

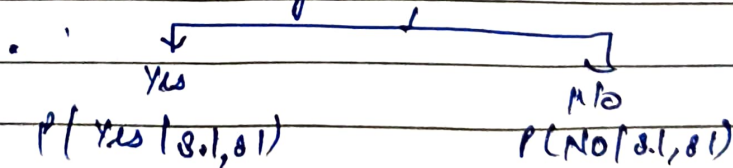
Date - 20-May

* concept of underflow

- representing decimal in computer using binary was always a challenge.
- The problem is related to above
- Underflow is condition when number is too small ^{max zero} $\rightarrow 0.000001$, so it becomes problem to store and apply operation as the number may lose precision
- It becomes problematic ex in case of biology (studying cells)
- How underflow is related to Naive Bayes

Ex • cgpa iq placement

• new query {8.1, 81} \rightarrow Y / No ?



$$P(Y|8.1, 8.1) = P(Y) (P(8.1|Yes)) (P(8.1|Yes))$$

This prob is related to \rightarrow cgra

This prob is related to placement

- Hence, if there were 2500 columns, there will be 2500 probabilities. and each probability lies between $0 < x < 1$. and if they were to multiply the number will be way very small, and consider it as zero (underflow)
- Same will happen with $P(\text{No}|\text{given})$, both probs will be zero due to underflow

Solution for underflow \Rightarrow log probabilities

- $\log(a, b) = \log a + \log b$

Hence $\cdot \log(p(a) p(b) p(c) \dots) = \log(p(a)) + \log(p(b)) + \dots$

- So even if we were to have 2500 columns we can apply and convert into log and after adding we get log of probability
- Whosever log probability is more, that class will be assigned.