PRABHAT CHANDA

(857) 328-3304 | prabhchanda@gmail.com | LinkedIn | Portfolio

PROFESSIONAL EXPERIENCE

Data Operations Co-op | Enel North America | Andover, MA **Feb 2024 – Jun 2024**

- Designed 4 interactive Tableau dashboards using SQL to extract Salesforce data, enabling real-time KPI tracking.
- Automated workflows and reduced manual processing time by 60% for 5 Energy Market utilities, including NYSEG.
- Ensured accuracy in demand response programs by validating 300 sites' records against company data using Excel.
- Boosted business growth and secured $10M in value by resolving 1K+ cases and procuring data from 400 utilities.
- Saved 200 hours weekly by performing root-cause analysis and automating work orders using 5 Shiny R scripts.
- Streamlined 2M records for CRM integration via Excel (PivotTable, VLOOKUP, VBA) and reduced errors by 25%.
- Delivered 2 analytics projects using Agile methodologies and enhanced data reporting with clear documentation.

Graduate Teaching Assistant | Northeastern University | Boston, MA **Sep 2023 – Dec 2023**

- Held weekly office hours for 70 students, evaluated 500+ assignments and developed 20 educational materials.
- Organized 3 hackathons and mentored 15 groups to develop Data Science solutions for real-world challenges.

Data Analyst | Chanda Hospital | Sonipat, India **Aug 2021 – Aug 2022**

- Centralized 2 years of healthcare records to reduce paperwork by 70% while maintaining data quality and integrity.
- Optimized inventory through statistical evaluations and data-driven strategies and saved 8% in procurement costs.
- Increased patient engagement by 15% through targeted marketing campaigns informed by demographic insights.
- Predicted demand using time series forecasting and stochastic processes and increased patient capacity by 20%.
- Delivered insightful reports and highlighted risks to internal stakeholders by analyzing data through Power BI.

PROJECTS

Sentiment Analysis Pipeline | GCP, ELK, CI/CD, Tableau, Docker, Airflow, Pinecone **Sep 2024 – Dec 2024**

- Built an MLOps pipeline in GCP using Airflow to automate the processing and evaluation of 338M Amazon reviews.
- Utilized Snorkel for weak supervision and leveraged statistical modeling to improve classification accuracy by 75%.
- Integrated CI/CD workflows and established real-time monitoring using ELK stack to enhance system observability.
- Implemented RAG pipelines to generate insights and delivered a Tableau dashboard for strategic decision-making.

Analysis of Corporate Layoffs | Python, Tableau, Statistics, Visualization, Reporting **Sep 2023 – Dec 2023**

- Analyzed workforce data for 164k employees and built 5 Power BI dashboards to identify key trends in the industry.
- Performed data pre-processing in Python and reduced errors by 95% using advanced statistical methods.

E-Commerce Database | SQL, Python, Mongo DB (NoSQL), Data Visualization **May 2023 – Aug 2023**

- Created a scalable database architecture using SQL and MongoDB and improved query performance by 75%.
- Executed optimized complex SQL queries to analyze data, uncover actionable insights and streamline operations.

Income Classification of UCI Census | Python, ML, Statistics, Algebra, Feature Engineering **Jan 2023 – Apr 2023**

- Developed 5 machine learning models from scratch and achieved highest accuracy of 79% with Naive Bayes.
- Improved logistic regression accuracy by 15% using L2 regularization and applied SMOTE to handle class imbalance.
- Constructed a neural network model to validate backpropagation and achieved 77% precision and 82% recall rate.

EDUCATION

Master's in Data Analytics | Northeastern University | Boston, MA **Sep 2022 – Dec 2024**

- **Coursework:** Databases, Data Mining, Data Warehousing, Statistical Learning (ML), Machine Learning Operations

Bachelor's in Computer Science | Vellore Institute of Technology | Vellore, India **Jul 2018 – May 2022**

SKILLS

Programming & Databases	Python, MySQL, MongoDB (NoSQL), PostgreSQL, Oracle, R Studio
Frameworks & Libraries	Scikit-Learn, TensorFlow, PyTorch, Keras, NLTK, Shiny, PySpark, Pandas
Data Science & Machine Learning	Regression, Classification, Clustering, Ensemble Learning, Neural Networks, NLP, A/B Testing, Hypothesis Testing, AI/ML Algorithms, Statistical Analysis
Data Analytics & Visualization	Tableau, PowerBI, Looker, Excel, Data Extraction, Reporting, DAX, Big Data, Business Intelligence & Analysis, EDA, Data Modeling, KPI Analysis
Data Engineering & Cloud	AWS, GCP, DataFlow, Docker, Airflow, MLFlow, FastAPI, GitHub Actions, Git, Elasticsearch, Logstash, Kibana, Jira, Salesforce, ELT/ETL, CI/CD, Snorkel