

Lab - 2

Use an appropriate data set for building the decision tree (ED3) and apply this knowledge to classify a new sample

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier
export_text, plot_tree
```

```
data = pd.read_csv("tenis.csv")
```

```
from sklearn.preprocessing import LabelEncoder
encoder = LabelEncoder()
```

```
for col in data.columns:
```

```
    data[col] = encoder.fit_transform(data[col])
```

```
X = data.drop(columns=['play'])
```

```
y = data['play']
```

```
clf = DecisionTreeClassifier(criterion='entropy')
```

```
clf.fit(X, y)
```

```
print(export_text(clf, feature_names=list(X.columns)))
```

```
plt.figure(figsize=(12, 8))
```

```
plot_tree(clf, feature_names=X.columns,
```

```
        class_names=encoder.classes_, filled=True,
```

```
        rounded=True, fontsize=12, proportion=False)
plt.show()
```

```

sample = pd.DataFrame([2, 1, 0, 0], columns =
    ['outlook', 'temp', 'humidity', 'windy'])
prediction = dt.predict(sample)
print("Prediction for new sample: ", encoder.inverse_
    transform([prediction[0]]))

```

Output:

prediction for new sample: No

Outlook ≤ 0.50

class: 1

Outlook > 0.50

humidity ≤ 0.50

Outlook ≤ 1.50

windy ≤ 0.50

class: 1

windy > 0.50

class: 0 → Answer

Outlook > 1.50

class: 0

humidity > 0.50

windy ≤ 0.50

class: 1

windy > 0.50

temp ≤ 1.0

class: 0

temp > 1.00

class: 1

~~1/2/3/4/5~~