MACHINE LEARNING & PATTERN RECOGNITION (CT851-E23)

Assignment (074-MSCSK)

Course Instructor: Dr. Dibakar Raj Panta **Deadline:** *15th Chaitra, 2075 (**Sharp**)*

---

1. Describe the basic modules in designing a pattern recognition system. Why there is a need of pattern recognition? Give examples
2. How do we evaluate the performance of a classifier or a machine learning model?
3. Explain features, training set and test set. Why it is better to have a big database for classifications or modeling ML algorithm?
4. How do we evaluate the performance of a classifier? What is a cross validation technique?
5. Explain overfitting and underfitting. Describe confusion matrix.
6. Briefly explain what is generalization in the context of pattern recognition problems?
7. Define the term supervised and unsupervised classification with example.
8. State the *Bayes Rule* and explain how it is applied to pattern classification problems. Show that in a multiclass classification task the Bayes decision rule minimizes the error probability.
9. What is meant by dimension in database? Explain principal component analysis (PCA) with figure.
10. Describe the basic steps that must be followed in order to develop a clustering task. To which category of clustering schemes does the *k*-means algorithm belong? What is its major advantage? Which are the factors that influence the computational duration of this algorithm?
11. Describe the perceptron learning algorithm and its properties.
12. Describe a Support Vector Machine. Define the optimization task solved in SVM learning.
13. Discuss Deep learning algorithms with example
14. Describe K- nearest neighbor and its importance.
15. **2.7** In a three-class two-dimensional problem the feature vectors in each class are normally distributed with covariance matrix

$$\Sigma = \begin{bmatrix} 1.2 & 0.4 \\ 0.4 & 1.8 \end{bmatrix}$$

The mean vectors for each class are $[0.1, 0.1]^T$, $[2.1, 1.9]^T$, $[-1.5, 2.0]^T$. Assuming that the classes are equiprobable, (a) classify the feature vector $[1.6, 1.5]^T$ according to the Bayes minimum error probability classifier; (b) draw the curves of equal Mahalanobis distance from $[2.1, 1.9]^T$.

16. What are the name and the formula of a widely used distance measure? Use the formula to compute the distance between the points $\boldsymbol{x} = (1, 2)^T$ and $\boldsymbol{y} = (2, 1)^T$.

Programming Task:

Topic: Linear Classifier or discriminant based on perception algorithm.

Basic concept:

The purpose of Discriminant Analysis is to classify objects into one of two or more groups based on a set of features that describe the objects. For example, we want to know whether a soap product is good or bad based on several measurements on the product such as weight, volume, smell, etc. The object here is soap. The class category or the group ("good" and "bad") is what we are looking for (it is also called dependent variable). Each measurement on the product is called features that describe the object (it is also called independent variable). If we can assume that the groups are linearly separable, we can use linear discriminant model to classify object based on their features.

Perception Function:

Let us consider two class cases. Let us consider linear discriminant function given below:

$$g(x) = w^T x + w_0$$

Where $w = [w1, w2 \ldots wl]^T$ is weight vector and w_0 is the threshold. If x_1 and x_2 are two points on the decision hyperplane then it follows:

$w^T (x_1 - x_2) = 0$. Now, let us assume that the two classes ω_1 and ω_2 are linearly separable then we can define x in two classes as

$w^T x > 0$ *then x belongs to* ω_1 and $w^T < 0$ *then x belongs to* ω_2

To optimize the task, we need cost function and the algorithm to minimize the cost function.

Now, we define cost function as : $J(w) = \sum \delta_x w^T x$

Let us suppose that Y is subset of training vectors which are misclassified by the weight vector w. To indicate this misclassification we choose the variable δ_x equals -1 if $x \in \omega_1$ and +1 if $x \in \omega_2$

When the cost function takes minimum value 0, all training features are correctly classified.

To minimize the cost function, we have to adopt iterative scheme defined as follows:

$$w(t+1) = w(t) - \rho_t \sum \delta_x x$$

Where ρ_t is positive real number. This is known as perception algorithm.

In summary, the algorithm is initialized from an arbitrary weight vector say w_0 and the correction vector $\sum \delta_x x$ is formed using the misclassified features. The weight vector is repeated until the algorithm converges to a solution that is all features are correctly classified.

Problem: A medical company tested two medicines say A and B to know the effectiveness of medicines to cure certain type of influenza. Medicine A is given to 200 populations for observation and classified as class 1 category. Medicine B is given to other 200 population and classified as class 2 category. The features selected to classify the medicines are side effect of medicines and progress on cure. These features are scaled in positive and negative values. For example positive value of side

effect means common side effect and negative value of side effect means uncommon side effect. Similarly positive value of progress on cure means recovery is in progress. Negative value indicates no progress on recovery.

Question 1:

Load medical data. The medical data consists three columns, the first column is side effect of medicine in terms of value. The second column is recovery data in terms of value. The last column is class type.There are 400 observations and each row is one observation or sample.

Plot the features for the medicine A and B and label the axes.

Question 2:

Implement in Matlab to Classify the medical data by using perceptron algorithm as described above.

Hints:

Follow the following steps:

1. Find the data size (training data). Matlab command :size
2. Create initial weight vector as per the general equation described above.Matlab command:e.g weight=ones(sizedata+1,1) , here we have to add constant .
3. Initialize test class data . It should be the same size as the data. Matlab command:Zeros
4. Create initial error vector. Matlab command : ones(1,number of iteration).
5. Make training data including class labels. (Initialize class level creating by using ones command) Matlab command:
6. Evaluate each sample in the training data(value=weight'*traindata)
7. Assign samples to classes(testclass(values>=0)=1 else 2)
8. Find incorrectly classified data(delta=trainclass-testclass)
9. Pick up indices , where error occurs(error=find(delta not equal to zero))
10. If no error stop
11. Compute correction term
12. Set learning rate.For example Rho=.005
13. Update weight vector.example weight=weight-rho*correction
14. Plot data according to sample class
15. Plot decision line according to equation as mentioned above.