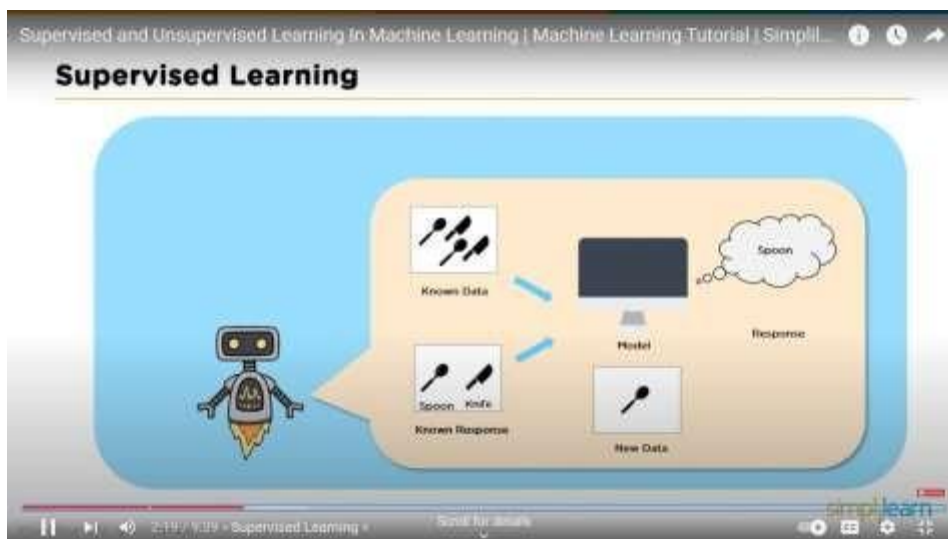


Developing a Flight Delay Prediction Model using Machine Learning

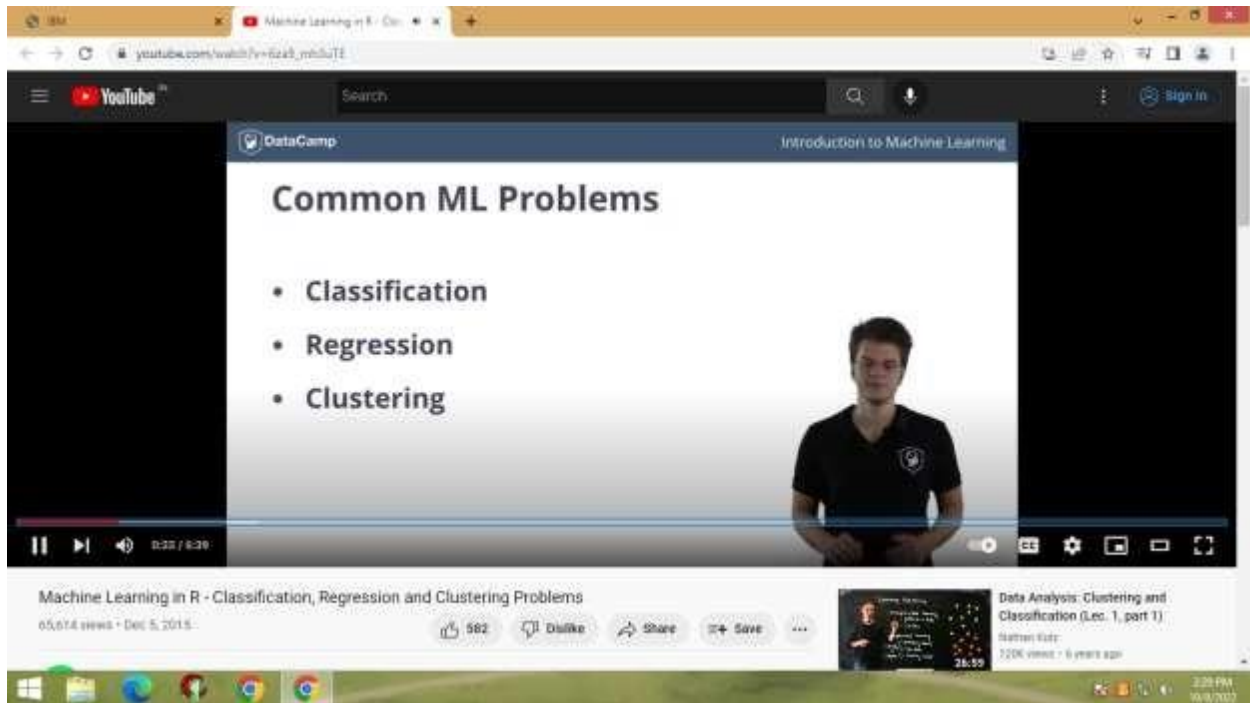
TEAM LEADER : PREETHI S
TEAM MEMBER 1 : MADHUBALA A
TEAM MEMBER 2 : ROHINI S
TEAM MEMBER 3 : AKSHAYA S

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 on the right side 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 · 1 dislike · Share · Save

Recommended videos:
Data Analysis: Clustering and Classification (Lec. 1, part 1)
Matthew Kuhn · 7,20K views · 6 years ago



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 on the right side 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 · 1 dislike · Share · Save

Recommended videos:
Data Analysis: Clustering and Classification (Lec. 1, part 1)
Matthew Kuhn · 7,20K views · 6 years ago

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 \Rightarrow `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	Unauthorized
3	403	Forbidden
4	404	Not Found
5	406	Not Acceptable
6	415	Unsupported Media Type
7	429	Too Many Requests

Decision Tree:

Entropy in Decision Tree Intuition:

Tutorial 37: Entropy In Decision Tree Intuition

DECISION TREE ENTROPY

Entropy

$f_1, f_2, f_3, o/p$

the purity of split

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

% of +ve class / % of -ve class

Training Example

9 yes / 0 no

5 yes / 0 no

4 yes / 3 no

3 yes / 0 no

2 yes / 0 no

1 yes / 3 no

0 yes / 3 no

2:50 / 8:52

Scroll for details

Tutorial 37: Entropy In Decision Tree Intuition

DECISION TREE ENTROPY

Entropy

$f_1, f_2, f_3, o/p$

the purity of split

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

$P_{+} / P_{-} = \% \text{ of +ve class} / \% \text{ of -ve class}$

$S = \text{Subset of Training Example}$

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

0.72 bits

3 yes / 0 no = 1 bits

Entropy

Means the purity of split

3 yes / 0 no

2 yes / 0 no

1 yes / 3 no

0 yes / 3 no

2 yes / 0 no

1 yes / 3 no

0 yes / 3 no

0.72 bits

8:05 / 8:52

Scroll for details

Decision Tree information:

Tutorial 38- Decision Tree Information Gain

DECISION TREE INFORMATION GAIN

ENTROPY

$f_1: 4Y/5N$
 $f_3: 3Y/3N$

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

② Information Gain

$Gain(S, A) = H(S) - \sum_{v \in \text{VAL}} \frac{|S_v|}{|S|} H(S_v)$
 $H(f_1) = 0.94$

$f_2: 6Y/2N$
 $H(f_2) = 0.91$

$f_3: 3Y/3N$
 $H(f_3) = 1$

$Gain(S, f_1)$
 $= H(S) - \frac{8}{14} H(f_2) - \frac{6}{14} H(f_3)$
 $= 0.91 - \frac{8}{14} \times 0.91 - \frac{6}{14} \times 1$
 $= 0.049$

$H(S) = -P_1 \log_2 P_1 - \dots$
 $-\frac{8}{14} \log_2 \frac{8}{14} - \frac{6}{14} \log_2 \frac{6}{14}$

0.06 / 12:39

Tutorial 38- Decision Tree Information Gain

DECISION TREE INFORMATION GAIN

ENTROPY

$f_1: 4Y/5N$
 $f_3: 3Y/3N$

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

② Information Gain

$Gain(S, A) = H(S) - \sum_{v \in \text{VAL}} \frac{|S_v|}{|S|} H(S_v)$
 $H(f_1) = 0.94$

$f_2: 6Y/2N$
 $H(f_2) = 0.91$

$f_3: 3Y/3N$
 $H(f_3) = 1$

$Gain(S, f_1)$
 $= H(S) - \frac{8}{14} H(f_2) - \frac{6}{14} H(f_3)$
 $= 0.91 - \frac{8}{14} \times 0.91 - \frac{6}{14} \times 1$
 $= 0.049$

$H(S) = -P_1 \log_2 P_1 - \dots$
 $-\frac{8}{14} \log_2 \frac{8}{14} - \frac{6}{14} \log_2 \frac{6}{14}$

11:11 / 12:39

Economic Recovery Indicators
Anticipating the India's GDP growth on a real-time basis
info.econdata.com

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	Q/P
G_1	D_1		Yes
G_2	D_2		Yes
			No
			No
			Yes
			...

① Entropy

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

\downarrow
64/32

② GINI IMPURITY

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

34/01

C_2 - Leaf Node

2:24 / 11:12

Scroll for details