In [ ]:

```
from sklearn.feature_extraction.text import TfidfVectorizer
```

In [ ]:

```
from sklearn.feature_extraction.text import TfidfTransformer
dataset = [
    "I enjoy reading about Machine Learning and Machine Learning is my PhD subject",
    "I would enjoy a walk in the park",
    "I was reading in the library"
]
```

In [ ]:

```
import pandas as pd
tfIdfVectorizer=TfidfVectorizer(use_idf=True)
tfIdf = tfIdfVectorizer.fit_transform(dataset)
df = pd.DataFrame(tfIdf[0].T.todense(), index=tfIdfVectorizer.get_feature_names(), columns=
df = df.sort_values('TF-IDF', ascending=False)
print (df.head(25))
```

In [ ]:

```
docA = "The car is driven on the road"
docB = "The truck is driven on the highway"

tfidf = TfidfVectorizer()

response = tfidf.fit_transform([docA, docB])

feature_names = tfidf.get_feature_names()
for col in response.nonzero()[1]:
    print (feature_names[col], ' - ', response[0, col])
```

In [ ]:

```

```