
MACHINE LEARNING PRACTICAL
LOGISTIC REGRESSION

AGENDA

- Data Collection || Reading Data
- Data Exploration
- Data Processing (Separating The Features From The Target)
- Splitting the data to training data & Test data
- Data Scaling (Data Standardization, Normalization)
- Build Model
- Model Evaluation
- Saving Model
- Build Test Method

✓ ABOUT DATA

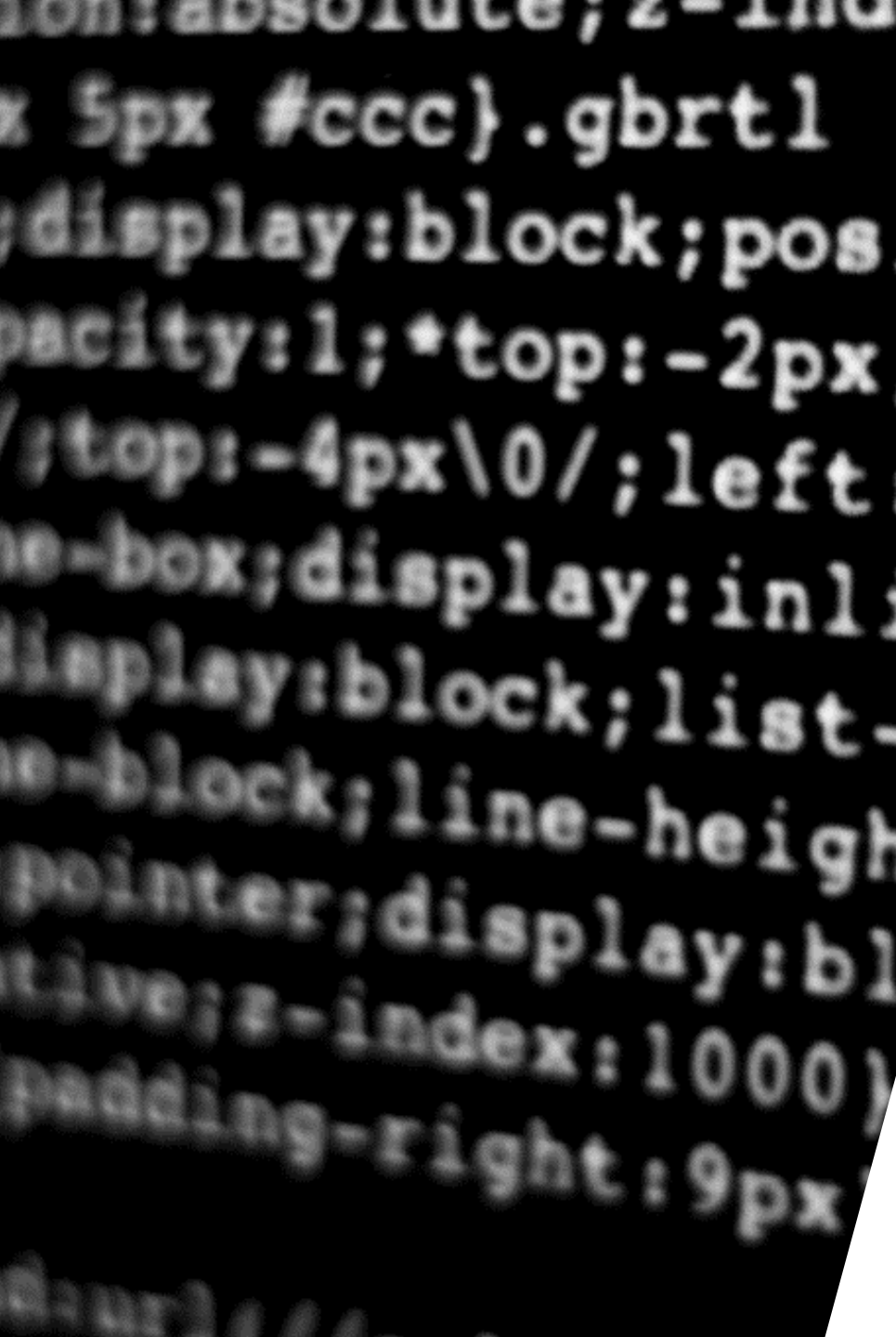
Parkinson’s Disease (PD) is a degenerative neurological disorder marked by decreased dopamine levels in the brain. It manifests itself through a deterioration of movement, including the presence of tremors and stiffness. There is commonly a marked effect on speech, including dysarthria (difficulty articulating sounds), hypophonia (lowered volume), and monotone (reduced pitch range). Additionally, cognitive impairments and changes in mood can occur

✓ SCALE DATA

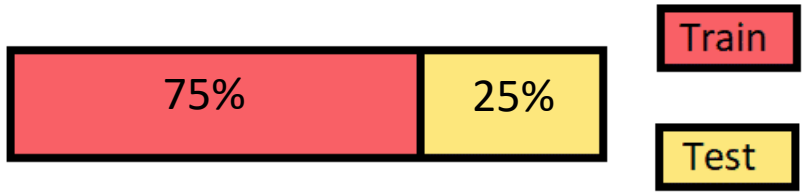
Separating The Features From The Target

Define X ➔ Input / Attributes of Data

Define Y ➔ Output of Dataset (Result)



TRAIN || TEST

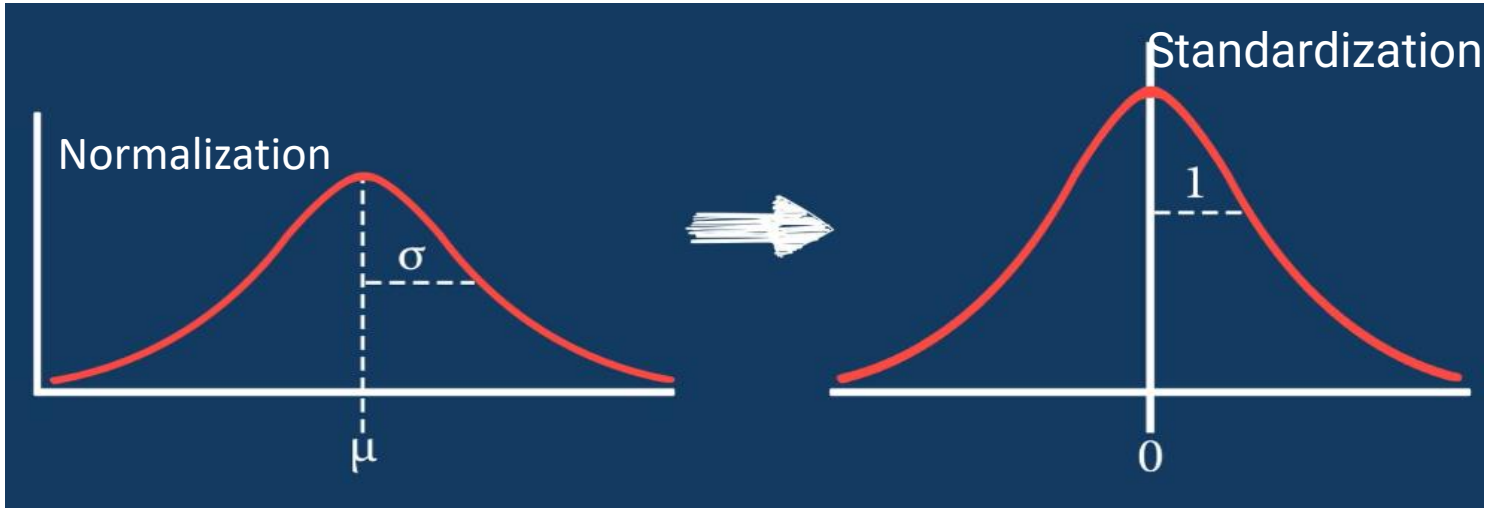


DATA STANDARDIZATION

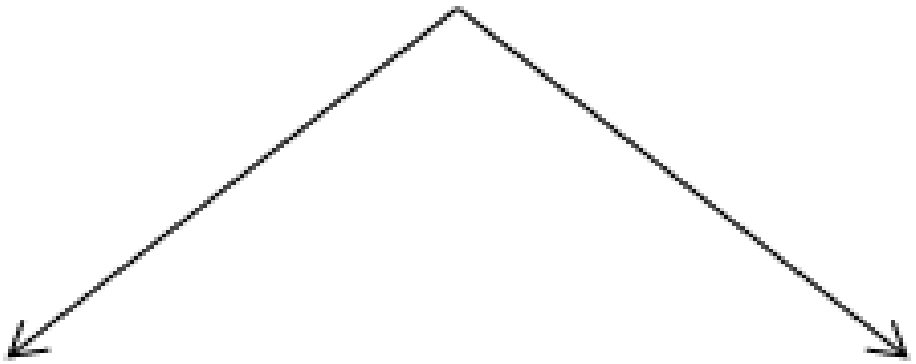
input (Gap Data) ➔ Run Slow

NORMALIZATION

Data Imbalance



Feature scaling



Normalization

$$X_{new} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Standardization

$$X' = \frac{\text{X} - \text{Mean}}{\text{Standard deviation}}$$

LABEL ENCODING

	name	gender	age	city
a	Abby	F	33	Berlin
b	Ben	M	16	Tokyo
c	Charlie	M	22	Sydney
d	Dave	M	65	York
e	Ella	F	18	Sydney

→

	name	gender	age	city
a	Abby	0	33	0
b	Ben	1	16	2
c	Charlie	1	22	1
d	Dave	1	65	3
e	Ella	0	18	1

```
df['Gender'].map({'M': 1, 'F': 0})
```

HOT ENCODING

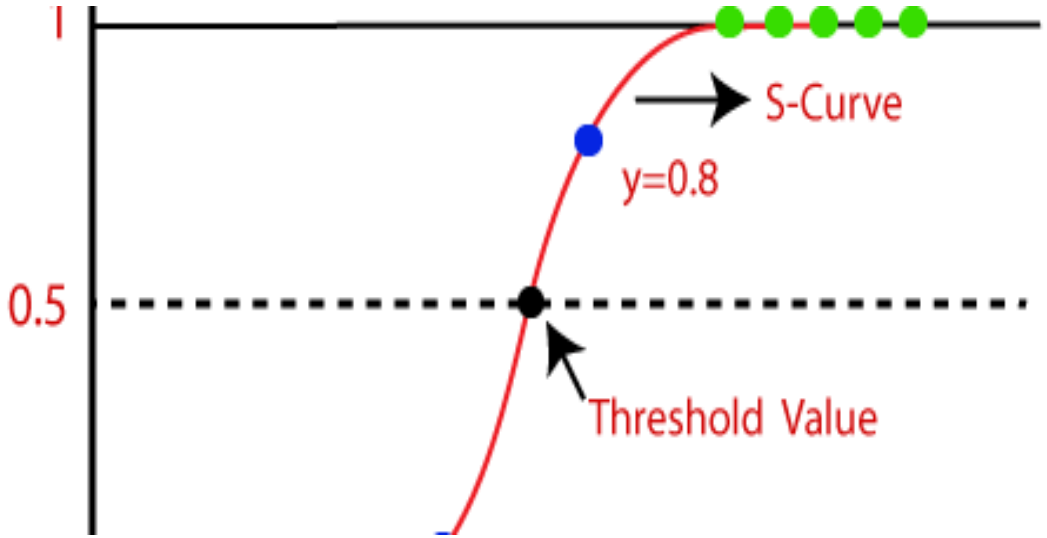
Color
Red
Green
Blue

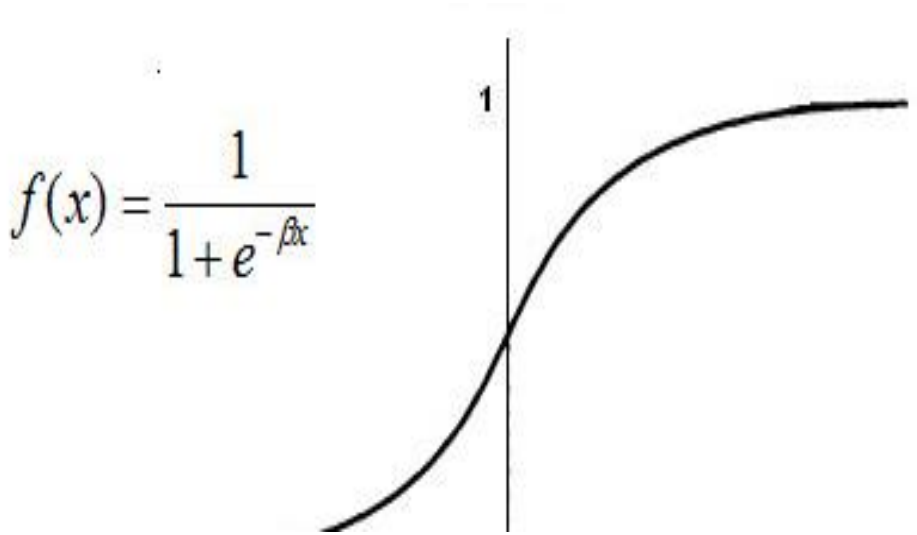
One-hot encoding →

d1	d2	d3
1	0	0
0	1	0
0	0	1

Distribution Attributes (Nationality(4) 4 Cols)

LOGISTIC REGRESSION

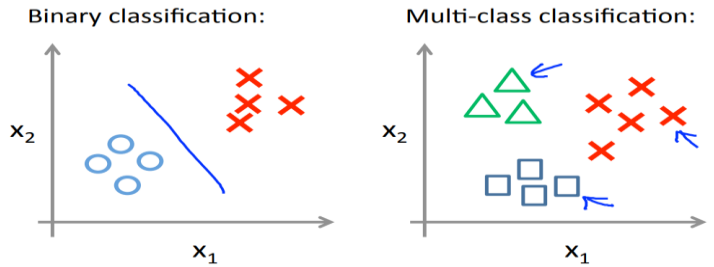




Logistic Function (Sigmoid)

- Logistic regression is used to handle the classification problems.
- Logistic Model Build on Probability , This type of statistical model
- Recall The Sigmoid Function

- Linear Regression (Squashing)
- Classification of (Binary Classifier & Multi Classifier)



SUMMARY

- Threshold classifier output $h_\theta(x)$ at 0.5:
  - If $h_\theta(x) \geq 0.5$, predict "y = 1"
  - If $h_\theta(x) < 0.5$, predict "y = 0"

Classification: y = 0 or 1

$h_\theta(x)$ can be > 1 or < 0

Logistic Regression: $0 \leq h_\theta(x) \leq 1$

Classification

THANK YOU !