|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| id | filename | amount | fee | transaction Type | has Error |
| 1 | file1 | 1000 | 100 | capture | 0 |
| 2 | file2 | 100 | 10 | capture | 0 |
| 3 | file3 | 0 | 100 | capture | 1 |
| 4 | file4 | 100 | 10 | refund | 1 |
| 5 | file5 | “-100” | 0 | refund | 1 |
| 6 | file6 | “100” | 0 | capture | 1 |
| 7 | file7 | -100 | 0 | refund | 0 |

We want to create a model that can detect files with error and files that can be accepted.

Based on the above data set please answer the following questions

1. Which classification algorithm you suggest and why?

2. Which features would you suggest to create the model and train it?

1. I will prefer to do it by Binary Classification model because binary classification tasks that are involve “0” and “1”. When the state found 0, it will not accept the file where as it found 1 then the file will be accepted.
2. I will suggest the target feature will be ‘has Error’ and the level ‘filename’. Depend on my target value output the level will accept or reject. So the ‘has Error’ will be the train model what is my understanding.

Once it is done then we can train another model based on the target column ‘transaction type’. It is also a binary model because it has two states.