# **Rene Infanta Marceline**

480-803-1802 | Tempe, AZ | rimarcel@asu.edu | linkedin.com/in/rene-marceline | github.com/rene-marceline

## **SUMMARY**

Data professional with 2+ years' experience in Data Engineering and 1 year in Data Science. Expertise in data pipelines, exploratory analysis, machine learning, visualization, and experimentation. Proven ability to optimize workflows and drive growth through data-driven solutions. Demonstrated leadership coordinating cross-functional teams, fostering collaboration, and mentoring to achieve business goals.

## **EDUCATION**

## Master of Science in Computer Engineering

Arizona State University | Tempe, Arizona, USA

Relevant Coursework: Algorithms, Data Visualization, Sensors and Machine Learning, Time Frequency Signal Processing, Python **Bachelor of Technology in Electronics and Communication Engineering** August 2016 - May 2020

Amrita Vishwa Vidyapeetham | Coimbatore, Tamilnadu, India

## **EXPERIENCE**

## Software Engineer - Data

Local Grown Salads

- Improved data accessibility by 40% and ensured secure data flow by leading API development with Python/Flask and React, integrating authentication and QR scan/recall features, enabling seamless data-driven decision-making for product enhancements.
- Significantly improved workflows within the LGS ecosystem by enhancing data integrity and retrieval speed by 30%, utilizing PostgreSQL and Firebase for efficient data management.
- Led and coordinated collaborative efforts among cross-functional teams, fostering synergy and maximizing efficiency in achieving project objectives aligned with business goals.

### **Data Scientist**

Blue Guardian

- Spearheaded the development of a machine learning pipeline, using LSTM models for detecting depression from speech data, achieving a 30% increase in model precision through meticulous data pre-processing and advanced hyper-parameter tuning.
- Initiated the development of speech diarization pipelines, implementing K-means clustering techniques to segment speech data into distinct speaker segments, resulting in a 25% reduction in processing time.

### Data Engineer

Synaptic dots Solutions Pvt. Ltd.

- Chennai, TN • Increased data processing speed by 20% by developing an ETL pipeline with Python and MongoDB, enabling real-time insights into customer behavior, and automated workflows using AWS (S3, Lambda), reducing manual efforts by 15%.
- Leveraged advanced data analytics to analyze Facebook ads data, providing data-driven strategic recommendations to clients.
- Boosted revenue by 10% through predictive modeling techniques, leveraging Tableau to create dashboards for swift decision-making based on key metrics, and conducting A/B testing to evaluate the impact of various models and workflows.
- Managed and mentored a team of interns, providing guidance and support on data engineering best practices and project execution.

## **TECHNICAL SKILLS**

Programming: Python, R, Javascript (D3, NodeJs, React), C++, HTML/CSS, PostgreSQL, SQL, MongoDB, Redshift Machine Learning: LLMs, NLP, Regression, Forecasting, Clustering, Predictive, Deep Learning, Boosting, Bagging, A/B Testing Libraries: Pandas, Pytorch, Tensorflow, Keras, NLTK, NumPy OpenCV, Scikit-learn, Matplotlib, Seaborn, CUDA, PySpark, Jupyter Other Skills: Tableau, Power BI, AWS, JIRA, Spark, Hadoop, Kafka, AzureML, CI/CD, Git, Amazon EMR, Airflow, Linux, Docker, Looker, Databricks, Kubernetes, Data Modeling, Data Analysis, ETL/ELT pipelines, Data Visualization, Data Management, MLOps

## PROJECTS

## **Youtube Trends Analysis**

- Executed an end-to-end analysis of YouTube trending videos using Airflow for orchestration, AWS infrastructure for secure data ingestion and storage, Glue for cataloging and preprocessing, and Athena & SQL for extracting insights.
- Developed an insightful OuickSight dashboard visualizing key YouTube metrics like view count, like/dislike ratio, and subscriber growth, revealing videos <10 mins with high engagement in Gaming/Comedy outperformed by 35%.

## **Prediction of Hotel Booking Cancellation**

- Conducted EDA on 14 features using Python to identify trends and patterns, and selected the features that most influence the target value, and built an XGBoost model predicting future demands with 92% accuracy after optimization.
- Constructed an interactive hotel booking dashboard with 3 KPIs and 6 reports using Streamlit in Python, enabling exploration of booking cancellation patterns.

## Credit Risk Analysis: Predicting AmEx Default Probability

- Trained an XGBoost model on a robust dataset comprising over 500K+ records, ensuring ample data representation and enabling the model to capture relationships essential for accurate credit risk default prediction.
- Employed PCA to identify the most predictive variables, resulting in a 20% improvement in AUC on validation datasets.

## PUBLICATIONS

• R. Marceline, S. R. Akshaya, S. Athul, K. L. Raksana and S. R. Ramesh, "Cloud Storage Optimization for Video Surveillance Applications", International Conference on Smart Systems & Inventive Technology (ICSSIT) [Link]

### July 2023 - February 2024

August 2021 - May 2023

Tempe, USA

Tempe, USA

July 2019 - July 2021

March 2024 - Present