

Bayesin teoreema

- Kehittäjä Thomas Bayes 1700 luvulla
- Ehdollinen todennäköisyys teoreema
- Kuvaa tapahtuman A todennäköisyyttä ehdolla B ja tapahtuman B todennäköisyyttä ehdolla A välistä suhdetta

Naive Bayes algoritmi

- Käytetään esim. spam -filtereinä
- Sentiment analyysi
- Yleensä ensimmäinen valinta tekstien analysoinnissa
- Jos on todennäköisempää, että viesti A on spämmi kuin oikea viesti ->viesti on spämmi

- Training-datasta luodaan bags of words
- Lasketaan kuinka usein tietty sana esiintyy spam viestissä(word frequency)
- Nopea ja kohtuu yksinkertainen matemaattinen algoritmi joka toimii hyvin varsinkin pienemmillä training datalla
- Olettaa että datan featuret ovat riippumattomia keskenään

Sananen matikkaa