Swapnil Ranshing
Data Scientist, Western Union, Pune, India
Portfolio Website: [Swapnil's Portfolio Website](#)

☎: +91- 836 974 6859
E-Mail: swapnilransing001@gmail.com
LinkedIn Profile: [Swapnil's LinkedIn Profile](#)

Summary

- 7.5 years of expertise in machine learning, deep learning, natural language processing (NLP), GenAI, model development and deployment for fraud and credit risk detection and payment fraud risk strategy development
- Postgraduate degree from IIT Bombay and [(Elite Silver) certificate for Data Science for Engineers from IIT Madras](#) and [certificate for Deep Learning specialization](#) from DeepLearning.AI and Coursera

Skills Summary

- **Machine Learning and Deep Learning Models:** Supervised and Unsupervised Learning, Ensemble Methods, Decision Trees, Random Forest, XGBoost, Deep Learning, Neural Networks, CNN, RNN, optimization algorithms
- **Advanced Models:** Large Language Models (LLM), Generative AI
- **Languages:** Python, SQL
- **Data Analysis:** Data Wrangling, Feature Engineering, Data Visualization, Statistical Analysis
- **Frameworks:** Sci-kit Learn, TensorFlow, Keras
- **Tools:** AWS SageMaker, Power BI, Git, GitHub

Work Experience

Western Union Technology Engineering Center- Pune   (Apr 2022- Present)

- Data Scientist, Digital Risk Decision Sciences
- **Responsibilities:** Machine learning model and strategy development for Fraudulent transaction detection
- **Projects:**
  - **Machine Learning Model and Risk Strategy Development for Fraudulent Transaction Detection**
    I. Developed and deployed **10 models** on highly imbalanced fraud data for various customer segments
    II. **Featurization and EDA**: Created around **700 features** consisting of velocity, age, count, decay, categorical and rating variables from large dataset. EDA is performed to understand the features correlation with target.
    III. **Feature Selection**: Small set of important features selected using weight of evidence and Information value, forward feature selection, correlation coefficient and vote counts from different ml models feature importance.
    IV. **Modeling**: Tuned hyperparameter of machine learning and deep learning algorithms and selected best performing model using AWS sagemaker studio processing job.
    V. **Risk Strategy Development and Deployment**: Risk scorecard developed using model probability distribution and risk optimization strategies are developed and productionized to improve risk metrics.
    VI. **Results**: Challenger models gave 30% additional fraud capture rate than champion at 20% operating range, 20% improved PR AUC and 3% lift in customer acquisition with reduced chargebacks in production.

  - **NLP and Deep Learning Model Development**
    I. Developed NLP application to effectively assist manual review team for transaction review decision.
    II. Developed **auto encoder** anomaly detection model which increased new customer acquisition by 10%.

  - **Data Analysis, Dashboard Development for portfolio improvement**
    I. **Developed 12 Power Bi dashboards** which resulted in quicker data analysis, improved decision for customer acquisition campaigns, KPI/model monitoring, fraud pattern, processor change, product updates etc.
    II. Improved customer approval by 200 bps and reduced chargebakcs by 20 bps by implementing risk strategies.

  - **Clustering model - Customer Segmentation** (Winner of the Innovation Contest, DRDS, WU – 2023)
    I. Developed 8 clustering models on entire digital transaction data which defined high fraud capture segments
    II. Results shown 18 bps faster funds payout, 3% additional fraud capture and 20 bps reduced manual review

Danfoss India Technical Center- Pune (Service Continuity)   (Aug 2021- Apr 2022)
Eaton India Innovation Center- Pune   (Aug 2017- Jul 2021)
- Senior Engineer, Center for digital prototyping and twins
o **Projects:**
  - **Regression Machine Learning Model for Coupled Electric Motor Pump Product**
    Optimized power loss and noise using regression algorithm and achieved <u>0.94 coefficient of determination value</u>
  - **Excavator System Design** (Article: Excavator Inauguration)
    Estimated models output within six sigma by tuning control parameters and <u>achieved 90% lab test conformity</u>.

Portfolio Projects [(Portfolio Link)](Portfolio Link)

- **Predicting Clients Loan Repayment Ability** [(GitHub Link)](GitHub Link) [(Kaggle Problem Statement Link)](Kaggle Problem Statement Link)
  I. **EDA and Featurization**: EDA is performed to understand variables correlation with target. Created around <u>1770 features</u> consisting of velocity, age, count, decay, categorical variables from provided datasets.
  II. **Feature Selection**: Small set of important features selected using weight of evidence and Information value, forward feature selection and count of votes from different ml models feature importance.
  III. **Modeling and Results**: Hyperparameter tuned machine learning and deep learning algorithms. Achieved <u>76% ROC AUC</u> with LightGBM model which captured 53% defaulters at 20% operating range.
  IV. **Risk Strategy Scorecard Development**: Risk scorecard developed using PDO calibration method.

- **NLP - Question Pair Similarity** [(GitHub Link)](GitHub Link) [(Kaggle Problem Statement)](Kaggle Problem Statement)
  I. Features consisting of TFIDF weighted WORD2VEC, word count arithmetic and fuzzy logic were created.
  II. **Modeling and Results**: Hyperparameter tuned machine learning and deep learning algorithms. Achieved accuracy= 75.39%, F1-score= 0.67 and log loss=8.87 with XGBoost model.

- **NLP - Multiclass Classification** [(GitHub Link)](GitHub Link) [(Kaggle Problem Statement)](Kaggle Problem Statement)
  I. Malware byte and asm files featurized using <u>unigrams, bigrams and image intensity.</u>
  II. **Modeling and Results**: Hyperparameter tuned machine learning and deep learning algorithms. Achieved <u>0.014719 oot log loss</u> and 0.368% misclassification with XGBoost model.

- **Chatbot Development Using Langchain** [(GitHub Link)](GitHub Link) [(Chatbot link)](Chatbot link)
  I. Developed LLM project to answer EPFO questions using google palm, huggingface embedding, FAISS database and prompt template.
  II. A question asked to the retrieval chain finds the similar questions from the vector database and produces the answer

- **Time Series Forecasting Using LSTM** – Stock Price Prediction [(GitHub Link)](GitHub Link)
  I. Created features using technical indicators and generated time series data using Keras time series generator.
  II. Trained data using sequential LSTM model and predicted on test data with 16.9 MAE.

Education Qualification

Examination	University/Board	Institute	Year	CPI/ %
M.Tech. (Mechanical Engg.)	IIT Bombay	IIT Bombay, India	2017	9.31
B.Tech. (Mechanical Engg.)	Pune University	VIT, Pune, India	2013	9.1

Certifications

📖Deep Learning Specialization   📖Data Science for Engineers   📖Programming, Data Structure and Algorithms
📖Amazon Web Services Machine Learning Essential Training   📖Learning Amazon SageMaker
📖Power BI Data Modeling with DAX   📖Power BI A-Z: Hands-On Power BI For Data Science!   📖DFSS GB