Satya Sai Manoj Pithani

Dallas, TX, 75080; +1 469.961.8963; manojsatya91@gmail.com, Portfolio, LinkedIn

EDUCATION

The University of Texas at Dallas **January 2023 - December 2024**
Master of Science, Business Analytics GPA 3.8
National Institute of Technology, Tiruchirappalli **May 2015 - May 2019**
Bachelor of Engineering GPA 3.1

SKILLS & TOOLS

Programming: MATLAB, Python, Java, SAS, Spark, SPSS Statistics, Machine Learning, CI/CD Pipelines, KNIME, PostgreSQL
Tools: Hadoop Hive, Looker, Tableau, Power BI, AWS S3, Redshift, Agile, NLP, MS Excel, MapReduce, Git, GitHub, Docker
Cognitive Skills: Critical, Strategic, Collaboration, Innovation, Decision-making, Dynamic, Communication, Analytical

PROFESSIONAL EXPERIENCE

Data Scientist, Amazon Inc, Hyderabad, India **January 2020 - September 2022**

- Implemented a Q&A system using a Large Language Model to streamline fraud inquiry responses, reducing investigation time by 70% and achieving an 85% accuracy rate in identifying fraudulent patterns and behaviors.
- Migrated 5TB of on-premises data to AWS Redshift and RDS using ETL/ELT processes, reducing migration complexity within 4 months, enabling 50% faster reporting for stakeholders.
- Created ML models to identify customer purchase trend analysis, improving recommendations to customers by 32%.
- Visualized predictions in Amazon Quick Sight, enabling stakeholders to mitigate stock shortages and overages patterns by 15% over a 2-year period enhancing inventory management.
- Enhanced ad-hoc analytical capabilities by SQL querying and automating ETL workflows through AWS S3, boosting data processing efficiency by 20% and scaling solutions within Amazon Quick Sight.
- Streamlined data analysis by automating 30+ monthly reports using Microsoft Excel array formulas, Pivot tables, and VBA, reducing manual handling time by 11.25% in 6 months and elevating reporting efficiency.
- Spearheaded customer feedback analysis by leveraging Tableau to pinpoint critical service improvement areas, driving a 21% increase in customer satisfaction metrics and optimizing service quality.

Data Analyst, Hindustan Zinc limited Vedanta, Udaipur, India **June 2019 - December 2019**

- Harnessed the KNIME analytics platform to perform A/B testing and generated strategic data-driven insights, improving (KPIs) financial performance by 7.6% in Q3 and prevented a 12% loss in silver production.
- Executed model risk assessments aligned with SR 11-7 by validating models, analyzing risk exposures, and establishing frameworks, reducing compliance breaches by 85% and improving operational reliability.
- Engineered predictive models using Decision Trees and Random Forests to forecast free flow market trends, driving data-informed decisions that boosted quarterly sales by 5% using Google analytics.

Data Engineer Internship, Indian Institute of Management, Ahmedabad, India **May 2018 - August 2018**

- Performed scalable Big Data workflows integrating HDFS with AWS cloud services, optimizing large-scale Healthcare data management for 5TB+ datasets, resulting in 11% improved processing efficiency.
- Designed and Streamlined ETL workflows with PySpark scripting and Impala to integrate data from HDFS, MongoDB (NoSQL databases), and AWS S3, reducing data integration time by 20%.
- Implemented NLP text classification models to analyze 5B+ hospital surveys, improving stakeholder insights by 40%.

ACADEMIC PROJECTS

Box Office Revenue Prediction | Python - NumPy, Pandas, Sklearn TensorFlow **March 2024 - May 2024**

- Preprocessed a dataset of 7,000 films, applying feature engineering resulting in an 8% increment in model accuracy.
- Developed and evaluated ML/deep learning models, including Random Forest, XGBoost and Neural networks, achieving best predictive accuracy with RMSE reduced by 30% and R-squared values exceeding 0.85.
- Leveraged the "Dollar Concept" to rank and create binary features for production companies, cast, and crew, for statistical modeling, pinpointing quantitative key influencers that helped film revenue potential by 10%.

Yelp Review Quality Analyzer with NLP | Python - Keras, NLTK, SpaCy, PyTorch **June 2023 - August 2023**

- Developed an XGBoost ML model to predict high-quality Yelp reviews, achieving an accuracy of 84% for true positives and 83% for true negatives, on a dataset of 8 million reviews, outperforming baseline models by 32%.
- Engineered 120+ NLP features, including BERT-based sentiment scores, and syntactic dependency patterns, reducing mean square error by 35% (from 16.6% to 10.79%) for 5-star rating predictions.