

YASH PRADHAN

Charlotte, NC | +1 (704) 439-7859 | ypradhan@charlotte.edu | **LinkedIn:** www.linkedin.com/in/iamyashpradhan/ | **GitHub:** <https://github.com/iam-yashpradhan> | **Medium:** https://medium.com/@iam_yashpradhan

EDUCATION

University of North Carolina at Charlotte
Master of Science in Data Science & Business Analytics

Charlotte, NC, USA
Expected Graduation: December 2024

SKILLS

- Programming:** Python (Pandas, Numpy, Streamlit, Pytorch, Scikit-learn, Tensorflow, PySpark, Keras, Seaborn, Statsmodel, Matplotlib, Plotly, FastAPI), R (Dplyr, Ggplot2, TidyR), Javascript (React, Node, Express), MySQL, MongoDB, PostgreSQL, Hadoop.
- Tools:** Power BI, Tableau, Google Data Studio, Palantir Foundry, Google Analytics, Amazon Web Services, Git, Docker, Google Cloud Platform, Snowflake, Databricks, Postman, REST APIs, Airflow, Power Automate.
- Core Competencies:** Data Science, Machine Learning, Deep Learning, Statistical & Predictive Modelling, API Development, CI/CD Pipelining, Data Analysis/Visualization, Cloud Technologies, Big Data, A/B Testing, Leadership, and Collaboration.

WORK EXPERIENCE

American Tire Distributors
Data Science Intern

NC, United States
June 2024 – August 2024

- Leveraged **Foundry** and **predictive models** to build workflows for optimizing supply chain processes, saving approximately **\$3.3 million** in FIFO inventory costs, and created dynamic dashboarding apps for key stakeholders to provide **functional reports** on each workflow, enhancing data visibility and decision-making.
- Integrated **OpenAI's GPT-4o large language model (LLM)** into an application, enabling the generation of valuable insights based on various application parameters, and enhancing the system's intelligence and user experience.
- Optimized data pipelines using **PySpark**, resulting in a **30% reduction** in processing time, significantly enhancing efficiency and performance.
- Automated alert mechanisms and health checks with **Power Automate**, ensuring **99% system uptime** and proactively identifying potential issues.

UNC Charlotte, School of Data Science
Graduate Research Assistant

NC, United States
May 2024 – Present

- Conducted **spatiotemporal regression analysis** on homicide research data based on county health demographics, utilizing advanced statistical techniques.
- Performed extensive **exploratory data analysis (EDA)** and feature engineering on socioeconomic factors, computing social vulnerability index to enhance model performance.
- Applied **natural language processing (NLP)** techniques, including **named entity recognition (NER)**, for information retrieval from news and newspaper data.
- Leveraged advanced classification techniques such as **CatBoost**, **Graph Neural Networks (GNN)**, **LightGBM**, **Bagging**, and **K-Nearest Neighbors (KNN)** to accurately classify outcome variables.
- Contributed to writing a **research paper**, effectively communicating findings and methodologies.

Graduate Teaching Assistant – Artificial Intelligence & Deep Learning (AI & DL)

January 2024 – April 2024

- Development and research of curriculum materials and assignments for AI and DL course, specializing in topics such as **CNNs**, **RNNs**, and **neural networks**, with a focus on feature engineering and data extraction techniques.

Znova
Senior Machine Learning Engineer

India
September 2022 – August 2023

- Developed a machine learning model for a CPG client, achieving a **15% increase in pricing accuracy**, equating to an estimated annual revenue boost of \$1.5 million.
- Improved brand equity estimation accuracy by 18% using **Ridge**, **Lasso**, and **Linear Regression**. Additionally, created a post-modeling KPI analysis web app with **Streamlit**.
- Implemented a churn rate reduction model, leading to a **71% decrease in customer churn** within the first week of deployment. Utilized **Support Vector Classifier**, **Logistic Regression**, and **Random Forest Classifier**.
- Streamlined data processing with **Extract, Transform, Load (ETL) pipelines**, automating 90% of manual tasks. Conducted Exploratory Data Analysis (EDA) using predefined functions.
- Utilized advanced forecasting techniques (**ARIMAX**, **XGBoost**, **Exponential Smoothing**) to reduce stock-outs by 80%.
- Optimized replenishment and demand management, resulting in a \$2 million annual cost saving for the FMCG organization in carrying costs.

- Implemented **natural language processing** for predictive models, driving an 11% revenue increase through accurate consumer behavior forecasting.
- Constructed and managed ETL pipelines, reducing data entry time by 85%, enhancing operational efficiency for faster **forecasting**, improved data warehousing, and a **25% increase in data-driven insights** through behavioral, cohort, RFM, and funnel analysis.
- Successfully deployed **AWS architecture**, accelerating data processing speed by 30%, enabling timely reporting dashboards and enhanced decision-making.
- Implemented **A/B testing methodologies**, contributing to a data-driven approach and yielding a 40% increase in user engagement and satisfaction.

PROJECTS

- [Detect AI-Generated Text](#): Detection system using BERT, linguistic features, Pinecone vector DB, and OpenAI's API.
- [Stock Prediction App](#): Interactive app using Python, Streamlit, YFinance, StatsModels, and Plotly, leveraging ARIMA time series forecasting to provide data-driven investment insights based on real-time financial data.
- [Demand Forecasting for Bike Sharing System](#): Time Series Analysis using SARIMAX.
- [Open AI API Implementation](#): Chat GPT API Implementation using Python.
- [Breast Cancer Classification](#): A Classifier that accurately classifies a histology image using Keras and Tensorflow.
- [Consumer Surplus Factor Dashboard](#): Optimization of Relative Price Index, Market Share Potential, Pricing Impact and Post Predictive Modelling analysis using Streamlit.
- [Vision Cap](#): An IoT device based on Computer Vision, Text Analysis and Sign Board Detection for Visually Impaired.