

Harshwardhan Raghunath Patil

Bloomington, IN | +18128227090 | mail@harshwardhanpatil.com | linkedin.com/in/harshwardhanpatil

Education

Master of Science in Data Science

August 2022 – December 2023

Indiana University, Bloomington

Indiana, USA

Coursework : Statistics, Machine Learning, Cloud Computing, Exploratory Data Analysis, Data Mining

Bachelor of Technology in Computer Science

August 2016 – March 2020

Shivaji University, Kolhapur

Maharashtra, India

Coursework : Operating System, Computer Networking, Distributed Systems, Database Management Systems

Experience

System Engineer

November 2020 – July 2022

Infosys Limited

Bangalore, KA, India

- Operated within a 9-member DevOps team to develop data pipelines that adhere to modern networking infrastructure.
- Created and maintained a comprehensive PL/SQL codebase to ensure accurate data mappings within 13 projects.
- Assisted in the modification of extract, load, and transfer job scripts, resulting in a 60% reduction in job execution time.
- Proposed and initiated automation of 3 operational activities, to improve efficiency and effectiveness.
- Applied Operational Excellence techniques resulting in a 40% improvement in productivity.

Projects

Credit Card Fraud Detection System

- Conducted data cleaning, preprocessing, and feature engineering to improve the quality and consistency of the data.
- Performed exploratory data analysis (EDA) on the dataset of about 3M credit card transactions using Python.
- Developed and implemented a Machine Learning model to detect and classify transactions with an accuracy of 85%.

Travel Recommendation System

- Cleaned and integrated data from Yelp and Tripadvisor to transform the data into more usable form using Python.
- Performed EDA to understand distributions and correlations between different attributes in the data.
- Implemented a machine learning model to recommend travel itineraries based on user preferences and historical data with an accuracy of 80%.

Fake News Analysis and Classification

- Analyzed BuzzFeed dataset of real and fake news to uncover patterns and trends using tidyverse and dplyr libraries in R.
- Evaluated 3 binary classifiers (logistic regression, random forest, and Naive-Bayes) that identify fake news by analyzing the terms (unigrams) in the title and body of articles.

World Happiness Exploration

- Conducted data cleaning and preprocessing on World Happiness Report dataset, including handling missing data, removing outliers and ensuring data consistency for analysis.
- Performed an Exploratory Data Analysis (EDA) on the factors influencing happiness score across 156 nations.

Skills, Activities and Interests

Languages: Python, R, SQL

Frameworks: Numpy, Pandas, Sklearn, Matplotlib, Seaborn, TensorFlow, Tidyverse, ggplot2, dplyr

Tools: Git, Jupyter Notebook, SQL Developer Studio

Platforms: Linux, Windows, AWS

Mathematics: Probability Theory, Descriptive and Inferential statistics, ANOVA, A/B Testing

ML Algorithms: Linear and Logistic regression, Random Forest, Naive-Bayes, K-Means, Neural Networks

Interests: Chess, Origami

Extracurricular Activities

Vice-President, Data Science Club, Indiana University, Bloomington

December 2022 – Present

- Represented Data Science Club at fair, successfully promoting activities and opportunities to potential members.
- Acquired hands-on experience in data science tools and techniques through club projects and initiatives.
- Developed relationships with industry professionals and organizations to enhance club activities and opportunities.