Tushar Chauhan

Data Scientist, Accenture (May'22-Present)
Gurugram, India

Mobile : +91 7755058077
Email : chauhant08@gmail.com

PROFESSIONAL SUMMARY

I am a Data Science professional with over four years of hands-on experience in the industry. My expertise lies in a wide array of ML and DL tasks, with a particular focus on supervised techniques. Over the course of my career, I have successfully tackled challenging projects encompassing diverse domains. Now, I am enthusiastic about furthering my knowledge in the cutting-edge techniques of the field, particularly in the realm of Natural Language Processing (NLP) and model deployment

EDUCATION

B.Tech - M.Tech dual degree in Chemical Engineering *(August'14-June'19)*
Indian Institute of Technology Kanpur

WORK EXPERIENCE

Data Scientist, Accenture India Private Limited *(May'22 - Present)*

Consumer goods recommendation system

- Collaborated with a 4-member team to design and implement an **AI-driven product recommendation system**, aiming to reduce manual effort and enhance user experience
- Analyzed sales patterns, identified **top-selling products**, and performed data preprocessing
- Engineered features, incorporating sales, quantity, frequency, recency, seasonality, geography, and promotion aspects for each retailer's products
- Developed a **Random Forest multi-output classifier** to predict the probability of being sold in the market
- Implemented **sales pattern-based strategies** and **business rules** tailored to meet market requirements

Data Scientist, FN Mathlogic Consulting Services Limited *(July'19-May'22)*

Iron ore detection

- Led and collaborated with a 2-member team to develop a **computer vision architecture** for extracting iron ores from **high-resolution satellite imagery**
- Gained comprehension of fundamental concepts of **TFRecords** and utilized them to optimize computations
- Conducted research and implemented various image segmentation techniques to achieve desired results
- Utilized transfer learning from pre-trained models like **RESNET, DENSENET, VGG, and MOBILENET**
- Implemented the **U-NET architecture** to perform accurate segmentation of iron-ores in the satellite imagery

Credit risk scorecard

- Developed a **credit risk detection model** and capture diverse purchasing patterns of various customer segments
- Created **XGBoost and Neural Network** models using customer's historical transaction data and credit history
- Engineered features at both customer level and transaction level to enhance good/bad capturing model capacity
- Worked with **databrick clusters** and **azure databricks** environment
- Fine tuned the model's **hyper-parameters** using **hyper-optimisation**

Heading Identification using NLP

- Developed a **NLP Model** to obtain proper heading **hierarchical structure** in research papers
- Extracted data from **HTML documents** and transformed into **structured data**
- Designed and implemented **Multi-Input NLP model** to process **textual** and **Numerical** data simultaneously to obtain more information
- Analyzed the **error cases** and based on the same created **lag/lead features** to enhance the model performance
- Created **subjective rules** on top of the model for better **generalization** of model outcome

TECHNICAL SKILLS

Python | PySpark | Gen AI | Computer Vision | NLP | MS Azure Databricks | MS Azure ML | Google Cloud Storage | Auto ML (GCP) | Vertex AI (GCP) | Dataproc Jupyter (GCP) | Google Big Query (GCP) | Pandas | Numpy | Scikit-Learn | Tensorflow | Keras | Pytorch | MS Excel

CERTIFICATIONS

Generative AI with Large Language Models
DeepLearning.ai, Issued February 2024