Kanha Bansal

kanhab97@gmail.com | +44 7442233219 | Glasgow | www.linkedin.com/in/kanhab97 | https://github.com/kanhab97

EDUCATION

M.Sc. Data Science

University of Glasgow

Sep 2022 - Present | Glasgow, UK

Coursework

- Machine Learning
- Natural Language Processing
- Deep Learning
- Big Data
- Web Science
- Data Visualisation
- Cyber Security

B.Tech in IT

CET Bhubaneswar

Aug 2015 - Jul 2019 | India

CGPA - 8.33/10 | GRADE 1

Coursework

- Data Structure and Algorithm
- Object-Oriented Programming
- Databases
- Software Engineering

SKILLS

Programming Language

- Python
- Java
- HTML/CSS
- SQL

Tools and Libraries

- Matplotlib
- NumPy
- PyTorch
- Scikit-Learn
- NLTK
- Pandas
- Django
- Tensorflow
- Keras
- Apache Spark
- Spacy
- GitHub
- Tableau
- Postman
- MySQL Workbench
- JIRA
- VersionOne

ACHIEVEMENTS

- **Excellence Award |** 2020-2022
  Awarded three times in a team of 50 for Innovative ideas by Deloitte
- **Merit Scholarship |** 2016-2019
  Granted merit scholarship 4 years in a row for being in top 6 in a class of 60 students

WORK EXPERIENCE

Business Technical Analyst

Deloitte Consulting

Jan 2020 - Aug 2022 — Hyderabad, India

- Developed an **LSTM-based NER model** to automatically recognise and extract relevant entities from the legacy data, such as customer names, addresses, policy numbers, claim details and many other fields from the fact sheets documents that helped **reduce manual efforts by 80%.**
- Developed and implemented a **CNN model** capable of recognising and extracting key details from scanned documents, such as insurance forms and invoices. This ML-based approach drastically reduced manual efforts, enabling faster and more accurate data migration.
- Implemented a classification model to automate data validation in the QA environment. By selectively targeting the groups classified as incorrect by the model, we achieved a **20% increase in the success rate** for production data migration. Also, there was a **40% reduction in manual efforts** as the groups predicted as correct didn't need additional validation.

Technology & Evolution Intern - JAVA Developer

Reliance Jio Info Comm Ltd

May 2017 - Jul 2017 — Mumbai, India

- Developed a Simulator using IOT protocols (LWM2M and MQTT) that connected to a remote device and got requisite geographical data like wind direction, and position of the device to generate an alarm when it crossed the prescribed threshold for different parameters.

PERSONAL PROJECTS

Deep Learning-based Classification of Colon Cancer Cell Nuclei

- Developed two deep neural network models with **98.5% accuracy for classifying cell nuclei in colon biopsy** images into different cell types. Utilised custom ConvNet and transfer learning with pre-trained torch-vision models, optimised hyper-parameters with Ray Tune, analysed performance metrics, and submitted accurate predictions to Kaggle competition

Semantic Discernment of UK Twitter: Unveiling Topics through Short Text Analysis and Grouped Data Evaluation

- Analysed a self-crawled dataset of 10,000 UK tweets using **topic modelling techniques** to uncover themes. Explored challenges in short text topic modelling and evaluated the impact of clustering grouped data. Assessed topic quality, compared approaches, and addressed tweet-specific topic formation issues.

Emotion Analysis Using NLP with Ml and DL techniques

- Employed k-means clustering to unveil underlying patterns, evaluated and compared multiple ML classifiers using Scikit Learn, fine-tuned parameters for optimal performance, and utilised **deep learning techniques with BERT-based models** to accurately analyse and classify emotions in text. Delivered valuable insights into the field of **emotion analysis** through rigorous experimentation and analysis.

Big Data Analysis and Pipeline Development using Apache Spark and Java

- Designed and implemented a batch-based text search and **filtering pipeline in Apache Spark using Java**, handling large sets of text documents and user-defined queries, while ensuring efficient processing and ranking of relevant documents based on the **DPH scoring model.**

Electric Vehicle Share System (Django, MySQL, HTML/CSS/JavaScript)

- Designed and developed a fully functional electric vehicle share system using **Django , Python, MySQL database, and a combination of HTML, CSS, JavaScript, and Bootstrap** for the frontend. Implemented features for vehicle reservation, rental, return, payment, and reporting, while enabling operators to manage vehicle tracking, charging, repairs, and inventory maintenance.