



Shubhankar Kumar

Master of Science
Mathematics and Scientific Computing
National Institute of Technology, Allahabad

+91-8210558907
shubhankarbittu9934@gmail.com
shubhankar.2021msc23@mnit.ac.in
[Github](#) | [HackerRank](#) | [mUnicampus](#)
[LinkedIn](#) | [Portfolio](#)

SUMMARY

Experienced AI & ML specialist with a Master's degree in Mathematics and Scientific Computing. Proficient in Python and R, with extensive experience in scikit-learn, Pytorch, and TensorFlow. Skilled in NLP, computer vision, and deep learning models. Committed to leveraging technology to solve complex problems and drive innovation.

EXPERIENCE

• eClerx

04/2024

Data Scientist(Analyst)

[Live Url](#) |

- Spearheaded projects on cutting-edge technology, including object detection, object tracking, OCR in natural scenes, image classification, and video analysis, resulting in a **30% increase** in accuracy for object detection algorithms and a **25% reduction** in processing time for OCR tasks.

•IBM

02/2023

Artificial Intelligence Intern

[Live Url](#) |

- Demonstrated proficiency with big data technologies, deep learning frameworks, and cloud computing platforms, including Spark, CUDA, TensorFlow, Keras, OpenCV, AWS resulting in **20% faster** model training times and **15% reduction** in cloud infrastructure costs.

•SORADIS GLOBTECH PRIVATE LIMITED

05/2022

Junior Web Developer

[Live Url](#)

- Implemented web technologies including HTML, CSS, JavaScript, MySQL, and Git, leveraging them to develop interactive and responsive web applications. Additionally, conducted web scraping using Python libraries such as BeautifulSoup and Selenium, utilizing CSS selectors to extract data from educational websites. Successfully gathered and organized data in XLS format, contributing to streamlined data analysis processes and enhancing decision-making capabilities.

PROJECTS

•Human Emotion Detection : Comprehensive Deep Learning Approach

01/2024

[Github Url](#) |

- * Deployed a comprehensive deep learning approach utilizing various models such as **LeNet**, **ResNet34**, **EfficientNet**, **MobileNetV2**, **ViT**, and **HuggingFace ViT**, achieving accuracies up to **89.29%**.
- * Executed benchmarking on TensorFlow and ONNX platforms, highlighting GPU-optimized models with inference times as low as **0.025s**. Specifically, the TensorFlow GPU model showcased **0.15s** inference time, **0.8s** CPU time, and a model size of **1000MB**, while the ONNX GPU model exhibited superior performance with an inference time of **0.025s** and a reduced model size of **328MB**, outperforming TensorFlow.

•PharmaOptiMix: Pharmaceutical Sales Optimization Project

12/2023

[Github Url](#) |

- * Performed pharmaceutical sales forecasting using diverse models, including **ARIMA**, **Auto ARIMA**, **Prophet**, and various **LSTM** architectures (**Vanilla**, **Stacked**, **Bi-Directional**), taking average Mean squared error (MSE).
- * Demonstrated expertise in both traditional and modern time-series analysis methods, surpassing **Naïve methods** for improved accuracy.
- * Proven track record of optimizing hyper-parameters and contributing to data-driven insights for strategic decision-making in the pharmaceutical industry. Instrumental in informing resource planning and marketing strategies with a forward-looking approach.

•PubMed 200k RCT Sequential Sentence Classification With SkimLit

11/2023

[Github Url](#) |

- * Our exploration of deep learning models for sequential sentence classification in medical abstracts, leveraging the PubMed 200k RCT dataset, yielded **promising outcomes**. The **BERT-based model** stood out with an impressive accuracy of **88%**, surpassing other architectures.

- * Notable performances were observed across various models, including the **NaiveBiase Model** (72% Accuracy), **Conv1D Model** (78% Accuracy), **Pretrained Token Embedding (Universal Sentence Embedding)** (75% Accuracy), **Conv1D Model using Character Level Embedding** (73% Accuracy), **Model with Both Token and Character Level Embedding** (76% Accuracy), and **Model with Token, Character, and Position Level Embedding** (81% Accuracy).
- * These findings emphasize the efficacy of **advanced embeddings** in capturing nuanced patterns in medical abstracts. Our success paves the way for future research, including generalization assessments and the exploration of **ensemble methods** for further improvements. We extend our gratitude to the authors of the "PubMed 200k RCT" paper, whose dataset and insights significantly contribute to advancing **NLP in the medical domain**.

•AI-driven Conversational Q&A System for ShubH Tees

10/2023

[Github Url](#) |

Designed and Managed an innovative conversational interface leveraging **Google Palm LLM**, **Hugging Face embeddings**, and **Streamlit** for seamless interaction with ShubH Tees' **MySQL** database. Integrated advanced technologies such as **Langchain framework**, **Chromadb** vector store, and **Few Shot Learning** for precise natural language query handling. Evidenced expertise in **NLU/NLG**, resulting in a feature-rich system with **Streamlit UI**, **Langchain** integration, and efficient **Chromadb** vector storage. Successfully delivered a cutting-edge Q&A solution, showcasing diverse technical skills and innovative problem-solving..

•Shubh Chatbot For Food Delivery For Restaurant Service

04/2023

[App Url](#) |

Developed a versatile restaurant services chatbot using Dialogflow, with a focus on menu information, reservation bookings, and customer support. **Menu Info:** Dishes, prices, and dietary details.**Reservations:** User-friendly booking system. **Order Tracking:** Real-time updates on food orders.**Special Offers:** Notifies about discounts. Developed a robust restaurant chatbot using Dialogflow, integrating backend systems for reservations and order processing. Recognized for enhancing user experience through intuitive features and real-time services..

EDUCATION

Degree/Certificate	Institute/Board	Year
PGP in DS, BI and AI	Aegis School of Data Science	2023-2024
Master's in Mathematics and Sc.Comp	National Institute of Technology , Allahabad	2021-2023

SKILLS

- **Programming:** Python, C/C++
- **Artificial Intelligence:** Q-Learning, Deep Q-learning, A3C
- **ML Algorithm:** Supervised, Unsupervised, Reinforcement
- **Deep Learning Algorithms:** ANN, CNN, RNN , Computer Vision
- **NLP:** Text Normalisation (Stemming and Lemmatization), Sentiment Analysis, Word Embedding , LSTM RNN ,BERT ,Transformer
- **Web Framework and Cloud Deployment:** Flask, Fastapi, Heroku , AWS* ,GCP*,Azure* , Spark ,Hadoop*
- **Tools/Frameworks:** Pandas , Keras, Numpy , Matplotlib ,Tensorflow, Pytorch, SciPy, OpenCV ,
- **Operating Systems:** Windows, Linux* * Elementary proficiency

KEY COURSES TAKEN

- **Computer Science:** Computer Programming (with Lab) , Data Structures and Algorithms (with Lab) , Database Management System , Machine Learning , Deep Learning , Artificial intelligence , Computer Vision.

ACHIEVEMENTS

- **IIT JAM 2021**, IIT JAM Entrance Qualifier conducted by IISc Bengaluru, **AIR-802** 2021
- **GATE 2022** , Secured **AIR 1007** among 0.15 million candidates appearing for the test 2022
- **13th Edition AGBA Winner 2023**, Aegis Graham Bell Awards is one of the most credible and largest innovation awards supported by the Ministry of Electronics and Information Technology, Government of India - NITI Aayog; Skill India - National Informatics Center Services Incorporated. 2023