Midterm Resources

October 20th

1 Naive Bayes

$c_{NB} = \text{argmax}_{c \in C} P(c_j) \prod_{x \in X} P(x|c)$ where C is the set of classes and X is the set of features.

Prior: $p(c_j)$

Likelihood: $p(x|c)$

Add-1 Smoothed Naive Bayes: $\hat{P}(w_i|c) = \frac{count(w_i,c)+1}{(\sum_{w \in V} count(w,c))+|V|}$ where V is the vocabulary.

2 N-gram Language Models

$P(w_1 w_2 ... w_3) \approx \prod_i P(w_i|w_{i-k}...w_{i-1})$ where k is the context window.

Bigram Maximum Likelihood Estimates: $P(w_i|w_{i-1}) = \frac{count(w_{i-1},w_i)}{count(w_{i-1})}$