Journal of Network Communications and Emerging Technologies (JNCET) www.jncet.org
Volume 8, Issue 1, January (2018)

Intelligent Inventory Miner Approach for Effective Inventory Forecasting and High Dimensional Inventory Data Management

A.Venugopal
Assistant Professor, Department of Computer Science, Sree Narayana Guru College, K.G.Chavadi, Coimbatore, Tamil Nadu, India.

Jisha. P
M.Phil Scholar, Department of Computer Science, Sree Narayana Guru College, K.G.Chavadi, Coimbatore, Tamil Nadu, India.

Abstract – Data mining is the most appropriate method for many applications. Inventory data mining is one of the most prominent applications in every business. Tracking and analyzing the market database is more complicated process. To overcome the inventory data management issues, an intelligent inventory mining technique IIMiner is proposed. These inventory data's are considered as time series data's with quantity and product brands. These time series datasets need more attention for business intelligence process. There are several techniques were used in the existing researches. Those techniques are based on the statistical analysis with the historical inventory transactional data. Inventory data forecasting and detecting anomaly from the time series data is performed using the IIMiner framework. To achieve the effectiveness, the IIMiner proposed three phases, which initially performs the Frequent Inventory Item Mining (FIIM). The sales forecasting using Temporal Naïve forecasting model is proposed, which is used to predict the inventory values. And finally the inventory anomaly has been detected from the time series data using a semi-supervised classification method based on the support vector machines. Here the weighted batch based anomaly detection is performed. The new anomaly detection approach is named as weighted batch-support vector machines (WB-SVM). Finally the experiments and results proved the proposed system utilized maximum accuracy and effectiveness in inventory management and visualization.

Index Terms – Data Mining, Business Intelligence, Big Data, Forecasting, Demand Prediction, Anomaly Detection.

1. INTRODUCTION

Data mining in Market Basket Analysis and Business intelligence has been a hot topic of research for several decades [1]. It includes the application of data mining techniques to inventory mining and Analysis. In the current scenario several organizational and supermarket databases are capable of providing a lot of information on product sales, stock and buying behaviors, which can be investigated in order to find the items that are most demand and also those that are purchased frequently [2][3]. Several researchers only concentrated on frequent items and its association or basic inventory mining, the current proposal deals with the association finding for both frequent and ability to find high demand products and its expected quantity. The current inventory management and analysis system has tremendous opportunities over the world. This needs a more accurate and highly scalable with user friendly features for product analysis and decision making.

Due to the high and dynamic dataset, the analysis and findings are tedious, this kind of huge datasets are called as big data [4]. The proposed study has tremendous opportunity to improve the inventory analysis in the market database and helps to find high and low demand products and its associations for right decision making.

2. PROBLEM DEFINITION

Owing the continuous development in the usage of computers in all places, several databases are being constantly generated. However, there is no effective technique to utilize these inventory databases efficiently and to find the valuable associations in and between them for effective inventory forecasting and management [5]. Association rule mining and Business intelligence concepts are used to find interesting association or correlation relationships among a large amount of data items.

With huge amounts of data constantly being collected and accumulated, many organizations are showing interest in mining inventories and handling them effectively from this large collection of business transaction records, as it can assist in many business decision making processes, such as catalog design, cross marketing and others. And the smart inventory management can be used to improve the business level and profit. Finding and calculating the inventory aging and anomaly entry from large databases is extremely convoluted. These databases contain numerous irrelevant and redundant data records which are not essential to extract the desired results. In addition, these irrelevant data considerably affect the quality of the inventory miner and hence there is a requirement to preprocess these records. This will be performed using effective feature selection algorithms. Mining inventories with its aging and demand forecasting in large databases play a vital role in the field of market analysis.

The rules which are generated using association rule and inventory mining tools can help to the user make the right decision on the market data. This paper addresses the above problem along with some research issues such as the issues associated with the discovery of frequent, closed and weighted itemsets from transactional data sets for effective demand forecasting, which also finds the anomaly in inventory streaming. The problem of mining high demand products by considering effective data mining is the popular research area. Every algorithm proposed in literature results [6] in slow computation due to its uncertain huge datasets.

3. PROPOSED SYSTEM AND METHODOLOGY

This chapter brings the details about the proposed system and the contributions of the proposed work, this also addresses the challenges in the previous work on positive rule mining, inventory mining, demand forecasting and provides a condensed and meaningful representation of *results* named IIMiner (Intelligent inventory Miner), which integrates the concept of closed itemset into high demand item mining. The contributions are four-fold and correspond to resolving the previous challenges.

3.1 Contributions of the proposed system:

Mining and managing inventories with frequent patterns and its quantitative rules are very tough task because there are an enormous number of such patterns that can be derived from a known data set. In specific, the major problems associated in mining inventories are as listed below:

(1) Identifying frequent products using Frequent Inventory Item Mining (FIIM), and

(2) Forecasting inventories from the past transactions and temporal details with the same algorithm

(3) Finding high demand products without utilizing much steps and time.

(4) Discovering frequent as well as high demand, weighted patterns in large dataset.

(5) Finding transaction anomalies from the time series dataset using WB-SVM (weighted batch support vector machine).

To address the above issues, several inventory mining tools and data mining techniques have been constructed and integrated in the SAP software's. However, the inventory mining process is difficult when the dataset is dynamic and uncertain. And although probabilities of item occurrence may be remapped to weights, the semantics behind probabilistic and weighted inventory mining is radically different. In fact, the probability of occurrence of an item within a transaction may be totally uncorrelated with its relative importance. For instance, an item that is very likely to occur in a given transaction may be deemed the least relevant one by a domain expert. A parallel effort has devoted to discovering the transactions and the quantity among data. The proposed system initially takes the inventory dataset and performs different levels to achieve the effective business intelligence and business decision making. The detailed steps of the proposed IIMiner are depicted in figure 1.0.

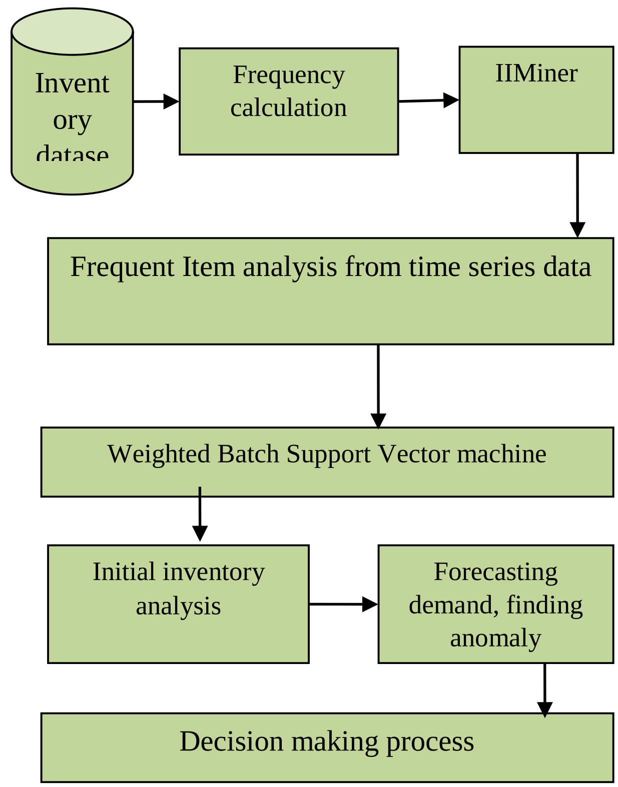

Figure 1.0 Architecture of proposed IIMiner

IIMINER is built upon the previous developed large scale data analysis system, Iminer and FIU-Miner, which is a Fast, Integrated, and User-friendly system to ease data analysis and decision support systems. The system is composed of different layers: *GUI (Graphical User Interface), Data Analysis Level, Task and System and Administration Level*.

GUI: This level contains numerous interactive interfaces for inventory management operations. Particularly, it provides a graphical chart based output to allow users to have an overview of the current inventory status and the forecast. In addition, several key indices of inventory, e.g., forecast qty and inventory-to-sales ratio, are presented in Inventory Index

Interface, assisting users in promptly querying the forecast details of a particular product.

Data Analysis Level. This level is the heart of the system. Beside basic data processing and exploration functionalities, Data Analysis Level consists of appropriate data mining solutions to the corresponding tasks of inventory management, including Inventory Forecasting, Anomaly Detection, and Inventory Aging Analysis..

Task and System Administration Level. Task and System Management Level provides a fast, integrated, and user-friendly system to configure complex tasks, integrate various data mining algorithms, and execute tasks in a distributed environment. All the data analysis tasks in Data Analysis Level can be configured as workflows and scheduled automatically.

4. IIMINER FRAMEWORK

The IIMiner framework involved with inventory data analysis, management, demand forecasting, and inventory aging analysis.

4.1 Data Exploration: Initially, the IIMiner performs Statistical Analysis and *candidate generation process, which* are capable of assisting data analysts in exploring inventory data efficiently and effectively. Statistical product analysis in different sizes and dimensions can quickly discover interesting time frame, product quantities, and its association between the other products. Data exploration process is a convenient approach to explore high-dimensional data so that data analysts can have a better view of the characteristics of the dataset.

4.2 Data Analysis: In the system, the data mining approaches in *Algorithm Library* can be organized as a configurable procedure in Operation Panel. *Operation Panel* is a unified interface to build workflows for executing such tasks automatically. User can check the desired process for the given product such as A, B…Z.

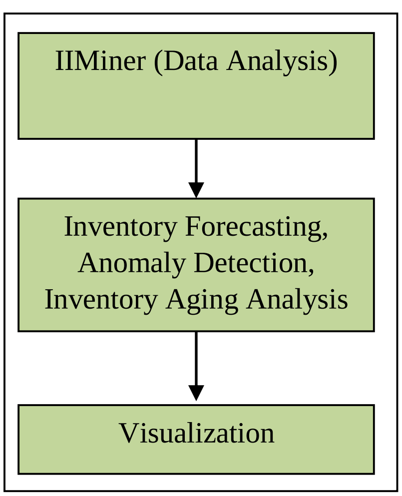

Figure 2.0data analysis process

As shown in figure 2.0 *Operation Panel* mainly contains three phases, *Inventory Forecasting*, *Anomaly detection*, and *Inventory Aging Mining*. These phases implement various mathematical models and advanced data mining algorithms to address the challenges of history data analysis in inventory management.

a. Inventory Forecasting: The primary goal of inventory forecasting is to minimize the inventory loading in the inventory management process. A common practice of inventory forecasting is to predict the demand for a particular item in the future and reserve the appropriate amount of items, based on the forecasting results. But, inventory data is a type of time series with large volume, long time span, and less regularity. These features bring up three challenges to inventory forecasting:

- Implementing an accurate interpretable inventory prediction;
- Mining association between items and finding the demand of the items
- Modeling the relationships among multiple time series data sets and predicting their future values simultaneously.

Although, in recent years, there has been an explosion of interest in mining time series, traditional approaches such as auto-regression(AR), linear dynamical systems (LDS), Kalman filter(KF) cannot solve above challenge directly. Hence, more accurate and optimal forecasting methods are needed. In IIMiner a new inventory forecasting models with semi-supervised temporal Naïve prediction model is designed and deployed to solve the prediction problems.

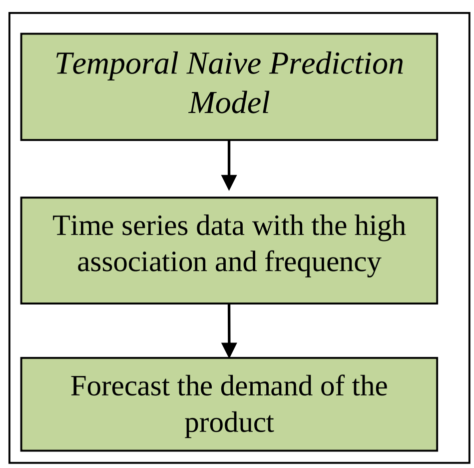

Figure 3.0 Temporal Naive Prediction Model

Temporal Naive Prediction Model : To implement an accurate and interpretable inventory forecasting, a two-step temporal naïve prediction model is proposed. First, it adopts machine learning techniques and temporal techniques combined with time series analysis methods to obtain a forecasting basis. Second, it takes into account multiple factors of inventory such as seasonality, trend, and special events for dynamic inventory forecasting.

The temporal naïve prediction model collects the time series data from the database and finds the forecasting value of the given product. The process is shown in figure 3.0 the time series data of product A is "2,3,1,4,8,10,12,18,28,32,…..45". and the forecast will be performed for each transaction temporally. The temporal naïve prediction is performed by the following equation Eq(1.1). The overall framework of dynamic forecasting model is shown in Fig 3.0.

$$v(t) = \sum_{l} \text{wpt}\, vp^{t}, \text{s.t.} \sum \text{lwpt} = 1 \qquad \text{Eq(1.1)}$$

where, v is the forecast value, p represented as a product, t is the time value. Suppose the forecasting value of the learning model p∈P and the time is represented as (t), and its weight at (***wt)*** time. Then the temporal time for the stock out time at t is represented in the equation. Initially (*t*=0), all the learning model have the same contribution to the forecasting result,

Instance	Value	Forecast
0	2	2
1	3	2
2	1	3
3	4	1
4	8	4
5	10	8
6	12	10
7	18	12

Figure 4.0Temporal Naive Prediction Model forecast results

The temporal naïve prediction model forecast results are shown in fig 4.0 with the actual and forecast values. To evaluate the results Mean Absolute Error(MAE) Mean Square Error (MSE), Mean absolute percentage error (MAPE), Mean Percent Error (MPE) are calculated. Temporal naïve based prediction model are integrated into an interactive interface IIMiner to provide multiple prediction of demand forecasting.

a. Temporal sales Prediction: Existing management systems often forecast the two time series stock in and stock out separately. Both of them are treated as independent ignoring their relationship. In practice, the amounts of stock in and stock out in an inventory are often dependent on each other. The amount of stock out (S_{out}) is usually subject to the amount of stock in (Sin) at the same or near time periods to prevent the situation that an item is out of stock. Also, the scheduled Sin primarily depends on the past S_{out} to avoid the situation that a unit is in excess of demand. So, two time series of S_{in} and S_{out} bear some inter-dependencies according to the characteristics of inventory management. The existing single time series predictive methods lack the capability of capturing the dynamic relationships between multiple time series or predicting their future values simultaneously. Little research attention has been paid to predicting the movement of a collection of related time series. In IIMiner, the model the interdependencies of time series data and integrate them into the process of time series prediction. The model can capture the dynamic relationships between multiple time series data set and predict their future values simultaneously. Specifically, in the domain of inventory management, the aggregated amount of S_{in} is often larger than the aggregated amount of S_{out} in a given period to avoid "out of stock", so the hold out options is added in the result. In addition, S_{in} and S_{out} should be close to each other to prevent a unit from "excess demand". Based on such an intuition, we transform the requirement of inventory management into model constraints and perform time series prediction under the constraints.

Determining the Forecasting Basis: In this step, the IIMiner employed the machine learning algorithms to capture the hidden patterns in the inventory time series of every product. Each algorithm is used to build an inventory forecasting regression model based on the past inventory transaction data and update on a single transaction basis. These algorithms includes time series analysis algorithm and weighted linear representation model.

a. Inventory Anomaly or Abnormality Detection

Monitoring inventory index for abnormality detection is a critical task of inventory management. This problem becomes difficult in the Big Data era since the data scales substantially, and the types of abnormalities get diversified. In the proposed system, in addition to providing traditional statistical methods, a classification-based abnormality detection method to identify abnormal items is deployed. Here, the abnormality is big fluctuation of sales or sale data. A precise and practical abnormality detection method using WB-SVM is proposed. The fundamental techniques and novelties are high-lighted as follows:

- Labeling anomalous points automatically. Original inventory time series do not contain the label information of abnormality. This searches for anomalous points from the data automatically and annotates them with different labels and highlights that, so that the problem is modeled as a semi-supervised problem.
- Adopting WB_SVM based classification method instead of regression-based method to reduce the computing effort. After labeling the data with class labels, abnormality detection can be formulated as a classification problem. Compared with common regression-based methods, the computing effort is reduced significantly.

- Changing the cost measure to be cost-sensitive. Existing abnormality detecting methods consider the equal costs for different types of detecting errors. However, in practice, abnormality detection is a cost-sensitive learning problem. The costs for different class instances are different.

Based on above ideas, IIMiner implements a classification based anomaly detecting model named WBSVM, which combines the temporal Weighted batch Support Vector Machine. The proposed method is used to capture the fluctuated points to form the active training dataset, and the weights of the change points are also determined according to the changing trend. Finally, Weighted batch Support Vector Machine (WSVM), which is a semi supervised learning technique is adopted to build the anomaly detection model.

WB-SVM Process: The proposed system uses SVM based classification algorithm for anomaly detection, this is a semi supervised technique, and which is the new classification method proposed with temporal and weighted batch processing. It develops on the basis of statistical model. The basic thought of WB-SVM is first input the sample and through the kernel function map to the higher dimensional eigenspace, then looking for the optimum boundary in the eigenspace through the maximizing classification interval and the classification interval is maximized and can be transformed into quadratic programming problem. Support vector machines based solution has been applied in the proposed inventory anomaly detection process. Several literatures discussed the one class SVM. The proposed system deals with the concept of weighted batch SVM. Inventory anomaly detection with a WB-SVM offers promising results. The system concentrated on the filtering step or combined with approaches that can deal with a small amount of uncertainty. The basic principle of WB-SVM is finding the optimal linear hyperplane in the feature space that maximally separates the inventory time series dataset into two classes. For linearly separable and non-separable data, it can be translated into quadratic programming (QP) and can get an only limit point; this is interesting when this applied for WB-SVM based anomaly detection. In the case of non-linear, WB-SVM can map the input to a high-dimensional feature space by using non-linear mapping and then the linear hyperplane can be found. After one class classification round, the proposed system will make the next searching round by the changed hyperplane.

WB-SVM is used to solve inventory anomaly classification problems. WB-SVM maps linear algorithm into non-linear sub space. And it performed after the clustering process. For this mapping purpose it uses batch function. This inventory anomaly and batch functions can be used at the time of training of the classifier to selects support vector along the statistics data of this function. These support vectors are used by SVM to classify data that outline the hyper plane in the feature space.

Support vector machine is to find out the multi kernel hyperplane which is the most distant away from any data, this can minimize error rate. The distance between data and hyperplane is shown in the Figure 4.1 Weighted batch Support vector machine must meet certain conditions:

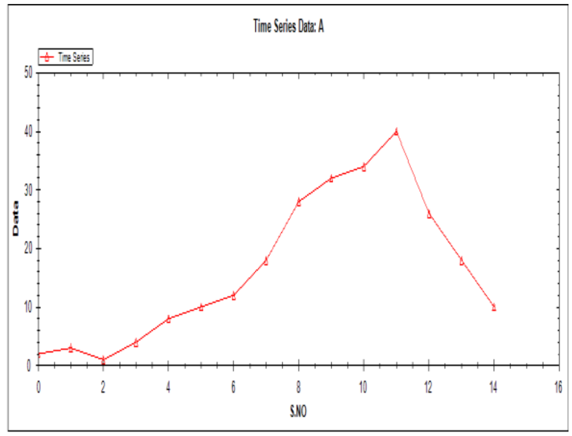

Figure: 5.0 WB-SVM classifications from the time series data

The distance between the data and hyper plane is shown in the above figure. The idea is that WB-SVM maps inputs vectors nonlinearly into the high dimensional feature space and construct the optimum separating hyperplane for inventory anomaly detection. In some cases, the technique uses linear and some time non linear. So this is known as optimal hyper plane selection.

5. RESULTS OF TEMPORAL NAÏVE FORECASTING MODEL

This first evaluate the performance of core functionalities of the system on a real world inventory big dataset collected from the government website. Compared with existing inventory analysis techniques, IIMiner can perform large-scale data management and data analysis. IIMiner shows a powerful decision support performance by real world inventory data.

Forecasting Accuracy. To illustrate the performance of proposed inventory forecasting models, this conduct two experiments for temporal naive inventory forecasting model respectively. Existing Forecasting Techniques:

i. Simple Moving Average (SMA):

A simple moving average (SMA) is an arithmetic moving average calculated to perform inventory forecasting for a number of time periods and then dividing this total by the number of time periods. The SMA is calculated using the following equation

$$F(t+1) = \frac{((Dt + D(t-1) + D(t-2) + ... + D(t-n+1))}{n}$$

Here the Dt is the dataset. ie the time series of a selected product. The following figure 6.0 shows the result of the SMA with the set of error calculation.

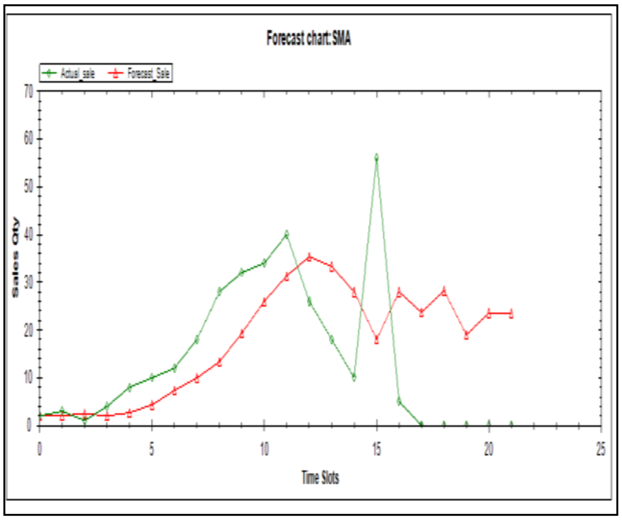

Figure 6.0 forecast results of SMA

ii. Exponential Smoothing (ES):

Exponential smoothing (ES) is a forecasting method used for smoothing time series data. Whereas in the simple moving average (SMA) the past observations are weighted equally and exponential functions are used to assign exponentially decreasing weights over time.

$$F(t+1) = (Alpha * D(t)) + (1 - Alpha) * F(t)$$

Here the Dt is the dataset. ie the time series of a selected product. The following figure 7.0 shows the result of the ES with the set of error calculation.

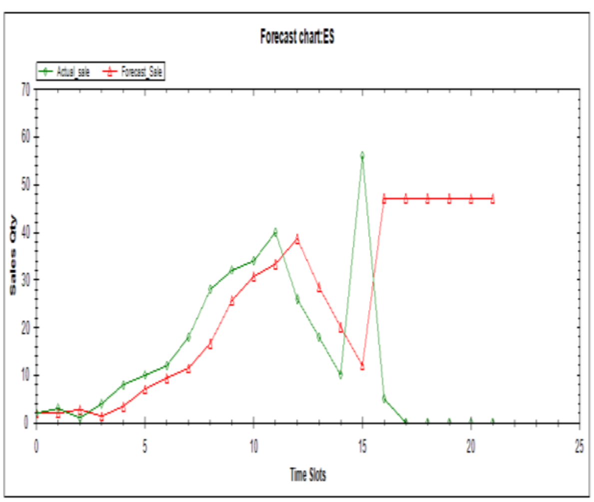

Figure 7.0 forecast results of ES

iii. Proposed Temporal Naïve Forecasting model:

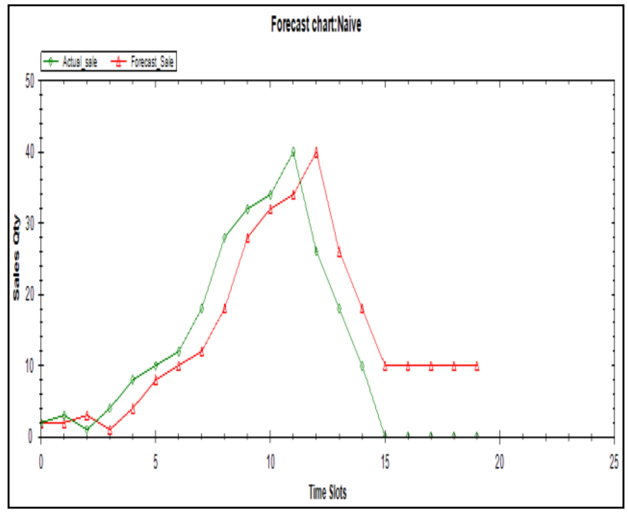

Figure 8.0 the forecast results of temporal naïve forecasting

After the deployment of IIMiner, this weekly collected the real-world data of stock out from several products. The Temporal naive inventory forecasting model predicts the amount of stock out from given date As shown in figure 8.0, compared with common statistical analysis approaches, i.e., Moving average and Contemporary comparison, the proposed temporal naive inventory forecasting model obtains the highest fitting degree with the true values of both sales trends and periodicity. Moreover, the model can give a reasonable interpretation to the analysis results.

6. CONCLUSION

The system proposed IIMiner has been developed to forecast inventory sales and finds its associations in the high dimensional inventory dataset. Mining frequent items from an inventory database using Finding Inventory Item Mining (FIIM) and temporal naïve forecasting model has been designed and developed. This helps to discover the items with high demand as well anomaly transaction on different parameter have been proposed. For fast anomaly data detection and handling, the system created a new algorithm named as WB-SVM which includes two improves the existing SVM with weighted batch features. This feature is very useful to handle high dimensional and big inventory databases. This improves the performance of existing iMiner with effective implementation. The proposed algorithm reduces the number of iterations and errors and maintains the datasets effectively. The experiments' and results shows the proposed system performs better than existing system.

REFERENCES

[1] Chaudhuri, Surajit, Umeshwar Dayal, and Vivek Narasayya. "An overview of business intelligence technology." *Communications of the ACM* 54, no. 8 (2011): 88-98.

[2] Ahmed, Syed Riaz. "Applications of data mining in retail business." In *Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004. International Conference on*, vol. 2, pp. 455-459. IEEE, 2004.

[3] Sani, Babangida, and Brian G. Kingsman. "Selecting the best periodic inventory control and demand forecasting methods for low demand items." *Journal of the Operational Research Society* (1997): 700-713.

[4] Fan, Wei, and Albert Bifet. "Mining big data: current status, and forecast to the future." *ACM sIGKDD Explorations Newsletter* 14, no. 2 (2013): 1-5.

[5] Mishra, Brojo Kishore, Deepannita Hazra, Kahkashan Tarannum, and Manas Kumar. "Business Intelligence using Data Mining techniques and Business Analytics." In *System Modeling & Advancement in Research Trends (SMART), International Conference*, pp. 84-89. IEEE, 2016.

[6] Venugopal, Jisha. P. " Data Mining Techniques for Business. Intelligence and Its Applications – A Survey." International Journal of Advanced Research in Computer and Communication Engineering *(IJARCCE)*, pp. 193 - 196. 2017.