Homework 2

Grading Policy:

In handwriting section, you need to provide detailed answers or derivations. Partial points will be given for correct reasoning. Please write your answers and derivations on A4 papers and hand in the homework in class on 2021/11/17 at 09:00 am.

In programming section of homework consists of “implementation” and “report.” In the implementation, you have to complete the blocks enclosed by "PLEASE WRITE HERE" comments. If your submitted python file cannot be successfully compiled due to the modification outside of the blocks mentioned above, some points will be deducted. In the report, answers are expected to be stated with your observations followed by explanations. Writing in English or Chinese are acceptable. Please upload your python file and report named hw1_StudentID.py and hw1_StudentID.pdf respectively to the eeclass before 23:30 pm on 2021/11/17. No late submission will be accepted.

Part 1: Handwriting

1. In Principal Components Analysis, given a y and it’s a low-dimensional projection x with the following equation: $\mathrm{y} = \mathrm{w}^{\mathrm{T}} x$, please answer the following questions:
   (a) [5pt] Please explain the connection between w and y (Use math).
   (b) [10pt] Assume the covariance matrix of y is $\begin{matrix} 16 & 0 & 2 \\ 0 & 25 & 0 \\ 2 & 0 & 4 \end{matrix}$ and we set $\mathrm{k} = 1$ in PCA, please determine the matric w.
2. [10pt] Define a multivariate Bernoulli mixture where inputs are binary and derive the EM equation.
3. [5pt] Generalize the Gini index and the misclassification error for K > 2 classes. Generalize misclassification error to risk, taking a loss function into account.
4. [10pt] Show that the derivative of the softmax, $\mathrm{y_i} = \frac{\exp(\mathrm{a_i})}{\sum_j \exp(a_j)}$, is $\frac{\partial y_i}{\partial a_j} = y_i(\delta_{ij} - y_j)$ where δ_{ij} is 1 if $\mathrm{i} = \mathrm{j}$ and 0 otherwise.

Part 2: Programming

In this part you need to train the Support Vector Machine (SVM) and Decision Tree classifier to train the 3 class of Iris plant from Iris dataset (the iris dataset is provided with HW materials).

This Iris dataset contains 3 classes of 50 instances each, where each class refers to a type of iris plant. There are 4 attributes: 1: sepal length, 2: sepal width, 3: petal length, 4: petal width (all in cm) and 5: class (Iris Setosa, Iris Versicolour, and Iris Virginica).

Implementation section:

- [5pt] Correctly implementing Decision tree classifier with tree diagram plot.
- [5pt] Correctly implementing SVM with decision boundary plot.

Report section:

1. [10pt] Implement Decision tree classifier with tree diagram plot while considering **depth** parameter as 2, 3 and 4 respectively; report the difference in performance (in terms of accuracy) and which one works better and why? (You can analyze through confusion matrix or with tree diagram)
2. [8pt] Implement and Experiment the performance (accuracy) of SVM classifier while choosing **Gamma** value as (1, 100 and 500) with fix Regularization parameter (**C**) as 1.
3. [7pt] Plot the decision boundary for above conditions (Q 2) i.e., Gamma=1, Gamma=100 and Gamma=500 with C=1, also explain what you analyze from it? Which is better? Why? (You can analyze through confusion matrix or with changes in decision boundary)
4. [8pt] Implement and Experiment the performance (accuracy) of SVM classifier while choosing **C** value as (1, 100 and 500) with fix **Gamma** variable as 1.
5. [7pt] Plot the decision boundary for above condition (Q 4) i.e., C=1, C=100, and C=500 with Gamma=1, also explain which one is better? Why? (You can analyze through confusion matrix or with changes in decision boundary)
6. [10pt] Compare the performance of these two algorithms [**Decision tree and SVM**] and give some insight about the pros and cons of using one algorithm over another and why? (In reference with **Iris dataset**)