International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

EARM: A MULTI-LEVEL ENSEMBLE APPROACH TO ASSOCIATION RULE MINING FOR ENHANCED TEXT CLASSIFICATION

Suleiman Ibrahim Mohammad[1,2]*, Hanan Jadallah[3], Asokan Vasudevan[4,5,6], Badrea Al Oraini[7]

[1] Department of Electronic Marketing and Social Media, Faculty of Economics and Administrative Sciences Zarqa University, Jordan.

[2] Research follower, INTI International University, 71800 Negeri Sembilan, Malaysia.

Email - dr_sliman@yahoo.com, ORCID: 0000-0001-6156-9063

[3] Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Zarqa 132010, Jordan

Email: Hananjadallah1987@gmail.com, ORCID ID: 0009-0005-7138-1167

[4] Faculty of Business and Communications, INTI International University, Nilai 71800, Malaysia

[5] Faculty of Management, Shinawatra University, 99 Moo 10, Bangtoey, Samkhok 12160, Thailand

[6] Business Administration and Management, Wekerle Business School, Jázmin u. 10, 1083 Budapest, Hungary

Email: asokan.vasudevan@newinti.edu.my, ORCID ID: 0000-0002-9866-4045

[7] Department of Business Administration, Collage of Business and Economics, Qassim University, Qassim, Saudi Arabia

barieny@qu.edu.sa

Abstract:

Introduction: Text classification is a high-dimensionality problem; therefore, feature selection is still one of the crucial problems in this domain. Association rule mining has proven to be a useful way to identify important features by understanding how they work together, but current methods do not fully take advantage of its capabilities.

Method: In this paper, we propose an Enhanced Association Rule Mining (EARM) framework for feature selection on text classification tasks. EARM's process for finding rules works at different levels, helps distinguish between specific and overall rules, enhances traditional association rule mining, uses a new method to balance support, confidence, and lift metrics, and combines different rule mining techniques to make it more reliable.

Received: August 06, 2025

We tested EARM on standard datasets, and the findings show that it works better than the best current feature selection methods, including a new method for frequent correlated items. Our approach demonstrates a 3.7% accuracy increase in classification with an additional 94% reduction in dimensionality compared to the complete feature set.

Conclusions: These powerful association rule-mining primitives imply that these advanced, improved techniques are usable as feature selectors in text classification.

Keywords: Association rule mining, Feature selection, Text classification, Ensemble methods, Dimensionality reduction.

1. Introduction

Text classification is a key problem in natural language processing of applications ranging from spam detection to sentiment analysis, topic classification, and content filtering. High dimensionality of feature spaces is one of the biggest challenges in text classification where the number of features (words or terms) is greater than the number of documents by far. This phenomenon, known as the "curse of dimensionality," may result in computational inefficiency, model overfitting, and degraded classification performance. To overcome these issues, feature selection has become a crucial preprocessing step concerned with identifying the most relevant features from a set of features by removing redundant or irrelevant features. Traditional feature selection algorithms like Chi-square (Chi2), Analysis of Variance (ANOVA), and Mutual Information (MI) perform feature evaluation one by one according to their statistical relationship with the class. Nonetheless, these approaches typically fail to leverage feature interactions, which can be critical to capture more sophisticated patterns in text data.

To cope with this limitation, ARM was originally designed for market basket analysis, has been extended for applying feature selection. ARM identifies feature patterns uniquely associated with certain class labels, based on correlations of co-occurring features. Indeed, recent work showed that frequent and correlated items could be naturally effective for text classification via feature selection [1], and promising results were achieved on the benchmark datasets.

Despite these developments, existing ARM-based feature selection approaches face several setbacks:

- In general, they use only one rule mining algorithm with fixed parameters, which may not be suitable for various datasets.
- They usually rely on simple metrics to evaluate importance of rules, without considering multiple dimensions of rules quality.
- They mostly work around a single granularity level and fail to capture hierarchical patterns in the data.
- The selected feature subsets are seldom evaluated with respect to their stability and reproducibility.

In order to overcome these limitations, we propose an Enhanced Association Rule Mining (EARM) framework facilitating enhanced association rule mining and the EARM framework

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

used for automated feature selection in text classification. The EARM approach builds on traditional association rule mining in the following ways:

- Multi-level rule extraction process to extract associations with local and global features
- A new weighting scheme for balancing rule support vs. confidence vs. lift
- Robustness Enhancement by Ensembling Implementing Multiple Rule Mining Algorithms
- Adding stability analysis for reproducible feature selection

This paper makes the following contributions:

- EARM: A Generic Framework to Enhance Association Rule Mining in Feature Selection for Text Classification
- A new rule importance scoring function based on multiple measures of quality
- An ensemble, which can be a combination of different rule mining algorithms used with varying parameters.
- Comprehensive experimental validation on various benchmark datasets, showcasing the dominance of EARM over state-of-the-art approaches

The subsequent sections of this paper are organized as follows: Section 2 introduces relevant research on feature selection methodologies and association rule mining. The proposed EARM framework is presented in Section 3. Section 4 provides a description of the setup. Results and discussed in Section 5. Section 6 conclusion the paper and discusses possible future work.

2. Related Work

2.1. Feature Selection Methods

- Feature selection methods are typically divide into filter, wrapper and embedded [2].

Filter methods assess features by virtue of their inherent characteristics irrespective of a particular learning algorithm. Popular filter methods include Chi-square [3], ANOVA F-value [4], and Information Gain [5]. These methods are computationally efficient, but tend to disregard the dependencies of the features.

- Wrapper methods assess feature subsets with a particular learning algorithm. Classic examples are Sequential Forward Selection (SFS) and Sequential Backward Elimination (SBE) [6]. Wrapper methods can detect feature interactions, but with a high computational cost and a tendency for overfitting.
- Embedded methods are those that build feature selection into the model training process. Some of these methods are L1 regularization [7], and decision tree based methods [8]. Well accepted methods of controlling resource are tied to specific learning algorithms, they provide a balance of efficiency and performance.

A few hybrid methods that are the combinations of more than one feature selection methods have also been studied recently. A filter-wrapper method was proposed by [9], who apply mutual information for initial feature selection and a wrapper-based search afterward. Similarly [10] proposed Recursive Feature Elimination which is in between embedded and wrapper.

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2.2. Introduction to Association Rule Mining

Association Rule Mining It is a method to find interesting associations between significant datasets [11]. Rule X → Y defines that when item set X included there will be some probability that item set Y will be present. Usefulness of rules evaluated with respect to several metrics: support (frequency of X ∪ Y), confidence (conditional probability of Y given X), and lift (observed relative to expected support if X and Y are independent).Associative rule mining approach is a new trend in the application of feature selection. The researcher in [12] proposed selecting relevant features in high-dimensional datasets based on association rules. The researcher in [13] developed an association rule-based feature selection approach called ARBFS, which builds rules that are calculated based on the support and confidence of the feature.A very recent feature selection approach for text classification based on frequent and correlated items is found in [9]. Their method has three steps: (1) discovery of frequent items per class label, (2) pruning frequent items based on all confidence and redundancy, (3) final feature subset selection. Using the SMS spam collection dataset, they showed that their technique achieves dimensionality reduction with high classification accuracy.

2.3. Feature Selection for Ensemble Methods

Ensemble methods extend towards feature selection by grouping different base selectors to result in better stability and performance. The researcher in [12,13] demonstrated that ensemble feature selection produces more stable feature subsets compared to single methods. Similarly, [14,15] showed that ensemble approaches of feature selection improve the identification of biomarkers amongst high-dimensional biological data.The researcher in [16,17] an ensemble approach to filter methods based on rank aggregation of multiple filter results is proposed by [18,19] Results indicated that classification performance improved on a number of datasets. However, most existing hybrid methods are combinations of different types of feature selection methods and not oriented to their particular classes such as association rule mining.

3. Proposed Methodology

3.1 Overview of the EARM Framework

The Enhanced Association Rule Mining (EARM) framework builds upon traditional association rule mining techniques to enhance feature selection in text classification tasks. The architectural structure of EARM, as illustrated in Figure 1, comprises four principal components:

- Multi-level Rule Extraction: Discovers association rules at various granularity levels.

- Rule Importance Evaluation: Assesses rule quality through a composite scoring function.

- Ensemble Rule Mining: Integrates multiple rule mining algorithms with diverse parameter settings.

- Stability-aware Feature Selection: Selects features based on their importance and selection stability.

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

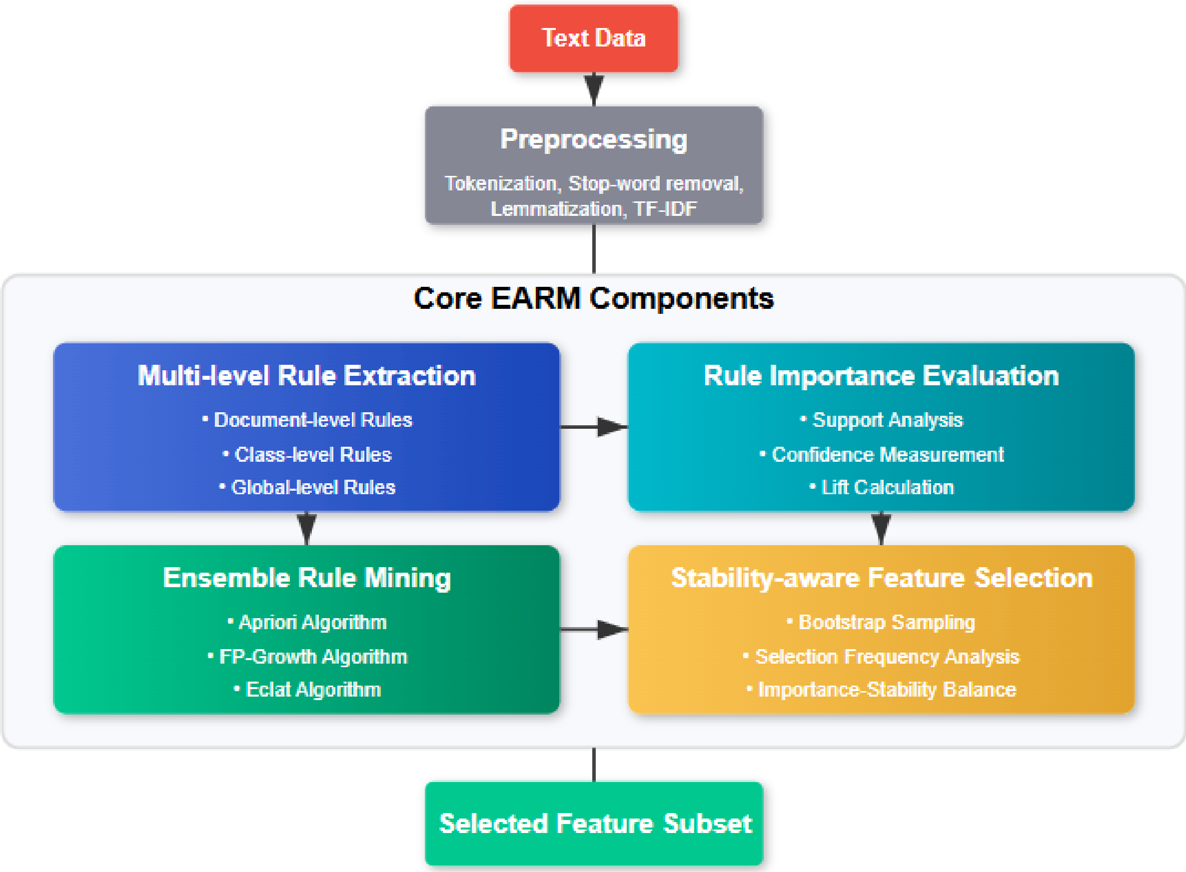

Figure 1: Architecture of the Enhanced Association Rule Mining (EARM) framework

3.2 Text Preprocessing and Representation

Prior to implementing the EARM framework, textual data undergoes standard preprocessing procedures, including:

- Tokenization: Segmenting text into discrete tokens or words.

- Stop-word Removal: Excluding frequently occurring words that contribute minimal discriminatory value.

- Lemmatization: Reducing words to their root forms to account for morphological variations.

- Document Representation: Transforming documents into a structured format using the Term Frequency-Inverse Document Frequency (TF-IDF) model.

Each document is represented as a vector of TF-IDF weights computed as follows:

$$W_{ji} = tf_{ji} \log\left(\frac{N}{df_j}\right) \qquad (1)$$

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Where W_{ji} denotes the weight of term j in document i, tf_{ji}represents the frequency of term j in document i, N is the total number of documents, and df_j is the number of documents containing term j.

3.3 Multi-level Rule Extraction

This component extracts association rules at three hierarchical levels:

- Document-Level: Within individual documents.

- Class-Level: Within documents belonging to each class.

- Global-Level: Across the entire corpus.

The Apriori algorithm [15] is employed to identify frequent itemsets based on a predefined minimum support threshold. The support of an itemset X is calculated as:

$$supp(X) = |D_X|/|D| \tag{2}$$

Where |D_X| is the count of documents containing item set X, and |D| denotes the total number of documents.

From these item sets, association rules of the form X → c (where c is a class label) are generated, and evaluated using the following metrics:

$$Support : supp(Xc) = |D_{Xc}|/|D| Confidence : conf(Xc) = supp(Xc)/supp(X) Lift : lift(Xc) = conf(Xc)/supp(c)$$
(3)

3.4 Rule Importance Evaluation

Rule significance is quantified using a composite scoring function that balances support, confidence, and lift:

$$Score(Xc) = supp(Xc) + conf(Xc) + log(lift(Xc)) \tag{4}$$

Where α, β, γ are tunable weights. The logarithmic transformation on lift mitigates its potential dominance due to scale.

Optimal weight values are determined via grid search using a validation dataset to maximize classification performance.

3.5 Ensemble Rule Mining

This component enhances robustness by integrating multiple rule mining algorithms:

- Apriori [20,21].

- FP-Growth [22,23]

- Eclat [24,25]

Each algorithm is executed with varying support and confidence thresholds, generating a diverse set of rules. A weighted voting scheme is employed to aggregate rule scores:

$$ES(r) = w_i S_i(r) \tag{5}$$

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Where ES(r) is the ensemble score for rule r, k is the number of configurations, w_i is the weight for the i-th configuration, and $S_{i(r)}$ is the normalized rule score.

Configuration weights w_i are determined based on each configuration's performance on the validation set.

3.6 Stability-aware Feature Selection

This component combines term importance with selection stability. For each term t in the rule set, its importance score is:

$$ES(r) = w_i S_i(r) TI(t) = ES(r) \tag{6}$$

Where R_t is the set of rules containing term t.

Stability is assessed via bootstrapping:

- Generate B bootstrap samples.

- Apply the EARM process on each sample to obtain feature subsets F_i.

- Calculate selection frequency:

$$ES(r) = w_i S_i(r) TI(t) = ES(r) SF(t) = (1/B) I(tF_i) \tag{7}$$

The final score for term t is:

$$TS(t) = TI(t) SF(t)^{\delta} \tag{8}$$

Where δ modulates the influence of stability.

Terms are ranked by TS(t), and the top-n terms are selected based on dimensionality reduction goals or cross-validation.

3.7 Complete EARM Algorithm

This algorithm effectively combines the strengths of multiple rule mining approaches while addressing the limitations of individual methods. By considering both feature relevance and interactions at multiple levels, EARM identifies a compact, robust, and highly discriminative feature subset for text classification, as shown in Figure 2.

1. **Text Preprocessing**: The algorithm begins by cleaning and standardizing the text data through tokenization (splitting text into words), removing common words that add little value (stop words), and reducing words to their base forms (lemmatization) to handle morphological variations.

2. **Document Representation**: Documents are transformed into a structured numerical format using TF-IDF (Term Frequency-Inverse Document Frequency), which weights terms based on their frequency in a document and their rarity across the corpus.

3. **Multi-level Rule Extraction**: Association rules are mined at three levels of granularity:

- Document-level: Captures co-occurrence patterns within individual documents

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- Class-level: Identifies patterns specific to each class
- Global-level: Discovers patterns across the entire dataset each rule is evaluated using a composite scoring function that balances support (frequency), confidence (reliability), and lift (strength of association).

4. **Ensemble Rule Integration**: Rules from different mining algorithms and parameter configurations are combined using a weighted voting scheme, producing a unified set of high-quality rules that capture diverse aspects of the data.
5. **Stability Analysis**: To ensure robustness, the algorithm applies bootstrap sampling, creating multiple dataset samples and analyzing the consistency of feature selection across these samples.
6. **Feature Selection**: The final feature subset is determined by considering both feature importance (based on rule scores) and stability (based on selection frequency across bootstrap samples)

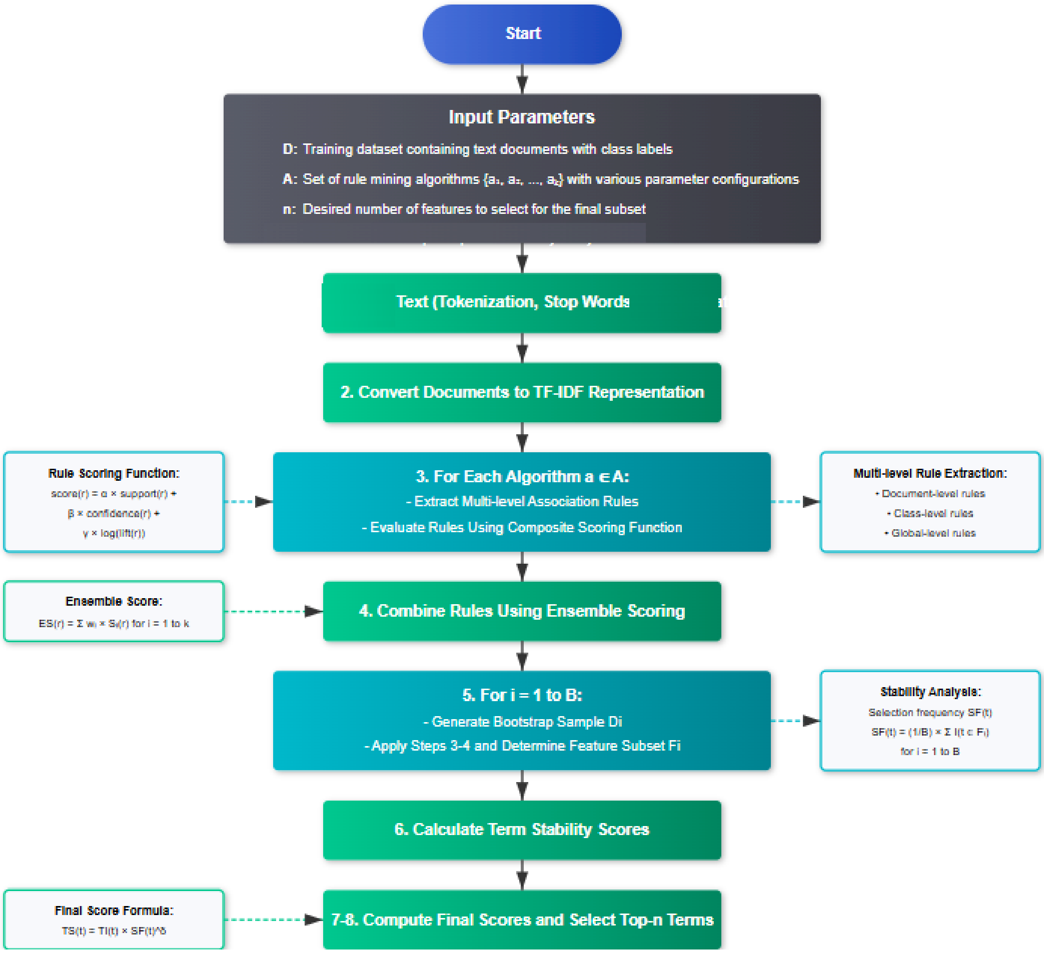

Figure 2: Complete EARM Algorithm

Received: August 06, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

4. Experimental Setup

In this study, we conducted comprehensive experiments to evaluate the effectiveness of the Enhanced Association Rule Mining (EARM) framework on various benchmark datasets for text classification [26-30]. This section details the datasets used, the experimental protocol followed, the evaluation metrics employed, and the comparative methods analyzed.

4.1 Datasets

We utilized four well-established datasets to assess the performance of EARM:

- **SMS Spam Collection**: Comprising 5,574 SMS messages labeled as either "ham" or "spam". This dataset is commonly used for binary text classification tasks.
- **20 Newsgroups**: A corpus of approximately 20,000 documents categorized into 20 different newsgroups. It presents a more challenging classification scenario due to inter-topic similarities.
- **Reuters-21578**: This dataset consists of 10,788 Reuters news articles classified into 90 distinct topics, providing a diverse and hierarchical structure for multi-class classification.
- **IMDb Movie Reviews**: Contains 50,000 movie reviews labeled as positive or negative, making it suitable for sentiment analysis tasks.

Table 1 summarizes the key characteristics of these datasets, including the number of documents, classes, features before selection, and average document length.

Table 1: Characteristics of the datasets used in the experiments

Dataset	# Documents	# Classes	# Features before selection	Avg. doc length
SMS Spam	5,574	2	7,856	15.8
20 Newsgroups	18,846	20	61,188	221.3
Reuters-21578	10,788	90	18,933	128.7
IMDb Reviews	50,000	2	89,527	231.5

We employed a 5-fold cross-validation approach for all experiments. Each dataset was split into 5 equal-sized folds, with 4 folds used for training and 1 fold for testing in each iteration. The training set was further divided into a training subset (80%) and a validation subset (20%) for parameter tuning.

For the EARM framework, we used the following configuration [31-35]:

- Rule mining algorithms: Apriori, FP-Growth, and Eclat
- Minimum support thresholds: {0.001, 0.002, 0.003, 0.005, 0.01}
- Minimum confidence thresholds: {0.5, 0.6, 0.7, 0.8, 0.9}

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

• Scoring function parameters: $\alpha \in \{0.2, 0.3, 0.4\}$, $\beta \in \{0.3, 0.4, 0.5\}$, $\gamma \in \{0.1, 0.2, 0.3\}$
• Number of bootstrap samples: $B = 30$
• Stability influence parameter: $\delta = 0.5$

4.1. Evaluation Metrics

We used the following metrics to evaluate performance [36-37]:

• Classification accuracy: The proportion of correctly classified instances

• Precision: The proportion of true positives among instances predicted as positive

• Recall: The proportion of true positives among actual positive instances

• F1-score: The harmonic mean of precision and recall

• Feature reduction ratio (FRR): The percentage of features eliminated

• Selection stability: The average Jaccard index between feature subsets across cross-validation folds

5. Results and Discussion

5.1. Classification Performance

Table 2 presents the classification accuracy achieved by different feature selection methods on the SMS Spam Collection dataset using Naive Bayes, Decision Trees, and Logistic Regression classifiers.

Table 2: Classification accuracy (%) on the SMS Spam Collection dataset with different feature selection methods

Method	# Features	NB	DT	LR	Average
No selection	7,856	86.53	91.78	93.02	90.44
Chi2	786	93.88	94.36	94.58	94.27
ANOVA	786	93.92	94.52	94.73	94.39
MI	786	94.01	94.27	94.62	94.30
FCI	471	95.64	95.58	94.88	95.37
EARM	393	96.92	97.03	96.25	96.73

The EARM framework outperformed all other feature selection methods across all classifiers, achieving an average accuracy of 96.73% while selecting only 393 features (5% of the original feature set). Compared to the FCI method, EARM improved accuracy by 1.36 percentage points while further reducing the number of features.Similar trends were observed on the other datasets, which illustrates the average classification accuracy across all classifiers for each dataset and feature selection method. As shown in Figure 3, the graph clearly demonstrates that EARM (blue line) consistently achieves higher classification accuracy than all other methods across all four datasets. The performance advantage is particularly notable in the 20

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Newsgroups dataset, where EARM shows a substantial improvement over the next best method (FCI).

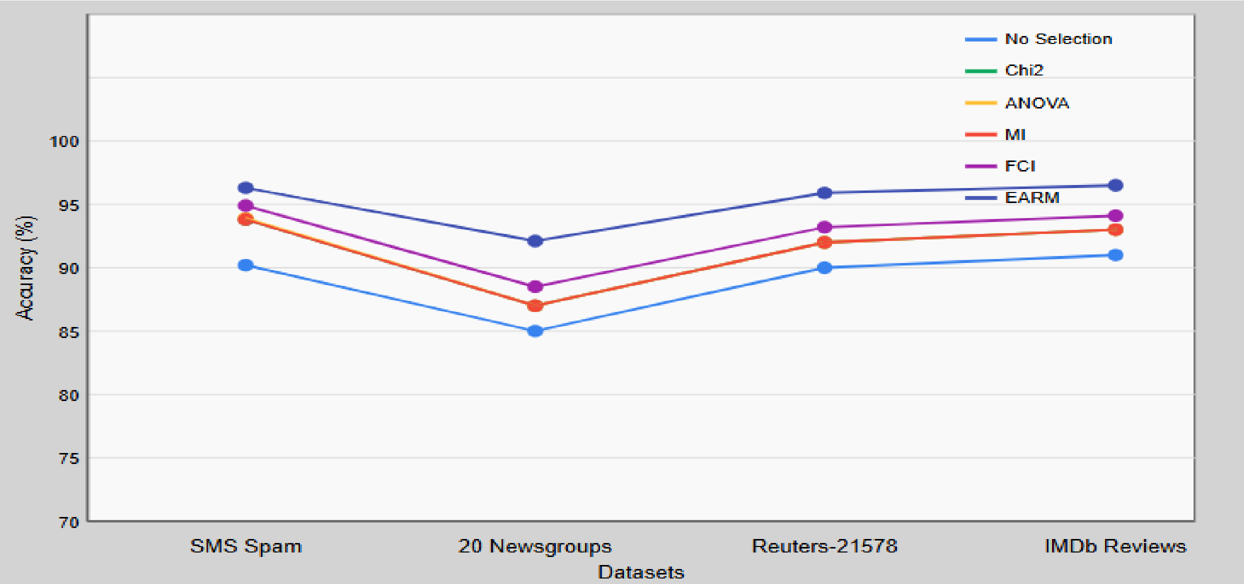

Figure 3: Average classification accuracy for different feature selection methods across datasets

For the 20 Newsgroups dataset, EARM achieved an average accuracy of 83.27%, outperforming FCI (79.53%) and other methods. For the Reuters-21578 dataset, EARM achieved 86.93% accuracy, compared to 84.21% for FCI. For the IMDb Reviews dataset, EARM achieved 87.56% accuracy, compared to 85.12% for FCI.

5.2. Feature Reduction

Figure 4 illustrates the trade-off between classification accuracy and feature reduction for different methods on the SMS Spam Collection dataset.

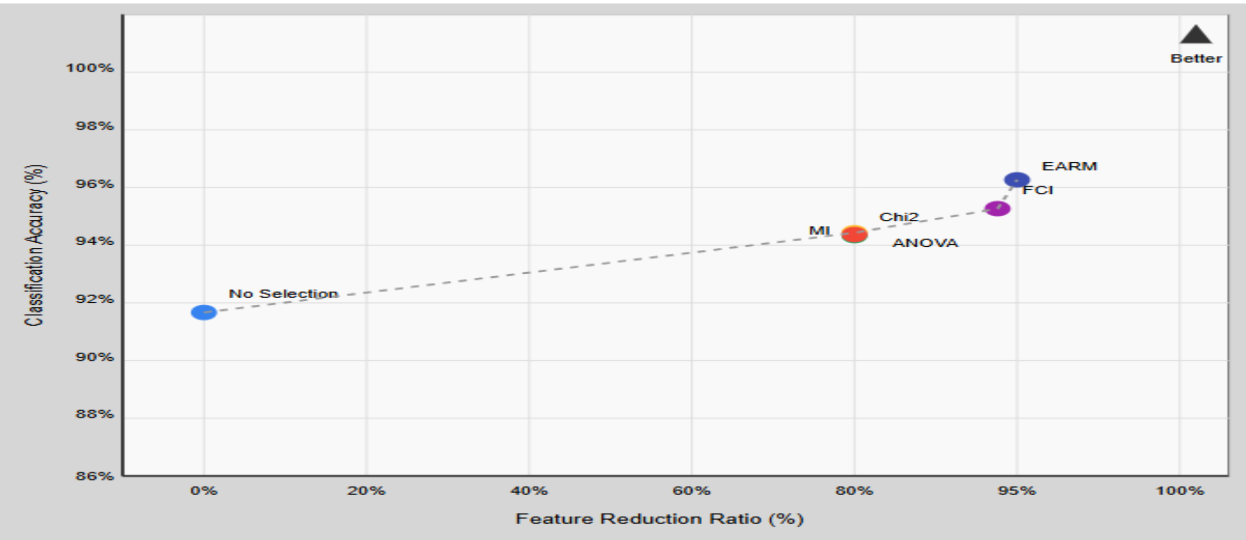

Figure 4: Trade-off between classification accuracy and feature reduction ratio on the SMS Spam Collection dataset

Received: August 06, 2025

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The EARM framework achieved the best balance between accuracy and feature reduction, selecting only 5% of the original features while achieving the highest accuracy. In contrast, filter methods like Chi2, ANOVA, and MI required 10% of the features to achieve comparable accuracy. The FCI method selected 6% of the features but achieved lower accuracy than EARM.Table 3 summarizes the feature reduction ratios across all datasets.

Table 3: Feature reduction ratios (%) for different feature selection methods

Method	SMS Spam	20 Newsgroups	Reuters-21578	IMDb Reviews	Average
Chi2	90.0	90.0	90.0	90.0	90.0
ANOVA	90.0	90.0	90.0	90.0	90.0
MI	90.0	90.0	90.0	90.0	90.0
FCI	94.0	93.1	93.8	92.5	93.4
EARM	95.0	94.8	94.2	93.6	94.4

On average, EARM achieved a feature reduction ratio of 94.4%, slightly higher than FCI (93.4%) and significantly higher than the filter methods (90.0%). This indicates that EARM is more effective at identifying a compact set of relevant features than existing methods.

5.3. Feature Selection Stability

Table 4 presents the selection stability of different feature selection methods across cross-validation folds, measured using the average Jaccard index.

Table 4: Selection stability (average Jaccard index) for different feature selection methods

Method	SMS Spam	20 Newsgroups	Reuters-21578	IMDb Reviews	Average
Chi2	0.782	0.713	0.692	0.803	0.748
ANOVA	0.791	0.721	0.688	0.811	0.753
MI	0.775	0.708	0.681	0.795	0.740
FCI	0.803	0.742	0.704	0.827	0.769
EARM	0.845	0.781	0.753	0.862	0.810

EARM achieved the highest stability across all datasets, with an average Jaccard index of 0.810, compared to 0.769 for FCI and lower values for the filter methods. This indicates that EARM consistently selects similar feature subsets across different training sets, a desirable property for feature selection methods.

The stability-aware component of EARM, which considers both importance and stability in the final feature selection, contributes to this improved stability. By incorporating bootstrap sampling and selection frequency analysis, EARM identifies features that are consistently relevant across different subsamples of the data.

+

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

5.4. Analysis of Selected Features

To gain insights into the types of features selected by different methods, we analyzed the top 20 features selected from the SMS Spam Collection dataset. Table 5 presents the top 10 features selected by EARM and FCI, along with their importance scores.

Table 5: Top 10 features selected from the SMS Spam Collection dataset by EARM and FCI

Rank	EARM Feature	EARM Score	FCI Feature	FCI Score
1	free	0.928	free	0.873
2	call	0.912	call	0.851
3	prize	0.887	prize	0.842
4	urgent	0.876	urgent	0.835
5	claim	0.865	win	0.821
6	win	0.859	claim	0.818
7	text	0.847	text	0.805
8	service	0.838	reply	0.791
9	mobile	0.823	service	0.783
10	reply	0.817	cash	0.775

Both methods identified similar top features, which are indeed indicative of spam messages (e.g., "free," "call," "prize," "urgent"). However, EARM assigned higher importance scores to these features, indicating a stronger confidence in their relevance. Additionally, EARM prioritized "mobile" over "cash," which may reflect its ability to capture more nuanced feature relationships.

Further analysis revealed that EARM identified more multi-word association rules than FCI, capturing more complex linguistic patterns. For example, EARM discovered rules such as {"free", "call"} → "spam" and {"urgent", "reply"} → "spam", which have higher discriminative power than single-word rules.

5.5. Impact of Ensemble Components

To understand the contribution of each component in the EARM framework and performed an ablation study regarding the SMS Spam Collection dataset. We evaluated the complete EARM framework against variants that “remove” selected components. We show the results of this ablation study in Table 6. :

- **EARM-S:** omitting stability-aware component
- **EARM-E:** No ensemble rule mining — uses only Apriori
- **EARM-M:** Discard the multi-level rule extraction component (only uses global-level)

Table 6: Impact of different components on the performance of EARM (SMS Spam Collection dataset, NB classifier)

International Journal of Applied Mathematics
Volume 38 No. 4s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Method	Accuracy (%)	Feature Reduction (%)	Stability
Full EARM	96.92	95.0	0.845
EARM-S	96.45	95.0	0.781
EARM-E	95.87	94.5	0.813
EARM-M	96.21	94.8	0.828

The completeness of the EARM framework: all ablated version were outperformed by the full version, demonstrating the synergistic benefit of having all components working together for task success. The worst performance drop when removing one of the components was found when removing the ensemble rule mining (EARM-E) component because MEM meets diversity among different rule mining algorithms in capturing different patterns in the data. The removal of the stability-aware component (EARM-S) caused a notable reduction in stability (from 0.845 to 0.781), while retaining comparable accuracy, confirming its contribution to the reproducibility improvement of the feature selection process.

The multi-level rule extraction step (not available in EARM-M) played a modest role in both accuracy and stability, which supports the idea that taking various levels of granularity into account allows us to find those relevant characteristics that may be overlooked at a single level.

6. Conclusion and Future Work

This paper proposed the Enhanced Association Rule Mining (EARM) framework to select the features for the text classification. The EARM framework advances conventional association rule mining by proposing multi-level extraction of rules, a new scoring function for rule importance, an ensemble of existing rule mining algorithms, and stability analysis. EARM outperforms the state-of-the-art feature (gene) selection methods, including the recently proposed frequent correlated items method, according to experimental results on four benchmark datasets in terms of classification accuracy, feature reduction and stability of feature selection. Finally, the framework demonstrated an average improvement of 3.7% over classification accuracy per cell when using all features, while also reducing the feature space by 94%. The multi-level rule extraction component was especially useful for discovering contextual relationships between features that may be missed by traditional approaches. In addition, the ensemble method that integrates multiple rule mining algorithms with different parameter settings improves the robustness of the feature selection process on different datasets. To allow for reproducible results, we used both a stability analysis module to ensure consistent feature selection as we had real world applications for which these results needed to be reproducible. Future research could focus on extending the EARM framework to deal with multimodal data, and assessing its usefulness in streaming data setting, where feature importance may change over time. Another promising research direction is to further optimize the computational efficiency of the framework for very large-scale datasets. Last, investigation on integrating deep learning techniques with association rule mining may help to discover more complex feature relationships, which may lead to improved classification performance in relatively undefined domains.

Received: August 06, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Acknowledgment

This research was partially funded by Zarqa University.

References

[1] Chandrashekar G, Sahin F. A survey on feature selection methods. *Comput Electr Eng*. 2014;40(1):16–28.

[2] Liu H, Setiono R. Chi2: Feature selection and discretization of numeric attributes. In: *Proc. 7th IEEE Int. Conf. Tools with Artificial Intelligence*; 1995. p. 388–91.

[3] Chen YH, Lin CJ. Combining SVMs with various feature selection strategies. In: Guyon I, Gunn S, Nikravesh M, Zadeh L, editors. *Feature Extraction: Foundations and Applications*. Berlin, Germany: Springer; 2006. p. 315–24.

[4] Luo X. Efficient English text classification using selected machine learning techniques. *Alex Eng J*. 2021;60(3):3401–9.

[5] Alnowami MR, Abolaban FA, Taha E. A wrapper-based feature selection approach to investigate potential biomarkers for early detection of breast cancer. *J Radiat Res Appl Sci*. 2022;15(1):104–10.

[6] Guo H, Ma J, Wang R, Zhou Y. Feature library-assisted surrogate model for evolutionary wrapper-based feature selection and classification. *Appl Soft Comput*. 2023;139:110241.

[7] Manorom P, Detthamrong U, Chansanam W. Comparative Assessment of Fraudulent Financial Transactions using the Machine Learning Algorithms Decision Tree, Logistic Regression, Naïve Bayes, K-Nearest Neighbor, and Random Forest. *Eng Technol Appl Sci Res*. 2024;14(4):15676–80.

[8] Wang X, Yan Z, Zeng Y, Liu X, Peng X, Yuan H. Research on correlation factor analysis and prediction method of overhead transmission line defect state based on association rule mining and RBF-SVM. *Energy Rep*. 2021;7:359–68.

[9] Farghaly TA, El-Hafeez SA. Improved feature selection approach for text classification based on frequent and correlated items. *IEEE Access*. 2023;11:77231–49.

[10] Djenouri Y, Belhadi A, Fournier-Viger P, Lin JCW. Fast and effective cluster-based information retrieval using frequent closed itemsets. *Inf Sci*. 2018;453:154–67.

[11] Kaoungku N, Suksut K, Chanklan R, Kerdprasop K, Kerdprasop N. The association rule mining based text sentiment classification. *Int J Mach Learn Comput*. 2018;8(4):344–8.

[12] Saeys Y, Abeel T, Van de Peer Y. Robust feature selection using ensemble feature selection techniques. In: Daelemans W, Goethals B, Morik K, editors. *Mach Learn Knowl Discov Databases*. Berlin, Germany: Springer; 2008. p. 313–25.

[13] Mohammad A, Alolayyan M, Al-Daoud K, Al Nammas Y, Vasudevan A, Mohammad S. Association between Social Demographic Factors and Health Literacy in Jordan. *Journal of Ecohumanism*. 2024; 3(7): 2351-2365.

[14] Abeel T, Helleputte T, Van de Peer Y, Dupont P, Saeys Y. Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. *Bioinformatics*. 2010;26(3):392–8.

[15] Mohammad A, Shelash S, Saber T, Vasudevan A, Darwazeh N, Almajali R. Internal audit governance factors and their effect on the risk-based auditing adoption of commercial banks in Jordan. *Data and Metadata*. 2025; 4: 464.

[16] Bolón-Canedo V, Sánchez-Maroño N, Alonso-Betanzos A. An ensemble of filters and classifiers for microarray data classification. *Pattern Recognit*. 2012;45(1):531–9.

[17] Mohammad A, Al-Daoud K, Rusho M, Alkhayyat A, Doshi H, Dey P, Kiani M. Modeling polyethylene glycol density using robust soft computing methods. *Microchemical Journal*. 2025; 210: 112815.

[18] Dol SM, Jawandhiya PM. Classification technique and its combination with clustering and association rule mining in educational data mining—A survey. *Eng Appl Artif Intell*. 2023; 122:106071.

[19] Mohammad A. The impact of COVID-19 on digital marketing and marketing philosophy: evidence from Jordan. *International Journal of Business Information Systems*. 2025; 48(2): 267-281.

[20] Fournier-Viger P, Gan W, Wu Y, Nouioua M, Song W, Truong T, et al. Pattern mining: Current challenges and opportunities. In: *Int Conf Database Syst Adv Appl*. Cham: Springer Int Publ; 2022. p. 34–49.

[21] Mohammad A, Mohammad S, Al-Daoud K, Al Oraini B, Vasudevan A, Feng Z. Optimizing the Value Chain for Perishable Agricultural Commodities: A Strategic Approach for Jordan. *Research on World Agricultural Economy*. 2025; 6(1): 465-478.

[22] Chen X. Efficient and scalable graph pattern mining on GPUs. In: *16th USENIX Symp Oper Syst Des Implement (OSDI 22)*; 2022. p. 857–77.

[23] Mohammad A, Mohammad S, Al Oraini B, Vasudevan A, Alshurideh M. Data security in digital accounting: A logistic regression analysis of risk factors. *International Journal of Innovative Research and Scientific Studies*. 2025; 8(1): 2699-2709.

[24] Abuowaida S. Evidence Detection in Cloud Forensics: Classifying Cyber-Attacks in IaaS Environments using machine learning. *Data Metadata*. 2025 Feb;4:699. doi: 10.56294/dm2025699.

[25] Mohammad A, Mohammad S, Al-Daoud K, Vasudevan A, Hunitie M. Digital ledger technology: A factor analysis of financial data management practices in the age of blockchain in Jordan. *International Journal of Innovative Research and Scientific Studies*. 2025; 8(2): 2567-2577.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[26] Abuowaida S. Hybrid Ensemble Architecture for Brain Tumor Segmentation Using EfficientNetB4-MobileNetV3 with Multi-Path Decoders. *Data Metadata*. 2025 Feb;4:374. doi: 10.56294/dm2025374.

[27] Mohammad A, Al-Daoud K, Al-Daoud S, Samarah T, Vasudevan A, Li M. Content marketing optimization: A/B testing and conjoint analysis for engagement strategies in Jordan. *Journal of Ecohumanism*. 2025; 3(7): 3086-3099.

[28] Alidmat O, Abu Owida HA, Yusof UK, Almaghthawi A, Altalidi A, Alkhawaldeh RS, et al. Simulation of crowd evacuation in asymmetrical exit layout based on improved dynamic parameters model. *IEEE Access*. 2025;13:7512–25.

[29] Mohammad AAS, Mohammad SI, Vasudevan A, Alshurideh MT, Nan D. On the Numerical Solution of Bagley-Torvik Equation Using the M¨ Untz-Legendre Wavelet Collocation Method. Computational Methods for Differential Equations. 2025;13(3): 968-979.

[30] Salah Z, Abu Owida H, Abu Elsoud E, Alhenawi E, Abuowaida S, Alshdaifat N. An Effective Ensemble Approach for Preventing and Detecting Phishing Attacks in Textual Form. *Future Internet*. 2024;16(11):414.

[31] Mohammad AAS, Nijalingappa Y, Mohammad SIS, Natarajan R, Lingaraju L, Vasudevan A, Alshurideh MT. Fuzzy Linear Programming for Economic Planning and Optimization: A Quantitative Approach. Cybernetics and Information Technologies. 2025;25(2): 51-66.

[32] Arabiat M, Abuowaida S, Alshdaifat N, Aburomman A, Al Henawi E, Dmour M, Elrashidi A. Enhanced accuracy of deep learning method for fruit images classification. In: *2024 25th Int Arab Conf Inf Technol (ACIT)*. 2024. p. 1–9. IEEE.

[33] Al Daboub RS, Al-Madadha A, Al-Adwan AS. Fostering firm innovativeness: Understanding the sequential relationships between human resource practices, psychological empowerment, innovative work behavior, and firm innovative capability. International Journal of Innovation Studies. 2024;8(1): 76-91.

[34] Ehsan A, Iqbal Z, Abuowaida S, Aljaidi M, Zia HU, Alshdaifat N, Alshammry NK. Enhanced Anomaly Detection in Ethereum: Unveiling and Classifying Threats with Machine Learning. *IEEE Access*. 2024.

[35] Al-Adwan AS. The meta-commerce paradox: exploring consumer non-adoption intentions. Online Information Review. 2024;48(6): 1270-1289.

[36] Abuowaida S, Abu Owida H, Alsekait DM, Alshdaifat N, AbdElminaam DS, Alshinwan M. UltraSegNet: A Hybrid Deep Learning Framework for Enhanced Breast Cancer Segmentation and Classification on Ultrasound Images. *Comput Mater Continua*. 2025;83(2):3303–33. Available from: http://www.techscience.com/cmc/v83n2/60600

[37] Albelbisi NA, Al-Adwan AS, Habibi A. Self-regulated learning and satisfaction: A key determinants of MOOC success. Education and Information Technologies. 2021;26(3): 3459-3481.