

```
import pandas as pd
import re
```

```
twitter_data = pd.read_csv("train.csv")
print(twitter_data.head())
```

	id	label	tweet
0	1	0	@user when a father is dysfunctional and is s...
1	2	0	@user @user thanks for #lyft credit i can't us...
2	3	0	bihday your majesty
3	4	0	#model i love u take with u all the time in ...
4	5	0	factsguide: society now #motivation

```
twitter_data.drop(['id'], axis=1, inplace=True)
```

```
twitter_data.head()
```

	label	tweet
0	0	@user when a father is dysfunctional and is s...
1	0	@user @user thanks for #lyft credit i can't us...
2	0	bihday your majesty
3	0	#model i love u take with u all the time in ...
4	0	factsguide: society now #motivation

```
import nltk
from nltk.stem import PorterStemmer
```

```
stemmer = PorterStemmer()
def clean_sentences(text):
    text = text.lower()
    text = re.sub(r"[^a-z0-9^!\.\/]", " ", text)
    text = " ".join(text.split())
    text = " ".join(stemmer.stem(word) for word in text.split())
    return text
```

```
x = twitter_data['tweet']
y = twitter_data['label']
```

```
x = x.map(lambda a:clean_sentences(a))
```

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,stratify=y,random_state=42)
```

```
from sklearn.feature_extraction.text import CountVectorizer
vectorizer = CountVectorizer()
vectorizer.fit(x_train)
x_train = vectorizer.transform(x_train)
x_test = vectorizer.transform(x_test)
```

```
from sklearn.svm import LinearSVC
from sklearn.linear_model import LogisticRegression
model = LogisticRegression(max_iter=1000)
model.fit(x_train, y_train)
```

```
LogisticRegression(max_iter=1000)
```

```
from sklearn.metrics import accuracy_score, precision_score, confusion_matrix
print("Accuracy : ",accuracy_score(y_test, model.predict(x_test)))
print("Precision : ",precision_score(y_test, model.predict(x_test)))
print(confusion_matrix(y_test, model.predict(x_test)))
```

```
Accuracy : 0.9593292454010762
Precision : 0.8333333333333334
[[7371  59]
 [ 266 295]]
```

✓ 0s completed at 12:25

