Preet Modi

Email : prmodi@iu.edu
Linkedin:https://www.linkedin.com/in/preet-modi-ba9332192/
Mobile : +18123182011
Website: https://impreetmodi.github.io/

Education

- **Indiana University Bloomington** — Bloomington, IN
  Masters in Data Science; GPA: 3.93 — *Aug 2022 - May 2024*
  Courses: Advanced Database Concepts, Big Data Management, Introduction to Statistics, Algorithms, Computer Science, Data Mining, Cloud Computing, Machine Learning, Software Engineering, Predictive Analytics(Kelly School of Business)
- **Dharmsinh Desai University** — Gujarat, India
  Bachelor of Information Technology ; GPA: 3.7 (8.29/10.0) — *Aug 2018 - May 2022*

Skills Summary

- **Languages:**: Python, R, SQL, Java, C, C++, HTML, CSS, JavaScript, C#, Linux, ReactJS
- **Database & Tools:**: SQL Server, PostgreSQL, Hive, MongoDB, Tableau, PowerBI, Airflow, Kafka, SAP, SAS, Excel, VS Code, AWS, GCP, PySpark, Databricks, Snowflake, Git, Azure, EC2, MATLAB
- **Data Science:**: ETL, Predictive Modeling, Regression, Classification Trees, Time Series Analysis, Data Warehousing, Natural Language Processing, Hypothesis Testing, Artificial Intelligence, Statistical Analysis, Data Visualization

Experience

- **Sacoma Specialty Products,LLC** — Edinburgh, IN
  Data Science Intern — *May 2023 - Current*
  - **SQL, Epicor, SAP, Excel**: Integrated Epicor and SAP systems with AWS services, leveraging custom Business Activity Queries (BAQs) to optimize supply chain operations
  - **Amazon Redshift, QuickSight**: Created a centralized data lake enabling data extraction, transformation, and loading processes.
- **Indiana University** — Bloomington, IN
  Data Analyst — *Oct 2022 - Current*
  - **Power BI, Advanced Excel and Tableau**: Working with Department of Residential Program and Services to analyze financial data metrics for dorms and eateries.
  - **DAX**: Utilized DAX(Data Analysis Expressions) to create custom calculations and perform advanced analysis for insightful decision-making
- **Institute for Plasma Research** — Gandhinagar, India
  Data Science Intern — *Dec 2021 - April 2022*
  - **Development of Analytics Dashboard for High Performance Computing (HPC) Cluster**: Built a customized, real-time interactive HPC dashboard for administrators to monitor performance metrics of HPC system having impressive capabilities, including a processing power of 1 petaflop, 10,000 CPU cores and 44 GPU cards.
  - **Python, Flask, Dash, Node JS and Tableau**: Utilized a technology stack comprising Python with Flask for the backend, Node.js for the frontend, and incorporated Dash for visualization, enabling the development of a comprehensive application with seamless integration between the different components.
- **JP Morgan Chase and Co.** — Remote, India
  Data Engineering Intern — *Oct 2021 - Dec 2021*
  - **Apache Airflow, Kafka, and Kubernetes**: Implemented scalable data pipelines, ensuring data quality and integrity.
  - **SAS, SQL, MS Power Tools**: Used SAS for statistical analysis, along with Power BI, SQL, and various MS Power tools, to perform data analysis, visualization, and reporting tasks.

Academic Projects

- **Determining the Causal Inference of a new pricing strategy on customer retention rates for an online subscription service (Netflix)**: Conducted **Predictive Analytics** TO predict customer churn and identify potential factors affecting customer retention. The project involved data collection, preprocessing, and performing A/B testing, followed by statistical analysis using Stata to interpret the results and determine the magnitude of the effect of the pricing strategy on customer retention rates. (Jan '23)
- **Exploratory Data Analysis for Bureau of Transportation Statistics Flight Performance**: Implemented a **data pipeline**, Developed a storage model in **NoSQL** server, Executed an algorithm using a parallel programming framework using **Hadoop**, Proposed a cleaning improvement solution, Explored a big data cloud platform environment and finally created an reliable data management plan. **K-Means Clustering** Algorithm was implemented. (Aug '22)
- **Claim Severity Prediction using Computer Vision and Machine Learning**: Developed a machine learning model that can accurately predict the severity of auto insurance claims using images of damaged cars. Used **convolutional neural networks (CNNs)** model to identify the extent of the damage and predict the parts that need to be replaced, as well as determine whether a car is repairable or a total loss.(Oct '22)

Publications

- **"Insurance Management with Premium Prediction ", Volume 9, Issue XII, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 1222-1238, ISSN: 2321-9653 (Impact Factor: 7.429)**: - DOI: https://doi.org/10.22214/ijraset.2021.39416
- **" An efficient Artificial Neural Network for Coronary Heart Disease Prediction ", Volume 9, Issue XII, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 1474-1483, ISSN: 2321-9653 (Impact Factor: 7.429)**: - DOI: https://doi.org/10.22214/ijraset.2021.39559