Multilabel Classification

February 13, 2022

1 Multilabel Classification

```python
[1]: import pandas as pd
     from pandas import DataFrame
     from sklearn.model_selection import train_test_split
     from sklearn.ensemble import RandomForestClassifier
     from sklearn.metrics import accuracy_score
```

```python
[2]: df_dataset = pd.read_excel("iris.xlsx")
```

```python
[3]: df_dataset
```

```
[3]:        Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm  \
     0       1            5.1           3.5            1.4           0.2
     1       2            4.9           3.0            1.4           0.2
     2       3            4.7           3.2            1.3           0.2
     3       4            4.6           3.1            1.5           0.2
     4       5            5.0           3.6            1.4           0.2
     ..    ...            ...           ...            ...           ...
     145   146            6.7           3.0            5.2           2.3
     146   147            6.3           2.5            5.0           1.9
     147   148            6.5           3.0            5.2           2.0
     148   149            6.2           3.4            5.4           2.3
     149   150            5.9           3.0            5.1           1.8

                 Species
     0       Iris-setosa
     1       Iris-setosa
     2       Iris-setosa
     3       Iris-setosa
     4       Iris-setosa
     ..              ...
     145  Iris-virginica
     146  Iris-virginica
     147  Iris-virginica
     148  Iris-virginica
     149  Iris-virginica
```

```
[150 rows x 6 columns]
```

```
[4]: y = df_dataset["Species"]
     X = df_dataset.drop(["Species"], axis = 1)
```

```
[5]: X
```

```
[5]:         Id  SepalLengthCm  SepalWidthCm  PetalLengthCm  PetalWidthCm
     0        1            5.1           3.5            1.4           0.2
     1        2            4.9           3.0            1.4           0.2
     2        3            4.7           3.2            1.3           0.2
     3        4            4.6           3.1            1.5           0.2
     4        5            5.0           3.6            1.4           0.2
     ..     ...            ...           ...            ...           ...
     145    146            6.7           3.0            5.2           2.3
     146    147            6.3           2.5            5.0           1.9
     147    148            6.5           3.0            5.2           2.0
     148    149            6.2           3.4            5.4           2.3
     149    150            5.9           3.0            5.1           1.8

     [150 rows x 5 columns]
```

```
[6]: y
```

```
[6]: 0          Iris-setosa
     1          Iris-setosa
     2          Iris-setosa
     3          Iris-setosa
     4          Iris-setosa
                   ...
     145     Iris-virginica
     146     Iris-virginica
     147     Iris-virginica
     148     Iris-virginica
     149     Iris-virginica
     Name: Species, Length: 150, dtype: object
```

```
[7]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.40,␣
     ↪random_state=35)
```

```
[8]: len(X_train)
```

```
[8]: 90
```

```
[9]: len(y_train)
```

```
[9]: 90
```

```
[10]: len(X_test)
```

[10]: 60

```
[11]: len(y_test)
```

[11]: 60

```
[12]: rf_model = RandomForestClassifier().fit(X_train, y_train)
```

```
[13]: y_pred = rf_model.predict(X_test)
```

```
[14]: y_pred
```

```
[14]: array(['Iris-versicolor', 'Iris-versicolor', 'Iris-virginica',
              'Iris-versicolor', 'Iris-setosa', 'Iris-virginica',
              'Iris-virginica', 'Iris-versicolor', 'Iris-versicolor',
              'Iris-setosa', 'Iris-versicolor', 'Iris-virginica', 'Iris-setosa',
              'Iris-virginica', 'Iris-setosa', 'Iris-virginica',
              'Iris-versicolor', 'Iris-setosa', 'Iris-setosa', 'Iris-setosa',
              'Iris-versicolor', 'Iris-versicolor', 'Iris-virginica',
              'Iris-versicolor', 'Iris-setosa', 'Iris-setosa', 'Iris-setosa',
              'Iris-virginica', 'Iris-setosa', 'Iris-virginica', 'Iris-setosa',
              'Iris-versicolor', 'Iris-virginica', 'Iris-setosa',
              'Iris-versicolor', 'Iris-virginica', 'Iris-setosa',
              'Iris-virginica', 'Iris-virginica', 'Iris-virginica',
              'Iris-versicolor', 'Iris-setosa', 'Iris-setosa', 'Iris-versicolor',
              'Iris-virginica', 'Iris-virginica', 'Iris-setosa',
              'Iris-virginica', 'Iris-virginica', 'Iris-versicolor',
              'Iris-virginica', 'Iris-versicolor', 'Iris-versicolor',
              'Iris-versicolor', 'Iris-setosa', 'Iris-versicolor',
              'Iris-versicolor', 'Iris-versicolor', 'Iris-setosa',
              'Iris-versicolor'], dtype=object)
```

```
[15]: accuracy_score(y_test, y_pred)
```

[15]: 1.0