Welcome Everyone

1. Temporal / Encoding data
2. Intro to ML ←
3. Maths for ML

Feature Engineering

1] Using Domain Knowledge

date → day, month, year

hour   min   → time of day

One, label encoding.

}as depth possible as we are aware of domain

2]

EDA $\Rightarrow$ doing analysis on provided data

Cusr   Income   USD Income

Label Encoding

Gender

Gender
M ⇒ number
F ⇒ number
O ⇒ number

⟹

Gender

0

1

2

2 > 0
other > Male

fit   (train)

transform   (test)

fit_transform

Gender

M $\longrightarrow$ 0

F $\longrightarrow$ 1

M $\longrightarrow$ 0

F $\longrightarrow$ 1

M $\longrightarrow$ 0

M $\longrightarrow$ 0

Airline		Airline_Ind	Airline_Jet	Airli_Vist
Indigo	0	1	0	0
Jet	1	0	1	0
Vistara	2	0	0	1
	⇓			

adding a order

Airline — Jet   > Airline — Jet

True   False

Fair ✓

Data Science Case Study

1] Hypothesis testing Practical ⇒ Hypothesis
2] Housing Dataset ⇒ EDA
3] Car dekho ⇒ EDA, feature
4] Airline ⇒ Temporal, FE, Encoding
                                    ↙   ↘
                                 Label   OHE
                                    ↙   ↘
                                 Target   ohe