Data mining methods in a metrics-deprived inventory transactions environment

E. A. Beardslee & T. B. Trafalis
Department of Industrial Engineering, University of Oklahoma, USA

Abstract

Over the past decade, several data mining techniques have come into use within a variety of data intensive fields. Mining transaction data to discover interesting patterns or help forecast future rare events based upon historical records is of key interest. Among the classic problems facing any supply chain is that of determining how much of a given item to keep on the shelf of a warehouse or a store, and at what level to initiate replenishment of this supply. This problem has been well examined and extensively researched. However, applying current data mining techniques to support the determination of effective inventory stock levels and economic re-order points has yet to be explored. In this paper, Naïve Bayes, Bayes Network, SVM, MLP, Logistic, and J48 decision tree data mining techniques are compared with respect to their effectiveness when tasked with predicting stock shortage conditions indicated by the reduced transaction history of a bench stock inventory. These comparisons have been made using training and test data sets drawn from a set of 1.8 million transactions. Each method is applied to the same set of transaction data that spans three years of bench stock inventory activity. This set of transactions contains orders, receipts, stock outages, and stock shortages but it does not include issue transactions from the inventory. Since the time of issue or removal of the item from the inventory is not known, determining the expected demand becomes very difficult. The effectiveness of the data mining predictive capability is measured with respect to the accuracy of the prediction and the lead-time provided to the inventory planner for the given prediction. The results show similar performance among most of the data mining algorithms using "actual event" prediction evaluation criteria.

Keywords: data mining, inventory, supply chain, Naïve Bayes, SVM, prediction, stock shortage, transaction, time-series, rare event.



WIT Transactions on Information and Communication Technologies, Vol 35, © 2005 WIT Press
www.witpress.com, ISSN 1743-3517 (on-line)

1 Introduction

Within the world of supply chain and inventory management, thousands of transactions occur daily that provide a time series mosaic of the broader activities supported by these functions. With the advent of very large data storage capabilities, powerful processors and advanced data search methods, discovering the previously unknown aspects of this mosaic have emerged revealing interesting new views. This discovery is facilitated and advanced primarily by the rapidly growing technology known as data mining [1]. Within the last decade, data mining has been used to advance the effectiveness and efficiency of transaction based systems such as supply chain and inventory management.

Among the classic problems facing any supply chain, that of determining how much of a given item to keep on the shelf of a warehouse or a store, and at what level to initiate replenishment of this supply has been well examined and extensively researched. However, applying the current data mining techniques to support the determination of effective inventory stock levels and economic re-order points has yet to be fully explored. One specific area that needs research is that of using a transaction history of orders, receipts, stock shortages, and stock outages to help inventory planners and stock item managers with the task of forecasting or predicting future stock shortages. Using data mining, in a "metrics deprived" transaction environment in which standard inventory planning techniques falter, it may be possible to discover patterns, associations, classes or categories within the transaction stream that provide a clue to the next stock shortage or stock outage.

Using several current data mining techniques, we plan to review each method's ability to provide reliable predictive capability to an inventory item manager. Each method will be applied to the same set of transaction data that spans three years of bench stock inventory activity. This set of transactions contains orders, receipts, stock outages, and stock shortages. It does not include issue transactions from the inventory. Therefore, the time of issue or removal of the item from the inventory is not known. The lack of this data point complicates the problem space by removing a key transaction element, without which, determining expected demand becomes very difficult.

The report that follows contains a review of some of the current research into the uses, problems and potentials of data mining large data sets in section 2; a description of the problem under analysis in section 3; and a detailed description of the methods used to explore and analyze the problem in section 4. In section 5, computational results are discussed. A set of possible future extensions and enhancements to this specific area of research and a conclusion are discussed in sections 6 and 7 respectively.

2 A brief review of selected current research

A great deal of research in the past decade has focused on the exploitation of the vast amount of data becoming available in nearly all facets of business, science, academia, and government. A variety of data mining techniques have come to



light that have shown promising results when applied to large data sets of transactions and operations for the purpose of discovering new and interesting trends and patterns in the data. A case study performed by Dhond et al [2] showed that by using neural network data mining techniques to analyze thousands of transactions, a model could be created that, if implemented, would result in significant cost reduction in retail pharmaceutical inventory. One key technique that Dhond et al [2] used during this data mining approach was changing the interval or resolution of the prediction time frame. A prediction time frame of one day resulted in unacceptable error rates; however increasing the time frame to one week reduced the prediction error significantly.

Prediction of some critical event or situation is a common goal of many data mining implementations, [2,3,4,5,6,7]. The class of problems which is under analysis in this paper has the primary goal of predicting a rare event based on the patterns in historical transaction data. In addition, the transactions and events that record some piece of data about the inventory are all time based. Therefore some examination of time-based data mining is important. Finding interesting relationships and patterns within historical data streams leads to time-series analysis and data mining in the temporal domain. In the research presented by Vilalta and Ma [4], the goal of predicting rare events in a time series was examined by gathering only the events that precede a target rare event with the additional use of a fixed time frame within which the events are recorded and categorized. Most learning algorithms assume even class distributions and adopt a discriminant-description strategy: they search for a separator that best discriminates examples of different classes, [4]. The authors of this paper describe and attempt to address the problem of mining and learning the rules that define a rare event within the data body and show that this requires special consideration.

3 Problem description

Given a set of inventory bin locations storing bench stock items (nuts, bolts, fasteners, rivets, washers, etc., as individual inventory items, having negligible value, often between $0.0001 and $0.50 each), where each item is used within a variety of maintenance tasks, the goal of the inventory management specialist supporting these maintenance operations focuses on keeping the required level of bench stock in each bin at an acceptable quantity such that the maintenance operation is not interrupted. It is critical for an inventory manager, or the inventory management systems, know how much stock is contained within any given stock location. Yet, obtaining this information is not trivial. Inventory systems over the past 20 years have moved toward tracking issues from stock locations through a variety of point-of-sale class technologies. However, the bench stock inventory manager does not have the luxury of a cooperative consumer that retail inventory managers enjoy; the removal of a bench stock item from the inventory location for use on the maintenance floor is rarely recorded.



In the framework of this research, a frequent, periodic manual review of the inventory, providing estimates of the stock levels, is the method of determining the stock levels throughout the bench stock inventory. This method is susceptible to the following problems: the inventory reviews are not typically done on a strict time interval, the estimates of the stock inventory levels are greatly influenced by the warehouse worker's experience and estimating capabilities, and the recommended re-order points may be ignored by the warehouse worker in the light of external maintenance pressures. The process of inventory review and the stock location barcode scans provide the base transaction data evaluated through the data mining process described within this research.

So the problem, defined within the constraints identified, can be stated; using the data points available, and a history of several months of transactions, can this data provide reliable predictions of impending stock shortages for a given stock location.

4 Method of analysis

In order to construct an effective and useful evaluation of a set of data mining techniques, a series of preparation steps were performed before data mining. The method of analysis described here starts with a detailed description of the transaction data and how it is extracted, transformed and condensed. Then, the data formatting and cleansing techniques will be reviewed. The WEKA Explorer data mining workbench, Witten and Frank [1], was the primary tool used within this research to conduct the data mining tests. The criteria for evaluating the data mining techniques will then be defined.

4.1 Data extraction

As previously described, the method of recording the current state of any given stock location or bin within this bench stock inventory is through manual review of the bin and electronic scan of a standard bar code identifying the stock in the bin. The warehouse workers only scan the bins if the quantity in the bin, in their estimate, is below the re-order point printed on the bin label. Otherwise, the bin is not scanned and no information regarding the inventory level of the bin is captured. If a bin is scanned and is not already on order, an order is created by the system and transmitted to the vendor contracted to supply the given part.

The only other information item captured in the scan process is the presence of an empty stock location. This process accounts for two of the data elements extracted into the data warehouse: the stock outage and the evidence of demand manifested through the request for order placement. Receipt transactions are recorded when the item ordered for a stock location is placed into the bin. Manual orders are often placed to expedite shipment from an alternate source of supply. This order is also recorded and again provides evidence of demand. Stock shortages are reported by the maintenance workers and recorded in the system as an immediate demand notification.



Because of the possibility of overlapping order placement, a single continuous time series describing the replenishment history for a stock location can not be accurately described by the individual correlated orders and receipts. Accurate description of the replenishment stream for a given stock location by the transaction data is critical to the data mining task. Therefore, during the data warehouse extract, transformation and load process (ETL process), a normalized time series is created or updated for each stock location. This time series is defined as a supply-chain segment, fig. 1. A supply-chain segment is initiated by any order immediately following a receipt for a given bin, and ends with the last receipt preceding the next order. The first supply-chain segment starts with the first order, and the most recent segment for a bin is considered incomplete. It is important to note that the time sequence of the bench stock orders and receipts is paramount. The replenishment state of the stock location must be viewed from the stock location itself and not the order processing apparatus.

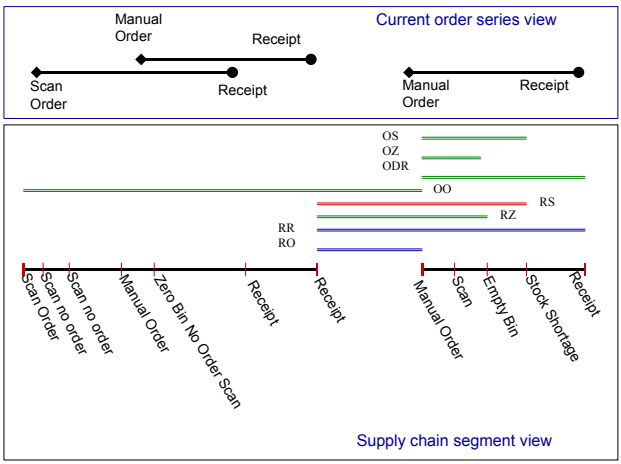


Figure 1: Key supply-chain segment attributes.

With each bin history, defined through a series of supply-chain segments, the key data elements to be examined by the data mining methods can be described in fig. 1. All of the time intervals are calculated in whole days. Table 1 defines the time based attributes.

Table 1: Training set date attributes.

Prior Segment Date	Following Segment Date	Attribute Described
Segment end (receipt)	Segment end (receipt)	Receipt to Receipt (RR)
Segment end (receipt)	Segment start (order)	Receipt to Order (RO)
Segment end (receipt)	Stock outage date	Receipt to Zero (RZ)
Segment end (receipt)	Stock shortage date	Receipt to Short (RS)
Segment start (order)	Segment start (order)	Order to Order (OO)

Following Segment Date	Following Segment Date	Attribute Described
Start segment (Order)	Segment end (receipt)	Order to Receipt (ODR)
Start segment (Order)	Stock outage date	Order to Zero (OZ)
Start segment (Order)	Stock shortage date	Order to Short (OS)

A ratio attribute of stock shelf time (RO) to stock shelf time plus lead time (RR) can be constructed from this data to provide a normalized comparison of relative replenishment stability. Another attribute derived from the supply-chain segment is the inter-segment daily use (IDU) data element. It is the number of received items in the prior segment minus the re-order point quantity, divided by the number of days from the end of the prior segment to the start of the following segment. This gives a relatively accurate view of the daily usage at that point in time. When this is evaluated in relation to the planned usage rate, it may give insight into the adequacy of the order generation parameters.

All of these attributes help describe the replenishment state of a given stock location. However, at any given point in time, a bin may require immediate attention in order to prevent a stock shortage. At other times the bin may not need any attention outside the normal warehouse management process. Determining the state or class in which a stock location may be now or will be soon, and presenting this concept in a simple way to the inventory managers is a key derived data attribute. The process of identifying these key features is critical because of its impact on the effectiveness of many data mining algorithms. "Feature selection is of paramount importance for any learning algorithm which when poorly done (i.e., a poor set of features is selected) may lead to problems associated with incomplete information, noisy or irrelevant features, not the best set/mix of features, among others [8]. In addition to the normal difficulties surrounding good feature selection, primary of which is the manual selection based upon a deep understanding of the learning problem and what the attributes actually mean [1], the significant target event (stock shortage and outage) we are searching for is a rare event in the transaction stream. The feature selection method employed to define the attributes and class definition within the supply-chain segment data was a combination of the method used by [4] on predicting rare events and the automatic methods available in WEKA and manually.

Two training sets of data were drawn from the data warehouse, one containing an unrestricted random sample of 34,754 supply-chain segments from the entire set of 1.8 million segments, and the second, a sample of 29,072 records containing supply-chain segments which denoted only stock shortages.

Test sets were extracted from the database by selecting a set of transaction days, randomly selecting an unrestricted set of supply-chain segments, and then randomly selecting a set of supply-chain segments restricted to those which contained a stock shortage in the first supply-chain segment following the selected query day. Each of the test sets contained the supply-chain segment ending on the selected day and the immediately preceding supply-chain segment for each selected stock location.

4.2 Data formatting

In order to use the WEKA data mining workbench, the data was formatted for input into the WEKA implementations of the data mining and machine learning algorithms. Before the final input files were prepared, the raw data from the training files was examined for errors and outliers within each attribute set. An



experienced data analyst in bench stock inventory management identified outliers through the analysis of the attribute distributions. The errors were removed from the training sets, as well as the outliers. The two training sets were then formatted as input files and examined by the six data mining algorithms within WEKA to create twelve data mining models for re-evaluation against the test data sets. The six data mining methods tested are implemented within WEKA as the following java classes:

1. weka.classifiers.bayes.BayesNet,
2. weka.classifiers.bayes.NaiveBayes,
3. weka.classifiers.functions.MultilayerPerceptron,
4. weka.classifiers.functions.SMO,
5. weka.classifiers.functions.Logistic,
6. weka.classifiers.trees.J48.

4.3 Data mining

After the models for each of the combinations of the two training sets and six data mining algorithms were built, each model was re-evaluated against the test sets. Seven test sets from each of the two types of data extracts (one associated with known future stock shortages and the other an unrestricted test set) were evaluated against each data mining model. The results were recorded and processed for further comparison analysis.

The predictive accuracy is not measured with respect to the class assigned to the data, but with respect to the actual event whose occurrence is being forecasted; in this case a future record of a stock shortage. In the case of evaluating the predictions, the classification errors were the focus of analysis. If a classifier changed a record classification from a prior inaccurate position contraindicating a stock shortage to a class annotating an imminent shortage, then the classifier was considered accurate to the degree of the classification change. Four classes were modelled by the data: A, B, C, and D. These classes are levels-of-attention that any given stock location should receive from an order expeditor. An attention class of A refers to a stock shortage condition. Class B refers to a stock outage or a high probability for stock shortage in the near future. Attention class C indicates that the stock location replenishment cycle may be conducive to a stock shortage. The class D stock locations are those not requiring any attention and are being handled by the supply chain systems adequately.

It is also important to note that it is highly likely that the data attributes do have dependencies between them. The dates within the transaction stream are temporally dependent and several “wellness” indicators are recorded at times corresponding to the action date timestamps in the transaction stream being mined. The class of each supply-chain segment is also defined by a complex combination of the data attributes. This fact impacts the validity of using the Naïve Bayesian method because of the attribute independence assumption. However, Bayesian networks allow for modelling of arbitrarily complex dependencies between attributes [9], which is why both the Naïve Bayesian and Bayesian Network methods were evaluated.



5 Results of analysis

Using the actual stock shortage information, the predictions of each of the data mining algorithms were analyzed using selected typical measures of data mining performance, and also using a prediction deviation measure. This measure rewarded predictions when the prediction was closer to the actual state of the inventory following the supply chain segment sample; and penalized the prediction when it moved away from the actual future state of the inventory. Because the stock shortage test data contained only sample data that would be considered positive results, the measures of data mining performance that are calculated using false positives and true negatives could not be used. The purpose of using the stock shortage test sets was to show, with relatively little data extraction effort, that the data mining methods tested could predict the occurrence of a stock shortage when the data immediately preceding the shortage did not necessarily indicate an imminent problem.

Table 2: Comparison results.

REGWQ Group	Mean	Mining Algorithm	REGWQ Group	Mean	Mining Algorithm
Stock Shortage Recall					
Stock Shortage Test Data			Unrestricted Test Data		
A	33.9	Bayes Network	A	43.0	Logistic
B A	30.5	J48Tree	A	43.0	MLP
B A	28.8	MLP	A	42.8	Bayes Network
B A	26.5	Naïve Bayes	A	42.4	J48Tree
B A	24.5	Logistic	A	38.8	Naïve Bayes
B	23.5	SMO	A	38.8	SMO
Success Rate Comparison with Unrestricted Test Data					
Unrestricted Training Set			Stock Shortage Training Set		
A	0.94	Naïve Bayes	A	0.93	SMO
A	0.93	SMO	A	0.93	Logistic
B A	0.91	Logistic	B A	0.89	MLP
B C	0.88	MLP	B A	0.88	Bayes Network
C	0.87	J48Tree	B	0.86	J48Tree
C	0.84	Bayes Network	C	0.82	Naïve Bayes
Precision Comparison with Unrestricted Test Data					
Unrestricted Training Set			Stock Shortage Training Set		
A	22.6	Naïve Bayes	A	16.7	SMO
B A	15.9	SMO	A	16.0	Logistic
B A	14.1	Logistic	B A	10.0	MLP
B	8.7	MLP	B A	8.9	Bayes Network
B	8.2	J48Tree	B A	8.2	J48Tree
B	7.0	Bayes Network	B	6.3	Naïve Bayes
Prediction Deviation Comparison with Unrestricted Test Data					
Unrestricted Training Set			Stock Shortage Training Set		
A	0.003	Naïve Bayes	A	0.003	SMO
B A	-0.009	SMO	B A	-0.009	Bayes Network
B C	-0.019	MLP	B A	-0.014	MLP
B C D	-0.023	Logistic	B C	-0.021	Logistic
C D	-0.030	Bayes Network	D C	-0.040	J48Tree
D	-0.041	J48Tree	D	-0.048	Naïve Bayes

With the tests conducted using the stock shortage restricted test data, the interaction was not significant for any combination of the mining methods and training data. However, when using the unrestricted test data, the Naïve Bayes algorithm was significantly affected for certain performance measures. In every metric where an interaction was found between the training set and the mining algorithm, the Naïve Bayes algorithm performed significantly better when the unrestricted training data was used to train the Naïve Bayesian model.

The Ryan-Einot-Gabriel-Welsch (REGWQ) multiple range test was used to perform the multiple means comparisons within the experiment. The results of the test runs are shown in table 2. The grouping denoted by the same letter between the different mining algorithms indicates that there is no statistically significant difference between the algorithms in the group.

Examining the stock shortage recall metric for the stock shortage test data, percentage predicted correctly is well above 20 percent. This is actually very useful when it may equate to preventing one of five stock shortages, predicted on a given day, through early intervention by an inventory manager. The stock shortage recall is even higher in the unrestricted test cases and all of the mining algorithms perform equally well. The sequential minimal optimization (SMO) implementation of the support vector machine (SVM) appears to be the most consistent performer when the unrestricted test set data is examined.

The results of the experiment also show that the Naïve Bayes method worked very well when the unrestricted training set was used. This was somewhat unexpected because of the assumed dependencies among the data attributes. In fact, the method intended to counteract these assumed dependencies, the Bayes Network, performed significantly poorer when the unrestricted training set was used.

6 Future research

Throughout this research a variety of other research avenues appeared and had to be set aside for the future. The degree of feature selection performed for this study was a first pass solution and additional analysis is needed toward selecting the most productive data attributes from the transaction stream for training and prediction. The algorithms supporting the data mining methods used here were all executed using the default tuning parameters provided within the WEKA data mining workbench version 3.4.3 [1]. Additional testing and experimentation is required of the data mining algorithms to establish an optimal set of algorithm parameters.

Additional research focusing on the temporal nature of the transactions with respect to demand seasonality encoded within the data may also produce interesting and useful results. Combining this research with a more extensive evaluation of specific data mining with the goal of predicting rare events could also lead to usable results.



7 Conclusion

Through the research presented here, six data mining methods were applied to a transaction data set generated, in the course of maintenance operations, from a large bench stock inventory management system. Samples of data were extracted from the data warehouse populated by these transactions and used as data mining classification training sets and also test data sets. Of particular interest was the evaluation of the data mining techniques with respect to their ability to use the classification models built using the training sets to identify stock location status represented in a single day transaction test set. The results clearly show that using any of the data mining techniques evaluated will provide predictions of near term stock shortages that are not manifest in the most current inventory status records. Such information is very important to inventory managers and to maintenance operations success. Development and enhancement of this research holds great promise toward bringing critical, timely information to those tasked with keeping maintenance and production activities, supported by large, complex inventories, flowing.

References

[1] Witten, I.H. & Frank, E., *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations,* Morgan Kaufmann, 1999.

[2] Dhond, A., Gupta, A., & Vadhavkar, S., Data mining techniques for optimizing inventories for electronic commerce, *Proc. 6th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, August 2000.

[3] Brause, R., Langsdorf, T., & Hepp, M., Neural data mining for credit card fraud detection", *Tools with Artificial Intelligence*, 1999, and *Proc. of the 11th IEEE Int. Conf.*, pp. 103-106, 9-11 November 1999.

[4] Vilalta, R. & Ma, S., Predicting rare events in temporal domains, *ICDM 2002 Proc, 2002 IEEE Int. Conf. on Data Mining*, pp.474 – 481, 9-12 December 2002.

[5] Wai-Ho Au, C., Keith C.C., & Yao, X., A novel evolutionary data mining algorithm with applications to churn prediction, *IEEE Transactions on Evolutionary Computation*, **7(6)**, pp.532-545, 2003.

[6] Wilson, A.J., Morgenstern, M.P., Pfahringer, B., & Leschi, C., Data mining bread quality and process data in a plant bakery, *Proc. of the 12th ICC Cereal and Bread Congress*, Harrogate, UK, May 2004.

[7] Yohda, M.; Saito-Arita, M., Okada, A., Suzuki, R., & Kakemoto, Y., Demand forecasting by the neural network with discrete fourier transform, *Data Mining*, 2002, *Proc. of the ICDM* 2002 and 2002 *IEEE International Conference*, pp.779 – 782, 9-12 December 2002.

[8] Piramuthu, S., The Hausdorff distance measure for feature selection in learning applications, *Proc. of the 32nd Annual Hawaii Int. Conf. on System Sciences* (*HICSS-32)*, 1999.

[9] Wang, Z., & Webb, G.I., Comparison of lazy Bayesian rule, and tree-augmented Bayesian learning, *Proc. of the 2002 IEEE Int. Conf. on Data Mining (ICDM 2002)*, pp.490 – 497, 9-12 December 2002.