AN ENHANCEMENT IN EFFECTS OF FINANCIAL DEVELOPMENT ON FOREIGN DIRECT INVESTMENTS BY APRIORI OVER GENETIC ALGORITHM

Rajesh.A[1], S.Ashok Kumar[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The purpose of this study is an enhancement in the effects of financial development on foreign direct investments by using Apriori algorithm in comparison with Genetic algorithm (GA). **Materials & Methods:** The dataset in this paper utilizes the publicly available dataset from National Financial Development to prove the effectiveness of the approach. The study's sample size was 280 (Groups 1 and 2 each had 140 participants), and the calculation was done using G-power 0.8 with alpha and beta values of 0.05 and 0.2, respectively, with a confidence interval of 95%.The enhancement in effects of financial development on foreign direct investments is performed by the Apriori algorithm whereas number of samples (N=10) and Genetic algorithm (GA) where number of samples (N=10).
Results: The accuracy rate of the Apriori algorithm classifier is 95.68 percent greater than the accuracy rate of the Genetic algorithm (GA), which is 93.263 percent. The results, which have a significance level of p=0.0341 ($p<0.05$), indicate that there is statistical significance for two groups.
Conclusion: The Apriori algorithm provides better results in terms of accuracy compared to the Genetic Algorithm (GA) for enhancing the impact of financial development on foreign direct investment.

Keywords: Financial Development, Novel Apriori algorithm, Genetic algorithm (GA), Accuracy rate, foreign direct investments, data mining technique

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India, pin: 602105
[2*]Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamil Nadu, India, pin: 602105

1. Introduction

Over the past two decades, foreign direct investment (FDI) has become increasingly important in developing countries, and an increasing number of developing countries have managed to attract significant and increasing amounts of inward FDI (Haudi, Wijoyo, and Cahyono 2020) . Economic theory has identified a number of channels through which FDI inflows can benefit the host economy. The conventional method lagged behind and had more trouble showing these benefits in practice (Azhar and Kn 2012). Most notably, a large number of applied papers have examined the relationship between FDI and GDP growth, but their results have been far from conclusive (Alfaro 2003). Despite the absence of any firm conclusions, and surprisingly, most countries continue to vigorously pursue policies aimed at encouraging greater FDI inflows (Alfaro et al. 2004). This paper examines the impact of foreign direct investment (FDI) on economic growth using a new Apriori-based data mining algorithm for FDI inflows into India. Experimental results show that the application of the Apriori algorithm is more promising in financial development on changes in FDI than the genetic algorithm (GA) (Korgaonkar 2012). IEEE Explore published 65 scientific articles and Google Scholar found 86 articles. A paper by Omar M. Al-Nasser and Xavier Garza Gomez (Al Nasser and Gomez 2009) examines the direct relationship between FDI and the development of the stock market and banking system using pooled data from 15 Latin American countries from 1978 to 2003. Rajdev Tiwari and Manu Pratap Singh (Tiwari and Singh 2010) formulated and validated a correlation-based genetic algorithm (GA) method for optimal attribute subset selection, where GA is used as an optimal search tool for attribute subset selection.Demirhan & Masca (Demirhan, Masca, and Others 2008) found that market size, infrastructure and openness to trade are positive significant while inflation and tax rate are negative significant determinants of FDI in developing countries. Furthermore, Mottaleb & Kalirajan (Mottaleb and Kalirajan 2010) found that Asian developing countries are favored by foreign investors and the lower-middle income countries across the continents are the top FDI recipient countries. Lipsey (Lipsey 2001) analyzed the FDI inflows among the developed countries and suggested different FDI determinants in developed countries compared to developing countries Our team has extensive knowledge and research experience that has translated into high quality publications(K. Mohan et al. 2022; Vivek et al. 2022; Sathish et al. 2022; Kotteeswaran et al. 2022; Yaashikaa, Keerthana Devi, and Senthil Kumar 2022; Yaashikaa, Senthil Kumar, and Karishma 2022; Saravanan et al. 2022; Jayabal et al. 2022; Krishnan et al. 2022; Jayakodi et al. 2022; H. Mohan et al. 2022).

Zhang, Li, Li, and Zhou (Zhang et al. 2010) noted that F.D.I. from foreign companies can contribute to technological modernization and management efficiency of domestic companies. Shiralasheti and Hugar (Siralasheti 2009) also noted that F.D.I. can motivate G.D.P. growth rates, growth in industry and services, reduction in unemployment, reduction in poverty, increase in standard of living, increase in foreign exchange reserves, increase in exports and improvement in the stock market. Kim et al. (M.-J. Kim, Min, and Han 2006) also predicted a stock index using a multi-classifier combination method based on GA to include classifiers derived from machine learning, experts, and users. Rui Chang (Chang and Liu 2011) presented a new optimization algorithm called apriori-improve based on insufficient apriori. (Parakh et al. 2020; Pham et al. 2021; Perumal, Antony and Muthuramalingam 2021; Sathiyamoorthi et al. 2021; Devarajan et al. 2021; Dhanraj and Rajeshkumar 2021; Uganya, Radhika and Vijayaraj 2021; Tesfaye et al. 2021). ; Nandini, Ejilarasan, and Rajeshkumar 2020; Kamat et al. 2020).

The main drawback of the existing genetic algorithm (GA) method is that the imbalance in data processing and the lack of necessary extracted features from FDI data sometimes cannot provide the expected accuracy of the classification result.To overcome this disadvantage, this study proposes a novel technique for the financial development of FDI using the Apriori data mining algorithm in comparison with Genetic algorithm (GA). To evaluate the performance of statistical models, the prediction error for each model will be compared and evaluated based on the accuracy and Mean Square Error (MSE).

2. Materials and Methods

This work was carried out in the DBMS laboratory of the Department of Computer Science, Saveetha School of Engineering. In this article, the dataset contains time series with annual measures of FDI as a percentage of GDP for 40 countries over the period 2005–2020. Of the 40 sample countries, 20 are developed countries and 20 are developing countries. The database is divided into 75% training and 25% testing. Group 1 is a Genetic Algorithm (GA) and Group 2 is a new Apriori algorithm. Calculation is performed using G-power 0.8 with alpha and beta quality of 0.05, 0.2 with 95% confidence interval (Azhar and Kn 2012). These data sets required a monitor with a resolution of 1024×768 pixels (7th generation processor, i5). , 4 8 GB RAM, 500 GB hard disk) and Python software with required library and tool functions Test setup for the proposed system to be

implemented using IBM SPSS V26.0 tool used for statistical analysis.

Genetic Algorithm (GA)

A genetic algorithm (GA) is a search algorithm based on the mechanics of natural selection and genetics, and they combine survival of the fittest among string structures to form a search algorithm (Kelly and Davis 1991). GA is particularly suitable for multiparameter optimization problems with an objective function subject to many hard and soft constraints. The basic idea of a GA is to start with a set of solutions to a problem and try to generate new generations of solutions that are better than the previous ones. GA works through a simple cycle consisting of the following four steps: initialization, selection, crossover, and mutation. The main feature of GA is the population of "chromosomes". Each chromosome acts as a potential solution to a target problem and is usually expressed as binary strings.

Pseudocode:

start
Generate the initial population randomly
Compute the fitness score
For each epoch **do**
Select 2k parents in the population
Generate k offspring by crossover operation
Select m highest score solutions as new population
Mutate n offspring in the new population
Compute fitness score
End for

Apriori Algorithm

Association rule learning is a popular and well-researched technique in data mining for identifying intriguing relationships between variables in sizable databases. The algorithm's name derives from the fact that it makes use of knowledge of common item set features (Han et al. 2004). The association analysis is carried out independently for each of the four class features. As part of the Attribute Analysis, we previously chose the attributes that are pertinent to the class attributes, thus this step is first. The Apriori Algorithm produced the following results. FDI influx expressed as a proportion of GFCF. The findings demonstrate that purchasing power is highly correlated with foreign direct investment when expressed as a proportion of gross fixed capital formation.

Class Attribute: FDI outflow as percentage of GFCF

The findings demonstrate a tight relationship between private credits and GDP per capita current prices and the outflow of foreign direct investment when expressed as a proportion of gross fixed capital formation.

Class Attribute: FDI stock inflow GDP per capita

The results show that the per capita inflow of stocks is closely associated with GDP deflator and it is associated with other attributes strongly when they are taken in combination.

Class Attribute: FDI stock outflow GDP per capita

The findings indicate that private credits and the ratio of deposited assets in banks to GDP are closely related to the per capita inflow of stocks. The range of the financial development variables that the FDI variables are related with makes it possible to anticipate the range in which the FDI can lie using the association rules.

Pseudocode:

Step1: Partition the cluster into groups let this term be n and k be the colors so the loop will be set for k times.
Now, select clusters one at a time.
Step2: Generate items, their items' support, transaction ids.
Step3: F1 = find_frequent_1_itemsets (T);
Step4: For (k = 2; Fk-1 $\neq \phi$; k++) {
Step5: Generate the Ck from the Fk-1
Step6: Ck = candidates generated from Fk-1;
Step7: get the item Iw with minimum support in Ck using F1, ($1 \leq w \leq k$).
Step8: x = Get _ item_ min_ sup (Ck, F1);
Step9: Get the target transaction IDs that contain item x.
Step10: Target = get_ Transaction_ ID(x);
Step11: For each transaction t in Target Do
Step12: Increment the count of all items in Ck that are found in Target;
Step13: Fk

Accuracy is the proportion of correctly classified cases overall cases. True - Positive is where the model could recognize all valid data correctly as a true class. False Positive is where false data is recognized as a true class by the model. A false negative is where the model could not recognize all valid data in true class. True Negative is where the model could be recognized as a false class from false cases.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \qquad (1)$$

$$MSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (ai - pi)^2} \qquad (2)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |ai - pi| \qquad (3)$$

Here, N represents the total number of records predicted, ai represents the actual value of a record and pi represents the predicting value of a record and Nc denotes the number of records that are correctly classified.

Statistical Analysis

Google Colab is utilised to produce the result. It was necessary to use Python software from Milano 2013 with the necessary library functions and tool

functions, as well as a display with a resolution of 1024 x 768 pixels (CPU 7th generation, i5, 4 8GB RAM, 500 GB HDD), to train these datasets. G-power 0.8, alpha and beta quality of 0.05 and 0.2, and a confidence interval of 96% are used in the calculation. Here, IBM SPSS V26.0 is the software programme used for statistical implementation (Healey 2014). When comparing the two groups using the SPSS software, the independent sample t test was used to determine the mean, standard deviation, and standard error mean statistical significance between the groups. This will provide the precise values for the two separate parameters that will be used with.

3. Results

The accuracy rate of the Apriori algorithm is compared with the genetic algorithm in Figure 1 using a straightforward bar graph (GA). When compared to the Genetic Method (GA), the Apriori algorithm has a greater accuracy rate (95.68) than the GA (93.263). Apriori algorithms produce variable outcomes, with standard deviations ranging from 100 to 150 for them and 200 to 300 for genetic algorithms (GA). Between the Apriori algorithm and the Genetic algorithm (GA), there is a significant difference (p0.05 Independent sample test). Apriori algorithm vs accuracy rate of genetic algorithms (GA) on the X-axis Y-axis: Average accuracy rate for identifying terms within 1 standard deviation (95% CI).

The comparison between the evaluation metrics for the Apriori algorithm and the genetic algorithm (GA) classifier is shown in Table 1. The Apriori algorithm has an accuracy rate of 95.68, whereas the Genetic Algorithm (GA) has a rate of 93.263.In all aspects of parameters Apriori algorithm provides better performance compared with the Genetic algorithm (GA) of an enhancement in effects of financial development on foreign direct investments.

Table 2 shows the statistical calculations for the Apriori algorithm and the Genetic algorithm, including Mean, standard deviation, and standard error The accuracy rate parameter used in the t-test. The mean accuracy rate of Apriori algorithm is 95.68 and Genetic algorithm (GA) is 93.263. The Standard Deviation of Apriori algorithm is 0.21029 and Genetic algorithm (GA) is 1.67283. The Standard Error Mean of Apriori algorithm is 0.10291 and Genetic algorithm (GA) is 0.82938.

Table 3 displays the statistical calculations for the comparisons between the Apriori algorithm and the Genetic algorithm for independent samples (GA). The accuracy significance level is 0.0341. The Apriori algorithm and Genetic algorithm (GA) are compared using an independent samples T-test with a threshold of significance of 0.33232 and a 95% confidence interval. This independent sample test includes lower and upper interval difference, mean difference, standard error difference, significance as 0.001, significance (2-tailed), and significance.

4. Discussion

In this study, the proposed model exhibits the Apriori algorithm and Genetic algorithm (GA), in which the Apriori algorithm has the highest values. The Apriori algorithm's accuracy rate is 95.68% higher compared with the Genetic algorithm (GA) that has an accuracy rate of 93.263% in analysis of an enhancement in effects of financial development on foreign direct investments.

A comparative study has been presented between Apriori algorithm and Genetic algorithm. Accuracy analysis has been performed to investigate the importance of each of the input parameters. Apriori algorithm provides better accuracy output when compared to the Genetic algorithm algorithm. Apriori algorithm is a powerful technique to predict the stock market price movement. The accuracy result produced by Apriori algorithm is better than the Genetic algorithm. Apriori algorithm can significantly improve classification accuracy and time efficiency. This shows that the maximum accuracy is obtained quickly in the Apriori algorithm. MSE is used to measure the performances for the models in making predictions. MSE is the average of the square errors between the prediction and the actual values. From the experiment result it is found that Apriori algorithm is used to obtain minimal rules.

Previous studies usually employed statistical and machine learning techniques to forecast future financial values. Kyoung jae kim et al (Kim and Han 2000) proposed a genetic algorithm (GA) approach to feature discretization and determination of connection weights for artificial neural networks (ANN) to predict stock price index. In this study, GA is used to improve both the learning algorithm and the complexity of the feature space. A genetic algorithm (GA) approach to instance selection in artificial neural networks (ANNs) for financial data mining was put forth by Kyoung-jae Kim (K.-J. Kim 2006). In the suggested approach, the GA simultaneously optimises a selection task for pertinent instances and connection weights between layers. For the purpose of forecasting stock price indices, James D. Thomas (Thomas and Sycara 2000) attempted to combine text learning and genetic algorithms. A new genetic algorithm-based approach is presented and applied to the purpose of forecasting the future performances of individual stocks by Mahfoud and Mani (1996). The system broadens the application of GAs beyond their conventional optimization domain to inductive machine learning or Classification.Mendes et al. (Mendes, Godinho,

and Dias 2012) proposed a trading system using genetic algorithms and technical indicators to maximize the Stirling ratio but the performance on testing data struggled to make profits. Yubo jia, (Jia et al. 2012) recommended an improved apriori algorithm based on data division and dynamic item-sets counting.

The limitation of the proposed Apriori algorithm is wasting time for scanning the whole database searching on the frequent itemsets. In future ,it can be enhanced for a variety of datasets directly for feasibility. The proposed approach takes less time and is easy to implement as compared to the previous method. This research can help users to find the right investment strategies with good profit. The results show that strategy design has higher efficiency and takes less time for execution.

5. Conclusion

The proposed model exhibits the Apriori algorithm and Genetic algorithm (GA), in which the Apriori algorithm has the highest values. In the analysis of an improvement in the effects of financial development on foreign direct investments, the Apriori algorithm has an accuracy rate that is 95.68% higher than the Genetic algorithm (GA), which has an accuracy rate of 93.263%.

Declaration

Conflicts of Interest

No conflict of interest in this manuscript

Authors Contributions

Author AR was involved in data collection, data analysis & manuscript writing. Author SAK was involved in conceptualization, data validation, and critical review of manuscripts.

Acknowledgment

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical And Technical Sciences (Formerly known as Saveetha University) for successfully carrying out this work.

Funding: We thank the following organizations for providing financial support that enabled us to complete the study.

1. Chipontime Technologies Pvt. Ltd. Bangalore.
2. Saveetha University
3. Saveetha Institute of Medical And Technical Sciences
4. Saveetha School of Engineering.

5. References

Demirhan, Erdal, Mahmut Masca, and Others. 2008. "Determinants of Foreign Direct Investment Flows to Developing Countries: A Cross-Sectional Analysis." *Prague Economic Papers* 4 (4): 356–69.

Jayabal, Ravikumar, Sekar Subramani, Damodharan Dillikannan, Yuvarajan Devarajan, Lakshmanan Thangavelu, Mukilarasan Nedunchezhiyan, Gopal Kaliyaperumal, and Melvin Victor De Poures. 2022. "Multi-Objective Optimization of Performance and Emission Characteristics of a CRDI Diesel Engine Fueled with Sapota Methyl Ester/diesel Blends." *Energy*. https://doi.org/10.1016/j.energy.2022.123709.

Jayakodi, Santhoshkumar, Rajeshkumar Shanmugam, Bader O. Almutairi, Mikhlid H. Almutairi, Shahid Mahboob, M. R. Kavipriya, Ramesh Gandusekar, Marcello Nicoletti, and Marimuthu Govindarajan. 2022. "Azadirachta Indica-Wrapped Copper Oxide Nanoparticles as a Novel Functional Material in Cardiomyocyte Cells: An Ecotoxicity Assessment on the Embryonic Development of Danio Rerio." *Environmental Research* 212 (Pt A): 113153.

Jia, Yubo, Guanghu Xia, Hongdan Fan, Qian Zhang, and Xu Li. 2012. "An Improved Apriori Algorithm Based on Association Analysis." In *2012 Third International Conference on Networking and Distributed Computing*, 208–11. ieeexplore.ieee.org.

Kim, Kyoung-Jae, and Ingoo Han. 2000. "Genetic Algorithms Approach to Feature Discretization in Artificial Neural Networks for the Prediction of Stock Price Index." *Expert Systems with Applications* 19 (2): 125–32.

Kotteeswaran, C., Indrajit Patra, Regonda Nagaraju, D. Sungeetha, Bapayya Naidu Kommula, Yousef Methkal Abd Algani, S. Murugavalli, and B. Kiran Bala. 2022. "Autonomous Detection of Malevolent Nodes Using Secure Heterogeneous Cluster Protocol." *Computers and Electrical Engineering*. https://doi.org/10.1016/j.compeleceng.2022.107902.

Krishnan, Anbarasu, Duraisami Dhamodharan, Thanigaivel Sundaram, Vickram Sundaram, and Hun-Soo Byun. 2022. "Computational Discovery of Novel Human LMTK3 Inhibitors by High Throughput Virtual Screening Using NCI Database." *Korean Journal of Chemical Engineering*. https://doi.org/10.1007/s11814-022-1120-5.

Lipsey, Robert E. 2001. "Interpreting Developed Countries' Foreign Direct Investment." In *Investing Today for the World of Tomorrow*, 285–325. Springer.

Mendes, Luís, Pedro Godinho, and Joana Dias. 2012. "A Forex Trading System Based on a

Genetic Algorithm." *Journal of Heuristics* 18 (4): 627–56.

Mohan, Harshavardhan, Sethumathavan Vadivel, Se-Won Lee, Jeong-Muk Lim, Nanh Lovanh, Yool-Jin Park, Taeho Shin, Kamala-Kannan Seralathan, and Byung-Taek Oh. 2022. "Improved Visible-Light-Driven Photocatalytic Removal of Bisphenol A Using V2O5/WO3 Decorated over Zeolite: Degradation Mechanism and Toxicity." *Environmental Research*. https://doi.org/10.1016/j.envres.2022.113136.

Mohan, Kannan, Abirami Ramu Ganesan, P. N. Ezhilarasi, Kiran Kumar Kondamareddy, Durairaj Karthick Rajan, Palanivel Sathishkumar, Jayakumar Rajarajeswaran, and Lorenza Conterno. 2022. "Green and Eco-Friendly Approaches for the Extraction of Chitin and Chitosan: A Review." *Carbohydrate Polymers* 287 (July): 119349.

Mottaleb, Khondoker Abdul, and Kaliappa Kalirajan. 2010. "Determinants of Foreign Direct Investment in Developing Countries: A Comparative Analysis." *Margin: The Journal of Applied Economic Research* 4 (4): 369–404.

Saravanan, A., P. Senthil Kumar, B. Ramesh, and S. Srinivasan. 2022. "Removal of Toxic Heavy Metals Using Genetically Engineered Microbes: Molecular Tools, Risk Assessment and Management Strategies." *Chemosphere* 298 (July): 134341.

Sathish, T., R. Saravanan, V. Vijayan, and S. Dinesh Kumar. 2022. "Investigations on Influences of MWCNT Composite Membranes in Oil Refineries Waste Water Treatment with Taguchi Route." *Chemosphere* 298 (July): 134265.

Vivek, J., T. Maridurai, K. Anton Savio Lewise, R. Pandiyarajan, and K. Chandrasekaran. 2022. "Recast Layer Thickness and Residual Stress Analysis for EDD AA8011/h-BN/B4C Composites Using Cryogenically Treated SiC and CFRP Powder-Added Kerosene." *Arabian Journal for Science and Engineering*. https://doi.org/10.1007/s13369-022-06636-5.

Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. "Algal Biofuels: Technological Perspective on Cultivation, Fuel Extraction and Engineering Genetic Pathway for Enhancing Productivity." *Fuel*. https://doi.org/10.1016/j.fuel.2022.123814.

Yaashikaa, P. R., P. Senthil Kumar, and S. Karishma. 2022. "Review on Biopolymers and Composites – Evolving Material as Adsorbents in Removal of Environmental Pollutants." *Environmental Research*. https://doi.org/10.1016/j.envres.2022.113114.

TABLES AND FIGURES

Table 1. Accuracy Values for Apriori and Genetic.

SI.No.	**Test Size**	**ACCURACY RATE**	
		Apriori algorithm	Genetic algorithm (GA)
1	Test1	93.23	92.10
2	Test2	93.54	92.23
3	Test3	92.36	93.19
4	Test4	94.34	92.92
5	Test5	93.12	91.92
6	Test6	94.56	93.01
7	Test7	95.35	92.85

8	Test8	92.36	92.28
9	Test9	93.35	92.58
10	Test10	94.54	92.34
Average Test Results		95.68	93.263

Table 2. The statistical calculation such as Mean, standard deviation and standard error Mean for Apriori algorithm and Genetic algorithm (GA). The accuracy rate parameter used in the t-test. The mean accuracy rate of Apriori algorithm is 95.68 and Genetic algorithm (GA) is 93.263. The Standard Deviation of Apriori algorithm is 0.21029 and Genetic algorithm (GA) is 1.67283. The Standard Error Mean of Apriori algorithm is 0.10291 and Genetic algorithm (GA) is 0.82938.

Group		**N**	**Mean**	**Standard Deviation**	**Standard Error Mean**
Accuracy Rate	GENETIC ALGORITHM (GA)	10	93.263	1.67283	0.82938
	APRIORI ALGORITHM	10	95.68	0.21029	0.10291

Table 3. The statistical calculations for independent samples test between Apriori algorithm and Genetic algorithm (GA). The significance value p=0.001 ($p<0.05$), which shows that two groups are statistically significant.

Group		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval (Lower)	95% Confidence Interval (Upper)
Accuracy	Equal variances assumed	7.234	0.0341	17.012	18	.000	11.873	0.983	11.7849	13.89495
	Equal variances not assumed			17.145	12.734	.000	11.453	0.235	10.6738	12.23849

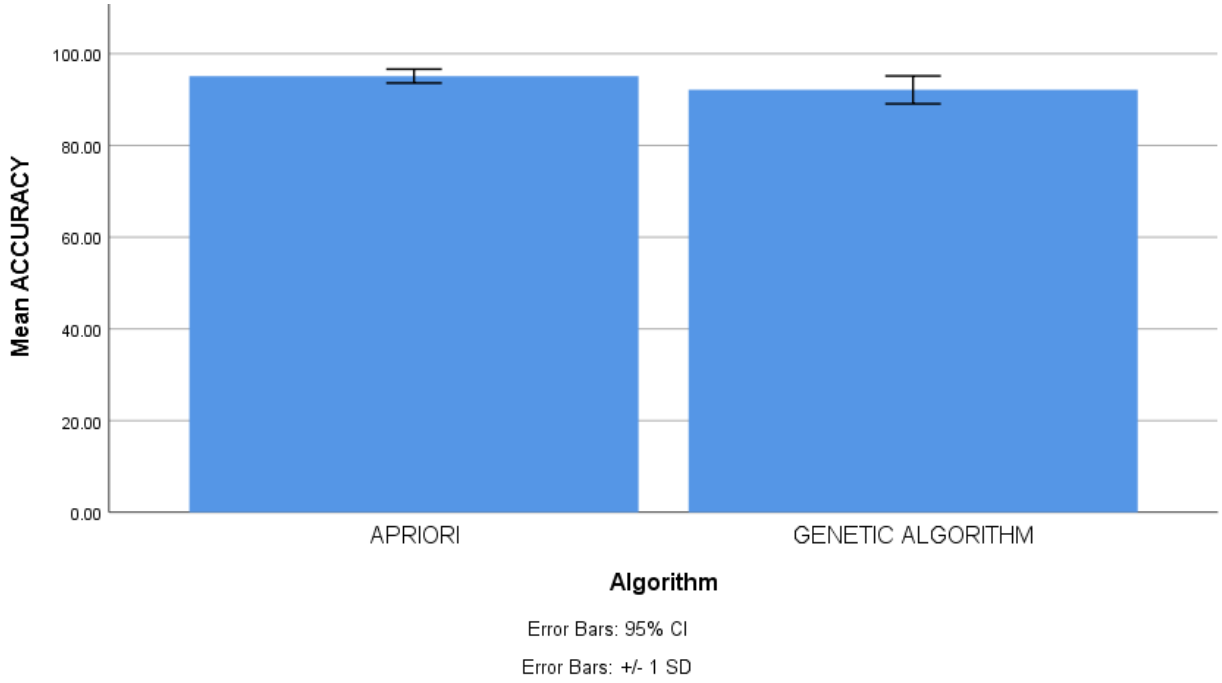

Fig. 1. Simple Bar graph for Apriori algorithm accuracy rate is compared with Genetic algorithm (GA). The Apriori algorithm is higher in terms of accuracy rate 95.68 when compared with Genetic algorithm (GA) 93.263. Variable results with its standard deviation ranging from 100 lower to 150 higher Apriori algorithm where Genetic algorithm (GA) standard deviation ranging from 200 lower to 300 higher. There is a significant difference between Apriori algorithm and Genetic algorithm (GA) ($p<0.05$ Independent sample test). X-axis: Genetic algorithm (GA) accuracy rate vs Apriori algorithm Y-axis: Mean of accuracy rate, for identification of keywords ± 1 SD with 95 % CI.