HackerEarth Machine Learning challenge: Adopt a buddy

Steps that I have taken are as follows :

1. Get the Data from the source.

2. Exploratory Data Analysis on the data

3. Handling missing data

4. Feature Engineering - creating new features

5. Preparing the data to be feed in ML

6. Handling categorical data

7. Scaling the data for better ML models performance

8. Training ML model

9. Evaluating model

10. Predicting the values from model

11. Exporting the data to CSV

Data Exploration

Data exploration is one the major part of analysis the data and the correlations between data. One of the major problem I faced during all my analysis is the "missing values". There were a lot of missing values in the "Condition" column, handling the missing data was very important in this column as this column was highly correlated with the target labels.

Feature Engineering

Feature Engineering is one the major part of analysing and preparing the data for ML model. As the dataset I was provided had a limited features, it was really important to create new features from the given features to get a good score. Created multiple new features from "issue_data" and "listing_date" like month, days, difference in these date, the time it took. Modified "pet_id" and extract important features. As the machine only understands only a numerical values it was important to use Label Encoder to encode the categories into numerical values. Scaling the data is always a good practise as it improves the ML models performance especially where "Euclidean distance" is being used.

Machine Learning Models

I tried many different models in this challenge like Decision Trees, Random Forest, K-Nearest Neighbours etc. I found **XBGClassifier** was scoring much more than the other ML models. After tuning the Hyper parameters of the model I successfully scored **90.92760**. For evaluating the model I used the Evaluation metric provided in the problem description.

Technology used

Numpy (For doing linear algebra), Pandas (For data manipulation), Scikit-Learns(For implementing ML Algorithms)