



---

**EDUCATION****University of Cincinnati, Carl H. Lindner College of Business, Cincinnati, Ohio**

August 2020

*Master of Science in Business Analytics*

GPA: 3.96/4

**The LNM Institute of Information Technology, Jaipur, India**

May 2014

Bachelor of Technology, Electronics and Communication

---

**EXPERIENCE****FedEx Services** [Tech stack: Data Mining, PYSPARK, ML, Databricks, BTYD, LGBM, MLflow, ADF, Terraforms, Agile]**Sr Data Scientist – Marketing**

Dec 2022 – April 25

**Key Projects Include:***Proactive Retention Model*

- Enhanced BTYD code by building covariates expansion extension to improve model outputs
- Used MLOps using ADF, and terraforms to build pipelines in production for the model to run at a regular cadence automatically
- Added MLflow experiments to track model's performance (precision), and revenue and volume trends
- Usecase resulted in a net increment in revenue of more than 150M in 2 years.

*Customer Lifetime Model*

- Built the New customer model for accounts created in the last 6 months
- Built the entire product end-to-end (including pipelines, Terraform)
- Added Triggers and alarms to ensure to track model's performances and created a dashboard to print the same every month.

**Circle K** [Tech stack: Data Mining, PYSPARK, ML, Databricks, Time Series, Clustering, UCM, xgBoost, Collaborative filtering]**Data Scientist**

Oct 2020-Nov 22

**Key Projects Include:***Localized Pricing Model*

- Developed elasticity and time series UCM model to implement localized pricing project across 21 Business Units
- Developed nonlinear optimization process for product pricing
- Created metrics to measure performance of the process based on expected vs observed lifts

*Assortment Model*

- Developed recommendation engine to specify what each store should add and drop from its shelves

**AIG Analytics and Services** [Tech stack: Data Mining, MSSQL, PYSPARK, ML, Random Forest, LGB, SVM, TF-IDF, Python, Snowflake,

Tableau, KPI's, Stakeholder's Reporting, A/B testing]

**Business Analyst (eq Data Scientist)**

July 2015 - October 2018

**Key Projects Include:***Policy Churn Model*

- Analyzing the changing attributes in a policy to figure out the features impacting the churn of a policy.
- Implemented an SVM model to predict the probability of a policy cancellation prior to the expiration date by applying TF-IDF on the various attributes of an insurance policy and analyzing the reasons for cancellation
- The final model had an accuracy, precision, recall and F-score all exceeding 95%

*Propensity Model*

- Identification of the potential customers having the highest probability to be converted using a gradient boosting model
- Using the same model to identify the most significant factors leading to customer adoption of a specified insurance product and performing A/B testing to prove the same
- Analyzed the future strategy based on the model insights resulting in AIG cutting down on marketing cost by 10%

*Fraud Detection Model*

- Identified fraudulent transaction patterns using Random Forest algorithm by training on historical claims data
- Added value to the model by reducing false positive rate, and integrating the data from the global CLUE search DB
- Helped AIG to increase the efficiency of the resource allocation in the Claims Department by 20%

---

**PROJECTS****Agentic Multimodal RAG(arXiv Corpus)** [Tech stack: Embeddings, Vector DBs, LangGraph, RAG, Guardrails, eval Harness, Langsmith, FastAPI, streamlit]

- Built an Agentic RAG system over 100 arXiv papers(text+figures, 5 topic domains), encompassing web search agent capabilities.
- Built LLM-as-judge metrics(context precision-recall, Faithfulness, relevancy, and answer correctness), validated via Cohen's k against hand labels.
- Performed a component-by-component ablation study (retrieval → rerank → agentic fanout → guardrails) with cost/latency tradeoffs.

**Optical Digit Recognition** [Tech stack: Computer Vision, SVM, CNN, Python, Keras, tkinter]

- Develop a Handwritten Digits classification Model to classify a given image of a handwritten digit into one of 10 classes representing integer values from 0 to 9, inclusively
- Approach – Achieved 99.42% accuracy by performing Hyper-parameter selection and tuning for Dense Networks on CNN architecture, and then created an API using tkinter in Python for the same