

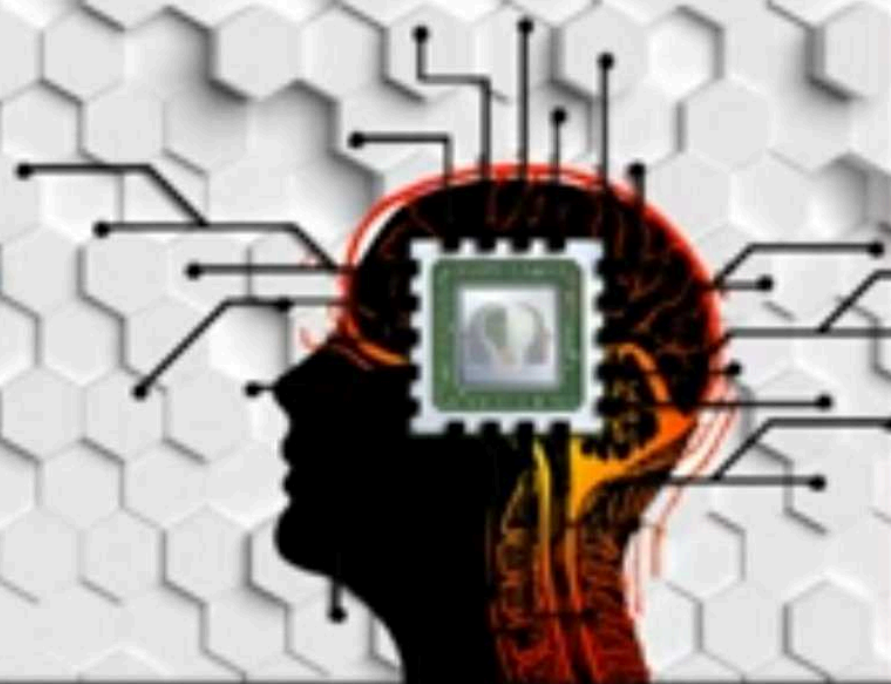
**Data Preprocessing  
Feature Engineering**

**&**

**Tutorial**

**E/हिं**

# Introduction of **Feature Engineering**



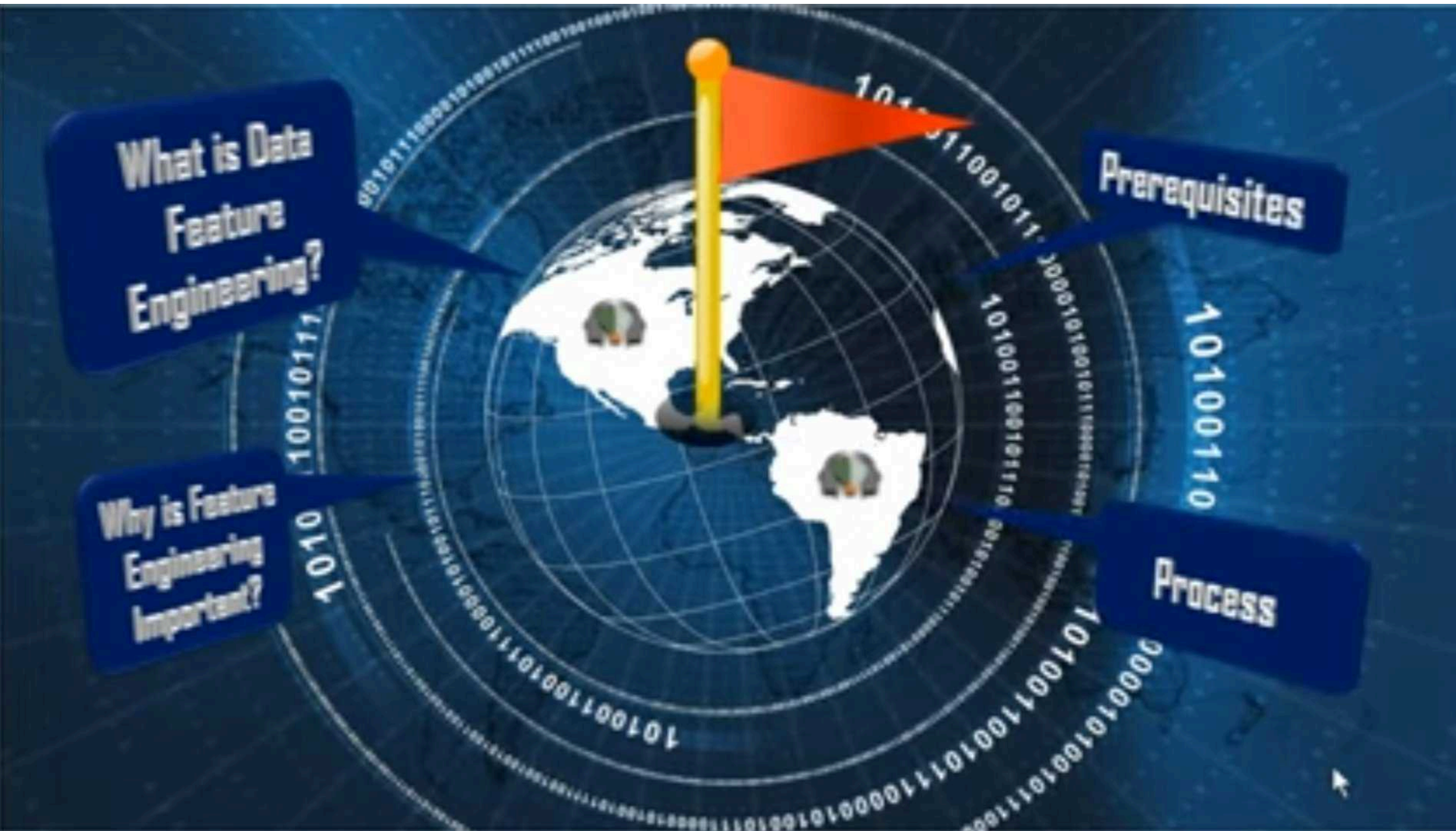
**Machine Learning & Data Science: Hands-on Python Course in Hindi**

What is Data  
Feature  
Engineering?

Why is Feature  
Engineering  
Important?

Prerequisites

Process





# What is Feature Engineering?

---



- **Feature** is an attribute or property shared by all of the independent units on which analysis or prediction is to be done.
- **Feature Engineering** is the process to create feature/extract the feature from existing features by domain knowledge to increase the performance of Machine Learning Model.



# What is Feature Engineering?

---



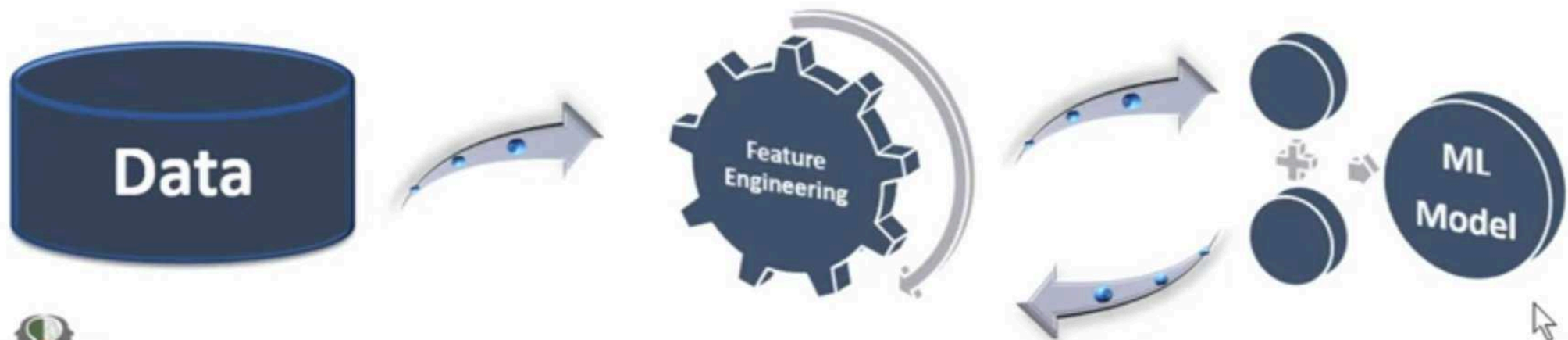
- **Feature** is an attribute or property shared by all of the independent units on which analysis or prediction is to be done.
- **Feature Engineering** is the process to create feature/extract the feature from existing features by domain knowledge to increase the performance of Machine Learning Model.
- Feature Engineering is an art.



# Why is Feature Engineering Important?



- Quality data always help to improve the accuracy and performance of Machine Learning model.
- Machine Learning algorithm follow the rule (*learn like Kids*)
  - GIGO – Garbage In Garbage Out





# What are the process of Feature Engineering?

---



- **Major Process of Feature Engineering:**

- Brainstorming or testing features
- Deciding what features to create
- Creating features
- Checking how the features work with your model
- Improving your features if needed
- Go back to brainstorming/creating more features until the work is done



# Examples of Feature Engineering?



Train Boarding Time	Train Reach at Station	Delayed or On Time	Delay Time in Minute
10.00 AM	10.15 AM	Delayed	00.15
05.00 PM	05.00 PM	On time	00.00



# Examples of Feature Engineering?



DateTime	Hour	Day	Month	Year	Day of Week
07-Apr-2020 12.00.00	12	7	4	2020	2
10-Apr-2020 23.04.00	23	8	4	2020	5



# Examples of Feature Engineering?



Age	Sex	Male	Female	Age range
23	Male	1	0	21-30
12	Female	0	1	11-20
34	Male	1	0	31-40
26	Female	0	1	21-30
15	Male	1	0	11-20
38	Male	1	0	31-40



# Prerequisites for Feature Engineering

---



- **Python Libraries**

- NumPy
- Pandas
- Matplotlib
- Seaborn
- Scikit Learn

- **Software**

- Anaconda
- Jupyter Notebook & Spyder



- **Mathematics**

- Statistics
- Probability
- Calculus
- Linear Algebra

- **Learning Platform**

- [www.YouTube.com/IndianAIProduction](https://www.YouTube.com/IndianAIProduction)

