# ENHANCING CLUSTERING-BASED CLASSIFICATION ALGORITHMS IN E-COMMERCE APPLICATIONS

[1] **AYMAN MOHAMED MOSTAFA, [1]MOHAMED MAHER, [1]M.M. HASSAN**

[1]Faculty of Computers and Informatics, Zagazig University, 44519, Egypt

E-mail:  [1]am_mostafa@zu.edu.eg

## ABSTRACT

Data mining algorithms are used for analyzing data from different sources and extracting useful information from a large volume of data. Algorithms of data mining are used in E-commerce companies to help them identifying online customer behavior to recommend appropriate products based on customers' needs. In this paper, our aim is enhancing the result of the classification techniques that applied to an online shopping agency dataset by using clustering techniques which applied to this dataset before entering it to classification techniques, so farthest first, expectation maximization (EM), and K-mean clustering algorithms are applied to an online shopping agency dataset to allocate related objects into the same cluster. After applying clustering algorithms, a group of data mining classification algorithms such as Bayes net, Naïve Bayes, K star, filtered classifier, decision table, J48, and JRIP are applied to the three clustering algorithms. A logistic model tree (LMT) classification algorithm is applied also to measure the performance parameters for each classifier. The experimental results achieved high rates in accuracy, precision, recall, F-measure, and ROC when compared to recent research paper.

**Keywords**: *Data mining, Classification, Clustering, Logistic Model Tree, and E-commerce*

## 1. INTRODUCTION

In a short period of internet appearance, a rapid evolution in information technology is emerged that help people for using and accessing information in e-commerce. E-commerce with the internet certainly consider an important input for the success of any enterprise and offering massive opportunities and worldwide markets [1].

E-commerce sites are important sales channels. The aspect of most business functions is changed due to an appearance of e-commerce in competitive enterprises. It is very important to use data mining methods to analyze data that is carried out by persons who have visited these websites. In general, e-commerce has enabled online transactions and making data is not easy in real time [2].

Data mining is the art that has some of the techniques which are used for extracting non-obvious, useful information from a huge database or a large amount of data. Data mining techniques are divided into several categories such as clustering techniques, classification techniques, association rules techniques [3].

Clustering is used to divide the similar data into a cluster and dissimilar data into another cluster. Clustering techniques are divided into several categories: partitioning algorithms, hierarchal algorithms, density-based clustering algorithms, farthest first cluster and filtered clustering [3].

Classification is used to find a model that described classes and concept. Classification techniques are divided into several categories: Bayes net, Naïve bays, IBK, K star, LWL, Classification via regression, filtered classifier, Randomizable filtered classifier, Input Mapped classifier, Decision table, JRip, J48, LMT [4].

Data mining inside of e-Commerce is used to make better decision making for a new idea or a new integrated technology. Data mining create a way for decision-makers to be able to make their decision more effective for the improvement of their business [5].

Weka is a software written in Java and contains several GUI such as Explorer, Experimenter, Knowledge Flow and Simple CLI. Weka has used data mining tools for comparing between clustering and classification techniques [6].

The aim of this paper is proving that when applying clustering algorithms to an online shopping agency dataset before applying classification algorithms will give best results compared to a recent research paper [7] which applied classification algorithms only to the same dataset.

The contribution of this paper is as the follows: introduction, related work, clustering algorithms, classification algorithms, dataset description, the experimental results, and conclusion.

## 2. RELATED WORK

This section has some related work that uses data mining inside of e-commerce and online store websites.

As presented in [4], a novel method is used for combining between clustering with classification algorithms due to improving classification accuracy. The experimental results and analysis show that classification algorithm accuracy can be improved by applying classification techniques after clustering algorithms.

As presented in [5], Some benefits and some challenges in data mining inside of e-commerce websites are presented in this paper. There are some benefits for e-commerce companies which used data mining algorithms on their websites such as analyzing purchases behavior for a customer, forecasting the new sales and allow these companies to make merchandise planning. On the other hand, there are some challenges f for e-commerce companies which used data mining algorithms on their websites such as the transformation of data, data mining scalability and spider identification.

As presented in [7], a comparative study of some classification algorithms is presented in this paper which applied to the online shopping agency dataset. The decision table classifier shows in the experimental result as the greater algorithm which can be applied on the online shopping agency to provide the best implementation of a powerful model that will allow a customer to determine their needs from products which produced by e-commerce websites.

As presented in [8], a comparative study of some clustering algorithms is presented in this paper by using weka tool. The conclusion of this paper after analyzing the results of testing the clustering algorithms shows that the accuracy of the k-mean algorithm is greater than the hierarchical clustering algorithm and the quality clusters are produced by using k-mean algorithm when using a large dataset.

## 3. CLUSTERING ALGORITHMS

The main objective of using clustering algorithms is to allocate related or similar objects into groups called clusters. Objects that are similar to each other are placed into the same cluster. As presented in Table 1, Farthest First Cluster, EM Cluster, and K-Mean Cluster algorithms are used on the online shopping agency dataset. These clustering algorithms are presented in [3, 6]. Selected Clustering Algorithms used on Dataset are shown in Table 1.

*Table 1: Selected Clustering Algorithms*

| Cluster Name | Description |
|---|---|
| EM Cluster | Expectation Maximization algorithm is used in statistical models into which the model depends on unobserved implicit variables for finding the maximum probability or minimum a posteriori estimates of parameters. |
| Farthest First Cluster | Farthest First is an algorithm that applies each cluster center in turn at the point that is far from the current cluster center. |
| K Mean Cluster | K Mean is an unsupervised learning method used for classifying data into k clusters and outcomes of those clusters are independent of each other. |

## 4. CLASSIFICATION ALGORITHMS

After applying clustering algorithms on the dataset and saving the result, a set of classification algorithms are applied on the clustered data to improve the accuracy of the classifier.

As presented in Table 2, the applied classifications algorithms are Bayes Net, Naïve Bays, K Star, Filtered Classifier, Decision Table, JRip, J48 [7, 9, and 10] and logistic model tree (LMT). Time and accuracy are taken to implement E-commerce model. Selected classification Algorithms used on Dataset are shown in Table 2.

*Table 2. Selected Classification Algorithms*

| Classifier | It's Description |
|---|---|
| Bayes Net | Bayesian Network is a method which received a great attention from users for using in the classification process. |
| Naïve Bays | Naïve Bays classifier is an algorithm that uses statistical methods for making a classification process on selected data. |
| K Star | K Star classifier is the class of a test instance is based on the class which training instances similar to it as determined by the similarity function. |
| Filtered Classifier | The filtered classifier is an algorithm that implemented for an arbitrary classifier on selected data which pushed through an arbitrary filter. |
| Decision Table | Decision table classifier is an algorithm that contains a class for creating and utilizing decision table classifier. |
| JRip | J Repeated Incremental Pruning classifier is an algorithm which contains a some of the rules of a class that is created by using reduced error Jrip (RIPPER). |
| J48 | J48 is a Classifier that accepts nominal classes only. |
| LMT | LogisticModelTree classifier is a classification tree with logistic regression function at the leaves. |

## 5.  DATA DESCRIPTION

The comparative study between some classification algorithms that was presented in [7] is applied on online shopping agency dataset to allow this enterprise identifying the best suitable classifier algorithm to this dataset which was the decision table classifier. Applying decision table classifier on this dataset will allow the enterprise implementing a powerful model that can help customers to determine their needs from products presented in E-commerce websites.

In this paper classification algorithms are applied after applying some clustering algorithms on the same dataset to improve the performance and accuracy of the results.  Dataset contains some attributes as follows: serial no., buyer_name, gender, age, educational_level, brand, product_name, item_description, category, quantity, price, item_Type, payment_method, no. of visits, rating, user_satisfaction of the product, best_deal, no. of likes, positive_comments, negative_comments, no. of posts, Facebook, Instagram, Twitter, and overall_user_process_satisfaction.

## 6. EXPERIMENTAL RESULTS AND ANALYSIS

The performance and accuracy of the implemented model can be improved by applying classification algorithms after applying clustering algorithm on online shopping agency dataset. Figure1 and Figure2 shows Weka explorer interface that includes clustering algorithm on the dataset and Weka explorer interface that include classification algorithm on the results of the clustered data (order_log.arff).

```
Facebook
  mean                    1008.4926  962.2634 1073.9922  991.2186 1017.0598
  std. dev.                565.084    536.91  478.9734  531.7276  567.1299

Instagram
  mean                     809.285   705.6923  781.5352  803.8003  732.7186
  std. dev.                444.3261  438.3155  411.795   443.217   392.5372

Twitter
  mean                     543.3357  590.8066  540.0474  517.3905  536.1313
  std. dev.                246.761   279.8042  266.0231  246.8309  254.7514

UserProcessSatisfaction
  mean                      49.7818   52.1904   43.5798   51.9835   48.6479
  std. dev.                 27.398    28.6691   27.9791   30.378    28.7927


Time taken to build model (full training data) : 12.6 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      58 ( 17%)
1      78 ( 23%)
2      60 ( 17%)
3      81 ( 24%)
4      66 ( 19%)


Log likelihood: -98.82457
```

*Figure 1: Clustering Results of Online Shopping Agency Dataset*

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances      343             100        %
Incorrectly Classified Instances     0               0         %
Kappa statistic                      1
Mean absolute error                  0.0362
Root mean squared error              0.0613
Relative absolute error             11.373 %
Root relative squared error         15.3546 %
Coverage of cases (0.95 level)      100      %
Mean rel. region size (0.95 level)  42.1574 %
Total Number of Instances           343

=== Detailed Accuracy By Class ===
```

| | TP Rate | FP Rate | Precision | Recall | F-Measure | MCC | ROC Area | PRC Area | Class |
|---|---|---|---|---|---|---|---|---|---|
| | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | cluster0 |
| | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | cluster1 |
| | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | cluster2 |
| | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | cluster3 |
| | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | cluster4 |
| Weighted Avg. | 1.000 | 0.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | |

```
=== Confusion Matrix ===

 a  b  c  d  e  <-- classified as
58  0  0  0  0 | a = cluster0
 0 78  0  0  0 | b = cluster1
 0  0 60  0  0 | c = cluster2
 0  0  0 81  0 | d = cluster3
 0  0  0  0 66 | e = cluster4
```



*Figure 2: Classification Results of Clustered Data*

The experimental results that were presented in [7] showed that the classification processes are achieved without applying clustering algorithms. As presented in Figure. 3, the decision table classifier achieved 87.1% for true_positive_rate (TP_rate), 7.8% for false_positive_rate(FP_rate), 88.3% for precision, 87.1% for recall, 87.2% for F-measure, and 95.5% for ROC area. The second-best results presented in [8] were achieved by filtered classifier algorithm that showed 86.8% for true_positive_rate (TP_rate), 7.9% for false_positive_rate(FP_rate), 87.8% for precision, 86.8% for recall, 87.2% for F-measure, and 95.2% for ROC area. Whereas the third best results were achieved by J48 algorithm that shows 86.1% for true_positive_rate (TP_rate), 8.6% for false_positive_rate(FP_rate), 86.2% for precision, 86.1% for recall, 85.9% for F-measure, and 88% for ROC area.
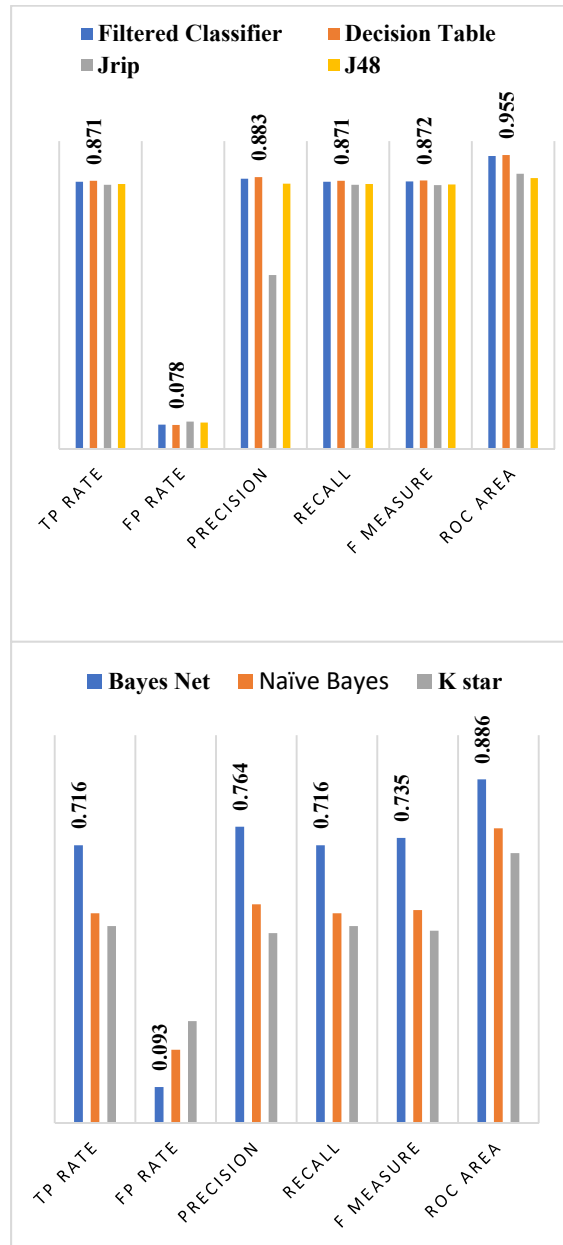


*Figure 3: Performance Parameters of Classifier Algorithms*

## 6.1 Performance Parameters Classifiers after Farthest First Clustering

The logistic model tree (LMT) classifier was not used in the previous experimental results that were presented in [8]. In this paper, the LMT classifier will be added to a group of classification algorithms after applying clustering algorithms to improve the performance and accuracy better than applying classification algorithms only on the same dataset.

As presented in Figure 4, the classification processes are achieved after applying Farthest First clustering algorithm. The LMT classifier achieved 93% for true positive rate (TP rate), 15.4% for false positive rate (FP rate), 93% for precision, 93% for recall, 92.8% for F-measure, and 98.1% for ROC area. Whereas the second-best results are achieved by Naïve Bayes classifier. The NB classifier achieved 80.8% for true positive rate (TP rate), 30.8% for false positive rate (FP rate), 80.5% for precision, 80.8% for recall, 80.6% for M-measure, and 85.9% for ROC area. The third best results are achieved by Decision Table classifier. The DT classifier achieved 80.2% for true positive rate (TP rate), 40.5% for false positive rate (FP rate), 79.1% for precision, 80.2% for recall, 78.8% for M-measure, and 79% for ROC area.
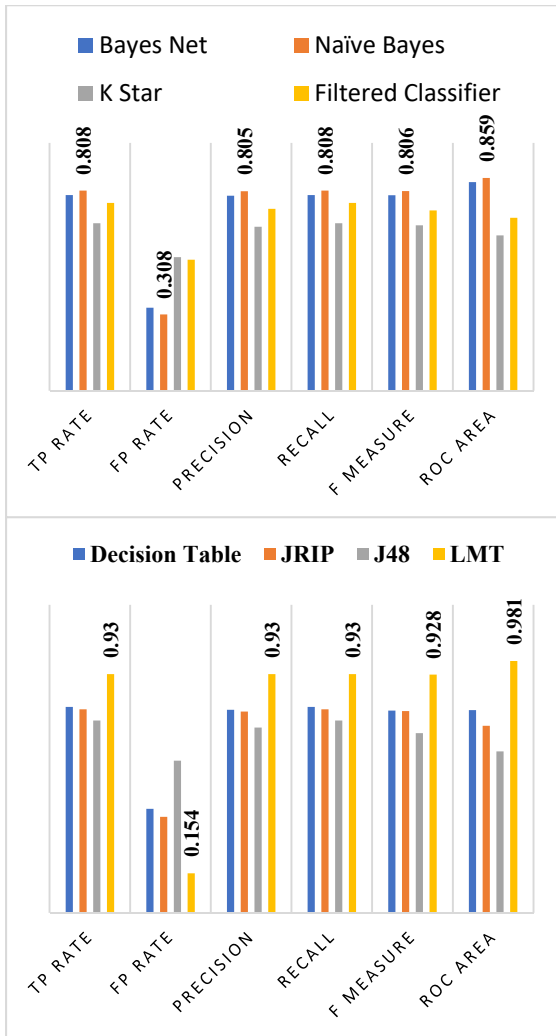
All experimental results of classification algorithms after applying Farthest First clustering are shown in Table 3.

*Table 3. Performance after Farthest First Cluster*

| Classifier Algorithm | TP-Rate | FP-Rate | Precision | Recall | F-Measure | ROC Area |
|---|---|---|---|---|---|---|
| Bayes Net | 0.79 | 0.33 | 0.78 | 0.79 | 0.78 | 0.84 |
| Naïve Bayes | 0.80 | 0.30 | 0.80 | 0.80 | 0.80 | 0.85 |
| K* - Star | 0.67 | 0.53 | 0.66 | 0.67 | 0.66 | 0.62 |
| Filtered Classifier | 0.75 | 0.52 | 0.73 | 0.75 | 0.72 | 0.69 |
| Decision table | 0.80 | 0.40 | 0.79 | 0.80 | 0.78 | 0.79 |
| JRIP | 0.79 | 0.37 | 0.78 | 0.79 | 0.78 | 0.72 |
| J48 | 0.74 | 0.59 | 0.72 | 0.74 | 0.70 | 0.62 |
| LMT | **0.93** | **0.15** | **0.93** | **0.93** | **0.92** | **0.98** |

## 6.2 Performance Parameters Classifiers after EM Clustering

As presented in Figure 5, the classification processes are achieved after applying Expectation Maximization (EM) clustering algorithm. The LMT classifier achieved 100% for true positive rate (TP rate), 0.0% for false positive rate (FP rate), 100% for precision, 100% for recall, 100% for F-measure, and 100% for ROC area.

Whereas the second-best results are achieved by Naïve Bayes classifier. The NB classifier achieved 98.5% for true positive rate (TP rate), 3% for false positive rate (FP rate), 98.7% for precision, 98.5% for recall, 98.5% for M-measure, and 100% for ROC area. The third best results are achieved by Bayes Net classifier.

This classifier achieved 98.5% for true positive rate (TP rate), 3% for false positive rate (FP rate), 98.6% for precision, 98.5% for recall, 98.5% for M-measure, and 100% for ROC area.



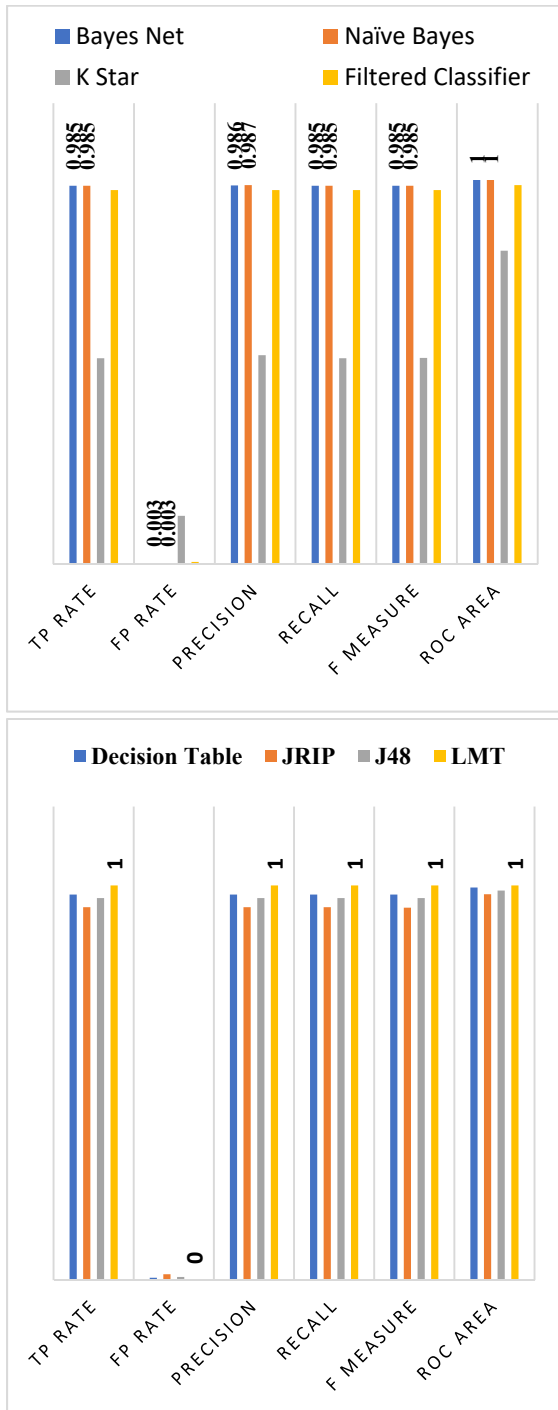*Figure 4: Performance Classifiers after Farthest First Clustering*

*Figure 5: Performance Classifiers after EM Clustering*

All experimental results of classification algorithms after applying Farthest First clustering are shown in **Table 4**.

*Table 4. Performance after EM Cluster*

| Classifier Algorithm | TP-Rate | FP-Rate | Precision | Recall | F-Measure | ROC Area |
|---|---|---|---|---|---|---|
| Bayes Net | 0.98 | 0.003 | 0.98 | 0.98 | 0.98 | 1.00 |
| Naïve Bayes | 0.98 | 0.003 | 0.98 | 0.98 | 0.98 | 1.00 |
| K* - Star | 0.53 | 0.126 | 0.54 | 0.53 | 0.53 | 0.81 |
| Filtered Classifier | 0.97 | 0.006 | 0.97 | 0.97 | 0.97 | 0.98 |
| Decision table | 0.97 | 0.006 | 0.97 | 0.97 | 0.97 | 0.99 |
| JRIP | 0.94 | 0.015 | 0.94 | 0.94 | 0.94 | 0.97 |
| J48 | 0.96 | 0.008 | 0.96 | 0.96 | 0.96 | 0.98 |
| LMT | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 |

## 6.3 Performance Parameters Classifiers after K-Mean Clustering

As presented in **Figure. 6**, the classification processes are achieved after applying K Mean clustering algorithm. The LMT classifier achieved 99.4% for true_positive_rate (TP_rate), 0.6% for false_positive_rate(FP_rate), 99.4% for precision, 99.4% for recall, 99.4% for F-measure, and 100% for ROC area. Whereas the second-best results are achieved by JRIP classifier. The JRIP classifier achieved 95% for true_positive_rate (TP_rate), 4.9% for false_positive_rate(FP_rate), 95.3% for precision, 95% for recall, 95% for M-measure, and 96.8% for ROC area. The third best results are achieved by Decision Table (DT) classifier. The DT classifier achieved 94.8% for true_positive_rate (TP_rate), 5.5% for false_positive_rate(FP_rate), 94.8% for precision, 94.8% for recall, 94.8% for M-measure, and 96.1% for ROC area.
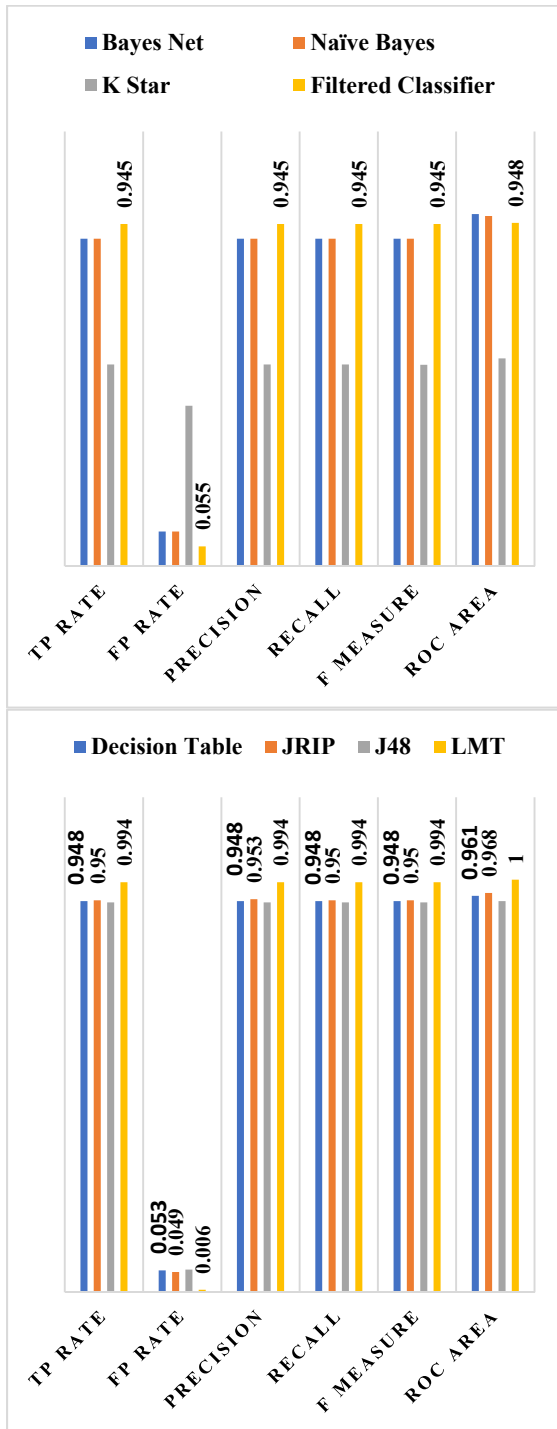
*Figure 6: Performance Classifiers after K-Mean Clustering*

All experimental results of classification algorithms after applying Farthest First clustering are presented in **Table 5**.

*Table 5. Performance after K-Mean Cluster*

| Classifier Algorithm | TP-Rate | FP-Rate | Precision | Recall | F-Measure | ROC Area |
|---|---|---|---|---|---|---|
| Bayes Net | 0.90 | 0.096 | 0.90 | 0.90 | 0.90 | 0.97 |
| Naïve Bayes | 0.90 | 0.096 | 0.90 | 0.90 | 0.90 | 0.96 |
| K* - Star | 0.55 | 0.443 | 0.55 | 0.55 | 0.55 | 0.57 |
| Filtered Classifier | 0.94 | 0.055 | 0.94 | 0.94 | 0.94 | 0.94 |
| Decision table | 0.94 | 0.053 | 0.94 | 0.94 | 0.94 | 0.96 |
| JRIP | 0.95 | 0.049 | 0.95 | 0.95 | 0.95 | 0.96 |
| J48 | 0.94 | 0.055 | 0.94 | 0.94 | 0.94 | 0.94 |
| LMT | **0.99** | **0.006** | **0.99** | **0.99** | **0.99** | **1.00** |

Based on the previous experimental results, **Table 6** explains the best TP, FP, precision, recall, F-measure, and ROC area based on applying classification algorithms after clustering step.

*Table 6. Best Clustering Performance*

| Rank | Classifier Algorithm | Cluster Algorithm |
|---|---|---|
| 1 | LMT | EM |
| 2 | LMT | K-Mean |
| 3 | NB | EM |
| 4 | Bayes Net | EM |
| 5 | JRIP | K-Mean |
| 6 | DT | K-Mean |
| 7 | LMT | Farthest First |

## 7.   CORRECT / INCORRECT PARAMETERS

The correct / incorrect parameters of the presented clustering algorithms are compared to the parameters of the classifiers presented in [7]. In this section the correct / incorrect parameters of classification algorithms are measured after the applying the three clustering algorithms. This is presented in the following section.

### 7.1 Correct/Incorrect Parameters after Farthest First Cluster

As presented in **Fig. 7**, the highest new correct of classifier accuracy after applying farthest first cluster achieved a rate of 93% by LMT classifier while the lowest new incorrect of classifier accuracy after applying farthest first cluster achieved 6.9% rate by LMT classifier. While the decision table (DT) classifier that was presented in [7] achieved the highest correct of the classifier accuracy after applying only classification processes on the dataset was equal to 87.13% and the lowest incorrect of classifier accuracy achieves 12.87%.
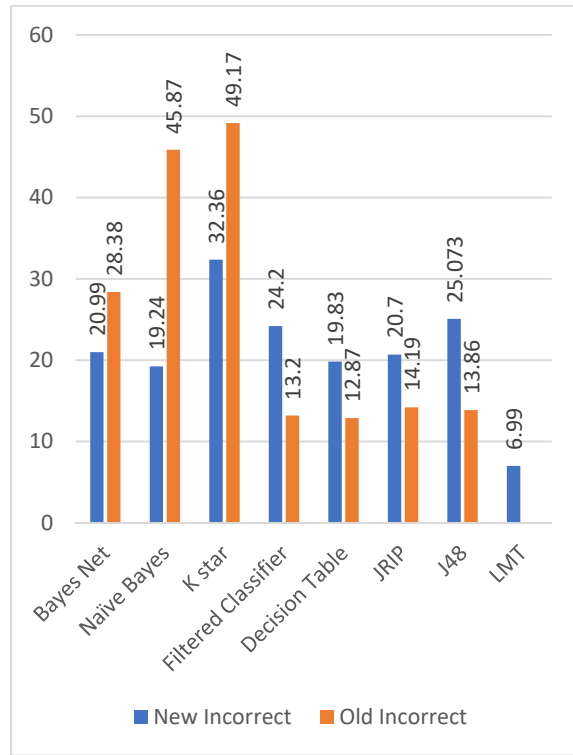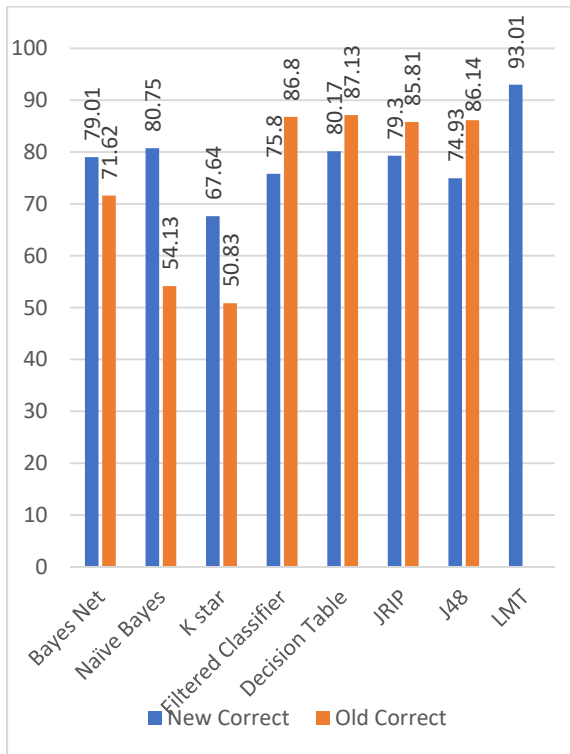
*Figure 7: Classifier Accuracy after Farthest First Cluster*

### 7.2   Correct/incorrect Parameters after EM Cluster

As presented in **Figure 8**, the highest new correct of classifier accuracy after applying EM clustering achieved 100% by LMT classifier and the lowest new incorrect accuracy achieved 0.0% by LMT classifier. While the decision table (DT) classifier that was presented in [7] achieved the highest correct of the classifier accuracy after applying only classification processes a rate of 87.13% and the lowest incorrect accuracy achieved 12.87%.
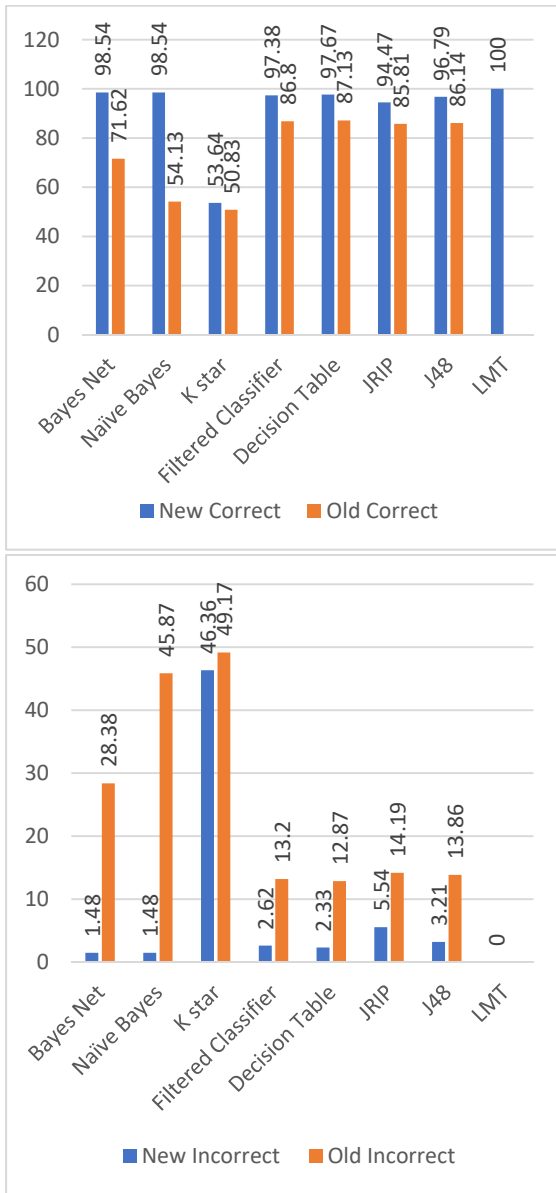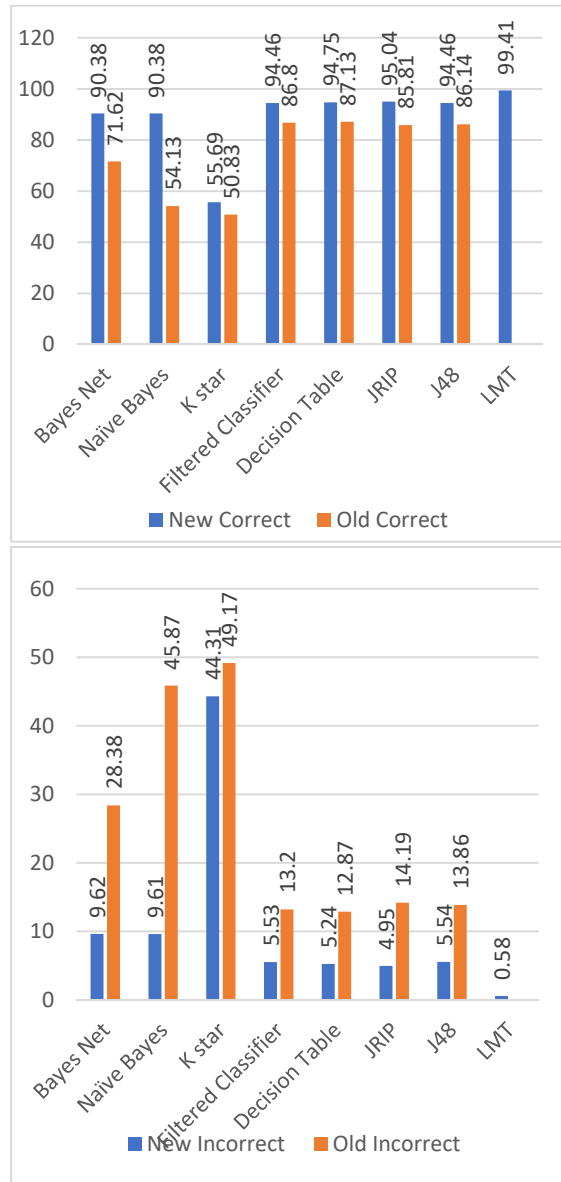
*Figure 8: Classifier Accuracy after EM Cluster*



*Figure 9: Classifier Accuracy after K-Mean Cluster*

### 7.3 Correct/incorrect Parameters after K-Mean Cluster

As presented in **Figure 9**, the highest new correct of classifier accuracy after applying K-Mean clustering achieved 99.41% by LMT classifier and the lowest new incorrect accuracy achieved 0.58% by LMT classifier also. While the decision table (DT) classifier that was presented in [7] achieved the same accuracy for both correct and incorrect parameters by 87.13% and 12.87% respectively.

## 8.    EXECUTION TIME PERFORMANCE

The execution time is measured by applying the three clustering algorithms on all classifiers. As presented in **Figure 10**, after applying Farthest First cluster, the best classifier time was recorded with Bayes Net, Naïve Bayes, and K Star with an execution time of 0 second. After applying EM cluster, the best classifier time was K Star and Naïve Bayes with execution time of 0 and 0.02 respectively. After applying K-Mean cluster, the best classifier time was recorded with Naïve Bayes,

K Star, and Filtered classifier with an execution time of 0 second.

But after applying only classification operations [7], the K Star classifier was the fastest classifier with a total time of 0 second while the latest classifier is Bayes Net that achieved a total time of 0.28 second.
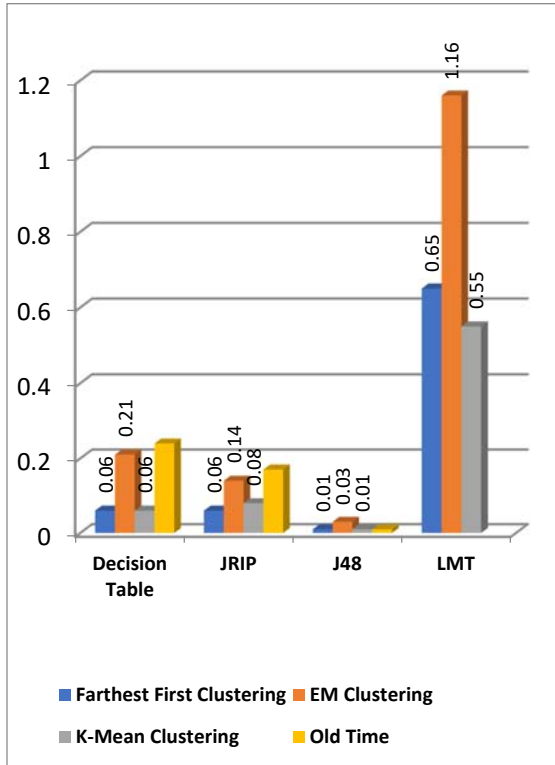


*Figure 10: Execution Time after Clustering Process*

From the experimental results presented in Figure 10, the LMT classifier achieved 0.65 second with Farthest First cluster, 1.16 second with EM cluster, and 0.55 second with K-Mean cluster.

## 9.  CONCLUSION AND FUTURE WORK

In this paper, eight data mining classification techniques are applied after three data mining clustering techniques severally to improve the performance and accuracy of classification techniques with the ability to make the comparative test to find the best result of combining clustering techniques with classification techniques. The result shows that LMT classifier applied after EM cluster algorithm gives the highest accuracy of 100% and k Star classifier after EM cluster

algorithm gives the lowest accuracy of 53%. This result can help e-commerce company by selecting the optimal clustering algorithm followed by classification algorithm which was EM clustering followed by LMT classifier suitable to order log dataset which allows the enterprise for implementing a powerful model. This can allow a customer to determine their needs from products which produced by E-commerce websites. Future work includes reducing the execution time of LMT classifier which gives the best results but takes the largest execution time more than other classifier algorithms.

## REFRENCES

[1] Hrudaya Ku. Tripathy and B.K.Tripathy, "A Rough Set Approach for Clustering the Data Using Knowledge Discovery in World Wide Web for E-Business", IEEE International Conference on e-Business Engineering, 2007, pp 717-722.

[2] Ahmad Tasnim Siddiqui and Sultan Aljahdali, "Web  Mining Techniques in E-Commerce Applications", International Journal of Computer Applications (0975 – 8887) Vol. 69, No.8, 2013, pp. 39-43.

[3] Archana and Jitendra Kumar "Comparative Study of Clustering Techniques", journal of current engineering research, Vol. 4, No. 3, May-June, 2014, pp.7-10.

[4] Yaswanth Kumar Alapati and Korrapati Sindhu, "Combining Clustering with Classification:A Technique to Improve Classification Accuracy", International Journal of Computer Science Engineering (4E), Vol. 5, No.06, Nov, 2016, pp. 336-338.

[5] Mustapha Ismail, Mohammed Mansur Ibrahim, Zayyan Mahmoud Sanusi and Muesser Nat, "Data Mining in Electronic Commerce: Benefits and Challenges", Int. J. Communications, Network and System Sciences, No. 8, 28 December, 2015, pp. 501-509.

[6] Narendra Sharma, Aman Bajpai and Mr. Ratnesh Litoriya, "Comparison the various clustering algorithms of weka tools", International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, Vol. 2, No. 5, May, 2012, pp. 73-80.

[7] Rana Alaa El-Deen Ahmeda, M.Elemam.Shehaba, Shereen Morsya and Nermeen Mekawie, "Performance study of classification algorithms for consumer online shopping attitudes and behavior using data mining", International Conference on Communication Systems and Network Technologies, 2015, pp. 1344-1349.

[8] Bharat Chaudhari and Manan Parikh, "A Comparative Study of clustering algorithms Using weka tools", International Journal of Application or Innovation in Engineering & Management (IJAIEM), Vol. 1, No. 2, October, 2012, pp. 154-158.

[9] Meenakshi and Geetika, "Survey on Classification Methods using WEKA", International Journal of Computer Applications (0975 – 8887), Vol. 86, No. 18, January, 2014, pp. 16-19.

[10] Rafet Duriqi, Vigan Raca and Betim Cico "Comparative Analysis of Classification Algorithms on Three Different Datasets using WEKA", 5th Mediterranean Conference on Embedded Computing, 2016, pp. 335-338.