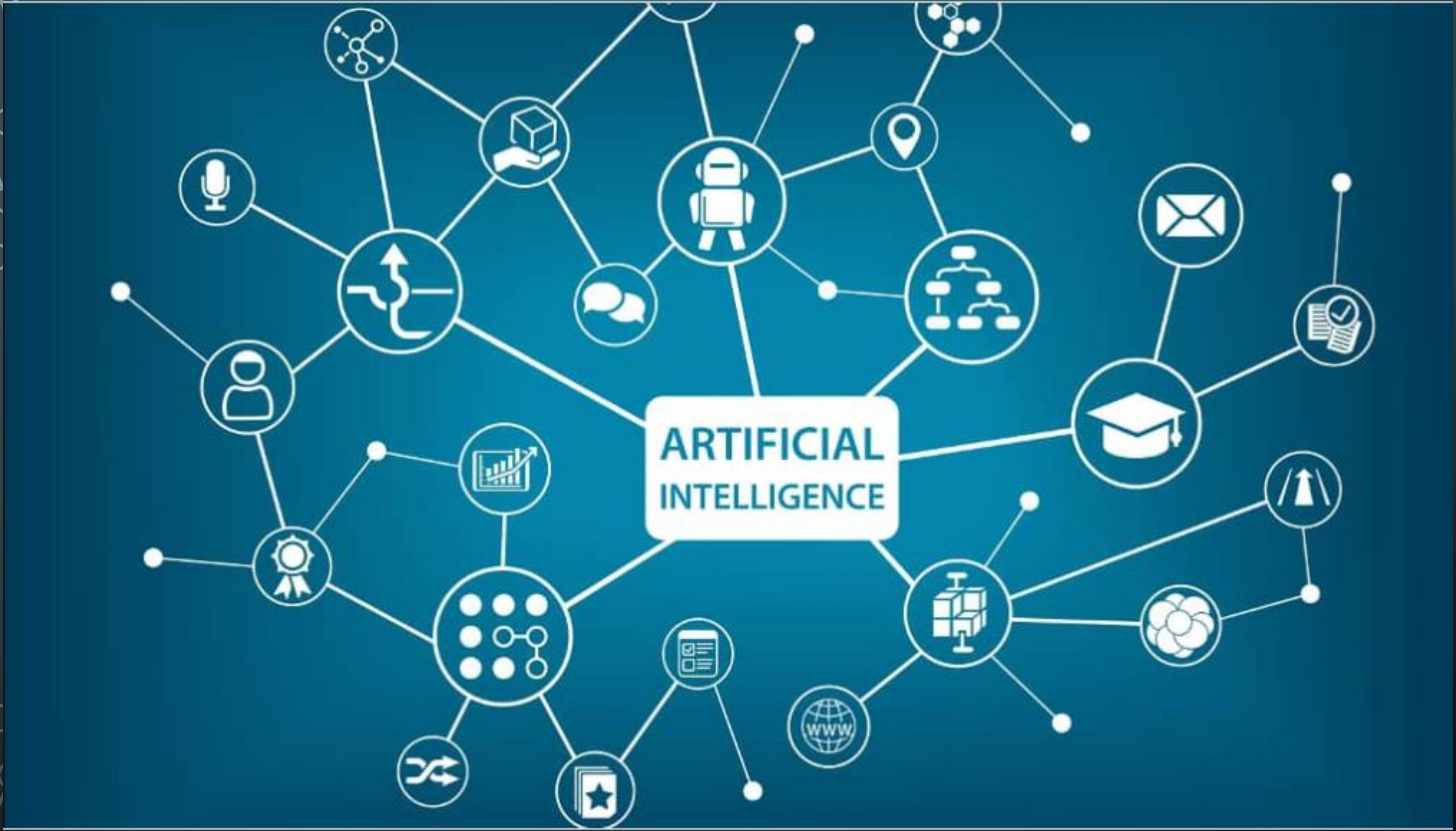


DECISION TREES

AGENDA

I. INTRODUTION TO DECISION TREE
II. WHAT IS DECISION TREE
III. DECISION TREE LEARNING ALGORITHM
IV. EXAMPLES OF DECISION TREE
V. AVANTAGE AND DESAVANTAGE
VI. CONCLUSION

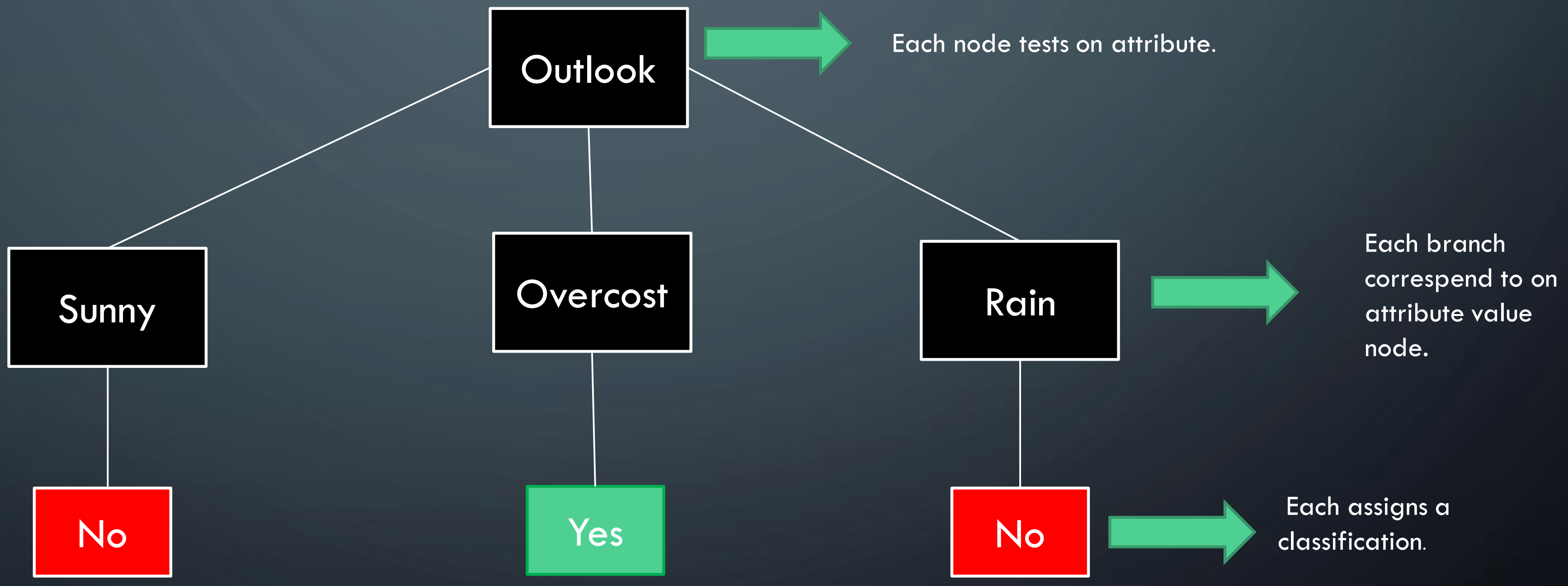
I. INTRODUCTION TO DECISION TREE

- A decision tree is a series of nodes , a directional graph that starts at the base with a single node and extends to the many leaf that represent the categories that the tree can classify.
- Another way to think of a decision tree is a flow chart , where the flow starts at the root node and ends with a decision made at the leaves .
- It is a decision support, it uses a tree-like graph to show the predication that result from a series of feature-based splits.

II. WHAT IS A DECISION TREE

- ❑A decision tree it represents a fraction that takes as input a vector of attribute values and returns a «decision» that is a single output value.
- ❑Decision tree algorithm falls under the category of supervised learning. The can be used to solve both regression R classification problems .
- ❑A decision tree reaches its classification by performing a sequence of tests .

WHAT IS A DECISION TREE



III. DECISION TREE LEARNING ALGORITHM

- Here we gonna use a ID3(Iterative Dichotomisen 3) is a one of common decision tree algorithm .
- Dichotomisen means dividing into two completely opposite things .
- Algorithm iterativly deivides attributes into two groups which are the most dominant attribute and others to constract a tree.
- Then , it calculates the Entropy & Information Gains of each attribute . In this way the most dominant attribute can be founded.

DECISION TREE LEARNING ALGORITHM

- After then , the most dominant one is put on the tree as decision node.
- Entropy and Gain scores would be calculated again among ohter attributes.
- Porcedure continues until reaching a decision for that branch.
- Calculate the Entropy of every attribute using the data Set(S).

DECISION TREE LEARNING ALGORITHM

- Entropy = $\sum$-P(I).$log_2\, P(I)$
- Split the set S into subsets using the attribute for which the resulting Entropy(after spliting) is minimum (or equivalent information is max) .
- Gains(S,A) = Entropy(S) -$\sum$[P(S/A).Entropy(S/A)]
- Make a decision tree node containing that attribute.
- Recure on subsets using remaining attributes .

IV. EXAMPLES OF DECISION TREE

EXAMPLE : «TO GO FOR OUTİNG OR NOT »

Day	Outlook	Temperature	Humidity	Wind	Decision
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcost	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Weak	Yes
7	Overcost	Cool	Normal	Strong	No
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	No
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcost	Mild	High	Strong	Yes
13	Overcost	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

EXAMPLES OF DECISION TREE

- Caclculate Entropy
- Decision column consist of 14 instances and includes two labels : Yes and No .
- There are 9 decisions labelled yes & 5 decisions labelled no

$$\text{Entropy(Decision)} = -\text{P(yes)}*\log_2 P(yes) - \text{P(no)}*\log_2 P(no)$$

$$\text{Entropy(Decision)} = -(9/14)*\log_2(9/14) - (5/14)*\log_2(5/14)$$

$$=0.940$$

CACLCULATE ENTROPY

- Calculate wind factor of decison (Weak and strong) .
- Gain(D,W) = Entropy(D) -∑[P(D/W)*Entropy(D/W)
- Wind attribute has two labels : weak amd strong .
- Gain(D,W) = Entropy(D)-[P(D/W=weak)*Entopy(D/W=weak]-[p(D/W=strong)*Entropy(D/W=strong)]
- Now we calculate Gain for weak & Strong Wind.

- Weak wind factor on decision
- Wind weak : 8/14 no : 2 , yes : 6
- Entropy(D/W=sweak) = -P(yes)*$\log_2 P(yes)$ –P(no)*$\log_2 P(no)$
- Entropy(D/W=weak) = -(2/8)*$\log_2(2/8)$ –(6/8)*$\log_2(5/8)$

=0,811

- Strong wind factor on decicion
- Strong wind : 6/14 no : 3 , yes : 3
- Entropy(D/W=strong) = -P(yes)*$\log_2 P(yes)$ −P(no)*$\log_2 P(no)$
- Entropy(D/W=strong) = -(3/6)*$\log_2(3/6)$ −(3/6)*$\log_2(3/6)$

  =1
- Gain(D,W) = Entropy(D)-[P(D/W=weak)*Entopy(D/W=weak]-[p(D/W=strong)*Entropy(D/W=strong)]

  =0,940-[(8/14)*0.811]-[(6/14)*1] = 0,048

- High humidity factor on decision
- High humidity : 7/14 no : 4 , yes :3
- Entropy(D/H=high) = -P(yes)*$log_2 P(yes)$ –P(no)*$log_2 P(no)$
- Entropy(D/H=high) = -(4/7)*$log_2(4/7)$ –(3/7)*$log_2(3/7)$

=0,97

- Normal humidity factor on decision
- Nomal humidity : 7/14 no : 1 , yes : 6
- Entropy(D/H=normal) = -P(yes)* $log_2 P(yes)$ –P(no)* $log_2 P(no)$
- Entropy(D/H=normal) = -(1/7)* $log_2(1/7)$ –(6/7)* $log_2(6/7)$

=0.58

- Gain(D,H) = Entropy(D)-[P(D/H=high)*Entopy(D/H=high]-[p(D/H=normal)*Entropy(D/H=normal)]

=0.940-[(7/14)*1)]-[(7/14)*0.58)]

=0,151

- Hot tempereature factor on decision
- Hot temperature : 4/14 no :2 , yes : 2
- Entropy(D/T=hot) = -P(yes)*$log_2 P(yes)$ –P(no)*$log_2 P(no)$
- Entropy(D/T=hot) = -(2/4)*$log_2(2/4)$ –(2/4)*$log_2(2/4)$

=(2/4)*(-1) – (2/4)*(-1)

= 1

- Mild temperature factor on decision
- Mild temperature : 6/14 no : 2 , yes : 4
- Entropy(D/T=mild) = -P(yes)*$log_2 P(yes)$ –P(no)*$log_2 P(no)$
- Entropy(D/T=mild) = -(2/6)*$log_2(2/6)$ –(4/6)*$log_2(4/6)$

=-(2/6)*(-1,58) – (4/6)*(-0,58)

=0,93

- Cool temperature factor on decison
- Cool temperature : 4/14 no : 1 , yes : 3
- Entropy(D/T=cool) = -P(yes)*$log_2 P(yes)$ −P(no)*$log_2 P(no)$
- Entropy(D/T=cool) = -(1/4)*$log_2(1/4)$ −(3/4)*$log_2(3/4)$ =0,815
- Gain(D,H) = Entropy(D)-[P(D/T=hot)*Entopy(D/T=hot]-

  [p(D/T=mild)*Entropy(D/T=mild)] - [p(D/T=cool)*Entropy(D/T=cool)]

  =0.940-[(4/14)*1)]-[(6/14)*0.93)]-[(4/14)*0,815

  =0,032

- Sunny outlook factor on deicision
- Sunny outlook : 5/14 no : 3 , yes : 2
- Entropy(D/O=sunny) = -P(yes)*$log_2 P(yes)$ –P(no)*$log_2 P(no)$
- Entropy(D/O=sunny) = -(3/5)*$log_2(3/5)$ –(2/5)*$log_2(2/5)$

=-(3/5)*(-0,74) – (2/5)(-1,32)

=0,44 + 0,52

= 0,96

- Overcost outlook factor on decison
- Overcost outlook : 4/14 no : 0 , yes : 4
- Entropy(D/O=overcost) = -P(yes)*$log_2 P(yes)$ −P(no)*$log_2 P(no)$
- Entropy(D/O=overcost) = -(0)*$log_2(0)$ −(4/4)*$log_2(4/4)$

  = -(0)- (1)*$log_2(1)$

  =0

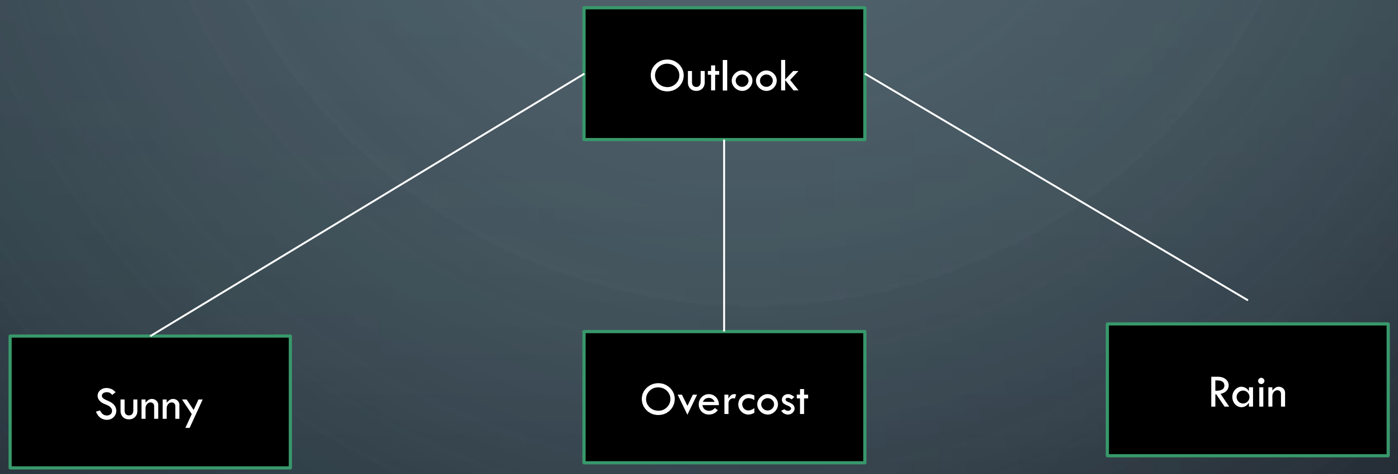
- Rain outlook factor on decison
- Rain oultook : 5/14 no : 2 , yes : 3
- Entropy(D/O=rain) = -P(yes)*$log_2 P(yes)$ −P(no)*$log_2 P(no)$
- Entropy(D/O=rain) = -(2/5)*$log_2(2/5)$ −(3/5)*$log_2(3/5)$= 0,96
- Gain(D,H) = Entropy(D)-[P(D/O=sunny)*Entopy(D/T=sunny]-

  [p(D/T=overcost)*Entropy(D/T=overcost)] - [p(D/T=rain)*Entropy(D/T=rain)]

  =0.940-[(5/14)*1)]-[(4/14)*0)]-[(5/14)*1)]

  =2,226

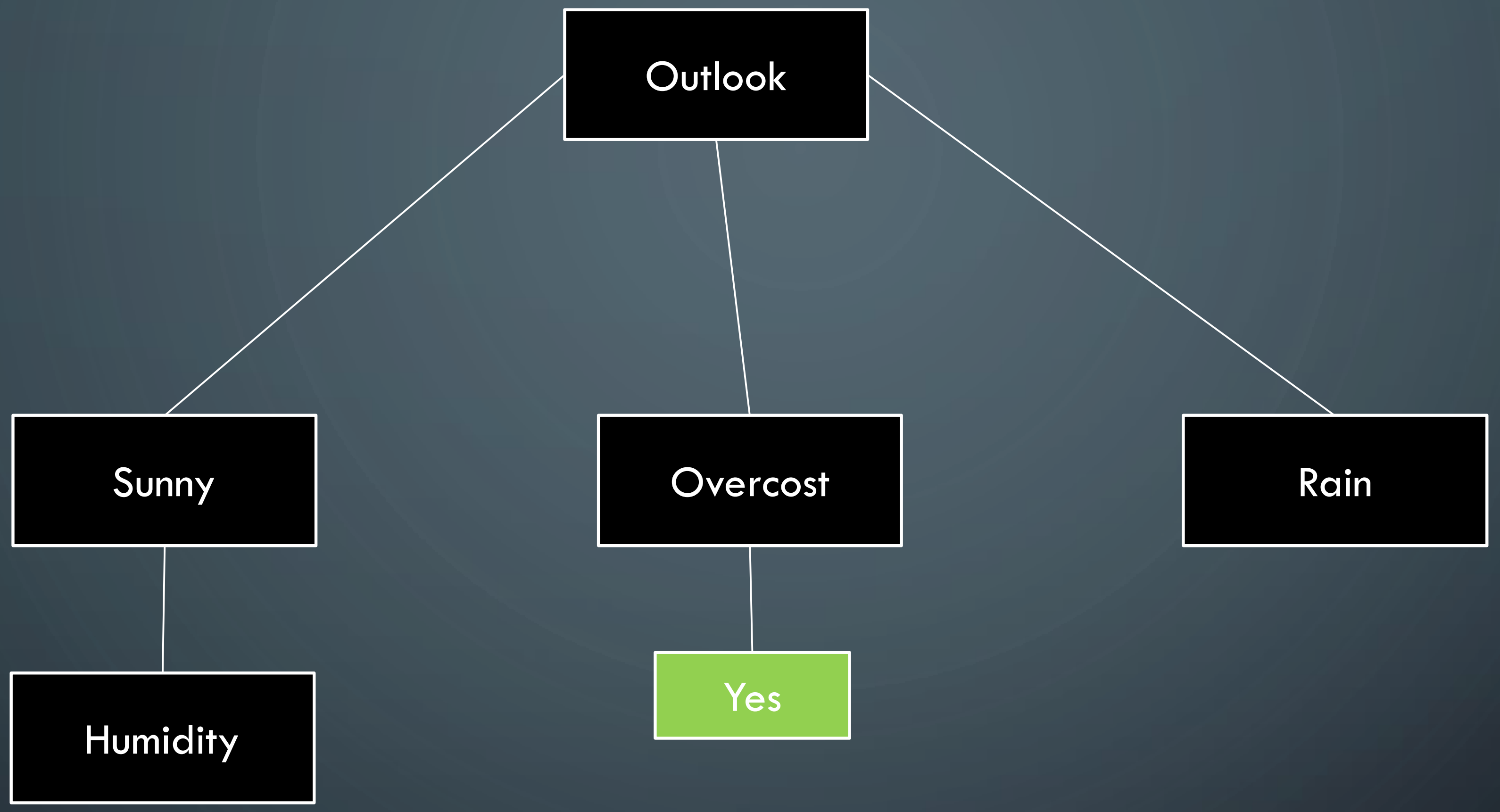
- After calculated all of gain finallyt we got that result :
- Gain(Decision , Outlook) = 2,226
- Gain(Decicion , Temperature) = 0,032
- Gain(Decision , Humidity) = 0,151
- Gain(Decision , Wind) = 0,048
- Outlook factor on decison produces the highest score that’s way outlook decision will appear on the root node of the tree.


Outlook
Sunny
Overcost
Rain

OVERCOST OULOOK ON DECİSİON

- The decision will be yes if outlook were overcost

Day	Outlook	Temp	Humidity	Wind	Decision
3	Overcost	Hot	High	Weak	Yes
7	Overcost	Cool	Normal	Strong	Yes
12	Overcost	Mild	High	Strong	Yes
13	Overcost	Hot	Normal	Weak	Yes


Outlook
Sunny
Overcost
Rain
Humidity
Yes

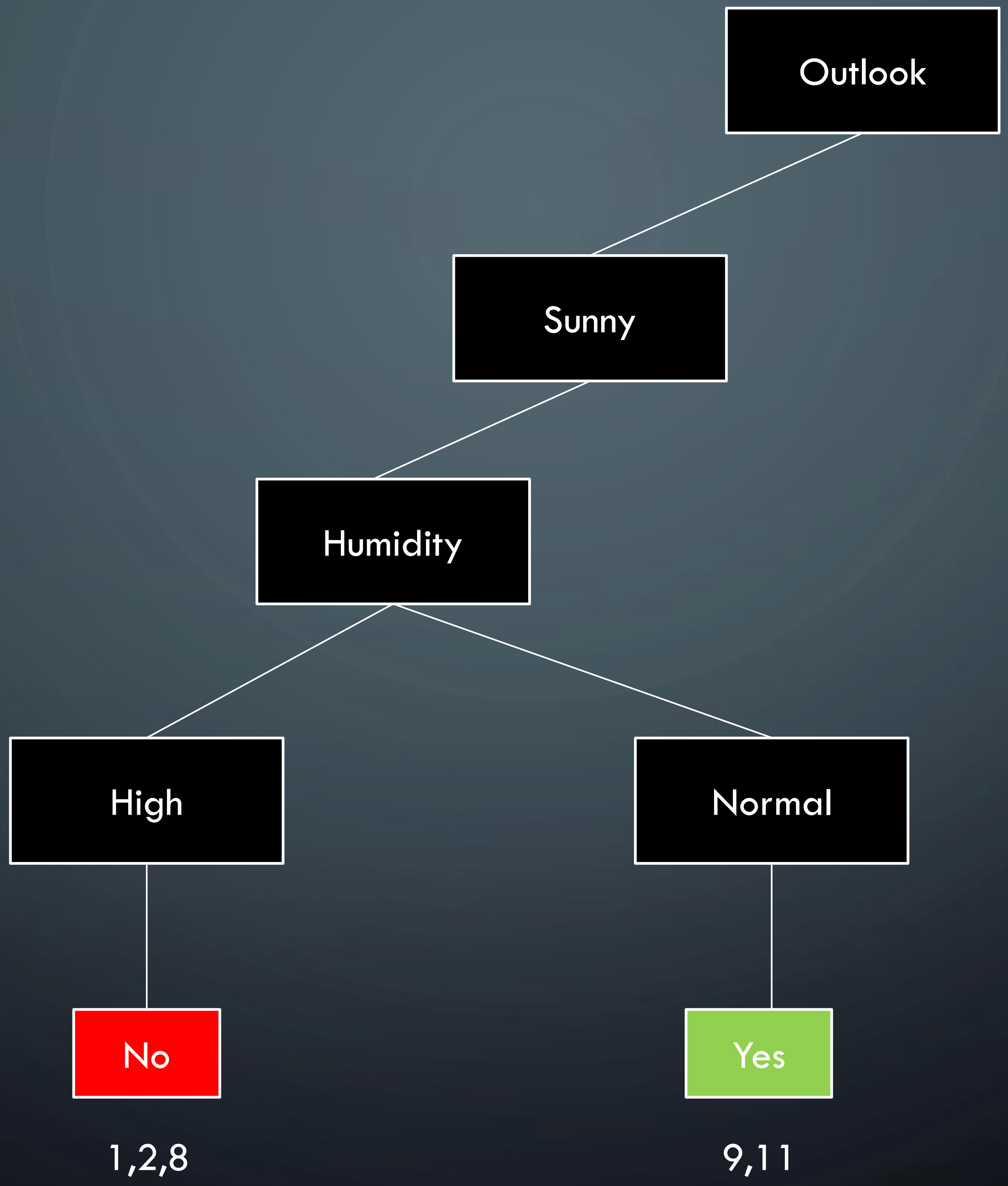
SUNNY OUTLOOK ON DECİSİON

- (outlook = sunny/temp)gain = 0,570
- (outlook = sunny/humidity)gain = 0,930
- (outlook= sunny/wind)gain = 0,019
- Humidity is a decision

SUNNY OUTLOOK ON DECİSİON

Day	Outlook	Temp	Humidity	Wind	Decision
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
8	Sunny	Mild	High	Weak	No

Day	Outlook	Temp	Humidity	Wind	Decision
9	Sunny	Cool	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes


Outlook
Sunny
Humidity
High
Normal
No
Yes
1,2,8
9,11

RAİN OUTLOOK ON DECİSON

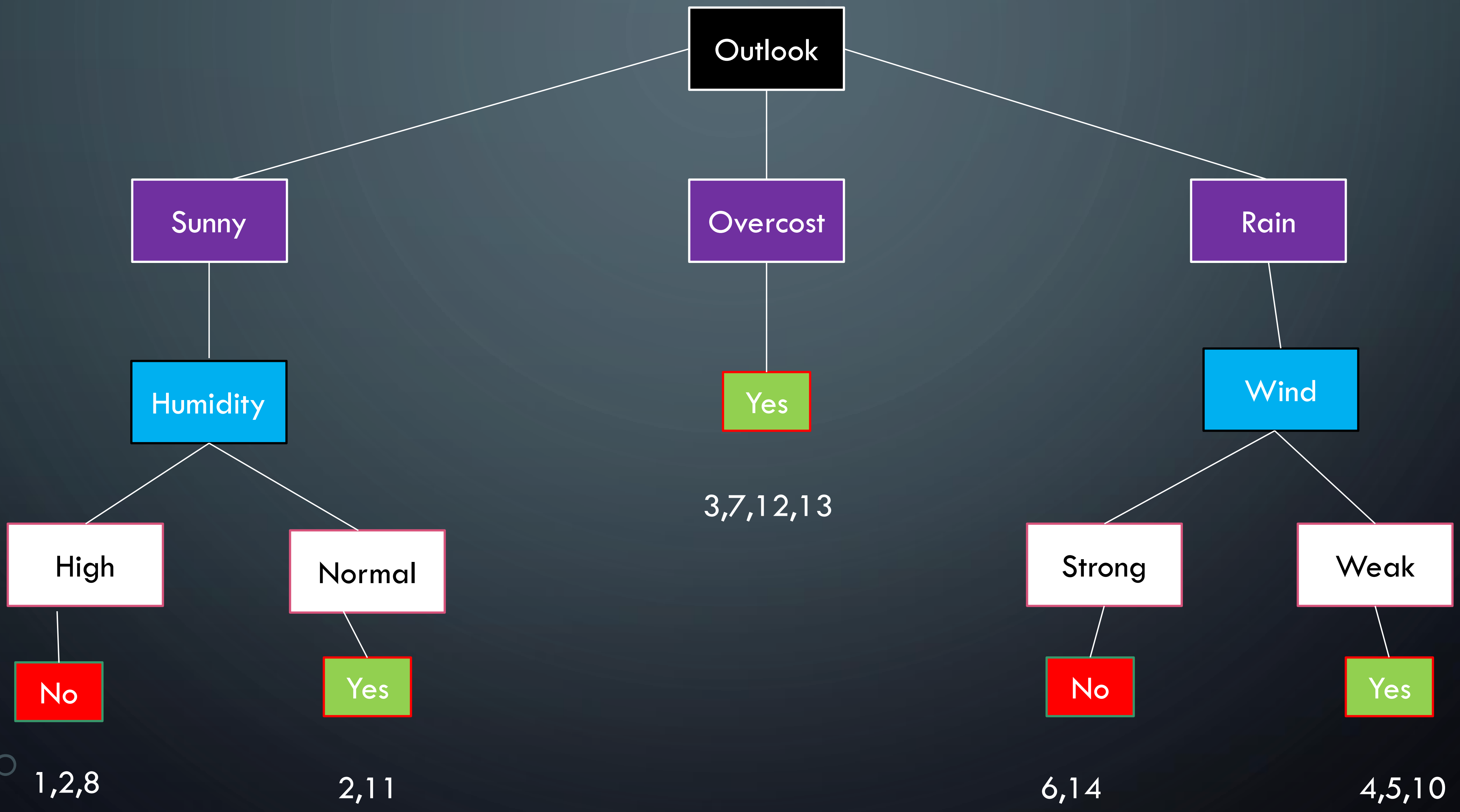
- ❑(Outlook = Rain /Temp)gain x
- ❑(Outlook = Rain/Humidity)gain x
- ❑(Outloot = Rain/Wind)gain √
- ❑Wind produces the highest score if outlook were rain

RAİN OUTLOOK ON DECİSON

Day	Outlook	Temp	Humidity	Wind	Decision
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
10	Rain	Mild	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No


Wind
Strong
Weak
No
Yes
6,14
4,5,10

COMPLETE DECİSİON TREE



DESADVANTAGE OF DECİSİON TREE

- Overfitting: Over fitting is a common flaw of decision trees. Setting constraints on model parameters (depth limitation) and making the model simpler through pruning are two ways to regularize a decision tree and improve its ability to generalize onto the test set.
- Predicting continuous variables: While decision trees can ingest continuous numerical input, they are not a practical way to predict such values, since decision-tree predictions must be separated into discrete categories, which results in a loss of information when applying the model to continuous values.

DESADVANTAGE OF DECİSİON TREE

- Heavy feature engineering: The flip side of a decision tree’s explanatory power is that it requires heavy feature engineering. When dealing with unstructured data or data with latent factors, this makes decision trees sub-optimal. Neural networks are clearly superior in this regard.

CONCLUSION

A different point of view is to regard I(node) as a measure of impurity. The more “mixed” the samples are at a node (i.e. has equal proportions of all class labels), the higher the impurity value. On the other hand, a homogeneous node (i.e. has samples of one class only) will have zero impurity. The Gain(A) value can thus be viewed as the amount of reduction in impurity if we split according to A. This affords the intuitive idea that we grow trees by recursively trying to obtain leaf nodes which are as pure as possible.

THANK YOU FOR LISTENING

- AHMAT MAHAMT B14120058
- FURKAN YİĞİT G171210092
- OĞUZHAN BUNCUK G171210108
- ENES ŞAHİN G171210300

RESOURCE

- https://www.edureka.co/blog/decision-tree-algorithm/
- https://www.google.com/search?q=Artificial+Intelligence&source=lnms&tbm=isch&sa=X&ved=2ahUKEwjwtYWRvuPoAhW1DWMBHexdD8UQ_AUoAXoECBIQAw&biw=1366&bih=576#imgrc=ZV3gTJCnJC8wmM
- https://www.google.com/search?q=entropy+formula&source=lnms&tbm=isch&sa=X&ved=2ahUKEwjR45zgh-ToAhUEwsQBHenxC6UQ_AUoAXoECBMQAw&biw=1366&bih=576
- https://www.dsi.unive.it/~atorsell/AI/mod1-12-decision-trees.pdf
- https://www.youtube.com/watch?v=WC_DhP3vyy8