

EDUCATION

University of Nottingham, United Kingdom

Masters in Data Science

Sep 2024-Present

Dr. A.P.J. Abdul Kalam Technical University, India

Bachelor of Engineering in Electronics and Communication

CGPA: 7.62

July 2022

PROFESSIONAL EXPERIENCE:

Project Engineer

Wipro, India

April 2022 - June 2024

- Engineered robust automated test scripts using Java and Selenium to ensure comprehensive coverage of critical application functionalities, effectively saving multiple hours of manual test time in each sprint.
- Implemented RESTful APIs using Spring Boot controllers to seamlessly retrieve and display data from databases, optimizing efficiency in data handling and presentation.
- Developed and implemented a Selenium Test Automation Framework to streamline script creation with TestNG, allowing for prioritization and controlled execution sequencing.
- Led the implementation of BDD testing framework using Cucumber, improving team collaboration and test coverage.

Machine Learning Intern

Internity Foundation, India

June 2021 - July 2021

- Engineered end-to-end NLP pipeline for movie reviews sentiment classification using Python, scikit-learn, and NLTK.
- Optimized model performance through comparative analysis: achieved 75.02% accuracy with Multinomial Naive Bayes using Bag of Words versus 74.70% with TFIDF vectorization.
- Implemented comprehensive text preprocessing including tokenization, lemmatization, and feature engineering for 25,000 movie reviews.

Data Science Intern

Internity Foundation, India

Dec 2020 - Jan 21

- Analyzed web server logs using Python to detect and geolocate security vulnerabilities.
- Implemented automated data collection and cleaning pipelines for URL log analysis.
- Enhanced web security through pattern recognition and anomaly detection.

RELEVANT SKILLS:

Java, Python, Data Structure, OOPS Concept, Algorithm, Data Science, MySQL, Spring Boot, Automation, Jenkins, Git, Azure, Selenium, BDD Cucumber, Web Scraping, RESTful API

PROJECTS:

Deepfake audio Detection with Neural networks using Audio features

Jan 2022 - Apr 2022

- Developed a CNN-based audio classification model for distinguishing between human speech and artificial voices.
- Achieved accuracies across different audio features: MFCC (75.97%), STFT (73.34%), FFT (71.48%), and Mel Spectrogram (70.86%).
- Achieved enhanced accuracy of 80.47% by implementing feature concatenation approach with ReLU activation.
- Trained and validated models using dataset of 28,700 ASV spoof audio samples.

Sentiment Analysis using Naive Bayes

June 2021 - July 2021

- Implemented an NLP model using the Multinomial Naïve Bayes algorithm for sentiment analysis on the IMDb dataset, incorporating movie ratings as one of the features.
- Trained the model on the IMDb dataset, consisting of 25,000 reviews, to learn sentiment patterns and associations with movie ratings.
- Employed stemming, bag of words, and tokenization as distinct functions to extract features from the text. These functions were instrumental in converting the textual data into vectors, serving as features for the Multinomial Naïve Bayes algorithm.

Stock Market Prediction using Lasso Regression With Time Lag

Feb 2021 - April 2021

- Built LASSO regression model for NIFTY50 stock prediction achieving RMSE of 14.78 and MAPE of 2.98% on test data.
- Implemented time series analysis incorporating technical indicators and market trends

- Applied the time lag technique to capture temporal dependencies in the stock market data.
- Engineered feature extraction pipeline for historical price data and technical analysis metrics

Web Server Log Analysis -

Dec 2020 - Jan 2021

- A data science project that focused mainly on web URL log data with the primary goal of detecting Cross-Site Scripting(XSS) attacks initiated by users on the URL.
- Executed web scraping techniques to gather relevant data and applied exploratory data analysis to meaningful insights.
- Implemented several functions to visualize the customer geo locations, identify most and least visited webpages, and detect XSS attack attempts on the URL.
- Engineered algorithms to detect the operating systems and browsers associated with user-initiated GET requests, contributing to enhanced understanding and security measures.

PUBLICATIONS

Deepfake Audio Detection with Neural Networks Using Audio Features

(IEEE-2022)

Stock Market Prediction with Lasso Regression Using Technical Analysis and Time Lag

(IEEE-2021)

CERTIFICATIONS

- ***Microsoft Certified: Azure Data Engineer Associate***
- ***Microsoft Certified: Azure Data Fundamentals***
- ***Microsoft Certified: Azure AI Fundamentals***
- ***Huawei Certified ICT Associate in Artificial Intelligence***
- ***Huawei Certified ICT Associate in Routing and Switching***

Nov 2023

Jun 2023

Sep 2023

Jun 2021

Aug 2020