

# Developing a Flight Delay Prediction Model using Machine Learning

TEAMID:PNT2022TMID26438

TEAM LEADER : SIVAKUMAR.M

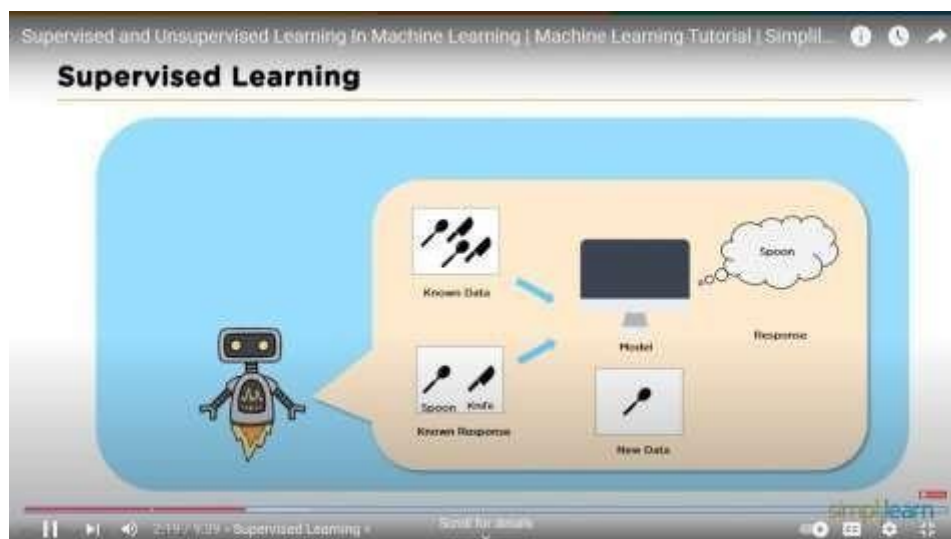
TEAM MEMBER 1 : SAIPRIYA.S

TEAM MEMBER 2 : SHIFLI S J

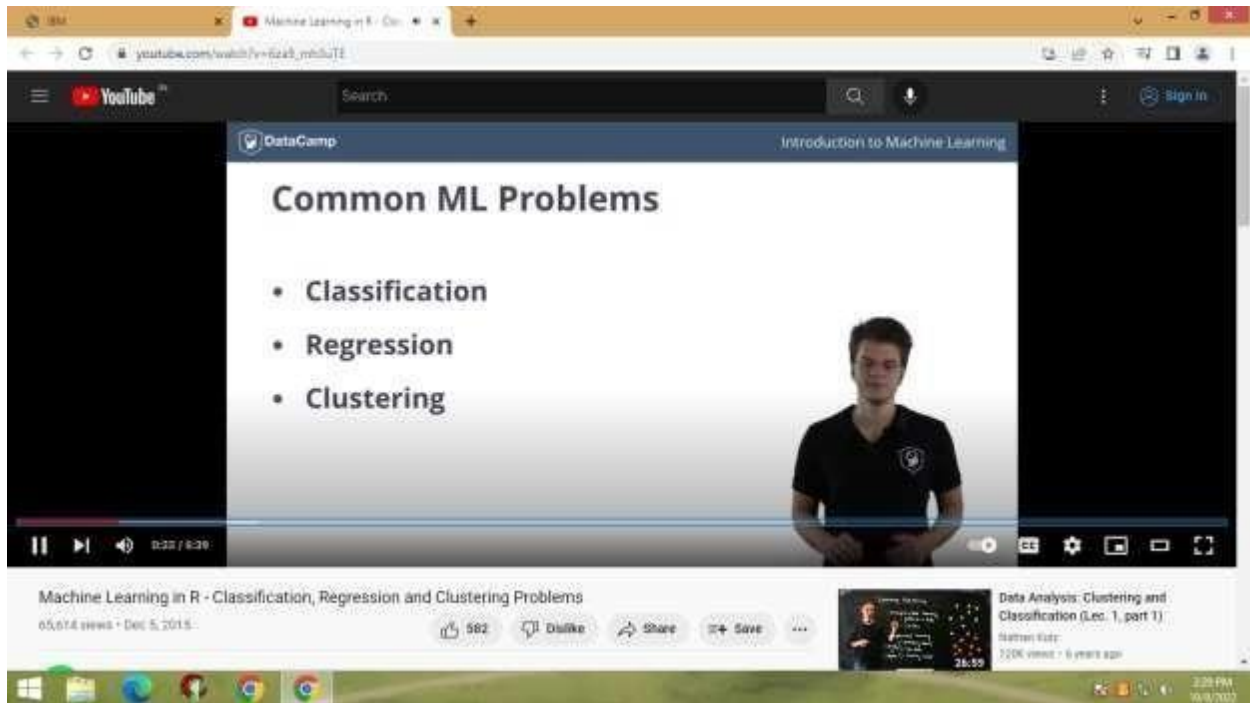
TEAM MEMEMBER 3 : SABAREES

## Prior Knowledge:

Supervised and unsupervised learning:

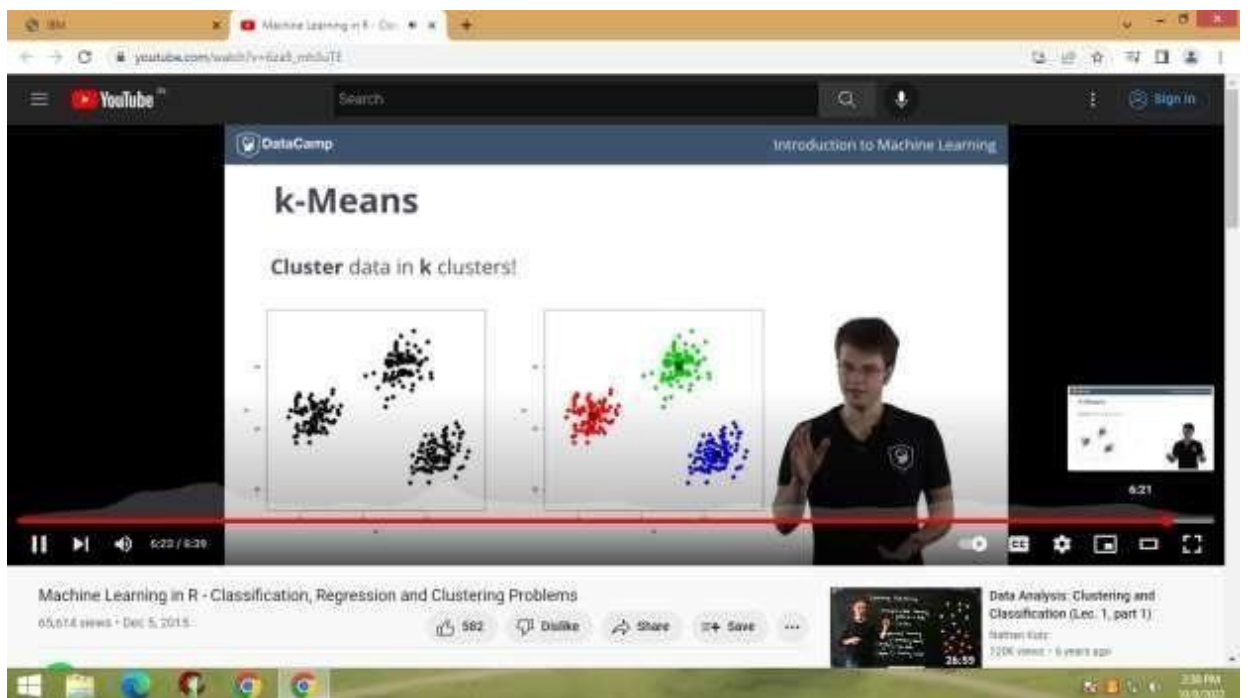


## Regression Classification and Clustering:



A screenshot of a YouTube video player. The video is titled "Machine Learning in R - Classification, Regression and Clustering Problems" and is from the channel "DataCamp". The video is at 0:22 / 8:30. The main content is a slide titled "Common ML Problems" with a bulleted list: "Classification", "Regression", and "Clustering". A presenter is visible in the bottom right corner of the slide. The video player interface includes a search bar, a sign-in button, and a list of recommended videos at the bottom.

Machine Learning in R - Classification, Regression and Clustering Problems  
05,674 views · Dec 5, 2015 · 582 likes · 582 dislikes · 582 shares · 582 saves



A screenshot of a YouTube video player. The video is titled "Machine Learning in R - Classification, Regression and Clustering Problems" and is from the channel "DataCamp". The video is at 6:22 / 8:30. The main content is a slide titled "k-Means" with the subtitle "Cluster data in k clusters!". The slide shows two scatter plots: the left one has black data points, and the right one has red, green, and blue data points. A presenter is visible in the bottom right corner of the slide. The video player interface includes a search bar, a sign-in button, and a list of recommended videos at the bottom.

k-Means  
Cluster data in k clusters!

Machine Learning in R - Classification, Regression and Clustering Problems  
05,674 views · Dec 5, 2015 · 582 likes · 582 dislikes · 582 shares · 582 saves

Flask:

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

```
from flask import Flask
app = Flask(__name__)


@app.route('/hello/<name>')
def hello_name(name):
    return "Hello %s!" % name

if __name__ == '__main__':
    app.run(debug = True)
```

WARNING: Do not use the development server in a production environment.  
Use a production WSGI server instead.  
\* Debug mode: off  
\* Running on <http://127.0.0.1:5000/> (Press CTRL-C to quit)  
127.0.0.1 - - [14/Dec/2019 11:44:47] "GET /hello/1" 200 -

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

## Flask – Redirect & Errors

 Standardized status codes

Prototype → Flask.abort(code)

Sl.no	Status Code
1	HTTP_300_MULTIPLE_CHOICES
2	HTTP_301_MOVED_PERMANENTLY
3	HTTP_302_FOUND
4	HTTP_303_SEE_OTHER
5	HTTP_304_NOT_MODIFIED
6	HTTP_305_USE_PROXY
7	HTTP_306_RESERVED

Sl.no	Code	Description
1	400	Bad Request
2	401	Unauthenticated
3	403	Forbidden
4	404	Not Found
5	406	Not Acceptable
6	415	Unsupported Media Type
7	429	Too Many Requests

## Entropy in Decision Tree Intuition:

[illegible]



## Decision Tree information:

Tutorial 38- Decision Tree Information Gain

### DECISION TREE INFORMATION GAIN

#### ENTROPY

$H(f_1) = -\frac{4}{5} \log_2 \frac{4}{5} - \frac{1}{5} \log_2 \frac{1}{5}$   
 $H(f_2) = -\frac{6}{14} \log_2 \frac{6}{14} - \frac{8}{14} \log_2 \frac{8}{14}$   
 $H(f_3) = -\frac{3}{13} \log_2 \frac{3}{13} - \frac{10}{13} \log_2 \frac{10}{13}$

#### ② Information Gain

$Gain(S, A) = H(S) - \sum_{v \in \text{val}} \frac{|S_v|}{|S|} H(S_v)$   
 $H(S) = 0.94$   
 $H(f_1) = 0.94$   
 $H(f_2) = 0.91$   
 $H(f_3) = 1$

$Gain(S, f_1)$   
 $= H(S) - \frac{4}{5} H(f_2) - \frac{1}{5} H(f_3)$   
 $= 0.94 - \frac{4}{5} \times 0.91 - \frac{1}{5} \times 1$   
 $= 0.049$

Tutorial 38- Decision Tree Information Gain

### DECISION TREE INFORMATION GAIN

#### ENTROPY

$H(f_1) = -\frac{4}{5} \log_2 \frac{4}{5} - \frac{1}{5} \log_2 \frac{1}{5}$   
 $H(f_2) = -\frac{6}{14} \log_2 \frac{6}{14} - \frac{8}{14} \log_2 \frac{8}{14}$   
 $H(f_3) = -\frac{3}{13} \log_2 \frac{3}{13} - \frac{10}{13} \log_2 \frac{10}{13}$

#### ② Information Gain

$Gain(S, A) = H(S) - \sum_{v \in \text{val}} \frac{|S_v|}{|S|} H(S_v)$   
 $H(S) = 0.94$   
 $H(f_1) = 0.94$   
 $H(f_2) = 0.91$   
 $H(f_3) = 1$

$Gain(S, f_1)$   
 $= H(S) - \frac{4}{5} H(f_2) - \frac{1}{5} H(f_3)$   
 $= 0.94 - \frac{4}{5} \times 0.91 - \frac{1}{5} \times 1$   
 $= 0.049$

0 to 1 bit

## Gini Impurity intuition in depth in Decision

Tutorial 39- Gini Impurity Intuition In Depth In Decision Tree

Press Esc to exit full screen

GINI IMPURITY DT

$f_1$	$f_2$	$f_3$	O/P
$C_1$	$D_1$		Yes
$C_2$	$D_2$		Yes
			No
			No
			Yes
			...

① Entropy

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

$= 0$

$\rightarrow (f_1)$   $4Y/13N$

$\rightarrow (f_2)$   $3Y/10N$

$\rightarrow (f_3)$   $leaf Node$

② GINI IMPURITY

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

2:24 / 11:12

Scroll for details