

COMS W4771 : Machine Learning - Problem Set #2

Dhananjay Shrouthy - ds3521@columbia.edu

October 2, 2016

Problem 1

(a) $\mu_{y,j} = \frac{\sum_{i=1}^n 1\{y_i=y\}x_{i,j}}{\sum_{i=1}^n 1\{y_i=y\}}$

(b) Training Error Rate = 21.62% ; Test Error Rate = 37.6%

(c) Training Error Rate = 5.779% ; Test Error Rate = 13.138%

(d) **Vocabulary words with 20 largest α_j values (decreasing order)** - firearms, occupied, israelis, serdar, argic, appressian, ohanus, melkonian, sahak, villages, cramer, armenia, cpr, sdpa, handgun, optilink, palestine, firearm, budget, arabs
Vocabulary words with smallest α_j values (increasing order) - athos, atheism, atheists, clh, teachings, revelation, testament, livesey, atheist, solntze, wpd, scriptures, theology, believers, ksand, alink, benedikt, jesus, prophet, mozumder

Problem 2

- (a) For $1 \leq c \leq 14$, the region in which the classifier f^* predicts one is $\frac{16 - \sqrt{64 - 24\ln(c)}}{6}, \frac{16 + \sqrt{64 - 24\ln(c)}}{6}$
- (b) For $c \geq 15$, there will be no region in which the classifier f^* will predict 1 for smallest expected penalty as for $c \geq 15$, the term $\sqrt{64 - 24\ln(c)}$ will have imaginary roots and not real roots.

Problem 3

- (a) The eigenvalues of $\Sigma + \sigma^2 I$ are $\lambda_1 + \sigma^2, \lambda_2 + \sigma^2, \dots, \lambda_d + \sigma^2$
- (b) The eigenvalues of $(\Sigma + \sigma^2 I)^{-2}$ are $(\lambda_1 + \sigma^2)^{-2}, (\lambda_2 + \sigma^2)^{-2}, \dots, (\lambda_d + \sigma^2)^{-2}$