

# Ishal Abhishek Mummidivarapu

Atlanta, GA | Mobile: +1 832-497-0178 | Email: [ishalabhishek2000@gmail.com](mailto:ishalabhishek2000@gmail.com) | LinkedIn : [Ishal Abhishek](#)

## SUMMARY

- Accomplished **Data Analyst with 2+ years** of experience in data science & AI/ML, contributing to dynamic projects across diverse industries including HR, Manufacturing, and IT.
- Proficient in **Python, R** and **SQL** for data analysis, with expertise in Pandas, NumPy, Matplotlib, and Seaborn for efficient data manipulation and visualization.
- Applied advanced machine learning models using **TensorFlow, PyTorch**, and **Scikit-Learn** to develop predictive analytics solutions for diverse business challenges.
- Experienced with Generative AI, integrating **LLMs**, including **GPT(3.5)**, and Llama with Langchain framework
- Proficient in data extraction, transformation, loading (**ETL**) processes with experience in **Python, SQL**, and **Big Data** frameworks like **Apache Spark** and **Hadoop**.
- Experience building and deploying complex data pipelines on cloud platforms like **AWS** and **GCP**.
- Translated complex findings into actionable insights, for better understanding of non-technical stakeholders.
- Proven ability to translate data insights into actionable solutions through interactive dashboards and visualizations using tools like **Tableau, Power BI**, and **Domo**.
- Applied ML techniques (regression, decision trees, ensembles) and time series forecasting in use cases.
- Designed and implemented **ETL** processes using Google Cloud, , AWS, Python and SQL for seamless data integration across projects.
- Managed and queried databases including **Oracle, MySQL**, and **PostgreSQL**, ensuring data integrity and accessibility for analytical purposes.
- Proficient in version control with **Git**, collaborating seamlessly with cross-functional teams in agile environments.
- Adapted seamlessly to Windows and Linux environments, ensuring compatibility with diverse IT systems.
- Demonstrated a track record of delivering measurable outcomes and driving positive impact in various data science projects.
- Experience collaborating with data scientists and stakeholders to drive informed decision-making.

## TECHNICAL SKILLS

<b>Programming Language:</b>	Python, SQL, R, SAS
<b>Frameworks &amp; Tools:</b>	Pandas, NumPy, Matplotlib, Seaborn, Plotly, ggplot2, R Studio, PyCharm
<b>Data Visualization:</b>	Tableau, Power BI, Domo Advanced Excel
<b>Machine Learning:</b>	TensorFlow, PyTorch, Keras, Scikit- Learn, NLP, Supervised and Unsupervised Learning, Predictive Modelling, Clustering Techniques, Neural Networks, Regression, Classification Models, Statistical Analysis, NLTK, SpaCy, Gensim, Hugging Face Transformers, Time Series Analysis, Large Language Models (LLMs)
<b>Cloud Platforms:</b>	AWS S3, EMR, Redshift, Google Cloud Platform, Snowflake
<b>Databases:</b>	Oracle, MySQL, PostgreSQL, SQL Server
<b>Big Data Technologies:</b>	Apache Spark, Apache Hadoop, Apache Kafka, Databricks, ETL/ELT
<b>Other Technical Skills:</b>	Microsoft SQL Server, SSIS, SSRS, SSAS, Docker, Kubernetes, Informatica, Talend, Google Big Query, Apache Hive, Data Quality and Governance, Advance Analytics, Data Mining, Data Visualization, Data Warehousing, Data Transformation
<b>Version Control Tools:</b>	Git, GitHub, Bitbucket
<b>Operating Systems:</b>	Windows, Linux

## CERTIFICATIONS

- |  |          |
|--|----------|
| • Domo Data Specialist                 | Oct 2023 |
| • Google Cloud Certified Data Engineer | Oct 2022 |

## EDUCATION

Northeastern University - Boston, MA, USA  
Master of Data Science (CGPA: 3.62)

Sept 2021 - May 2023

PROFESSIONAL EXPERIENCE

Randstad Sourceright, Atlanta

May 2022 – Dec 2022

Data Engineer

- Enhanced a **Python-driven ETL** pipeline by incorporating advanced **SQL** functionalities such as window functions to boost 20% in processing efficiency, expediting the delivery of insights.
- Maintained an automated data pipeline that ingests data from Vendor Management Systems' API into **PostgreSQL** and **GCP Big Query**
- Collaborated with data scientists to create interactive dashboards on **Domo** and **Tableau**
- Streamlined database operations by migrating to GCP Big Query, resulting in a **25%** decrease in data retrieval times. This enabled enhanced support for complex data science algorithms, improving predictive analytics.

Northeastern University, Boston

Sept 2021 - Jan 2022

Data Scientist

- Developed **machine learning models** to analyze energy consumption patterns captured by Sense devices, leveraging techniques such as regression and clustering to uncover insights and identify usage trends.
- Compiled **Power BI** dashboards to visualize energy consumption of over 90,000 houses in California using SENSE's data.
- Forecasted electricity consumption trends using time-series forecasting models such as **ARIMA** and **Exponential Smoothing**
- Implemented interactive filters in the dashboard, enabling stake holders to explore data based on location, time, and energy usage.

Coral Reefs, India

May 2020 – Jul 2021

Data Analyst

- Developed and implemented **AWS lambda functions** to extract marketing data from service S3 buckets, utilizing SQS for file name storage and transferring data to target S3 buckets, enhancing data accessibility and reliability.
- Developed scripts to incrementally load CSV data into **Redshift** data warehouse, ensuring timely and accurate availability of marketing performance metrics for analysis.
- Integrated **Google Ads API** data extraction with AWS Lambda functions and data pipelines, streamlining marketing data collection and facilitating comprehensive campaign performance analysis within AWS infrastructure.
- Designed algorithms and data pipelines to aggregate and integrate marketing metrics such as conversion rate (**CVR**), click-through rate (**CTR**), cost per click (**CPC**), and cost per acquisition/conversion (**CPA**), enabling comprehensive analysis of marketing campaign performance.
- Utilized **SQL** queries to perform data extraction from Redshift data warehouse for analysis, ensuring accurate retrieval of marketing performance metrics.
- Conducted data analysis and applied predictive modeling and time series techniques to forecast marketing metrics and identify opportunities for optimizing marketing spend and improving campaign **ROI**.
- Developed **Power BI** dashboards for dynamic and interactive visualization of Key Performance Indicators (**KPIs**) , providing stakeholders with actionable insights for decision-making.

PROJECTS

Gym Attendance Analytics & Prediction

Jun 2023 – Sep 2023

- Optimized gym management with Python AWS Lambda function, cron-triggered, and integrated with AWS RDS for data scraping; employed ARIMA modeling for trend forecasting, improving operational efficiency and capacity planning.

Fraud Detection System with Snowflake Marketplace Integration

Apr 2022 – Sep 2022

- Improved fraud detection capabilities through Snowpark API for Python-Snowflake integration and IP Geolocation in Snowflake, resulting in significant gains in data processing efficiency and strengthened privacy compliance.

Periocular Region Based Person Recognition

Jan 2022 – May 2022

- Advanced Python, TensorFlow, and OpenCV, processing 45k images with Feature Fusion and Histogram of Gradients, resulting in an impressive 96.5% accuracy rate, markedly enhancing model efficiency and predictive analytics.

Exploring Airbnb Listings: Data Extraction, Analysis and Prediction

Sep 2021 – Dec 2021

- Elevated Airbnb listing analytics with Random Forest Regressor, R, and Python, achieving a 95% price prediction accuracy and integrating geospatial tokenization in Power BI dashboard, streamlining strategic insights for market trends.