12/3/25

LAB - 2

Use an appropriate dataset for building the decision tree (ID3) and apply this knowledge to classify a new sample.

```
import pandas as pd
import numpy as np
from collections import Counter
import math

def entropy(y):
    counts = Counter(y)
    probabilities = [count/len(y) for count in
                                  counts.values()]
    return -sum(p*math.log2(p) for p in probabilities)

def information_gain(data, feature, target):
    total_entropy = entropy(data[target])
    values = data[feature].unique()
    weighted_entropy = sum((len(data[data[feature]==
    v])/len(data))* entropy(data[data[data[feature]==v]
    [target]) for v in values)
    return total_entropy - weighted_entropy

def id3(data, features, target):
    if len(set(data[target]))==1:
        return data[target].iloc[0]
    if len(features)==0:
        return data[target].mode()[0]
    gains = {feature: information_gain(data, feature,
            target) for feature in features}
    best_feature = max(gains, key=gains.get)
```

```
        tree = {best_feature: {}}
        for value in data[best_feature].unique():
            subset = data[data[best_feature] == value]
            remaining_features = [f for f in features if
                                  f != best_feature]
            tree[best_feature][value] = id3(subset, remaining_
                                  features, target)
    return tree


def print_tree(tree, indent = ""):
    if not isinstance(tree, dict):
        print(indent + "→" + str(tree))
        return
    for key, value in tree.items():
        print(indent + str(key))
        for sub_key, sub_tree in value.items():
            print(indent + "L " + str(sub_key))
            print_tree(sub_tree, indent + "  ")
#Load Dataset
file_path = '/content/tennis.csv'
data = pd.read_csv(file_path)

#Apply ID3 Algorithm
features = list(data.columns[:-1])
target = 'play'
decision_tree = id3(data, features, target)
print_tree(decision_tree)
```

Output:

```
outlook
├── sunny
│     └ humidity
│           ├── high : no
│           └── normal : yes
├── overcast : yes
└── rainy
      └ windy
            ├── False : yes
            └── True : no
```

12/3/25