Nafisa Anjum

Date of birth: 07/01/1997 | **Nationality:** Indian | **Gender:** Female | **Email address:**

nafisa.anjum@mailbox.tu-dresden.de | **Website:** https://github.com/nafisa17 | **LinkedIn:**

https://www.linkedin.com/in/nafisaanjum/ |

Address: Hans-Dankner-Straße 5A, 01069, Dresden, Germany (Home)

ABOUT ME

I am a Computer Science graduate with expertise in handling unprocessed text and curated photonic data, as well as building multivariable models for complex data analysis. My experience spans projects in cancer research and deep learning, where I developed and refined predictive models to drive insights and improve efficiency. My work has included both independent projects and collaborative academic research, emphasizing accuracy and model performance in high-dimensional data environments.

WORK EXPERIENCE

05/2024 – CURRENT Dresden
GRADUATE RESEARCH ASSISTANT- SYSTEMS ENGINNER HOLYPOLY

Designed a monitoring system using netdata

11/2023 – 07/2024 Dresden , Germany
GRADUATE RESEARCH ASSISTANT ELSE KRÖNER FRESENIUS ZENTRUM FÜR DIGITALE GESUNDHEIT

- Explored different Large language models for different medical applications
- Conducted my research project - Prediction of Important Biomarkers from H&E-stained Whole Slide Images Using Transformer-Based Multiple Instance Learning Model

15/03/2022 – 01/04/2023 Dresden, Germany
GRADUATE RESEARCH ASSISTANT FRAUNHOFER IPMS

- Exploring the effect of light intensity on a micromirror (photonic) array.
- Visualising the heatmap from accumulated measurement data to gain insight into the properties of each individual mirror.
- Data processing based on the relative movement of the micro mirrors in Z plane using Python libraries.

Website https://www.ipms.fraunhofer.de/en/Components-and-Systems/Components-and-Systems-Actuators/Optical-Actuators/spatial-light-modulators.html

EDUCATION AND TRAINING

10/2021 – CURRENT Dresden, Germany
MASTER OF SCIENCE Technische Universität Dresden

with a specialization in Computational Life Science (CLS).

Field of study Computational Modelling and Simulation (CMS)

2016 – 2020 Kolkata, India
BACHELOR OF TECHNOLOGY Maulana Abul Kalam Azad University of Technology

Thesis Efficient Celebrity Profiling in Twitter Social Network (Using Deep Learning, NLP, and multi-class classification)

PROJECTS

03/2024 – 07/2024
Prediction of Important Biomarkers from H&E-stained Whole Slide Images Using Transformer-Based Multiple Instance Learning Model

- Research Focus: Explores transformer-based multiple instance learning to predict crucial biomarkers (MSI, TMB, HRD) from H&E-stained whole slide images (WSI) in endometrial carcinoma and stomach adenocarcinoma.
- Key Findings: Demonstrates the viability of predicting biomarkers directly from H&E-stained WSIs, offering a cost-effective approach to precision medicine. Highlights the critical role of feature extractor choice, with the UNI model showing superior results in predicting MSI and TMB for stomach adenocarcinoma.
- Impact: Reinforces the potential of AI in healthcare, making biomarker prediction more accessible and refining treatment recommendations in oncology.

Supervisor : Dr. rer. nat. Narmin Ghaffari Laleh

06/2023 – 09/2023
Napari-parallel

The aim of this project is to develop a distributed system using Dask and napari packages to reduce the computational cost of large-scale image data manipulation and to create a user guide on how to apply this system for biologists. The written plugin significantly reduces the time required to perform complex computational tasks such as Voronoi tessellation, edge detection, or filtering for a big set of images. The relevance of the work lies in the need to process the digital results of the experiment in the form of graphic images.
This was an academic team project under the Bio-image Analysis Group, Physics of Life, TU Dresden under the supervision of Dr. Robert Haase
Team mates : Artem Tomilo, Himanshu Manoj Kaloni

Link https://github.com/nafisa17/napari-parallel

2019 – 2020
Efficient Celebrity Profiling in Twitter Social Network

- Applied deep learning in the context of Efficient Celebrity Profiling on Twitter to classify tweets into six domains (Business, Sports, Entertainment, Education, Technology & Politics)
- Implemented text mining techniques to extract the features from tweets, such as sentiment scores, number of mentions, and frequency of hashtags and words.
- Implemented NLP algorithms to analyze the extracted data.
- Created a data dictionary for every document that shows the number of occurrences for each word within the document.
- The tweets of an unidentified celebrity were gathered, processed to eliminate irrelevant information, and compared to a data dictionary. This comparison was used to identify the category or genre that this celebrity belonged to.

2018
SPAM /HAM classification

- The goal was to create a confusion matrix to evaluate the performance of the classifier. The precision, recall, and F1-score were 0.93, 0.96, and 0.95 respectively for all three models.
- I used the classification report to compare the performance of all three models on the test set. The results showed that Random Forest Classifier had the highest accuracy score (0.9526), followed by SVM (0.9396) and Decision Tree (0.9396).
- Overall, this project demonstrated Random Forest Classifier as the best model for text message classification in terms of accuracy and precision.

SKILLS AND CERTIFICATIONS

Technical Skills

- Python (Programming Language)
- Deep Learning
- Artificial Intelligence (AI)
- ChatGPT / Prompt engineering
- Natural Language Processing (NLP)
- Machine Learning Algorithms
- Data Science
- C/C++
- Microsoft Word/PowerPoint/Excel

Certifications

- Machine Learning-Stanford Online
- The joy of computing using python- Indian Institute of Technology, Madras (NPTEL-online platform)
- Introduction to Generative AI - Google Cloud Training Online