

January 17, 2023

DSML: CC Fundamentals.

# Feature Engineering - 1.

## Recap:

- (a) Hypothesis testing -  $\alpha$ , p-value, Null, Alternate.
- (b) Z-test } 1 sample, 2 sample, independent, paired,
- (c) t-test } left-tailed, right-tailed, 2-tailed.
- (d) R-S test
- (e) ANOVA
- (f) Kruskal
- (g) Pearson  $r$
- (h)  $\chi^2$  test



Class begins @ 9:05 p.m.

## agenda:

- 1] Feature Engineering - 2.
- 2] Experiment Design - 1
- 3] Computer vision. - Opencv - 1
- 4] Natural language processing - 2.  
(spacy)

What is a feature?

$$\rightarrow X = \{ \bar{x}_i \in \mathbb{R}^d \}_{i=1}^n$$

Example 1: Height, Weight, Obesity.

	<u>height (cm)</u>	<u>weight (kg)</u>	<u>Obesity</u>
①	165.2	75	
②	170.2	62	
③	144	45	

co



Obesity

Example 2:

I PL matches.  $\rightarrow$  20 over match.

	<u>Overs</u>	<u>Runs</u>	<u>Wickets</u>	<u>Team(Bat)</u>	<u>Team(Bowl)</u>	<u>Win</u>
Match 1	1					
	2					
	3					
	4					
	5					
	$\vdots$					
	20					

