

Mukul Kumar

Github: github.com/CS-savvy

LinkedIn: [linkedin.com/in/mukulkr](https://www.linkedin.com/in/mukulkr)

Email: mukul.kr99@gmail.com

Mobile: +91-909-7209-033

PROFILE

I am a Machine learning and Data science enthusiast with comprehensive research expertise, with 5 filed patents and authored 2 research publications. With 6 years of expertise in developing and deploying deep learning models, I have a thorough grasp of programming and computer science fundamentals.

SKILLS SUMMARY

- **Languages:** Python, C, C++, SQL, Bash
- **Algorithms:** Diffusion, GPT, Graphs, CNN, GNN, GANs, LSTM, Transformers
- **Frameworks:** PyTorch, TensorFlow, DGL, Keras, OpenCV, HuggingFace, ONNX, Scikit, SpaCy, Flask, Streamlit
- **Tools:** Docker, TensorFlow Serving, TensorRT, Triton Inference Server, GIT, MySQL
- **Platforms:** Linux, Web, Windows, MacOS, AWS, GCP
- **Soft Skills:** Leadership, Event Management, Writing, Time Management

EXPERIENCE

- **XA Group** Remote
Senior Data Scientist (Full-time) *September 2023 - present*
 - **Roles and Responsibilities:** Research, development, and maintenance of AI solutions.
 - **Task Involves:** Prepare dataset, research on latest NN architecture, conduct experiments, document experiments, and findings, and Maintain AI models.
- **Blackstraw Technologies Private Limited** Remote
Senior Data Scientist (Full-time) *March 2019 - August 2023*
 - **Roles and Responsibilities:** As a consultant, work with the client's technical team to solve any obstacles connected to data science and machine learning. In addition to my core work as a developer, I actively participate in recruiting as a technical panelist.
 - **Project - Information Extraction and decoding from receipt-like documents:** Despite numerous perspective distortions in the scanned document, multiple computer vision, and natural language processing models were applied to scanned images of receipts to properly extract all data. NLP techniques like Text classification, NER, coreference resolution, dependency parser, and Knowledge graph embedding have been used to structure the extracted data.
 - **Tech Stack Used:** Python, C++, Pycuda, PyTorch, TensorFlow, Deep Graph Library (DGL), TF-serving, TensorRT
 - **Impact:** The solution has been deployed in multiple countries. There have been 4 patents and 1 research paper published related to this solution.
 - **Project - Social Distancing Monitoring system:** Using CCTV cameras installed on the targeted property, an end-to-end system was created to identify violations of the social distancing rule.
 - **Tech Stack Used:** Python, Nvidia Deepstream, TensorRT, GStreamer, FFmpeg, Docker.
 - **Impact:** The method was implemented at a facility with over 40 CCTV cameras, and a patent application for the solution was filed in India.
- **CrowdANALYTIX Solutions Pvt Ltd** Bengaluru
Data Science Intern (Full-time) *September 2018 - March 2019*
 - **Roles and Responsibilities:** Here, I collaborated with the team to design and implement numerous deep learning models for dataX.ai (an online platform for fashion product catalog creation).
 - **Task Involves:** Preparing dataset, creating training model on Linux-based GPU servers, translating to serving format, and deploying on TensorFlow-serving.

PUBLICATION & PATENTS

- **Research Paper: Graph Attention Networks for Efficient Text Line Detection on Receipt-Layout Documents:** Paper was presented at Document Intelligence Workshop in KDD conference, Washington DC, USA - August 14, 2022
- **Research Paper: A Novel Transformer-based Approach for High Accuracy Classification of Parkinsons:** Paper was accepted on 31st July 2023 - MDPI biomimetics
- **Patent: Methods, System, Apparatus and Articles of Manufacture to Detect Lines on Documents:** Patent ID - US81259087 — Filed Jun 7, 2022
- **Patent: Automatic document content extraction and decoding:** Patent ID - US81254014 — Filed Jun 24, 2021
- **Patent: Automated Extraction Of The Purchased Items In Images Of Receipts Based On Computer Vision And Deep Learning Techniques. The Goal Is To Extract Each Of The Individual Items In The Printed Receipt Mainly Composed Of The Product Description, Quantity and Price.:** Jun 24, 2021
- **Patent: Automated Receipt Decoding In Dds. Product Matching And Dictionaries:** Jun 24, 2021
- **Patent: Method And System To Detect Distance Between Entities:** Patent ID - 202021034381 — Filed Aug 11, 2020

CERTIFICATIONS AND ACHIEVEMENTS

- Qualified GATE CSE 2018 - score: 346
- Accomplished a two-month software engineering internship at Oil and Natural Gas Corporation Ltd. in Dehradun, India. - June - July 2017
- Certificate: DeepLearning.AI TensorFlow Developer - April, 2020
- Certificate: Google Cloud Fundamentals: Core Infrastructure - Feb 2017

EDUCATION

- **COLLEGE OF ENGINEERING ROORKEE** Roorkee, India
Bachelor in technology in Computer Science & engineering *8 Aug 2014 - 25 July 2018*