

# PREET MODI

+1 (812) 318-2011 | [prmodi@iu.edu](mailto:prmodi@iu.edu) | <https://www.linkedin.com/in/preetjmodi/> | [Portfolio](#)

## EDUCATION

### Master of Science in Data Science

Indiana University Bloomington (Kelley School of Business) | **GPA: 3.85/4**

August 2022 - May 2024

Bloomington, Indiana

Relevant Coursework: Data Mining, Applied Algorithms, Software Engineering, Machine Learning

### Bachelor of Technology in Information Technology

Dharmsinh Desai University | **GPA: 4/4** | CSI-President

August 2018 - May 2022

Gujarat, India

## TECHNICAL SKILLS

**Languages:** Python, SQL, R, Java, C, C++, HTML, CSS, JavaScript, C#, Linux, VBA, React JS

**Databases:** MySQL, SQL Server, MongoDB, PostgreSQL, NoSQL, AWS, GCP, Azure, Oracle, EC2, T-SQL

**Data Science :** NumPy, Pandas, TensorFlow, Keras, Spark, Databricks, Salesforce, Looker, Qlik, Cognos

**Statistics:** SAS, Predictive Modeling, Regression, Classification, Time Series Analysis, Hypothesis Testing, Stata

**Other Tools:** Power BI, CRM, Dynamics 365, SAP, GitHub, Data Warehouse, UI/UX, Artificial Intelligence, GIS

## WORK EXPERIENCE

### Research Data Scientist

Indiana University Bloomington | *Power BI, SQL, IBM Cloud, Docker, Data Visualization*

Aug 2022 - Present

Bloomington, Indiana

- Analyzed data, processed large datasets (>10M records), and created novel visualizations for Carnegie Classification.
- Conducted financial analysis for IU Residential Services, achieving a 15% reduction in housing and dining expenses.
- Designed real-time dashboards, performed data modeling, data reporting using SSRS, and wrote DAX for analysis.
- Served as Associate Instructor for Big Data Management course, conducting labs for 90 students.

### Data Science Intern

Sacoma Specialty Products | *SQL, Epicor, Redshift, SAP, MySQL, Quicksight, Alteryx*

May 2023 - Sept 2023

Edinburgh, Indiana

- Integrated ERP systems with AWS, designed databases, and implemented BAQs to ensure seamless data migration.
- Deployed a scalable data pipeline in Amazon Redshift following ETL processes in Alteryx, optimizing data processing.
- Executed queries in SQL and deployed a Random Forest predictive model to forecast manufacturing orders in ERP.
- Built dashboard to visualize key performance indicators (KPIs), resulted in a 20% increase in supply chain efficiency.

### Data Analyst (High-Performance Computing)

Institute for Plasma Research (IPR) | *Shell Scripting, Dash, Plotly, Hadoop, Hive, Agile*

December 2021 - May 2022

Gujarat, India

- Developed an HPC analytics web-application for cluster having over 10,000 CPU cores and 44 GPU cards.
- Application built using Python and React, following Agile SDLC methodology, and utilized JIRA for project management.
- Automated CI/CD pipelines for app deployment, enhancing system performance and enabling HPC admins to monitor resources, leveraging problem-solving, communication, and cross-functional teamwork for actionable insights.

## PROJECTS

### Topic Modeling on Credit Card Fraud Detection | *XGBoost, Python, MySQL, Selenium, Flask, ChatGPT API*

- Utilized NLP tools including spaCy, NLTK, and Selenium for text preprocessing and topic modeling.
- Leveraged DataIQ for efficient data preprocessing; employed XGBoost model for fraud detection with a 90% accuracy.

### Kansas City Housing Dashboard | *Python, AWS, Power BI, EC2, Jenkins, MVC, Javascript*

- Led the development of a Kansas City Housing web platform using AWS, Django and microservices framework
- Deployed a pipeline on Jenkins for automated deployment and testing of front-end and back-end code

### Zillow Data Modeling and House Rate Prediction | *Synapse, Python, Azure SQL Database, ArcGIS, SVM*

- Optimized database queries and developed data pipeline in Synapse to analyze and visualize housing market trends.
- Implemented Support Vector Machines (SVM) model for house rate prediction, achieving an accuracy of 85%. Integrated ReactJS for frontend development, incorporating ArcGIS and Google Maps API to enhance user interface (UI).

## PUBLICATIONS

### Insurance Management with Premium Prediction | DOI

- International Journal for Research in Applied Science and Engineering Technology (IJRASET)(Impact Factor: 7.429)

### An efficient Artificial Neural Network for Coronary Heart Disease Prediction | DOI

- International Journal for Research in Applied Science and Engineering Technology (IJRASET)(Impact Factor: 7.429)