CSE512 Fall 2018 Machine Learning - Homework 5

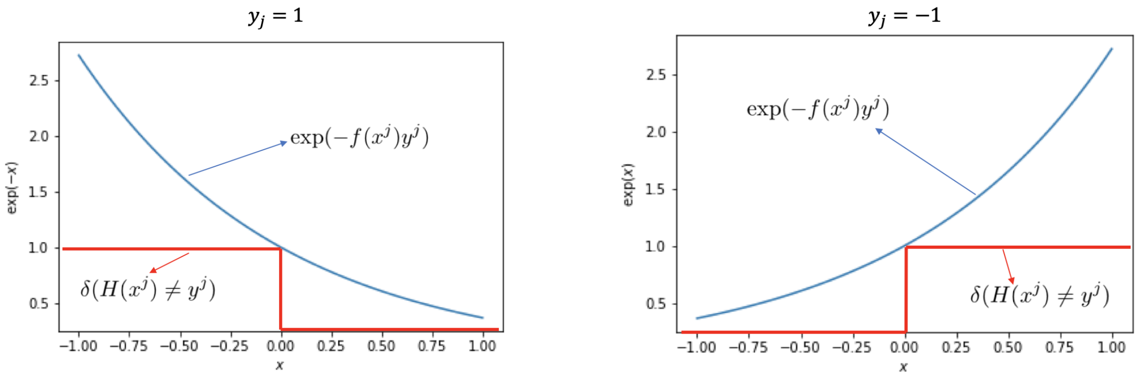
Your Name: Caitao Zhan

Solar ID: 111634527

NetID email address: caitao.zhan@stonybrook.edu

Names of people whom you discussed the homework with: Ting Jin

1 Boosting



(a) $y_j = 1$ (b) $y_j = -1$

Figure 1: 0/1 loss and exponential loss

1. Figure 1 illustrates that equation (1) will always hold either $y_j = 1$ or $y_j = -1$
$$\delta(H(x^j) \neq y^j) \leq \exp(-f(x^j)y^j)) \tag{1}$$
Therefore, the summation of each term will also hold
$$\frac{1}{N}\sum_{j=1}^{N}\delta(H(x^j) \neq y^j) \leq \frac{1}{N}\sum_{j=1}^{N}\exp(-f(x^j)y^j))$$

2. We use the weight update rule as follow,
$$\begin{aligned}
w_j^1 &= \frac{1}{N} \\
w_j^2 &= \frac{1}{N}\frac{\exp(-\alpha_1 y_j h_1(x_j))}{Z_1} \\
w_j^3 &= \frac{1}{N}\frac{\exp(-\alpha_1 y_j h_1(x_j))\exp(-\alpha_2 y_j h_2(x_j))}{Z_1 Z_2} \\
&\ldots \\
w_j^{t+1} &= \frac{1}{N}\frac{\exp(-y_j f(x_j))}{\prod_{t=1}^{N} Z_t}
\end{aligned}$$

The weights of the (t+1)th step sum up to 1, so we have:
$$\begin{aligned}
&\sum_{i=j}^{N} w_j^{t+1} = 1 \\
\Rightarrow &\frac{1}{N\prod_{t=1}^{N} Z_t}\sum_{j=1}^{N}\exp(-y_j f(x_j)) = 1 \\
\Rightarrow &\frac{1}{N}\sum_{j=1}^{N}\exp(-y_j f(x_j)) = \prod_{t=1}^{N} Z_t
\end{aligned}$$

3. (a)

$$\begin{aligned}
Z_t &= (1-\epsilon_t)\exp(-\alpha_t) + \epsilon_t \exp(\alpha_t) \\
\frac{\partial Z_t}{\partial \alpha_t} &= -(1-\epsilon)\exp(-\alpha) + \epsilon_t \exp(\alpha_t) = 0 \\
\Rightarrow \alpha_t &= \ln \sqrt{\frac{1-\epsilon_t}{\epsilon_t}} \\
\Rightarrow Z_t^{opt} &= (1-\epsilon_t) e^{-\ln \sqrt{\frac{1-\epsilon_t}{\epsilon_t}}} + \epsilon_t e^{\ln \sqrt{\frac{1-\epsilon_t}{\epsilon_t}}} \\
&= 2\sqrt{\epsilon_t(1-\epsilon_t)}
\end{aligned}$$

(b)

$$\begin{aligned}
Z_t &= 2\sqrt{\epsilon_t(1-\epsilon_t)} \\
&= 2\sqrt{(\frac{1}{2}-\gamma_t)(\frac{1}{2}+\gamma_t)}
\end{aligned}$$

We have the fact: $\ln(1-x) \leq -x$, when $0 \leq x < 1$. Let $x = 4\gamma_t^2$, where $0 \leq \gamma_t < \frac{1}{2}$

$$\begin{aligned}
\ln(1-4\gamma_t^2) &\leq -4\gamma_t^2 \\
\ln \sqrt{1-4\gamma_t^2} &\leq -2\gamma_t^2 \\
\ln 2\sqrt{\frac{1}{4}-\gamma_t^2} &\leq -2\gamma_t^2 \\
2\sqrt{\frac{1}{4}-\gamma_t^2} &\leq e^{-2\gamma_t^2} \\
\Rightarrow Z_t &\leq e^{-2\gamma_t^2}
\end{aligned}$$

(c)

$$\begin{aligned}
\epsilon_{training} &\leq \exp(-2\sum_{t=1}^{T}\gamma_t^2) \\
&\leq \exp(-2\sum_{t=1}^{T}\gamma^2) && \gamma_t \geq \gamma, \gamma > 0 \\
&\leq \exp(-2T\gamma^2)
\end{aligned}$$

When $T \to \infty$, $\exp(-2T\gamma^2) \to 0$
Therefore, adaboost will achieve zero training error after large enough steps.

2 Clustering with K-means

1. Report total with group sum of squares(twgsos), p1, p2, p3 with different k.

K	twgsos	p1	p2	p3
2	5.3648e+08	0.7982	0.5481	0.6731
4	4.6111e+08	0.6788	0.8683	0.7736
6	4.3135e+08	0.5518	0.9443	0.7481

Table 1: Question 2.5.1

2. Iteration = 8, when K = 6.

3. In question 2.5.3 and 2.5.4, my random seed is 0. I repeat 10 times for each K and do an average.

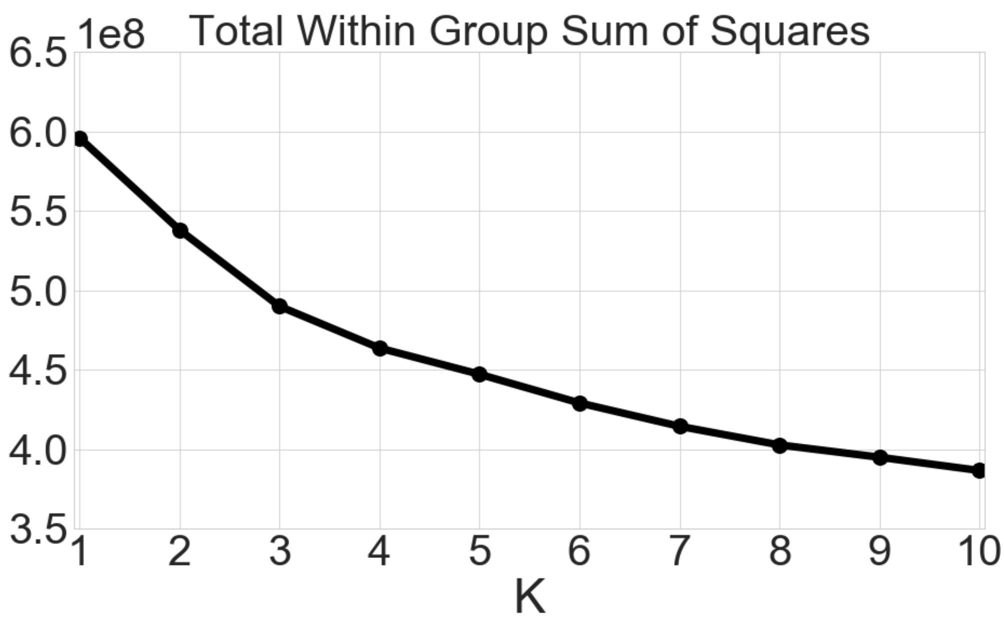


Figure 2: Total within group sum of squares as a function of K

4. When K=4, p3 reaches its highest point.

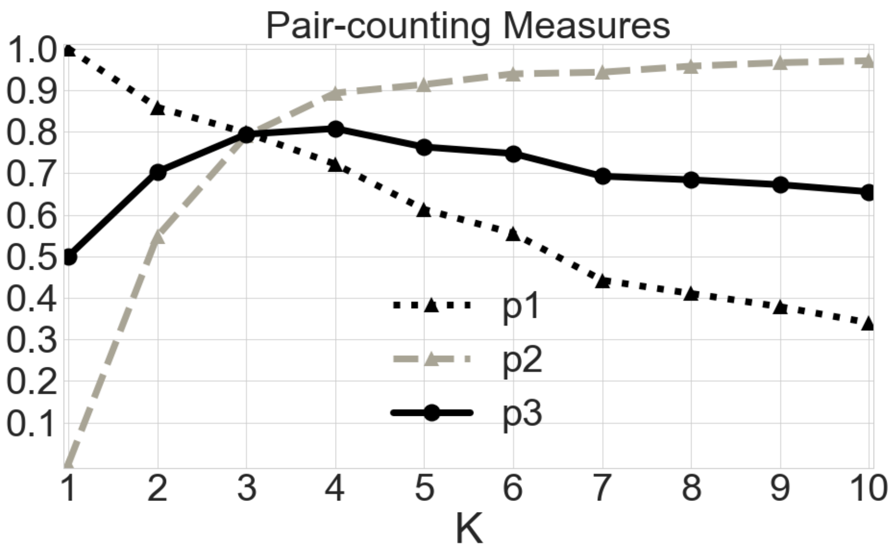


Figure 3: p1, p2, p3 as a function of K

3 Scene Classification

1. See source file HW5_BoW.m
2. **15.6443%**.
   For default kernel parameters and 5-fold cross validation, just use '-v 5' as the libsvm options: svmtrain(trLbs, trD, '-v 5'); The accuracy is very very bad.
3. I did a simple grid search. See the Table 2

Accuracy (%)	C=0.1	1	10	20	40	80	160
γ=0.1	15.6443	15.6443	26.3928	33.1458	46.9893	53.9111	64.2093
1	15.6443	25.9989	59.3134	65.6162	69.668	73.6635	77.7153
10	24.3669	58.5256	76.0833	80.8666	85.0872	86.888	88.4074
20	32.9769	65.8413	80.8104	84.9184	86.7192	88.5763	88.9702
40	45.9764	70.0619	84.8059	86.3815	88.3512	88.2386	88.6325
80	52.6168	74.789	86.888	87.6759	87.7884	87.7321	87.7321
160	59.0884	78.6156	87.1131	87.6759	87.6759	87.6759	87.6759

Table 2: 5 fold CV accuracy when tuning C and γ using grid search

4. See source file cmpExpX2Kernel2.m
5. I did a simple grid search and my best accuracy is **94.4288%**, when $\gamma = 1.4$, $C = 20$.
6. Kaggle submit.

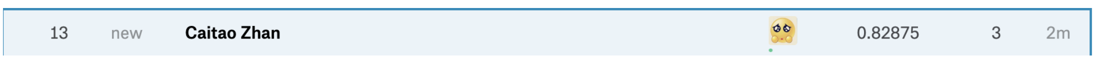


Figure 4: Submit