

Prediction model based on NLP and NN for financial data outcome revelation

Hiral R. Patel^{1*}, Satyen M. Parikh² and Ajay M. Patel³

¹DCS, Ganpat University, Ganpat Vidyanagar-384012, Gujarat, India

²FCA, Ganpat University, Ganpat Vidyanagar - 384012, Gujarat, India

³AMPICS, Ganpat University, Ganpat Vidyanagar - 384012, Gujarat, India
hrp02@ganpatuniversity.ac.in

Available online at: www.isca.in

Received 30th April 2017, revised 16th July 2017, accepted 20th July 2017

Abstract

The financial market is too vigorous in nature. As per the growth of world, financial market became the most valuable investment component for the people. The common mentality for the investors is to maximise the capital amount in very short time so to maximise the profit is ultimately to decide the investment strategy for increase the ROI. This approach requires a lot of analytical work to meet the investment objective. It also requires the portfolio rebalancing or switching strategy to optimise the revenue. The proposed model based on predictive approach to recommend the future value of selective stock from given data and the model performs the analytical task based on financial news and technical financial data. For analysing the news impact proposed approach deal with TF-TDF text mining technique and semantic analysis of Natural Language Processing to predict the impact of news. It also used technical data and to forecast future value the historical data of stock is necessity. So the modified neural network based on back propagation methodology used as forecasting machine learning methodology in presented model to predict the future value.

Keywords: Back Propagation, Financial Data, Forecasting, Neural Network, NLP, ROI, Semantic Analysis.

Introduction

The financial market is too vigorous in nature. As per the growth of world, financial market became the most valuable investment component for the people. The common mentality for the investors is to maximise the capital amount in very short time so to maximise the profit is ultimately to decide the investment strategy for increase the ROI. This approach requires a lot of analytical work to meet the investment objective¹. It also requires the portfolio rebalancing or switching strategy to optimise the revenue².

The proposed model based on predictive approach to recommend the future value of selective stock from given data and the model performs the analytical task based on financial news and technical financial data³. For analysing the news impact proposed approach deal with TF-TDF text mining technique and semantic analysis of Natural Language Processing to predict the impact of news⁴. It also used technical data and to forecast future value the historical data of stock is necessity. So the modified neural network based on back propagation methodology used as forecasting machine learning methodology in presented model to predict the future value⁵.

Methodology

Data Collection: The proposed model uses news data and financial data history. So the implemented approach have

created the agents to collect the data from various authenticate sources of news and financial data³.

Data Pre-processing: As per the learning methodology conventions, before applying the any technique, data should be statistically proved for implemented model⁴.

Analytical Study: The analytical study approach focus on quantitative and qualitative study. For the qualitative analysis, to calculate the news impact TF-IDF with semantic analysis techniques is used.

For quantitative analysis, using news impact consideration modified neural network with back propagation method developed to forecast the future value. This methodology used for classification as well as predictive analytical study⁴.

Result Data Selection: The model is generalised approach to get the future product value. To justify the results the specific product selection is done for representing the work⁶.

Each sector has positive and negative affected stocks so to predict the value based on the news that may positive or negative impacted.

Here for outcome revelation, selected stock with different tools with techniques and proposed model performance compared in the terms of scalability, efficiency, throughput and accuracy⁷.

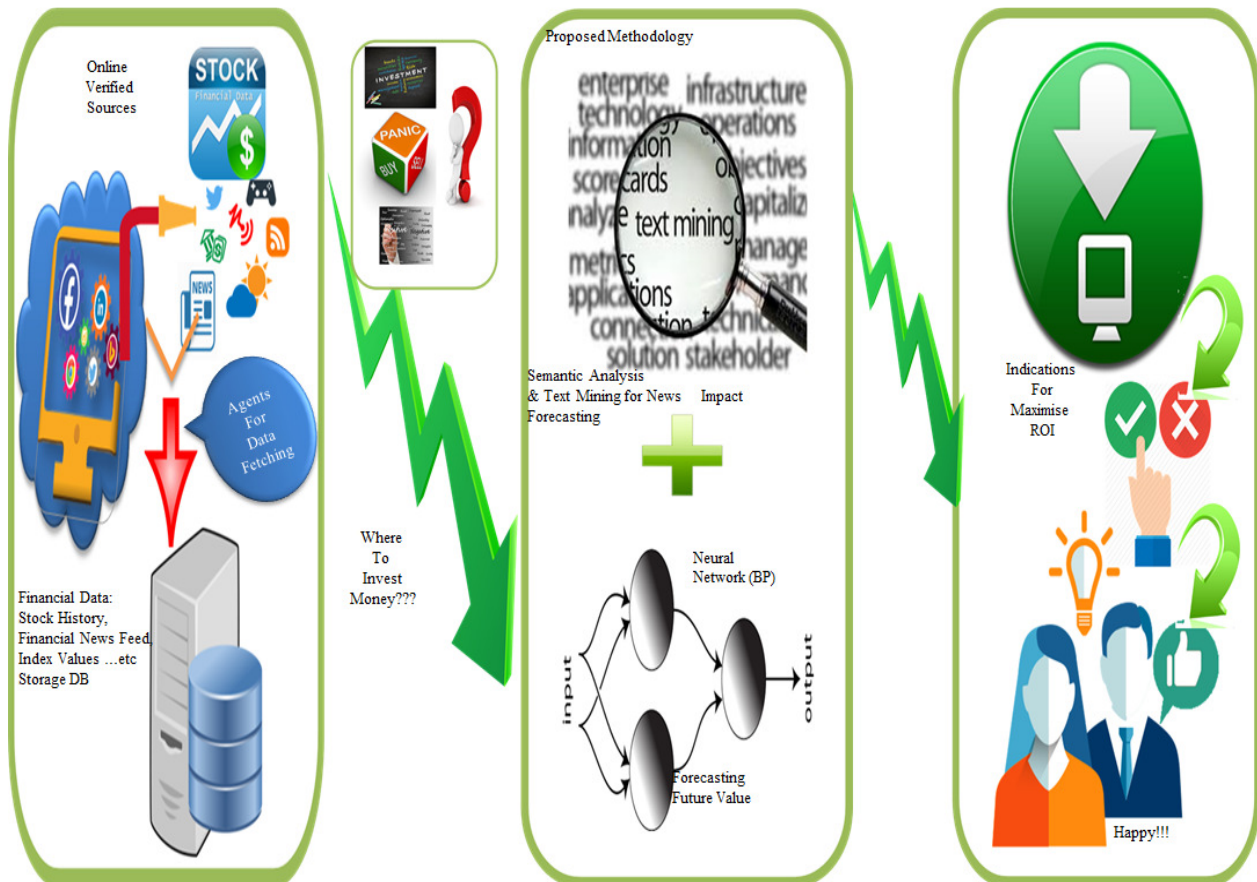


Figure-1: Proposed Model for Prediction of Stock.

Results and Discussion

The financial market stocks are categorised using market capital value in different categories. i. Small CAP Stock- The market capital below 2 billion dollar stocks are under this category. So from BSE small cap stock list, given result discussion consider three variant stocks based on Oil sector are consider like Jindal Drilling and Inds. Ltd., and ONGC scripts^{8,9}. ii. Mid CAP Stock- The market capital between 2 billion dollar to 10 billion dollar stocks are under this category. So from BSE small cap stock list, given result discussion consider three variant stocks based on Oil sector are consider like GAIL, HPCL and Petronet LNG scripts^{8,10}. iii. Large CAP Stock- The market capital above 10 billion dollar stocks are under this category. So from BSE small cap stock list, given result discussion consider three variant stocks based on Oil sector are consider like Aban Offshore Ltd., The Chennai Petroleum Corp., Indraprasth Gas, and Indian Oil scripts⁸.

Outcome Revelation: The neural network is specially used to work with non-linear and dynamic nature base data¹¹⁻¹³. The developed model results also compared with existing tools like Aluda Neuro intelligence, XLMiner, Sipina Research, Tanagra and Weka which supporting neural network different

techniques. According to the outcomes given by the different tools, this model implement the Neural Network with back propagation method is implemented using C sharp language^{8,12}. The Table-1 shows the performance of different tool and the implemented methodology.

Conclusion

This experimental study performs comparative analytical study with developed approach. As per the study, techniques are playing effective role in different criteria. No doubt in past developed approach, regression was best out fitting model for predict the trend as per time series being effect, Statistical approach is best for linearity and continuous being effect and the machine learning based neural network performs best for most dynamism and non – linear based effect so the classification based modelling techniques are used as best forecaster to deal with random behaviour. Among mentioned classifier predictor, neural network used as best trainer for non-linear data for short term prediction to maximum coverage of dynamism in fluctuation. So the above table shows the comparative results and concludes that modified implemented approach gives better result in terms of accuracy, scalability, efficiency and through put.

Table-1: Outcome Revelation of Proposed Model.

Script Type	Script Name	Applied Tool	Method	Measure	Through Put [MS]
Small Cap	Aban Offshore Ltd.	Aluda Neuro.	Gradient Descent	R2=0.99	914
			Levenberg	R2=1	902
			Back Propagation	R2=0.87	1000
		XLMiner	NN	Accuracy=69%	141
		Sipina	Single Layer Perceptron	60%	550
			Multilayer Perceptron	61%	555
		Tanagra	Multilayer Perceptron	68%	281
		Weka	10 Fold MLP	62%	110
			20 Fold MLP	60%	150
	Proposed	Back Propagation	78%	120	
	Jindal Drilling and Inds. Ltd.	Aluda Neuro.	Gradient Descent	R2=0.99	843
			Levenberg	R2=0.98	740
			Back Propagation	R2=0.86	838
		XLMiner	NN	Accuracy=70%	135
		Sipina	Single Layer Perceptron	61%	660
			Multilayer Perceptron	63%	575
		Tanagra	Multilayer Perceptron	68%	281
		Weka	10 Fold MLP	73%	110
			20 Fold MLP	70%	160
	Proposed	Back Propagation	77%	105	
	ONGC	Aluda Neuro.	Gradient Descent	R2=0.99	905
			Levenberg	R2=0.89	860
			Back Propagation	R2=0.88	981
		XLMiner	NN	Accuracy=70%	195
		Sipina	Single Layer Perceptron	60%	560
			Multilayer Perceptron	63%	575
		Tanagra	Multilayer Perceptron	69%	301
Weka		10 Fold MLP	70%	490	
		20 Fold MLP	72%	510	
Proposed	Back Propagation	77%	127		
Mid Cap	Chennai Petroleum Corpn	Aluda Neuro.	Gradient Descent	R2=0.99	835
			Levenberg	R2=1	678
			Back Propagation	R2=0.76	761
		XLMiner	NN	Accuracy=67%	151
		Sipina	Single Layer Perceptron	61%	555
			Multilayer Perceptron	63%	575
		Tanagra	Multilayer Perceptron	69%	271
		Weka	10 Fold MLP	64%	124
			20 Fold MLP	61%	159
Proposed	Back Propagation	80%	140		

Script Type	Script Name	Applied Tool	Method	Measure	Through Put [MS]
	Indraprasth Gas	Aluda Neuro.	Gradient Descent	R2=0.98	930
			Levenberg	R2=1	899
			Back Propagation	R2=0.8	999
		XLMiner	NN	Accuracy=67%	151
		Sipina	Single Layer Perceptron	62%	540
			Multilayer Perceptron	63%	565
		Tanagra	Multilayer Perceptron	69%	271
		Weka	10 Fold MLP	70%	130
			20 Fold MLP	69%	150
	Proposed	Back Propagation	80%	140	
	Indian Oil	Aluda Neuro.	Gradient Descent	R2=0.99	501
			Levenberg	R2=1	501
			Back Propagation	R2=0.8	501
		XLMiner	NN	Accuracy=69%	141
		Sipina	Single Layer Perceptron	60%	550
			Multilayer Perceptron	61%	555
		Tanagra	Multilayer Perceptron	68%	281
		Weka	10 Fold MLP	62%	110
			20 Fold MLP	60%	150
	Proposed	Back Propagation	79%	140	
	Large Cap	GAIL	Aluda Neuro.	Gradient Descent	R2=0.98
Levenberg				R2=1	871
Back Propagation				R2=0.76	881
XLMiner			NN	Accuracy=68%	144
Sipina			Single Layer Perceptron	61%	553
			Multilayer Perceptron	65%	565
Tanagra			Multilayer Perceptron	67%	291
Weka			10 Fold MLP	72%	111
			20 Fold MLP	70%	121
Proposed		Back Propagation	79%	142	
HPCL		Aluda Neuro.	Gradient Descent	R2=0.99	991
			Levenberg	R2=1	881
			Back Propagation	R2=0.78	777
		XLMiner	NN	Accuracy=68%	144
		Sipina	Single Layer Perceptron	62%	530
			Multilayer Perceptron	66%	565
		Tanagra	Multilayer Perceptron	69%	284
		Weka	10 Fold MLP	63%	120
			20 Fold MLP	61%	140
Proposed		Back Propagation	80%	135	
Petronet LNG		Aluda Neuro.	Gradient Descent	R2=0.99	888

Script Type	Script Name	Applied Tool	Method	Measure	Through Put [MS]
			Levenberg	R2=1	876
			Back Propagation	R2=0.8	897
		XLMiner	NN	Accuracy=70%	143
		Sipina	Single Layer Perceptron	61%	570
			Multilayer Perceptron	66%	585
		Tanagra	Multilayer Perceptron	69%	288
		Weka	10 Fold MLP	67%	130
			20 Fold MLP	66%	150
		Proposed	Back Propagation	79%	140

References

- Hiral R.P. and Satyen P. (2012). Automated News based ULIP Fund Switching Model. International Conference GCEMP 2012, Awarded as Best Technical Paper, GFJMR ISSN 2229-4651, 3.
- Hiral R.P., Suthar A.B. and Satyen P. (2012). A Comparative Study on Financial Stock Market Prediction Models. *The International Journal of Engineering And Science*, (IJES), 1(2), 188-191. ISSN: 2319 – 1813 ISBN: 2319 – 1805 , Indexed in ANED (American National Engineering Database).
- Hiral R.P., Suthar A.B. and Satyen P. (2014). A Proposed prediction model for forecasting the financial market value according different factors. *International Journal of Computer Technology & Application (IJCTA)* ISSN 2229-6093, 5(1), 131. Impact Factor – 2.015 IC Value 5.17.
- Hiral R.P. and Satyen P. (2014). A Technical and Fundamental Parameters analysis for Financial Market Prediction Using Semantic Analysis.
- Hiral R.P. and Satyen P. (2015). Dynamic IS based Asset Allocation on Crude Trend Analysis – exploring a Hedging Concept. GCEMP-15, GFJMR ISSN 2229-4651, 6.
- Patel H.R. and Satyen P. (2016). Comparative analysis of different statistical and neural network based forecasting tools for prediction of stock data. Proceedings of the Second International Conference on Information and Communication Technology for Competitive Strategies, ACM, 126. <http://dl.acm.org/citation.cfm?id=2905055>. 2905186.
- Patel H.R. and Satyen P. (2016). Comparative Analytical Study for News Text Classification Techniques Applied for Stock Market Price Extrapolation. International Conference on Smart Trends for Information Technology and Computer Communications, 239-243. <http://www.springer.com/in/book/9789811034329>.
- Hiral R.P., Darji D.N. and Satyen P. (2016). Prediction Model for Stock Market using News based different Classification, Regression and Statistical Techniques (PMSMN). ICT in Business Industry & Government (ICTBIG), International Conference on, IEEE Publication, 1-5. <http://ieeexplore.ieee.org/document/7892636/>.
- Satyen P. (2009). Deriving Business Intelligence Through Collaborative Cooperative Multi Agent Model For Mutual Fund Asset Allocation. Proceeding of the IEEE national Conference on Computing and Communication Systems CoCoSys-09 at University Institute of Technology, Burdwan 02-04 Jan 2009.
- Satyen P. (2008). A Comparative Study of Data Mining Techniques and Selection Issue. National Conference (Innovative Dimension for Business and Information Technology 23-24 Feb-2008). ISBN: 978-81- 906446-0-0.
- Satyen P. (2008). Data Mining Supported Integrated Intelligent Advisory Model (IIAM) For Financial Growth. International Conference (ICETAETS) on Emerging Technology and Applications in Engineering, Technology and Science (13-14 Jan-2008) 978-81-906220-1-1
- Satyen P. (2007). Role of Multi-Agent System in Real Time Business Intelligence. National Conference ENVISION-2007.
- Parikh S.M. (2008). Analysis And Modeling A Distributed Co-Operative Multi Agent System For Scaling-Up Business Intelligence. Saurashtra University.