# Sarvesh Mishra [Kaggle Expert]

Kaggle • Github • LinkeDIn

sarveshmishra.bhu20@gmail.com
+91 8957883380

## Educational qualification

Bachelor of Science, Banaras Hindu University — *2020-2023*
*CGPA - 7.6/10*

## Data Science Course & Certification

iNeuron Full Stack Data Science Masters — *06/2023 - 02/2024*

### *Technical skills*

- **Machine Learning** : Supervised and Unsupervised
- **Deep Learning**: ANN, CNN
- **Language and Framworks** : Python, R, Scikit-learn, Flask
- **Database:** SQL, Mysql, Snowflake
- **Libraries** : Sklearn , Numpy , Pandas , Beautifulsoup
- **Natural Language Prcessing (NLP)** : NLTK, Word2Vec
- Statistics, Linear algebra, Probability
- **Data analysis & visualization tools**: Excel, PowerBI
- **Others:** AWS, DSA

### *Projects*

**Kaggle Mohs Hardness Competition** — *View Project*

- Kaggle hosted a competition where various attributes of minerals were provided, and participants had to determine the hardness of minerals using machine learning algorithms.
- Criteria of evaluation was median absolute error
- Used XGBoost, RandomForest Regressor, Gradient boosting algo and ensemble learning to predict the hardness
- Applied GridsearchCV for best parameters
- Performed Data cleaning, data scaling and addressed overfitting of data
- As a result, concluded the competition with a rank of 734th out of 1500+ competitors

**Customer Churn Prediction System Using Machine Learning and Flask** — View Project

- **Objective:** In this machine learning project, the primary objective was to develop a robust customer churn prediction system capable of forecasting whether a customer is likely to cancel their subscription or remain a loyal subscriber. The project employed classification machine learning algorithms to address this critical business challenge, providing valuable insights for customer retention strategies.
- Developed meaningful insights in data exploration part using pandas
- Developed a web application using Flask to deploy the machine learning model locally.
- Created a user-friendly interface for users to input relevant data and receive predictions regarding customer churn.
- Integrated the trained model into the Flask application for real-time predictions.

**Laptop Price Predictor Using Regression Model** — View Project

- **Objective:** The objective was to develop a regression predictive model to predict the absolute price of laptops based on various input variables. The project utilized Python, machine learning models, Pandas for data manipulation, regression techniques, and statistical analysis.
- Employed Pandas for data cleaning tasks, such as imputation of missing values or removal of irrelevant information. Engaged in feature engineering to create new relevant features or transform existing ones to improve the model's performance
- Evaluated the performance of the regression model using relevant metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).
- Achieved a predictive accuracy score of 86%, indicating the model's ability to estimate laptop prices.

**Expense Data Analysis and Visualization Using Power BI** — *View Project*

- **Objective:** This project focuses on leveraging Power BI and Excel to analyze and visually represent expense data, providing stakeholders with actionable insights for informed decision-making. The primary objective is to present expense-related information in a meaningful and easily understandable format. The project employs various visualization techniques, including Bar Charts, Pie Charts, and Line Charts, to effectively communicate key findings.

**Diabetes Data Analysis and Prediction** *View Project*

- **Objective:** The primary objective was to analyze and visualize a dataset related to diabetes, aiming to identify the factors that contribute to the occurrence of diabetes.
- Conducted thorough exploration of the dataset to understand its structure and characteristics. Utilized statistical and graphical methods to summarize and interpret the main features of the data. Visualized relationships between different independent variables and the target variable (Outcome) to identify potential patterns and correlations.
- Implemented a simple machine learning model using regression techniques. Chose an appropriate regression algorithm based on the nature of the problem and dataset.

## Achievements

Achieved Kaggle 2x Expert status through consistent participation and high-performance rankings in data science competitions. Demonstrated expertise in applying advanced machine learning techniques and collaborating with a global community of data scientists

4 Star SQL on HackerRank

## About me:

With a solid foundation in mathematics, physics, and geology acquired during my BSc studies at Banaras Hindu University, I have developed a strong analytical mindset and a keen interest in problem-solving. Now, I am enthusiastically transitioning into the field of data science, driven by my passion to tackle complex data-related challenges and harness the power of data-driven insights to make a meaningful impact.