

Numair Shaikh

✉ numairsh77@gmail.com | ☎ +91-7276060606 | 🏠 A104, Alcon Stylus, Pune, MH, India

🌐 LinkedIn | 🐙 GitHub | 🌐 Portfolio Website

EDUCATION

Pune Institute of Computer Technology

Bachelor of Engineering, Computer Engineering

GPA - 9.69/10.0

Relevant Courses Taken -

- Machine Learning
- Information & Cyber Security
- Design & Analysis of Algorithms
- Engineering Mathematics (I, II & III)
- Computer Organisation & Architecture
- Data Structures
- Object Oriented Programming
- Discrete Mathematics

Pune, India

Aug 2018 - June 2022

WORK & INTERNSHIP EXPERIENCE

eQ Technologic

Software Engineer

Aug 2022 - Present

Pune, India

- Spearheaded the design and development of Out-of-the-Box (OOTB) APIs to support easy migration and integration between proprietary systems in eQ's flagship integration product eQube-MI.
- Developed an end-to-end CAD migration pipeline between Teamcenter PLM and Windchill PLM in a successful proof-of-concept delivered to a major argo-mechanical customer.

AlgoAnalytics

Internship

Sept 2021 - Apr 2022

Pune, India

Project Intern, Natural Language Processing (NLP)

- Contributed to the development of **FABRIC**, a SaaS platform which processes the textual information collected from news, social media and internal company documents such as quarterly reports to assist, augment and automate financial analysis of listed Indian IT conglomerates.
- Developed a pipeline to extract and process text data from diverse sources, and implemented FAISS for efficient indexing, enabling rapid retrieval of processed textual, and Dense Passage Retrieval (DPR) to optimize user queries, facilitating quick access.
- Implemented a text summarization mechanism using the RoBERTa model a sentiment analysis model using FinBERT to assess the sentiment of financial text summaries.
- Worked on a Q&A system to convert natural language queries into SPARQL queries, for effective and fast querying of knowledge graphs.

Thoughtworks

Internship

June 2021 - Sept 2022

Pune, India

Intern, Application Development

- Contributed to the design and development effort design, and development of a proof of concept for the Alumni Connect Portal project at Thoughtworks.
- Contributed to the development of the frontend of the application using React.js, developing several key pages.
- Contributed to the backend development of the application, utilizing Spring Boot for efficient and scalable server-side functionality.
- Developed an authentication mechanism by integrating the application's access with Okta SSO and LinkedIn SSO, enhancing security, while allowing easy access to verified alumni.
- Developed an OTP verification service using SendGrid and integrated it into the Spring Boot backend, providing a secure multi-factor authentication for first-time users.

TECHNICAL SKILLS

Programming Languages:

C/C++, Java, Python, JavaScript

Libraries/Frameworks:

React.js, Redux.js, Node.js, Spring Boot, Flask

Machine Learning Frameworks:

Tensorflow, PyTorch, scikit-learn, Keras

Databases:

MySQL, PostgreSQL, MongoDB, GraphDB

Version Control:

Git, SubVersion (SVN)

Generating Knowledge Graphs for Explainable Artificial Intelligence and Querying them by Translating Natural Language Queries to SPARQL*First Author**Accepted for publication in the [International Journal of Information Technology \(IJIT\)](#) - Springer In-Press*

In this research work, a comprehensive workflow is presented for the generation of Knowledge Graphs for Explainable AI, along with a system to translate natural language input to SPARQL queries for easy search and retrieval.

Query-Based Summarization and Sentiment Analysis for Indian Financial Text by leveraging Dense Passage Retriever, RoBERTa, and FinBERT*First Author**Accepted and Presented as poster paper at the [20th International Conference on Natural Language Processing To be published under ICON, 2023 proceedings in ACL Anthology](#)*

In this research, a workflow for summarizing financial news and documents is discussed. Post retrieval of financial news and texts from the internet, specifically from sources/aggregators, DPR is used for document retrieval and RoBERTa to create concise and accurate summaries based on user queries. Furthermore, sentiment analysis using FinBERT is added to provide sentiment scores. Through the integration of these approaches, robust system is presented, which is capable of extracting summaries that are personalized to user queries, while additionally conducting sentiment analysis to capture the overall market mood.

[ICON 2023 - Book of Abstracts](#), please refer to paper number 2 under poster presentations.

Unmasking Disinformation: Detection of Fake News Online using Learning Techniques*Submitted to the [International Conference on Emerging Smart Computing & Informatics, IEEE](#)****Awaiting Decision***

Examined various machine learning and deep learning models utilized for the purpose of fake news detection from Kaggle Fake-News dataset. Observed that DiltilBERT, emerged as the best performer for the identification of fake news, achieving a remarkable 99.23% accuracy on testing data.

PERSONAL & ACADEMIC PROJECTS

Sentiment Analysis on Cryptocurrency News and Tweets[GitHub](#)

Fine-tuned DistilBERT for sentiment analysis on labelled tweets and news articles related to cryptocurrency news.

The model was trained on a dataset containing 82283 records and achieved an accuracy of 94.50% during testing.

Tech-Stack: Flask, Tensorflow, Keras, BERT-based models.

Project Management System for final-year project management.[GitHub](#)

This project aimed to develop a comprehensive project management application tailored for my institution (PICT). The objective was to enable students to seamlessly provide updates, reviews, and reports to their assigned internal project guides. Additionally, an automated guide allotment system, based on domain expertise and areas of interest, was implemented.

Tech-Stack: React.js & Redux (Frontend), Node.js (Backend) and MongoDB

Sentiment Analysis on Tweets

Fine-tuned and compared FinBERT and XLNet models for sentiment analysis on tweets.

Developed a web-application for sentiment analysis of real-time twitter data.

Tech-Stack: React.js, Flask, Tensorflow, Keras, BERT-based models.

Grocery Store (Mobile Application)[GitHub](#)

Created a mobile application for grocery stores on the Android platform, designed specifically for small traders and shops.

The application facilitates the listing of essential staple foods, fruits, and vegetables, providing a user-friendly and accessible solution for managing inventory and sales in the grocery sector.

Tech-Stack: Java, Android Studio, Firebase

CERTIFICATIONS

- [The Complete 2021 Web Development Bootcamp](#), Udemy, June 2021
- [Exploratory Data Analysis](#), Coursera (Authorized by Johns Hopkins University), July 2020
- [Reproducible Research](#), Coursera (Authorized by Johns Hopkins University), July 2020
- Machine Learning, Coursera, May - June 2020

CO-CURRICULARS

- Secured the Runner-Up position amongst 100+ entries for the final-year project at [Impetus & Concepts \(InC\)](#) hosted at Pune Institute of Computer Technology.
- Prepared and presented a poster for my research paper, titled *Query-Based Summarization and Sentiment Analysis for Indian Financial Text by leveraging Dense Passage Retriever, RoBERTa, and FinBERT* at the [20th International Conference on Natural Language Processing](#).

TEST SCORES

- Graduate Record Examination (GRE) - 320
Quantitative Reasoning - 168, Verbal Reasoning - 152
Analytical Writing - 4.0
Test Date - 12th Dec 2023
- International English Language Testing System (IELTS) - 8.5
Listening - 9, Reading - 8.5, Writing - 7.5, Speaking - 8
Test Date - 4th Nov 2023