

What is Data Science ?

'Full Stack Data Science' or 'End to End Data Science'

Big data → Hadoop (Engineering side) How to store & use huge amount of data

✓ Data Engineering → SQL queries → collect data, merge & give it to us

✗ Machine Learning → What do you do with that data → model

Deployment. → interface i.e creating and application

i.e: Create a application for customer to use

→ Flask, AWS, Heroku, Microsoft Azure.

Lifecycle of Data Science

Starts with a problem that adds value to the business

No Data Science problem



only Business problem that is solved by Data Science.

ex: google earns from advertisement i.e Customer Specific

↳ 'click on the advertisement'

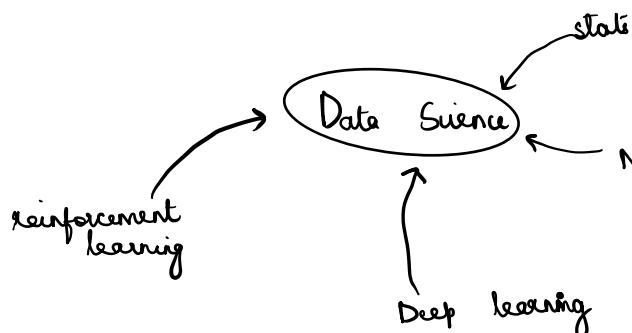
therefore they need to know you i.e your profile and preference.

per click they get 300 rupee.

Convert Business problem to Data Science problem

Data Science is a small problem of A.I

↳ Automation / reduce human effort using Data



step 1 : Think of possible causes ←

Step 2 : check our data

is our assumption
correct ?

step 3 : Where to get data ?

data Storage / hive

data Engineer → SQL query to get the data

Problems Data Quality

 Data Mismatch

 Record Missing

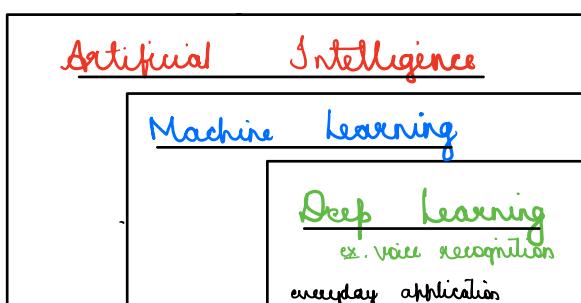
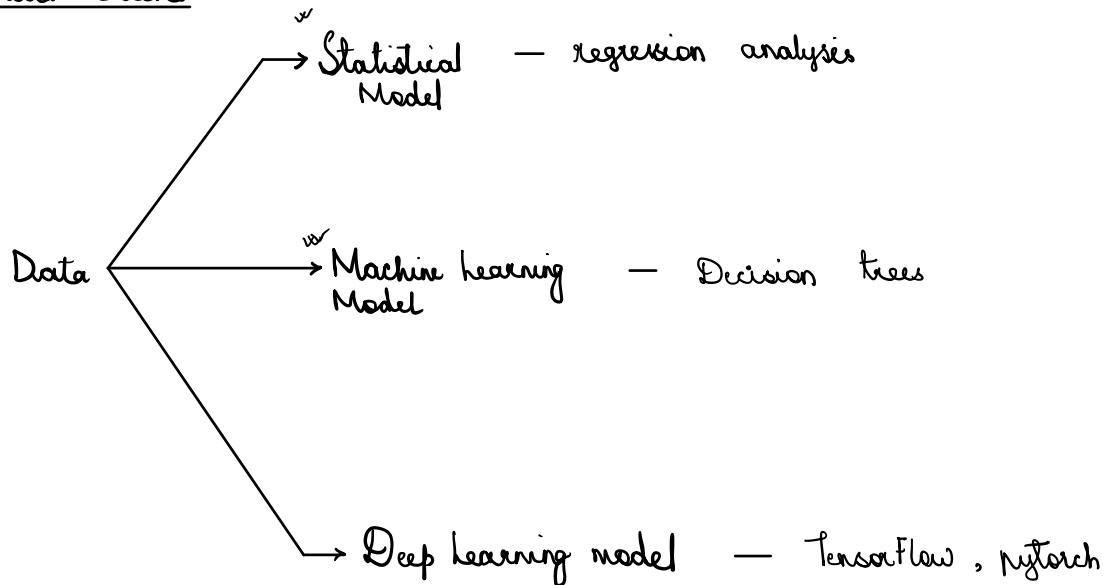
Extract Transfer Load → ETL

Database

Organisation collect data on different level



Data Science



Deployment : Heroku, AWS, Flask, Microsoft Azure

This is what the customer sees.

How do you come up with the reasons

Subject Matter Expert (SME)

→ will give you reasons so you can collect data to corroborate with the assumption

I Probability

ex. Ludo

$$\begin{array}{l} 6+1 \\ 5+2 \\ 4+3 \\ 3+4 \\ 5+2 \\ \hline 6 \end{array} \quad 6/36 = \underline{\underline{1/6}}$$

$$\Rightarrow \frac{\text{No. of cases}}{\text{total no. of cases}} = \frac{6}{6 \times 6} = \frac{6}{36} \Rightarrow \text{Prob} = \underline{\underline{1/6}}$$

Probability should not be greater than 1

Probability : $\frac{\text{Possible cases}}{\text{total no. of cases}}$

$$0 \leq P \leq 1$$

ex 2. 2 coins \rightarrow HT } HH = $\frac{2}{4} = \frac{1}{2}$.

TH
HT
TH
TT

ex 3. 2, 3, 5 \rightarrow $\frac{3}{6} = \frac{1}{2}$ $P(A) = \underline{\underline{\frac{1}{2}}}$

events

Mutually Exclusive events (MEE)

→ when something happens the other absolutely cannot happen.

ex 1. Toss a coin

$$A = \{ \text{outcome is H} \} \quad \therefore A \text{ & } B \text{ are mutually exclusive events}$$

$$B = \{ \text{outcome is T} \}$$

ex 2. Toss a dice

$$A = \{ \text{outcome is an even number} \}$$

$$B = \{ \text{outcome is an odd number} \}$$

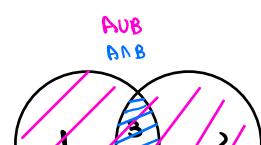
$\therefore A \text{ & } B \text{ are mutually exclusive events}$

Probability of mutually exclusive events

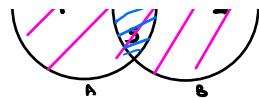
set : It's a collection of well defined and unique objects.

ex. $A = \{1, 3, 5\}$

$B = \{2, 3, 5\}$



$$A \cup B = \{1, 2, 3, 5\} \Rightarrow \text{Union}$$



$$A \cap B = \{3, 5\} \Rightarrow \text{Intersection}$$

$A^c = U - A$, everything other than A

$B^c = U - B$, everything other than B

for mutually exclusive events

$$P(A \cup B) = P(A) + P(B)$$

$$\downarrow A = \{2, 4, 6\}$$

$$B = \{1, 3, 5\}$$

$$A \cup B = \{1, 2, 3, 4, 5, 6\} =$$

$$P(A) = 3/6 = 1/2$$

$$P(B) = 3/6 = 1/2$$

$$P(A \cup B) = 6/6 = 1$$

$$P(A) + P(B) = 1/2 + 1/2 = 2/2 = 1 = P(A \cup B)$$

Not mutually exclusive

$$A = \{2, 3, 5\}, \quad B = \{1, 3, 5\}$$

$$P(A) = 3/6 = 1/2 \quad P(B) = 3/6 = 1/2$$

$$P(A \cup B) = \{1, 2, 3, 5\} = 4/6 = 2/3$$

$$1/2 + 1/2 = 1 \neq 2/3 = P(A \cup B)$$

\therefore for not mutually exclusive events : $P(A \cup B) \neq P(A) + P(B)$

Independent Events

Occurrence of one event doesn't affect the occurrence of the other event.

and
 $P(A \cap B) = P(A) \cdot P(B)$

occurrence of one event doesn't affect the occurrence of other .

for independent events

$$P(A \cap B) = P(A) \cdot P(B)$$

Prob : S_1

S_2

$\rightarrow A = \{S_1 \text{ will pass}\} \Rightarrow A^c = \{S_1 \text{ will fail}\}$

$$P(A) = 1/2$$

$$P(A^c) = 1/2$$

$\rightarrow B = \{S_2 \text{ will pass}\}, B^c = \{S_2 \text{ will fail}\}$

$$P(B) = 2/3$$

$$P(B^c) = 1/3$$

$$\text{Probability both fail} \Rightarrow P(A^c \cap B^c) = 1/2 \cdot 1/3 = 1/6$$

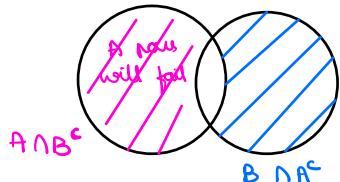
$$\text{Probability both pass} \Rightarrow P(A \cap B) = 1/2 \cdot 2/3 = 2/6 = 1/3$$

$$\text{exactly one will pass} \xrightarrow{(1)} \text{A pass \& B fails} = 1/2 \cdot 1/3 = 1/6$$

$$\textcircled{2} \ A \text{ fail } \& \ B \text{ pass} = \frac{1}{2} \cdot \frac{2}{3} = \frac{2}{6} = \frac{1}{3}$$

either $\textcircled{1}$ or $\textcircled{2}$ ie $\textcircled{1} \cup \textcircled{2}$

$$\frac{1}{6} + \frac{1}{3} = \frac{1}{6} + \frac{2}{6} = \frac{3}{6} = \frac{1}{2} //$$



$$\begin{aligned} P((A \cap B^c) \cup (B \cap A^c)) &= P(A \cap B^c) + P(A^c \cap B) \\ &= P(A) \cdot P(B^c) + P(A^c) \cdot P(B) \\ &= \frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} \\ &= \underline{\underline{\frac{1}{2}}} \end{aligned}$$

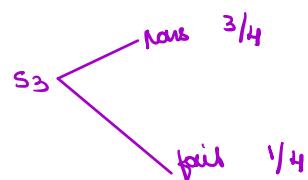
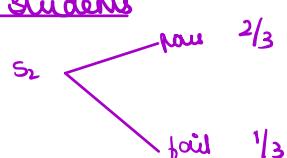
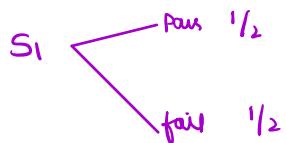
At least one will pass \rightarrow one will pass or Both will pass
 \rightarrow therefore Union

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

go through definitions

solve the following problem

Prob : There are three students



- i) all of them will fail
- ii) all of them will pass
- iii) exactly one of them will pass

iv) exactly two of them will pass.

solve : i) all fail

S_1 fail, S_2 fail & S_3 fail

Independent events

$$\therefore P(S_{1f} \cap S_{2f} \cap S_{3f}) = \frac{1}{2} * \frac{1}{3} * \frac{1}{4} = \frac{1}{24}$$

ii) all will pass

S_1 pass, S_2 pass & S_3 pass

Independent events

$$\therefore P(S_{1p} \cap S_{2p} \cap S_{3p}) = \frac{1}{2} * \frac{2}{3} * \frac{3}{4} = \frac{1}{4}$$

iii) Exactly one will pass

$$S_1 \text{ pass } \& S_2, S_3 \text{ fail} = \frac{1}{2} * \frac{1}{3} * \frac{3}{4} = \frac{1}{8}$$

$$S_2 \text{ pass } \& S_1, S_3 \text{ fail} = \frac{1}{2} * \frac{2}{3} * \frac{1}{4} = \frac{1}{12}$$

$$S_3 \text{ pass } \& S_1, S_2 \text{ fail} = \frac{1}{2} * \frac{1}{3} * \frac{2}{4} = \frac{1}{12}$$

$$\frac{1}{24} + \frac{1}{12} + \frac{1}{12} = \frac{1}{24} + \frac{2}{24} + \frac{12}{24} = \underline{\underline{\frac{15}{24}}}$$

iv) Exactly two will pass

$$S_{1f}, S_{2p}, S_{3f} = \frac{1}{2} * \frac{2}{3} * \frac{1}{4} = \frac{1}{12}$$

$$S_{1p}, S_{2f}, S_{3f} = \frac{1}{2} * \frac{1}{3} * \frac{1}{4} = \frac{1}{24}$$

$$S_{2p}, S_{3p}, S_{1f} = \frac{1}{2} * \frac{2}{3} * \frac{1}{4} = \frac{1}{12}$$
$$= \frac{1}{12} + \frac{1}{24} + \frac{1}{12}$$

$$= 0.4583$$

Matrix algebra

Calculus