

Recommending Stock Markets Using Web Media Content

Manoj V

Department of Computer Science and Engineering Velammal Engineering College
Chennai, India

Email: vmanoj1604@gmail.com

Revanth B

Department of Computer Science and Engineering Velammal Engineering College
Chennai, India

Email: revanthrey98@gmail.com

-----Abstract-----

Stock market volatility is influenced by information release, dissemination, and public acceptance. With the increasing volume and speed of social media, the effects of Web information on stock markets are becoming increasingly salient.

However, studies of the effects of Web media on stock markets lack both depth and breadth due to the challenges in automatically acquiring and analyzing massive amounts of relevant information. In this study, we systematically reviewed 29 research articles on quantifying the interplay between Web media and stock markets from the fields of Finance, Management Information Systems, and Computer Science. In particular, we first categorized the representative works in terms of media type and then summarized the core techniques for converting textual information into machine-friendly forms. Finally, we compared the analysis models used to capture the hidden relationships between Web media and stock movements.

Our goal is to clarify current cutting edge research and its possible future directions to fully understand the mechanisms of Web information percolation and its impact on stock markets from the perspectives of investors cognitive behaviors, corporate governance, and stock market regulation.

Keywords- **Web Media, Analysis Model, Stocks, Social Media.**

I. INTRODUCTION

Analyze the financial data for making prediction in stock markets by using big data analytics and recommend the user investment in various category stock markets based on web media.

IN traditional finance, the efficient market hypothesis states that a stock price is always driven by “unemotional” investors to equal the firm’s rational present value of expected future cash flows. Specifically, stock investors are constantly adjusting their beliefs on the potential market performances of stocks, although they typically disagree on the matter.

This disagreement among competing market participants leads to discrepancies between the actual price and the intrinsic value, causing a stock price to fluctuate around a stock’s intrinsic value, i.e., new information has intricate influences on asset prices. Although traditional finance and modern behavioral finance have different views on how information shapes stock movements, both believe that the volatility of the stock market comes from the release, dissemination and absorption of information.

The overall goal is to analyze the financial data for making prediction in stock markets by using big data analytics and recommend the user investment in various category stock markets based on web media.

II. PROPOSED SYSTEM

The research on media-aware stock movements began with financial reports and news articles. With the popularity of Web 2.0, new media sources, such as blogs, Tweets/ micro blogs, discussion boards, and social news, have emerged and played important roles in affecting stock markets. As a pilot study, found that the emotions of tweets affected stock trends for a brief period after the release of the tweets. In contrast to traditional news, social media allows users to express their opinions and feelings via comments, votes and so forth. Such user engagement efficiently enhances information dissemination and increases the value of the information. We are using Google API to connect to the internet and get the news articles to analyze the stock trends and recommended the correct stocks to the users.

III. RESEARCH METHODOLOGY

3.1. Big Data and Environment

Huge Collection of data is retrieved from open source datasets that are publicly available from major Application Providers like Money Control. Big Data Schemas were analyzed and a Working Rule of the Schema is determined. The CSV (Comma separated values) and TSV (Tab Separated Values) files are Stored in HDFS (Highly Distributed File System) and were read through Master and manipulated using Java API that itself developed by

us which is developer friendly, light weighted and easily modifiable.

3.2. Preprocessing and User From

A preprocessing is a backend job running in hadoop clusters and also called as long running jobs as it is scheduled to process bulk data so that the application would makes use of the results produced for updating. Dataset mapping process done in the preprocessing phase the entire sector both NSE and BSE companies will be mapped date wise, month wise and year wise data. The risk factor of each sector and company will be complete over the preprocessing time. The user stock market investment based on the user credential such as adhar number, Account details, annual income, working status, marital status, loan status, etc. all these fields will be available on form16.

3.3. Market trend:

Our application provides history of market data in NSE and BSE, so that user will get idea about market trend mutual funds. We have a various sector such Finance, Management Information Systems, and Computer Science User can view the history of all the sector by using our prediction mechanism among the sectors we have different companies user can view history of each companies in the basis of date wise, month wise and year wise.

3.4. Recommendation Based On Web Media:

Based on Web Media content such as News articles and social media and discussion board where there will be articles about the stocks which perform well and also about other stocks will be there where we use Google API to extract the information from the web and based on that we recommend the stocks to the users based on risk factors and taking into all considerations.

limitations presented by latency and throughput. Modern techniques, such as in-memory databases, which rely on main memory for a data store medium, are faster than disk-optimized database systems, but are still limited by today’s memory capability. In addition, in-memory database still lack a non-volatile storage medium to provide long-term persistent storage. To handle QF big data, SSD (solid state disks)-backed storage could be more efficient than HDD (hard disk drive)-backed storage.

Hadoop is a free, Java-based programming framework that supports the processing of large data sets in a Parallel and distributed computing environment. It makes Use of the commodity hardware Hadoop is Highly Scalable and Fault Tolerant. Hadoop runs in cluster and eliminates the use of a Super computer. Hadoop is the widely used big data processing engine with a simple master slave setup.

In the mutual fund industry, predictive analytics plays a key role in providing data-driven decisions for managing the resources under an NSE and BSE. The growth in the holdings by the investors decides the growth of net assets under management of an NSE and BSE, and is adversely affected by the offsets in redemptions.

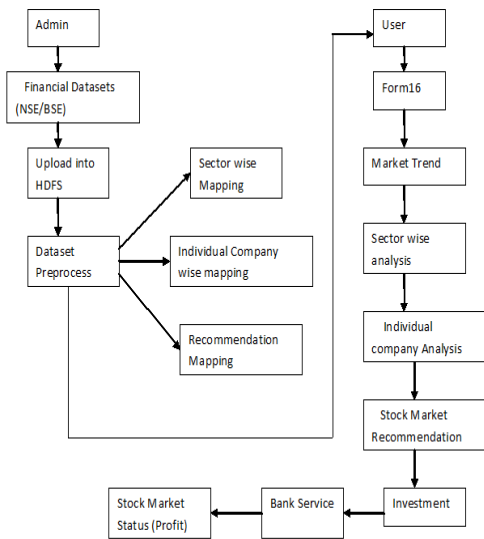
The attributes that trigger redemption by investors are complex in nature to identify and analyze. These attributes include financial transaction patterns by the investor, market conditions and sentiments, macroeconomics variables, scheme level features, and demographic factors. Predicting the redemption behavior requires sophisticated platform that can capture multiple factors that affect the redemption behavior.

However, big data predictive analytics using advanced analytics platform can analyze these massive amounts of transaction data and other time trend variables at a macro level. This platform can investigate these factors for near real-time data and can provide highly accurate predictions for the redeeming investors in the future at a investor-level.

This major process of the web application is

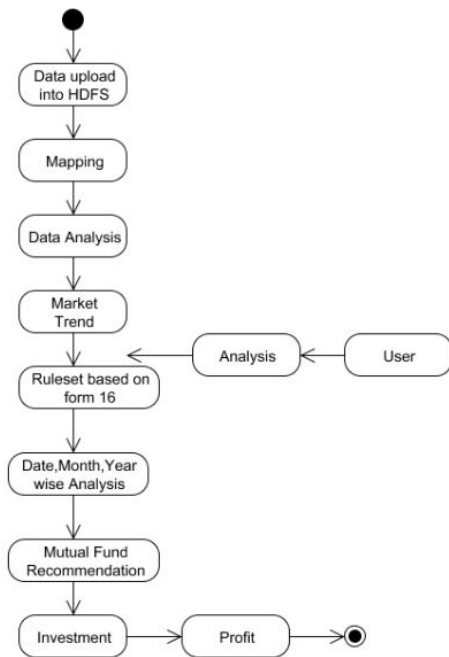
- Processing past histories of each individual company.
- User Form Processing to each about each individual user.
- Web Media Processing to know the present trend in the market.

IV. ARCHITECTURE



The performance of data processing systems based on system configuration, such CPU, memory, network and storage. Advancements in storage have lagged due to

V. ACTIVITY DIAGRAM



VI. ACKNOWLEDGMENT

The authors gratefully acknowledge the works of Qing Li, Yan Chen, Jun Wang, Yuanzhu Chen and Hsinchun Chen.

REFERENCES

- [1] The White House, "Making college affordable," <https://www.whitehouse.gov/issues/education/higher-education/making-college-affordable>, 2016.
- [2] Complete College America, "Four-year myth: Making college more affordable," <http://completecollege.org/wp-content/uploads/2014/11/4-Year-Myth.pdf>, 2014.
- [3] H. Cen, K. Koedinger, and B. Junker, "Learning factors analysis—a general method for cognitive model evaluation and improvement," in *International Conference on Intelligent Tutoring Systems*. Springer, 2006, pp. 164–175.
- [4] M. Feng, N. Heffernan, and K. Koedinger, "Addressing the assessment challenge with an online system that tutors as it assesses," *User Modeling and User-Adapted Interaction*, vol. 19, no. 3, pp. 243–266, 2009.
- [5] H.-F. Yu, H.-Y. Lo, H.-P. Hsieh, J.-K. Lou, T. G. McKenzie, J.-W. Chou, P.-H. Chung, C.-H. Ho, C.-F. Chang, Y.-H. Wei et al., "Feature engineering and classifier ensemble for kdd cup 2010," in *Proceedings of the KDD Cup 2010 Workshop*, 2010, pp. 1–16.
- [6] Z. A. Pardos and N. T. Heffernan, "Using hmms and bagged decision trees to leverage rich features of user and skill from an intelligent tutoring system dataset," *Journal of Machine Learning Research W & CP*, 2010.

- [7] Y. Meier, J. Xu, O. Atan, and M. van der Schaar, "Personalized grade prediction: A data mining approach," in *Data Mining (ICDM), 2015 IEEE International Conference on*. IEEE, 2015, pp. 907–912.
- [8] C. G. Brinton and M. Chiang, "Mooc performance prediction via clickstream data and social learning networks," in *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2015, pp. 2299–2307.
- [9] KDD Cup, "Educational data minding challenge," <https://pslcdatashop.web.cmu.edu/KDDCup/>, 2010.
- [10] Y. Jiang, R. S. Baker, L. Paquette, M. San Pedro, and N. T. Heffernan, "Learning, moment-by-moment and over the long term," in *International Conference on Artificial Intelligence in Education*. Springer, 2015, pp. 654–657.