GD and BCGD Optimization Methods in Semi-supervised Learning

Zesen Huang
zesen.huang@studenti.unipd.it

1. Introduction

In this report, we artificially create a two-dimensional data set containing labeled and unlabeled data, and then predict the unlabeled data through the learning of labeled data. How to use GD and BCGD optimization methods in semi-supervised learning process is discussed by constructing loss function and finding its minimum value. In this project, we used Python to write all the programs and diagrams.

2. Dataset

First, we generate 1000 two-dimensional data points (x, y), including 100 label data and 900 unmarked data. The label data is divided into two classes, one labeled with 1 and one labeled with -1. In this project, the same feature represents the same label.

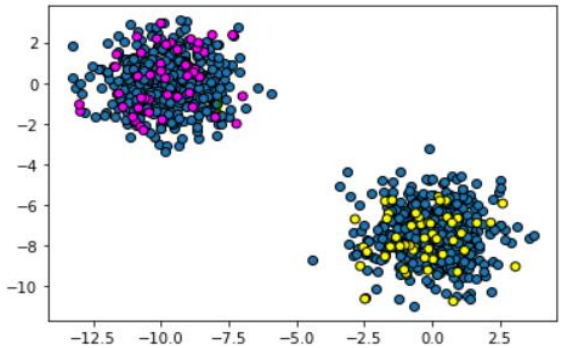


Figure 1 Red and yellow points reparent the label data, and the blue points reparent unmarkedd data

3. Similarity measure

3.1 cosine similarity

First use the cosine similarity to determine the weight between the label and the unmarked, unmarked and unmarked points.

$$cosine\_similarity(u, v) = \frac{\sum_{i=1}^{n} u_i v_i}{\sqrt{\sum_{i=1}^{n} u_i^2}\sqrt{\sum_{i=1}^{n} v_i^2}}$$

The following figure shows the distribution of weights of 4 points for randomly extracted which generated from cosine similarity.

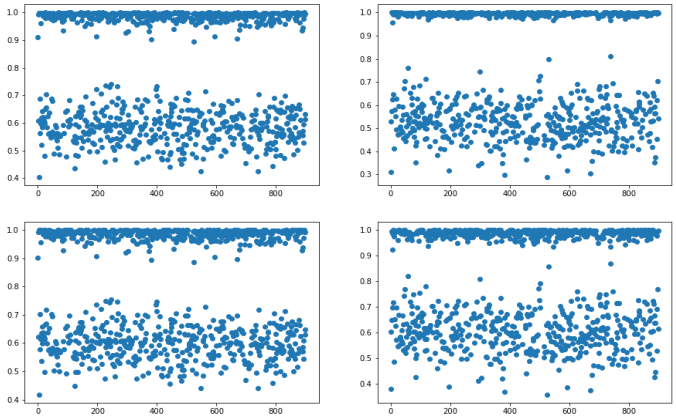


Figure 2 The weight distribution of 4 randomly selected points

Ideally, the smaller the weight, the different the class, and the larger the weight, the same. So ideally, each row should be very large, such as going straight to 1, and very small to go straight to 0, with little (or very little) other values in between. The dividing line of weight determined by cosine similarity is not obvious and is not suitable.

3.2 Euclidean distance

Since cosine similarity is not suitable for determining weights, try using Euclidean distance.

$$Euclidean\ distance = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

To make the difference between the two weights more obvious, we square the Euclidean distance once. The following figure shows the distribution of weights of 4 points for randomly extracted which generated from Euclidean distance.

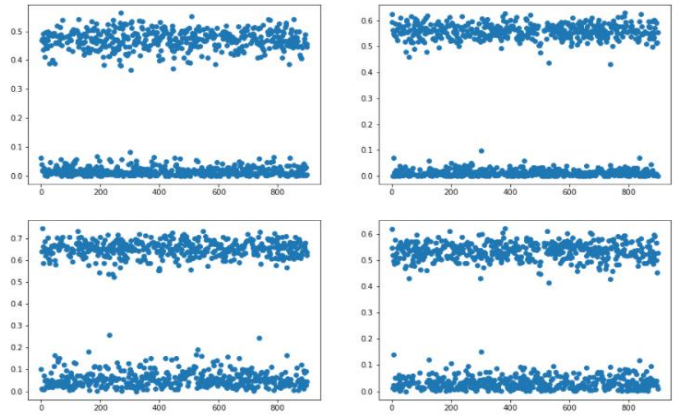


Figure 3 The weight distribution of 4 randomly selected points

As can be seen from the figure, it is better to use the square of Euclidean distance to determine the weight.

4. Establish a loss function

The following loss function is established to evaluate the error of the forecast label.

$$min\sum_{i=1}^{l}\sum_{j=1}^{u} w_{ij}(y^j-\bar{y}^i)^2+\frac{1}{2}\sum_{i=1}^{u}\sum_{j=1}^{u}\overline{w_{ij}}(y^i-\bar{y}^j)^2$$

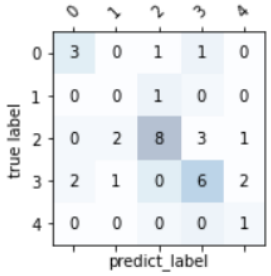
w_{ij} is the similarity between labeled examples i and unlabeled examples j.

$\overline{w_{ij}}$ is the similarity between unlabeled examples j.

Our goal is to minimize the loss function, so the first step is to calculate the gradient of the function.

Next, use sklearn's Bayesian model for training and prediction, and then calculate the accuracy rate and trainning set's confusion matrix.

Here is the confusion matrix:



1.1. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [4]. Where appropriate, include the name(s) of editors of referenced books.

1.2. Illustrations, graphs, and photographs

All graphics should be centered. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text, and choose line widths which render effectively in print. Many readers (and reviewers), even of an electronic copy, will choose to print your paper in order to read it.

You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic. When placing figures in LATEX, it's almost always best to use \includegraphics, and to specify the figure width as a multiple of the line width as in the example below

```
\usepackage[dvips]{graphicx} ...
\includegraphics[width=0.8\linewidth]
{myfile.eps}
```

References

[1] FirstName Alpher, , and J. P. N. Fotheringham-Smythe. Frobnication revisited. Journal of Foo, 13(1):234–778, 2003.

[2] FirstName Alpher, , FirstName Fotheringham-Smythe, and FirstName Gamow. Can a machine frobnicate? Journal of Foo, 14(1):234–778, 2004.

[3] FirstName Alpher. Frobnication. Journal of Foo, 12(1):234–778, 2002.

[4] Actual Author Name. The frobnicatable foo filter, 2014. Face and Gesture (to appear ID 324).

[5] Actual Author Name. Frobnication tutorial, 2014. Some URL al tr.pdf.