2021 2nd Global Conference for Advancement in Technology (GCAT)
Bangalore, India. Oct 1-3, 2021

Future Sales Prediction For Indian Products Using Convolutional Neural Network-Long Short Term Memory

1st Pooja Kaunchi
dept of Electronics and Telecommunication
(Master of Technology)
Vishwakarma Institute of Information Technology
(Pune University)
Pune, India
kaunchi.pooja87@gmail.com

2nd Tushar Jadhav
dept of Electronics and Telecommunication
(Associate Professor)
Vishwakarma Institute of Information Technology
(Pune University)
Pune, India
tushar.jadhav@viit.ac.in,
jtushar13@gmail.com

3rd Yogesh Dandawate
dept of Electronics and Telecommunication
(Professor)
Vishwakarma Institute of Information Technology
(Pune University)
Pune, India
yogesh.dandawate@viit.ac.in,
yhdandawate@gmail.com

4th Pankaj Marathe
(Application Engineer)
Web-Minds IT Solutions
Pune, India
hr@webmindsitsolution.com

2021 2nd Global Conference for Advancement in Technology (GCAT) | 978-1-6654-1836-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/GCAT52182.2021.9587668

***Abstract*— Sales forecasting is an important aspect of modern market intelligence. A reliable revenue forecast will help a company preserve capital on unnecessary product, prepare better for the future, and increase profit. The estimation of grocery sales is associated with predicting the potential sales of stores such as supermarkets, retail outlets, and bakeries. It enables businesses to effectively distribute capital, forecast realistic sales income, and prepare a stronger strategy for the store's potential development. Conventional forecasting system fails to compete with big data and the precision of revenue forecasting. These problems may be solved by using different data processing strategies. This paper focuses on product sales predictive analytics challenges centered on historical sell data and time-series analysis. Future sales Pre-diction is done by using a hybrid combination of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). Proposed model is tested on a real-time big-mart dataset obtained from the local market shops. According to the Predictive approach used, the estimates obtained indicated that the number generated does reflect the actual data where the maximum degree of precision was 97 percent and the minimum degree of precision was only beyond 82 percent.**

Keywords*— Sales Forecasting, CNN, LSTM Model, Time series analysis

I. Introduction

Sales are the backbone of any enterprise, and predictive analytics is critical to the success of every enterprise. Better prediction aids in the development and improvement of corporate plans by growing industry awareness. Accurate sales prediction is critical for any enterprise, whether that was a small and medium business or a large corporation. Sales predictive analytics is the method of predicting potential short-term or long-term results using the company's sales data from previous times. That was one of the principles of good financial management. Sales Forecasting is being used to forecast demand for various goods sold at the Big Mart Stores or different stores in different towns. As the number of goods and channels continues to increase, estimating these by intuition appears difficult. A traditional revenue prediction delves thoroughly into past circumstances or events and then extends assumptions about consumer acquisition, identifying insufficiency and advantages until establishing management and sales strategies over the next year.

Sales predictive analytics issues that are both interesting and challenging are fairly popular. Sales predictive analytics is an essential component of modern market intelligence. Sales may be conceived of as a time series. Various forecasting models, such as Convolutional Neural Network (CNN), Auto-Regressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), and others have been built during that period. The advanced prediction approach in machine learning, or the artificial neural approach to estimate, has conceptual help to established, accurate knowledge projections. Moving Average and Auto-regression are the foundations of conventional prediction models.

Here, We have collected the DATASET from big-mart stores to implemented the system. In contrast, the Proposed Method will also test the performance of feed-forward artificial neural networks with time series inputs.

The presented paper is organized as follows. Section II presents the literature survey of the prediction of the future sales prediction for Indian Products using Convolutional Neural Network-Long Short Term Memory (CNN-LSTM) and Deep learning. Section III represents the Methodology of product sales prediction. The Experimental result is explained in Section IV and finally, Section V which concludes the Convolutional Neural Network-Long Short Term Memory (CNN-LSTM) based methodology and its performance.

II. Literature Survey

Machine Learning is defined as the computer program which learns by itself from its experience without any human interference. Research on future sales prediction for Indian Products using Convolutional Neural Network-Long Short

Term Memory (CNN-LSTM) has been done and some of them have been discussed below:

In paper [1], author S. Cheriyan examined the idea of sales data and sales forecasting partially. The numerous methods and metrics for sales forecasting are listed in the research work's later section. A best-fit statistical model for the revenue pattern prediction is proposed based on a performance assessment. Pan, H [2] suggests a method to forecast product demand using a convolutional neural network to analyse e-commerce details. This article incorporates the intrinsic existence of the related product details with the actual freight log data, which could be translated into a particular "data frame" structure. B. Singh [3] aims to predict future revenues at Amazon.com Inc., based on past revenue results. They consider three Prediction models based on previous data patterns, Auto-Regressive Integrated Moving Average (ARIMA) and neural network auto regression models.

Yuan, H [4] describe the significant element of our study to the creation of a novel technique for retrieving shoppers' preferences about product review issues to improve sales prediction results. Y. Li. [5] developed a forecast overall sales using a hybrid Gated Recurrent Units (GRU)-Prophet Framework including an attention mechanism. The Prophet framework and the GRU model including attention function have been used in this hybrid model to represent linear and nonlinear aspects, respectively. Bera S. [6] takes into account a real-world dataset in which the sales income of a restaurant is estimated. A second phase regression the model developed on the traditional regression analysis of linear regression, ridge regression, and decision tree regressor.

Tiwari Ret [7] explains an innovative implementation of Diagonal Re-current Neural Network (DRNN) to the challenge of predictive analytics. The suggested framework has temporal features because it is constituted of feedback adjustable links that really provide information. Makkar S [8] investigates the approaches used mostly by retail chain stores for a range of items at different store sites by inter-acting with a dataset from a retail chain of stores.

"By analysing Retail sales, we will learn that for what purpose those characteristics of a product perform as well as how they impact its selling." To assist Big-Mart in achieving this aim, a forecasting model will be developed to determine the main drivers that can maximize revenue in each store and also what improvements should be introduced to the item or store's functionality.

Long Short-Term Memory (LSTM) is artificial neural networks that are programmed to identify patterns in data sequences such as numerical time series data. For instruction, Long Short-Term Memory (LSTM) employs the back-propagation algorithm. LSTMs will recall crucial facts from the feedback they got because of their intrinsic memories, allowing them to be very accurate about anticipating what will happen next.

III. METHODOLOGY

The objective of the whole analysis is to examine the estimation of big mart store sales by using a machine learning methodology to evaluate the sales in the next month. For forecasting product sales, this approach combines loss convergence of several Convolutional Neural Network (CNN) Layers to long short-term memory (LSTM). For instance, loss convergence of several Convolutional Neural Network (CNN) layers will remove the inter time-series data and merge it into a square forecast error pattern curve describing each network and a product sales decision based on a mathematical model. The pattern in this is then predicted by the long short-term memory layer. Then, for even more precise estimation, the long short-term memory layer forecasts the pattern in the square estimated error rate gradient and sets several baseline model responses based on various system patterns.

Long short-term memory (LSTM) is a form of Recurrent Neural Net-work (RNN) with greater layer storage, making it ideally adapted for longer-term projections. The architecture is made up of cells in which information is collected, analysed, and afterward transferred towards the next cell. Each cell includes various gates which decide whether the data is processed, erased, or transferred to the next cell. The LSTM is a forerunner to standard neural networks, While Convolutional neural networks has their advantages, they suffer from either a major problem in that they are unable to retain data from previous time-steps, resulting in the loss of potential association information. The RNN solved this problem because it has a structural component known as the "secret state."

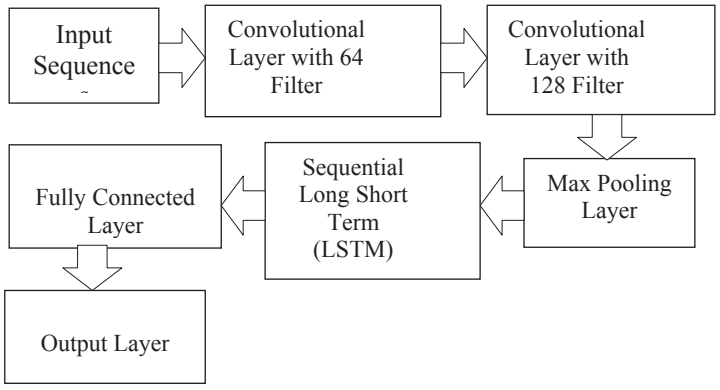

Fig.1. Architecture of CNN-LSTM framework for Product Sales predictions

LSTM (Long Short Term Memory):

The Ordinary LSTM cannot be used explicitly on sequences of spatial feedback. So, in order to accomplish tasks that require selected features to forecast things, we could have a more comprehensive structure. This is where the CNN-LSTM concept comes into play. CNN Long Short-Term Memory Network (CNN-LSTM) is an LSTM framework built primarily for sequence prediction tasks with sequential input. To overcome the drawback of traditional sequence network, 3 gates are added into the cell of the network to facilitate the notion of memory. A memory is kept and updated when the cell reads inputs at every period. LSTMs with four gates: forget (f), input (i), memory (c) and output gate (o).

Given an old memory Ct−1, the new cell memory Ct is computed as:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

Forget Gate: decides which information is to be eliminated from the current memory

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

Memory Gate: generates new candidate memory.

$$\tilde{C}_t = \tanh\,(W_c x_t + U_c h_{t-1} + b_c)$$

Input Gate: This gate determines how much information of the candidate memory will be injected into the updated one.

$$i_t = \sigma\,(W_i x_t + U_i h_{t-1} + b_i)$$

Output Gate: determines how much of the cell memory is extracted out.

$$o_t = \sigma\,(W_o x_t + U_o h_{t-1} + b_o)$$

LSTM Training:

To solve the disadvantage of regular RNNs, three gates are added to the network's cell to allow the concept of memory.

1. Whenever the unit accepts impulses for each step size and information is stored and modified.
2. LSTMs comprised of four gates: forget (f), the input I, memory (c), and out-put (o) (o).
3. Assuming an old memory Ct1, compute the new cellular memory Ct as follows:

Forget Gate: The forget gate determines which data is to be removed from the present memory.

Memory Gate: This function creates fresh candidate memory.

Input Gate: This gate controls how much data from the candidate memory is fed into the update memory.

Output Gate: The outer loop controls how many of the cell memory is retrieved.

Convolutional layer: A filter or kernel matrices is being used to move across subsets of the provided inputs. The outcome of sliding is known as a feature map. This operation objective is to study a specific interpretation out from current data.

Pooling layer: The pooling layer is responsible for extracting the most prominent (in the instance of max pooling) or combining all the data (average pooling).

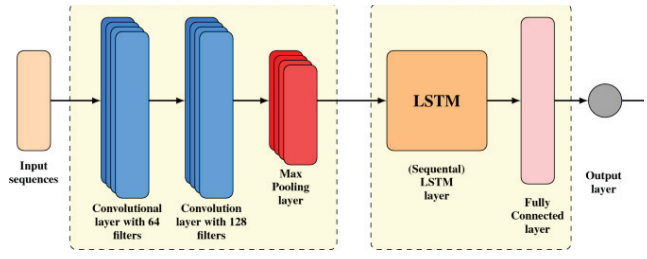

Fig. 2. Hybrid Architecture of CNN-LSTM

We have to import certain Keras modules in an attempt to develop the LSTM.

1. Initialization sequence for the neural network.
2. Dense to allow for the addition of a densely linked neural network.
3. LSTM for the addition of a Long Short-Term Memory layer.
4. Dropout allows us to add dropout layers to avoid over fitting.

IV. Experiment and Results

Convolutional Neural Network-Long Short Term Memory (CNN-LSTM) is capable of capturing much broader constraints than LSTMs. From such a time series forecasting standpoint, this could refer to frameworks that can detect trends over a longer time scales. LSTMs and other recurrent network architectures excel at discovering connections between persistent pieces of data over different time frames. To predict the future sales we have selected some parameters. Such as Shop ID, Product ID, Offer. These parameters will summarize the prediction of product sales. Below Table 1 represents the dataset with respect to their experimental results.

Dataset:

We have collected the DATASET from Big-mart store. Below Table 1 represents some instances from DATASET. Here, we have taken 7 major classes and every class has their own subclass. We have 1000 instances for training. We divide it into twofold i.e., one for training and another for testing. Training fold contains 700 samples and the testing fold contains 300 Samples.

TABLE I. Big-Mart Dataset

Labels	Material Description	Sales Quality	Sales Value	%Off	Category
1268362	Cinthol Cool Talc 300g	7	878.85	2	Health And Beauty
1312292	Whisper Ultra Clean XL + Wings San Pad 44	369	115133	5	Health And Beauty
1326057	Sofy Bodyfit Antibacteria XI San Pad 54	240	65270.5	3	Health And Beauty
1304119	Bournvita Refall 750g Pouch Pack	351	90793	10	Beverages
1072928	Society Leaf Tea Pouch 1 KG	198	86190	3	Beverages
1152004	Red Label Natural Care Box 500G	183	50875	5	Beverages
1348407	Surf Excel Matic Liquid Pouch Fl 2 Ltr	147	57310	5	Cleaning Aid
1334219	Ariel Matic Front Load 4kg AND 2kg Free	62	55880	2	Cleaning Aid
132724	Harpic Orginal 1LX 2 And 500	164	55104	5	Cleaning Aid

	ML Free				
1263925	Amul Taaza Toned Milk Tetrapack 1Ltr	870	53940	10	Dairy & Frozen
1021052	Amul Masti Buttermilk Tetrapack 1 Ltr	60	11270	5	Dairy & Frozen
1004442	Nescafe Classic Jar 50G	7	8385.95	0	Dairy & Frozen
1385612	Unibic Fruitn Nut Cookied 500g	82	6988.27	3	Ready Food
1386860	Parle Monaco Pizza 300gm	228	6968.81	5	Ready Food
1386176	Lal Mysore Pak 160gm BOGO West	91	6953.78	1	Ready Food
1358083	Value_KC Almond 1kg Pp	525	306865	12	Staples
1388200	Sg Cashew 500g	689	240461	10	Staples
1356154	Value_KC Almond 1kg Pp	785	234715	10	Staples
1004403	Maggi Noodles Masala 560g	1130	85875.23	10	Instant Food
1333672	Kelloggs Extra Muesli Fruit & Nut 750g	153	49600	3	Instant Food
1018208	Maggi Nutri-Licious Atta Mas 24x290g	437	34224.6	5	Instant Food

TABLE II. COMPARISON OF PROPOSED METHOD WITH WELL-KNOWN LSTM MODEL

Method	**Health And Beauty**	**Beverages**	**Cleaning Aid**	**Dairy & Frozen**	**Ready Food**	**Staples**	**Instant Food**
Actual	512	231	128	763	537	23	129
CNN	414	157	152	849	487	13	102
Proposed	509	229	124	743	531	19	118
LSTM	412	128	131	983	452	11	89

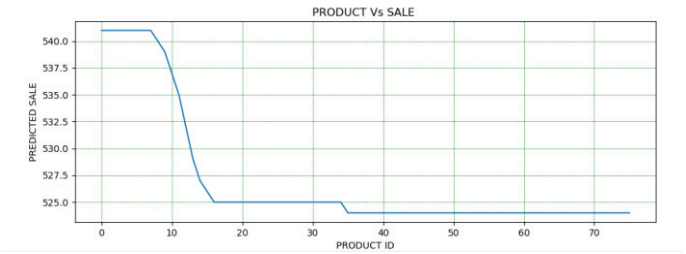

Fig. 3. Prediction results of Health And Beauty Class

Above figure shows the Prediction results of Health And Beauty Class for first shop with offering 10% discount. Here X-axis indicate product ID and Y-axis indicate its sale.

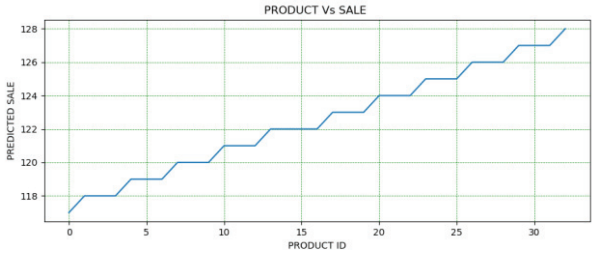

Fig. 4. Prediction results of Health And Beauty Class

Above figure shows the Prediction results of Health And Beauty Class for first shop with offering 2% discount. Here X-axis indicate product ID and Y-axis indicate its sale. If we decrease discount from 10% to 2% then sale also reduces.

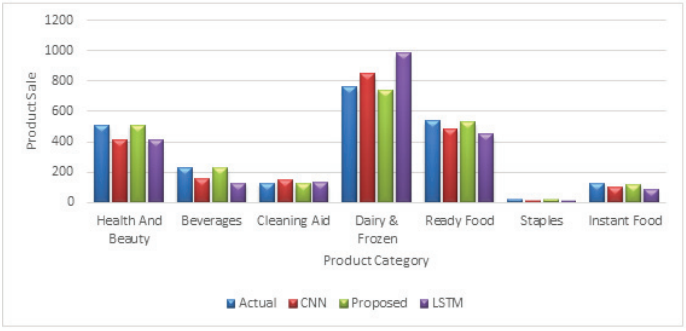

Fig. 5. Prediction results of Proposed and traditional Method

Proposed method outperform than the traditional LSTM model. Here, the actual sold items are indicated by blue colour, where the CNN result is represented by red colour. Finally, the proposed method is indicated by green colour, and purple colour represents the LSTM results.

V. CONCLUSION

The objective of the whole analysis is to examine the estimation of big mart store sales using a machine learning methodology to evaluate sales in the next month. Throughout this study, we employed some of its most exact predicting technologies employing a Long Short-Term Memory module, that assists users, experts, or any-body involved in selling goods by presenting them with a clear understanding of the future sales. According to the Predictive approach used, the estimates obtained indicated that the number generated does reflect the actual data where the maximum degree of precision was 97 percent and the minimum degree of precision was only beyond 82 percent.

REFERENCES

[1] S S. Cheriyan, S. Ibrahim, S. Mohanan and S. Treesa, "Intelligent Sales Prediction Using Machine Learning Techniques," 2018 International Conference on Computing, Electronics & Communications Engineering (iCCECE), 2018, pp. 53-58

[2] Pan, H., Zhou, H. Study on convolutional neural network and its application in data mining and sales forecasting for E-commerce.

[3] Yuan, H., Xu, W., Li, Q. et al. Topic sentiment mining for sales performance prediction in e-commerce. Ann Oper Res 270, 553–576 (2018)

[4] B. Singh, P. Kumar, N. Sharma and K. P. Sharma, "Sales Forecast for Amazon Sales with Time Series Modeling," 2020 First International Conference on Power, Control and Com-puting Technologies (ICPC2T), 2020, pp. 38-43

[5] Y. j. Li, Y. Yang, K. Zhu and J. Zhang, "Clothing Sale Forecasting by a Composite GRU-Prophet Model With an Attention Mechanism," in IEEE Transactions on Industrial Infor-matics

[6] Bera S. (2021) An Application of Operational Analytics: For Predicting Sales Revenue of Restaurant. In: Das S., Das S., Dey N., Hassanien AE. (eds) Machine Learning Algorithms for Industrial Applications. Studies in Computational Intelligence, vol 907. Springer, Cham. LNCS Homepage, http://www.springer.com/lncs, last accessed 2016/11/21.

[7] Tiwari R., Kumar R., Srivastava S., Gera R. (2021) Diagonal Recurrent Neural Network Based Prediction Model for Sales Forecasting. In: Malik H., Iqbal A., Joshi P., Agrawal S., Bakhsh F.I. (eds) Metaheuristic and Evolutionary Computation: Algorithms and Applica-tions. Studies in Computational Intelligence, vol 916. Springer, Singapore.

[8] Makkar S., Sethi A., Jain S. (2021) Predictive Analytics for Retail Store Chain. In: Gupta D., Khanna A., Bhattacharyya S., Hassanien A.E., Anand S., Jaiswal A. (eds) International Conference on Innovative Computing and Communications. Advances in Intelligent Sys-tems and Computing, vol 1166. Springer, Singapore.

[9] V. Shrivastava and N. Arya, "A study of various clustering algorithms on retail sales data", Int. J. Comput. Commun. Netw, vol. 1, no. 2, 2012.

[10] F. Sigrist and C. Himschall, Gradient Tree-Boosted Tobit Models for Default Prediction, 2018.

[11] Xu-Ying Liu, Jianxin Wu and Zhi-Hua Zhou, "Exploratory Undersampling for Class-Imbalance Learning", IEEE Transactions on Systems, Man and Cybernetics, Vol. 39(2), pp. 539-550, April, 2009.

[12] Deroski, Saso and Bernard Enko. "Is Combining Classifiers with Stacking Better Than Se-lecting the Best One?", Machine learning, Vol. 54(3), pp. 255- 273, March, 2004

[13] Domingos Pedro, "A Few Useful Things to Know About Machine Learning", Communi-cations of the ACM, Vol. 55, pp. 78-87, October, 2012.

[14] Communications of the ACM, Vol. 55, pp. 78-87, October 2012.

[15] Bose, Indranil, and Radha K. Mahapatra, ―Business Data Mining - A Machine

[16] Learning Perspective‖, Information & Management, Vol. 39(3), pp. 211-225,February 2001.

[17] Pat Langley and Herbert A. Simon, ―Applications of Machine Learning and Rule Induction‖, Communications of the ACM, Vol. 38(11), pp. 54-64, November 1995.

[18] Augusto Ribeiro, Isabel Seruca, and Natrcia Duro, ―Improving Organizational Decision Support: Detection of Outliers and Sales Prediction for a Pharmaceutical Distribution Company‖, Procedia Computer Science, Vol. 121, pp. 282-290, December 2017.

[19] Akshay Krishna, Akhilesh V, Animikh Aich, Chetana Hegde, ―Sales forecastingof Retail Stores using Machine Learning Techniques‖, 3rd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2018

[20] Gopalakrishnan T ―Prediction of Sales Value in online shopping using Linear Regression‖, 2018 4th International Conference on Computing Communication and Automation (ICCCA)

[21] Samaneh Beheshti-Kashia,b, Hamid Reza Karimic, Klaus-Dieter

[22] Thobenb, Michael Lu¨tjenb, and Michael Teuckeb, ―A survey on retail sales forecasting and prediction in fashion markets‖, Systems Science Control Engineering: An Open Access Journal, 2015 Vol. 3, 154–161

[23] Sandeep Chavan, ―Predicting Online Product Sales using Machine Learning‖, International Journal of Engineering Research Technology (IJERT) ISSN: 2278-0181 Vol. 9 Issue 04, April-2020

[24] Erkin Guler, 2 Taner Ersoz and 1 Filiz Ersoz, ―Applying Data Mining Technique to Sales Forecast‖,

[25] Norulhidayah Isa1 , Nur Syuhada Mohd Yusof2 , Muhammad Atif Ramlan3-31,

[26] The Implementation of Data Mining Techniques for Sales Analysis using Daily Sales Data ‖, International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.1.5, 2019

[27] Heramb Kadam, 2Rahul Shevade, 3 Prof. Deven Ketkar, 4Mr. Sufiyan Rajguru, ―A Forecast for Big Mart Sales Based on Random Forests and Multiple Linear Regression‖, IJEDR 2018 − Volume 6, Issue 4— ISSN: 2321-9939

[28] Inedi. Theresa, 2Dr.Venkata Reddy Medikonda, 3K.V. Narasimha Reddy, ‖ Prediction of Big Mart Sales Using Exploratory Machine Learning Techniques‖, International Journal of Advanced Science and Technology Vol.29, No. 6, (2020), pp. 2906 – 2911

[29] Bohdan M. Pavlyshenko, ―Machine-Learning Models for Sales Time Series Forecasting‖ A. Sayli, I. Ozturk, and M. Ustunel, "Brand loyalty analysis system using KMeans algorithm", Journal of Engineering Technology and Applied Sciences, vol. 1, no. 3, 2016.

[30] M.N. Maingi, A Survey on the Clustering Algorithms in Sales Data Mining.

[31] V. Shrivastava and N. Arya, "A study of various clustering algorithms on retail sales data", Int. J. Comput. Commun. Netw, vol. 1, no. 2, 2012.

[32] F. Sigrist and C. Himschall, Gradient Tree-Boosted Tobit Models for Default Prediction, 2018.

[33] Xu-Ying Liu, Jianxin Wu, and Zhi-Hua Zhou, ―Exploratory Undersampling for Class-Imbalance Learning‖, IEEE Transactions on Systems, Man and Cybernetics, Vol. 39(2), pp. 539-550, April 2009