

13 / 04 / 2018

D S T Q Q S S

1) É um modelo probabilístico que define que a probabilidade da próxima palavra w_m em uma frase é $P(w)$ ou $P(w_m | w_1, \dots, w_{m-1})$.

2) A regra da cadeia em probabilidade permite o cálculo de qualquer membro da distribuição conjunta de um conjunto de variáveis aleatórias utilizando probabilidade condicional

3) É uma medida estatística para calcular a probabilidade de dois eventos ocorrem juntos e ao mesmo tempo. Pode-se calcular a probabilidade conjunta de um elemento em função da regra da cadeia, esta pode ser calculada usando probabilidade condicional também chamada de regra geral do produto

$$P(w_1, \dots, w_m) = P(w_1) \cdot P(w_2 | w_1) \cdot P(w_3 | w_1, w_2) \cdot P(w_m | w_1, \dots, w_{m-1})$$
$$P(w_1 w_2, \dots, w_m) = \prod_i P(w_i | w_1, w_2, \dots, w_{i-1})$$

$$\text{Ex: } P(\text{"Sou uma programadora Python"}) =$$
$$P(\text{Sou}) \cdot P(\text{Uma} | \text{Sou}) \cdot P(\text{programadora} | \text{Sou uma})$$
$$\cdot P(\text{Python} | \text{Sou uma programadora})$$

4) Propriedade Markoviana diz que a distribuição de probabilidade do próximo estado depende apenas do estado atual e não da sequência de eventos que precedem.

$$\text{Ex: } P(\text{Python} | \text{Sou uma programadora}) = P(\text{Python} | \text{programadora})$$

5) O modelo N-gram é um modelo estatístico que permite fazer uma predição do próximo elemento de uma sequência de elementos que ocorreu anteriormente.

- desvantagem n grande: maior n° classes, crescimento exponencial.
- vantagem n grande: maior confiabilidade da inferência
- desvantagem n pequeno: menor n° de classes