Experiment No.: 05

Title: Write a program to implement Multiclass Classification.

.

Objectives: Multiclass Classification

Theory:

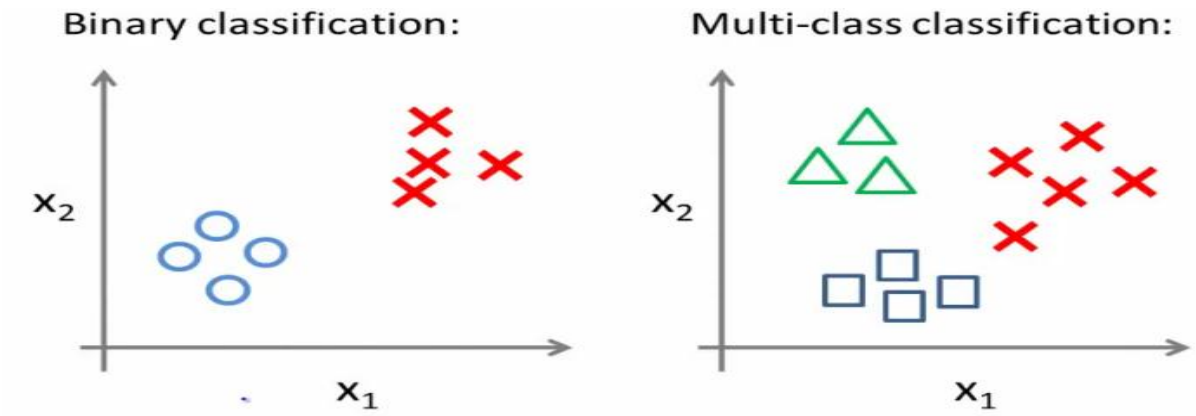
Multiclass Classification

In machine learning, **multiclass** or **multinomial classification** is the problem of classifying instances into one of three or more classes.

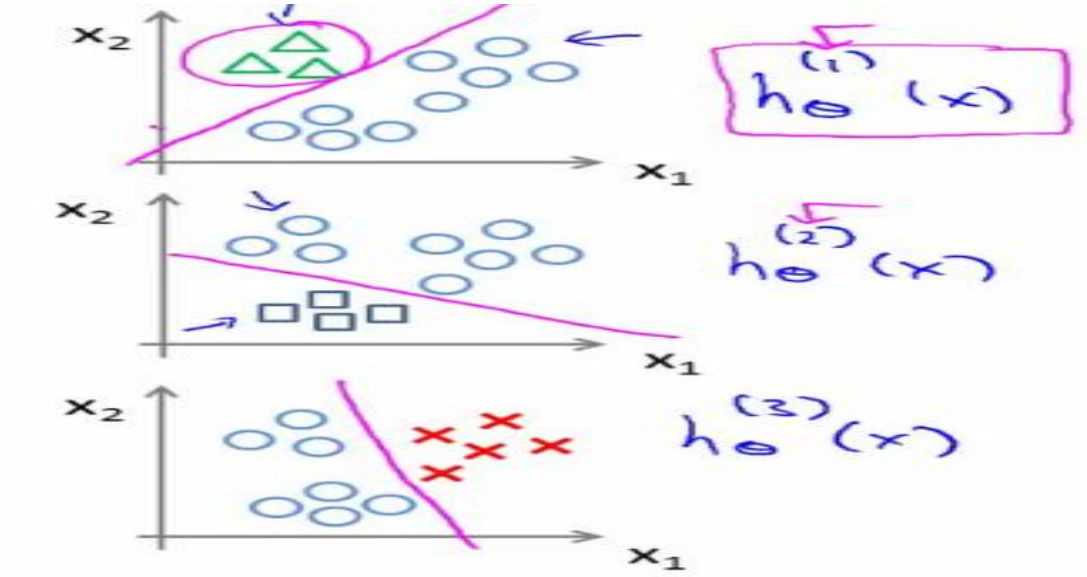
Given a dataset of m training examples, each of which contains information in the form of various features and a label. Each label corresponds to a class, to which the training example belongs to. In multiclass classification, we have a finite set of classes. Each training example also has n features.

For example, in the case of identification of different types of fruits, "Shape", "Color", "Radius" can be features and "Apple", "Orange", "Banana" can be different class labels.

Multiclass Classification: one vs all



- Given a dataset with three classes, how do we get a learning algorithm to work?
- Use one vs. all classification make binary classification work for multiclass classification
- Split the training set into three separate binary classification problems
  - **i.e. create a new fake training set**
    - **Triangle (1) vs crosses and squares (0) $h_\theta^1(x)$**
      - **$P(y=1 \mid x_1; \theta)$**
    - **Crosses (1) vs triangle and square (0) $h_\theta^2(x)$**
      - **$P(y=1 \mid x_2; \theta)$**
    - **Square (1) vs crosses and square (0) $h_\theta^3(x)$**
      - **$P(y=1 \mid x_3; \theta)$**



- **Overall**
- Train a logistic regression classifier $h_\theta^{(i)}(x)$ for each class i to predict the probability that y = i
- On a new input, *x* to make a prediction, pick the class *i* that maximizes the probability that $h_\theta^{(i)}(x) = 1$
- **For example**
- Triangle (1) vs crosses and squares (0)
  $h_\theta^1(x)=P(y=1 \mid x_1; \theta)=>>0.7$
- Crosses (1) vs triangle and square (0)
  $h_\theta^2(x)=P(y=1 \mid x_2; \theta)=>>0.4$
- Square (1) vs crosses and square (0)
  $h_\theta^3(x)=P(y=1 \mid x_3; \theta)=>>0.8$
- **New predicted class=Max($h_\theta^1(x)$, $h_\theta^2(x)$, $h_\theta^3(x)$)**

Algorithm-

1] Import all necessary libraries.
2] Read data set into pandas dataframe
3] Split dataset into training set and testing set using train_test_split function
4] Create logistic Regression object
4] Train model using fit function
5] Use build model for prediction