

Mutual Fund Rating Prediction using Proportional Odds Logistic Regression with Imbalanced Class

Ashoka Wilson Dsouza, Ismail B

Abstract: Mutual funds ratings given by rating agencies, are very popular and helps new/first time investors to select and invest in funds based on the ratings a fund takes without going through the detailed portfolio. However sometimes these ratings could be biased or incorrect or in favor of specific fund and it could affect an investor decision. New investors face a lot of problems while investing and choosing mutual funds due to poor professional advice and lack of right tools and resources to assess a fund's true performance. To overcome the problem of incorrect rating and to help an investor to choose the funds wisely using machine learning, we have attempted to predict the rating and classify mutual funds using proportional odds logistic regression which classifies funds into rating classes from 1 to 5 with 5 being the high rated fund and 1 being the low rated fund. While some prior studies have suggested methods of using clustering to classify based on performances using Supervised/Unsupervised learning, this paper deals with supervised learning for predicting the ratings using the mutual fund financial ratios and also handles imbalanced classes. To handle imbalance class problem in a multi-class setting, we propose a new class balancing hybrid methodology of using EM and Gauss-Smote sampling that significantly improves the rating prediction.

Keywords: Classification, Gauss-Smote, Imbalanced Classes, Mutual Fund Rating, Proportional Odds Model

I. INTRODUCTION

Mutual funds are one of the most preferred investments by new investors to build wealth since it's managed by Fund managers who make daily decisions on buying and selling of the securities, and the risk being diversified. Whether an investor own them through a brokerage firm or buys them directly, one needs to have a good understanding on their performances and returns. There is significant increase on number of mutual funds offered over the last few years. This has led to a number of rating and classification systems that provide investors with useful information about the return and quality of funds. Majority of these ratings and classifications are created by private, national or global rating organizations. For example, rating agencies such as Value Research, AMFI, CRISIL, Morningstar etc., group funds into categories according to their theme, capital, portfolio planning, growth prospects, etc. This information is then used

along with financial ratios to construct a more descriptive rating system commonly referred to as Star Rating.

Every rating agency has its own unique assessment system and we also have associations of mutual funds that maintain and publish their own ratings. However these ratings made by agencies are sometimes questionable. There may be problems arising due to wrong classification and rating which in general do not coincide. There are also few disadvantages around the rating system since they are often subjective in nature and need time to be elaborated. Most retail investors simply keep buying the latest table-toppers marked by star rating and rely on publicly available information from websites. A mutual fund rating is usually based on the measurement of a fund's historical performance [12] when compared to funds in same category. The fund ratings are determined utilizing financial ratios and the stars are assigned based on the percentile. For example, the top percentile funds are allotted a 5 star rating, while the bottom funds are given 1 star rating.

Many people believe that investment flows in mutual funds is closely linked to a fund's ratings (Franecki 2000). Ratings are used by most of investors to select mutual funds, so there are several studies as Faff, Parwada and Poh (2007), Del Guercio and Tkac (2008) that support the evidence about investor's investment decisions is mostly based on ratings. Star ratings are the most popular among mutual fund investors, because ratings are goals based on historical performance and easy to understand.

In this paper we use proportional odds classification model using machine learning which can automatically give rating to the funds and help investors to make appropriate decisions while selecting the funds based on the rating predicted by the model without being biased and reducing the human error. It also helps an investor to find the rating of a fund and choose the table toppers without relying on rating agencies. There is need of Research work to assess the performance and correct ratings of different mutual funds in India and keep the mutual fund investors fully aware of it.

Classification using machine learning includes determining the actual class to which an object belongs and grouping it accordingly. Several classifiers are binary, leading to a yes / no decision while others are multi-class, capable of classifying an object into one of many classes. Logit/Probit regression is used to model binary responses while Ordinal logistic regression is used to model multilevel ordinal variables.

Revised Manuscript Received on January 31, 2020.

*Correspondence Author

Ashoka Wilson Dsouza*, Department of PG Studies and Research in Statistics, Mangalore University, Konaje, India.

Email: ashokdesouza@gmail.com

Dr Ismail B., Professor of Statistics, Statistical Unit, Yenepoya (Deemed to be University), Derlakatte, Mangalore, India.

Email: prof.ismailb@gmail.com ; statistics@yenepoya.edu.in

In ordinal regression the dependent variable is ordered or ranked. In our case of mutual fund, the ratings variable is ordered (from 1 to 5, with 5 being the best fund and 1 being the poor fund) and is the dependent variable while other financial ratios such Standard Deviation, Sortino Ratio, Sharpe Ratio, Alpha, Beta and R-Squared are independent features. We then use proportional odds model as a classifier to predict an outcome and group funds by rating. There are many conventional methods for ordinal classes that can be employed to solve ordinal data problems. But, the use of machine learning classification techniques designed specifically for ordinal class data yields simpler and faster results and with better performance classifiers. Also since our dataset has class imbalance problem, we propose a new methodology of using EM and Gauss-Smote generated samples for preprocessing which improves the rating prediction significantly over SCUT[1] sampling.

This paper is structured as follows. Section 2 contains a description of the brief literature review. In Section 3, the proposed data preprocessing technique and results are discussed with comparison to prediction performance while Section 4 discusses our future plans and conclusion.

II. LITERATURE REVIEW

Several studies are conducted by the researchers which is related to the performance evaluation and classification of mutual funds using machine learning methods. For instance, a study carried out by Copen et al. (1996) examined how investors make investment decisions for mutual funds and it indicated that investors had very little knowledge of investment strategies. DiBartolomeo and Witkowski (1997) used iterative approach from Sharpe's method (1992) of style analysis for classification of equity mutual funds and their results show that almost 40% equity funds are misclassified and this can impact the rating of fund. Blume Marshall (1997) studies the method that Morningstar uses in assigning ratings and then explores in detail the implications for the ranking of domestic equity mutual funds. Indro, Jiang, Patuwo and Zhang used an artificial neural network to predict mutual funds' performance (Indro et al., 1999). They find that the Neural Network generated better forecasting results than classical linear models for funds. Ludwig and Piovoso (2005) apply Decision Trees, Neural Networks and Naïve Bayes to compare money managers. They use input features such as 1-, 2- and 5-year returns, percentage of outperforming quarters, tracking error and various financial ratios. The resulting accuracy from predicting subsequent managers' performance is above 65% and exceeded the performance of a simple scoring model. Francesco Lisi and Edoardo Otranto (2008) proposed a solution using three step procedure for classification of mutual funds in terms of risk and gain. They first characterize time series funds in terms of returns and in the second step, classes of homogeneous funds are clustered with respect to the risk levels. In the third step the previous two methods are merged and the results showed very similar to Morningstar rating. Also the clustering algorithm proposed by Lytkin et al. (2008) showed the results are based on return levels only for different kinds of functions. Del Guercio and Tkac (2009) showed how change in Morningstar rating can substantially influence investor decisions towards mutual

fund investment allocation. Alimi et al (2012), proposed a method of clustering similar mutual funds based on the historical investment performance similarities using k-means and Ward method. They considered six characteristics including rate of return, turnover rate, variance, semivariance, Treynor Ratio and Sharpe Ratios financial ratios for multi-objective portfolio optimization and classified mutual funds based on performance similarities. Axel Pers and Adam Orre (2017) classified mutual funds using ANN. They investigate how well artificial neural networks can predict future mutual fund performance and help financial advisors. When it comes to mutual fund classification and rating, we need to have a fair comparison of performance and ranking. Also we need to further explore and identify the correct classification and ratings of individual mutual funds without being biased.

III. PROPOSED METHODOLOGY

The present research is an attempt to automatically classify and rate, selected large cap equity mutual funds of selected Indian companies into rating classes using machine learning techniques. The following are the various steps followed

A. Data collection and pre-processing:- The data for this study is collected from Value research online website which consists of 98 Indian large cap equity mutual funds only along with ratings and financial ratios such as Standard Deviation, Sortino Ratio, Sharpe Ratio, Alpha, Beta and R-Squared that evaluate the performance of each fund. The outcome/dependent variable is rating which is ordered from 1 to 5 classes with 1 being poor fund performance and 5 being the best fund performance. The funds are rated by Value Research Online using both returns and risk based on financial ratios. The fund ratings are assigned as per the following class distribution:

Rating	Percentage Distribution
High (5)	Top 10%
Above Average(4)	Next 22.5%
Average (3)	Middle 35%
Below Average (2)	Next 22.5%
Low (1)	Bottom 10%

In our dataset with total of 98 funds, there are 10 funds with Rating 1, 26 funds with Rating 2, 31 funds with Rating 3, 24 funds with Rating 4 and 7 funds with Rating 5. We notice the problem of class imbalance with class 1 and 5 which are minority classes while class 2, 3 and 4 are majority classes. Since we have large number of majority class examples in our data, a classification algorithms tends to favor class with majority proportion of observations while incorrectly classifying the minority class examples, thereby decreasing the prediction power. Hence for imbalanced datasets, we need to correctly balance the class distribution first and then use the balanced data for classification. Ramanan et al. (2007) proposed the OVO and OVA approaches for multiclass imbalanced problems.



Hence to overcome class imbalance and improve the classification performance, sampling methods are used to pre-process the data. Sampling methods for imbalanced data may be grouped into two categories, under-sampling and oversampling, while under-sampling is used to reduce the number of instances/samples from majority class, oversampling increases the number of minority class instances. Some of the commonly used class balancing sampling techniques are Random under-sampling, Random Oversampling, Clustering, Under-sampling with ensemble learning, SMOTE, ROSE, ADASYN etc. and mixed strategies approaches such as performing oversampling using ROSE and undersampling using clustering. In our paper we have employed two methods i) the SCUT algorithm[3] and ii) EM and Gauss-Smote[9], to handle the class imbalance and to find a better solution among the two. The SCUT algorithm developed by Astha Agrawal et al. (2015) oversamples minority class examples through SMOTE [3] and under-samples majority class examples using cluster analysis based on EM algorithm. In addition, it handles both within-class and between-class imbalance. In SCUT, cluster-based under-sampling technique is employed, using the Expectation Maximization (EM) algorithm (Dempster et al., 1977) and oversampling using the SMOTE algorithm (Chawla et al., 2002). We then propose a new mixed strategy method EM+Gauss and test the data by combining EM and Gauss-Smote in which the cluster analysis using EM algorithm (Dempster et al., 1977) is employed for under-sampling the majority classes and Gauss-Smote (Hansoo Lee et al., 2017) is used to oversample the minority class. We then compare our experimentation results between imbalanced class, SCUT and EM+Gauss. The methods are employed using R-software.

B. Model Building and Results

We use Proportional Odds Model (POM) [13] described by McCullagh (1980) to classify the ordinal responses. The ordinal scale is ranked and is different from the nominal scale. The POM is the most popular model for ordinal data (Bender and Grouven, 1998) and extends the logistic regression model to situations where the response variable assumes ordinal values. The POM model uses cumulative probabilities upto a threshold, thereby making the whole range of ordinal categories binary at that threshold which then classifies into respective classes. On running the POM model for the given dataset with imbalanced class, we get a prediction accuracy of 60% and other evaluation metrics [16] by individual classes shown in Table-I below. We observe that the overall accuracy of the imbalanced model is poor which is also explained by the uneven proportion of performance measure for each classes as shown in Table-I below. This result clearly shows the effect on the overall accuracy due to larger proportion of the skewed majority class. Also we drop the Sortino Ratio variable from all further analysis due to high correlation with Sortino ratio

Table- I: Evaluation Metrics for imbalanced class

Evaluation Metrics	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	1.00	0.71	0.67	0.45	0.33
Specificity	1.00	0.96	0.71	0.84	0.96
Precision	1.00	0.83	0.36	0.63	0.5

Recall	1.00	0.71	0.67	0.45	0.33
F1 Measure	1.00	0.77	0.47	0.53	0.4

To test the assumptions of POM model, we first run the POM model on the data set and then, check the assumption of parallel lines which can be done using the Brant test[2] for parallel regression assumption (Brant, 1990). The Brant Test is run using R and the output is interpreted as follows. If the probability (p-value) is greater than your alpha level (0.05), then your dataset satisfies the proportional odds assumption and if the assumption of parallelism is rejected, then it is better to consider using multinomial regression.

Brant Test for parallel lines-The null hypothesis is H0: Parallel Regression Assumption holds

	chi2	df	Probability> chi2
Omnibus	9.702572	18	0.941
Standard Deviation	3.236661	3	0.357
Sharpe Ratio	3.142999	3	0.370
Sortino Ratio	1.604765	3	0.658
Beta	1.674649	3	0.643
Alpha	1.594072	3	0.661
R-squared	4.822623	3	0.185

We note for our dataset since all the p-values are greater than α (Chi2=0.05), the proportional odds model can be used to model our data. The Omnibus is for the whole model, the rest for the individual coefficients.

EM and SMOTE (SCUT):To further improve the prediction accuracy of ratings we then employ the SCUT algorithm developed by Astha Agrawal et al. (2015) which oversamples minority class instances through the SMOTE and under samples majority classes using cluster analysis. Our data set contains 5 classes with approximate overall class mean 20. We have 10 funds with Rating 1, 26 funds with Rating 2, 31 funds with Rating 3, 24 funds with Rating 4 and 7 funds with Rating 5. We notice the problem of class imbalance with class 1 and 5 which are minority classes while class 2, 3 and 4 are majority classes. Hence we apply SMOTE for class 1 and 5 which are minority classes to get 20 samples each generated by SMOTE oversampling and for class 2, 3 and 4 we run the EM algorithm using the MClust function in R which creates clusters and a total of 20 instances are randomly selected from each cluster. We now merge all the classes and a new balanced data set of 100 instances (in which each class has 20 instances) is obtained. Now on running the POM algorithm on the balanced data classes obtained using SCUT, we get an overall accuracy of 70% which shows an improvisation over imbalanced class.

Table- II: Evaluation Metrics for SCUT

Evaluation Metrics	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	1.00	0.50	0.67	0.75	0.60
Specificity	1.00	0.88	0.89	0.92	0.95
Precision	1.00	0.50	0.40	0.60	0.86
Recall	1.00	0.50	0.67	0.75	0.60
F1 Measure	1.00	0.50	0.50	0.67	0.71



EM and Gauss:The SMOTE, algorithm generates synthetic samples by using a combination of k-nearest neighbors and random numbers generated from uniform distribution. However SMOTE causes over-generalization problem when there is large gap between classes. Hence we have Gaussian-based SMOTE(Hansoo Lee et al) method that assures more diversity while generating artificial samples for minority classes. Hansoo Lee et al 2017[9] showed that Gaussian-based SMOTE algorithm shows improved performance than the SMOTE algorithm. Hence we now combine and proposed a new hybrid class balancing technique which combines EM algorithm for undersampling and Gauss-Smote for oversampling to resolve the class imbalance problem in our dataset. First 20 Gaussian data samples are generated for minority classes of 1 and 5 and then combined with EM algorithm samples for majority classes as drawn earlier. We now again merge all the classes and a new balanced data set of 100 instances (in which each class has 20 instances) is obtained using EM and Gauss sampling. Now on using POM model on the new data set, we observe that the new combined sampling method of EM and Gauss-smote gives a better accuracy of 76.6% than EM and SMOTE with 70% accuracy and original imbalanced classes with 60% accuracy. Also we note from Table-III that the Sensitivity, Specificity, Precision, Recall and F1 measures of the EM+Gauss show higher performance than SCUT. The results therefore confirm that the method proposed could solve the problem of class imbalance and predict rating of mutual funds with greater accuracy than the SCUT algorithm for mutual funds.

Table- III: Evaluation Metrics for EM + Gauss

Evaluation Metrics	Class: 1	Class: 2	Class: 3	Class: 4	Class: 5
Sensitivity	0.83	0.50	0.67	0.78	1.00
Specificity	1.00	0.92	0.81	1.00	1.00
Precision	1.00	0.60	0.29	1.00	1.00
Recall	0.83	0.50	0.67	0.78	1.00
F1 Measure	0.91	0.55	0.40	0.88	1.00

Table IV below shows the overall accuracy of model for each class balancing technique along with its p-values and class intervals. Since all p-values are statistically significant for $p < 0.05$, we conclude that our results are valid and statistically significant.

Table-IV: Model Summary of p-values, accuracy, Confidence intervals for each Class Balancing Technique

Evaluation Metric	Class Balancing Techniques		
	Imbalanced Class	SCUT	EM and Gauss
Accuracy (%)	60	70	76.6
p-value	0.007876	0.00004433	0.0000001798
95% CI	(0.406, 0.7734)	(0.506, 0.8527)	(0.5772, 0.9007)
No Information Rate	0.366	0.33	0.3

IV. CONCLUSION AND FUTURE WORK

The proportional odds model is a convenient and easy-to-implement model for modeling ordered categorical outcomes. In our study we have used the POM model to categorize and classify the ratings of mutual funds which are

ordered. We then predict the ratings for a list of test data and observe that the POM model is appropriate for our data since it can predict ratings with 60% accuracy for imbalanced classes while the prediction accuracy is substantially improved on balancing the classes using SCUT and EM- Gauss. We then observe the new combined model of EM and Gauss-smote gives a better accuracy of 76% than SCUT with 70% accuracy. The model can be further improvised by working on decision boundaries for each class and fine tuning class imbalance which can improve the prediction accuracy. Also we can set our own benchmarks for financial ratios and train the classifier instead of considering the data from rating agencies. For this study we have considered only large cap equity mutual funds and this study can be further extended to small, midcap, multicap funds etc. too to classify and predict the rating of a fund which can help investors on self-decisions rather than going by ratings given by rating agencies. It can also help the rating agencies to evaluate the ratings given by them and cross-validate.

REFERENCES

1. A.Astha, V.L.Herna & P. Eric, (2015), "SCUT: Multi-Class Imbalanced Data Classification using SMOTE and Cluster-based Undersampling", KDIR, SciTePress 226-234.
2. Brant, R. (1990) "Assessing proportionality in the proportional odds model for ordinal logistic regression". Biometrics, vol. 46, 1171-1178.
3. Chawla, N., Bowyer, K, Hall and Kegelmeyer, W. P. (2002). "Smote: synthetic minority over-sampling technique", Journal of artificial intelligence research, vol. 16(1):321-357.
4. D. Acharya, G. Sidana, "Classifying mutual funds in India: Some results from Clustering", Indian Journal of Economics and Business, vol. 6, no. 1, pp. 71-79, 2007.
5. F. Pattarina, S. Paterlini, T. Minerva (2014), "Clustering financial time series: an application to mutual funds style analysis", Computational Statistics & Data Analysis, vol. 47, 353-372.
6. Hereil, Pierre, M, Philippe, M, Nicolas & Roncalli, Thierry. (2010). "Mutual Fund Ratings and Performance Persistence". SSRN Electronic Journal.
7. Indro, D. C. & Jiang, C. X. & Patuwo, B. E. & Zhang, G. P., 1999. "Predicting mutual fund performance using artificial neural networks," Omega, Elsevier, vol. 27(3), 373-380.
8. Jaime, Cardoso and Joaquim. d. Costa, (2007), "Learning to classify ordinal data: the data replication method," Journal of Machine Learning Research, vol. 8, 1393-1429.
9. Lee, Hansoo & Kim, Jonggeun & Kim, S. (2017). "Gaussian-Based SMOTE Algorithm for Solving Skewed Class Distributions". International Journal of Fuzzy Logic and Intelligent Systems. 17. 229-234.
10. Lisi, Francesco & Otranto, Edoardo. (2008). "Clustering Mutual Funds by Return and Risk Levels". Mathematical and Statistical Methods for Actuarial Sciences and Finance. Vol 10. 978-988
11. Marathe, Achla & Shawky, Hany. (1999). "Categorizing mutual funds using clusters". Advances in Quantitative Analysis of Finance and Accounting. 7.
12. Michael C. Jensen, (1968), "The Performance of Mutual Funds in the period 1945-1964", The Journal of Finance, 389-416
13. McCullagh P, (1980), "Regression models for ordinal data," Journal of the Royal Statistical Society, Series B, vol. 42, Issue no. 2, 109- 142.
14. Dhume, P.S.Shenvi and Prof. B. Ramesh (2011), "Performance Analysis of Indian Mutual Funds with a Special Reference to Sector Funds", The Indian Journal of Commerce, Vol. 64, No. 9.
15. S. Takumasa, M. Tohgoroh, Mutoh, Atsuko, Inuzuka, Nobuhiro. (2015). "Clustering Mutual Funds Based on Investment Similarity". Procedia Computer Science.60.881-890
16. T, Grigorios & K, Ioannis. (2009), "Multi-Label Classification: An Overview". International Journal of Data Warehousing and Mining. Vol 3. 1-13.



AUTHORS PROFILE



Mr. Ashoka Wilson Dsouza, is currently pursuing his Ph.D. (Statistics) from Mangalore University. He received his M.Sc. (Statistics) from Mangalore University. His areas of research interest are Machine learning, Limited Dependent Variables in Econometrics and Time Series Analysis. He is also a guest faculty in the department of Big Data Analytics at AIMIT- St Aloysius College.



Dr. Ismail B., is Retired Professor from Mangalore University. Currently he is working as Professor in Department of Statistics at Yenepoya (Deemed to be University). He received his Ph.D. (Statistics) from Mangalore University. He was Common wealth Fellow 2000 and worked as Honorary Senior Research Fellow, Department of Statistics, University of Glasgow, Scotland. His research interest is on Nonparametric Regression, Econometrics and Financial Time series. He has worked on modeling discontinuous phenomenon. Estimation of jump size and testing for discontinuities in the field of regression analysis.