Rishi Sharma
AI / ML / DS

Contact: - +91 8860551924
rishisharma2628@gmail.com

BRIEF

Data science professional with 6+ years of experience in predictive analytics, consulting and data modelling using different supervised and unsupervised methodologies, using python and tensorflow. Versed with handling clients from various geographies and domains to drive better decision-making, provide meaningful insights hidden in data.

SKILLS (CONSULTING)

- Proficient at problem analysis and solving
- Conveying ideas with good communication skills
- Independent learner with strong team-work ethics
- Competitive attitude and problem solving
- Confident, articulate, professional speaking abilities

SKILLS (TECHNICAL)

- **PROGRAMMING**: - Python, SQL, Pyspark, HTML, Git, Bash
- **MACHINE LEARNING:** - Predictive modeling, Regression, Random Forest, XGBoost, Recommendation systems, Neural Networks, Deep Learning, LLM, Computer Vision, Natural Language Processing (NLP), Optical Character Recognition (OCR), Object Detection/Segmentation, KMeans Clustering
- **DATABASE**: - Postgres, Apache Hive, AWS DynamoDB
- **CLOUD** :- Amazon Web Services (AWS), Google Cloud (GCP), Docker
- **LLM :-** Langchain, OpenAI, GPT, GPT4, Davinci, Prompt Engineering

EXPERIENCE

Machine Learning Consultant | Deloitte, Gurgaon (2.5+ years)

Project 1 – LLM Based Audit Summarizer

- Created comprehensive audit summaries for client files, enhancing data findings and summary for reports.
- Created the backed for the online app by making use of python, FastAPI framework and AsyncIO
- Leveraging techniques like map reduce and vectorization in Langchain for faster operations and reduced latency.
- Used Langchain and OpenAI models (e.g., GPT-3.5, GPT-4, Davinci) for task involving summarization and categorization.

Project 2 – Chabot with Integrated Recommendation System

- Created a voice enabled chatbot on AWS Lex and Poly to take orders from customers in drive through, for a global restaurant chain
- Designing & implementing the flow for the chatbot including providing the utterances for training, creating the intent slots & managing edge cases
- Training a recommendation system to suggest menu items to customers for faster order processing and better user experience
- Recommendations from model led to a 7% increase in average order value, helping the client provide better service to its customers.

Project 3 – Credit Risk Modelling – Loan Defaults

- Created a machine learning models to perform classification on high risk credit defaulters with no loan history.
- Making use of user data such as bank transactions, employment details, assets under user, etc to create the data model.
- Making use of user specific data such as age, income, occupation, collateral, etc. to predict metrics such as PD, LGD and EAD for each user.
- Modelling led to a reduction in default rate by 3% saving a total of potential $50K for client across 7 bank branches.

Data Scientist | Optum (UnitedHealth Group), Gurgaon/ Bangalore (2 Years)

Project 1– NLP Based Service Navigation

- Creating machine learning models to perform classification for the user submitted UHG insurance claims and paperwork.
- Generating word embeddings using libraries such as gensim word2vec, while utilising transformers, spacy, flair, Huggingface - GPT, BERT
- Segmenting user documents using OpenCV and converting image to text using OCR Engine like Tesseract and Amazon Textract.
- Integrating the model in production pipeline over AWS. Cutting down operational cost by $ 100K dollars across 3 claim centres.

Project 2– Insurance Recommendation Model

- Building Machine Learning models for health care providers to shortlist customers for better sales & lower customer acquisition cost.
- Analysing RX and Medical claims history of members, from hive database along with other related details like Demographics, Online activity, AWV and other factors that could aid in the improvement of acceptance rate.
- Creating classification models based on the above user data. Recommendation from model led to 4% volume uplift. Received recognition from client's end.

Associate Software Engineer | Accenture, Gurgaon (1.5 Years)

Project 1– Automating data pipelines

- Creating ETL pipelines to store and update client's data across data lakes using creating cron jobs
- Generating word embeddings using libraries such as gensim word2vec, while utilising transformers, spacy, flair, Huggingface - GPT, BERT
- Segmenting user documents using OpenCV and converting image to text using OCR Engine like Tesseract and Amazon Textract.
- Integrating the model in production pipeline over AWS. Cutting down operational cost by $ 100K dollars across 3 claim centres.

EDUCATION

1) **Maharaja Agrasen Institute of Technology, Delhi** — **Aug 2013 – June 2017**
   B. Tech (Electronics and Communication) — **(Aggregate Percentage – 74%)**
2) **St Mary's School, Safdarjung Enclave, Delhi** — **Apr 2011 – Mar 2013**
1) 12th (PCM + Computer Science) — **(Aggregate Percentage – 91%)**