Architecture

News Article Sorting

Written By	Ashwani Devi
Document Version	1.0
Last Revised Date	

Document Version Control

Change Record:

Version	Date	Author	Comments
1.0	26-06-2023	Ashwani Devi	Architecture

Reviews:

Version	Date	Reviewer	Comments

Approval Status:

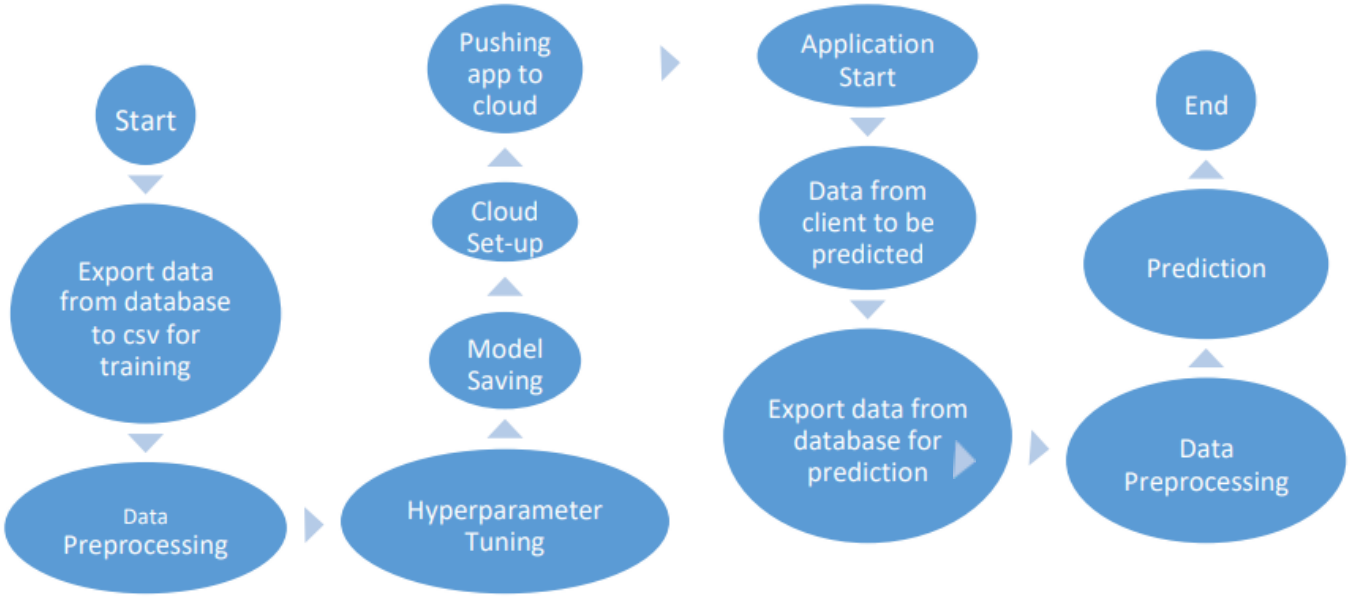
Version	Review Date	Reviewed By	Approved By	Comments

Contents

Document Version Control 2

1. Architecture 4
2. Architecture Description 5
2.1 Data Description
5 2.2 Export Data from DB to CSV for Training 5
2.3 Data Preprocessing 5
2.4 Data Preprocessing 5
2.5 Hyperparameter Tuning 5
2.6 Cloud Set-up 5
2.7 Push App to Cloud 5
2.8 Data from Client Side for Prediction Purpose 6

2. Architecture



3.Architecture Description

3.1 Data Description

We will be using the BBC News Classification Data Set present at Kaggle website. This Data set is satisfying our data requirement. Total 2225 instances present in different batches of data.

3.2 Export Data from database to CSV for Training

Here we will be exporting all batches of data from the database into one csv file for training.

3.3 Data Preprocessing

We will be exploring our data set here and do EDA if required and perform data preprocessing depending on the data set. We first explore our data set in Jupyter Notebook and decide what pre-processing and Validation we have to do such as imputation of null values, dropping some column, etc and then we have to write separate modules according to our analysis, so that we can implement that for training as well as prediction data.

3.4Hyperparameter Tuning

Now, we will do hyperparameter tuning for all the models and try to increase the performance of the model.

3.5 Model Saving

After performing hyperparameter tuning for models, we will save our models so that we can use them for prediction purposes.

3.6 Cloud Setup

Here we will do cloud setup for model deployment. Here we will also create our flask app and user interface and integrate our model with flask app and UI.

3.7 Push app to cloud

After doing cloud setup and checking the app locally, we will push our app to the cloud to start the application.

3.8 Data from client side for prediction purpose

Now our application on cloud is ready for doing prediction. The prediction data which we receive from client side will be exported from DB and further will do same data cleansing process as we have done for training data using modules we will write for training data. Client data will also go along the same process of **Exporting data from DB, Data pre-processing, Data clustering** and according to each cluster number we will use our **saved model** for prediction on that cluster.