Università
Ca'Foscari
Venezia

MASTER'S DEGREE IN ECONOMICS AND FINANCE

Machine Learning applied to Credit Rating for Italian Listed Companies

Graduand:
Leonardo Pollesel
877583

Supervisor:
Ch.ma Prof.ssa Monica Billio

Co-Supervisor:
Ch.mo Prof. Ionut Florescu

Academic year: 2019 / 2020

ABSTRACT

In this paper I use multiple Machine Learning techniques combined with financial data and ratios to build a model able to predict the future credit rating of American Listed companies and transfer the learning to Italian Listed companies.

Credit rating assess the financial stability of a company evaluating the ability to pay back its debts.

Keywords: *Machine Learning, MLP, Random Forest, CNN, Compustat, Bureau van Dijk, Bloomberg*

Table of Contents

INTRODUCTION

Credit rating is used to evaluate the credit risk of debtor accessing the ability to pay back debts, are used for companies to evaluate its financial instrument or the company itself. Credit rating give insights to investors to the creditworthiness of a company providing better investment decisions.

However, credit rating agencies disclose credit ratings on a quarterly base period.

I will use financial data and financial ratios to access the credit rating of companies aiming at predicting the credit rating in order to be a step ahead competitors and take advantage of the time gap between the disclose if financial statements by companies and credit ratings by companies, due to lack of reporting for Italian companies I had implement models based on yearly data.

I will use Machine Learning techniques to find a relation between financial data of companies and their credit ratings, starting from American listed companies I will build multiple models and given a good level of accuracy I will transfer the learning of those models to Italian listed companies.

The process involves the use of different models starting from Multilayer Perceptron, Random Forest and Convolutional Neural Networks.

DATA

Data Sources

To have all the data needed for my analysis I had to use different sources:

- Compustat: American listed companies (financial data, financial ratios and credit ratings)
- Bureau van Dijk (AIDA): Italian listed companies (financial data and financial ratios)
- Bloomberg: Ratings for Italian listed companies (S&P Domestic Long-Term Credit Rating)

Variables Used

To predict credit rating, I used financial data, financial ratios and credit ratings for both American and Italian companies.
To build the model I have used yearly observations, ranging from 2009 to 2017, of 930 American and 378 Italian listed companies for a set of 19 variables:

1. Accounts Payable - Trade
2. Assets - Other
3. Assets – Total
4. Capital Surplus/Share Premium Reserve
5. Cash and Cash Equivalents
6. Current Assets - Total
7. Current Ratio
8. Depreciation and Amortization
9. Earnings Before Interest
10. Employees
11. Gross Profit
12. Interest Expense - Total
13. Inventories - Finished Goods
14. Inventories - Raw Materials
15. Inventories - Total
16. Receivables - Total
17. Stockholders Equity - Parent

18. Total Debt/EBITDA
19. Total Debt/Equity

Data Preparation

Starting from Compustat I had to transform the data for American companies from monthly to yearly for financial ratios and credit ratings and finally merge all the three datasets for the financial data, ratios and ratings into one in order to have it ready for machine learning, I have done this process through Python.
Bureau van Dijk is a database for Italian companies, here the procedure was different since the database is giving data in a format different than Compustat where a variable is defined in multiple columns where each column has a specific date for that variable, so I had to transform it in order to have a column with dates and a single column for a single variable, this process has been done using R.
To attach the ratings for Italian companies I have searched each of the 378 companies on Bloomberg and copied the ratings available.

Data Cleaning

With the two datasets for American and Italian companies ready with financial data, financial ratios and ratings, I have cleaned them first dropping all the observations without a rating because are not usable since without a benchmark, then I have filled all the missing values with -1, a value not represented in the dataset.
In order to use the ratings for machine learning purposes I had to transform it into numerical values using a mapping scheme as follows:

AA+	0
AA	
AA-	
A+	
A	
A-	
BBB+	1
BBB	

BBB-	2
BB+	
BB	
BB-	3
B+	
B	
B-	
CCC+	4
CCC	
CCC-	
CC	
C	
D	
SD	
N.M.	

Table 1: Credi Rating Mapping

This mapping leads to five broad classes of rating balanced in order to have an equal portion of ratings in my dataset falling inside one class.

The classes represent:

- 0: prime grade, high grade, upper medium grade
- 1: lower medium grade
- 2: non-investment grade
- 3: highly speculative grade
- 4: substantial risk grade, extremely speculative grade, default imminent and in default grade

MACHINE LEARNING METHODOLOGY

Multilayer Perceptron

Multilayer Perceptrons are the simplest type of neural networks. They are made of one or multiple layers of neurons (nodes) all interconnected with each other.
The visible layers are the input layer and the output layer, in between there is at least one hidden layer, all the layers are fully connected with each other where each neuron of one layer connects with a certain weight to every neuron of the subsequent layer.
This model has a linear activation function in all neurons mapping the inputs to the outputs of each neuron and each layer is connected with a certain weight to every neuron of the following layer.

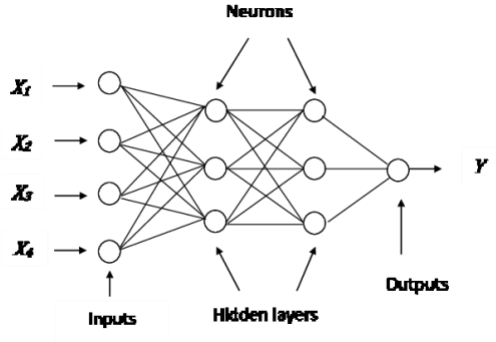

Figure 1: Multilayer Perceptron Model

MLP are used for prediction problems where a label or a class is predicted given a set of inputs.
In my analysis, financial data and financial ratios will be given to MLP as inputs in order to predict to have as output the class of rating where each company falls.

Random Forest

Random forest is an algorithm based on a large number of decision trees all working together in order to give a better predictive performance than could be obtained from any of the decision tree alone.
This process selects at each split in the learning process a random subset of features where if one or more features strongly predicts the response variable will be selected in many of the trees.

The model uses a method called bagging to create multiple subsets from the original data ab it randomly selects different subsets of features to build each decision tree. Creating multiple decision tree based on different subsets of data and different features it increases the probability of finding the right pattern leading to the classification of the data.

This process leads also to the concept of features importance where the most important features will be selected multiple times, this is defined as the relative contribution to the decision-making process of the algorithm.

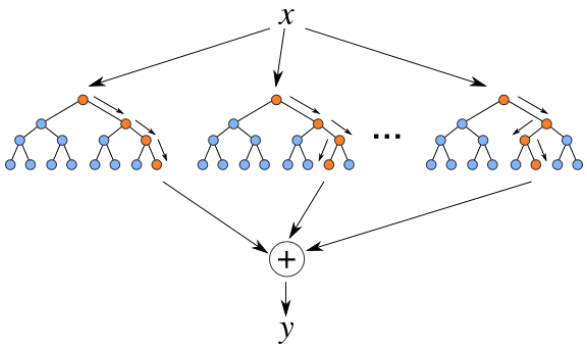

Figure 2: Random Forest Model

Using this method, I was able to draw conclusions about what features contribute most in the decision making of the model and finally use a selection of the most important features for a further analysis and a better prediction in a final model.

Convolutional Neural Network

Convolutional Neural Network is a deep learning algorithm, it is similar to MLP because it consists of an input and output layer with in between multiple hidden layers, but it differs because at least one of the hidden layers is a convolutional one and the neurons in one layer do not connect to all the neurons in the next layer but to a smaller portion of it. The term convolution refers to the mathematical combination of multiple functions merging multiple sets of information, this type of operation is performed on the input data.

The output of the convolution will be passed through an activation function, in my case I've used the Rectified Linear Unit (ReLU) activation function:

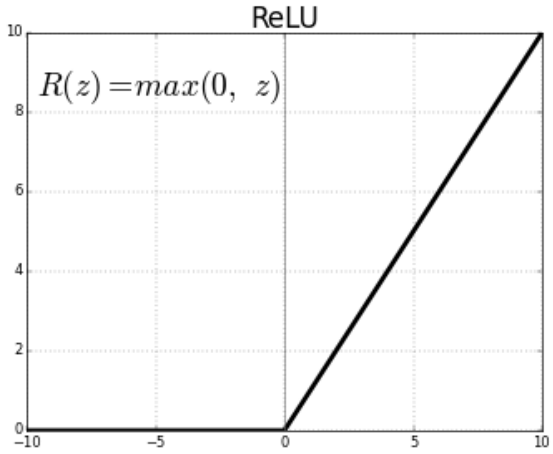

Figure 3: ReLU Activation Funciton

With the ReLU activation function the output is assumed to be zero when it is smaller than zero, this is used in order to have no influence on our result from the missing value filled with -1 while preparing the dataset.

After the convolutional layers to finalize the classification process there are some fully connected layers like in an MLP model giving the final classification for our set of variables.

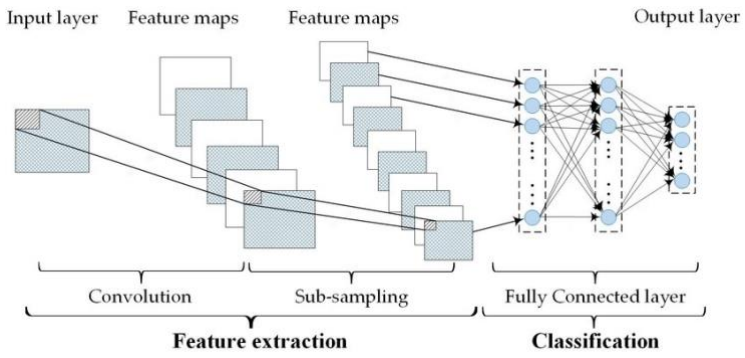

Figure 4: Convolutional Neural Network Model

MODELS ON AMERICAN LISTED COMPANIES

Before transferring the learning form American companies to Italian companies for predicting credit rating I tested MLP and CNN only on American companies. If I have a good level of accuracy, I can further deepen my analysis and transfer the learning to Italian listed companies.

Multiplayer Perceptron

The first step into my analysis is to apply an MLP model only on the Compustat dataset for all American model, I have fitted this model using a different date as split for the dataset into training and testing in order find out the best date to use as separation giving the best performance:

Date Split	Train Accuracy	Test Accuracy
2010	77.32 %	60.54 %
2011	77.53 %	62.33 %
2012	79.26 %	64.90 %
2013	77.25 %	64.38 %
2014	78.33 %	68.70 %
2015	75.47 %	69.71 %
2016	73.59 %	70.60 %

Table 2: MLP Results on American Listed Companies

As can be seen from the table above the best accuracy training the model is given splitting it on 2012 so taking as training all the observations before 2012 and as testing the observations after 2012. The train accuracy percentage tells that the model is able to predict correctly 79.26% of the ratings assigned to the observations used to train the model but at the same time the test accuracy show that using the model trained before it's able to evaluate correctly new observations (from 2013 to 2017 in my case) only 64.90% of the times.

The best accuracy for the testing instead is given by the model trained on data from 2009 to 2016 and tested in data from 2017 to 2018 with a percentage of 70.60% but at the same time the training part lost in terms of accuracy.

The overall best model using MLP is given by slitting the dataset on 2016, with a training accuracy of 73.59% and a testing accuracy of 70.60%.
This is a good result but as shown in the next section it can be further improved using a Convolutional Neural Network.

Convolutional Neural Network

A further deep analysis into American companies has been carried out implementing a Convolution Neural Network with the expectations of a better performance compared to the one of a Multilayer Perceptron model.
The same procedure has been used also for CNN splitting the dataset into training and testing on different dates and fitting multiple models based on the split date chosen:

Date Split	Train Accuracy	Test Accuracy
2010	97.81 %	62.60 %
2011	97.94 %	65.51 %
2012	96.57 %	67.36 %
2013	96.83 %	70.49 %
2014	97.06 %	75.00 %
2015	97.13 %	80.41 %
2016	96.07 %	86.85 %

Table 3: CNN Results on American Listed Companies

The table above shows the date chosen to split the Compustat dataset into training and testing, the training accuracy of the model along with the testing accuracy.
As can be seen there is a high improvement with CNN with the best model given by splitting the dataset in 2016 with a training accuracy of 96.07% and a testing one of 86.85%.
This result shows that the CNN is a performing model to be used for predicting credit rating so I can further implement it to transfer the learning from American to Italian companies.
For the sake of completeness, I will also compare it with MLP.

TRANSFER LEARNING TO ITALIAN LISTED COMPANIES

With the good result on American companies I have transferred the learning of the model to Italian companies with the aiming to predict the credit rating for Italian listed companies using a model trained on American listed companies.

This process has been done using a Multilayer Perception model, a Convolutional Neural Network model and a combination of feature selected with the feature importance given by Random Forest and a Convolutional Neural Network fitted on the feature selected with Random Forest.

Multilayer Perceptron and Convolutional Neural Network

The first step transferring the learning from American to Italian companies has been implemented fitting a model on the entire American dataset and tested on the entire Italian dataset.

I have fitted a Multiplayer Perceptron model and a Convolutional Neural Network and compared the accuracy of the models telling me how good the model performs and the loss function showing how the prediction error behaves training the model.

As expected during the training process the prediction accuracy increases and the loss function decreases:

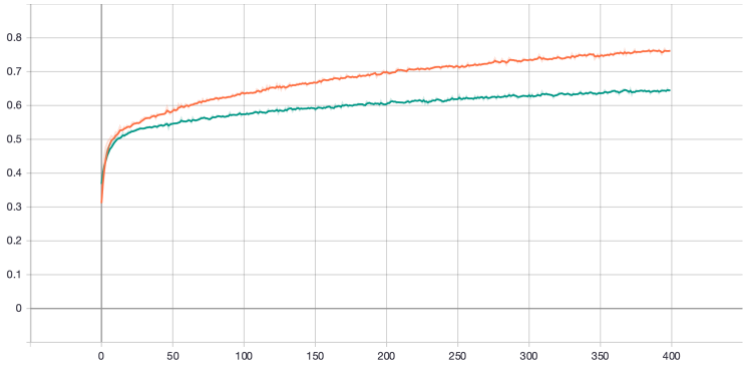

Figure 5: Train Accuracy (green: MLP, orange: CNN)

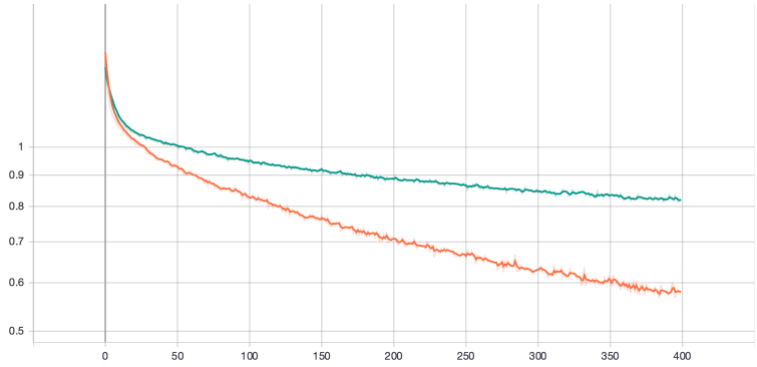

Figure 6: Loss Function (green: MLP, orange: CNN)

The graphs above are showing on the left the prediction accuracy of MLP and CNN and on the right the prediction error of both models.

Here is clear that CNN has better performance than MLP as the prediction accuracy is much higher and the prediction error decreases faster end is much lower.

The final results show for MLP and accuracy on training (American companies) of 67.35% and a testing accuracy (Italian companies) of 43.06%.

For CNN the results are slightly better with a training accuracy of 87.06% and a testing accuracy of 34.72%.

This result is in line with expectations regarding the training of the model but transferring the learning to Italian companies poorly predicts ratings for Italian companies.

To further improve the model, I will use feature selection from Random Forest.

Random Forest Feature Selection and Convolutional Neural Network

To improve the accuracy of the model I have used the feature selected from random forest and ordered them in from the most important one to the least important one.

Using only a selection of features I can avoid using variables that are not giving contribution to the decision-making process of the model having a more precise and more accurate model.

Below are shown the features with its relative importance in prediction credit rating:

Variable	**Importance**
Stockholders Equity - Parent	0.447161

Total Debt/EBITDA	0. 091476
Capital Surplus/Share Premium Reserve	0. 090335
Employees	0. 045834
Depreciation and Amortization	0. 040951
Assets - Other	0. 040038
Earnings Before Interest	0. 038552
Assets - Total	0. 033965
Total Debt/Equity	0. 027884
Gross Profit (Loss)	0. 025507
Inventories - Total	0. 019640
Current Assets - Total	0. 018732
Receivables - Total	0. 017662
Accounts Payable - Trade	0. 016885
Current Ratio	0. 015836
Inventories - Finished Goods	0. 010458
Inventories - Raw Materials	0.009409
Cash and Cash Equivalents - Increase/(Decrease)	0.008813
Interest Expense - Total (Financial Services)	0. 000861

Table 4: Random Forest Feature Importance

Having the feature importance, I have fitted a Convolutional Neural Network using a selection of the best features in the process of predicting credit ratings. I have trained the model using the best 7 features up to the best 10 to have a clear view of the best subset of variables to use.

N° of Features	Train Accuracy	Test Accuracy
7	87.47 %	29.16 %
8	91.11 %	41.66 %
9	88.14 %	29.16 %
10	92.58 %	22.22 %

Table 5: Convolutional Neural Network with Random Forest Feature Importance

With feature selection I was able to further improve the model having a final training accuracy of 91.11% and a testing accuracy 41.66% with the model using the best 8

features with the aim of transferring learning from American to Italian companies for the prediction of credit ratings.

CONCLUSIONS AND FUTURE WORK

I can conclude that the financial data and financial ratios used are useful for predicting credit ratings.

To summarize building a Convolutional Neural Network for American listed companies gave me a good level of accuracy in predicting credit ratings.

Transferring the learning from a model trained on American listed companies to Italian listed companies does not give a high level of accuracy.

This could be because of high differences in the structure of companies, furthermore the model can perform better if the data reported in Bureau van Dijk were more accurate with less missing values.

Further improvements can be done by selecting an higher range of variables from Compustat and Bureau van Dijk or from a different database for Italian companies since as mentioned before Bureau van Dijk is lacking of some data.

One further test is to try transferring the learning from European companies rather than American one since they have a structure closer to Italian companies.

Finally, if a database of Italian listed is large enough there is the possibility to avoid transferring the learning from American listed companies.

REFERENCES

Abdou, Hussein, et al. "Prediction of financial strength ratings using machine learning and conventional techniques." *Abdou, HA, Abdallah, WM, Mulkeen, J., Ntim, CG, & Wang, Y.(2017)'Prediction of financial strength ratings using machine learning and conventional techniques', Investment Management and Financial Innovation* 14.4 (2017): 194-211.

Quinlan, J. Ross. "Induction of decision trees." *Machine learning* 1.1 (1986): 81-106.

Breiman, Leo. "Classification and Regression Trees." (2017).

Golbayani, Parisa, Dan Wang, and Ionut Florescu. "Application of Deep Neural Networks to assess corporate Credit Rating." *arXiv preprint arXiv:2003.02334* (2020).

Kvamme, Håvard, et al. "Predicting mortgage default using convolutional neural networks." *Expert Systems with Applications* 102 (2018): 207-217.

Khashman, Adnan. "Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes." *Expert Systems with Applications* 37.9 (2010): 6233-6239.

Ye, Yun, Shufen Liu, and Jinyu Li. "A multiclass machine learning approach to credit rating prediction." *2008 International Symposiums on Information Processing*. IEEE, 2008.

Index of Figures

Index of Tables