DHRUV GARG 16 BCE 1190

CLASSIFICATION USING 2D PERCEPTRON

LAB 2

Perceptron theory:

Perceptron formula is given as (with bias):

$$\sum_{i=0}^{N} W_i x_i + b$$

The misclassified functions are corrected based on the **updation function:**

$$w \leftarrow w + \Delta w$$

There is a threshold w_0 which determines the **activation function.** If the summation of $w_i x_i$ is greater than the threshold, the perceptron is activated (+1), else it is inactive (-1).

PART 1: IRIS Dataset

After pre-processing the dataset, the first step is to **identify linearly separable attributes** to perform classification on. From the **pairplots** plotted after dataset cleaning, I identified that setosa and virginica were linearly separable with respect to sepal width and petal length. Hence, we implement the 2D perceptron on this. **Implementation 1** involved implementing our own code of the perceptron. It was a simple implementation, not providing the amount of tweaking that sklearn's Perceptron provides. This implementation allowed us to change the "eta" or **learning rate coefficient** (set at 0.01) and the **maximum iterations**(set at 10). The decision boundary was right in between the two classes.

The final decision boundary was reached in the **4th epoch**, after there were zero misclassified points. For **implementation 2**, we had Sklearn's class Perceptron, from linear_model. It was also given the **same configuration**: eta was 0.01 and max_iter was 10. <u>However, the decision boundary obtained here was more pushed towards one class (not desirable).</u>

PART 2: ABSENTEEISM FROM WORK Dataset

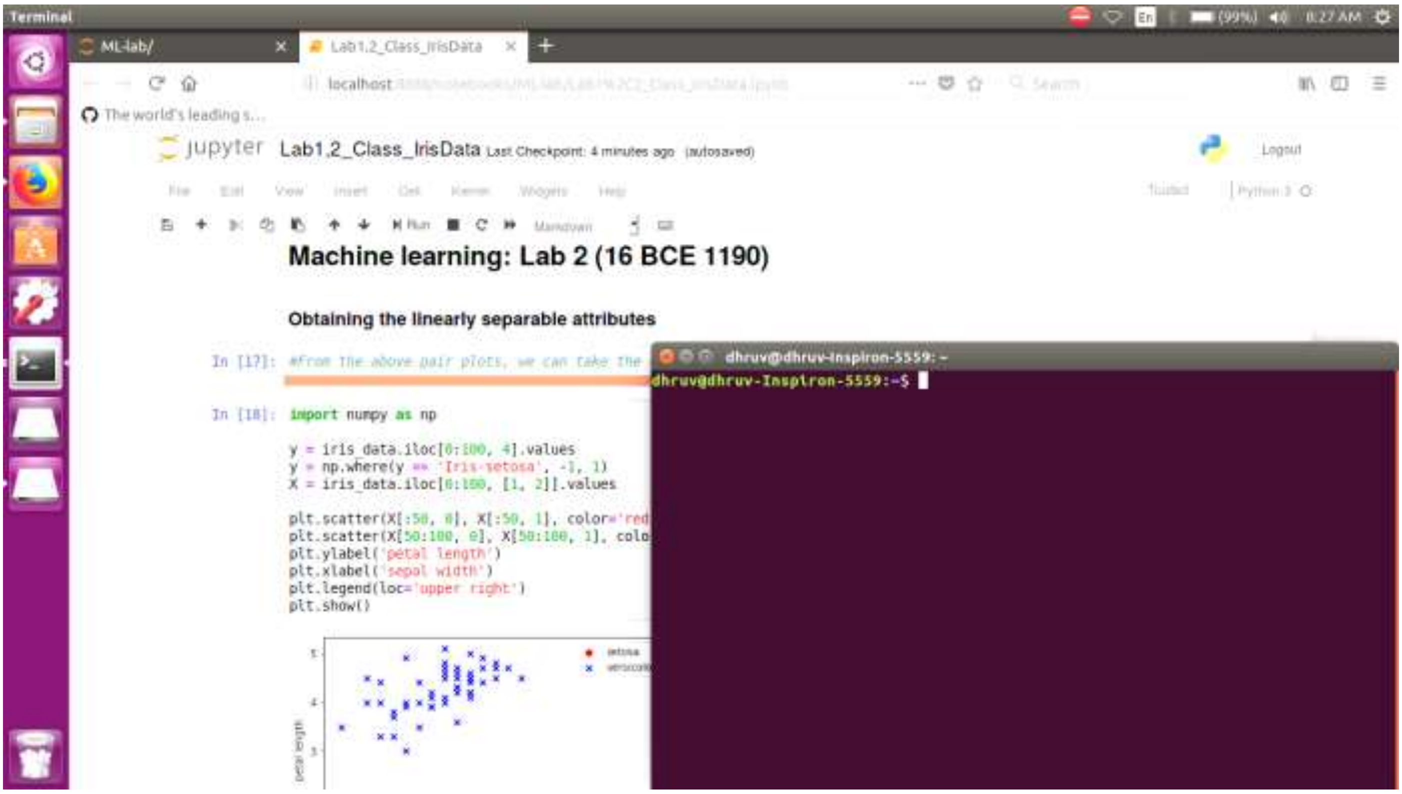
After cleaning the dataset, we proceeded to 2D perceptron classification. For classification, **pairplots** were first plotted to explore the possibility of linearly separable attributes. The **hue** was chosen from an attribute (education) that had a <u>decent number of distinct values, not too few (1 or 2) and not too many.</u> It was observed that 2 categories of "education" were linearly separable with respect to attributes: weight and absenteeism in hours. The **linearly separable valued records** were **extracted from the dataset** and a new temporary dataset was generated. Now, I was able to separate graduate workers from doctorate workers by a linear function. **Implementation 1** involved the book's implementation of the perceptron. The final decision boundary was reached in **25 iterations**, at a **learning rate of 0.01**.

Implementation 2 of the Perceptron involved Sklearn's Perceptron model. It was also able to **converge within 40 iterations**, at a slightly **steeper learning rate of 0.**1. <u>The decision boundaries obtained in both the implementations were very similar.</u>

DHRUV GARG 16 BCE 1190

SCREENSHOTS

Iris-dataset



Chosen dataset – Absenteeism at work