

A Review of Clustering Technique to Analyse Big Data in Finance

¹Er. Sheenu Sachdeva, ² Pooja Majoka

¹ Assistant Professor, Department of CSE, JCDDM College of Engineering, Sirsa, Haryana, India

² M.Tech. Scholar, Department of CSE, JCDDM College of Engineering, Sirsa, Haryana, India

ARTICLE INFO

Article History:

Accepted: 01 June 2023

Published: 12 June 2023

Publication Issue

Volume 9, Issue 3

May-June-2023

Page Number

444-447

ABSTRACT

A review of existing researches is stated here. Different researches are proposed research considering K-different clustering mechanism. These algorithms have been applied to various real-life applications running in serial, parallel, and high-performing computing environments. But these researches and traditional data mining techniques have their own limitations. The aim of researches is to judge the efficiency of different data mining algorithms on dataset and determine the optimum clustering algorithm. The performance analysis depends on many factors encompassing test mode, distance function and parameters. There are several researches which used K-MEAN clustering and optimization mechanism. The issues related to Kmean clustering would be resolved in this research. Research has introduced the more effective cluster mechanism to classify the data set. Therefore it is essential to propose a data mining technique to deal with big data in Finance. This research work would be helpful to know about data mining of big data. Data related to Equity and Mutual fund is considered in proposed model.

Keywords : Kmean, Big data

I. INTRODUCTION

K Mean Clustering

The k-means clustering algorithm is a data mining and machine learning tool used to cluster observations into groups of related observations without any prior knowledge of those relationships. By sampling, the algorithm attempts to show in which category, or cluster, the data belong to, with the number of clusters being defined by the value k. The k-means algorithm is one of the simplest clustering techniques and it is commonly used in medical imaging, biometrics, and

related fields. The advantage of k-means clustering is that it tells about your data (using its unsupervised form) rather than you having to instruct the algorithm about the data at the start (using the supervised form of the algorithm). It is sometimes referred to as Lloyd's Algorithm because the standard algorithm was first proposed by Stuart Lloyd in 1957.

Role of AI In Share Marketing

The artificial intelligence is suppose to put intelligence in decision making in order to choose appropriate share script at right time considering their 52 week low/high & present value. The AI has provided several

optimization techniques that are suitable to predict value of share.

II. Literature Review

T. Soni Madhulatha et. al. [1] discussed an overview on clustering methods. Cluster has been known as a group of objects. In a cluster the objects are of same class. In the different words, it can be said that the similar kind of objects have been grouped in a cluster. Similarly the mismatched objects have been grouped in a separate cluster.

W.Sarada, Dr.P.V.Kumar [2] reviewed clustering techniques and performed comparative study of those techniques. Clustering has been determined as a process in which a group of abstract objects has been converted into classes of similar objects. A cluster of data objects is used as a group. To do the cluster evaluation, first partition is to arrange the data into groups. It has been done on the base of data similarity.

Bhoj Raj Sharma et. al.[3] presented paper on clustering algorithms. Clustering analysis has been applied in several applications. It has been used in market research, pattern recognition, data analysis, and image processing. To assist the marketers, the clustering has been used. It is applicable to discover the different groups at the base of user.

Shweta Srivastava et. al. [4] has proposed the clustering method evaluation for microarray data international. Clustering allows the dealer of seller of goods to improve the clients base and work on the target sectors. 2. Clustering assist to identify the groups of houses on the basis of defined factors. Such factors may be their value, type as well as the geographical locations also.

Muhammad Husain Zafar et. al. [5] has offered the clustering dependent review of categorization algorithms. Clustering has been applied to review the earth-quake. On the base of sector hit by an earthquake, clustering is able to assist to know and analyze the next possible location. The possible

location where earthquake may be occurred, is shown using clustering.

Asir Antony Gnana Singh et. al. [6] proposed the research work on efficiency evaluation on clustering concepts. Clustering has been determined as a process in which a group of abstract objects has been converted into classes of similar objects. Several methods clustering methods are there such as Partitioning technique, Hierarchical technique, Density-based technique, Grid-Based technique, Model-dependent technique, and constraint-based technique.

III. Methodology

Research Methodology is providing the essential training to select mechanisms, scientific tool and technique, material to implement the research work. Different type of research methodologies have been explained below: Research methodology is the specific procedures or techniques used to identify, select, process, and analyze information about a topic. In a research, the methodology section allows the reader to critically evaluate a study's overall validity and reliability.

3.1 Objective

The several objective of this research has been discussed below

1. To study the existing techniques and mechanisms used to manage Big Data in Finance.
2. To investigate the limitation of existing Finance systems dealing Big Data.
3. To proposed the better solution in order to resolve the issues in Finance Systems.
4. To take the Data related to Equity and Mutual funds
5. To consider different factors such as face value, market cap, promoters holding , fi holding, domestic holding, 52 week low, 52 week high , dividend, P/E
6. To perform comparative analysis of tradition work with proposed work in order to represent how proposed model is better than previous.

7. To withdraw conclusion and discuss the scope of research according to results and discussion.

3.2 Problem Statement: In the traditional work there are different clustering algorithms which have advantages and disadvantages. The present research work is discussing such limitations. There are a variety of algorithms which are used in Finance systems for clustering such as hierarchical, partitioned; density based clustering according to the factors: methodology, structure, model, application or suitability, usefulness. But it is analyzed the tradition Finance systems are not sufficient and have their own limitations. Therefore it has become essential to propose an innovative and fast Finance System that would be efficient to deal with big data.

Mvo (Multi-Verse Optimizer): Multi-verse optimizer (MVO) is another innovation but effective nature-inspired optimization algorithm. It was developed by Mirjalili et al. They considered only two user-defined parameters implementing this algorithm. The main inspiration behind this algorithm is based on cosmological concepts. It has been considered as an innovative meta-heuristic optimization method which is also known as Multi verse Optimizer (MVO).

IV. Proposed Work

In the proposed work, the Data related to Equity and Mutual fund is considered. After getting data, it is classified in to different clusters. Same data is stored in a cluster. Each cluster has different type of data set but the data within a cluster have similarly. There are different factors such as face value, market cap, promoters holding , fi holding, domestic holding, 52 week low, 52 week high , dividend, P/E are considered to analyze the better result in future. Using this proposed module, it will be possible to determine the shares or funds that will provide maximum profit.

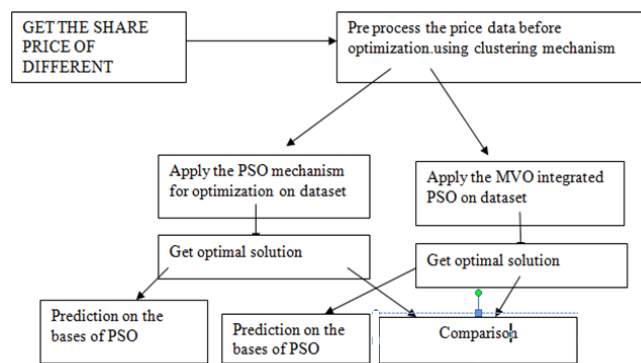


Fig Proposed work

V. REFERENCES

- [1]. T. Soni Madhulatha. AN OVERVIEW ON CLUSTERING METHODS. IOSR Journal of Engineering Apr. 2012, Vol. 2(4) pp: 719-725.
- [2]. W.Sarada, Dr.P.V.Kumar, A REVIEW ON CLUSTERING TECHNIQUES AND THEIR COMPARISON , International Journal of Advanced Research in Computer Engineering &Technology (IJARCET) Volume 2 Issue 11, November 2013
- [3]. Bhoj Raj Sharma.Clustering Algorithms: Study and Performance Evaluation Using Weka Tool.International Journal of Current Engineering and Technology ISSN 2277 - 4106 © 2013.
- [4]. Shweta Srivastava. Clustering Techniques Analysis for Microarray Data. International Journal of Computer Science and Mobile Computing A Monthly Journal of Computer Science and Information Technology IJCSMC, Vol. 3, Issue. 5, May 2014
- [5]. Muhammad Husain Zafar.A Clustering Based Study of Classification Algorithms.International Journal of Database Theory and Application Vol.8, No.1 (2015), pp.11-22.
- [6]. Asir Antony Gnana Singh. Performance Analysis on Clustering Approaches for Gene ExpressionData.International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 2, February 2016 .

- [7]. Kalyani M Raval. Data Mining Techniques
- [8]. V.Ramesh, P.Parkavi, P.Yasodha. Performance Analysis of DM Techniques for Placement Chance Prediction .
- [9]. Neelamadhab Padhy¹, Dr.Pragnyaban Mishra .Survey of Data Mining Applications and Feature Scope .June 2012.
- [10].David Jensen and Jennifer Neville. Data Mining in Social Networks.
- [11].Aarti Sharma, Rahul Sharma,Vivek Kr. Sharma,VishalShrivatava. Application of Data Mining – A Survey Paper.2014.
- [12].Kun-Ming Yu, Che-Yu Lin, Wen Ouyang Jiayi Zhou. An OpenCL Candidate Slicing Frequent Pattern Mining algorithm on graphic processing units. 2011 IEEE International Conference on Systems, Man, and Cybernetics

Cite this article as :

Er. Sheenu Sachdeva, Pooja Majoka, "A Review of Clustering Technique to Analyse Big Data in Finance", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 9, Issue 3, pp.444-447, May-June-2023.