% cost savings and reduce the efforts of MDM users by 50%, in 3 sprints
- Leveraged Machine Learning to develop a prediction model for MRO with an accuracy of 90% for 95% of materials
- Transformed the medication search service for RF Hospital utilizing Natural Language Processing and Trie data structure for ultra-fast search times with a mean response time of 30 milliseconds
- Awarded with the R-Sammaan Recognition Awards by four senior leaders and received various LinkedIn recommendations for designing and accomplishing projects that solve complex business use-cases, and delivering optimal results

ACADEMIC PROJECTS

Job Title Analysis (Stony Brook University) [Link] **Oct 2021 – Dec 2021**

- Devised a Machine Learning model to predict salaries having an R^2 score of 0.901 utilizing Natural Language Processing techniques on textual data and combining it with ordinal features
- Performed clustering on Job Titles and skills using Fuzzy String Matching, TF-IDF Vectorizer, and DBSCAN Clustering to identify similar jobs and achieved a Silhouette Coefficient of 0.134

Website Fingerprinting using Deep Learning (Stony Brook University) [Link] **Oct 2021 – Dec 2021**

- Implemented Convolutional Neural Network and Random Forest based Automated Website Fingerprinting on TOR network to identify attack-prone features with an accuracy of 94.02%

Personalized Web Search based on User Profiling (KJ Somaiya College of Engineering) **Jul 2017 – Mar 2018**

- Implemented Topic Modeling using Latent Semantic Indexing (LSI) and word2vec models; dynamically created hierarchical clusters of browsing history to re-rank the Search Engine Results Page for personalized results
- Published the paper “Mining User's Browsing History to Personalize Web Search” in 2018 ICICCT, IEEE [Link]