Bayesian Linear Regression

Nipun Batra

July 2, 2019

IIT Gandhinagar

MLE, MAP, Bayesian

Bayes Rule - 1

- $P(A|B)P(B) = P(B|A)P(A)$

Bayes Rule - 1

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider an example from Wikipedia:

Bayes Rule - 1

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider an example from Wikipedia:
    - A particular drug is 99% sensitive and 99% specific

Bayes Rule - 1

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider an example from Wikipedia:
    - A particular drug is 99% sensitive and 99% specific
    - i.e. test will produce 99% true positive results for drug users and 99% true negative results for non-drug users

Bayes Rule - 1

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider an example from Wikipedia:
  - A particular drug is 99% sensitive and 99% specific
  - i.e. test will produce 99% true positive results for drug users and 99% true negative results for non-drug users
  - 0.5% of people are users of the drug

Bayes Rule - 1

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider an example from Wikipedia:
  - A particular drug is 99% sensitive and 99% specific
  - i.e. test will produce 99% true positive results for drug users and 99% true negative results for non-drug users
  - 0.5% of people are users of the drug
  - Question: What is the probability that a randomly selected individual with a positive test is a drug user?

Bayes Rule - 2

- Test will produce 99% true positive results for drug users and 99% true negative results for non-drug users $\implies$

Bayes Rule - 2

- Test will produce 99% true positive results for drug users and 99% true negative results for non-drug users $\implies$
    - $P(Test = +|User = Drug) = 0.99$, or, $P(+|User) = 0.99$

Bayes Rule - 2

- Test will produce 99% true positive results for drug users and 99% true negative results for non-drug users $\implies$
  - $P(Test = +|User = Drug) = 0.99$, or, $P(+|User) = 0.99$
  - and $P(-|\overline{User}) = 0.99$

Bayes Rule - 2

- Test will produce 99% true positive results for drug users and 99% true negative results for non-drug users $\implies$
  - $P(Test = +|User = Drug) = 0.99$, or, $P(+|User) = 0.99$
  - and $P(-|\overline{User}) = 0.99$
- 0.5% of people are users of the drug $\implies$ $P(User) = 0.005$

Bayes Rule - 2

- Test will produce 99% true positive results for drug users and 99% true negative results for non-drug users $\implies$
  - $P(\mathit{Test} = + | \mathit{User} = \mathit{Drug}) = 0.99$, or, $P(+|\mathit{User}) = 0.99$
  - and $P(-|\overline{\mathit{User}}) = 0.99$
- 0.5% of people are users of the drug $\implies$ $P(\mathit{User}) = 0.005$
- Question: What is the probability that a randomly selected individual with a positive test is a drug user?
  $\implies$ $P(\mathit{User}|+) =?$

Bayes Rule - 2

- Test will produce 99% true positive results for drug users and 99% true negative results for non-drug users $\implies$
  - $P(Test = +|User = Drug) = 0.99$, or, $P(+|User) = 0.99$
  - and $P(-|\overline{User}) = 0.99$
- 0.5% of people are users of the drug $\implies$ $P(User) = 0.005$
- Question: What is the probability that a randomly selected individual with a positive test is a drug user? $\implies$ $P(User|+) =?$
- $P(User|+) = \frac{P(+|User)P(User)}{P(+)} =$

Bayes Rule - 2

- Test will produce 99% true positive results for drug users and 99% true negative results for non-drug users $\implies$
  - $P(Test = +|User = Drug) = 0.99$, or, $P(+|User) = 0.99$
  - and $P(-|\overline{User}) = 0.99$
- 0.5% of people are users of the drug $\implies$ $P(User) = 0.005$
- Question: What is the probability that a randomly selected individual with a positive test is a drug user?
  $\implies$ $P(User|+) =?$
- $P(User|+) = \frac{P(+|User)P(User)}{P(+)} =$
- $\frac{P(+|\text{User})P(\text{User})}{P(+|\text{User})P(\text{User})+P(+|\overline{User})P(\overline{User})} = \frac{0.99\times 0.005}{0.99\times 0.005+0.01\times 0.995} \approx .332$

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
  - A = Parameters (θ)

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
  - A = Parameters (θ)
  - B = Data ($\mathcal{D}$)

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
  - A = Parameters (θ)
  - B = Data ($\mathcal{D}$)
- We can rewrite the Bayes rule as:

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
  - A = Parameters (θ)
  - B = Data ($\mathcal{D}$)
- We can rewrite the Bayes rule as:
  - $P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}$

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
  - A = Parameters (θ)
  - B = Data ($\mathcal{D}$)
- We can rewrite the Bayes rule as:
  - $P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}$
  - Posterior:

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
  - A = Parameters (θ)
  - B = Data ($\mathcal{D}$)
- We can rewrite the Bayes rule as:
  - $P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}$
  - Posterior:
  - Prior:

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
    - A = Parameters (θ)
    - B = Data ($\mathcal{D}$)
- We can rewrite the Bayes rule as:
    - $P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}$
    - Posterior:
    - Prior:
    - Likelihood

Bayes Rule for Machine Learning

- $P(A|B)P(B) = P(B|A)P(A)$
- Let us consider for a machine learning problem:
  - A = Parameters (θ)
  - B = Data ($\mathcal{D}$)
- We can rewrite the Bayes rule as:
  - $P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})}$
  - Posterior:
  - Prior:
  - Likelihood
  -