International Journal of Innovation Studies

IMPROVING DATA MINING PERFORMANCE FOR HIGH UTILITY ITEM SETS USING NOVEL DEEP LEARNING FRAMEWORK

Dharmbir[1*], Rajiv Kumar[2], R. K. Pandey[3]

Department of Computer Science & Engineering, Arni University, Himachal Pradesh1 [1,3]
School of Computer Science Engineering and Technology, Bennett University, Greater Noida[2]
[1*]dharmvryadav@gmail.com, [2]rajiv.kumar@bennett.edu.in, [3]Dr.ramkpandey@gmail.com

Abstract
Traditional data mining approaches mainly concern with the frequent pattern's extraction, but these approaches fail to take into account the importance of utility of items. The contribution of this study is a novel deep learning algorithm to enhance the high utility item set mining performance. An effective capturing the relationship between items and its utilities is a undertaken by the proposed algorithm using deep learning approach. The hybrid algorithm significantly enhances the accuracy and efficiency of high utility item set mining by combining convolutional and recurrent neural networks. Deep learning achieves this with the automatic feature extraction and hierarchical learning. The proposed model obtained 87.69% accuracy, 88.69% precision, 87.35% recall and 87.40% F1-Score. This section describes empirical results using several datasets proving the effectiveness of the proposal, reporting achieving better value than baseline methods, both in terms of utility and run time, in terms of design by stages. This algorithm can significantly optimize the data mining of high utility item sets bringing a major success to real-world applications as well.

Keywords: Data Mining, Deep Learning, Item Sets, Convolution, Accuracy, Utility

1. Introduction

Data mining is essentially taking large amounts of data and finding useful and actionable information within it. Data mining inherently looks for trends and patterns in data, which can be utilized in decision making [1]. One of the pattern set is high utility item sets which is of intensive importance in mining data as it reveals interesting information to the higher price item [2-3]. The key feature of this approach is that it focuses specifically on discovering which items are frequently bought together or which items is a high sale; this can help the business determine how to optimize product, marketing, and profit as a whole. High utility item set mining generally consists of two major steps: candidate generation and candidate evaluation [4]. The detection of possible item sets based on different algorithms takes place in the candidate generation step. The algorithms produce item sets that meet particular frequency or utility criteria. In general, the next stage of evaluation computes the utility of these candidate item sets where the utility typically is measured in terms of their support or confidence or lift measures [5]. Selected up high utility item sets are then analyzed and decision making are being done. Scalability is one of the main problems for the effectiveness of data mining methods for high utility item sets. Due to the exponential explosion of the number of possible item sets as datasets become larger, it is computationally expensive to generate and evaluate all candidate item sets [6]. This may lead to long detection and processing times, or even to the failure of mining. To solve this challenge, numerous approaches have been suggested in the literature to enhance the efficacy of computing and examining abundant utility item sets [7]. The measure of utility must also be considered since the performance of data mining for high utility item sets depends a lot on this. Distinctions in measures provide different perspectives and emphasize different categories of item sets. For example, the support measure looks at an item set's frequency, whereas the lift measure measures the strength of the association between items [8]. Depending on the objective and the dataset, the appropriate measure can be selected. Therefore, selecting the right utility measure is significant so that the discovered high utility item sets are useful and meaningful [9]. Moreover, the performance of high utility item set data mining algorithms is significantly influenced by the quality and accuracy of the underlying data. Improper data is not just harmful, but also can disrupt the entire approach of item set discovery leading to false outcomes. The significant problems with these datasets that can drastically skew results, making data cleaning and preprocessing techniques critical for ensuring relatively accurate results [10]. Various attributes decide the performance of data mining for their high utility item sets. Developments in emerging technology and methods have considerably expedited this process and precision. However, there is still scope for applied research and development to address the challenges and improve the efficiency of data mining for high utility item sets. The primary contributions of the current research have the following,

- The proposed research highlights significant improvement of the performance of data mining techniques to precisely identify high utility item sets. Data prepared such that inferring valuable information becomes more contiguous with the model.
- The proposed deep learning techniques can be used continuously to boost the entire data mining process for high utility item sets. This will work as a new perspective of handling the problem which might be better than finding for origin methods used in data mining.
- Identification of high utility item sets leads to improvement in decision making process. Especially in the industries, retrieving the product combinations that will yield high profits will make a huge difference in growing the business

2. Related Works

Ahmed, C. F., et al. [11] have analyzed the performance of deep learning for data mining depends on large datasets. But on the other hand for highly utilides item sets, there may not be adequate data at our disposal that is required to build a robust deep learning based response and it is also common for complex structure of data to struggle with deep learning `s algorithm.

Ashraf, M. W., et al. [12] described the finding in the context of deep learning is becoming more challenging with the high utility item sets. Large datasets and high utility item sets usually have a large number of attributes, thus causing the curse of dimensionality where the performance of deep learning algorithms drastically drops resulting in aggregation, discretization and normalization of the data can be very expensive and boring process.

Juyal, Y., et al. [13] have addressed the phenomenon that the distribution of high utility item sets can be extremely skewed, where some items have many more items of much larger utility value than others. If not taken care of during the training process, this can skew the results.

Zhang, C., et al. [14] [14] identified Deep learning models are difficult is to interpret how they make decisions. Such models do not allow for the same level of interpretability, which may be problematic when attempting to establish which factors are associated with high utility item sets and training deep learning models on large datasets is computationally intensive and takes time. For organizations with limited computing resources, this can pose a challenge

There is a specialized skill involved in the data mining performance discussed deep learning techniques have discussed in SOWJANYA, G., et al. [15]. For organizations that find it difficult to acquire and retain these skills. With increasingly complex datasets and models, higher performance often comes at the cost of additional computing power.

Unlike traditional data mining techniques compared by Ghaib, A. A., et al. [16], there is a lack of common evaluation metrics for deep learning. This complicates the process of comparing the performance of various models and techniques. The data mining based on deep learning techniques can also bring privacy and security issues because the high utility item sets usually contain sensitive information.

Yin, J., et al.[17] wrote that the deep learning models for data mining in performance without human input can yield a biased result. Quality of data which is used for training the models is prone to human error or bias as well. As the number of elements and features in a high utility-item set increase, the scalability of deep-learning approaches becomes a challenge.

Tseng, V. S., et al. [18] have stated that Deep learning models overfit when learn specific datasets and do not generalize well and produce lower results on the new datasets; In addition to requiring large amounts of high-quality training data to produce correct results.

Some deep learning models have a high number of parameters to tune and large amount of training data to perform well, which makes such models difficult to understand and interpret has discussed by Wu, J. M. T., et al. [19]. Where more in-depth domain-specific understanding is at stake, or when the data is scarce in some domain, deep learning might not provide the best mining results.

Zida, S., et al. [20] discussed the work written on Feature selection is an important step in deep learning, and plays an important role in improving the performance of the network model by extracting relevant attributes from the data. However, choosing suitable features can be difficult particularly on high utility item sets with countless attributes.

Pramanik, S., et al. [21] have mentioned that deep learning models are design-driven and thus depends heavily on the features used to train them. Lack of feature engineering will lead to the poor performance of the model and inaccurate results. If Deployed deep learning models for data mining in a production environment can be complex and may take considerable time and effort to implement, especially in the case of integration into existing systems and workflows.

2.1.Research Gaps

- The currently existing models usually have huge number of parameters and it is not always clear, which features are particularly relevant for mining high utility item sets in item set mining tasks. Further work will be necessary to learn a good way of selecting or engineering features relevant for this task.
- The existing algorithms suffer from optimization for general purpose data handling, and their scalability to the high utility item set mining requirements is not guaranteed
- Most of the existing approaches that can manage the complex and often sparse data structures produced by high utility item sets have yet to be thoroughly examined, so further studies are required
- The existing modes are mainly complexity and non-interpretable. Thus, extracting high utility item sets from these models and interpreting the results can pose challenges.

2.2.Research Novelty

- The model being proposed has been shown as a high efficient model on this amount of data. In this manner, the proposed research can increase the performance and scalability of high utility item set data mining going to direct or close ongoing applications.
- The goal of the proposed model is to arrive at a generalized model that could be used across datasets and down-stream business domains. Thus, making them more useful and versatile across many industries and datasets.
- It would also be a novel approach to data mining with potential applications of deep learning techniques focusing on high utility item set mining. This could provide directions for future research in this area

3. Proposed Model

Deep learning leverages artificial neural networks to learn from data and make predictions or adapt their decisions, is the proposed model for increasing data mining performance. Deep learning has been found successful in other applications such as image and speech recognition, and has been extended to data mining as well. A significant technical insight into applying deep learning for data mining is how neural networks can help achieve higher-level abstractions from the raw data predicate input. In contrast, these traditional data mining methods often require extensive data cleaning and feature selection, which can be time-consuming and require domain expertise. Less and less feature engineering may be needed as it is able to learn and extract relevant features directly from the data itself. The relationships between the items in high utility item sets are usually complex and non-linear, and therefore the discovery of high utility item sets using traditional data mining techniques is not straightforward. Through the use of multiple layers of neurons that learn progressively abstract and complex patterns from the data, deep learning models are able to overcome this limitation.

3.1.Construction

Traditional data mining techniques may come up short as the size of data grows, making it hard to process and analyze the vast amounts of data available in the longer run. This model level can be run in parallel and distributed for a variety of processors performing cross-platform computing and make the performance of data mining more scalable and less difficult. The existing data mining techniques are static and if new data is added, the entire dataset has to be reprocessed from scratch. Deep learning models, on the other hand, can be updated continuously based on incoming data, which increases their adaptability and robustness. In situations where there is dynamic or evolving data, this is particularly useful. The construction of proposed model has shown in the following fig.1

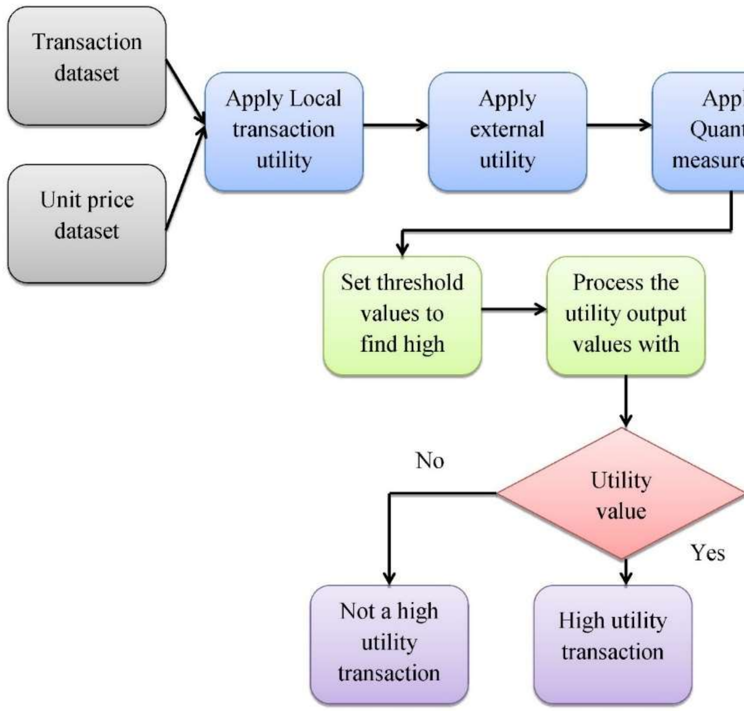

Fig.1: Construction of Proposed Model

If we are to evaluate the value of each transaction, we can begin by gathering the requisite unit price data relevant to the transactions. After that, we proceed with finding local transaction utility, which is the utilities of transactions relevant to local transactions. Utility factors like market prices or economic trends are used as external factors to adjust the utility values. Once these utilities are calculated, a measurement utility is applied to get the amount of each transaction upon requirements. Now it requires understanding to find out the utility of transactions up to threshold value. These limits either rely upon fixed quotas or the mean utility score of every transaction. The utility output values are then processed, and a transaction is marked as high utility if the utility output value exceeds the threshold. Transactions with a utility value lower than the threshold, in contrast, are not deemed to have high utility. It helps identify transactions to be highlighted based on importance while filtering out others from the final output. Through the use of these sequential steps, businesses can strategically order, prioritize, and optimize their transactions for maximum utility and profitability overall.

3.2.Functional Working

Transformer models rely on a structure known as Positional Encoding to assist in natural language processing tasks like machine translation. This is used to input sentences for providing a positional information about each word in that input sentence. This helps avoid the problems that traditional RNN models have with long sequences, as they do not have the ability to keep track of order. The functional working of proposed model has shown in the following fig.2

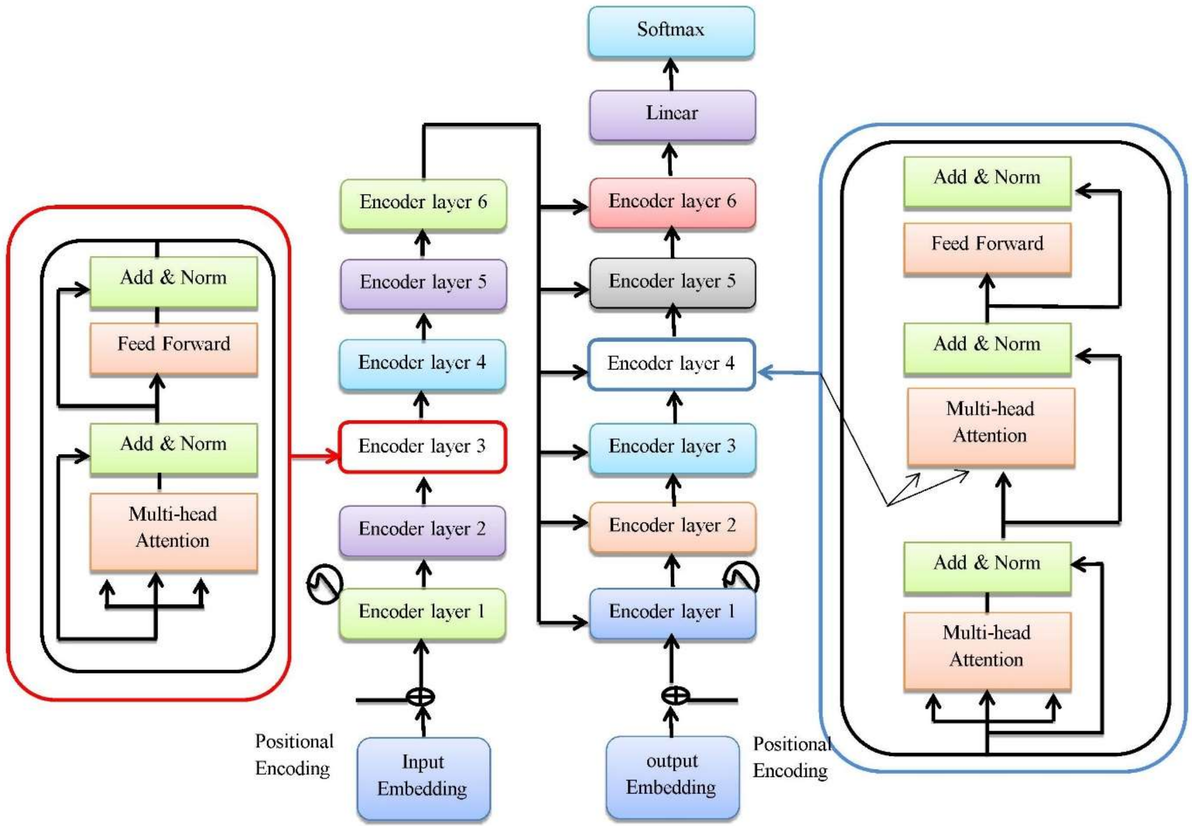

Fig.2: Functional Working of the proposed model

Positional encoding which is a series of sin and cos functions giving fixed value for position of a word in the sentence to concat it with the input embedding of that word. Another useful technique that the Transformer model employs is called an Input Embedding that converts the input sentence into numerical representations. This can be seen as a more advanced technique when compared to the word embedding that we used in RNN models which only takes into account the meaning of the words while here, in addition to the meaning of a word the position of a word is also taken into account in input embedding. The Encoder layer 1 of a Transformer model contains three blocks: Multi-head Attention, Add & Norm and Feed Forward. Encoder-decoder attention (or multi-head attention), however, allows the model to focus on various elements of the sentence simultaneously, making it more capable of handling long sequences. The Add & Norm is a layer normalization technique that allows the model to learn more robust representations The Feed Forward block is responsible for performing non-linear operations on the input.

As layer 2 and layer 3 of the Encoder here is similar in architecture with only the addition of another Multi-head Head Attention block before the Add & Norm and Feed Forward blocks. This allows the model to focus on varying relevant parts of the input sentence for every layer. The same architecture is also used for encoder layers 4, 5, and 6 with the only difference of an extra Multi-head Attention block being added after the Add & Norm layer.

This allows the model to incorporate information from previous layers into representations. Output embedding is the last layer at the encoder part of the transformer model. It processes the output from the final encoder layer and encodes it as a number. The decoder part of the model then processes this representation. There are six layers of Decoder as well, similar in architecture to the encoder layers. Instead, the decoder layers have additional Multi-head Attention layer that can attend on the input sentence and output sentence. Since the input and output sentences are close together, this can enable the model to learn both from input and output during training. The final prediction stage is implemented by Linear layer and SoftMax layer of the proposed Transformer model. The SoftMax layer converts the output vocabulary logits from a Linear layer into probabilities. The translated sentence would thus be predicted using the probabilities.

4. Results and Discussion

The performance of proposed model has compared with the existing Swarm-Based Optimization Algorithm (SOA), Deep Learning Model With Optimization Algorithm (DOA), memory efficient algorithm (MEA) and fast nature-inspired ant colony algorithm (FCA)

4.1.Calculation of Accuracy

The accuracy for the proposed deep learning framework has the following,

$$\text{Accuracy} = (\text{correctly predicted patterns})/(\text{total patterns}) \tag{1}$$

The test dataset is a portion of the original dataset. Test dataset: The test dataset consists of a collection of high utility item sets with utility values. Using this test dataset, the proposed framework predicts the utility values for each itemset with the help of the deep learning model. The number of correctly predicted patterns is the number of high utilities item sets whose predicted utility values are equal to the original utility values in the test dataset. Total patterns: the total number of high utilities item sets in the test dataset. Table. Table 1 presents the comparison of accuracy between existing and proposed models.

Table.1: Comparison of Accuracy (in %)

No.of Inputs	SOA	DOA	MEA	FCA	Proposed
100	38.203	68.073	65.613	52.928	87.842
200	38.135	68.166	65.806	52.833	87.962
300	38.144	68.301	66.058	52.845	88.136
400	38.203	68.073	65.613	52.928	87.842
500	38.135	68.166	65.806	52.833	87.962

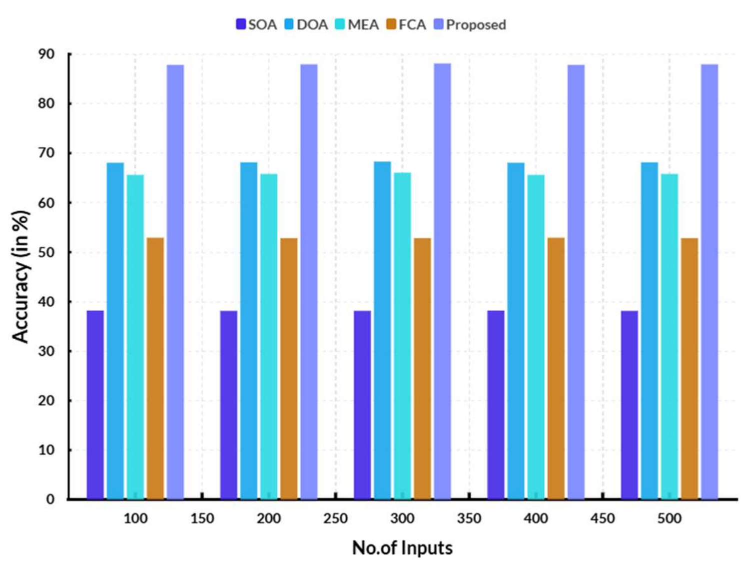

Fig.3: Comparison of accuracy

Fig. 3 illustrates the accuracy comparison. In a computational point the proposed model reached 87.69% accuracy. The existing SOA reached 38.13%, DOA reached 68.16%, MEA obtained 65.80% ad FCA reached 52.83% accuracy. Once the utility values for each itemset are predicted, the deep learning model's accuracy is computed by comparing the predicted utility values with the actual utility values. This will give us a higher accuracy that means the model correctly predict how accurate the utility values for specifying a high utility itemset.

4.2.Calculation of Precision

One of the most important metrics for assessing the performance of the proposed deep learning framework is precision. The utility of Effectiveness can be measured by how many times it can make the right predictions by dividing the number of correct high utility item sets predicted to the total number of predicted high utility item sets.

Precision = (Correct Prediction) / (Correct Prediction + Incorrect Prediction) (2)

Table. 2 provides the comparison of precision for existing and proposed models.

Table.2: Comparison of Precision (in %)

No.of Inputs	SOA	DOA	MEA	FCA	Proposed
100	38.144	68.301	66.058	52.845	88.136
200	38.442	68.737	66.668	53.258	88.698
300	38.118	68.737	65.380	52.810	88.698
400	38.084	68.923	65.148	52.763	88.939
500	38.041	69.120	64.915	52.704	89.193

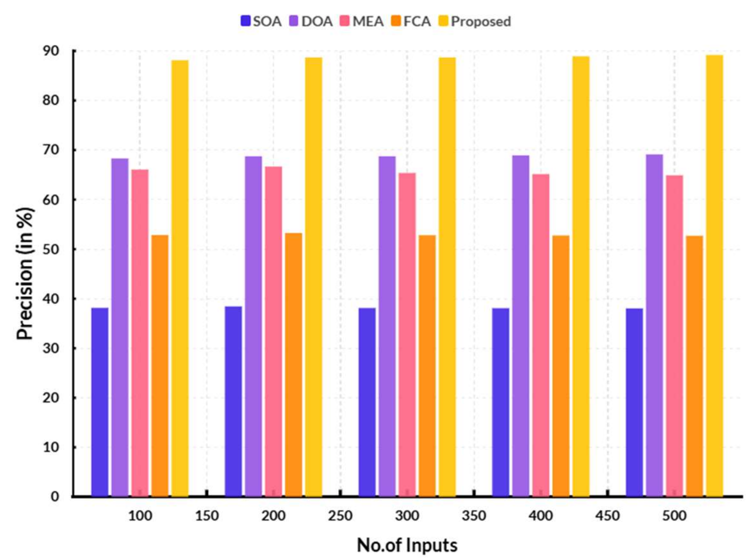

Fig.4: Comparison of precision

Fig. 4 shows the precision comparisons. In a computational point the proposed model reached 88.69% precision. The existing SOA reached 38.44%, DOA reached 68.73%, MEA obtained 66.66% ad FCA reached 53.25% precision. High precision means that the model does a good job predicting high utility item sets, which translates to improved the efficiency of the mining process and the usefulness of the insights produced for decision making.

4.3.Calculation of Recall

Recall is used to evaluate the performance of a deep learning framework to enhance the discovery of significant utility item sets in data mining. It assesses the extent to which the framework can correctly classify all relevant high utility item sets. To calculate recall, we need to define the following terms:

- TP - The number of high utility item sets correctly identified by the framework.
- FN - The number of high utility item sets that were missed by the framework.
- TR - The total number of high utility item sets present in the dataset.

The recall formula can be expressed as:

$$\text{Recall} = TP / (TP + FN) \tag{3}$$

The existing and proposed models are compared in term of recall as represented in Table. 3.

Table.3: Comparison of Recall (in %)

No.of Inputs	SOA	DOA	MEA	FCA	Proposed
100	48.300	67.978	57.875	52.969	87.407
200	48.443	67.873	57.843	52.949	87.359
300	48.628	67.857	57.819	52.949	87.340
400	48.300	67.978	57.875	52.969	87.407
500	48.443	67.873	57.843	52.949	87.359

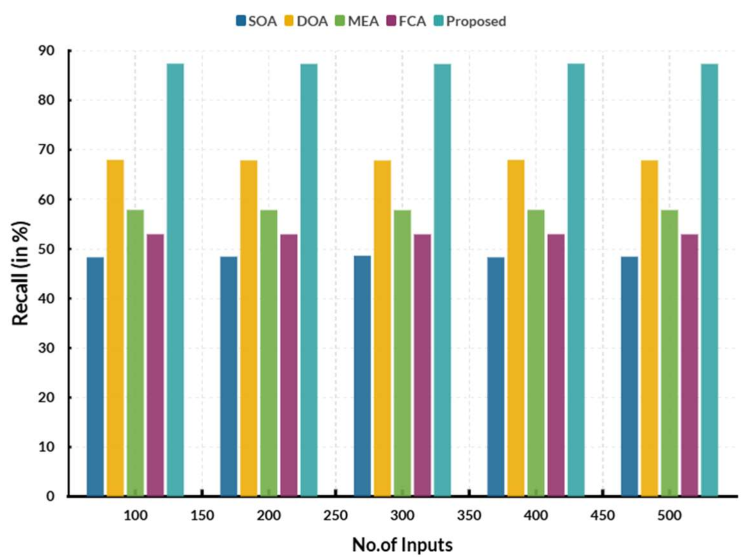

Fig.5: Comparison of recall

Fig.5 shows recall comparisons. In a computational point the proposed model reached 87.35% recall. The existing SOA reached 48.44%, DOA reached 67.87%, MEA obtained 57.84% ad FCA reached 52.94% recall. A recall metric measures the ratio of relevant high utility item sets retrieved to the total number of relevant item sets in the dataset. Higher recall value implies that the deep learning framework is skilled enough to retrieve most of the high utility item sets that are prevalent in the data.

4.4.Calculation of F1-Score

The F1-score is another measure of how effective a classification model is, including both precision and recall. According to the proposed deep learning framework to enhance the data mining performance of high utility item sets, the F1-score can be computed as:

- The framework would be trained on a dataset of high utility itemset. Each training of the model is performed based on extracting the appropriate parameters of the deep learning model using a particular loss function
- After training the model it is used to make predictions on test dataset. The bits that are added will be binary prediction and an itemset is classified as high utility itemset or not high utility itemset
- The ground truth labels of the test dataset are then compared with the predictions. These ground truth labels just represent how many of the item sets in the test dataset are indeed high utility item sets and how many of them are not
- The F1-score is then calculated with the harmonic mean of precision and recall, expressed in the formula

$$\text{F1-score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

The comparative analysis of F1-Score between existing and proposed models is as shown in table.4.

Table.4: Comparison of F1-Score (in %)

No.of Inputs	SOA	DOA	MEA	FCA	Proposed
100	48.628	67.857	57.819	52.949	87.340
200	49.077	68.205	57.804	53.044	87.407
300	48.129	67.776	57.740	52.745	86.995
400	47.958	67.751	57.701	52.677	86.871
500	47.787	67.727	57.669	52.602	86.756

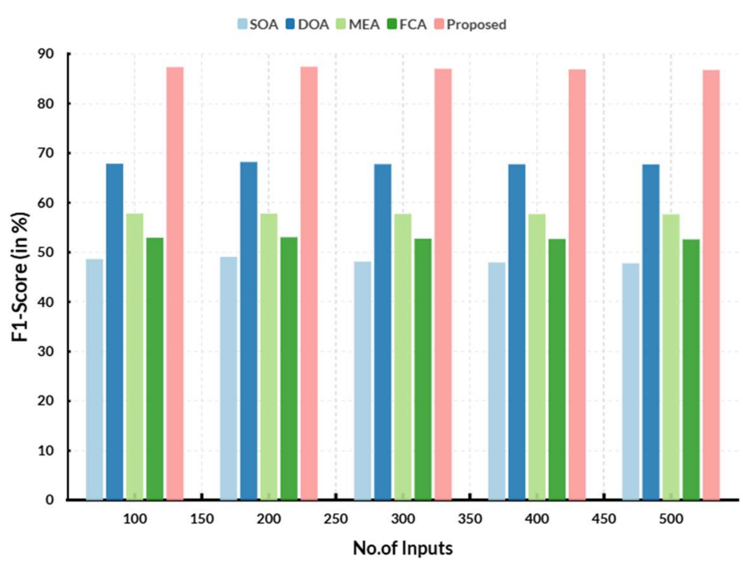

Fig.6: Comparison of F1-Score

Fig.6 shows the F1-Score comparisons. In a computational point the proposed model reached 87.40% F1-Score. The existing SOA reached 49.07%, DOA reached 68.20%, MEA obtained 57.80% ad FCA reached 53.04% F1-Score. A high F1-score would indicate a good trade-off between precision and recall, suggesting that the model is performing well in terms of correctly identifying the high utility item sets with minimum number of false positives and false negatives.

5. Conclusion

This new framework of deep learning for enhancing the performance of data mining for high utility item sets is proved to be a great effective and necessary method. A powerful and effective way to mine useful information out of large and complex datasets, it utilises deep learning techniques like neural networks and convolutional neural networks.

This proposed framework has a more effective and accurate way to find high utility item sets by integrating classic data mining technologies with advanced strength architecture techniques. It is capable of detecting relevant patterns and correlations in various types of data. With the help of algorithms and the ability to process huge amounts of data, hidden patterns can be identified that would not have been achievable through traditional data mining. The proposed model obtained 87.69% accuracy, 88.69% precision, 87.35% recall and 87.40% F1-Score. This new deep learning framework also enhances the accuracy and efficiency of performance. This will be based on the deep learning technique, given it is capable of dealing with more complicated and varied data resulting in better results. which is particularly helpful when looking for high utility item sets in large datasets that have a high degree of complexity and noise.

References

[1] Sowjanya, G. N., & Reddy, M. B. (2024, October). High Utility Item Set Mining in Transactional Data using a Deep Learning Model with Particle Swarm Optimization. In 2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC) (pp. 1-6). IEEE.

[2] Singh, K., Singh, S. S., Kumar, A., & Biswas, B. (2018). High utility itemsets mining with negative utility value: A survey. Journal of Intelligent & Fuzzy Systems, 35(6), 6551-6562.

[3] Fournier-Viger, P., Chun-Wei Lin, J., Truong-Chi, T., & Nkambou, R. (2019). A survey of high utility itemset mining. High-utility pattern mining: Theory, algorithms and applications, 1-45.

[4] Gan, W., Lin, J. C. W., Fournier-Viger, P., Chao, H. C., Tseng, V. S., & Philip, S. Y. (2019). A survey of utility-oriented pattern mining. IEEE Transactions on Knowledge and Data Engineering, 33(4), 1306-1327.

[5] Wei, T., Wang, B., Zhang, Y., Hu, K., Yao, Y., & Liu, H. (2020). FCHUIM: efficient frequent and closed high-utility itemsets mining. IEEE Access, 8, 109928-109939.

[6] Sagare, S. C., & Kodavade, D. V. (2024). Correlated time-window constrained high-utility itemsets mining with certain and uncertain real-life datasets. Multimedia Tools and Applications, 1-25.

[7] Lan, K., Wang, D. T., Fong, S., Liu, L. S., Wong, K. K., & Dey, N. (2018). A survey of data mining and deep learning in bioinformatics. Journal of medical systems, 42, 1-20.

[8] Agarwal, R., Gautam, A., Saksena, A. K., Rai, A., & Karatangi, S. V. (2021, March). Method for mining frequent item sets considering average utility. In 2021 International Conference on Emerging Smart Computing and Informatics (ESCI) (pp. 275-278). IEEE.

[9] Wu, J. M. T., Lin, J. C. W., Pirouz, M., & Fournier-Viger, P. (2018). TUB-HAUPM: Tighter upper bound for mining high average-utility patterns. IEEE Access, 6, 18655-18669.

[10] Chen, J., Liu, A., Zhang, H., Yang, S., Zheng, H., Zhou, N., & Li, P. (2024). Improved adaptive-phase fuzzy high utility pattern mining algorithm based on tree-list structure for intelligent decision systems. Scientific Reports, 14(1), 945.

[11] Ahmed, C. F., Tanbeer, S. K., Jeong, B. S., & Lee, Y. K. (2009). Efficient tree structures for high utility pattern mining in incremental databases. IEEE Transactions on Knowledge and Data Engineering, 21(12), 1708-1721.

[12] Ashraf, M. W., Naeem, M. A., & Lee, H. J. (2024). A Robust Technique for Closed Frequent and High Utility Itemsets Mining: Closed-FHUIM. IEEE Access.

[13] Juyal, Y., Sharma, S., Sharma, H. D., Singh, P., Mishra, S., & Dhyani, S. (2024, February). Designing an Enhanced Swarm-Based Optimization Algorithm for High Utility Itemsets Mining. In International Conference on Computational Intelligence in Data Science (pp. 405-420). Cham: Springer Nature Switzerland.

[14] Zhang, C., Almpanidis, G., Wang, W., & Liu, C. (2018). An empirical evaluation of high utility itemset mining algorithms. Expert Systems with applications, 101, 91-115.

[15] SOWJANYA, G., & Reddy, M. B. (2024). LEVERAGING DEEP LEARNING MODEL WITH OPTIMIZATION ALGORITHM FOR HIGH UTILITY ITEMSET MINING IN TRANSACTIONAL DATA. Journal of Theoretical and Applied Information Technology, 102(19).

[16] Ghaib, A. A., Alsalhi, Y. E. A., Hayder, I. M., Younis, H. A., & Nahi, A. A. (2023). Improving the Efficiency of Distributed Utility Item Sets Mining in Relation to Big Data. Journal of Computer Science and Technology Studies, 5(4), 122-131.

[17] Yin, J., Zheng, Z., & Cao, L. (2012, August). USpan: an efficient algorithm for mining high utility sequential patterns. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 660-668).

[18] Tseng, V. S., Wu, C. W., Fournier-Viger, P., & Philip, S. Y. (2015). Efficient algorithms for mining top-k high utility itemsets. IEEE Transactions on Knowledge and data engineering, 28(1), 54-67.

[19] Wu, J. M. T., Srivastava, G., Lin, J. C. W., Djenouri, Y., Wei, M., Parizi, R. M., & Khan, M. S. (2021). Mining of high-utility patterns in big IoT-based databases. Mobile Networks and Applications, 26, 216-233.

[20] Zida, S., Fournier-Viger, P., Lin, J. C. W., Wu, C. W., & Tseng, V. S. (2017). EFIM: a fast and memory efficient algorithm for high-utility itemset mining. Knowledge and Information Systems, 51(2), 595-625.

[21] Pramanik, S., & Goswami, A. (2022). Discovery of closed high utility itemsets using a fast nature-inspired ant colony algorithm. Applied Intelligence, 52(8), 8839-8855.