

## Contents

1 ex 1

1

1 ex 1

```
import pickle
from sklearn.feature_extraction.text import HashingVectorizer
from sklearn.naive_bayes import BernoulliNB
from sklearn.metrics import accuracy_score
from sklearn.tree import DecisionTreeClassifier

data = pickle.load(open("sklearn-data.pickle", "rb"))

vectorizer = HashingVectorizer(stop_words="english", n_features=2**6, binary=True)
x_train = vectorizer.fit_transform(data["x_train"])
x_test = vectorizer.fit_transform(data["x_test"])

classifier = BernoulliNB()
classifier.fit(x_train, data["y_train"])
y_pred = classifier.predict(x_test)

dec_tree_classifier = DecisionTreeClassifier()
dec_tree_classifier.fit(x_train, data["y_train"])
dec_tree_pred = dec_tree_y_pred = dec_tree_classifier.predict(x_test)

print("BernoulliNB accuracy:" ,accuracy_score(data["y_test"], y_pred))
print("decision tree accuracy:" ,accuracy_score(data["y_test"], dec_tree_y_pred))

# print(data["x_train"])
```

This gives an accuracy of about 70%

Dropped ex 2 in favour of staring at the wall during lockdown, but figured I might as well upload what i have