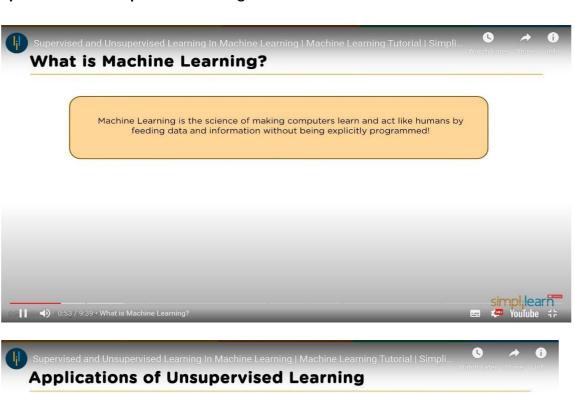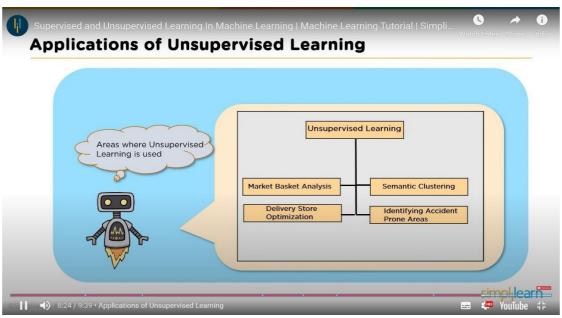# PRIOR KNOWLEDGE

Project Name	Developing a flight delay prediction model using machine learning.
Team id	PNT2022TMID00047

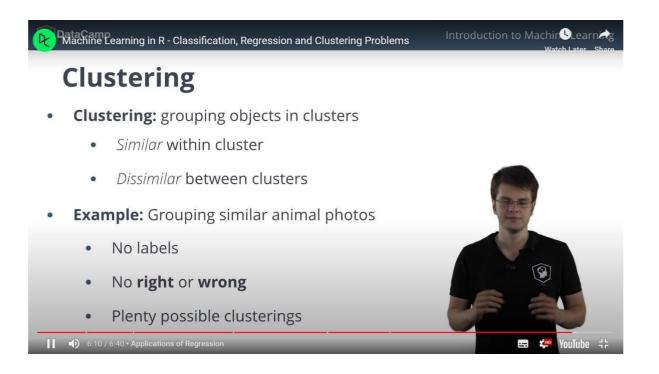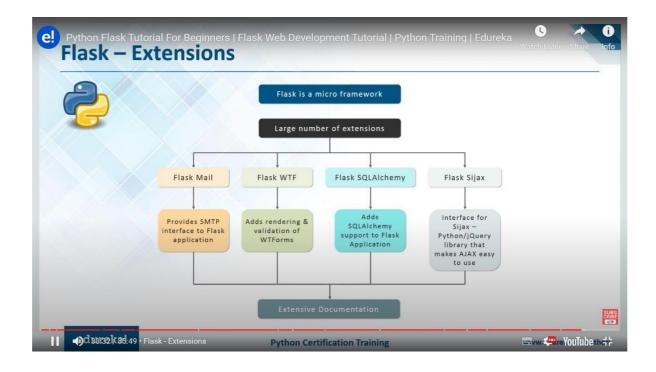**Supervised and unsupervised learning:**

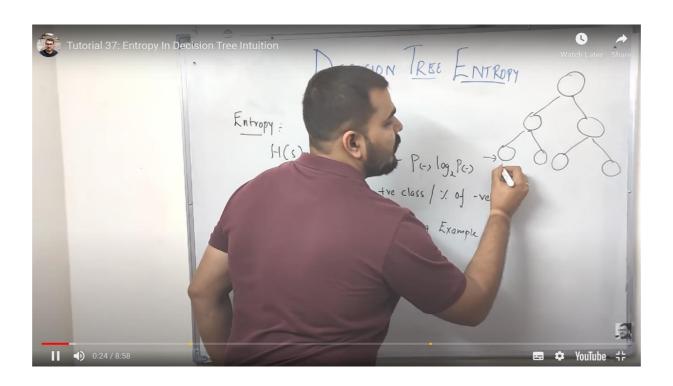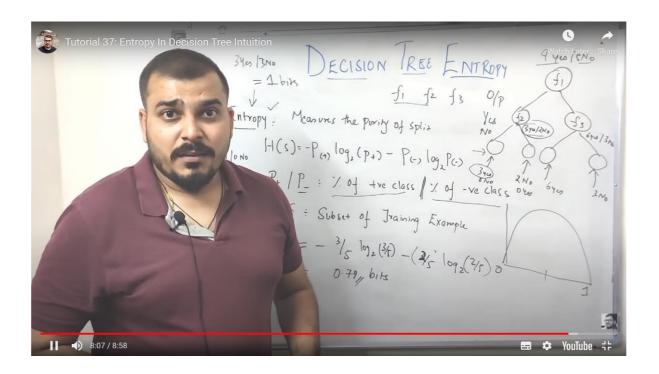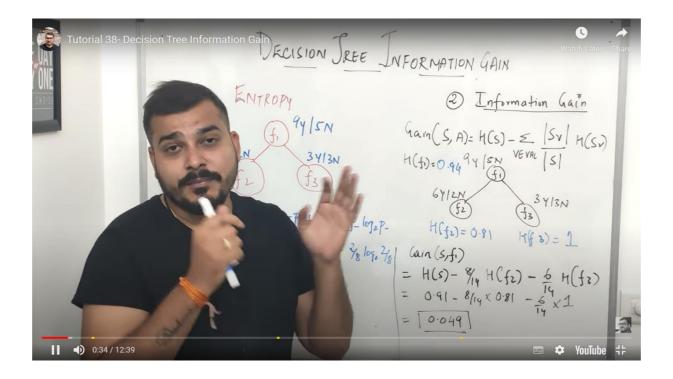**Regression Classification and Clustering :**

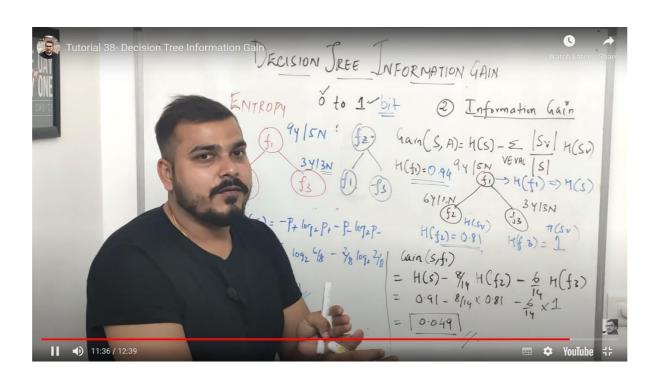**Flask:**

Decision Tree:

Entropy in Decision Tree Intuition:

Decision Tree information gain:

Gini Impurity intuition in depth in Decision tree:

GINI IMPURITY DT

$$\begin{array}{ccc} f_1 & f_2 & f_3 \\ c_1 & D_1 & \text{Yes} \\ c_2 & D_2 & \text{Yes} \\ & & \text{No} \\ & & \text{No} \\ & & \text{Yes} \end{array}$$

① **Entropy**

$$H(S) = -P_+ \log_2 P_+ - P_- \log_2 P_-$$

② **GINI IMPURITY**

$$GI = 1 - \sum_{i=1}^{n} (P)^2$$
$$= 1 - [(P_+)^2 + (P_-)^2]$$

GINI IMPURITY DT

Entropy

$$H(S) = -P_+ \log_2 P_+ - P_- \log_2 P_-$$

② **GINI IMPURITY**

$$GI = 1 - \sum_{i=1}^{n} (P)^2$$
$$= 1 - [(P_+)^2 + (P_-)^2]$$
$$= 1 - [(3/6)^2 + (3/6)^2]$$
$$= 1 - [0.25 + 0.25]$$
$$= 0.5$$

$6 Y | 3N$

$3Y | 3N \quad c_1$

$3Y | 0N \quad c_2 \rightarrow$ leaf Node

$$-3/6 \log_2 3/6 - 3/6 \log_2 3/6$$