MANIDEEP PIPPARI
Data Analyst

📞 **872-328-6416** ✉ **manideep.p@mailjobtech.com** in **LinkedIn** ⌾ **GitHub** ⌾ **Portfolio**

SUMMARY

Data Analyst with 4+ years of experience leveraging AutoML platforms, AWS Redshift, and Python scikit-learn to build predictive models and fraud detection systems. Expertise in Power BI with DAX, SQL automation, and churn analytics delivering $450K revenue growth and 12% retention improvement through ML-powered insights.

SKILLS

Methodologies: SDLC, Agile, Waterfall
Programming Languages: Python, SQL, R
Libraries: Pandas, NumPy, Pytorch, TensorFlow, Scikit-Learn, seaborn, Matplotlib, SciPy, Plotly
Database: Microsoft SQL Server, PostgreSQL, MySQL, MongoDB
Data Visualization Tools and Platforms: Tableau, Power BI, Advanced Excel (VLOOKUP, Pivot Tables)
ETL Tools: Alteryx, Talend, SSIS (SQL Server Integration Services), dbt
Statistical Analysis: A/B testing, Descriptive Statistics, Predictive Analytics, Hypothesis Testing, Regression Analysis, Time Series Analysis
Cloud Platforms: AWS (S3, EC2, Redshift), Microsoft Azure, Snowflake
Big Data Technology: Hadoop, Apache Spark, Apache Kafka, Databricks
Version Control: Git, GitHub
Soft Skills: Critical thinking, Problem-solving, Attention to detail, Communication Skills, Team Collaboration, Stakeholder Management

PROFESSIONAL EXPERIENCE

Data Analyst | PNC, Los Angeles, CA   **August 2024 - Current**
- Extracted and analyzed customer transaction data using **SQL** queries on **AWS Redshift data warehouse**, identifying spending trends that informed marketing segmentation strategy and contributed to $450K revenue increase.
- Automated weekly compliance reports by developing **SQL** views and scheduled SSRS reports connected to **AWS RDS**, decreasing manual report generation time while ensuring regulatory accuracy.
- Leveraged **AutoML platform (DataRobot)** to build fraud detection models analyzing transaction patterns across 200K+ accounts, automatically testing 50+ **ML algorithm**s and deploying top-performing model that flagged suspicious activity with 89% accuracy, reducing fraud investigation time by 30%.
- Implemented **AI-powered** data quality monitoring using **Tableau Prep** with automated anomaly detection, identifying data inconsistencies across 15 KPIs and reducing manual data validation time by 40% while improving dashboard accuracy.
- Performed statistical significance testing on **A/B test** results for email campaigns using **Excel** and **Python**, presenting findings to marketing leadership that shaped customer engagement strategy.
- Conducted **data quality** assessments across customer databases, documenting 180+ data inconsistencies in **Confluence** and partnering with IT teams to resolve systemic data issues.

Data Analyst | Vivma Software Inc, India   **September 2020 – August 2023**
- Built churn prediction model using **Python (scikit-learn, pandas)** analyzing behavioural data from 50K SaaS customers, delivering risk scores that enabled customer success team to improve retention by 12%.
- Designed **Power BI** dashboards visualizing subscription metrics and revenue forecasts with **DAX calculations**, connecting to **Azure SQL Database** for weekly business reviews to C-suite that guided product roadmap and pricing decisions.
- Performed **exploratory data analysis** on user activity logs using **Python** to identify usage patterns, creating data-driven feature adoption reports that informed product team's quarterly sprint planning.
- Established data validation framework using **SQL** constraints and **Python** scripts, implementing automated quality checks on customer data feeds that reduced downstream analytical errors by 18%.
- Maintained **Azure PostgreSQL database** queries and created data extraction scripts for analytics team, version-controlling code in **Git** and documenting processes to enable team scalability and knowledge sharing.

PROJECTS

Global Sustainable Energy Analysis and 2030 Forecast: Conducted analysis of global and country-specific energy consumption, renewable adoption, CO2 emissions from 2000s-present using OWID data; examined GDP correlations; built predictive Linear, Polynomial, Random Forest models forecasting 2030 renewable share with highest Random Forest accuracy; visualized global disparities via choropleths.

Spotify ETL Data Engineering Pipeline — End-to-End (AWS + Snowflake + Databricks + Airflow): Engineered an end-to-end automated Spotify ETL pipeline integrating AWS (Lambda, EventBridge, S3, Glue, Athena), Databricks (PySpark, Delta Lake), Snowflake (ELT, tasks, stored procedures), and Apache Airflow (Docker). Built a full Bronze-Silver-Gold lakehouse, used scalable ingestion, orchestration, data quality validation, partitioning, deduplication, and schema management, and delivered BI-ready curated datasets powering Power BI dashboards.

EDUCATION

Master of Science in Business Analytics | Sacred Heart University   **December 2024**
Bachelor of Business Administration | Dr Narayana College of Commerce, India   **December 2020**

CERTIFICATIONS

- **Introduction to Data Analytics by IBM and Coursera**   **G**enerative AI Mastermind by Outskill