

DEVELOPING A FLIGHT DELAY PREDICTION MODEL USING MACHINE LEARNING

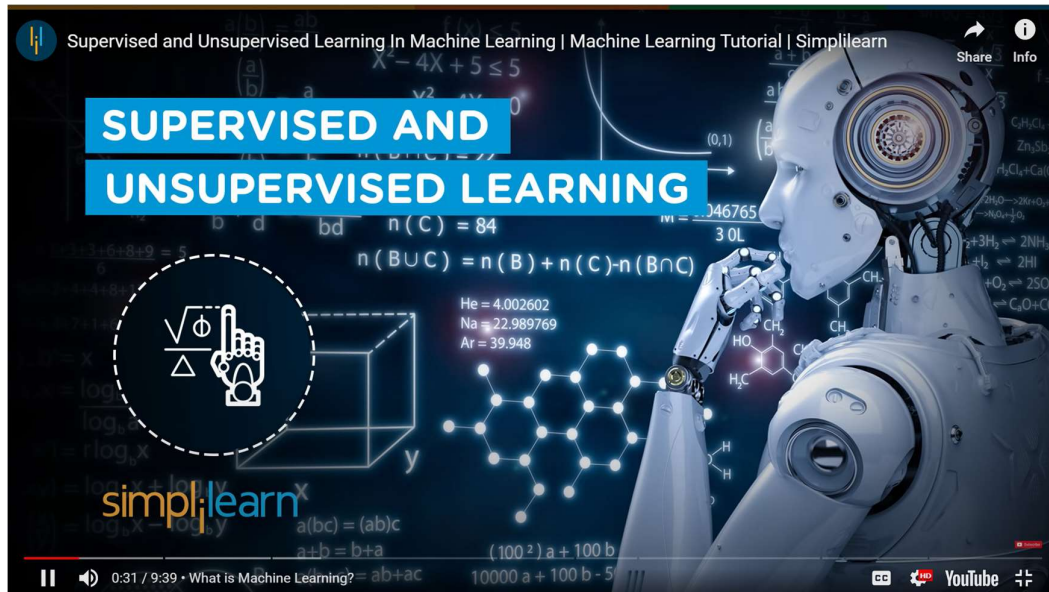
TEAM LEADER : Dhanushpriya T

TEAM MEMBERS :

1. M P S Gayathri
2. Jecintha T
3. Kaviyaa K V

Prior Knowledge:

Supervised and unsupervised learning:



Supervised and Unsupervised Learning In Machine Learning | Machine Learning Tutorial | Simplilearn

Share Info

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];
```

8:27 / 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

Classification Problem

Goal: predict category of new observation

Earlier Observations $\xrightarrow{\text{Estimate}}$ CLASSIFIER

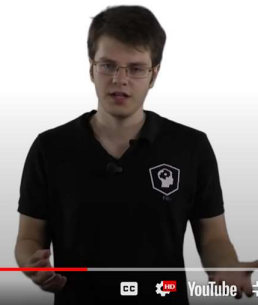
0:57 / 6:39 • Classification

YouTube



Clustering

- **Clustering:** grouping objects in clusters
 - *Similar* within cluster
 - *Dissimilar* between clusters



5:17 / 6:39 • Applications of Regression

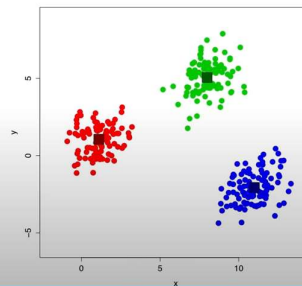
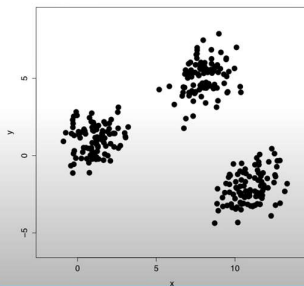


YouTube



k-Means

Cluster data in **k** clusters!



6:26 / 6:39 • Applications of Regression



YouTube



Flask:

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

Introduction to Flask

Flask!!

python™
powered

Enthusiasts named Pocco!

I'm learning Flask!

Mute (m)

25:37 / 38:50 • Introduction to Flask

Python Certification Training

www.edureka.co YouTube

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

Flask – Static Files

Web application will require a static file such as JS or CSS file

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

@app.route('/result')
def result():
    dict = {'phy':50,'che':60,'maths':70}
    return render_template('result.html', result = dict)

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

Python

```
<!doctype html>
<html>
  <body>

    <table border = 1>
      {% for key, value in result.iteritems() %}

        <tr>
          <th> {{ key }} </th>
          <td> {{ value }} </td>
        </tr>

      {% endfor %}
    </table>

  </body>
</html>
```

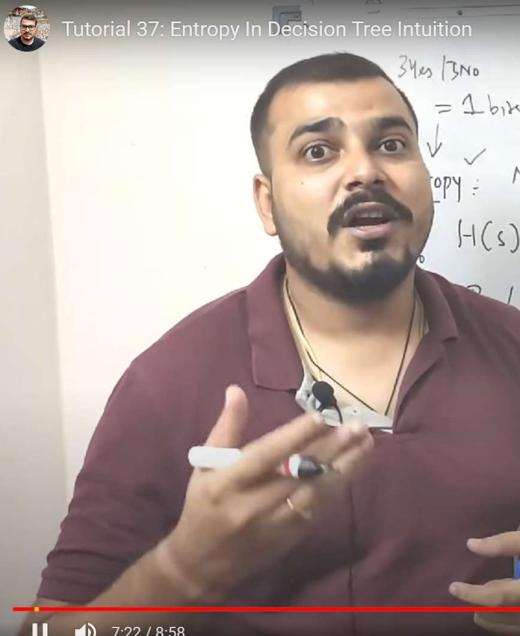
HTML

25:37 / 38:50 • Flask - Static Files

Python Certification Training

www.edureka.co YouTube

Entropy in Decision Tree Intuition:



DECISION TREE ENTROPY

34% / 30%
= 1 bits
↓ ✓
Entropy: Measures the purity of split

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

P_{+} / P_{-} : % of +ve class / % of -ve class

Subset of Training Example




$\frac{3}{5} \log_2(\frac{3}{5}) - (\frac{2}{5} \log_2(\frac{2}{5}))$
= 0.918 bits

94% / 6%
f1 f2 f3 o/p

```

graph TD
    f1((f1)) --> f2((f2))
    f1 --> f3((f3))
    f2 --> o1(( ))
    f2 --> o2(( ))
    f3 --> o3(( ))
    f3 --> o4(( ))
    o1 --> o1a((34% / 30%))
    o2 --> o2a((24% / 20%))
    o3 --> o3a((64% / 36%))
    o4 --> o4a((24% / 20%))
    
```

7:22 / 8:58

Decision Tree information gain:

Tutorial 38- Decision Tree Information Gain

DECISION TREE INFORMATION GAIN

ENTROPY

$H(S) = -P_1 \log_2 P_1 - P_2 \log_2 P_2 - \dots - P_n \log_2 P_n$
 $= -\frac{6}{8} \log_2 \frac{6}{8} - \frac{2}{8} \log_2 \frac{2}{8}$
 $= 0.91$

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

② Information Gain

$Gain(S, A) = H(S) - \sum_{v \in \text{VAL}} \frac{|S_v|}{|S|} H(S_v)$
 $Gain(S, f_1) = H(S) - \frac{6}{8} H(f_2) - \frac{2}{8} H(f_3)$
 $= 0.91 - \frac{6}{8} \times 0.81 - \frac{2}{8} \times 1$
 $= 0.049$

Tutorial 38- Decision Tree Information Gain

DECISION TREE INFORMATION GAIN

① ENTROPY

$H(S) = -P_1 \log_2 P_1 - P_2 \log_2 P_2 - \dots - P_n \log_2 P_n$
 $= -\frac{6}{8} \log_2 \frac{6}{8} - \frac{2}{8} \log_2 \frac{2}{8}$
 $= 0.91$

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

② Information Gain

$Gain(S, A) = H(S) - \sum_{v \in \text{VAL}} \frac{|S_v|}{|S|} H(S_v)$
 $Gain(S, f_1) = H(S) - \frac{6}{8} H(f_2) - \frac{2}{8} H(f_3)$
 $= 0.91 - \frac{6}{8} \times 0.81 - \frac{2}{8} \times 1$
 $= 0.049$

Gini Impurity intuition in depth in Decision Tree:

Tutorial 39- Gini Impurity Intuition In Depth In Decision Tree

GINI IMPURITY DT

f_1	f_2	f_3	%p.
C_1	D_1		Yes
C_2	D_2		No
			No
			Yes
			...

① Entropy

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

② GINI IMPURITY

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

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

$$= 1 - \left[\left(\frac{3}{6}\right)^2 + \left(\frac{3}{6}\right)^2\right]$$

$$= 1 - [0.25 + 0.25]$$

$$= 0.5$$

Diagram illustrating a decision tree split:

```

graph TD
    Root(( )) -->|f1| C1((C1))
    Root -->|f1| C2((C2))
    C1 -->|3Y/3N| Leaf1(( ))
    C2 -->|3Y/0N| Leaf2(( ))
    Leaf2 -->|dead Node| End(( ))
  
```

Entropy calculation for the split:

$$= -\frac{3}{6} \log_2 \frac{3}{6} - \frac{3}{6} \log_2 \frac{3}{6}$$

$$= 1 \checkmark$$

YouTube