NIRAJAN KHADKA

Mississauga, ON | Nirajan.Tech@gmail.com | **437-344-5361** | www.github.com/NirajanKhadka| www.linkedin.com/in/nirajan-khadka/

EDUCATION

Lambton College, Mississauga, Ontario, Canada **April 2023 – December 2024**
PG in Business Analytics | GPA: 3.6/4.0

Lovely Professional University, Phagwara, Punjab, India **August 2018 – July 2022**
B. Tech in Computer Science and Engineering | CGPA: 8/10
Specialized in Machine Learning and Big Data

SKILLS

Languages and Tools: Python, SQL (MySQL, PostgreSQL), R, Git
Data Visualization: Power BI (DAX), Tableau, Matplotlib, Seaborn
Data Analysis: Excel, Pandas, Statsmodels
Techniques: A/B testing, Hypothesis testing, Forecasting, EDA, Statistical Modeling
Cloud and Deployment: AWS (EC2, S3, Lambda, QuickSight, CloudWatch, IAM), Docker, Fast API, Streamlit, Flask
Machine Learning: Scikit-Learn, Tensorflow
Certifications: AWS Certified Cloud Practitioner, Google Advanced Data Analytics Professional Certificate

WORK EXPERIENCE

Freelance Data Analyst | Self-Employed **December 2022 – July 2024**
- Guided 30+ clients through data science projects, using **Python, SQL, and Power BI** for analytics, modeling, and dashboards.
- Created 10+ interactive dashboards in Power BI and Tableau, including sales analytics, driving data-informed decisions.
- Facilitated 50+ virtual sessions on data workflows, increasing project approval rates by 25%, and earning a 4.9/5.0 client rating.

Machine Learning Intern | Widhya | Hyderabad, India **January 2021 – February 2021**
- Executed quantitative modeling and EDA using Scikit-learn to forecast COVID-19 spread, achieving 92% prediction accuracy.
- Developed a flight delay prediction model with decision trees, achieving a 99% ROC-AUC score, improving scheduling insights.
- Designed a stock price prediction model using linear regression, achieving a low MSE of 1.6 on historical stock data (AAPL).

Machine Learning Intern | Ignitus | Pittsburg, Unites States **April 2020 – September 2020**
- Authored k-NN learning module content for an LMS, incorporating case studies like customer segmentation and fraud detection.
- Designed hands-on exercises, including predictive modeling and anomaly detection, aligned with financial and retail industry.
- Conducted data analysis using Pandas to identify learning behaviors, enabling improvement in LMS content engagement.

KEY PROJECTS

Learning Platform Engagement Analysis (Link) **MySQL, Excel, Python, Pandas, Scikit-learn, Matplotlib, Seaborn**
- Analyzed user engagement data (632 records) with SQL, Excel, and Python, revealing a 15% engagement drop among paid users from Q2 2021 to Q2 2022.
- Formulated and executed SQL queries for data extraction and preprocessing, including the creation of time-series views and periods.
- Conducted hypothesis testing in Excel, confirming with 95% confidence that paid users, on average, have more engagement.
- Utilized Python for outlier removal and developed Regression models, achieving 0.30 MSE in forecasting future engagement.

Credit Card Loan Default Prediction (Link) **Python, Pandas, Matplotlib, Seaborn, Scikit-learn, AWS Lambda**
- Automated data cleaning for 32K+ records using Python, employing median imputation and scaling resulting in 95% data quality and reducing data preparation time by 25%.
- Performed EDA with Pandas and Seaborn, identifying income-to-loan ratio as a primary default risk factor through feature analysis.
- Designed and optimized a machine learning pipeline, achieving 96.3% precision and 0.81 F1-score in predicting loan defaults.
- Deployed the model on AWS Lambda with FastAPI and Docker, enabling real-time predictions with a 300ms inference time.

Customer Churn Power BI dashboard (Link) **Python, Pandas, Matplotlib, Seaborn, Scikit-learn, Power BI**
- Built a machine learning model with 93% accuracy using XGBoost and Random Forest on 2K+ customer records, engineering features like customer tenure and usage frequency.
- Identified key churn indicators from customer data using statistics, enabling targeted retention strategies for high-risk segments.
- Created an interactive Power BI dashboard with churn trends, predictive model outputs enabling real time decisions.
- Optimized models via hyperparameter tuning improving prediction accuracy from 93% to 97% for churn identification.