Text Classification

Xiang Pan

09/11/2023

Outline

- MLE and Naive Bayes
- Regularization In ML
- Regularization in NLP

MLE: Basic Concept

ERM

$$\min \sum_{i=1}^{N} \ell(x^{(i)}, y^{(i)}, \theta)$$

MLE

$$\max \sum_{i=1}^{N} \log p(y^{(i)} \mid x^{(i)}; \theta)$$

NB: Sentence Example

Text	Reviews
"I liked the movie"	positive
"It's a good movie. Nice story"	positive
"Nice songs. But sadly boring ending. "	negative
"Hero's acting is bad but heroine looks good. Overall nice movie"	positive
"Sad, boring movie"	negative

NYU

ref:https://www.geeksforgeeks.org/applying-multinomial-naive-bayes-to-nlp-problems/

NB: Sentence Example

1. Get the words list
2. Calculate the Conditional Probability P(word| class)
3. Inference: Calculate the probability P(Class | Sentence)
   a. P(Class | Sentence) = P(Sentence | Class) * P(Class) / P(Sentence)
   b. P(Sentence | Class) = P(w_0|Class) * P(w_1|Class)
   c. P(Positive|Sentence) > P(Negative|Sentence)?

NYU

NB: Sentence Example

Text	Reviews
"I liked the movie"	positive
"It's a good movie. Nice story"	positive
"Nice songs. But sadly boring ending. "	negative
"Hero's acting is bad but heroine looks good. Overall nice movie"	positive
"Sad, boring movie"	negative

1. Calculate the positive words count and negative words count
2. Calculate the conditional probability
   a. P(word_i | class_j)
3. Do the inference.

NYU

NB: Sentence Example

Text	Reviews
"I liked the movie"	positive
"It's a good movie. Nice story"	positive
"Nice songs. But sadly boring ending. "	negative
"Hero's acting is bad but heroine looks good. Overall nice movie"	positive
"Sad, boring movie"	negative

"overall liked the movie"

P(overall | positive) = 1/17
P(liked/positive) = 1/17
P(the/positive) = 2/17
P(movie/positive) = 3/17

NB: Sentence Example

Text	Reviews
"I liked the movie"	positive
"It's a good movie. Nice story"	positive
"Nice songs. But sadly boring ending. "	negative
"Hero's acting is bad but heroine looks good. Overall nice movie"	positive
"Sad, boring movie"	negative

P(liked/negative) = 0/7

How should we handle the zero probability here?

**P(liked/negative)=
(liked times in negative sentence + 1) / (|negative_word| + |total_word|)
(0 + 1) / (7+ 21)**

Should we count the words with same stem as one word ?

Liked and Like?

Should we consider words with same sub word?

p(sad|negative) = 1/7
p(sadly|negative) = 1/7

-> p(sad|negative) = 2/7 ?

NYU

8

NB: Sentence Example

P(positive|"Nice songs. But sadly boring ending. ")
= P("Nice songs. But sadly boring ending. "|positive) * P(positive) / P("Nice songs. But sadly boring ending. ")
= P("Nice"|positive) * P("songs"|positive) * P("But"|positive) * P("sadly"|positive) * P("boring"|positive) * P("ending"|positive) * P(positive) / P("Nice songs. But sadly boring ending. ")

MLE and NB: Sentence Example

Can we improve?
- Stem and Stopwords
- Laplace smoothing
- Change the NB assumption
    - N-gram
    - Contextualized Models

Regularization in ML

Ref: https://en.wikipedia.org/wiki/Regularization_%28mathematics%29#/media/File:Regularization.svg
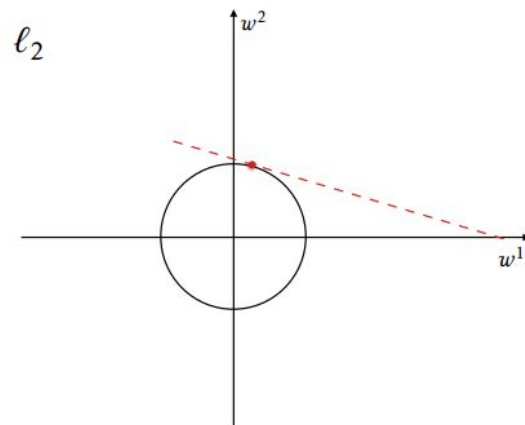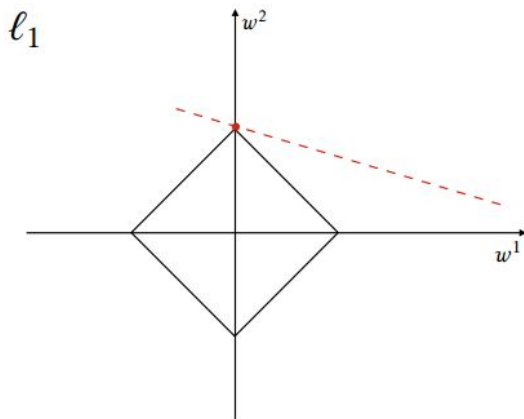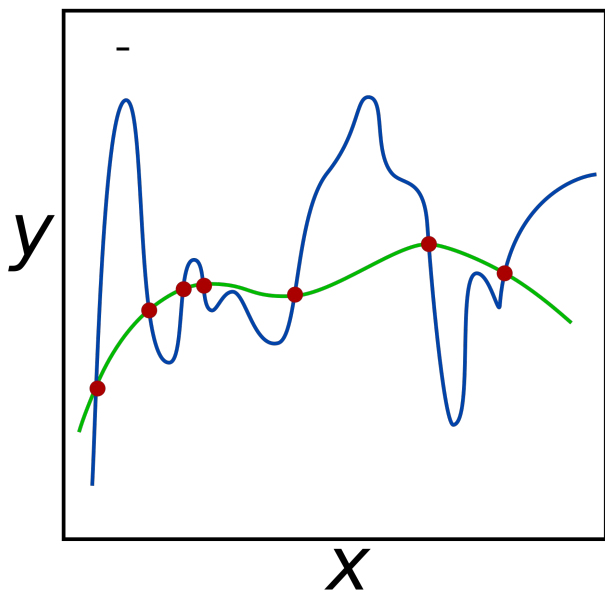
Regularization?

- Explicit regularization
    - Dropout
    - Data Augmentation
        - Back Translation
        - Style Change
        - Word Order shuffling
        - Synonyms change
        - …
- Implicit regularization
    - Early Stopping
    - Model Structure

NYU