```
import numpy as np
import pandas as pd
from sklearn import preprocessing
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split
import requests
import io
#from google.colab import drive
#drive.mount('/content')
#from google.colab import files
#uploaded = files.upload()

#url = "https://www.kaggle.com/xwolf12/datasetandroidpermissions/download/b3nIEWm6zcroShVi
#s = requests.get(url).content
#df = pd.read_csv(io.StringIO(s.decode('utf-8')), error_bad_lines=False)

df = pd.read_csv('/train.csv', sep=';')
# df.describe()
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 398 entries, 0 to 397
Columns: 331 entries, android to type
dtypes: int64(331)
memory usage: 1.0 MB
```

```
X = np.array(df.drop(columns=['android', 'type']))
Y = np.array(df['type'])

X_train, X_test, Y_train, Y_test = train_test_split(X,Y,test_size=0.2)
clf = KNeighborsClassifier()
clf.fit(X_train, Y_train)

accuracy = clf.score(X_test, Y_test)
print(accuracy)
```

```
0.95
```