Cairo University
Faculty of Engineering
Computer Engineering Dept.

# Sheet 3: Text Classification

1) Assume the following likelihoods for each word being part of a positive or negative movie review, and equal prior probabilities for each class.

| | pos | neg |
|---|---|---|
| I | 0.09 | 0.16 |
| always | 0.07 | 0.06 |
| like | 0.29 | 0.06 |
| foreign | 0.04 | 0.15 |
| films | 0.08 | 0.11 |

What class will Naive bayes assign to the sentence "I always like foreign films."?

2) 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.

3) Train two models, multinomial naive Bayes and binarized naive Bayes, both with add-1 smoothing, on the following document counts for key sentiment words, with positive or negative class assigned as noted.

| doc | "good" | "poor" | "great" | (class) |
|---|---|---|---|---|
| d1. | 3 | 0 | 3 | pos |
| d2. | 0 | 1 | 2 | pos |
| d3. | 1 | 3 | 0 | neg |
| d4. | 1 | 5 | 2 | neg |
| d5. | 0 | 2 | 0 | neg |

Use both naive Bayes models to assign a class (pos or neg) to this sentence:
A good, good plot and great characters, but poor acting.
Do the two models agree or disagree?

4) Consider a model that is trained to predict whether an email is Spam or Not Spam. After training the model, we apply it to a test set of 500 new emails (also labeled) and the model produces the following confusion matrix:

| | | True Class | |
|---|---|---|---|
| | | Spam | Not Spam |
| Predicted Class | Spam | 70 | 30 |
| | Not Spam | 70 | 330 |

a. Compute the precision of this model with respect to the Spam class.
b. Compute the recall of this model with respect to the Spam class.
c. Suppose we have two users (*A* and *B*) with the following preferences:
   - *A* hates seeing spam emails in her inbox! However, she doesn't mind periodically checking the "Junk" directory for genuine emails incorrectly marked as spam.
   - *B* doesn't even know where the "Junk" directory is. He would much prefer to see spam emails in his inbox than to miss genuine emails without knowing!

   Which user is more likely to be satisfied with this classifier? Why?

5) We trained a logistic regression model to use for a sentiment analysis task with two possible output classes: "positive" and "negative". The decision boundary is given by:

$$\text{decision}(x) = \begin{cases} 1 & \text{if } P(y=1|x) > 0.5 \\ 0 & \text{otherwise} \end{cases}$$

Show that this corresponds to a linear decision boundary in the input space.

6) Consider a trained binary logistic regression model that uses two features x1 and x2. The associated weights are learned such that w1=0.2, w2=0.4 and b=0.5.
   a. Write the output probability equations of this classifier.
   b. This classifier is used to determine if a given document is a question (class y=1) or a statement (class y=0). The features used are as follows:
      i. x1=1 if the document contains '?' and 0 otherwise.
      ii. x2=1 if the document contains the word "What" and 0 otherwise.

      Consider the following three test documents:

      doc1: The weather is great. →gold class: statement
      doc2: What does NLP stand for? →gold class: question
      doc3: What do you think →gold class: question

      Compute the class classification for each document.
   c. Compute the accuracy, precision, recall and F1-score of the obtained results.