Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 02, February -2018

ANALYZING FINANCIAL DATA AND MUTUAL FUNDS RECOMMENDATION BY USING BIG DATA ANALYTICS (PROPOSED ANALYSIS OF STOCK MARKET)

V Sangeetha[1].S Nalini[2], S Subiksha[3]

[1]*Student, Jeppiaar Maamallan Engineering College,*
[2]*Assistant Professor, Jeppiaar Maamallan Engineering College,*
[3]*Student, Jeppiaar Maamallan Engineering College,*

Abstract- *Big data analytics is the process of examining large and varied data sets -- In this paper, we analyze the financial data for making prediction in mutual funds by using big data analytics and recommend the user investment in various category mutual funds based on user details. Hadoop is a free , java-based programming framework which is widely used big data processing engine. It supports the processing of large data sets in a parallel and distributed computing environment. Our technique focused on all strategies such as computer hardware ,data processing ,prediction method, etc .we mainly consider the market dataset on National Stock Exchange(NSE) and Bombay Stock Exchange(BSE) based companies sectors such as Banking, Automobiles, Energy, Software and Moving Consumer Good(FMCG).Huge Collection of data is retrieved from open source datasets that are available from major Application Providers. The analysis would be completed before the user search any sector Mutual funds for investment, the process of financial data's Map Reduce function Grouping classified in data wise, month wise and year wise. We will get find out each and every companies Profit level in all the year within a seconds. Preprocessing is a backend job that entire sector is mapped. User's financial details are available on form 16 which is already filled by user in login phase, we calculate the risk factor for each company and recommend the investment. Our application going to guide the user Mutual Fund Investment which is based on user requirements .It is fully based on user credentials not a company profit over the year.*

Keywords: *Big data, parallel processing, Form 16*

I. INTRODUCTION

In the mutual fund industry, anticipating analytics plays a key role in keep 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 retrieval behavior requires sophisticated platform that can capture several factors that affect the retrieval 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 keep highly accurate predictions for the redemptive investors in the future at a investor-level.

II. RELATED WORKS

In this paper, the analysis patter appears on NYSE (New York Stock Exchange) and TAQ (Trade and Quota) mainly they focused in the market strategies and predict individual company profit by using Algorithmic trading technique[2]. It is the process of programming computers to place electronic trades according to predefined strategies. Algo-trading handles high-volume big data, and also high-velocity data processing [11][12]. In today's markets, a stock can experience 500 quote changes and 150 trades in one microsecond. Consequently, prices fluctuate by the milli-or even micro-second[9]. Placing a large number of orders at high speeds based on programmed strategies is crucial for profit generation. In particular, the rate of market data access and data-driven strategy processing closely determines algo-trading development In other words, more efficient big data infrastructure can create profitable opportunities[5].

The performance of data processing systems based on system configuration, such CPU, memory, network and storage. Advancements in storage have lagged due to 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[4]. To handle QF big data, SSD (solid state disks)-backed storage could be more efficient than HDD (hard disk drive)-backed storage[10].The problems are:

A. Analysis Strategies based on system configuration.
B. Market data need to be parsed by the symbol and time not by category and sector.
C. Data compression and decompression happened during the parallel processing
D. Computer hardware needs to be upgrades every quarter for getting financial computing service.

III. SYSTEM DESIGN

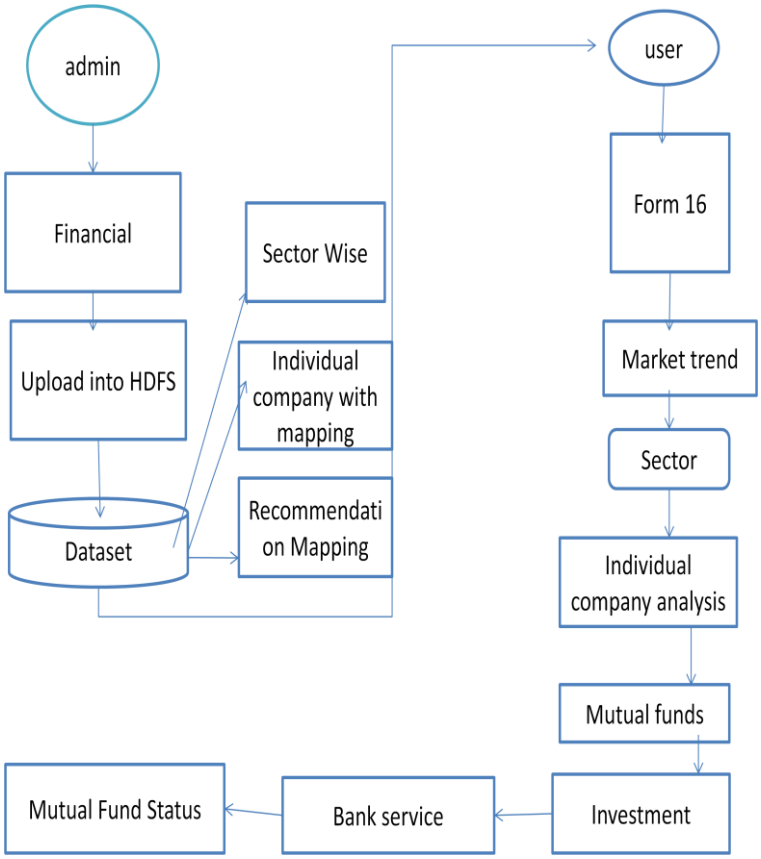

Figure 1.1Architecture diagram

IV.PROPOSED ANALYSIS OF STOCK MARKET:

We develops the application for analyze the financial data's by using big data analytics and Map Reduce program. Our technique focused all strategies such as computer hardware, data processing, prediction method, etc. we mainly consider the market dataset on National Stock Exchange (NSE) and Bombay Stock Exchange (BSE) based companies for data analysis among the NSE and BSE there are five sector datasets such as Banking, Automobiles, Energy & Power, Software and Fast Moving Consumer Good (FMCG).

The analysis process will be complete the user before search any sector Mutual funds Investment, the process of financial data's Map Reduce function Grouping the data's based symbol, time, category and sector. where as the time based grouping classified in date wise, month wise and year wise analysis. In Individual companies grouped by the sector and Individual sector will be grouped by the category finally the category is an over all market analysis. We can get find out each and ever companies profit level in all the year within a seconds.

Our application going to guide the user Mutual Fund Investment, It is based on user requirements. We get all the basic credential from the user then the system analyze and find out the investment. Which is consider the user details and also

consider all the top most companies. The mutual fund recommendation fully based on user credentials not a company profit over the year.

a)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.

The analysis process will be complete the user before search any sector Mutual funds Investment, the process of financial data's Map Reduce function Grouping the data's based symbol, time, category and sector. where as the time based grouping classified in date wise, month wise and year wise analysis. In Individual companies grouped by the sector and Individual sector will be grouped by the category finally the category is an over all market analysis. We can get find out each and ever companies profit level in all the year within a seconds.

Our application going to guide the user Mutual Fund Investment, It is based on user requirements. We get all the basic credential from the user then the system analyze and find out the investment. Which is consider the user details and also consider all the top most companies. The mutual fund recommendation fully based on user credentials not a company profit over the year.

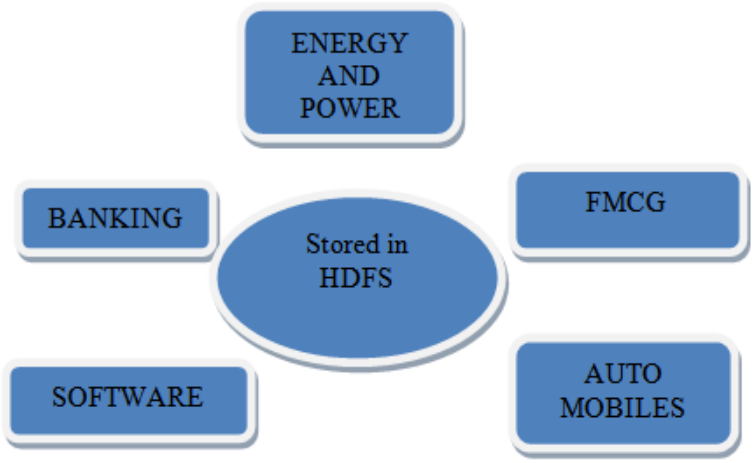

FIGURE 1.2Market analysis

b)PREPROCESSING AND USER FORM16:

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 updation. 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 mutual funds investment based on the user credential such as Account details, annual income, working status, marital status, loan status, etc. all these fields will be available on form16.

PAN NUMBER		
BANK NAME		
ACCOUNT NUMBER		
ANNUAL INCOME		
MARITAL STATUS		
SECTOR		
LOAN DETAILS		

C) MARKET TREND AND INDIVIDUAL PORTFOLIO ANALYSIS:
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 banking, automobiles, energy & power, software and fast moving consumer good. 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.

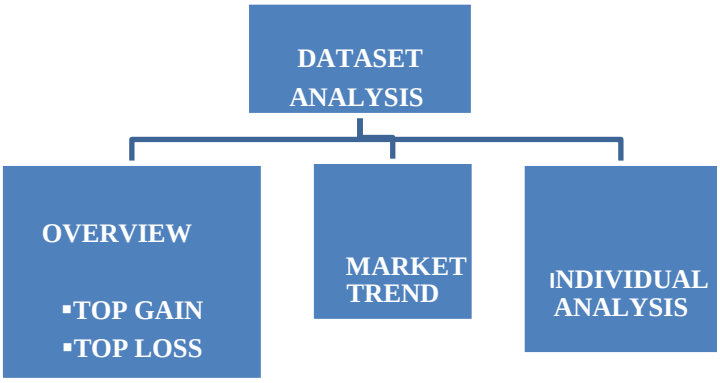

FIGURE1.3 Preprocessing system analysis

d)RECOMMENDATION AND MUTUAL FUND INVESTMENT:
The Mutual fund recommendation Classified in two types, one is Systematic Investment Plan (SIP) another one is EQUITY. SIP having three types of plan such as weekly, month and quarter. The equity having types of plan such as ELSS, Sector funds, Global funds, Equity diversified and Hybrids. User has already submitted the form16 during login phase which is having entire financial detail about the user, in

the preprocessing module we calculate the risk factor for each company so system will compare these two things here, then it will recommend the investment in any one of the company.

V. CONCLUSION

In our System analyze the financial data's and provide the mutual funds recommendation to the user, suggest various mutual funds investment based on analysis which one is top most level in the share market and verified whether those companies funds satisfied the user credential.

VI. REFERENCES

[1] J. J. Angel, L. E. Harris, and C. S. Spatt, "Equity trading in the 21st century," *The Quarterly Journal of Finance, vol. 1, pp. 1-53, 2011.*

[2] C. Alexander, *Market models: a guide to financial data analysis: John Wiley & Sons, 2001.*

[3] A. J. McNeil, R. Frey, and P. Embrechts, *Quantitative Risk Management: Concepts, Techniques and Tools:: Princeton University Press, 2015*

[4] C. Beattie and B. Meara, *"How Big Is Big Data? Big Data Usage and Attitudes among North American Financial Services Firms,"* in *Oliver Wyman*, ed, 2014.

[5]M. Fox, "Thomson Reuters releases "Big Data in Capital Markets" survey results," in *Thomson Reuters*, August 12, 2014.

[6]Carrion, "Very fast money: High-frequency trading on the NASDAQ," *Journal of Financial Markets*, vol. 16, pp. 680-711, 2013.

[7]www.nse.com and www.bse.com

[8]J. Coumaros, S. d. Roys, L. C. Leroyer, J. Buvat, and O. Auliard, "Big Data Alchemy," Capgemini Consulting 2014.

[9] W. Härdle, T. Kleinow, and G. Stahl, *Applied Quantitative Finance*: Springer Berlin, 2001]

[10] H. Chen, R. H. Chiang, and V. C. Storey, "Business Intelligence and Analytics: From Big Data to Big Impact," *MIS Quarterly* **related**, *vol. 36, pp. 1165-1188, 2012.*

[11] Alok kumar vishwakarma has implemented his work "cloud computing in International Journals of research in Engineering and technology *rterly, vol. 36, pp. 1165-1188, 2012*

[12] Hsinchen chen has worked related to "Business intelligence and analytics from big data to big impact" in IEEE transformation of big data

