Divyanshu Bhardwaj

Boston, MA | 857-334-9377 | bhardwaj.di@northeastern.edu | LinkedIn | GitHub

Education

Master of Science in Information Systems, Data Science (*GPA 3.92*) *Graduating May 2023*

Northeastern University, Boston

Relevant Courses – Deep Learning & Reinforcement, *Big Data Analytics and intelligent systems*, *Data Science Engineering MT*

Bachelor of Engineering in Electronics *Jun 2018*

University of Mumbai, Mumbai

Relevant Courses – OOPS, *Structured Programming Approach*

Technical Skills

Languages : Python, R, Java, PL/SQL, PySpark, JSON
Databases : MySQL, Oracle 11g, PostgreSQL, SSMS, Google Cloud Storage, Redshift, S3
Libraries : NumPy, Pandas, scikit-learn, matplotlib, seaborn, OpenCV, sympy, SciPy, TensorFlow, Pytorch
Mathematics : Statistics, A/B testing, Multi-arm Bandit, Bayesian Inference, Markov Chain Monte Carlo
Tools : Tableau, Power BI, Spark, CoLab, JIRA, GIT, Excel (VLOOKUP, Macros), GCP, AWS, SageMaker, EC2, Selenium Webdriver

Experience

Northeastern University, *Boston, USA*

Graduate Research Assistant- Under Prof Shawn Bhimani *Aug 2022 - Present*

- Working on huge text of corpus, released by companies to prevent modern slavery in global supply chain using machine learning
- Carried out text processing and cleaning, extracted topics using Gensim LDA and N-grams; also developing metrics for analysis

Graduate Teaching Assistant (Intro to Statistics and Data Analysis) *Sept 2022 - Present*

- Assessed student's learning to determine weaknesses & needs, improved their coding skills through assignments
- Supporting the professor to mentor and manage 30+ students' performance reviews through dashboards

Motorola Solutions, *Chicago, USA*

Data Scientist Intern *May 2022 - Aug 2022*

- Collaborated with cross-functional team to develop a solution by identifying relevant data sources, structuring problems
- Implemented **collaborative filtering** using **Matrix Factorization**, Achieved an RMSE score of **1.58** using ALS technique
- Optimized the algorithm for correlating incidents reducing the manual effort by **40%** for **2M+ records**
- Developed a solution to identify events that can be reduced in FM System, **90M+ events a year**, overall, **11**% reduction
- Extracted required data from MySQL server and created dashboards in tableau to improve the business needs

L&T Infotech (LTI), *Mumbai, India*

Data Scientist *Jan 2020 - Jun 2021*

- Exercised data mining, web scraping on multiple instances and discovered areas that needed manual effort reduction
- Planned, designed, and constructed a python-based automation for clients, **resulting in $150K cost savings**
- Automated SOX team process, reducing the manual effort of **120 hours a week** and **30% manpower reduction**
- Monitored KPIs and DelEx parameters, interfaced with clients to evaluate business needs & proposed convenient solutions
- Developed a query to extract reports from a database with **2 million rows of data** using python

Software Engineer/SOX Analyst *Aug 2018 - Dec 2019*

- Enhanced complex SQL queries to efficiently extract large volumes of data from various sources
- Collaborated with a team of 10 members to implement processes that maintained SOX compliance
- Deployed an auto ticket triaging tool which reduced resource efforts that had to work on 24/7 shift

Technical Projects

Image classifier for the Intracranial Hemorrhage Detection (*Keras, TensorFlow, Conv2D*) *May 2022 - Jun 2022*

- Built a CNN model with RSNA dataset that classifies hemorrhages using 3 convolutional layers & 2 fully connected layers
- Applied MaxPooling, BatchNormalization and Dropout techniques to increase the accuracy on validation data to 89.55 %

Business Meeting Summarization (*GCP Cloud Function, Firebase, FastAPI, Airflow, SQL*) *Jan 2022 - May 2022*

- Orchestrated ETL pipelines and automated video suggestions using Apache Airflow, Established API Endpoints using FastAPI
- Provided the model as a Service & Formulated SQL queries to join from multiple databases and visualize results in Bigquery

Yelp Reviews Topic Modeling: (*Python, NLP, spacy, Gensim*) *Oct 2021 - Dec 2021*

- Implemented sentimental Analysis on Topic Modeling to analyze the difference in positive & negative reviews using LDA algorithm
- Evaluated feature extraction techniques - count and tf-idf vectorizer for optimal sentiment predictions