

Text Representation

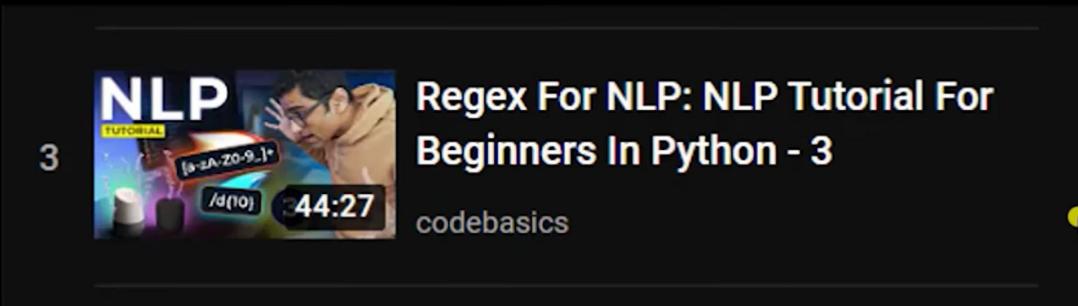


NLP Techniques

Rules & Heuristics

Machine
Learning

Deep Learning



3  44:27

Regex For NLP: NLP Tutorial For
Beginners In Python - 3

codebasics

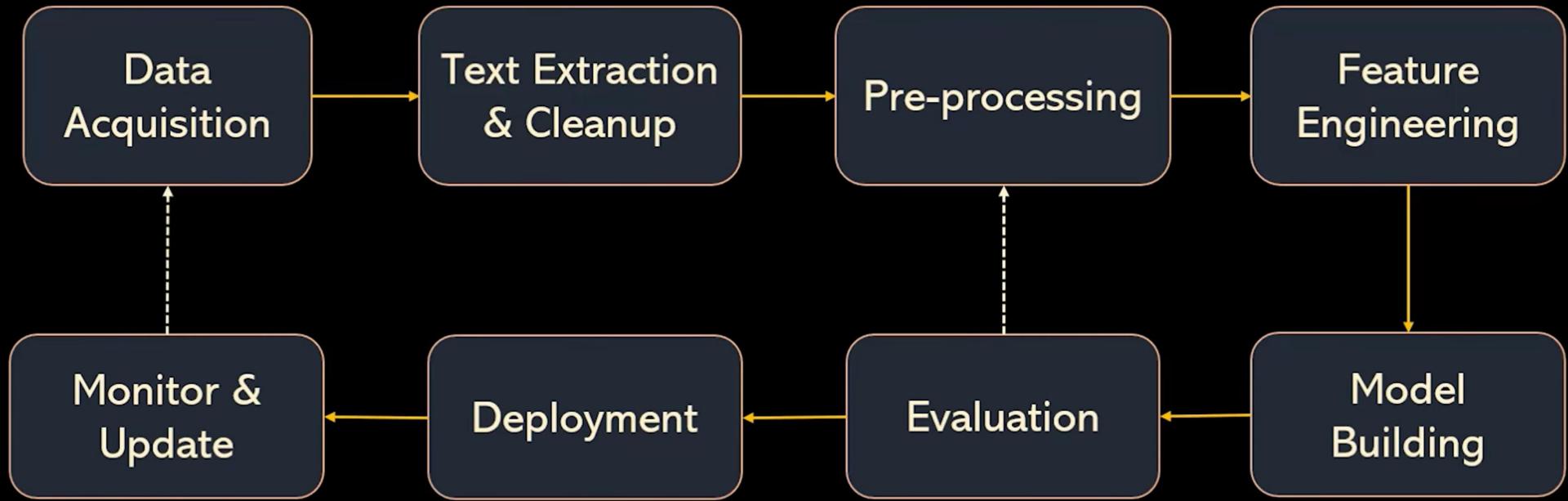
NLP Techniques

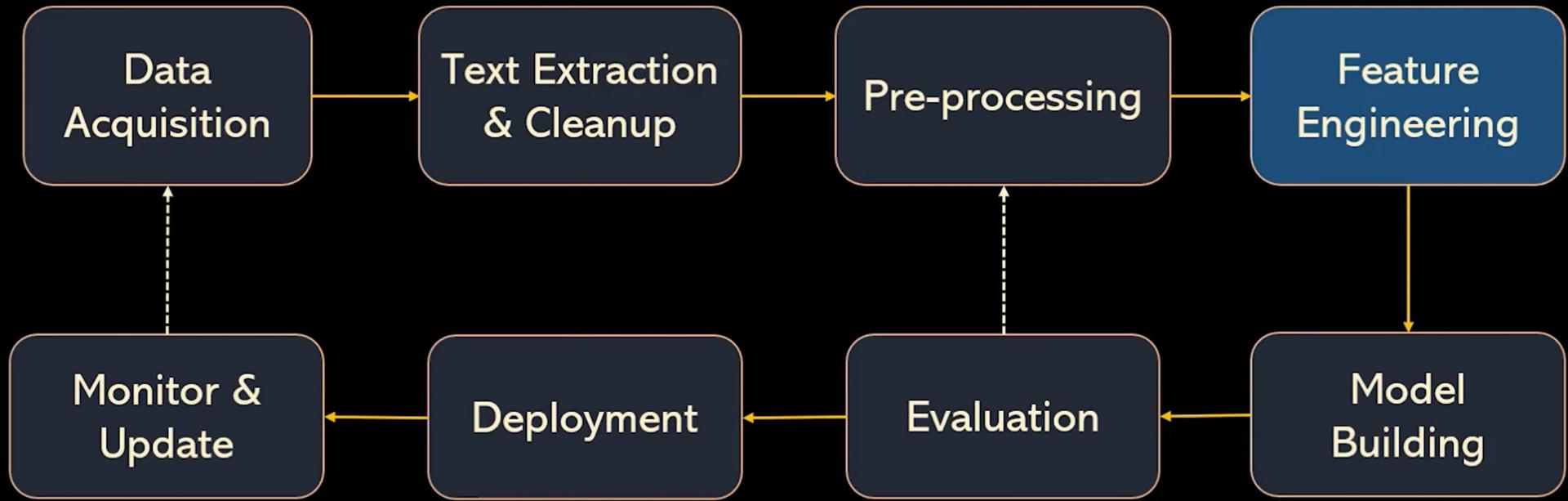
Rules & Heuristics

Machine
Learning

Deep Learning













Area

Facilities

Age

Location

Price



Features

Area

Facilities

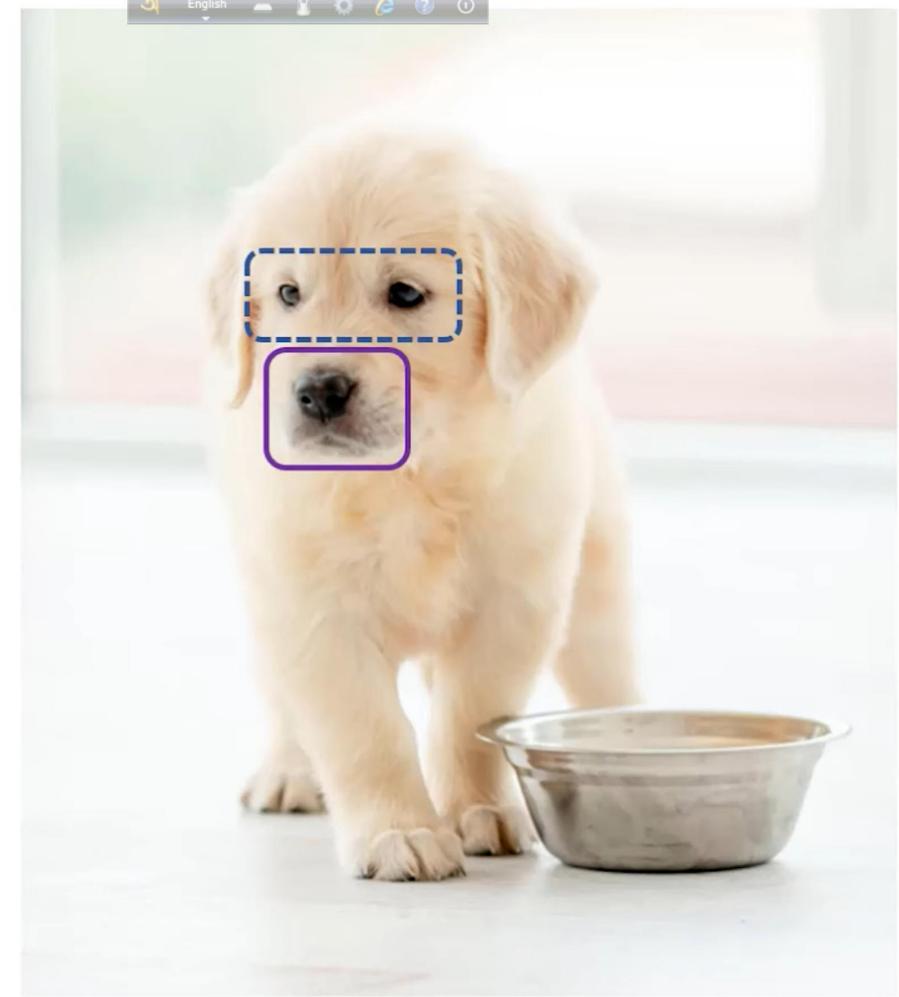
Age

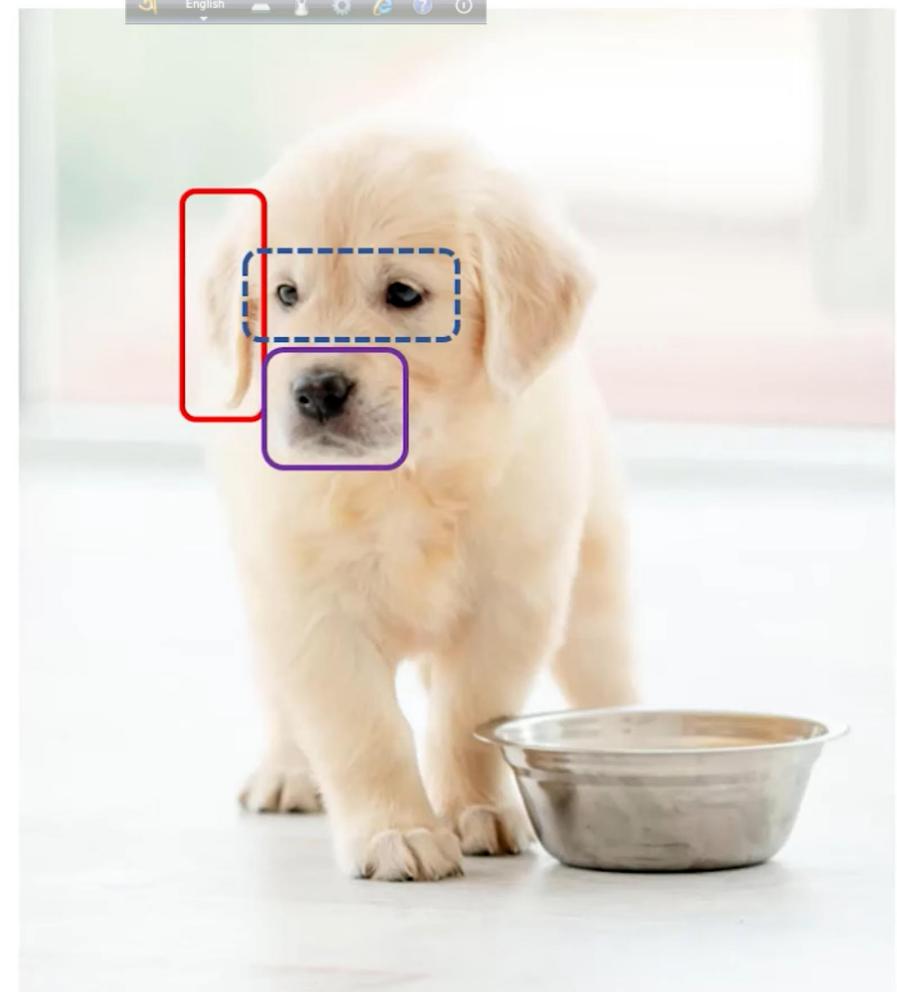
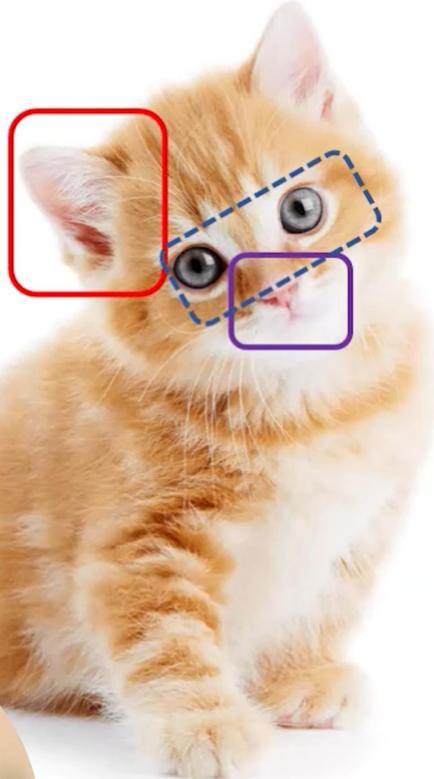
Location

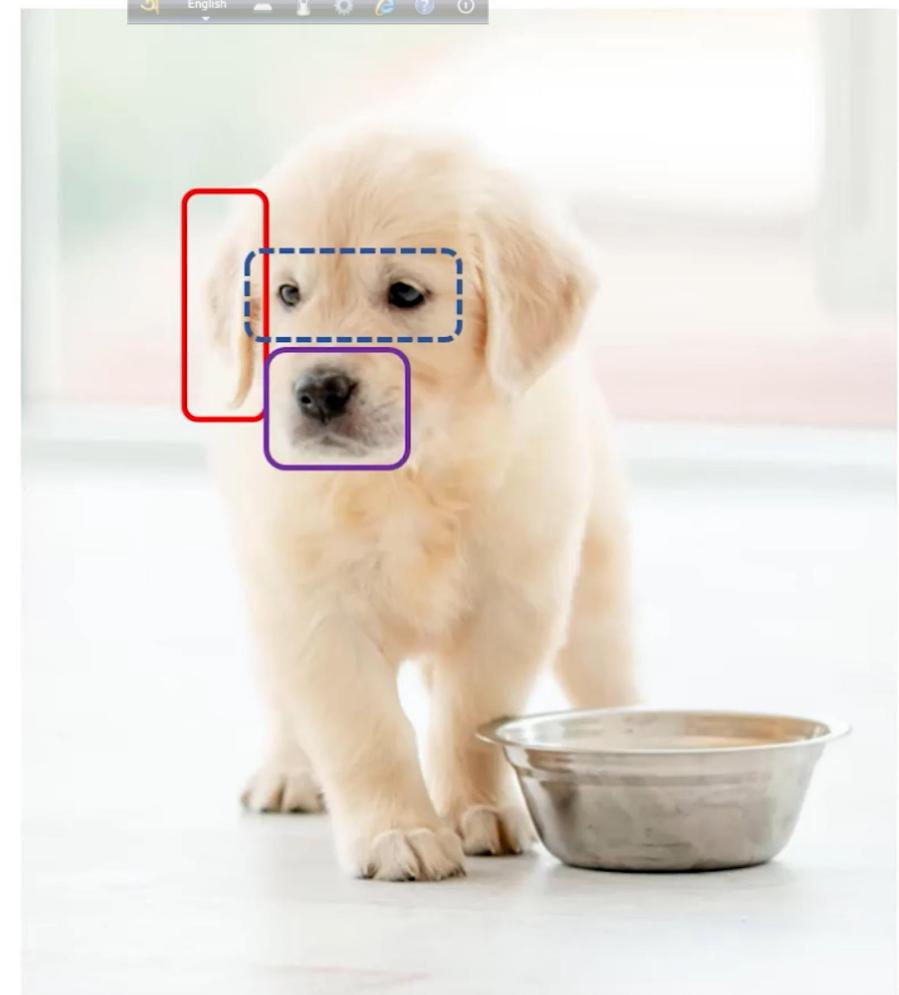
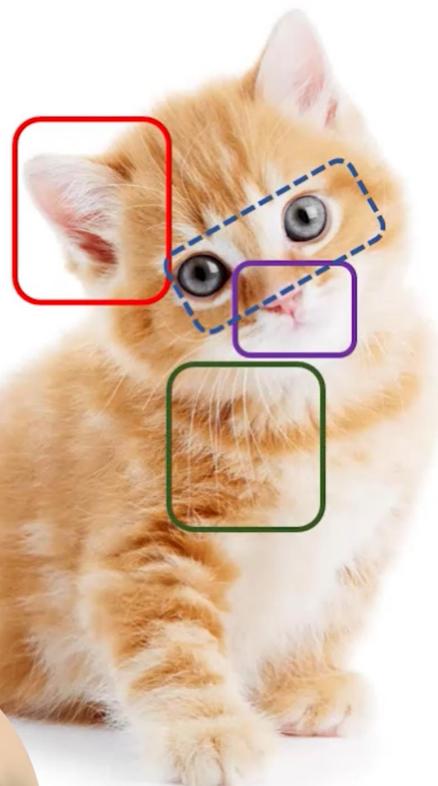
Price



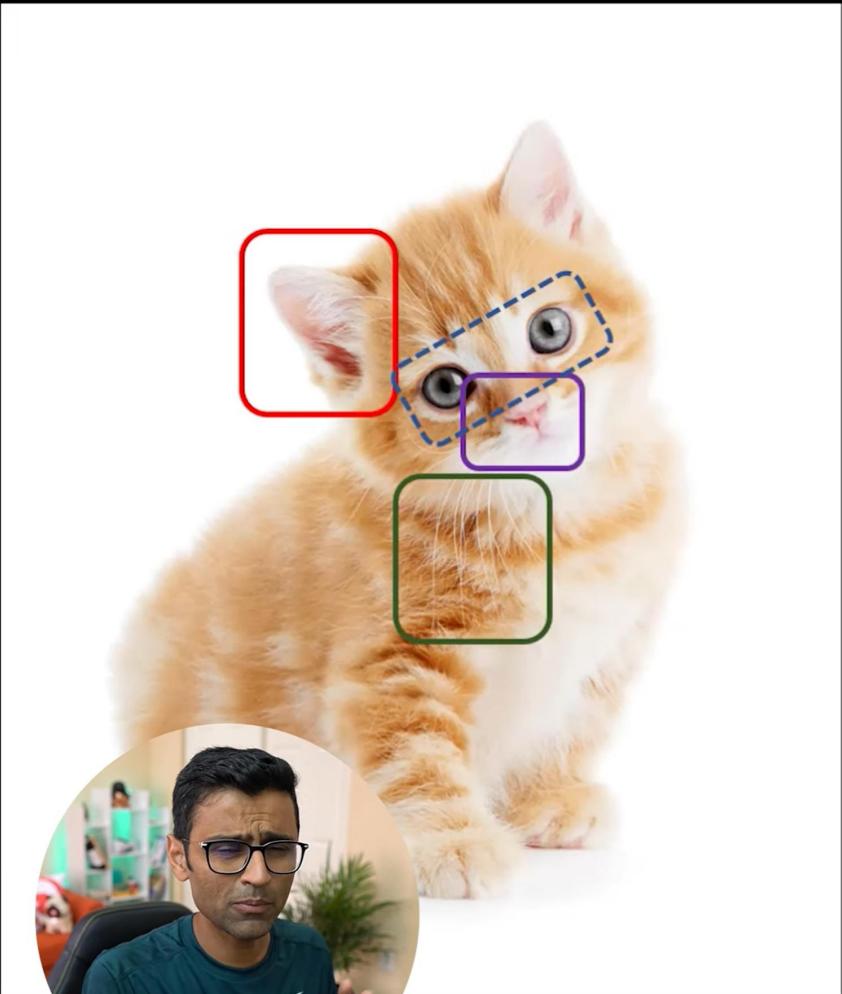








Features = Eyes, Nose, Whiskers, Ears



English

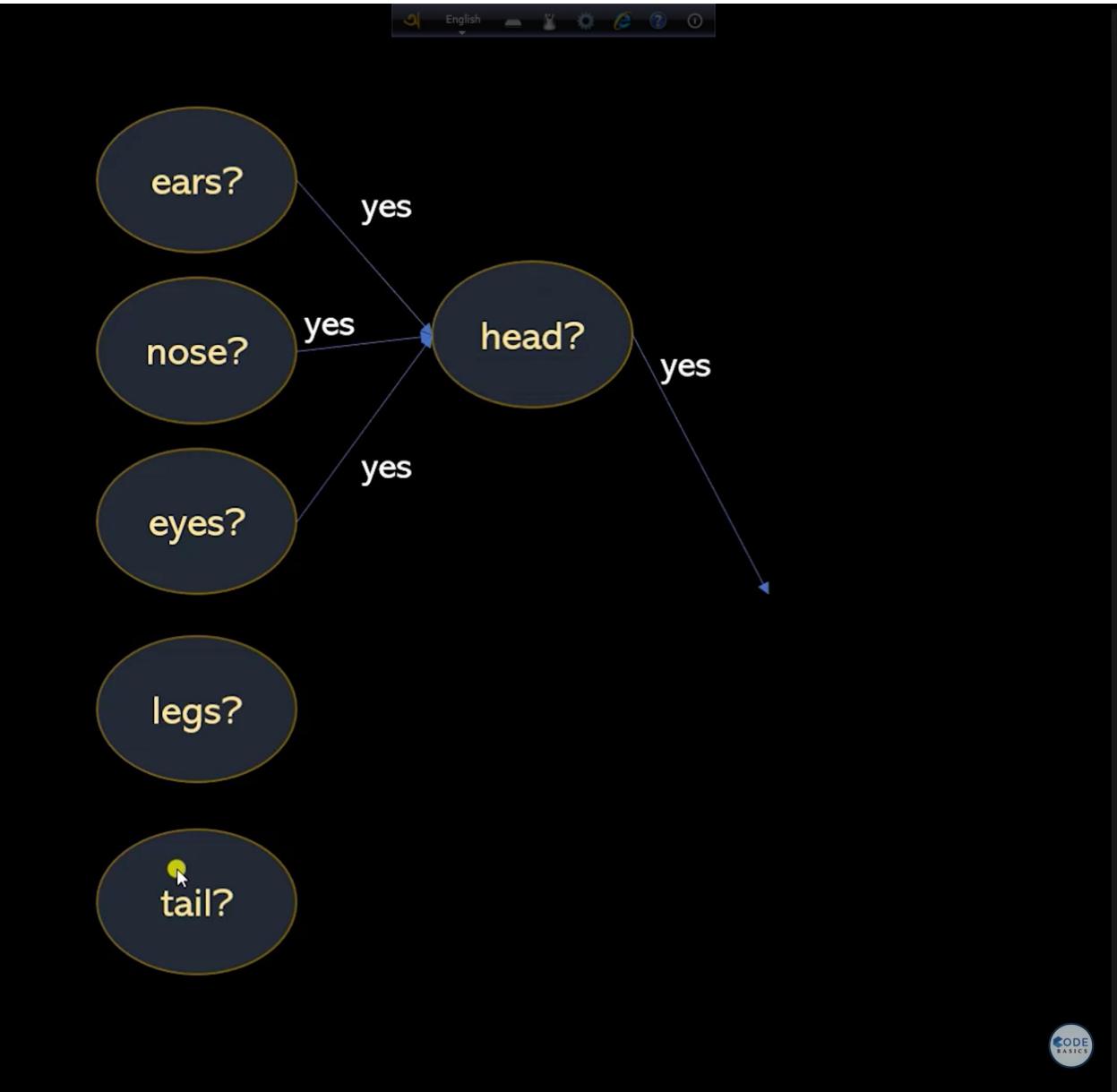
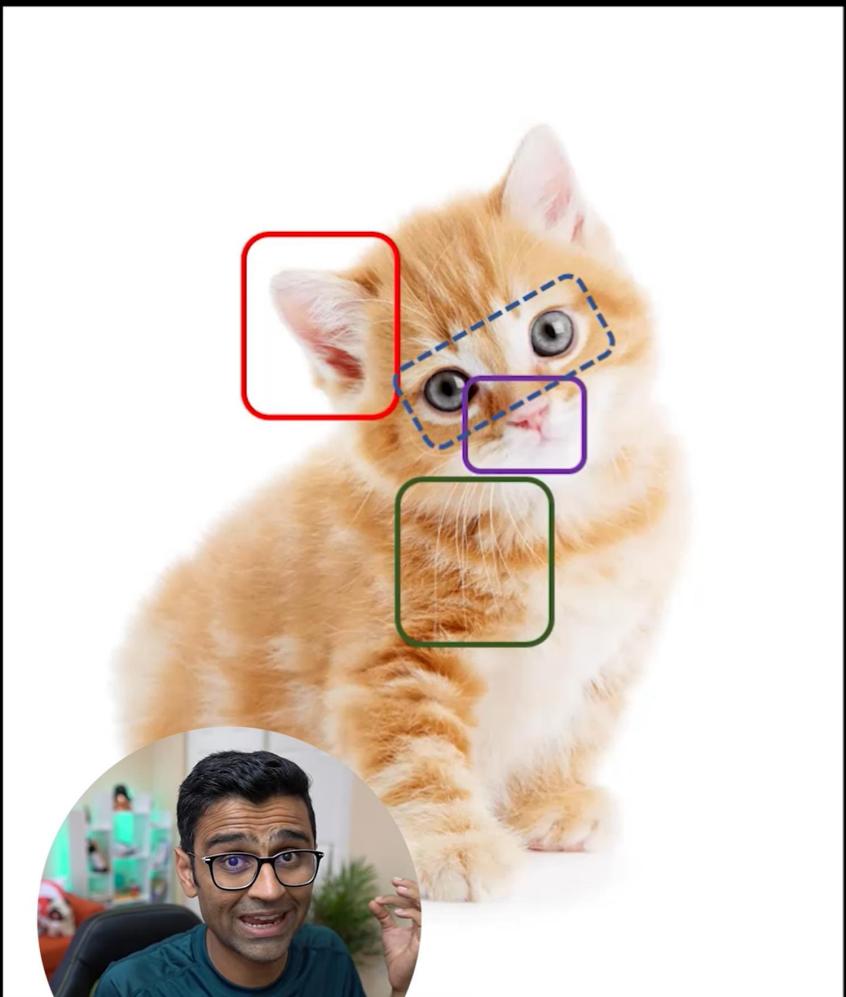
ears?

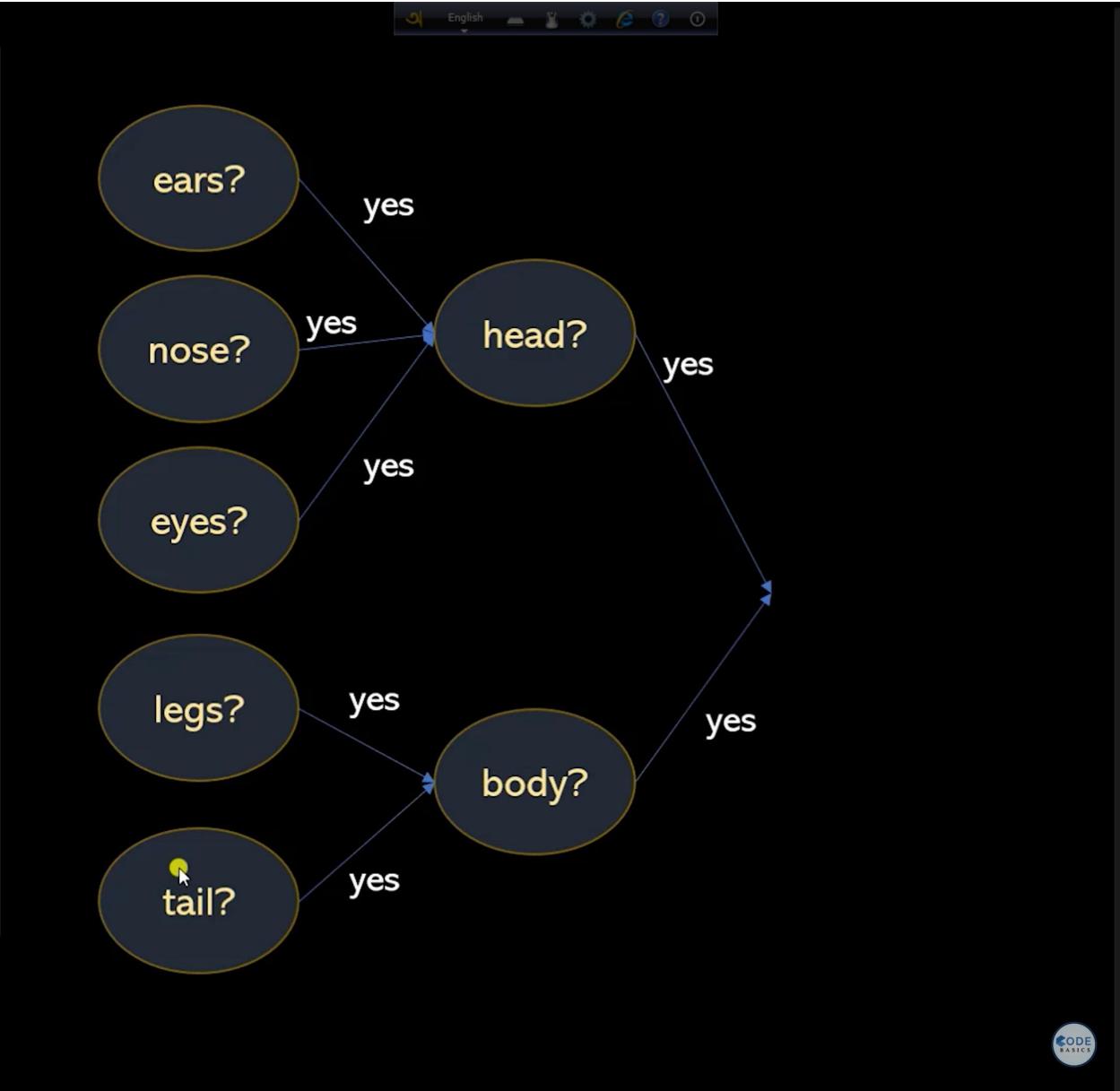
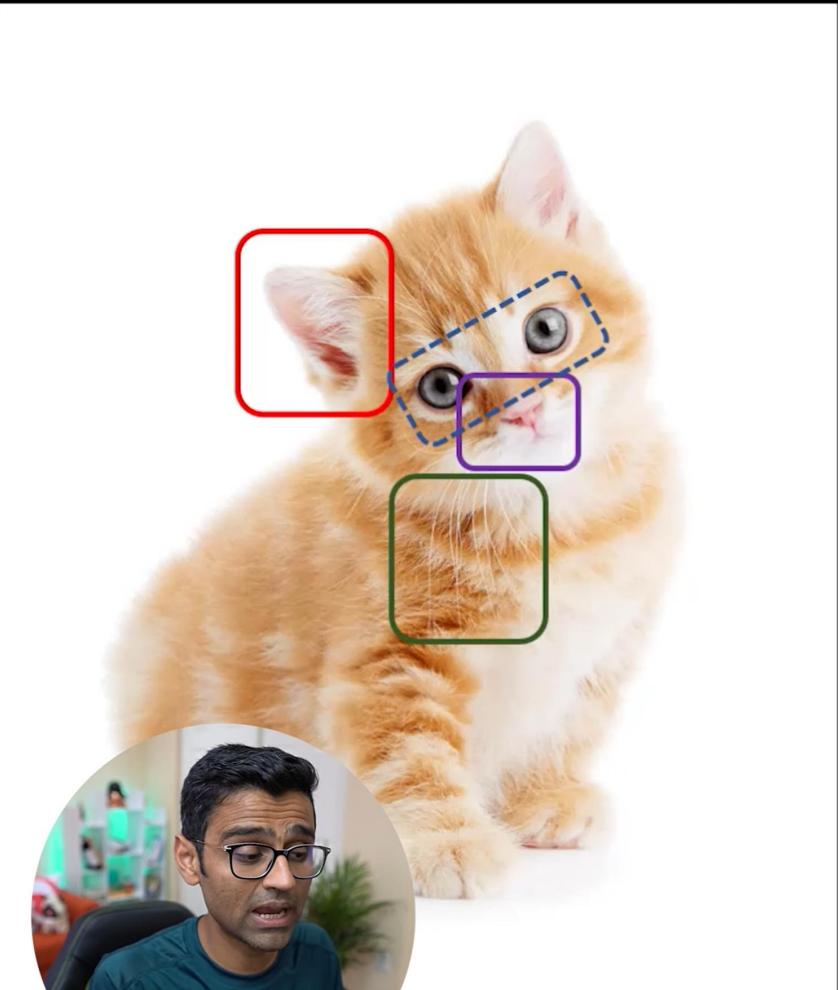
nose?

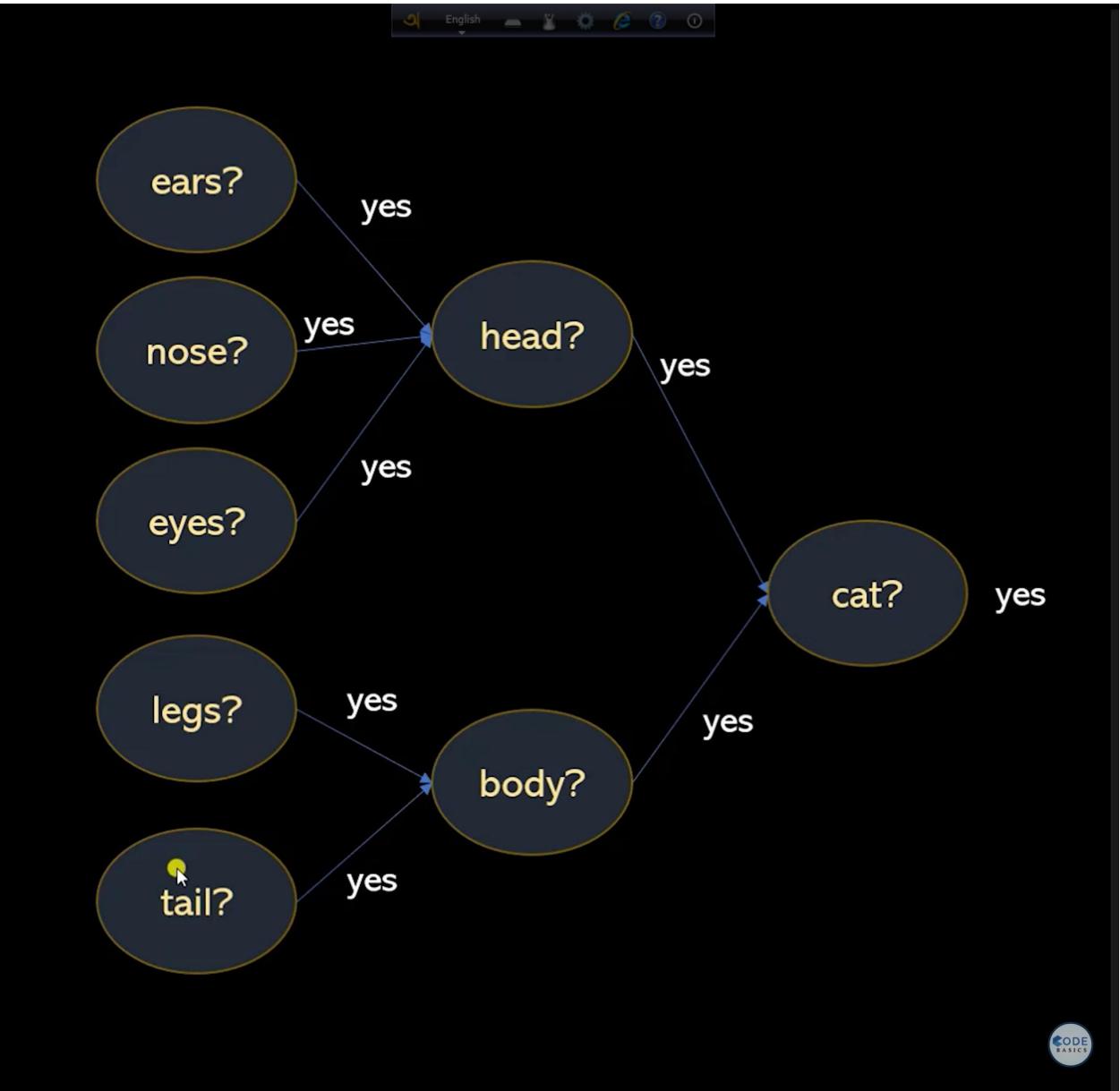
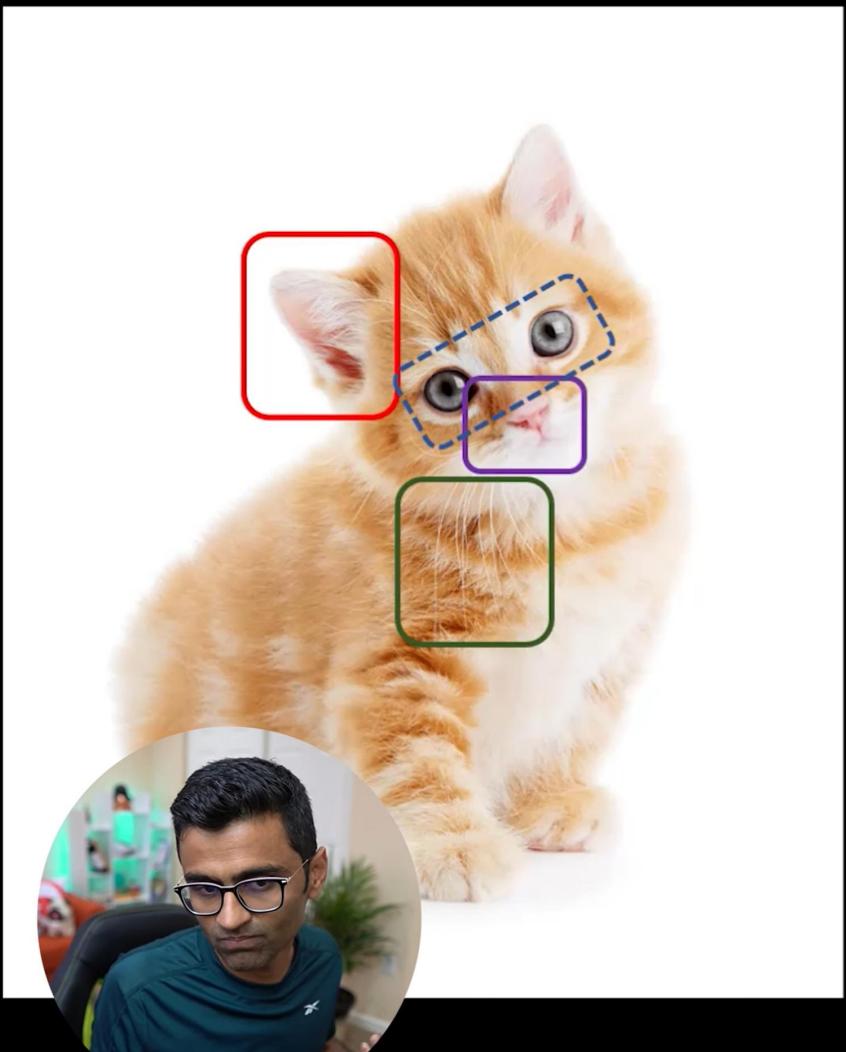
eyes?

legs?

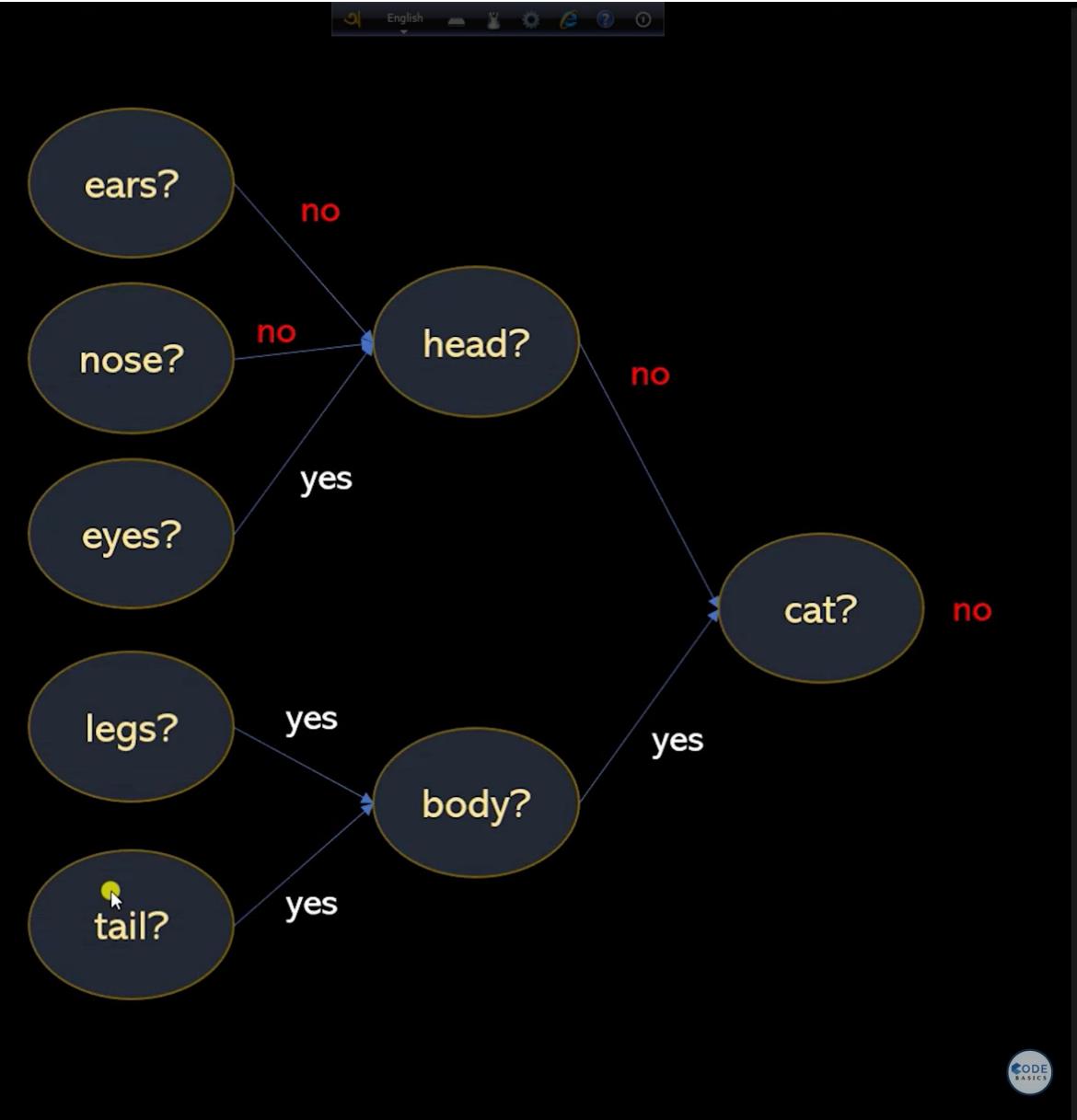
tail?









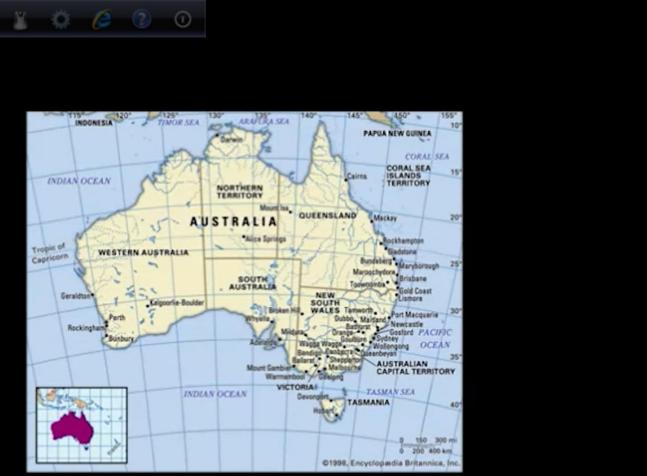




Dhoni



Cummins



Australia





Dhoni



Cummins



Australia

Person: 1

Healthy/fit: 0.9

Location: 0

Has two eyes: 1

Has government: 0

Person: 1

Healthy/fit: 0.87

Location: 0

Has two eyes: 1

Has government: 0

Person: 0

Healthy/fit: 0.7

Location: 1

Has two eyes: 0

Has government: 1



Image credits: cnn.com, Britannica.com and outlookindia.com



Dhoni



Cummins



Australia

$$\begin{bmatrix} 1 \\ 0.9 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0.87 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0.7 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$



Image credits: cnn.com, Britannica.com and outlookindia.com



Dhoni ----->





Dhoni

Person: 1
Healthy/fit: 0.9
Location: 0
Has two eyes: 1
Has government: 0

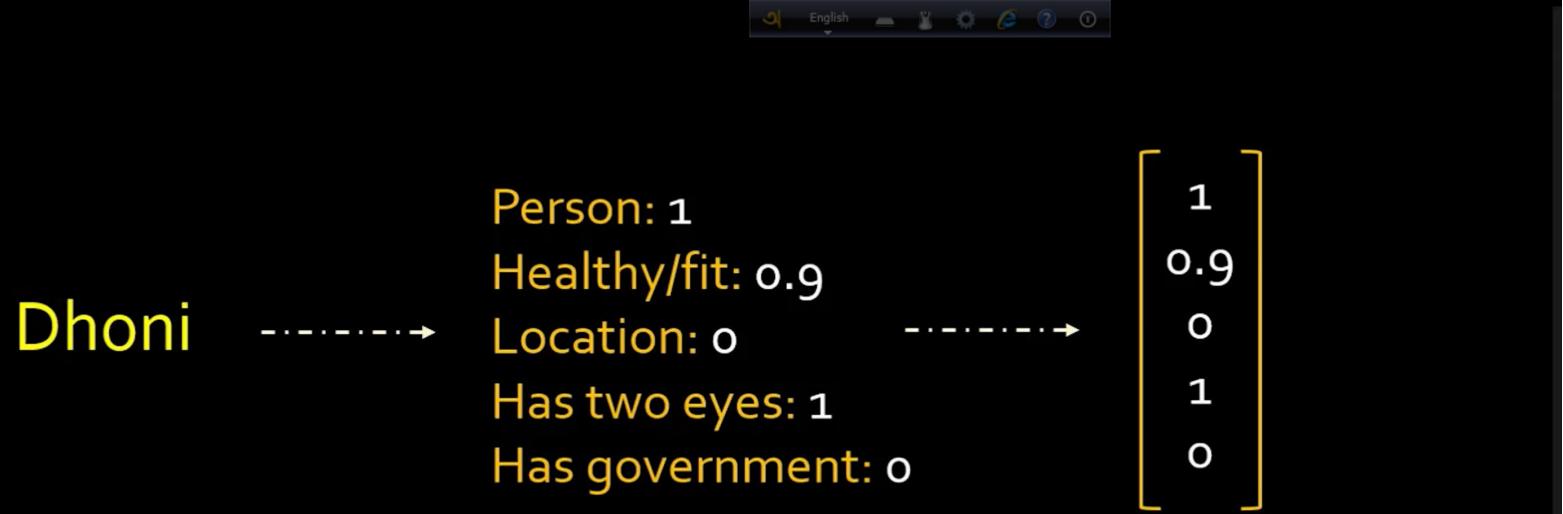




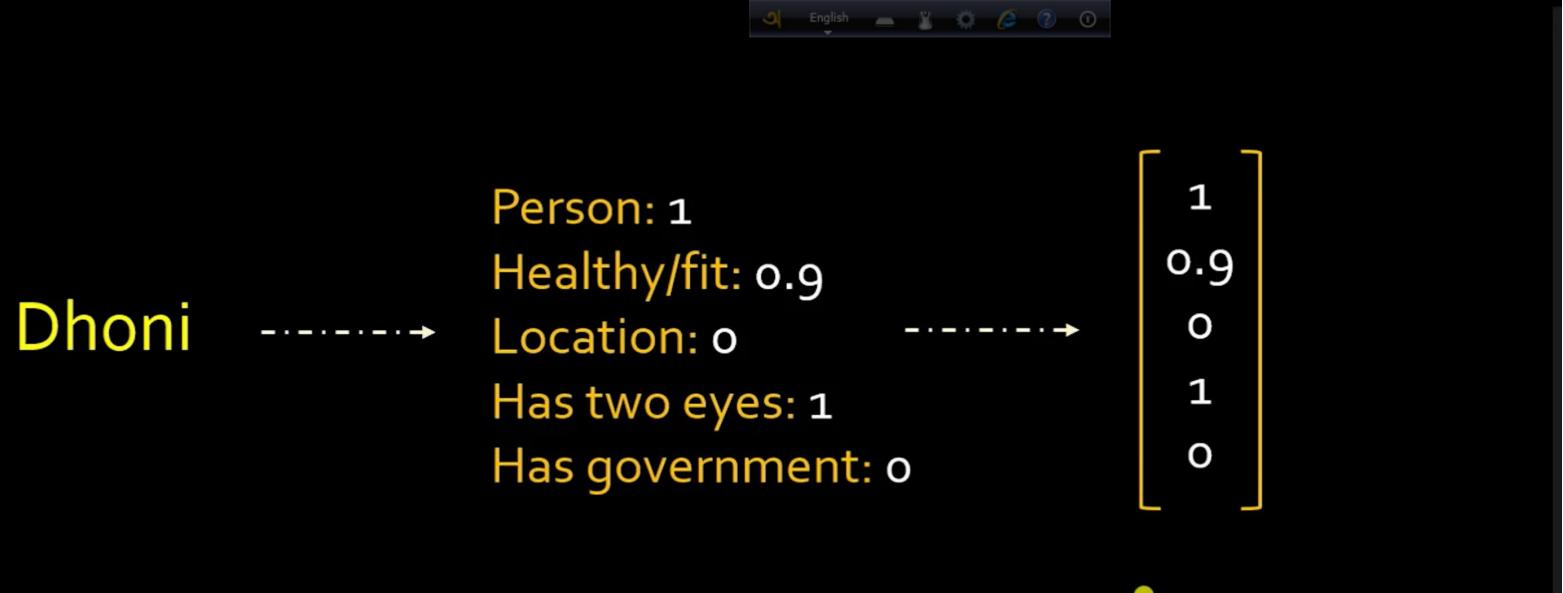
Dhoni

Person: 1
Healthy/fit: 0.9
Location: 0
Has two eyes: 1
Has government: 0

$$\begin{bmatrix} 1 \\ 0.9 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

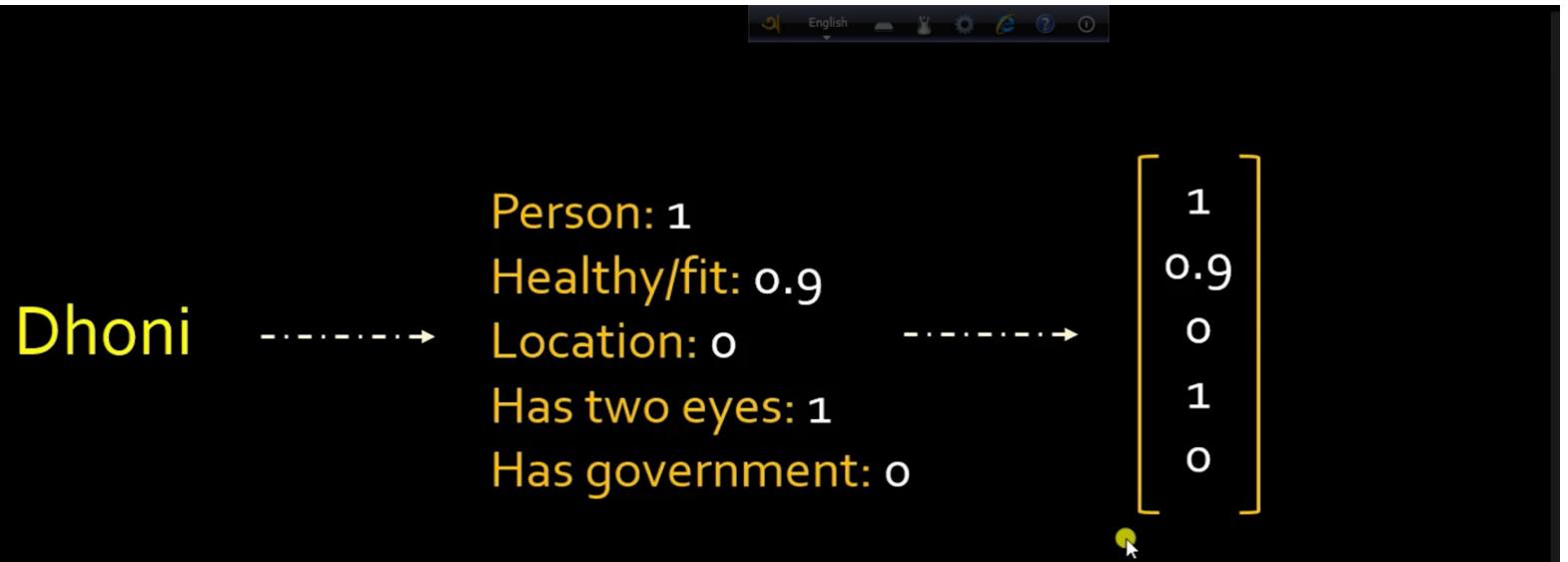


Feature Engineering is a process
of extracting features from raw
data



In NLP, we also call it **Text Representation**





In NLP, we also call it **Text Representation**



Representing text as a vector is also known as **Vector Space Model**

There are various approaches of converting text into vector



There are various approaches of converting text into vector

One Hot
Encoding

Bag Of Words

TF-IDF

Word
Embeddings



Often in NLP, feeding a good text representation to an ordinary algorithm will get you much farther compared to applying a top-notch algorithm to an ordinary text representation.



Quote from: Practical NLP Book