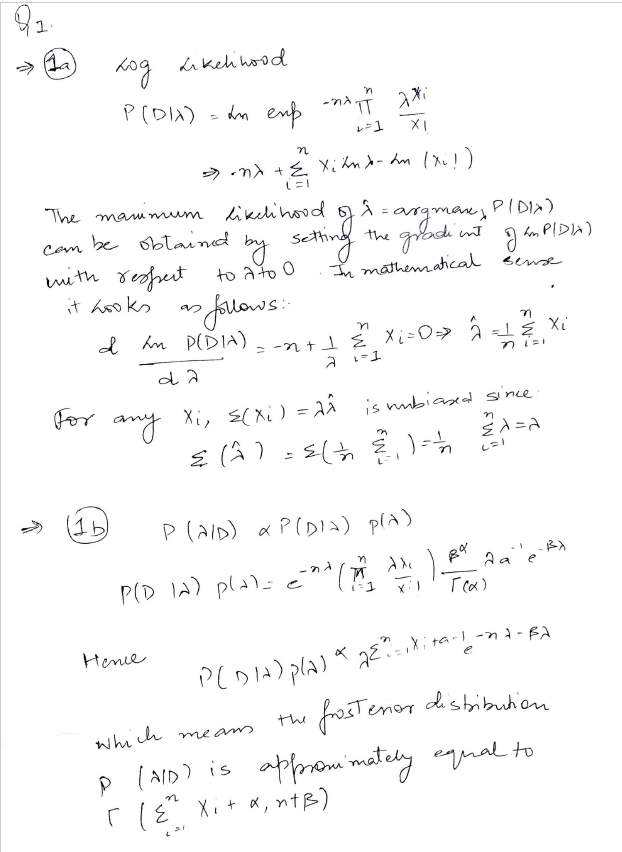
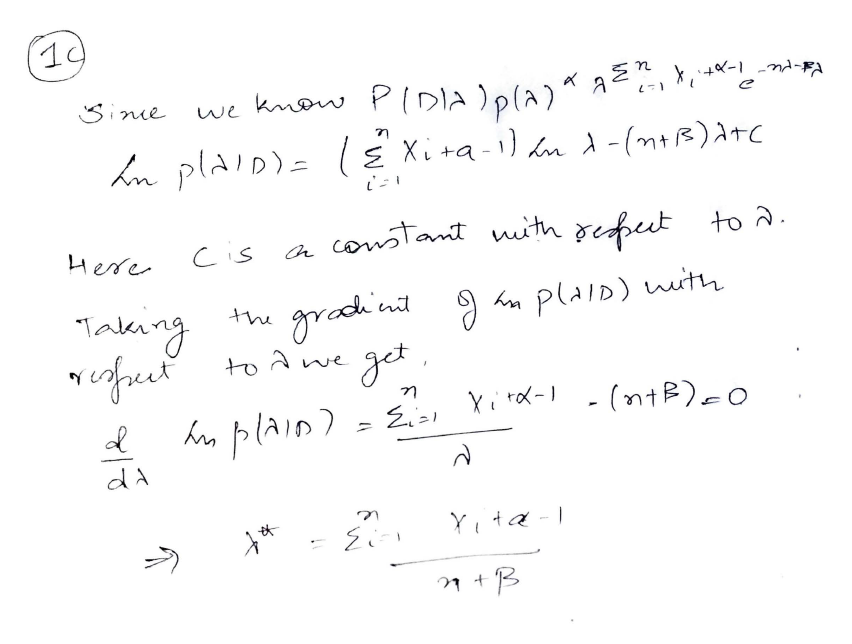
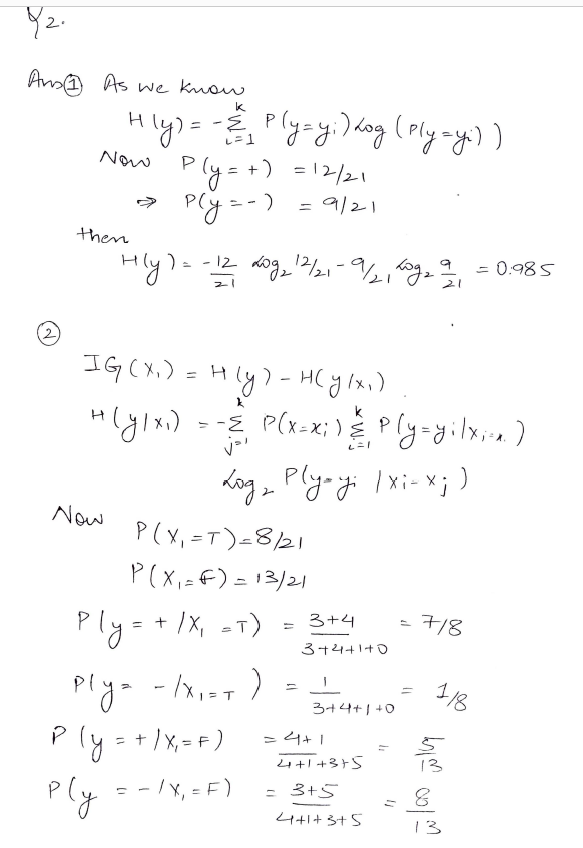
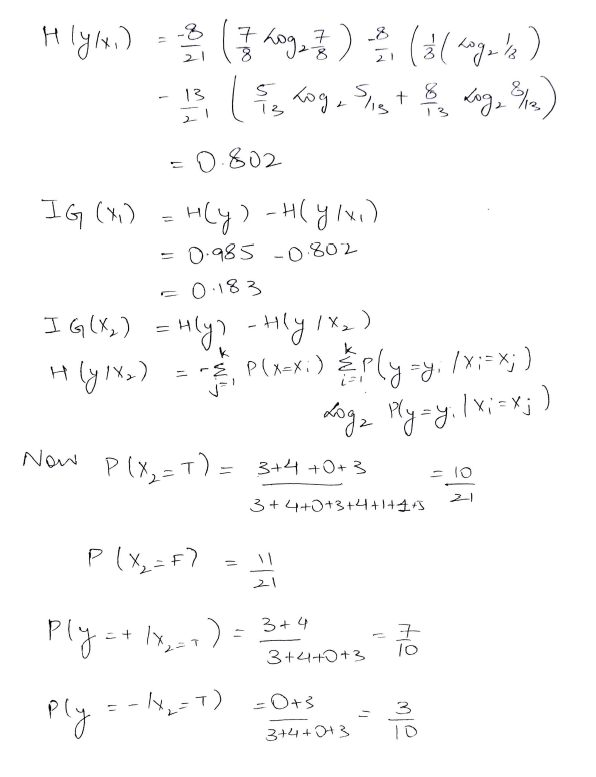
1. Extra credit answer included

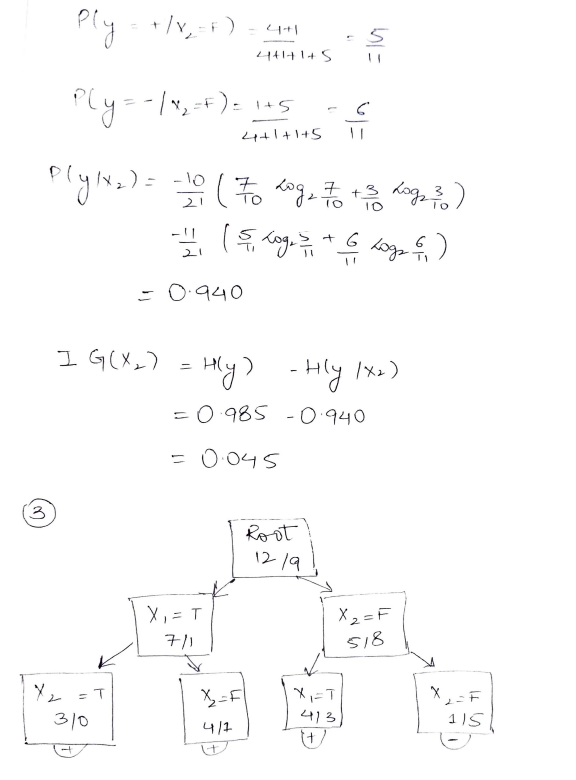




\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**2.**





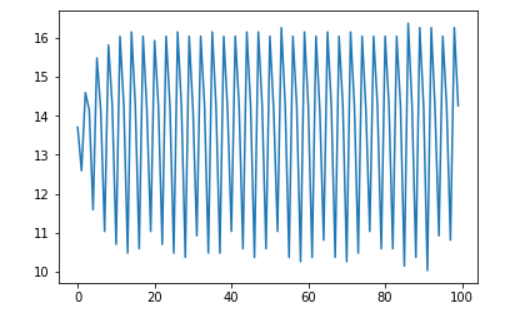
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

**Q3.**

* A

Compared to smaller learning rate, larger learning rate is generally less time consuming, requires lesser change in weights, and requires less loops to execute. But, the problem with the large learning rate model there might not be able to converge. Since, it would add big fractions in training weights.

* B



As seen through the loss function vs iteration plot. We can confirm that the loss is fluctuating between two constant values instead of converging to a certain boundary. Hence, it can be confirmed that the algorithm does not converge.  
Also it is evident that the data is not linearly separable otherwise the perceptron had given a convergent solution.

* C

Due to normalization the classifier value here is either 0 or 1. On the other hand for confidence we are using sigmoid function

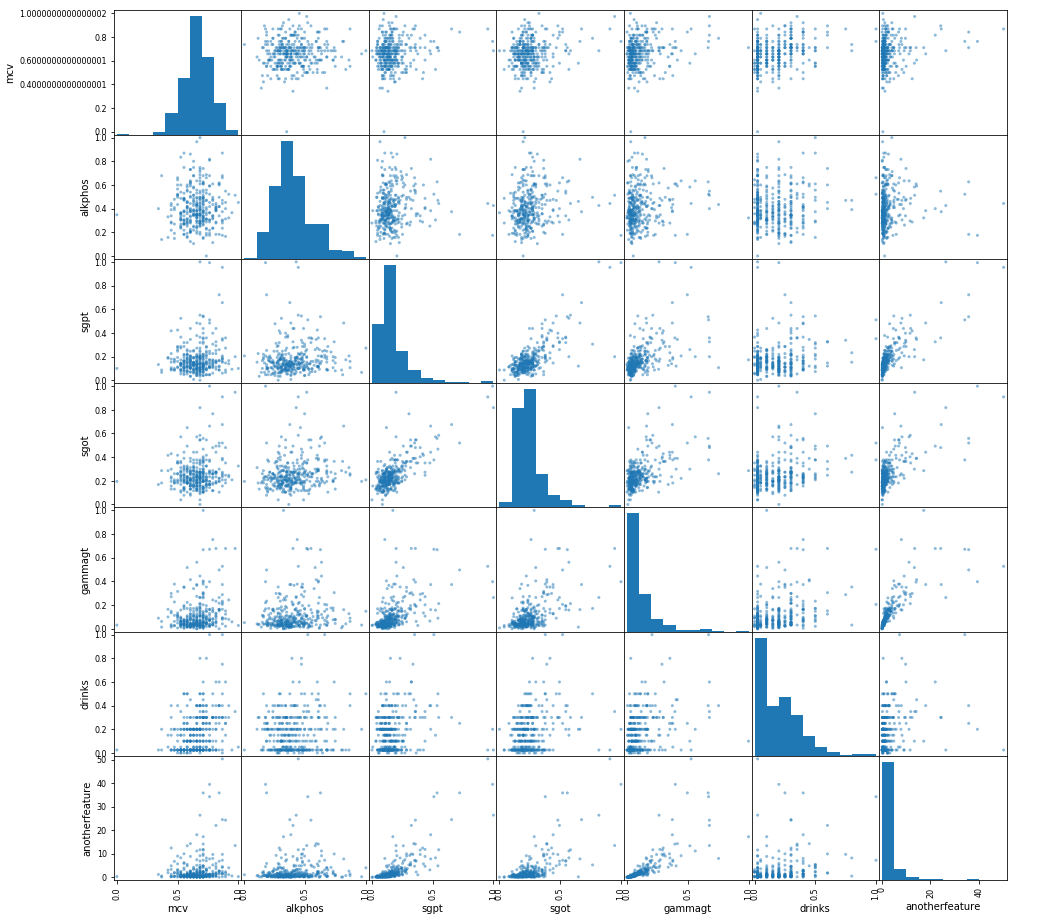
The table depicts the case as described above:

|  |  |  |  |
| --- | --- | --- | --- |
|  | actual | predictions | confidence |
| 4 | 1 | 1 | 1 |
| 140 | 0 | 1 | 0.999998635 |
| 249 | 1 | 1 | 1 |
| 250 | 1 | 1 | 0.999999528 |
| 167 | 0 | 1 | 1 |
| 305 | 1 | 1 | 0.921847935 |
| 1 | 1 | 0 | 3.16E-09 |
| 42 | 1 | 1 | 0.888777859 |
| 35 | 1 | 1 | 1 |
| 268 | 1 | 1 | 0.999994696 |

It is evident from the scatter plot that the Top right values are highly accurate.

* 4.

Scatter plot matrix is essential to understand the scenario:



X\_new = X

X\_new['synthetic\_feature']=X.sgpt \* X.gammagt \* 100

To increase the accuracy a synthetic feature has been used. And after running the algorithm there was a improvement i.e. 0.6011 instead of previous 0.5977.