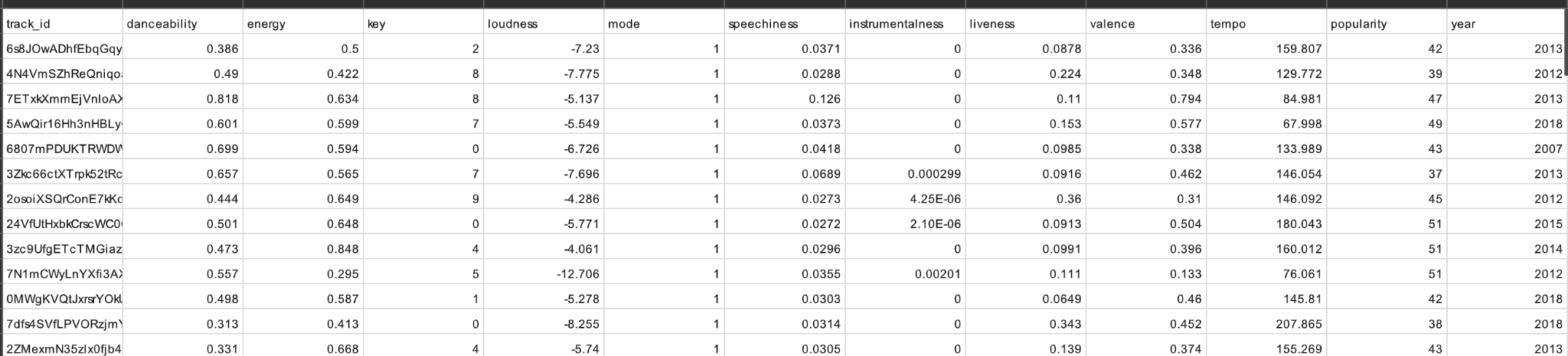
# ข้อมูลทั่วไปของข้อมูลทั้งหมด



* ตัวแปร (feature) ทั้งหมด = 10 ตัว ได้แก่

'danceability', 'energy', 'key', 'loudness', 'mode', 'speechiness', 'instrumentalness', 'liveness', 'valence', 'tempo'

Feature ที่ยังไม่ได้ใช้ = ‘year’ และ ‘popularity’

* ตัวแปรตาม (target label) ทั้งหมด = 10 ตัว ได้แก่

| ชื่อ label | ชื่อ dataframe | name |
| --- | --- | --- |
| missed | df\_miss | ความคิดถึง |
| sad | df\_sad | 10ปีกับรักแสนเศร้า |
| cry | df\_cry | ร้องไห้หนักมาก |
| discourage | df\_dis | เป็นท้อ |
| love | df\_love | เพลงรักไม่เคยลืม |
| working | df\_work | ฟังเพลินตอนทำงาน |
| sea | df\_sea | อารมณ์ทะเล |
| chillout | df\_chill | เพลงไทยสายชิลล์ |
| tired | df\_tired | เหนื่อยนักพักก่อน |
| chillhop | df\_hop | ชิลล์ฮอป |

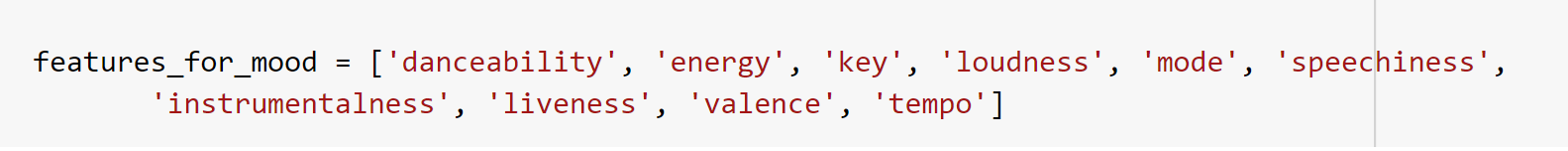
# รายงานผลการทำ classification

1. Data

1.1 ตัวแปรที่ใช้ (feature)

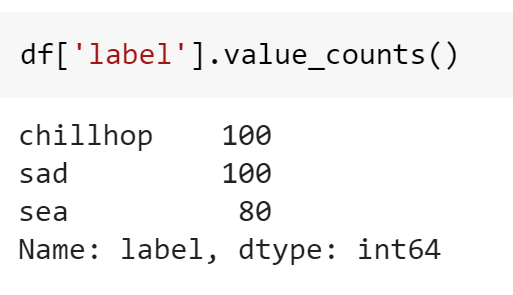
*List ของตัวแปรที่เราสามารถเอามาใช้ได้มีดังนี้ ทั้งนี้นิสิตสามารถปรับลดได้ เพื่อให้ได้ accuracy ที่มากที่สุด*

*ถ้านิสิตใช้ตัวแปรที่ต่างจากนี้ ----- โปรดระบุ-------*

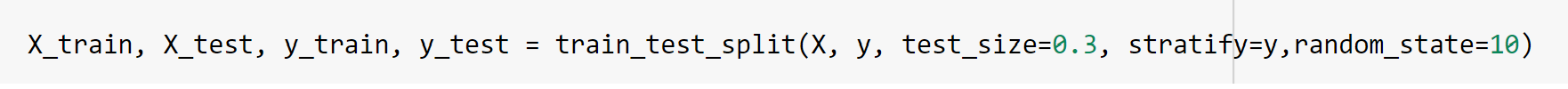


1.2 ตัวแปรตาม (target label) ที่ใช้

*ถ้านิสิตใช้ตัวแปรที่ต่างจากนี้ ----- โปรดระบุ-------*



1.3 การแบ่ง train/test



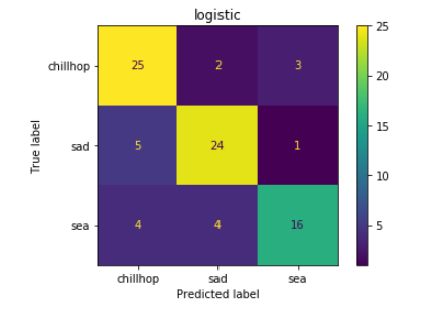
| data | จำนวน data |
| --- | --- |
| train | 196 |
| test | 84 |

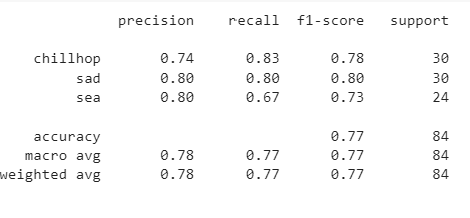
1. Model ที่เลือกใช้ และผลที่ได้

*โมเดลที่อาจารย์ใช้ใน code ที่ให้ไปนั้นสามารถปรับเปลี่ยนได้ เพื่อ accuaracy สูงสุด*

2.1 LogisticRegression

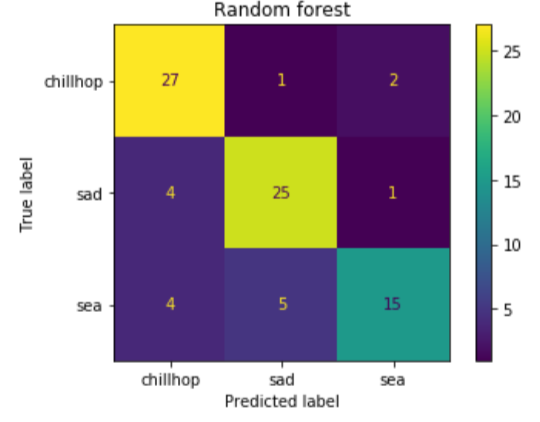
| Parameter ในโมเดล | ค่าที่ใช้ |
| --- | --- |
| multi\_class | ovr |
| max\_iter | 500 |
| C | 1 |
| penalty | l1 |
| solver | liblinear |

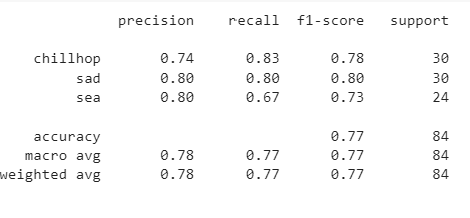




2.2 RandomForest

| Parameter ในโมเดล | ค่าที่ใช้ |
| --- | --- |
| criterion | gini |
| max\_features | log2 |
| n\_estimators | 500 |





2.3 xgboost

| Parameter ในโมเดล | ค่าที่ใช้ |
| --- | --- |
| objective | multi:softmax |

