

### 3. SYSTEM ARCHITECTURE

Following is the system architecture:-

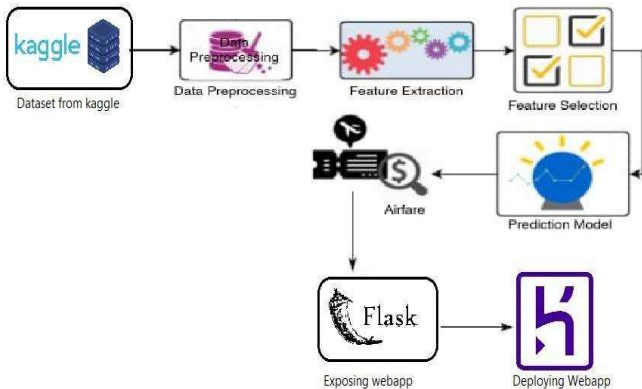


Chart -1: Architecture

### 4. ALGORITHM

**Step1: Downloading Dataset from Kaggle :** Kaggle lets in customers to discover and put up data sets, discover and construct fashions in a web-primarily based totally facts-technology data science environment, paintings with different facts scientists and device mastering engineers, and input competitions to resolve data science challenges.

**Step2: Data Processing:** Data preprocessing is a data mining technique. It is used to convert the raw records in a beneficial and useful format. In the dataset, many attributes include the identical information. Directly merging the tables creates many replica fields. Also, the records pronounced via way of means of the airways can also additionally consist of faulty values because of human error, foreign money conversion error, etc. Hence, as it should be designed records preprocessing workflow is vital to generate correct enter records for you to build the machine learning model.

**Step3: Feature Extraction:** Feature Extraction goals to lessen the wide variety of functions or features in a dataset through developing new functions from the prevailing ones (after which discarding the unique functions). These new decreased set of functions must then be capable of summarize maximum of the statistics contained withinside the unique set of functions.

**Step4: Feature Selection:** Feature selection is the system of decreasing the wide variety of input variables while making a predictive model. It is ideal to lessen the wide variety of input variables to each lessen the computational value of modeling and, in a few cases, to enhance the overall performance of the model.

**Step5: Predicting model:** Algorithm used in our model is Random forest. Random forest is supervised machine learning algorithm. It is a collection of multiple decision trees whose results are aggregated into one final result. As

the name suggests, "Random Forest is a classifier that contains a multiple number of decision trees on various subsets of the given dataset and takes the average to improve the predictive or final accuracy of that dataset.

### 5. CONCLUSIONS

This paper reported on a preliminary study in "airfare prices prediction". We gathered airfare data from a Kaggle website and showed that it is feasible to predict prices for flights based on historical fare data. The experimental results show that ML models are a satisfactory tool for predicting airfare prices. Other important factors in airfare prediction are the data collection and feature selection from which we drew some useful conclusions. From the experiments we concluded which features influence the airfare prediction at most.

### ACKNOWLEDGEMENT

To complete any type of project work is team work. It involves all the technical/nontechnical expertise from various sources. The contribution from the experts in the form of knows-how and other technical support is of its vital importance. We are indebted to our inspiring guide **Prof. K.V. Patil** our project coordinator.

We are also very thankful to **Prof. Dr. D. A. Godse** for her valuable guidance.

We have great pleasure in offering big thanks to our honorable Principal **Prof. Dr.S. R. Patil**. Last but not least, we would like to thank all the direct and indirect help provided by all the staff and our entire class for successful completion of this project.

### REFERENCES

- [1] Article – Moez Ali. "Deploying machine pipeline on the cloud using docker," <https://towardsdatascience.com/deploy-machine-learning-pipeline-on-cloud-using-docker-container-bec64458dc01>
- [2] Books-Bachir Chihani, Ph.D and Rafael Benevides, docker CLI & Dockerfile Cheat Sheet.
- [3] <https://analyticsindiamag.com> (22/07/2019). Why Use Docker In Machine Learning? We Explain With Use Case
- [4] <https://www.analyticsvidhya.com/blog/2021/04/rapid-fire-eda-process-using-python-for-ml-implementation/>. Rapid-Fire EDA process using Python for ML Implementation
- [5] <https://www.javatpoint.com/machine-learning-random-forest-algorithm> Random Forest Algorithm