# PADMESH NAIK

linkedin.com/padmesh-naik    |    508-762-6909    |    padmeshnaik22@gmail.com    |    github.com/padmeshnaik

## EDUCATION

**Worcester Polytechnic Institute (WPI), Worcester, MA**
Master of Science in Computer Science, GPA: 3.80/4.0                                    August 2022 - May 2024

**University of Mumbai, Mumbai, India**
Bachelor's in Computer Engineering; CGPI: 8.16/10                                    August 2016 - October 2020

## Skills

**Languages and Databases:** Python, Java, SQL, Javascript,  MongoDB, PostgreSQL, Snowflake, Cassandra, Amazon Redshift, HDFS

**Frameworks and Libraries:** Apache Spark, Pandas, NumPy, AWS, Azure, Flask, Django, Rest API, NodeJS, Apache Kafka, Hadoop, Docker, Kubernetes, Jenkins, Git, Tableau, Scikit-learn, Pytorch, Tensorflow

**Methodologies and Practices :** MLOps, Continuous Integration / Continuous Deployment (CI/CD), Version Control, Agile, Scrum

## PROFESSIONAL EXPERIENCE:

**Software Developer Intern (Data Focus Role)**                                    Jan 2024 – April 2024
Clearly Energy, Severna Park, MD

- Developed a scalable and interactive mapping application using Angular.js and OpenLayers for dynamic geospatial visualizations, containerized the application in Docker and deployed using Jenkins for automated deployment (CI/CD).
- Collaborated with the product team to assess data requirements, ensuring the developed solutions aligned with user expectations.
- Built a flexible data architecture using Django Models enabling customizable user profiles for tailored and personalized data views.
- Optimized data retrieval processes using SQL indexing, which significantly improved web responsiveness and user satisfaction.

**Software Data Engineer Intern**                                    May 2023 - August 2023
Building Assure PBC, Concord, MA

- Led the development of a rule-based decision model, streamlining alerts and responses to real-time data coming from IoT sensors.
- Implemented a Python-based data pipeline using AWS Lambda for serverless processing, integrated with MongoDB and Apache Kafka for real-time streaming and data management, resulting in a 25% enhancement in application performance and scalability.
- Integrated monitoring systems using AWS CloudWatch, enhancing the accuracy and reliability of real-time data inputs.

**Data Scientist**                                    September 2020 - August 2022
Tata Consultancy Services, Mumbai, India

- Built ETL pipelines using Snowflake and Python, automating data extraction, transformation, and loading processes, significantly reducing data processing time, enhancing data quality, and ensuring seamless integration for advanced analytics and reporting.
- Analyzed 20 years of purchase data from a Fortune 500 company utilizing Tableau and SQL, identified significant determinants of customer behavior and made modifications to the pricing strategy, resulting in a 12% increase in customer retention.

**Graduate Assistant**                                    January 2023 - Present
Worcester Polytechnic Institute, Worcester, MA

- Collaborating on faculty-led research projects focused on Fintech Technologies, implementing data pipelines using AWS services like Lambda functions, Redshift and S3 to enhance data processing capabilities and support advanced data analytics.
- Building regression models, incorporating feature engineering and model tuning to analyze financial data and derive insights.

## PROJECTS:

**BigMart Outlets Sales Prediction**

- Designed and implemented a robust MLOps pipeline for BigMart sales prediction using Docker, MLflow, Airflow, and Jenkins, ensuring automated tracking, versioning, orchestration, and continuous integration/deployment of machine learning models.

**Movie Ranking System**

- Developed a data processing system to analyze a movie reviews dataset using Hadoop, leveraging MapReduce for data aggregation and movie ranking, and deployed a multi-node Hadoop cluster on a virtual machine for scalable data processing.

**News Recommendation System**

- Built an advanced news recommendation system using Autoencoders, implementing data preprocessing with Python, MySQL and Amazon Redshift to enhance personalized content delivery, improve accuracy, and optimize performance.