

## PRIOR KNOWLEDGE

Project Name	Developing a flight delay prediction model using machine learning.
Team id	PNT2022TMID29843
Team members	Sridhar K V Sriram S M Tamilarasan D Vittal J S

### Supervised and unsupervised learning:

Supervised and Unsupervised Learning In Machine Learning | Machine Learning Tutorial | Simpli...

**What is Machine Learning?**

Machine Learning is the science of making computers learn and act like humans by feeding data and information without being explicitly programmed!

0:53 / 9:39 • What is Machine Learning?

simplilearn YouTube

Supervised and Unsupervised Learning In Machine Learning | Machine Learning Tutorial | Simpli...

**Applications of Unsupervised Learning**

Areas where Unsupervised Learning is used

```
graph TD; UL[Unsupervised Learning] --- MBA[Market Basket Analysis]; UL --- DSO[Delivery Store Optimization]; UL --- SC[Semantic Clustering]; UL --- IAPA[Identifying Accident Prone Areas];
```

8:24 / 9:39 • Applications of Unsupervised Learning

simplilearn YouTube

## Regression Classification and Clustering :

 DataCamp

Machine Learning in R - Classification, Regression and Clustering Problems

Introduction to Machine Learning 

Watch Later 


# Common ML Problems

- **Classification**
- **Regression**
- **Clustering**





0:26 / 6:40 • Introduction

YouTube

 DataCamp


Machine Learning in R - Classification, Regression and Clustering Problems

Introduction to Machine Learning 

Watch Later 

# Clustering

- **Clustering:** grouping objects in clusters
  - *Similar* within cluster
  - *Dissimilar* between clusters
- **Example:** Grouping similar animal photos
  - No labels
  - No **right** or **wrong**
  - Plenty possible clusterings



6:10 / 6:40 • Applications of Regression

YouTube

## Flask:

Python Flask Tutorial For Beginners | Flask Web Development Tutorial | Python Training | Edureka

# Introduction to Flask

Watch Later Share Info

The diagram shows the Python logo at the top left. Below it, a central flow shows 'Libraries' and 'Modules' leading to a 'Web Developer' (represented by a person at a computer). This leads to an 'APPLICATION' (represented by a megaphone). Below this, a comparison shows 'Life without Flask!' (a person looking frustrated at a computer) versus 'Using Flask!' (a person smiling at a computer), with a green arrow pointing from the former to the latter. A banner at the top asks 'What is a Web Framework?'.

What is a Web Framework?

Libraries Modules Web Developer APPLICATION

Life without Flask! Using Flask!

Python Certification Training

www.edureka.com YouTube

Python Flask Tutorial For Beginners | Flask Web Development Tutorial | Python Training | Edureka

# Flask – Extensions

Watch Later Share Info

The diagram shows the Python logo at the top left. Below it, a central flow shows 'Flask is a micro framework' leading to 'Large number of extensions'. This branches into four categories: 'Flask Mail' (Provides SMTP interface to Flask application), 'Flask WTF' (Adds rendering & validation of WTForms), 'Flask SQLAlchemy' (Adds SQLAlchemy support to Flask Application), and 'Flask Sijax' (Interface for Sijax – Python/jQuery library that makes AJAX easy to use). All four categories lead to 'Extensive Documentation'.

Flask is a micro framework

Large number of extensions

Flask Mail Provides SMTP interface to Flask application

Flask WTF Adds rendering & validation of WTForms

Flask SQLAlchemy Adds SQLAlchemy support to Flask Application

Flask Sijax Interface for Sijax – Python/jQuery library that makes AJAX easy to use

Extensive Documentation

Python Certification Training

www.edureka.com YouTube

## Decision Tree:

### Entropy in Decision Tree Intuition:

Tutorial 37: Entropy In Decision Tree Intuition

Watch Later Share

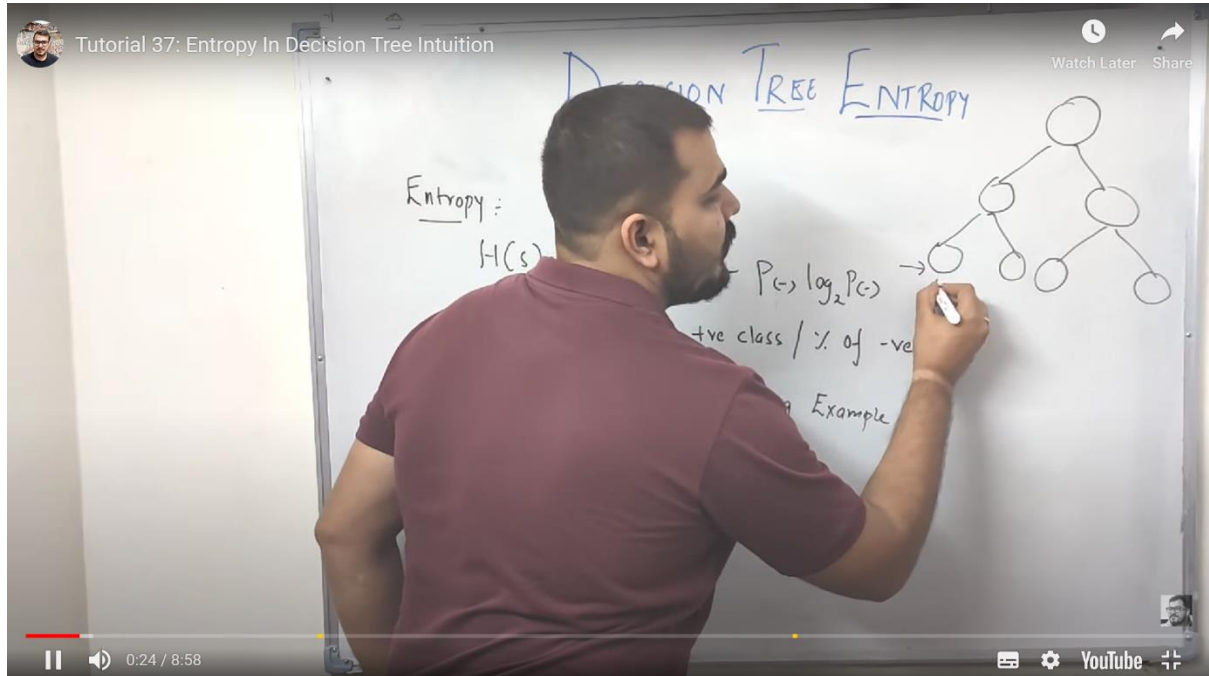
Entropy:

$H(s)$

$P_+ \log_2 P_+$

→

Example



0:24 / 8:58

YouTube

Tutorial 37: Entropy In Decision Tree Intuition

Watch Later Share

DECISION TREE ENTROPY

$f_1$   $f_2$   $f_3$  O/P

Yes No

Entropy: Measures the purity of split

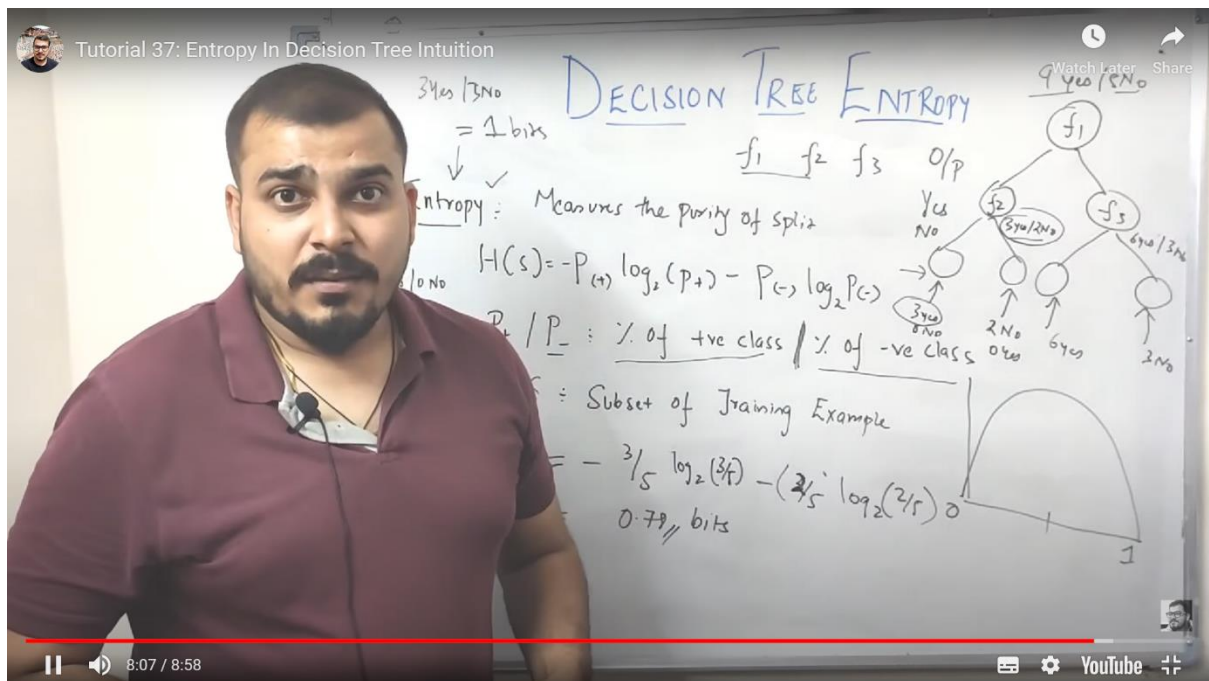
$H(s) = -P_+ \log_2(P_+) - P_- \log_2(P_-)$

$P_+ / P_-$  : % of +ve class / % of -ve class

S : Subset of Training Example

$= -\frac{3}{5} \log_2(\frac{3}{5}) - (\frac{2}{5} \log_2(\frac{2}{5}))$

0.79 bits



8:07 / 8:58

YouTube

## Decision Tree information gain:

Tutorial 38- Decision Tree Information Gain

### DECISION TREE INFORMATION GAIN

**ENTROPY**

94/5N  
34/3N  
64/2N  
34/3N

$H(f_1) = 0.94$   
 $H(f_2) = 0.81$   
 $H(f_3) = 1$

$H(S) = -P_1 \log_2 P_1 - P_2 \log_2 P_2 - P_3 \log_2 P_3$   
 $H(S) = -\frac{9}{8} \log_2 \frac{9}{8} - \frac{3}{8} \log_2 \frac{3}{8}$

**② Information Gain**

$$Gain(S, A) = H(S) - \sum_{v \in \text{VAL}(A)} \frac{|S_v|}{|S|} H(S_v)$$

$$Gain(S, f_1) = H(S) - \frac{9}{14} H(f_2) - \frac{6}{14} H(f_3)$$

$$= 0.91 - \frac{9}{14} \times 0.81 - \frac{6}{14} \times 1$$

$$= \boxed{0.049}$$

0:34 / 12:39

Tutorial 38- Decision Tree Information Gain

### DECISION TREE INFORMATION GAIN

**ENTROPY**

94/5N  
34/3N  
64/2N  
34/3N

$H(f_1) = 0.94$   
 $H(f_2) = 0.81$   
 $H(f_3) = 1$

$H(S) = -P_1 \log_2 P_1 - P_2 \log_2 P_2 - P_3 \log_2 P_3$   
 $H(S) = -\frac{9}{8} \log_2 \frac{9}{8} - \frac{3}{8} \log_2 \frac{3}{8}$

**② Information Gain**

$$Gain(S, A) = H(S) - \sum_{v \in \text{VAL}(A)} \frac{|S_v|}{|S|} H(S_v)$$

$$Gain(S, f_1) = H(S) - \frac{9}{14} H(f_2) - \frac{6}{14} H(f_3)$$


$$= 0.91 - \frac{9}{14} \times 0.81 - \frac{6}{14} \times 1$$

$$= \boxed{0.049}$$

11:36 / 12:39



### Gini Impurity intuition in depth in Decision tree:



# Tutorial 39- Gini Impurity Intuition In Depth In Decision Tree

[Watch Later](#)
[Share](#)

Y

Yes

Yes

No

No

Yes

...

Y

Yes

Yes

No

No

Yes

...

## GINI IMPURITY DT

### ① Entropy

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

### ② GINI IMPURITY

$$GI = 1 - \sum_{i=1}^n (P_i)^2$$

$$= 1 - [(P_+)^2 + (P_-)^2]$$

0:37 / 11:12

YouTube

[illegible]