SRIP Project 2 Documentation MLE: Learning the Classifier from Data

Task allotted

1. In the virtual-labs repository, pattern-recognition-iiith lab, the task was to resolve Issue No: 243
2. Issue No: 243 was to Convert the following experiment to JavaScript. Link to the experiment:- http://cse20-iiith.vlabs.ac.in/exp6/Experiment.html?domain=Computer%20Science&lab=Pattern%20Recognition%20Lab

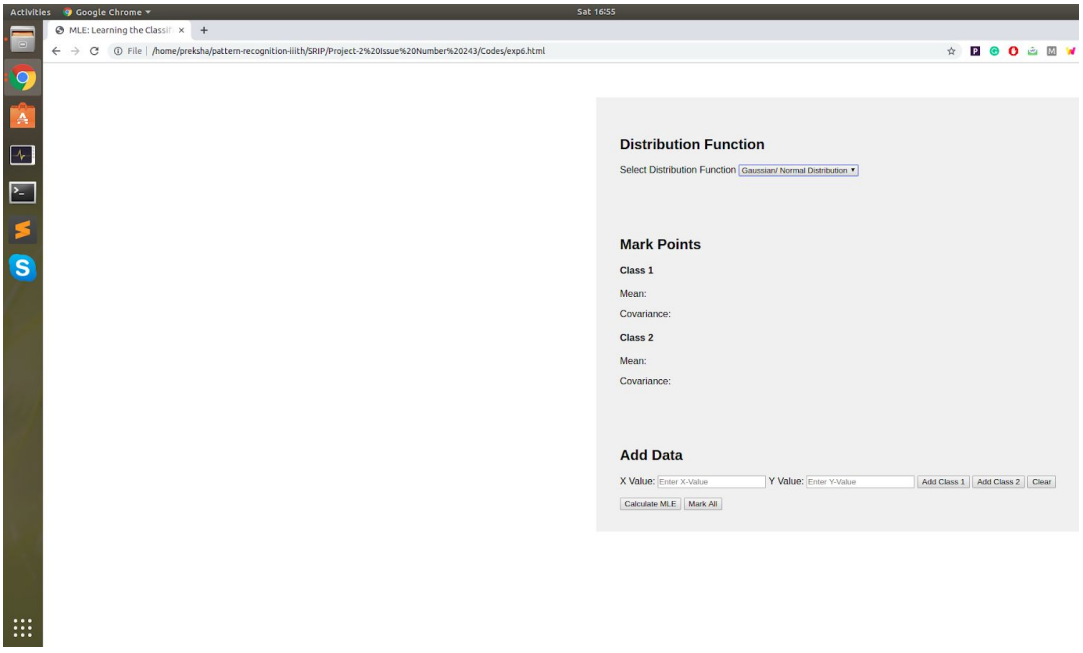
Experiment Explanation

In machine learning, different models are used to describe the results. Each model contains a set of parameters. These parameters define how to model looks like. Therefore good parameters would give a good model i.e. specific values of the parameters would give a good prediction for the model that would describe a given phenomenon.
Maximum likelihood estimation is a method that determines values for the parameters of a model. The Maximum Likelihood parameter Estimation determines the parameters that will maximise the probability of the sample data.
It is a method that will find the values of μ(mean) and σ(standard deviation) that result in the curve that best fits the data. The values that we find are called the maximum likelihood estimates (MLE).

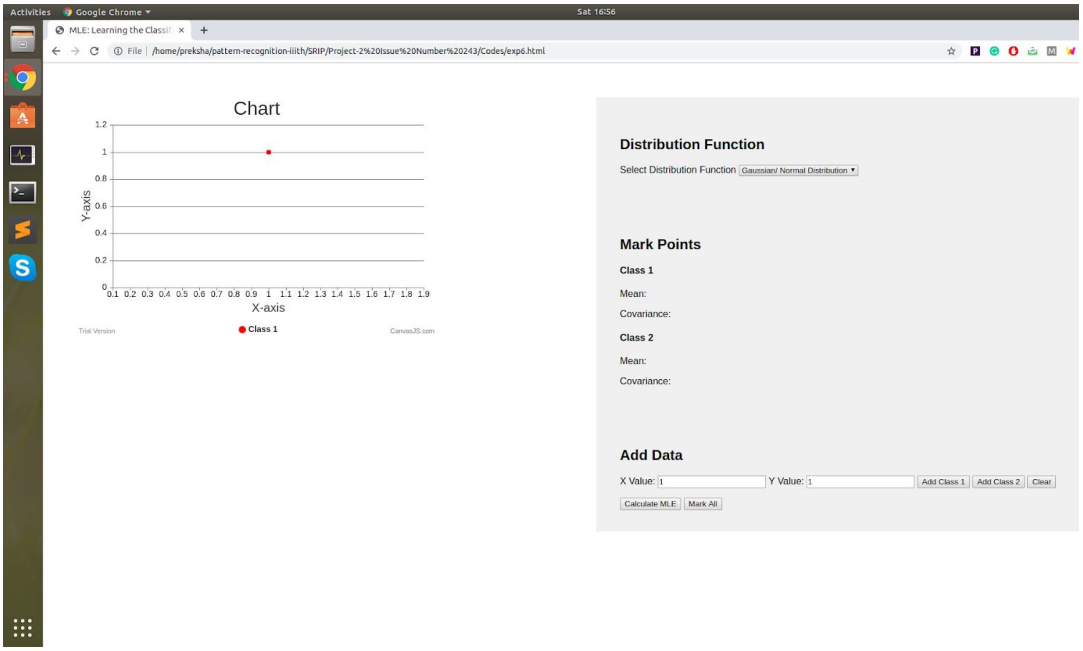
How to Run the Experiment

1. The forked repository(https://github.com/prekshap24/pattern-recognition-iiith) contains a folder named “SRIP”.

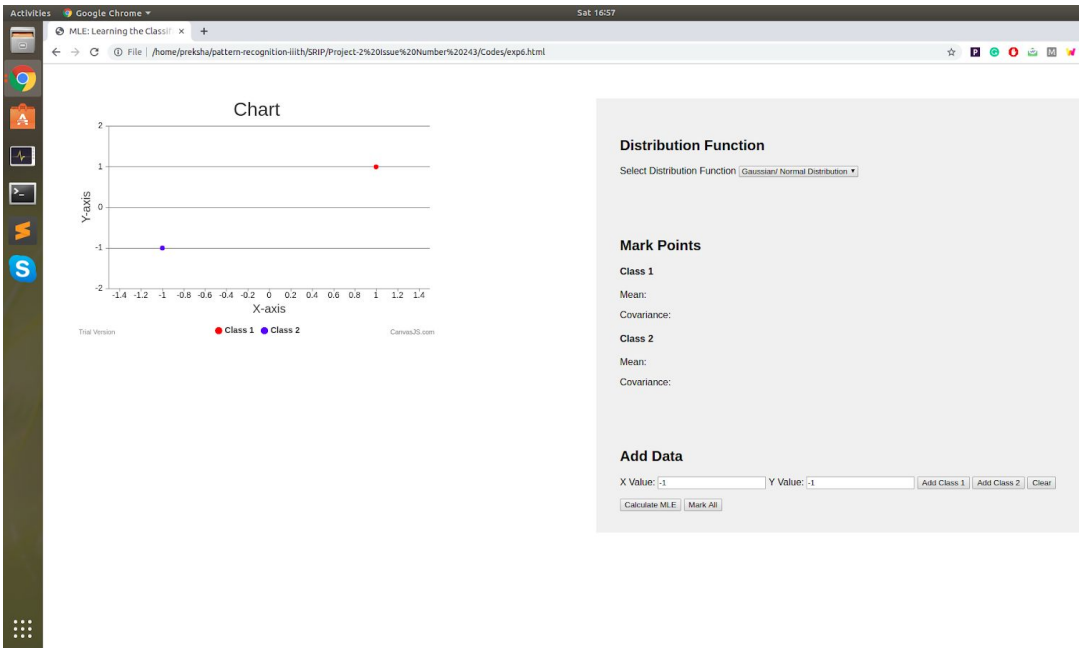
2. SRIP folder contains folders named as Codes and Libraries. Codes contain all the files containing code for the experiment written in JavaScript, HTML, CSS. Libraries contain JavaScript libraries used in the codes.
3. The Codes folder contains 3 files. To run the experiment, simply run the exp6.html file by clicking on it.
4. The experiment will open in the browser.



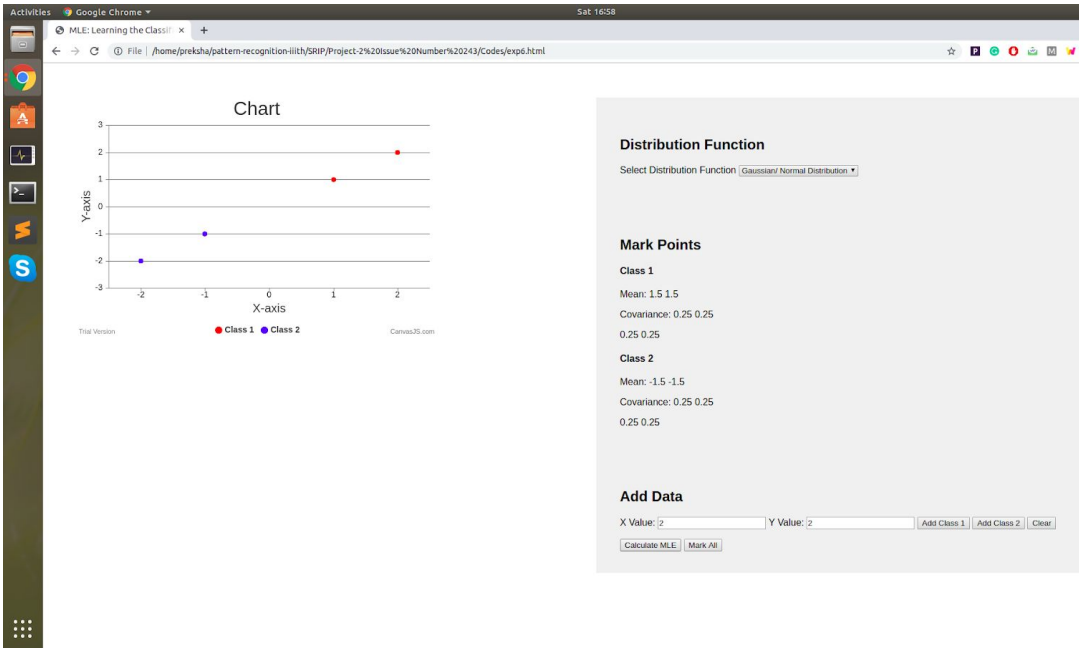
5. Select the distribution function using the drop-down menu at top right .
6. Add the X value and Y value for Class 1 and click on Class 1.



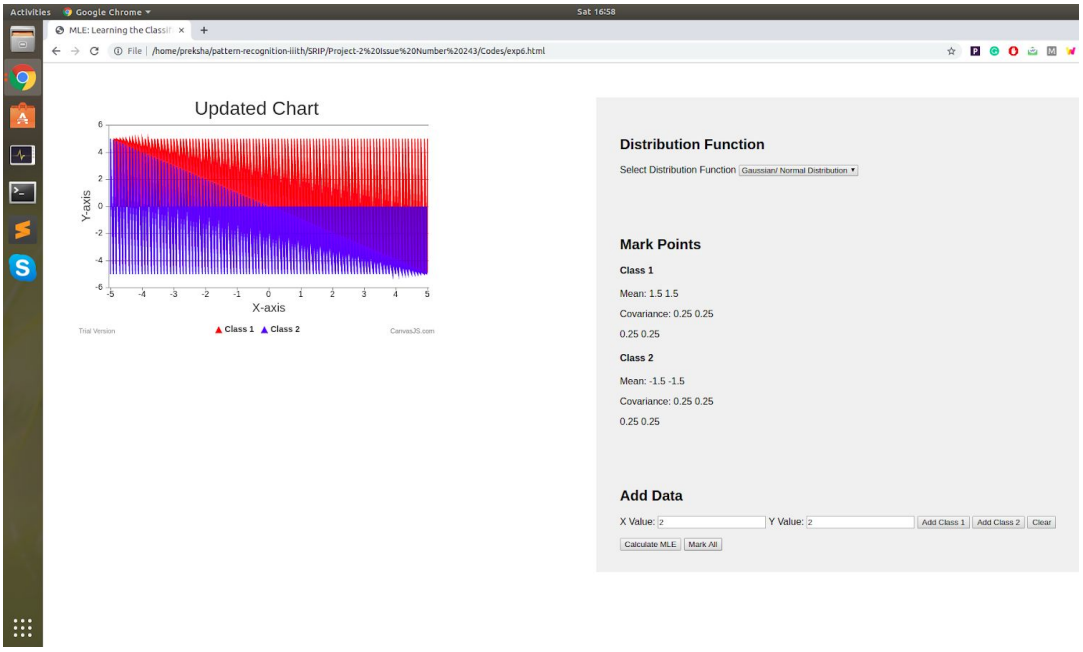
7. Similarly add X value and Y value for Class 2 and click on Class 2.



8. Estimate the maximum likelihood parameters



9. Click on Mark all Button to see the classifier



10. If you have selected Gaussian/ Normal Distribution curve, click on the ‘Calculate MLE’ and ‘Mark All’, repeatedly until the classifier learns completely and classifies completely.
11. If Uniform Distribution is selected, clicking on ‘Calculate MLE’ and ‘Mark All’ give the results in only one iteration.

Formulas Used

1. The formulas used for finding the mean of x data points and y data points for class 1 and class 2 are

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

2. The formula used for finding the covariance of the x data points with x data points, x data points with y data points, y data points with x data points and y data points with y data points, for class 1 and class 2 is

$$COV(x,y) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

3. Probability formula for Gaussian distribution is

$$P(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

4. The probability density function and cumulative distribution function for a *continuous* uniform distribution on the interval [a, b] are

$$P(x) = \begin{cases} 0 & \text{for } x < a \\ \dfrac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{for } x > b \end{cases}$$

$$D(x) = \begin{cases} 0 & \text{for } x < a \\ \dfrac{x-a}{b-a} & \text{for } a \le x \le b \\ 1 & \text{for } x > b. \end{cases}$$