Unit II

|  |  |
| --- | --- |
| **S.NO** | **Question** |
| 1 | Explain with an example on training and testing the naive Bayes with add-one smoothing |
| 2 | Illustrate Laplace smoothing with an example |
| 3 | Discuss Naive Bayes Classifiers? |
| 4 | We are given the following corpus  <S> I am Sam</S>  <S> Sam I am</S>  <S> I am Sam</S>  <S> I do not like green eggs and Sam</S>  Using a bigram language model with add-one smoothing, what is P(Sam | am)? Include <S> and </S>in your counts just like any other token. |
| 5 | Compute the probability of sentences like I want English food or I want Chinese food in the Berkeley Restaurant Project corpus |
| 6 | Perplexity is the metric for evaluating language models. Discuss. |
| 7 | Write a short notes on and Interpolation. |
| 8 | Assume the following likelihoods for each word being part of a positive or negative movie review, and equal prior probabilities for each class.    What class will Naive bayes assign to the sentence “I always like foreign films.”? |
| 9 | Given the following short movie reviews, each labeled with a genre, either comedy or action:  1. fun, couple, love, love comedy  2. fast, furious, shoot action  3. couple, fly, fast, fun, fun comedy  4. furious, shoot, shoot, fun action  5. fly, fast, shoot, love action and a new document D: fast, couple, shoot, fly  compute the most likely class for D.  Assume a naive Bayes classifier and use add-1 smoothing for the likelihoods. |
| 10 | Discuss in detail about how to Train the Naive Bayes Classifier |
| 11 | Explain with an example about binarization for the binary naive Bayes algorithm |
| 12 | Discuss the difference between Laplace smoothing and Add-k smoothing. |