**ANN MARIA CHIRACKAL GEORGE**

NY | Green Card Holder | [annchirackal@gmail.com](mailto:annchirackal@gmail.com) | Mobile: +1 267-251-5974 | [LinkedIn](http://www.linkedin.com/in/annchirackal) | [GitHub](https://github.com/annchirackal)

**DATA SCIENTIST PROFILE SUMMARY**

* Data Scientist with over **4 years of experience** applying **Machine Learning, Deep Learning, and NLP** to solve complex problems and conducting **A/B testing** for data-driven decision-making. Additionally, bringing **8+ years of experience in the tech industry- Data Engineer and Data Analyst**, with a strong background in **Supply Chain Analytics** and business optimization.
* Proficient in **Python, SQL,** and TensorFlow, with hands-on expertise in data mining, and uncovering hidden patterns. Skilled in **exploratory data analysis (EDA), statistical modeling, and predictive analytics** to derive meaningful insights and drive innovation.
* Adept at **developing compelling data stories**, answering open-ended business questions, and translating data-driven insights into **actionable strategies**. Strong ability to **collaborate with stakeholders and executives**, ensuring alignment between technical solutions and business goals.
* Looking for data science , applied scientist, gen ai roles around .

**SKILLS**

* **Programming Language:** Python, R, SQL, SAS
* **IDEs:** Visual Studio Code, PyCharm, Jupyter Notebook
* **Statistical Methods:** Hypothesis Testing, ANOVA, Time Series,STL Decomposition
* **Machine Learning:** Regression Analysis, Bayesian Method, Decision Tree, Random Forests, Support Vector Machine, Neural Network, Sentiment Analysis, K-Means Clustering, KNN, Classification, SVM, Naive Bayes, Natural Language Processing (NLP), LLM, CNN, XGBoost
* **Packages:** NumPy, Pandas, Matplotlib, SciPy, Polars, Scikit-Learn, PyTorch, TensorFlow, Keras, Spark, Langchain
* **Visualization/Other Tools:** Tableau, Power BI, Jira, Microsoft Excel
* **Cloud Technologies:** AWS, Azure
* **Database:** MySQL, SQL Server, Oracle
* **ELT/ETL Tools**: DBT, SAP HANA, SAP BW
* **Data Analytic Skills:** Data Cleaning, Data Wrangling, EDA, Data Visualization, Predictive Analytics, Pattern Recognition, Data Integrity, Quantitative Data, Data Science, Statistical Analysis, Data Analytics, Data Modeling, Big Query, Snowflake, Data Analysis, Data Mining,Mathematics,Gen AI
* **Soft Skills** : Critical Thinking, Communication Skills, Presentation Skills, Problem-solving, Decision Making.

**PROFESSIONAL EXPERIENCE**

**Consultant Senior Data Scientist Feb 2023 - present**

**SDG Group, Clients: Colgate-Palmolive and Classic Collision**

* Developed a machine learning model using time series analysis, employing the FB Prophet model for predicting independent variables and a decision tree regressor for forecasting future sales, achieving 75-80% forecast accuracy for the next 9 months. This empowered the supply chain team with actionable insights, significantly enhancing inventory planning and operational efficiency.
* Developed a price elasticity model using linear regression coefficients to assess the effects of price changes, promotions, and competitor pricing on sales, achieving a 10% improvement in the current pricing strategy and identifying customers highly sensitive to price changes.
* Executed multiple A/B tests, leading to a 3-4% revenue increase, demonstrating strong proficiency in experimental design and its impact on financial performance.
* Built and deployed a Data Quality Framework using dbt’s built-in and custom tests, reducing data inconsistencies by 40% and enhancing decision-making accuracy across business units.
* Optimized vehicle repair estimation recommendations using a hybrid AI approach by combining Apriori-based rule mining with XGBoost classification, identifying frequent insurer adjustments and predicting likely modifications before submission. Reduced negotiation loops by 40%, accelerated approval times, and improved insurer relationships by proactively aligning estimates with insurer expectations.
* Engineered an end-to-end accounting and finance data pipeline using DBT, SQL, and Python to extract, process, and validate financial data. Applied business logic, stored intermediate tables in Snowflake for verification, and seamlessly integrated the processed data into NetSuite via API calls. Improved data accuracy, automated financial reconciliation, and optimized accounting workflows
* Built observability dashboards to monitor model performance, ensuring accuracy, stability, and reliability. Tracked key metrics, monitored data quality ,identified data anomalies, and provided actionable insights to improve model effectiveness and maintain optimal performance.
* Developed a testing architecture to validate pipeline enhancements, ensuring reliability and stability. Enabled seamless deployment of improvements with confidence by detecting issues early and maintaining data integrity.

**Graduate Teaching Assistant Jan 2022 - Dec 2022**

**New Jersey Institute of Technology**

* Mentored ~150 students in statistics, machine learning, and deep learning coursework. Graded student assignments, hosted tutoring office hours, and provided additional administrative help to teaching professors.

**Data Analyst Team Lead May 2014 - Aug 2021**

**Accenture, Client:Hostess Brands**

* Led a cross-functional team of 3 consultants to automate operational reports via email, providing timely insights on delayed deliveries, low inventories, and other critical issues. This automation streamlined workflows, replaced manual table checks and enhanced efficiency by 90%.
* Developed and monitored 20+ data pipelines that extract data from ERP systems, apply business logic, and prepare it for reporting. Ensure data accuracy, optimized processing efficiency, and maintained reliable storage, making data available for live reporting
* Developed Tableau report for live inventory tracking**,** automating checks for raw materials used for production. This system alerted users to discrepancies in real-time, reducing manual adjustments and improving inventory tracking accuracy by 99%.
* Educated stakeholders on transitioning from Microsoft Excel to interactive dashboards using Tableau, leading to an 80% adoption rate. Enhanced analytical capabilities, improved decision-making, and streamlined reporting through dynamic, automated visualizations.
* Migrated 10+ reports from BO to Tableau, enabling stakeholders to access real-time data independently, eliminating delays from scheduled jobs and manual report downloads.
* Conducted technical knowledge transfer sessions and mentored new interns and team members, accelerating their onboarding, and enhancing team integration, ensuring they received 100% of the support needed to understand their roles.

**EDUCATION**

**Master of Science in Data Science**

New Jersey Institute of Technology, United States

**Bachelor of Technology in Electronics and Communication Engineering**

Mahatma Gandhi University,India

**MACHINE LEARNING AND AI PROJECTS**

**Binary Classification: Motor Insurance Fraud Detection** [**Git Repository**](https://github.com/annchirackal/Insurance-Fraud-Detection---End-to-End-Project)

Developed a web app to detect insurance fraud using K-Means for data clustering, XGBoost, and Random Forest for label detection, optimizing recall to minimize false negatives. Implemented a Flask interface for seamless model retraining and prediction, achieving 95% accuracy in fraud detection.

**LLM & Prompt Engineering: Resume Reviewer App Using Google Gemini** [**Git Repository**](https://github.com/annchirackal/LLM-Resume-Reviewer-App-Using-Google-Gemini)

Successfully prompt engineered a Gemini model to analyze resumes against job descriptions, extract missing keywords, and provide a percentage match, utilizing Google Gemini and deploying the solution with Streamlit.

**Fashion Recommendation Using H&M Data** [**Git Repository**](https://github.com/annchirackal/Recommendation-System-NCF-vs-Matrix-Factorization)

Built a recommendation system using H&M data with a hybrid approach combining collaborative and content-based filtering. Achieved 60% accuracy in the top 12 recommendations, enhancing personalized user experiences and improving product relevance.