KUNAL KHANDELWAL

khandelwal.kunal@outlook.com | linkedin.com/in/khandelwalkunal779/ | khandelwalkunal779.github.io/

PROFESSIONAL SUMMARY

Data Scientist & AI Engineer with experience in building Agentic AI systems, LLM-integrated solutions, and scalable data pipelines using PySpark and SQL. Skilled in OCR-based document automation, anomaly detection, and model evaluation in real-world settings.

EDUCATION

Indian Institute of Technology Roorkee (IITR) **November 2020 - July 2024**
Bachelor's, Civil Engineering *GPA: 8.186*

PROFESSIONAL EXPERIENCE

ICICI Lombard GIC Ltd **South Mumbai, Mumbai, MH, India**
Associate Data Scientist *July 2024 - Present*

- Developed an AI model evaluation framework for a speech-to-speech chatbot, benchmarking performance across metrics like ROUGE, cosine similarity, and other NLP-based scores to measure semantic and contextual deviations from ground truth.
- Designed, developed, and deployed a RESTful API to automate Form 29-30 approvals in used car motor insurance using OCR and LLM integration, enabling intelligent document processing and reducing manual intervention.
- Conducted anomaly detection on motor insurance claims using MySQL, identifying fraud-prone patterns such as unusually high repair costs for low-value vehicle models, aiding risk mitigation.
- Performed cross-source data aggregation and analysis of private car insurance quotes to evaluate broker performance across regions, delivering insights that informed strategic decisions on market positioning.
- Built a scalable market share estimation pipeline using PySpark on Databricks, powered by structured web data extracted via a custom Playwright-based web crawler, enabling data-driven decision-making in the motor insurance domain.
- Optimized web crawler performance with batch scripting and multiprocessing, achieving end-to-end automation and reducing runtime from 11 hours to 40 minutes, significantly improving system throughput and efficiency.
- Created an automated email workflow in Databricks using Microsoft Graph API, responding to 400+ vendor emails daily for NCB confirmation by extracting policy details from email content and querying a dynamic data mart.

The University of Tokyo **Hongo, Bunkyo, Tokyo, Japan**
Research Intern *May 2023 - July 2023*

- Predicted the demand for CFRTP by 2050 in the Wind Energy Sector through rigorous Statistical Models and Data Analysis.
- Led a collaborative effort to develop a Python-based algorithm that predicts the Tensile Strength of recycled CFRP materials, showcasing Programming and Algorithm Design Skills while advancing Sustainable Materials Research.

PROJECTS & OUTSIDE EXPERIENCE

Deep Learning based Prediction of Collision at Highway Intersection **Roorkee, UT, India**
TEG, IIT Roorkee *December 2022 - March 2023*

- Developed a model that could analyze driver behaviour with respect to pedestrian safety at pedestrian crossings.
- SSMs based on various vehicle dynamics data, such as speed, acceleration, and lateral position with respect to the lane, were used to identify potential safety issues and classify a driver's behaviour as risky or safe.
- The LSTM algorithm was used to model the behaviour of the vehicle based on the time-series trajectory data.

Artificial Intelligence based Photodetection **Roorkee, UT, India**
Team Robocon, IIT Roorkee *September 2021 - October 2021*

- Devised a method to define the path of a moving rover from a live camera feed using the OpenCV library in Python.

SKILLS

Core Competencies: Data Science, Machine Learning, Deep Learning, Natural Language Processing (NLP), LLM, Agentic AI, OpenCV, Data Analysis, Statistics, Mathematics
Technical Tools & Frameworks: Tensorflow, Keras, Pandas, NumPy, Jupyter, Power BI, Postman, GitHub
Cloud & Platforms: Microsoft Azure, Databricks, Apache Spark, Azure Functions, Azure Blob Storage, MySQL, RDBMS
Programming & Systems: Python, C/C++, HTML/CSS, Bash, Data Structures & Algorithms, Operating Systems, Computer Networking, REST APIs