***Model Deployment :***

Model deployment can be defined as the model that is kept in a production environment or a server where it takes input from the user and gives output in real-time. Suppose you have to build a model that predicts whether a customer will churn or not from a telecom company.

The model will be trained on several features. The model will be able to make predictions in real-time when you provide input of these fields. You have to give entries of features on which the model is trained. It would then be able to make predictions.

There are mainly ***two different models*** of model deployment:

* Batch Mode
* Real-time Mode

When we upload a .csv or any tabular structure, we predict with multiple values. This means that the values are predicted in a batch instead of one value at a time. This is called Batch Mode where predictions are made by the model in different batches.

In Online or real-time prediction, one-by-one approach is followed.

To deploy the model in production, we first need to ***save our model***.

When we build the model on our local system and make predictions, the model gives prediction. However, once we close the python file, everything gets destroyed.

So, it becomes important to save the model so that we do not need to re-train the model.

This is called ***Pickling or Serialization*** in python. Model saving is important for both the modes of model deployment.

To save your model, we will make use of a pickle library that allows you to save and load your model.

We can use HeroKu,AWS,Google cloud,Python aNywhere,Kubernetes Engine,etc for deployment purpose

***Instructions for deployment:***

1. Upload all the python code needed for deployment in Github.(including Procfile)
2. Login to Heroku ,create new app
3. In deployment method,we need to connect to Github and enter the data stored repository name .
4. We can choose between automatic and Manual deployment.
5. Link will be created.Using the link we can access to the model and start prediction.