The Use of Big Data Analytics to Predict the Foreign Exchange Rate Based on Public Media

Abstract

- In recent years, people have become more reliant on new media (websites, forums, and social media) to access and share information. Most social media information is considered public information.
- For many people, such extensive public information can help with making decisions— for instance, participants in the exchange rate market can use this information to their best interests at little cost.
- However, the variety and volume of data provided by new media, along with the speed at which such data appears, are continuously increasing. The veracity of new media data is also an issue. Furthermore, forecasting the daily movement of exchange rates has been such a difficult task that most econometric models are incapable of forecasting short term exchange rates with outstandingly higher accuracy than a naive random walk model.1

Existing system

- it is challenging to claim that big data analytics (BDA) can exceed the random walk mechanism and the notion of market efficiency in predicting the daily movement of the USD/TWD exchange rate.
- To address this challenge, we propose a BDA mechanism that incorporates machine-learning modeling to provide an effective way of dealing with more complex relationships.2

Disadvantages

- exchange rates from the Taiwan Futures Exchange, we set up seven types of public information by permuting the sources of news websites, forums, and social media.
- The selected financial news websites are cnYes and moneyDJ, the forum is ForeignEX of PTT, and the social media platform is Facebook. We apply the text mining technique3 to the obtained textual information about the exchange rate to derive the corresponding numerical representations.
- We then apply machine-learning modeling to learn the relationship between the derived numerical representations and the daily movement of exchange rates.

Proposed system

- Exchange rates can demonstrate the phenomenon of concept drifting and outliers. The term “concept drifting” means the concepts are not stable and change with time.
- That is, as time passes, the trend embedded in the observation data usually changes. Outliers are the observations far away from the fitting function deduced from a subset of given observations.
- Outlier detection becomes more challenging when the data are subject to concept drifting. Shin-Ying Huang and her colleagues addressed this challenge by proposing the resistant learning with envelope module (RLEM) and moving window technique.
- Therefore, we adopt the RLEM and moving window technique in Stage 1 of the proposed machine-learning model to filter out potential outliers. As the desired outputs are binary, Stage 2 adopts the reasoning neural network7 (RNN) to conduct the training. To speed up the training, the machine-learning model is implemented using TensorFlow and GPU.

Advantages

- adopt the RLEM and moving window technique in Stage 1 of the proposed machine-learning model to filter out potential outliers. As the desired outputs are binary, Stage 2 adopts the reasoning neural network7 (RNN) to conduct the training.
- To speed up the training, the machine-learning model is implemented using Tensor Flow and GPU

Hardware Requirements

- Processor :Intel Pentium IV 1GHz
- RAM :256MB (Min)
- Hard Drive :5GB free space
- Monitor :1024 * 768, High Color inch
- Mouse :Scroll Mouse(Logitech)
- Keyboard :104 keys

Software requirements

- OS : Windows XP/7/8
- Front End : Visual Studio 2010/ netbeans 7.1
- Back End : SQL Server 2005/ heidisql 3.2
- Browser : Any Web Browser

conclusion

- Our research differs from previous research in its emphasis on the predictive powers of important qualitative variables collected from social media.
- Alternatively, the conventional forecasting methods make use of available market data, including aggregate macroeconomic time series data and trading volume or price data.
- Empirical findings based on quantitative data have not found consensus. In other words, little can be said about which conventional methods have dominated forecasting performance.
- Due to the lack of a fair benchmark, subjective judgments complicate a comparison of our forecasting results with others. Even without such a comparison,our method should complement the conventional forecasting methods.

References

1. R.A. Meese and D. Rogoff, “Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?,” *J. International Economics, vol. 14, no. 1-2, 1983, pp. 3*24.
2. D. Zhang, S. Simoff, and J. Debenham, “Exchange Rate Modelling Using News Articles and Economic Data,” *Australasian Joint Conf. on Artificial Intelligence (AI 2005: Advances in Artificial Intelligence), 2005, pp. 467–476.*
3. R. Feldman and J. Sanger, *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data, Cambridge Univ. Press, 2007.*
4. M.G. Masud et al., “Classification and Novel Class Detection in Concept-Drifting Data Streams Under Time Constraints,” *IEEE Trans. Knowledge and Data Engineering, vol. 23, no. 6, 2011, pp. 859–874.*
5. R.H. Tsaih and T.C. Cheng, “A Resistant Learning Procedure for Coping with Outliers,” *Annals of Mathematics and Artificial Intelligence, vol. 57, no. 2, 2009, pp.* 161–180.