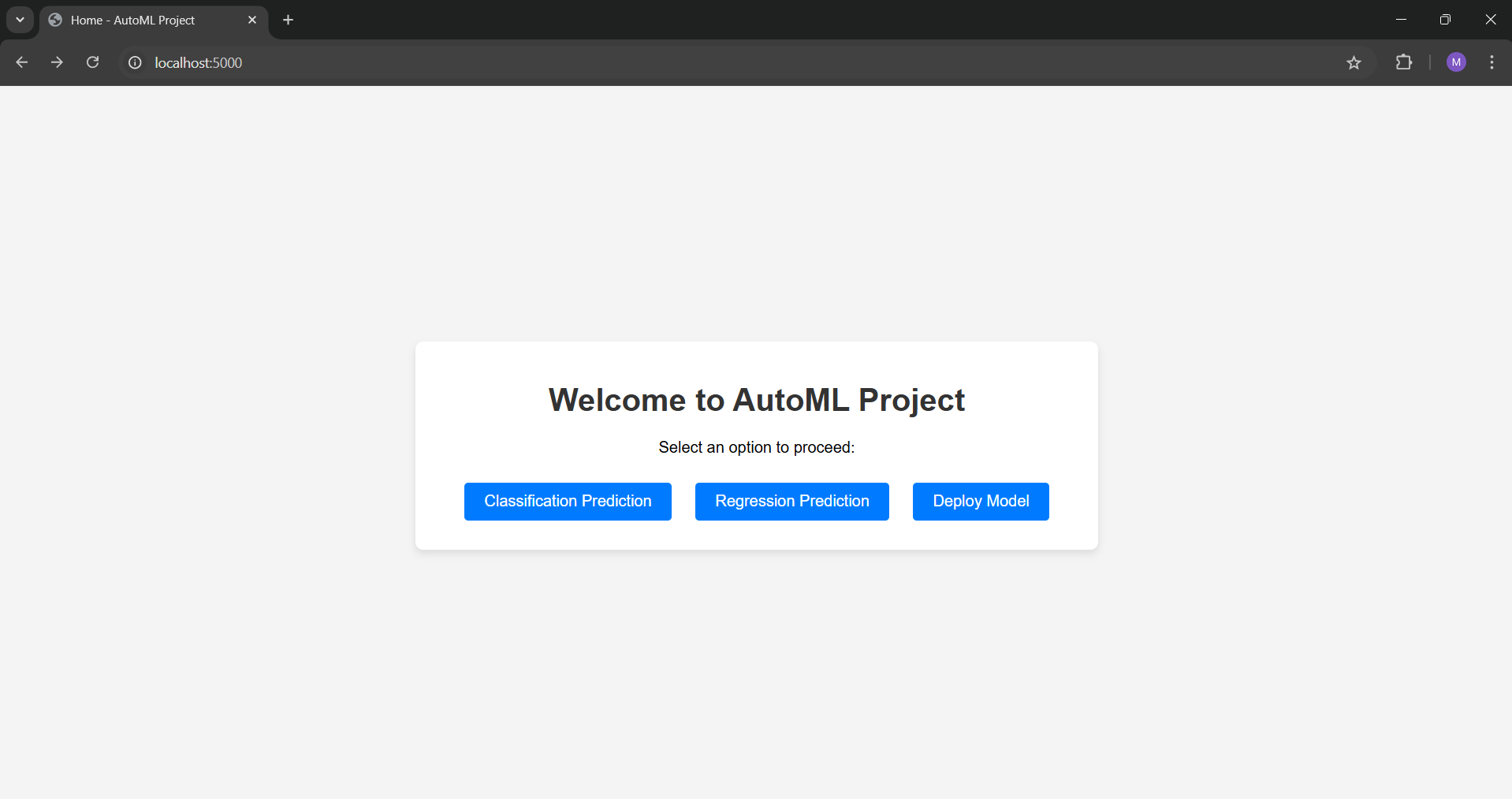
Automated Machine Learning Solution

*WIREFRAME DOCUMENTATION*

Homepage



The home page provides the user 3 options:

1. Classification Prediction:

Shall be used for classification problems.

* This performs all the necessary preprocessing.
* Performs data analysis.
* Trains 12 different models and evaluates them. Then, selects the model with highest accuracy
* Sends to the user the data analysis file, predicted file as well as the trained model

1. Regression Problem:

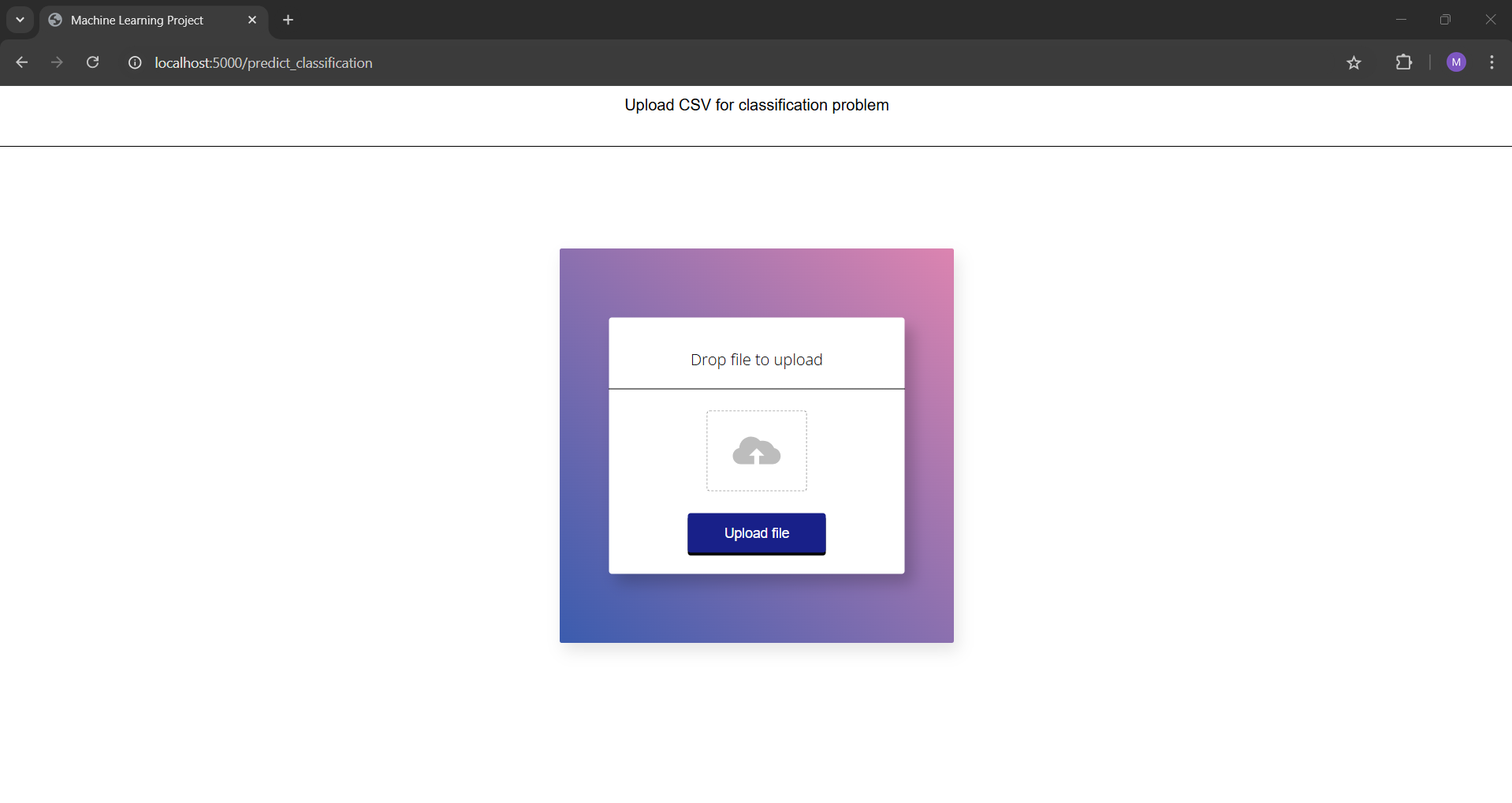
Shall be used for a regression problem.

* This performs all the necessary preprocessing.
* Performs data analysis.
* Trains 12 different models and evaluates them. Then, selects the model with lowest root mean squared error.
* Sends to the user the data analysis file, predicted file as well as the trained model.

1. Deploy Model:

Deploys the model on AWS or GCP according to users choice.

/predict\_classification

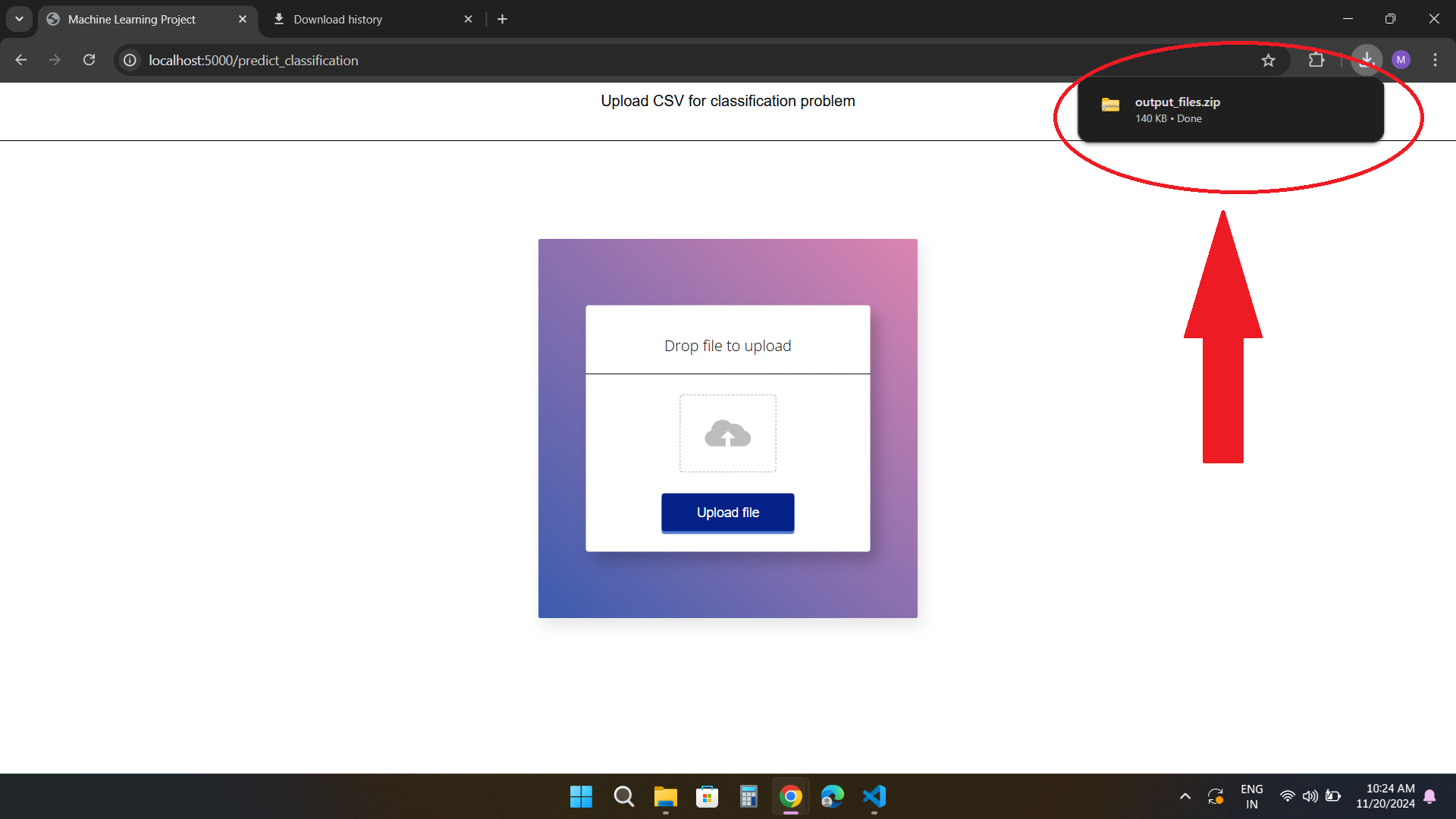


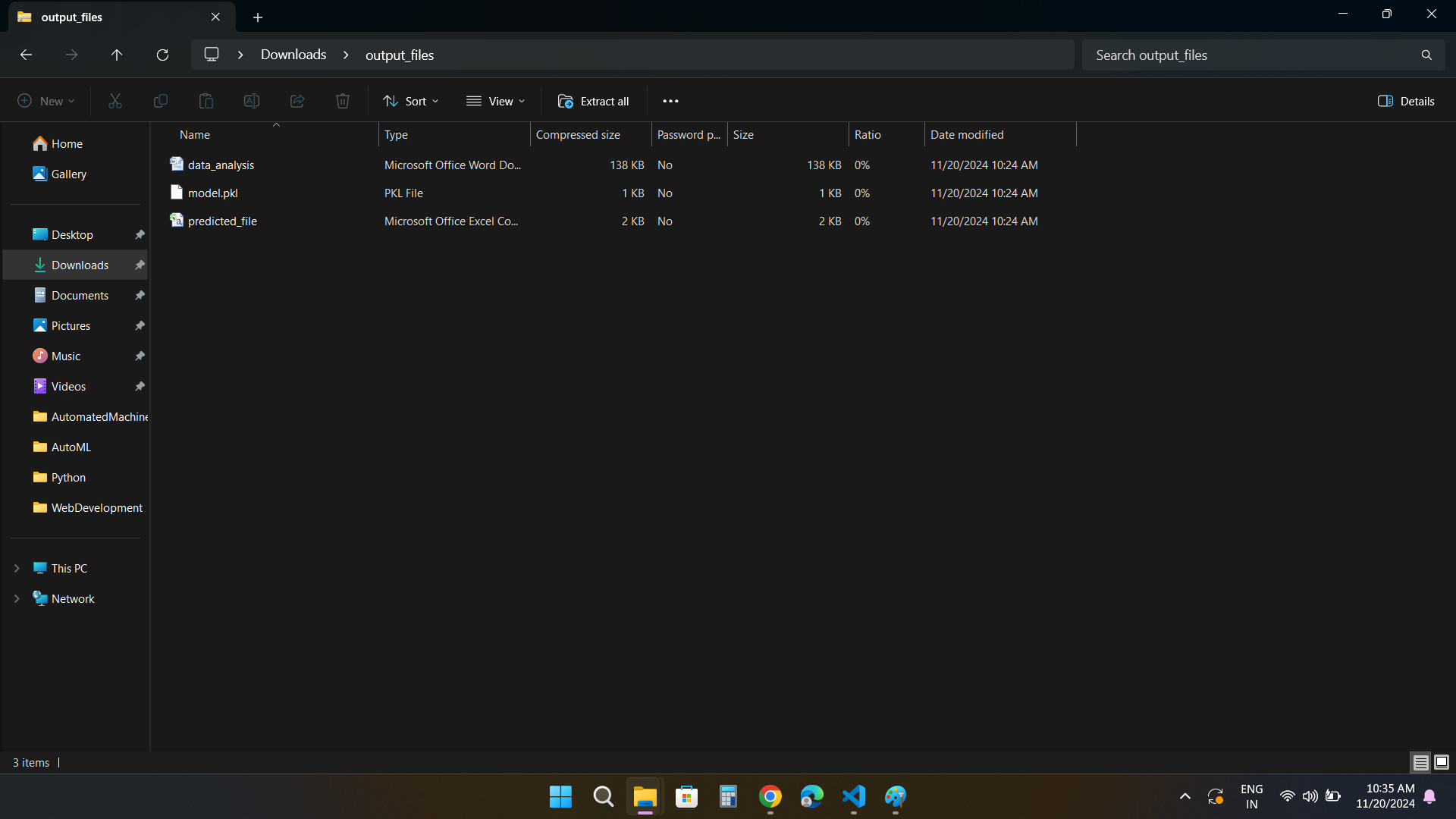
Here, the user shall upload the file either by dropping it in the dropzone or by selecting it from local files.

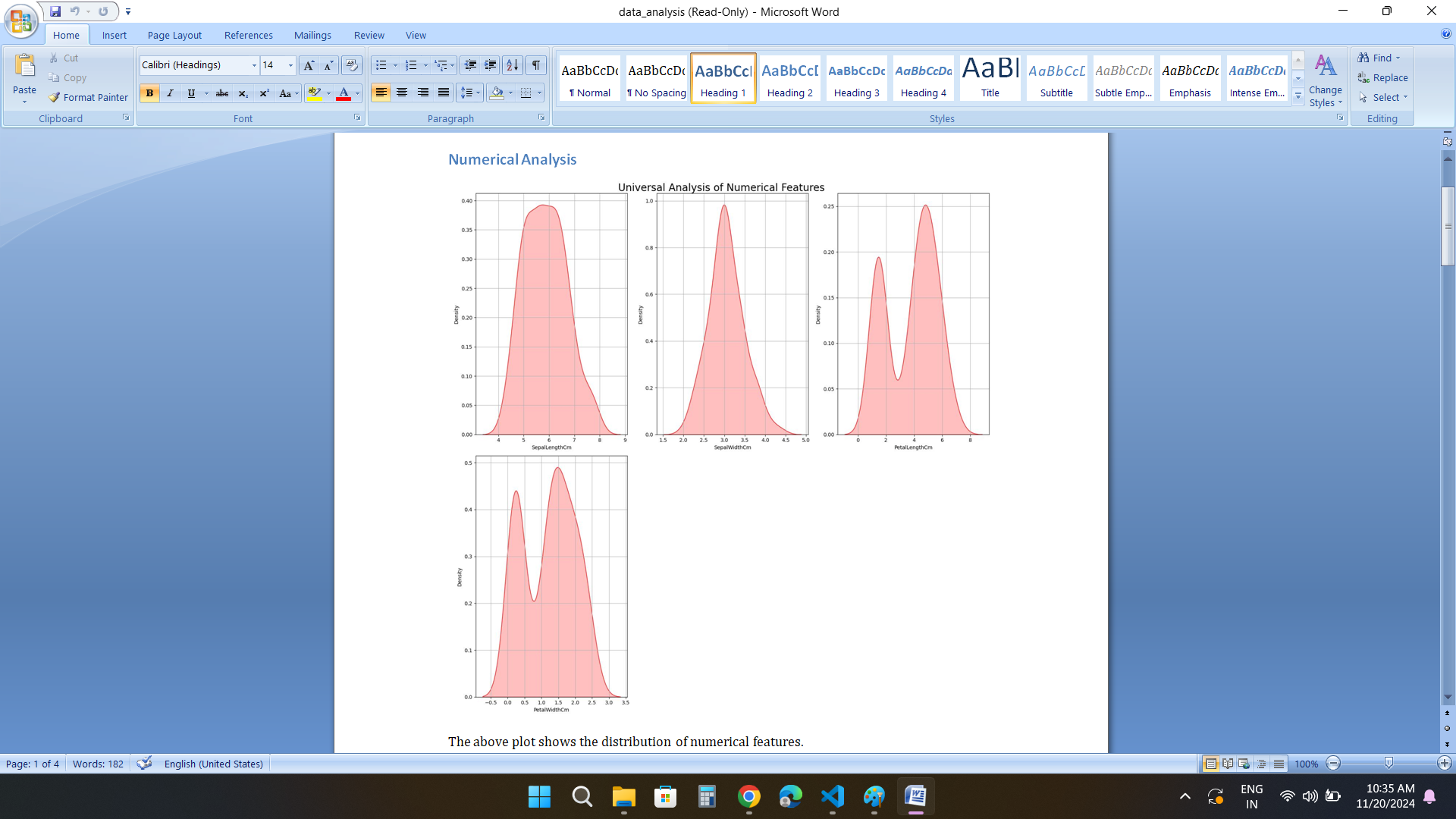
The uploaded file is then preprocessed and analysis is performed.

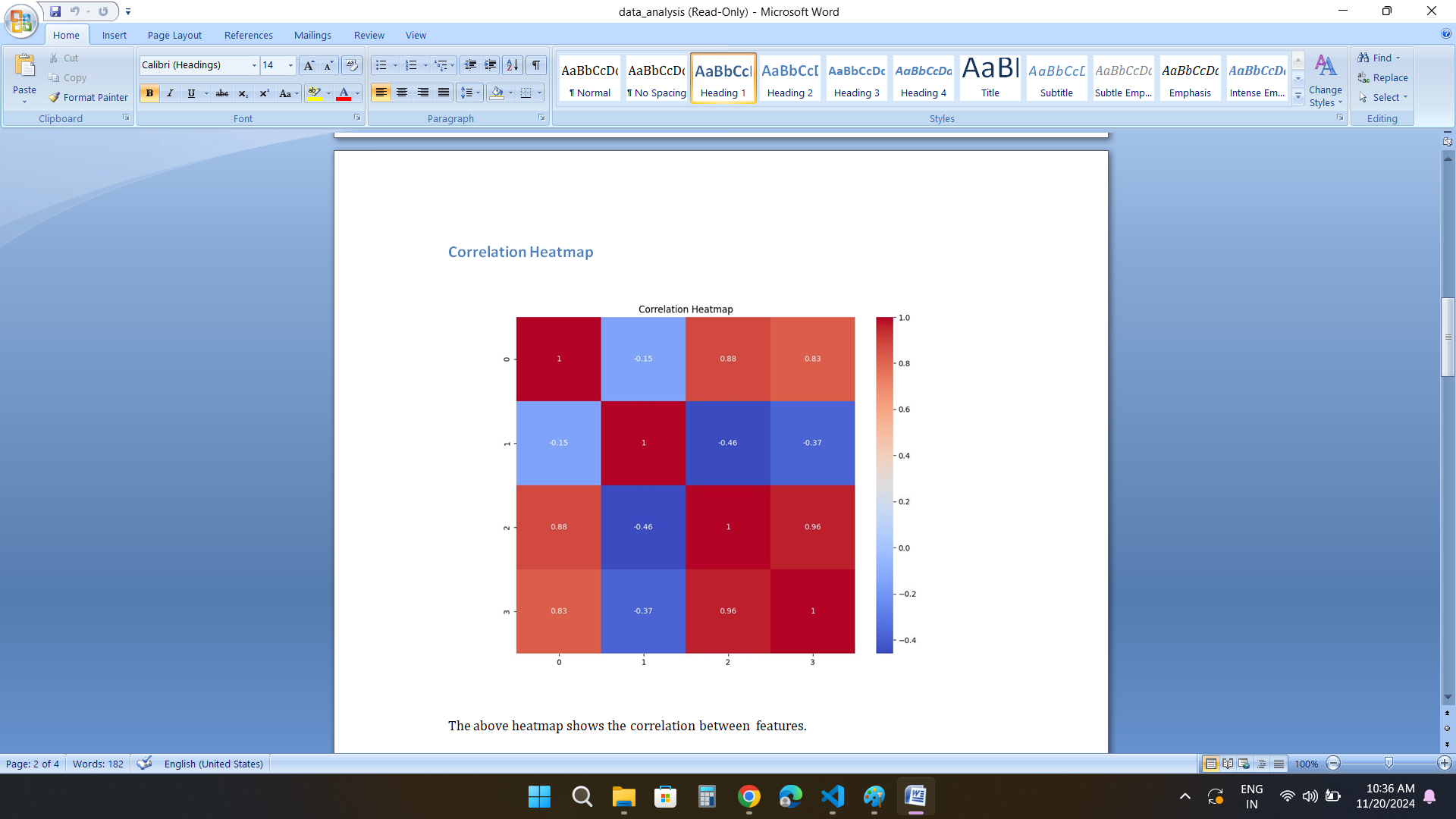
13 Classification models are trained and the one with highest accuracy is then selected.

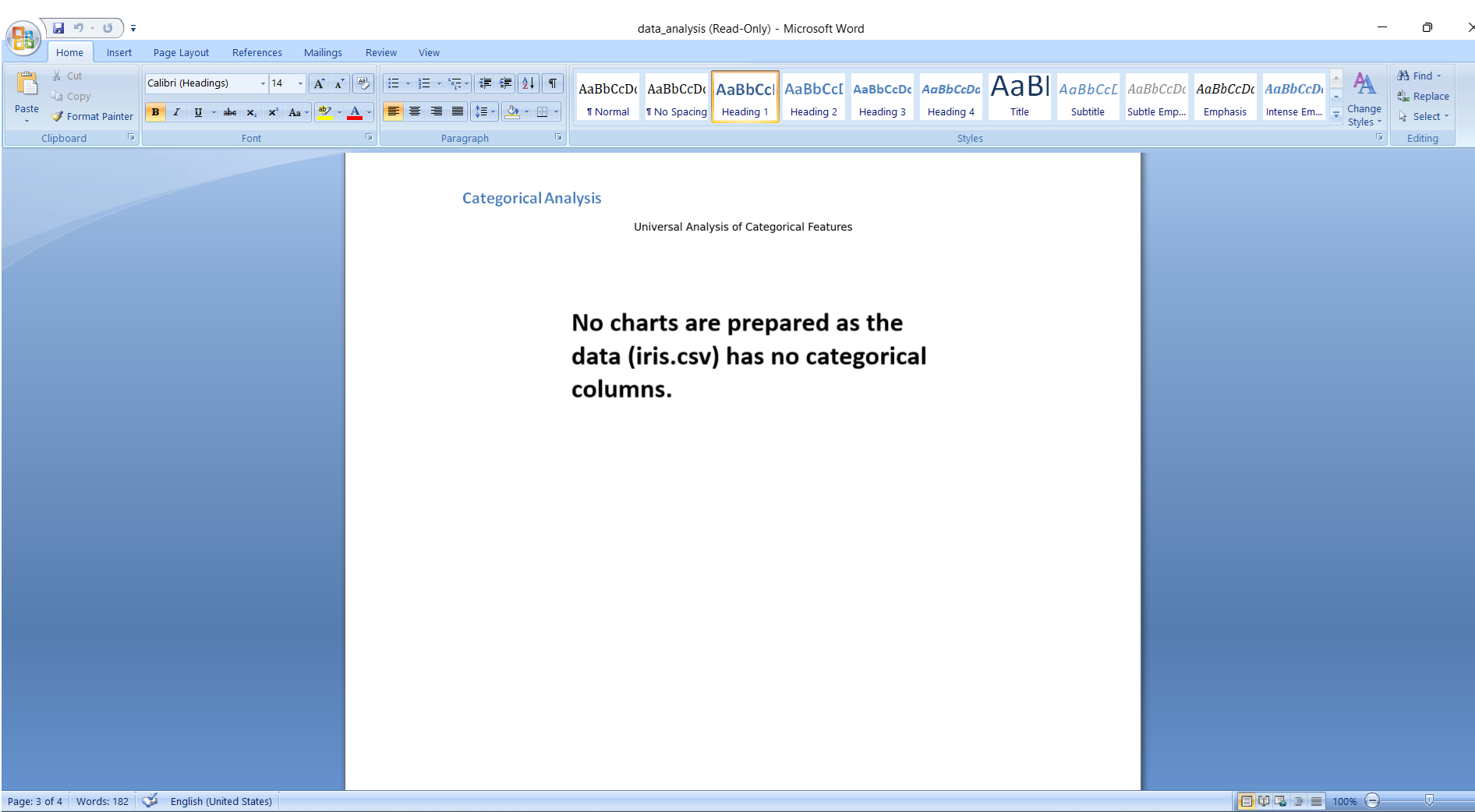
A zip file is downloaded, which contains data analysis, predicted file and the trained model.

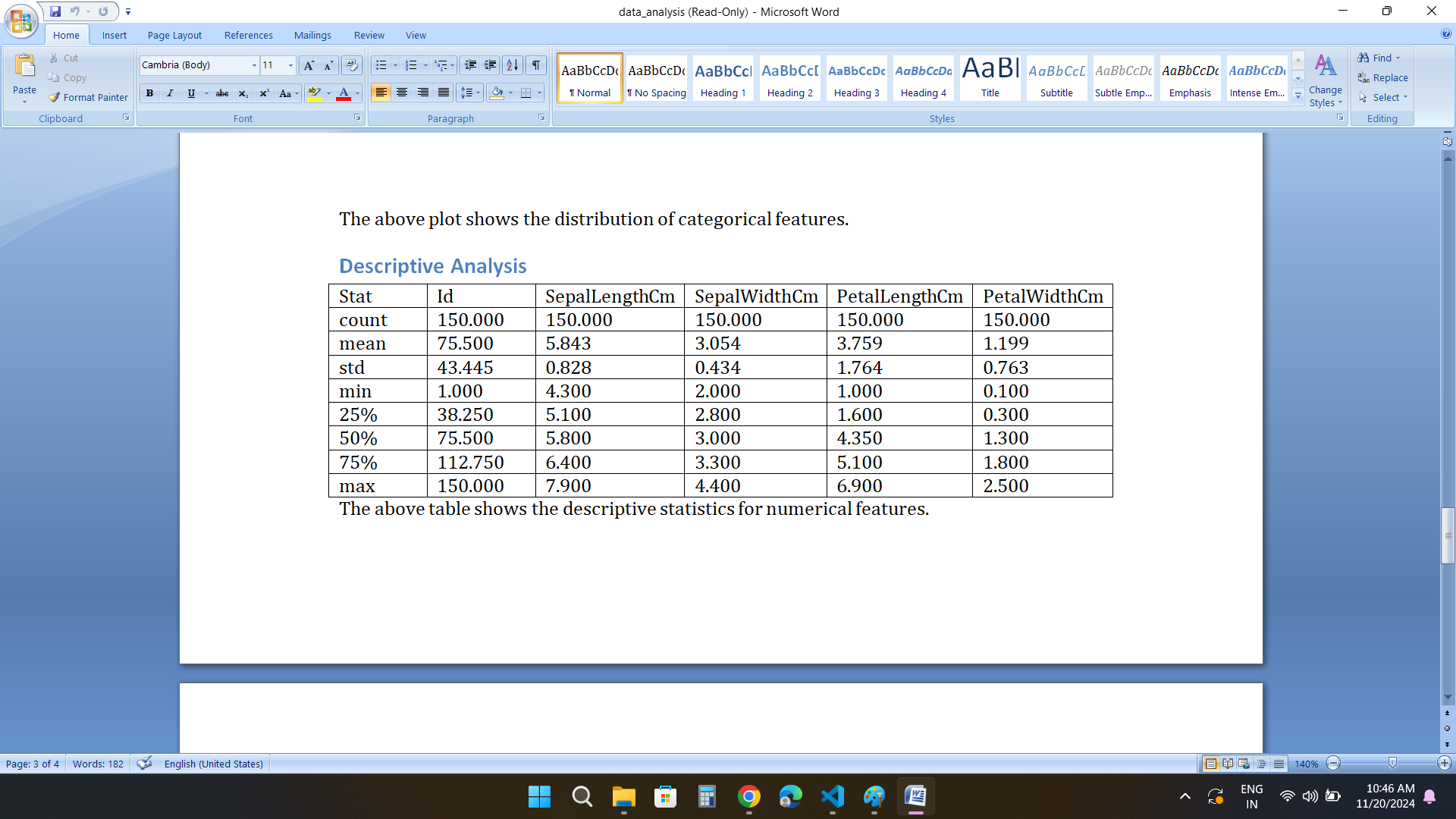


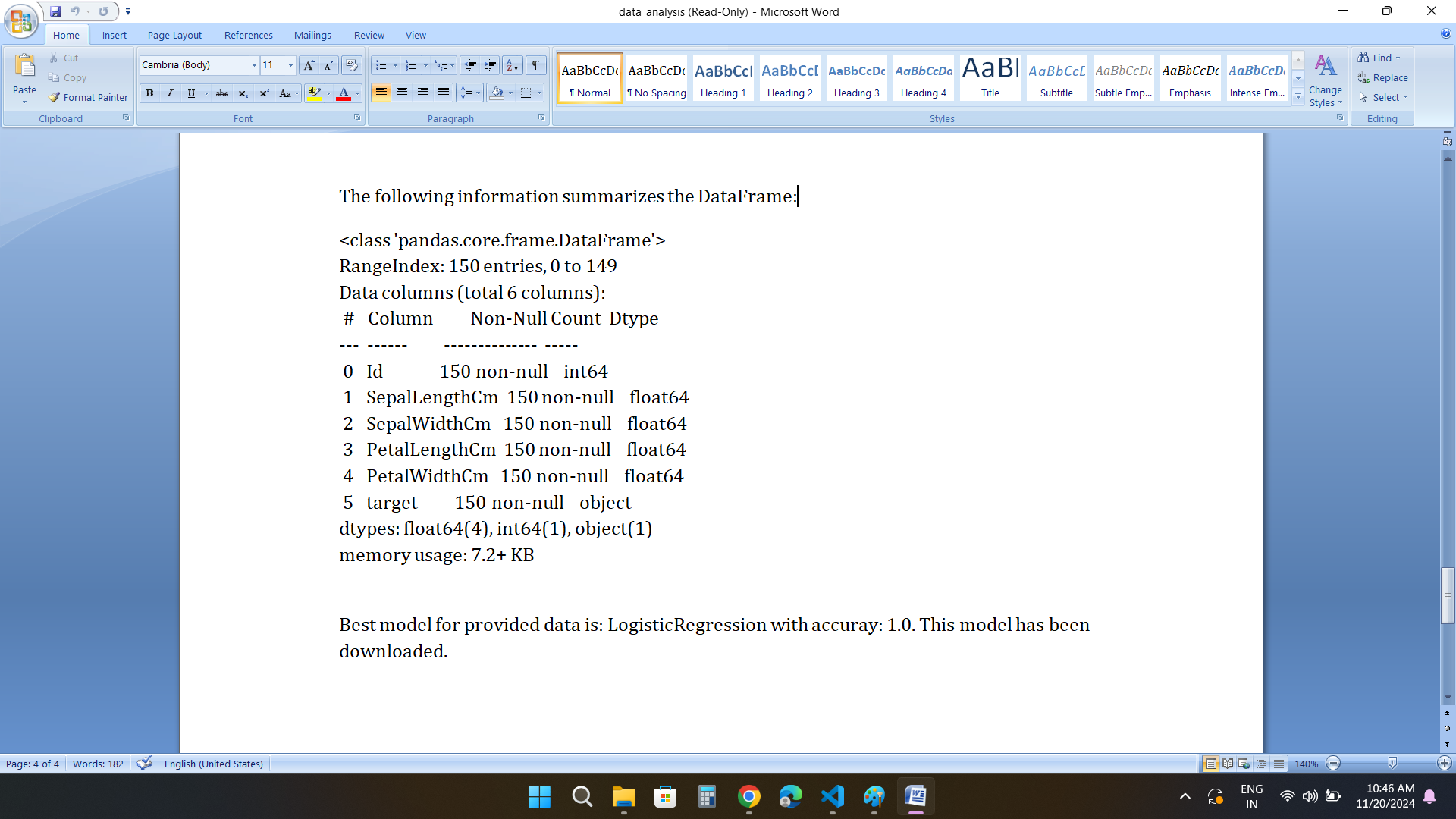


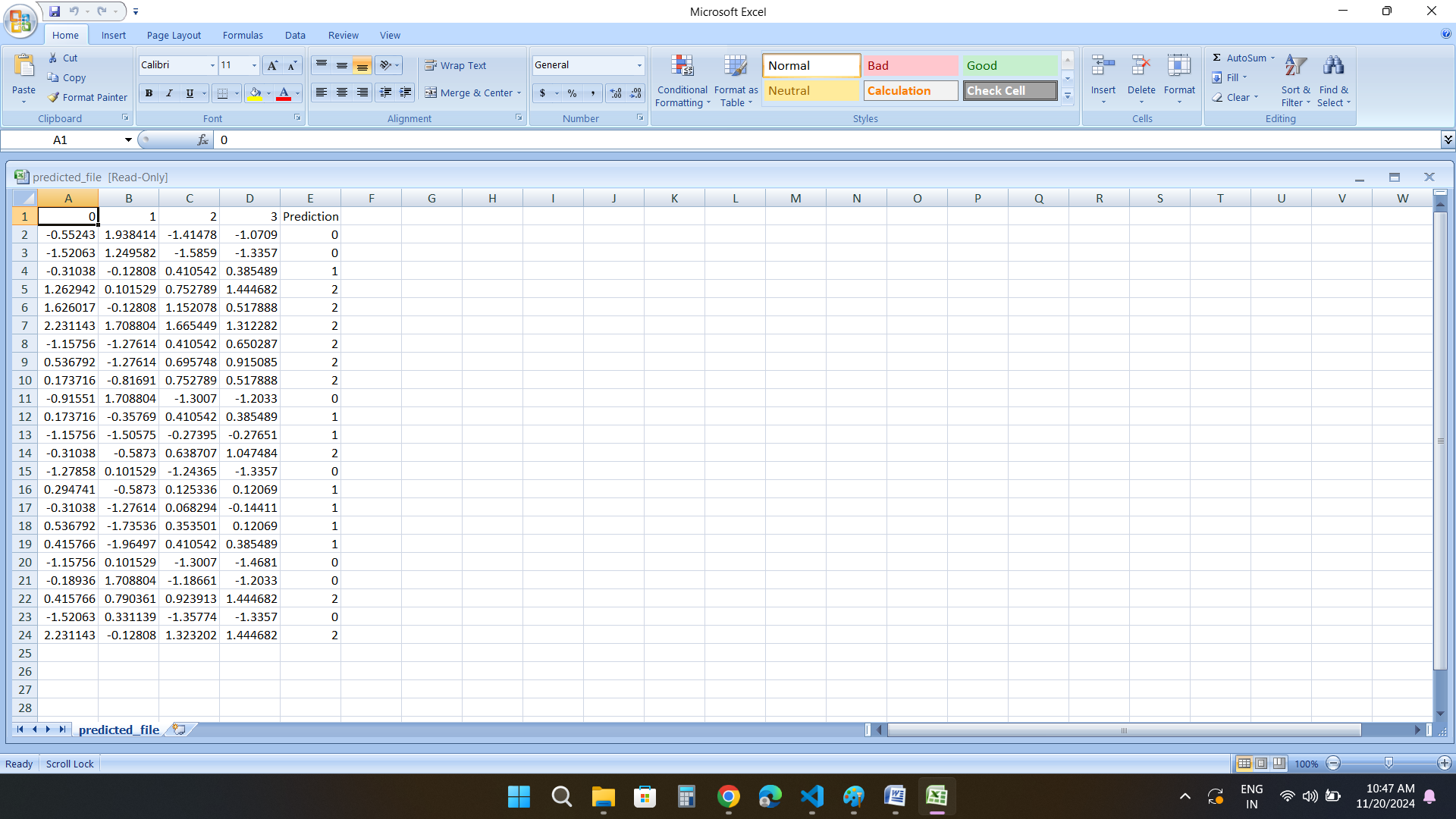






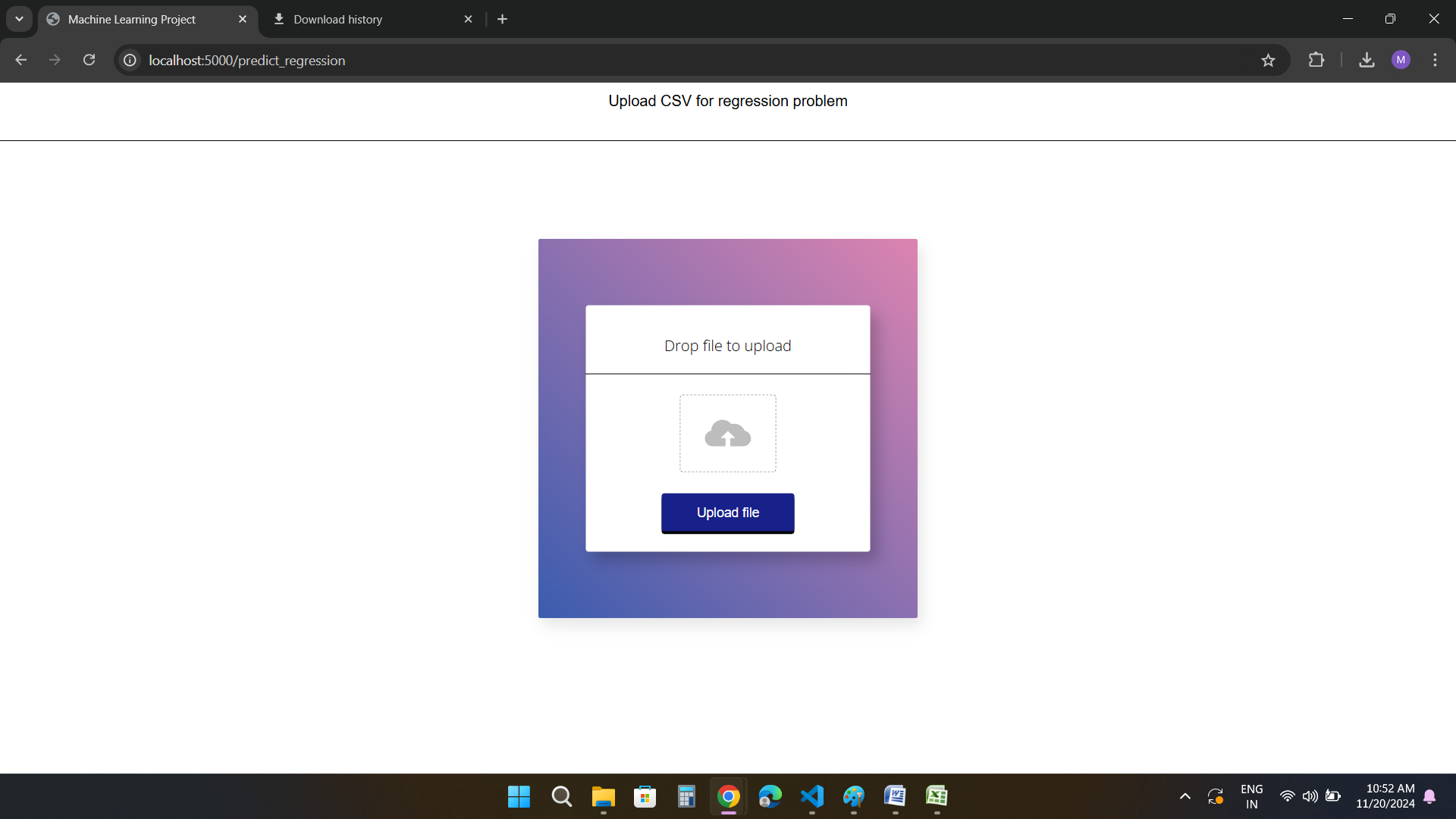






Above is the predicted file.

/predict\_regression

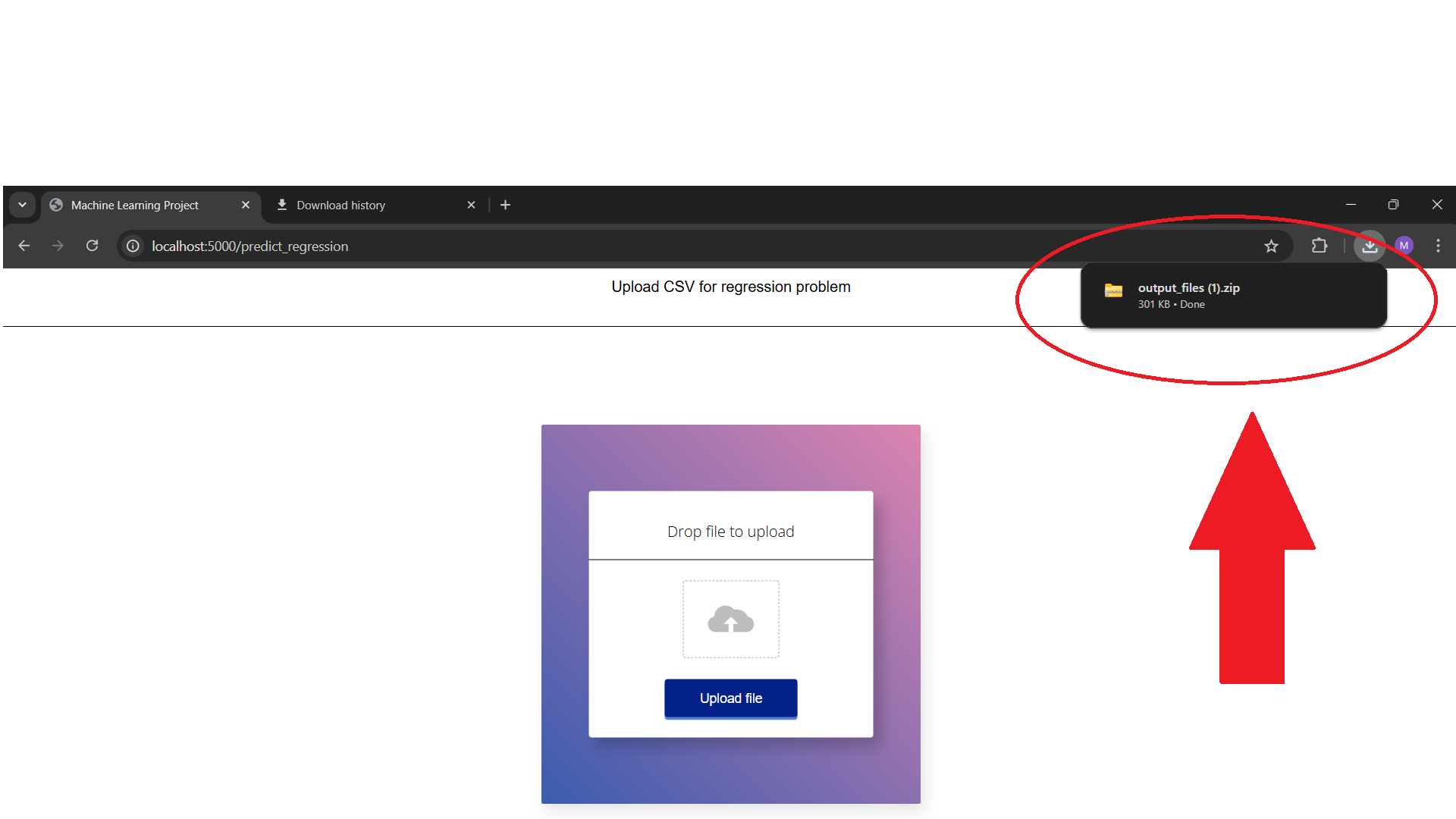


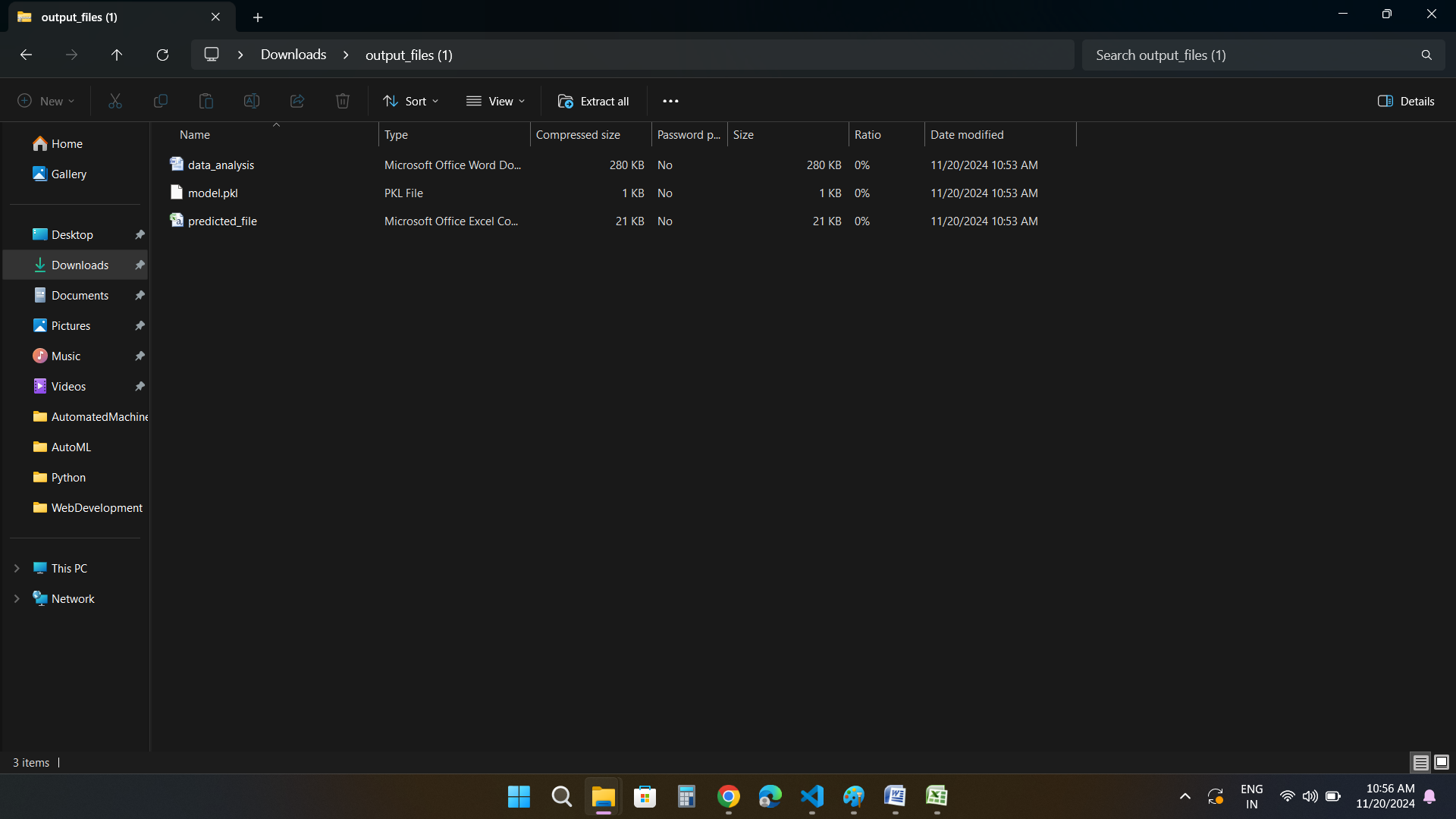
Here, the user shall upload the file either by dropping it in the dropzone or by selecting it from local files.

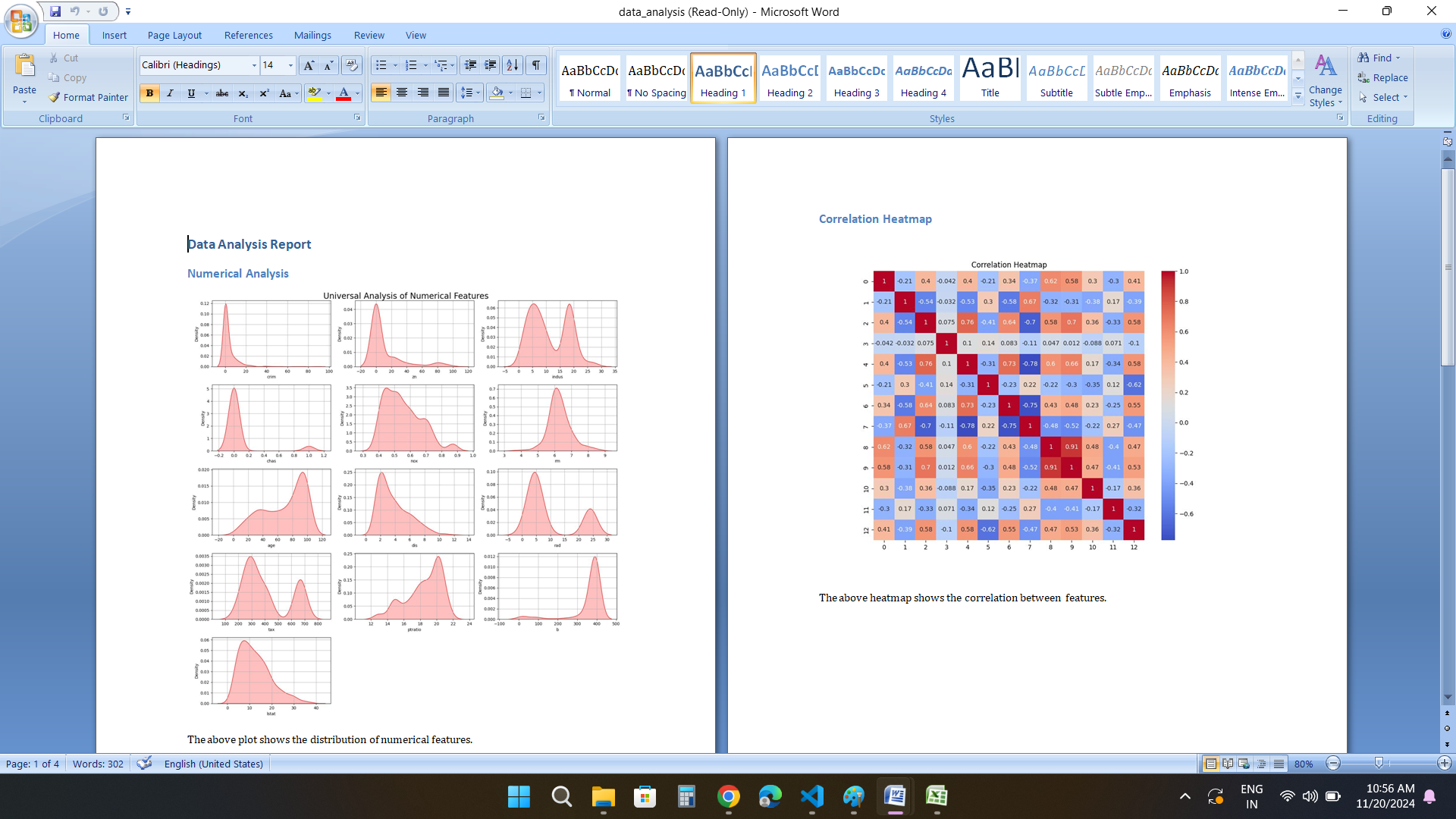
The uploaded file is then preprocessed and analysis is performed.

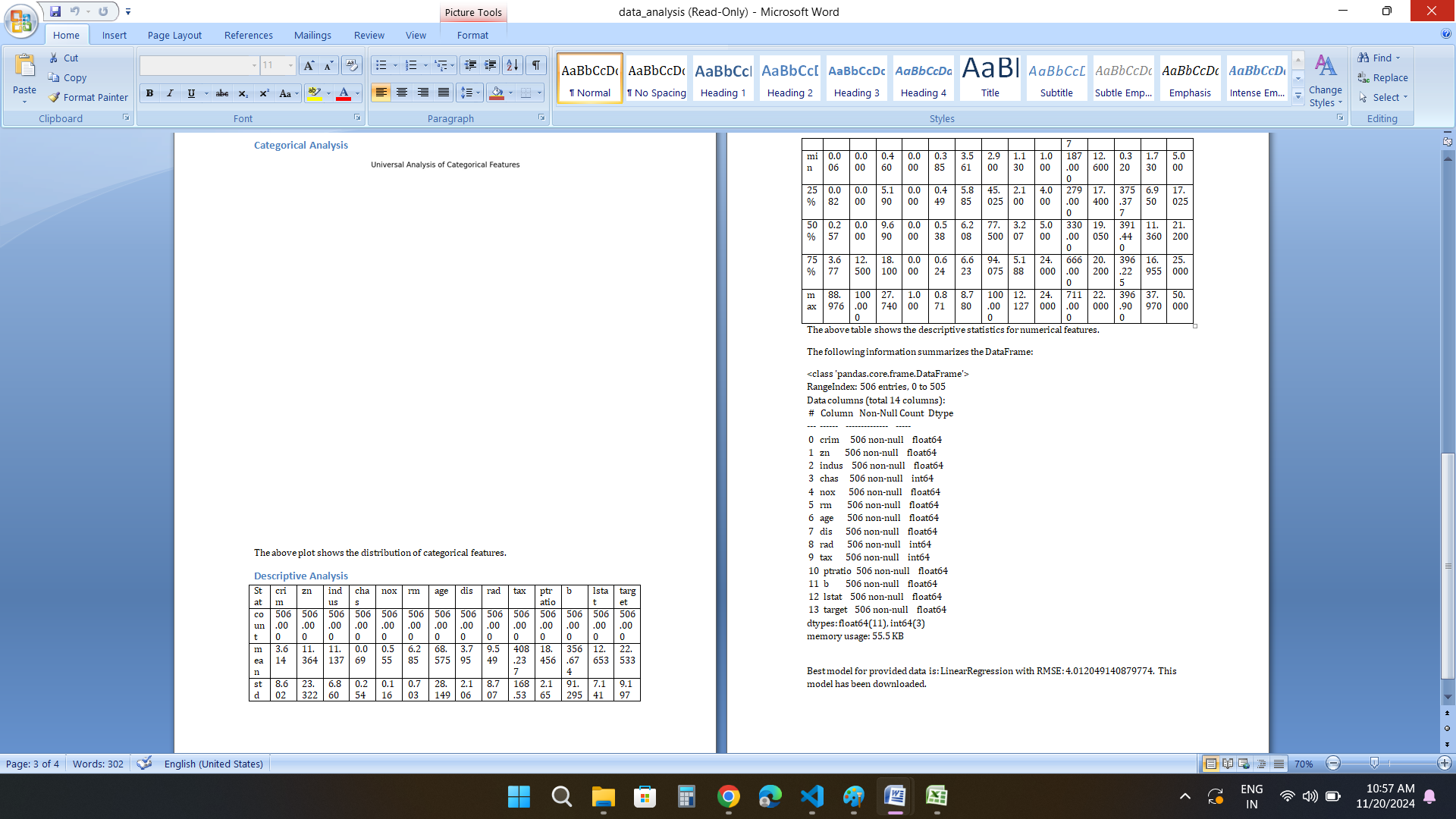
12 Classification models are trained and the one with highest accuracy is then selected.

A zip file is downloaded, which contains data analysis, predicted file and the trained model.

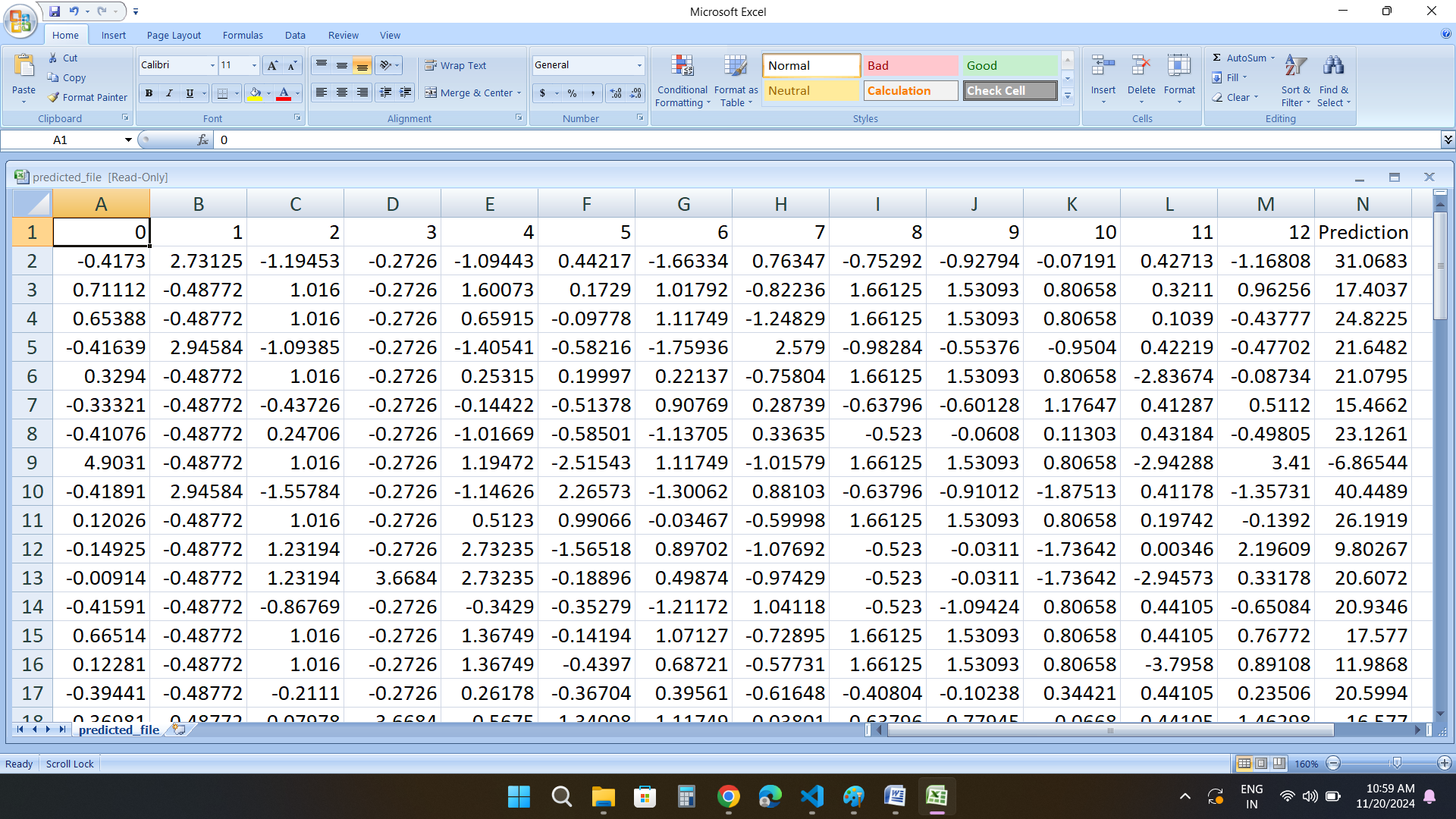






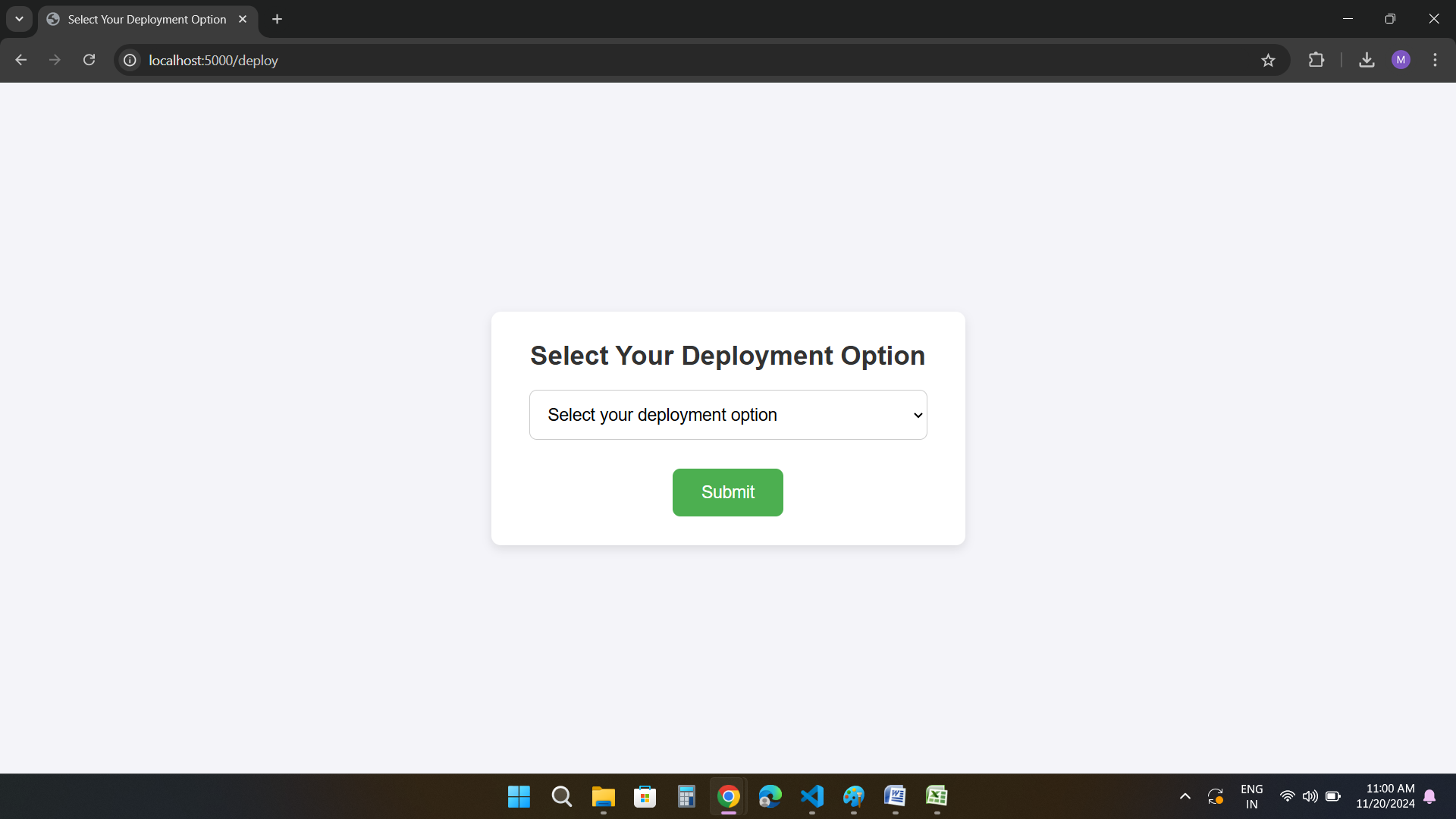


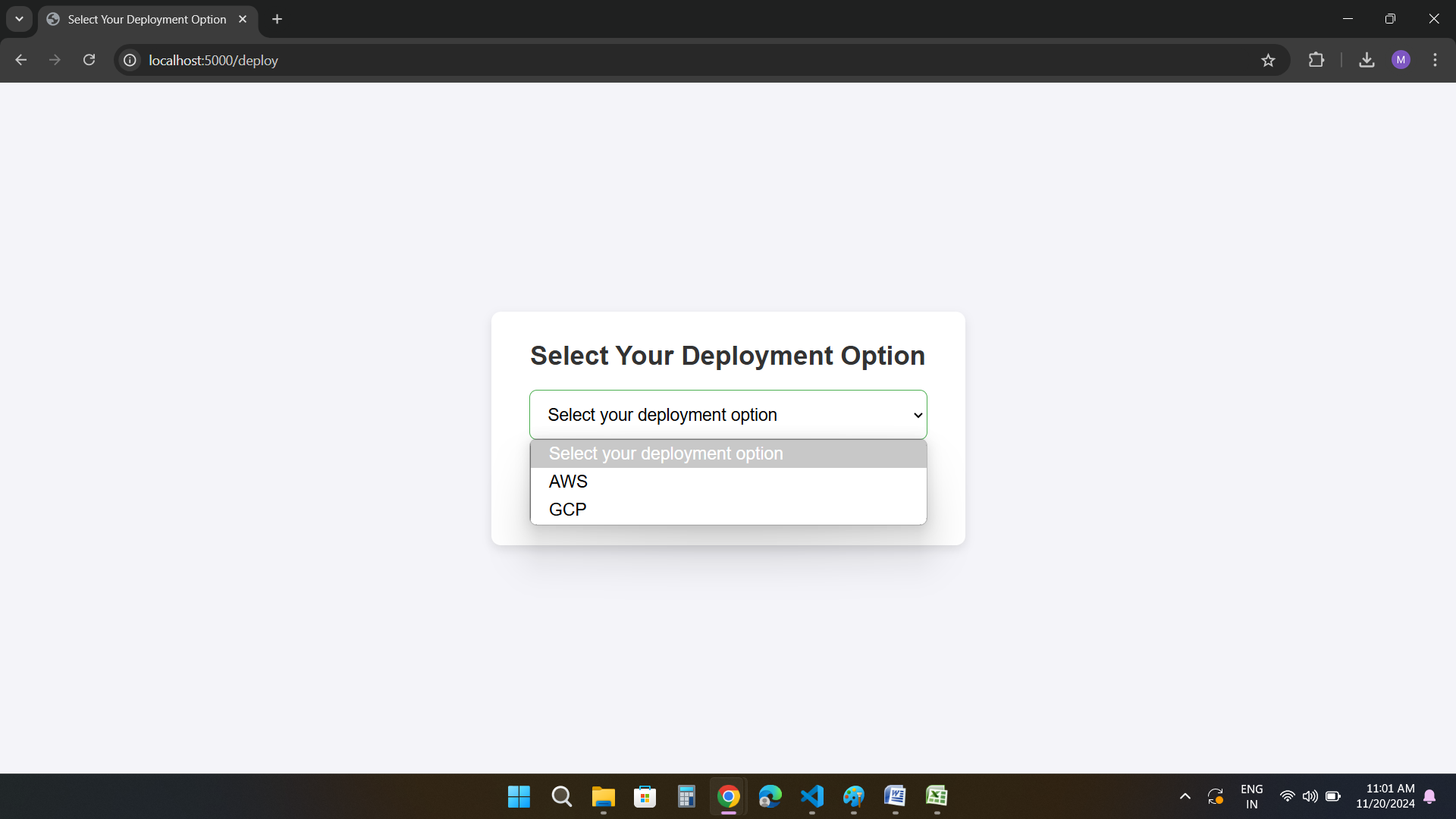
No categorical plots are created as there are no categorical columns.



Above is the predicted file.

/deploy





**The code has been tested on the following datasets:**

**Classification problems:**

1. Loan approval: <https://www.kaggle.com/datasets/taweilo/loan-approval-classification-data>
2. Iris dataset: <https://www.kaggle.com/datasets/uciml/iris>
3. Diabetes dataset: <https://www.kaggle.com/datasets/hasibur013/diabetes-dataset>
4. Mountains vs. Beaches: <https://www.kaggle.com/datasets/jahnavipaliwal/mountains-vs-beaches-preference>

## Stroke Prediction Dataset: <https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>

1. Titanic Dataset: <https://www.kaggle.com/datasets/yasserh/titanic-dataset>

## Fitness Club Dataset for ML Classification: <https://www.kaggle.com/datasets/ddosad/datacamps-data-science-associate-certification>

## Star Type Classification / NASA: <https://www.kaggle.com/datasets/brsdincer/star-type-classification>

**Regression problems:**

## Laptop Price - dataset: <https://www.kaggle.com/datasets/ironwolf437/laptop-price-dataset>

## Student Performance (Multiple Linear Regression): <https://www.kaggle.com/datasets/nikhil7280/student-performance-multiple-linear-regression>

## Walmart Sales: <https://www.kaggle.com/datasets/mikhail1681/walmart-sales>

## Boston housing dataset: <https://www.kaggle.com/datasets/altavish/boston-housing-dataset>

## Medical Cost Personal Datasets**:** <https://www.kaggle.com/datasets/mirichoi0218/insurance>

The end.