Implementation for the

Matan Antebi[1,2] and David Shriker[1,3]

[1] *Tel - Aviv University, Electrical Engineering Faculty*
[2] *ID - 300513678*
[3] *ID - 203718044*

February 2, 2020

This project deals with the areas of stability and generalization in machine learning theory, based on the subjects learned in: the course - Advanced Topics in Machine Learning, under the supervision of Dr. Lior Livni and based on Olivier Bousquet article - Stability and Generalization. The article of Bousquet was exploring different approches of estimates the accuracy learning algorithms and that was based on sensitivity analysis. Sensitivity analysis aims at determining how much the variation of the input can influence the output of a system. The input is a trainning set with m elements. Each element is combined by two (x_i, y_i), were y_i is the i-th label of the i-th element. The impact on the input to make some variant changing in the output may be in two ways: The first by removing the i-th element from the training set, the second is by replacing the i-th element from the training set by drawing another independent data from the dataset. We implemented convergence in two senses: the first - Stability of Bounded SVM regression, the second - Stability of soft margin SVM classification.

1 Introduction

The main subject in the paper of Oliver Bousquet,*Stability and Genaeralization , are learning algorithms. Those algorithms takes pairs of inputs dataset, means that each input has instance label. The output maps instances to the corresponding labels. Oliver explore how much the variation of the input can influence the output of a system - how changes in the composition of the learning set influence the function produces by the algorithm. The outcome of such approach is a way of getting the bounds of the difference between two errors, the emperical and the generalization errors.*
From the paper we chose to make simulation for Stability of bounded SVM regression & Stability of soft margin SVM classification. ***Our goal is to prove the bounds derived in the paper on the generalization error of stable learnig systems do exist****.*
In section 2 we will be noting definition of Generalization, stability, SVM model. Section 3 is about the data we worked with.

- *Systhesized data.*
- *Mnist data set.*

2 Definitions

2.1 Basic Notations - from Stability & Genaralization

Inputs: $\mathcal{X}, \mathcal{Y} \subset \mathbb{R}$.
Training set: $\mathcal{S} = \{z_1 = (x_1, y_1), \cdots, z_m = (x_m, t_m)\}$, z_i*are i.i.d.*
Removing the i-th element:

$$\mathcal{S}^{\setminus i} = \{z_1, \cdots, z_{i-1}, z_{i+1}, \cdots, z_m\}$$

Replacing the i-th element by drawing an independent data from $\mathcal{S}$:

$$\mathcal{S}^{i} = \{z_1, \cdots, z_{i-1}, z_i^{'}, z_{i+1}, \cdots, z_m\}$$

We measure accuracy by defining loss and cost function:

$$l(f, z) = c(f(x), y)$$

More basics notation which will be mantioned later on are defined below:
$\mathbb{E}_z[cdot], \mathbb{P}_z[\cdot]$ *both respectively are the espectation and probability when* z *sampled according to an unknown distribution.*

2.2 Generalization

In machine learning, generalization usually refers to the ability of an algorithm to be effective across a range of inputs and applications. A machine learning algorithm is used to fit a model to data. Training the model is kind of like infancy for humans, which means, examples are presented to the model and the model tweaks its internal parameters to better understand the data. Once training is over, the model is unleashed upon new data and then uses what it has learned to explain that data.
If you over-train the model on the training data, then it will be able to identify all the relevant information in the training data, but will fail miserably when presented with the new data. We then say that the model is incapable of generalizing (overfitting the training data).

2.3 Generalization & Empirical Errors

In supervised learning applications in machine learning theory, generalization error is a measure of how accurately an algorithm is able to predict outcome values for previously unseen data.
Because learning algorithms are evaluated on finite samples, the evaluation of a learning algorithm may be sensitive to sampling error. As a result, measurements of prediction error on the current data may not provide much information about predictive ability on new data. Generalization error can be minimized by avoiding overfitting in the learning algorithm. The performance of a machine learning algorithm is measured by plots of the generalization error values through the learning process, which are called learning curves.
*The main quantity measure to be able to get the performance of the learning algorithm is **generalization error** which define below:*

$$R(A, S) = \mathbb{E}_z[l(A_s, z)]$$

*As known from the generalization error(or basicly unknown), R cannot be computed since the distribution (mantioned in the paper as $\mathcal{D}$) is unknown. Therefore, to compute such error we will define an estimator - **empirical error**:*

$$R_{emp}(A, S) = \frac{1}{m}\sum_{i=1}^{m} l(A_s, z_i)$$

*Another estimator is defined in the paper (**leave one out error**):*

$$R_{loo}(A, S) = \frac{1}{m}\sum_{i=1}^{m} l(A_{\setminus i}, z_i)$$

So as we said, our goal in the project was to implement a simulation of a SVM model (for regression and classification) and to show that the result are tight to the bound given in Bousquet paper, for any $\epsilon > 0$:

$$\mathcal{P}_S[|R_{emp}(A, S) - R(A, S)| > \epsilon].$$

2.4 Stability

Before we'll define stability, let explain the main idea of stability analysis.
Stability analysis enables us to determine how the input variations (dataset) are going to impact the output of our system. In our case, the system is a learning algorithm that ingests data to learn from it. A supervised learning algorithm takes a labeled dataset that contains data points and the corresponding labels. The process of training involved feeding data into this algorithm and building a model.
We need to estimate the model performance,therefore the accuracy metric tells us how many samples were classified correctly, but it doesn't tell us anything about how the training dataset influenced this process. Ideally, we want the model to remain the same and perform its job with the same accuracy.
Therfore, stability of a learning algorithm refers to the changes in the output of the system when we change the training dataset. A learning algorithm is said to be stable if the learned model doesn't change much when the training dataset is modified. It is important of putting an upper bound to the model. Changing the data-set shouldn't change more than a certain threshold regardless of what subset you choose for training. If it satisfies this condition, it's said to be "stable".
In the paper Bousquet mentioned several stabilities, and as we mentioned above earlier the variant of the input for all those satbilities have two possibles for the input:

1. *Choosing a different subset for training.*
2. *Presence of noise in the dataset.*

***** **Does to add*******

- *Hypothesis stability - theorm*
- *Pointwise Hypothesis stability - theorm*
- *Error stability - theorm*
- *Uniform stability theorm*

2.5 overfitting & Regularization

*In statistics, therefore taken to machine learning, **overfitting** is the production of an analysis that corresponds too closely or exactly to a particular set of data, and may therefore fail to fit additional data or predict future observations reliably.*
An overfitted model is a statistical model that contains more parameters than can be justified by the data.
***Regularization** is a form of regression, that constrains the coefficient estimates towards zero. The process of regularization is extually adding some information in order to prevent overfitting.*
In classification a regulizer is used when the data-set is finite. It used by added to the a loss function and helps to get the right solve solution for the minimization of the loss function with hyper-parameter λ.

2.6 SVM

In machine learning, support-vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier.
An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on the side of the gap on which they fall.
When data are unlabelled, supervised learning is not possible, and an unsupervised learning approach is required, which attempts to find natural clustering of the data to groups, and then map new data to these formed groups.

***** **need to explain exactly what David has done in Models** *****

3 Data sets

***** **How to explain about the synthesized data********

***MNIST** database is a large database of handwritten digits that is commonly used for training various image processing systems. This database combined from training set and test set.*

- ***Training**: 60,000 images.*
- ***Test**: 10,000 images.*

3.1 Subsection

4 Section

Bibliography

Arnold, A. S. et al. (Mar. 1998). “A Simple Extended-Cavity Diode Laser”. In: *Review of Scientific Instruments* 69.3, pp. 1236–1239. URL: http://link.aip.org/link/?RSI/69/1236/1.

Hawthorn, C. J., K. P. Weber, and R. E. Scholten (Dec. 2001). “Littrow Configuration Tunable External Cavity Diode Laser with Fixed Direction Output Beam”. In: *Review of Scientific Instruments* 72.12, pp. 4477–4479. URL: http://link.aip.org/link/?RSI/72/4477/1.

Wieman, Carl E. and Leo Hollberg (Jan. 1991). “Using Diode Lasers for Atomic Physics”. In: *Review of Scientific Instruments* 62.1, pp. 1–20. URL: http://link.aip.org/link/?RSI/62/1/1.