

Ritika Gupta

Data Scientist

Results-driven business analytics graduate passionate about solving complex business problems using data, analytics, and data visualizations. Collaborative team player with a fast-learning mindset, eager to implement new algorithms and technologies to drive data-driven insights and contribute to organizational success.

WORK EXPERIENCE

Tata Consultancy Services

Systems Engineer | March 2017 – July 2021

Project – NEST Pensions (BFSI Domain)

- Played a pivotal role as a development team member, proficiently implementing new requirements using Java, PL/SQL, and Python, while expertly resolving defects and production issues. Actively engaged in design workshops and effectively prepared build plans for streamlined project execution.
- Actively learnt and implemented a fund recommendation system for our client and completed a successful POC which eventually got deployed. (Using Clustering algorithms such as K-Means, DBScan, Hierarchical clustering etc.)
- Proficient in SQL, adeptly wrote numerous packages and procedures from scratch to effectively handle intricate business logic and address unique edge cases, successfully resolving historical defects.
- Lead team of 3 members to successfully develop a new Complaint Management Tool (CMS).
- Developed and integrated a text similarity-based complaint ticket resolution system in CMS using NLP libraries (such as word2vec, NLTK etc.), resulting in a 20% decrease in average complaint resolution time.
- Created prototype of interactive reporting dashboards in both Tableau & PowerBI for CMS, enabling tracking of KPIs related to the Complaint Management System (CMS)
- Awarded Best Team Player Award for successful and bug free production delivery .

Digite Infotech Limited

Data Science Consultant - Intern | Jan 2022 – Jun 2022

Project – kAIron – Contextual AI Chatbot

- Actively contributed to the enhancement of their new "kAIron" product, a context-aware AI deployable chatbot based on the RASA framework, by implementing advanced features and functionalities .
- Carried out a POC on multilingual language models and implemented a new feature for handling multiple language translation for kAIron using cross lingual embeddings transformers and Google Translate api service.
- Accelerated the initial product marketing campaign by creating and delivering a range of business use case templates, leveraging tools like Airtable integrations and Google API integrations

ritika.gg1112@gmail.com 

Mobile - 8240688805

Kolkata, IN

[LinkedIn](#)

[GitHub](#)

[Kaggle](#)

SKILLS

- **Languages** – Python , Core Java, R, PL/SQL
- **Framework/Lib** – NumPy, Pandas, Scikit-learn ,Tensorflow, Keras, Matplotlib, Seaborn, Plotly
- **ML** – Regression, Classification, Clustering, Dimensionality Reduction, Ensemble Technique, Feature Selection
- **NLP** – NLP Tasks(Word Embeddings : NLTK, TFIDF, Word2Vec, Fasttext , Spacy , Gensim , DistilBERT, Classification ,NER, Sentiment Analysis), Chatbot Frameworks(RASA, Dialogflow), Transformers
- **DL** – CNN, ANN
- **Visualizations** – PowerBI ,Tableau
- **IDE** – Pycharm, Jupyter Notebooks, Google Colab, Eclipse, R Studio
- **Deployment**– Streamlit, Flask APIs, Fast API, Docker basics, MongoDB, Gitlab, Git

EDUCATION

MBA (Business Analytics & Finance)

CGPA-9.31

NIT, Tiruchirappalli

Aug 2021 - May 2023

- Awarded for Rank 1

B.TECH (Electronics & Communication Engg.)

CGPA-9.24

WBUT, West Bengal

May 2012 - May 2016

- Awarded for Department Topper for securing highest CGPA across 4 yrs

PROJECTS

1. Bankruptcy Prediction by integrating Instance Hardness based Under sampling and Supervised Learning Methods in highly Imbalanced dataset

[Github](#)

- Used the instance hardness (IHT) based under-sampling method in highly imbalanced dataset for classification using supervised learning method such as SVM, XGboost, Isolation Forest, Balanced Random Forest, AdaBoost with Random Under Sampling(RUS) and Easy Ensemble
- Combined various feature selection techniques to reduce the dimensionality of dataset.
- Average balanced accuracy of all 6 models increased from 75% to 77% after applying IHT under-sampling.

2. Real-Time Tweet Sentiment Analysis

[Github](#)

- This app extracts tweets from twitter apis based on input hashtag ,keyword or twitter user handle . It shows sentiment distribution of tweets and highest liked tweet , RT . Top twitter handles based on tweets, user mentions etc.
- It also displays word cloud for corresponding sentiment.
- It creates wordcloud for each of the three sentiments based on selected sentiment (positive, neutral, negative)
- It was deployed on Heroku

3. Portfolio Optimization to Diversify Risk using Unsupervised technique – Hierarchical Risk Parity(HRP)

[Github](#)

- Optimizing the allocation of assets based on risk parity principles, HRP outperforms the traditional Markowitz minimum variance(MVP) approach in terms of diversification and risk-adjusted returns.
- Implemented and compared the results of MVP vs HRP on Nifty 500 stock prices across the period of 2020-2023. Using metrics such as Sharpe Ratio and Std. Deviations.

4. Integration of Self-Organizing Maps(SOM) & Probabilistic Neural Network(PNN) for Bankruptcy Detection

(Research Paper under review) | [Github](#)

- Using an integrated model of unsupervised cluster SOM and combining it with supervised classifier PNN to predict bankruptcy of firms using various financial features as input.
- Explored various combinations of Kernels(RBF, Laplacian, Gaussian, Polynomial) and lattices(Hexagonal, Rectangle,Square) to understand which combination gives better results using a variety of classification metrics.

Higher Secondary(10+2)
CICSE *Percentage-84.71%*
R.V.M Girl's School, Kolkata
2011 - 2012

Secondary(10)
CICSE *Percentage-85.71%*
R.V.M Girl's School, Kolkata
2009 - 2010

CERTIFICATIONS

- Business Analytics and Data Mining Modeling Using R Part I &II [Cleared two proctored exams.](*NPTEL IIT Roorkee*)
- AWS Certified Cloud Practitioner
- DeepLearning.AI Tensorflow Developer Professional Certificate (*Coursera*)
- Kaggle Notebook Expert
- TCS Certified Power BI & Tableau Intermediate Develop