ANALYZING USER PROFILE AND MUTUAL FUNDS RECOMMENDATION

Mrs. T. AniBernish[1], Ms. J. Angel[2], Ms. S. Jayashree[3], Ms. D. Kamatchi[4]

[1]*Assistant Professor, Department of Computer Science and Engineering, Panimalar Institute of Technology, Tamil Nadu, India.*
[2]*UG Scholar, Department of Computer Science and Engineering, Panimalar Institute of Technology, Tamil Nadu, India.*
[3]*UG Scholar, Department of Computer Science and Engineering, Panimalar Institute of Technology, Tamil Nadu, India.*
[4]*UG Scholar, Department of Computer Science and Engineering, Panimalar Institute of Technology, Tamil Nadu, India.*

ABSTRACT

Tremendous Data insinuates datasets whose size are past the limit of common database programming instruments to get, store, regulate and separate. Hadoop is the broadly utilized enormous information handling motor with a straightforward ace slave setup. In the shared reserve industry, prescient examination assumes a key part in giving information driven choices to dealing with the assets under an NSE and BSE. Huge Information prescient examination utilizing progressed investigation stage can dissect the enormous measures of exchange information and other time slant factors at a large-scale level. This stage can explore these elements with close continuous information and can give exceedingly precise forecasts to the recovering financial specialists later on at a speculator level.

Keyword:– *Big Data, Mutual Funds, User Credentials, Analysis, Recommendatio*

---

1. INTRODUCTION

Hadoop is a free, Java-based programming structure that backings the preparing of expansive datasets in a Parallel and circulated figuring condition. It makes Use of the commodity hardware Hadoop is Highly Scalable and Fault Tolerant. Hadoop runs in the 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 common store industry, prescient investigation assumes a key part of giving information-driven choices to dealing with the assets under an NSE and BSE. The development of the possessions by the financial specialists chooses the development of net resources under the administration of an NSE and BSE and is unfavorably influenced by the counterbalances in reclamations. The characteristics that trigger recovery by financial specialists are mind-boggling in nature to recognize and examine. These characteristics incorporate budgetary exchange designs by the speculator, economic situations, and assessments, macroeconomics factors, conspire level highlights, and statistic factors. Anticipating the recovery conduct requires an advanced stage that can catch numerous elements that influence the reclamation conduct. Be that as it may, enormous information prescient investigation utilizing progressed examination stage can break down these monstrous measures of exchange information and other time incline factors at a full-scale level. This stage can examine these variables for close continuous information and can give exceptionally exact expectations to the reclaiming financial specialists later on at a speculator level.

2. PROBLEM DEFINITION

Investigation Methodologies in view of framework setup. Market data need to be parsed by the symbol and time, not by category and sector. Data compression and decompression happened during the parallel processing. Computer hardware needs to upgrade every quarter for getting financial computing service.

3. PROPOSED SYSTEM

In the existing system 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 an Algorithmic trading technique. 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. 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. 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.

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, the in-memory database still lacks 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.

In the proposed system, the application is being developed to analyze the financial data's by using big data analytics and Map-Reduce program. The proposed technique focused all strategies such as computer hardware, data processing, prediction method, etc. The proposed system 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 examination procedure will be finished the client before look through any area Common assets Venture, the procedure of budgetary information's Guide Diminish work Gathering the information's based image, time, classification and division. while the time-based gathering arranged in date savvy, month shrewd and year astute investigation. In Singular organizations assembled by the segment and Individual segment will be gathered by the classification at long last the classification is a general market examination. The benefit level of every last organization throughout the years can be found in a moment.

The proposed application going to control the client Shared Reserve Speculation, It depends on client prerequisites. The client's essential qualifications are assembled from the client then the framework breaks down and discovers the speculation. Which is consider the client subtle elements and furthermore think about all the best generally organizations? The common reserve suggestion completely in light of client certifications not an organization benefit throughout the year.

3.1 Advantages of Proposed System

1. Prediction of share prices before actual changes occur in share prices. So that timely selling or buying of shares can be done for higher profit margin.
2. Earlier decision-making ability for buying or selling shares.
3. Financial threads detection in quick time.

4. MODULES

- Big Data and Environment
- Preprocessing and User Form16
- Market trend & Individual Portfolio Analysis
- Recommendation and Mutual Fund Investment

4.1 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.

4.2 Preprocessing and User Form 16

A preprocessing is a backend work running in Hadoop groups and furthermore called for long-running employments as it is planned to process mass information with the goal that the application would make utilization of the outcomes delivered for the update. Dataset mapping process is done in the preprocessing stage the whole area both NSE and BSE organizations will be mapped date savvy, month shrewd and year insightful information. The hazard factor of every division and friends will be finished over the preprocessing time. The client shared assets venture in view of the client certification, for example, Aadhar number, Record subtle elements, yearly salary, working status, conjugal status, advance status, and so forth every one of these fields will be accessible on form16.

4.3 Market Trend and Individual Portfolio Analysis

The proposed application provides the history of market data in NSE and BSE so that user will get an idea about market trend mutual funds. The system considers various sectors such as banking, automobiles, energy & power, software and fast moving consumer good. The user can view the history of all the sector by using the systems prediction mechanism among the sectors. The application has different companies user can view the history of each company on the basis of date wise, month wise and year wise.

4.4 Recommendation and Mutual Fund Investment

The Mutual fund recommendation Classified into two types, one is Systematic Investment Plan (SIP) another one is EQUITY. SIP having three types of the plan such as weekly, month and quarter. The equity having types of the plan such as ELSS, Sector funds, Global funds, Equity diversified and Hybrids. The user has already submitted the form16 during login phase which is having entire financial detail about the user, in the preprocessing module the system calculates the risk factor for each company so the system will compare these two things here, then it will recommend the investment in any one of the company.

5. WORKING PRINCIPLE

In the proposed system going to take two different types of Indian market such as National Stock Exchange (NSE) and Bombay Stock Exchange (BSE) within the market having various kind of sector such as Banking, Automobiles, Energy & Power, Software and Fast Moving Consumer Good (FMCG).where as each sector having multiple companies such as,

Banking	**Automobile**	**Software**	**Energy & Power**	**FMCG**
Axis Bank	Hero Motors	Infosys	Adani Power	Colpal
Canara Bank	Mahindra	HCL	BPCL	Dabur
Federal Bank	Exide Industry	Wipro	Gail	Godrej
Bank of India	Bajaj-Auto	TCS	HP	Gskcons
Bank of Baroda	Ashok Leyland	Tech Mahindra	Indian oil	HUL

The proposed framework going to Dissect three years of offer market points of interest of the rundown of the organizations. The examination framework will center around market, area, and friends independently. The framework discovers which advertise performs well in most recent three years by utilizing information investigation. On the off chance that assume the outcome indicating NSE advertise after that framework discover which parts performing admirably in NSE showcase at that point break down all the organization under the market. This is the general investigation of the offer market dataset base paper idea finished at this level.

The proposed framework not exclusively doing investigation part, the application for the most part centered around Common Assets Suggestion. Along these lines, the framework needs to dissect singular division and friends for a long time. At long last, the framework goes out on a limb factor for every part and friends. This hazard factor is only the normal estimation of every last year.

Risk factor 1:

```
if(average2015>average2016&&average2015>average2017)
{
        if(average2016>average2017)
        {
```

```
                    risk=70;
                }
                else
                {
                    risk=50;
                }
        }
```

Risk factor 2:

```
        else
        if(average2016>average2015&&average2016>average2017)
        {
                if(average2015>average2017)
                {
                    risk=50;
                }
                else
                {
                    risk=50;
                }
        }
```

Risk factor 3:

```
        else if(average2017>average2015&&average2017>average2016)
        {
                if(average2016>average2015)
                {
                    risk=30;
                }
                else
                {
                    risk=50;
                }
                }
        }
```

The application will recommend the company for mutual funds investment based on this risk factor. The user has to submit form16 for the application which collects the User credentials. The system has already collected two types of mutual funds available in the share market such as,

SIP	EQUITY
Weekly wise	Sector Funds
Monthly wise	Global Funds
Quarter-wise	Hybrid Funds
	Equity Diversified
	ELSS

The Form16 having user marital status, loan status, annual income, working environment, etc., After the user submits the form16, the system does one of the rules set mining process then the system will find out which funds suitable for the user also which sector and company suitable for mutual funds investment.

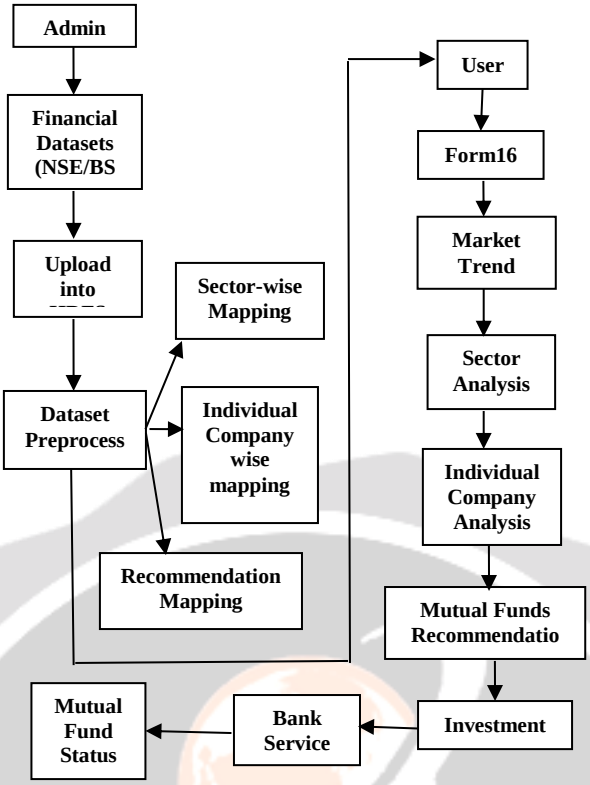
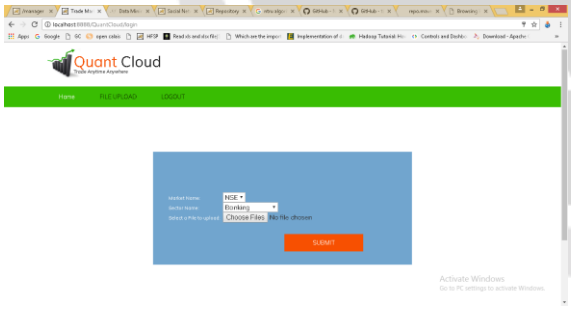


Fig 1: Block Diagram

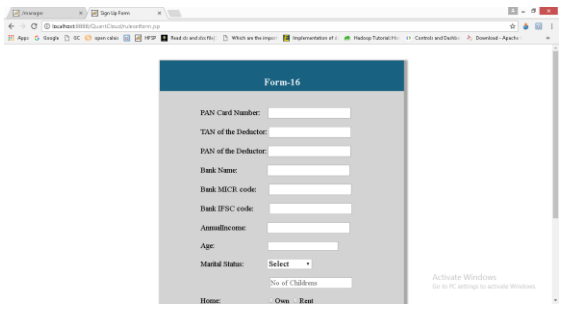
6. EXPERIMENTAL RESULTS

The below displayed are the results of the module implementation. These screenshots show the User Interface through which the modules are being developed.

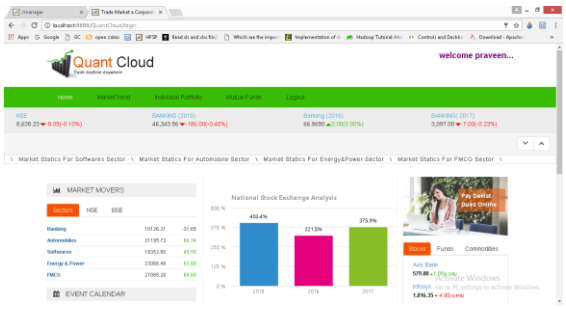
Phase 1: Updating the system with NSE and BSE details.



Phase 2: Generation of User form16.



Phase 3: Analysis of the data of three years.



6. CONCLUSION

In this manner, the framework dissects the budgetary information's and give the common reserve's proposal to the client, recommend different shared assets venture in view of an investigation which one is the highest level in the offer market and confirmed whether those organizations stores fulfilled the client certification.

7. REFERENCES

[1] W. Härdle, T. Kleinow, and G. Stahl, Applied Quantitative Finance: Springer Berlin, 2002.

[2] E. Thorp, "A perspective on quantitative finance: Models for beating the market," The Best of Wilmott, p. 33, 2005.

[3] B. Fang and P. Zhang, "Chapter 11: Big Data in Finance," in Big Data Concepts, Theories, and Applications, S. Yu and S. Guo, Eds., ed: Springer Switzerland, 2016. DOI: 10.1007/978-3-31927763-9_11

[4] X. Shi, P. Zhang, and S. U. Khan, "Quantitative Data Analysis in Finance," in Handbook of Big Data Technologies, A. Zomaya, and S. Sakr, Eds., ed: Springer International Publishing, 2016. DOI: 10.1007/978-3-319-49340-4_21

[5] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, et al., Big data: The next frontier for innovation, competition, and productivity: McKinsey Global Institute, 2011.

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

[7] 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.

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

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

[10] A. J. Menkveld, "High-frequency trading and the new market makers," Journal of Financial Markets, vol. 16, pp. 712-740, 2013.