


# PREDICTION USING DECISION TREE ALGORITHM

## GRIP – THE SPARKS FOUNDATION

PRESENTED BY:

RISHIKA HAZARIKA  
DATA SCIENCE AND BUSINESS ANALYTICS INTERN

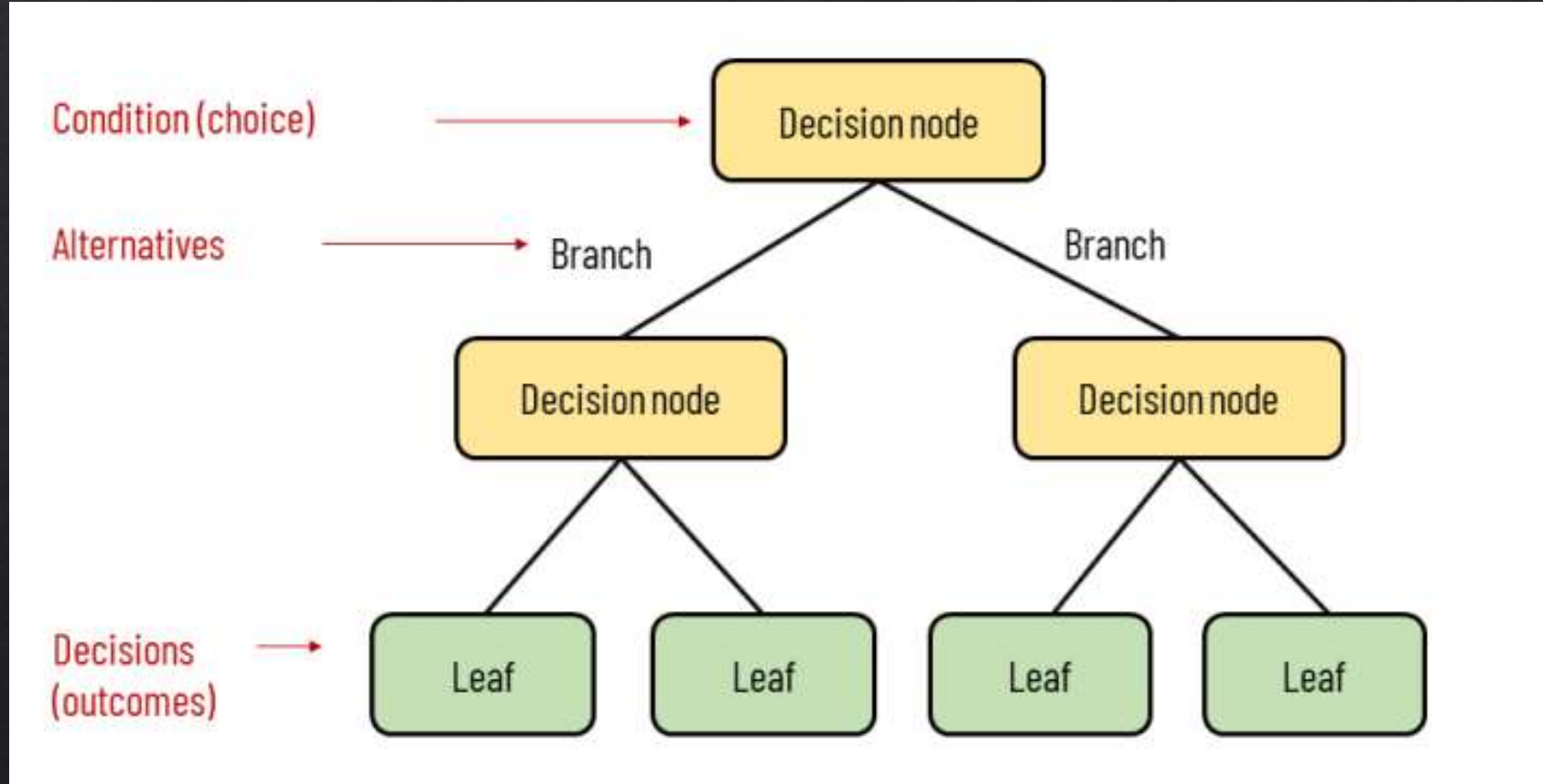
## TASK 6:

A decorative graphic on the left side of the slide features a central large blue hexagon with the text "#6" in white. Surrounding this central hexagon are several smaller hexagons in various shades of blue and teal. These smaller hexagons contain icons: a lightbulb, a thumbs-up, a network of nodes, a smartphone, a magnifying glass, a gear, and a speech bubble.

### Prediction using Decision Tree Algorithm (Level – Intermediate)

- Create the Decision Tree classifier and visualize it graphically.
- The purpose is if we feed any new data to this classifier, it would be able to predict the right class accordingly.
- Dataset : <https://bit.ly/3kXTdox>
- Sample Solution : <https://bit.ly/2G6sYx9>
- Task submission:
  1. Host the code on GitHub Repository (public). Record the code and output in a video. Post the video on YouTube
  2. Share links of code (GitHub) and video (YouTube) as a post on YOUR LinkedIn profile
  3. Submit the LinkedIn link in Task Submission Form when shared.
  4. Please read FAQs on how to submit the tasks.

## DECISION TREE:

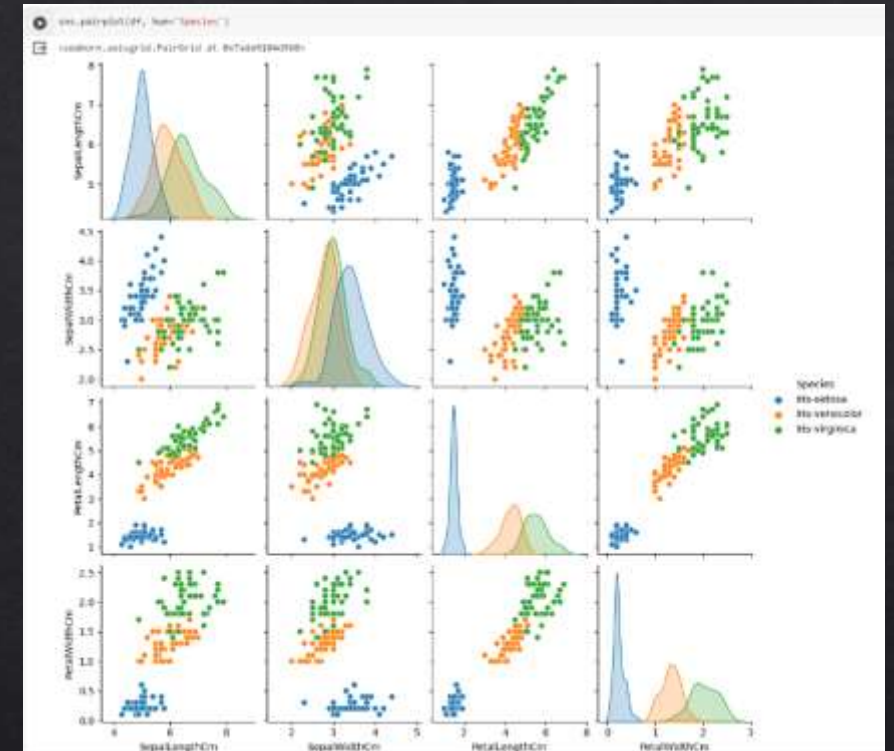


## INTERPRETATION OF SCATTERPLOTS:

- Sepal Width vs Sepal Length: There is a general trend, where sepal width increases, sepal length also increases. However the relationship is not perfectly linear, but has some degree of scatter.
- Sepal Width vs Petal Width: Follows a general trend but the relationship between sepal width and petal width is not perfectly linear.
- Sepal Length vs Petal Length: The scatterplot shows that sepal length and petal length are strongly positively correlated. This means that as sepal length increases, petal length also tends to increase.
- Sepal Width vs Petal Length: This scatterplot also demonstrates a strong positive correlation, which suggests that as sepal width increases, petal length also tends to increase.

Additionally, we can also derive that *Iris-setosa* generally has smaller petals compared to *Iris-versicolor* and *Iris-virginica*.

*Iris-virginica* generally has longer sepal and petal lengths compared to the other two species.





## INTERPRETATIONS OF HEATMAP:

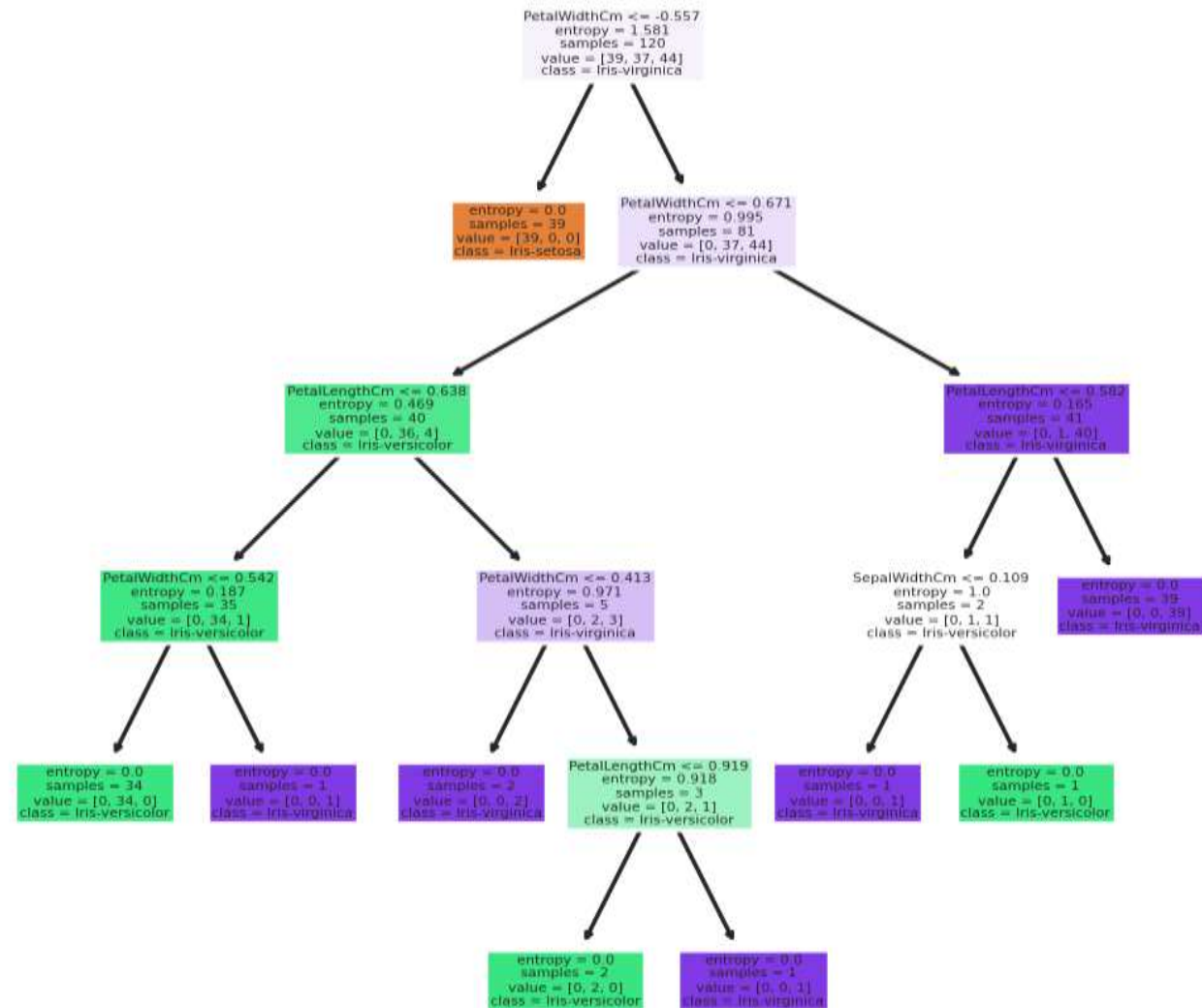
The values in the squares indicate the correlation coefficient.

- The diagonal line from the top left to the bottom right represents the correlation of each variable with itself. Since every variable is perfectly correlated with itself, all the numbers on the diagonal are 1.
- the correlation between Sepal Length and Petal Length is 0.9, which means that there is a strong positive correlation between these two variables.
- the correlation between Sepal Width and Petal Width is -0.4, which means that there is a strong negative correlation between these two variables.

etc.



## DECISION TREE II:



THANK YOU