IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 15, No. 4, August 2026, pp. 3240~3251
ISSN: 2252-8938, DOI: 10.11591/ijai.v15.i4.pp3240-3251

Resource optimization via identification of articles with aspects of priority and high utility

Garima Srivastava[1,2], Vaishali Singh[1], Sachin Kumar[3]
[1]Department of Computer Science Engineering, Maharishi School of Engineering and Technology, Maharishi University of Information Technology, Lucknow, India
[2]Department of Computer Science Engineering, Amity School of Engineering and Technology, Amity University Uttar Pradesh, Lucknow, India
[3]Department of Electronics and Communication Engineering, Amity School of Engineering and Technology, Amity University Uttar Pradesh, Lucknow, India

Article Info

Article history:

Received Apr 5, 2025
Revised May 20, 2026
Accepted Jun 19, 2026

Keywords:

Faster high-utility itemset miner
Frequent itemset mining
High utility pattern
Score
Triples
Vanilla neural network

ABSTRACT

Owing to the linear economic growth, the manufacturing of consumer products has increased manifold, products in diversified packaging and formats are now available in abundance, sometimes even with less or no requirement, resulting in losses. Addressing the gap that exists between productivity and requirements can minimize losses and unnecessary burden on manufacturing units. To minimize losses, a hybrid algorithm is proposed, using the advantages offered by classification and high utility patterns to identify the dominant aspects of articles with a high probability of sale. Articles are identified in a two-step process, identification of samples with context as prominent parameter by vanilla feedforward neural network (VNN) as a first step. The second step comprises the determination of utility pattern mining in articles using faster high-utility itemset miner (FHN), sentiment score obtained provides the utility pattern of the product. Proposed hybrid VNN + FHN, along with convolutional neural network (CNN), long short-term memory (LSTM), and transformer-based prediction model (TPM) were used for assessing the utility-driven product analysis. The hybrid VNN + FHN proposed displays the best adaptability by extracting 158 patterns with a utility of 162.88 units. The algorithm outperforms CNN in utility, LSTM and TPM in speed, and CNN in pattern count, making it a better choice for resource optimization.

Corresponding Author:

Garima Srivastava
Department of Computer Science Engineering, Maharishi School of Engineering and Technology
Maharishi University of Information Technology
Lucknow, India
Email: gsrivastava1@lko.amity.edu

1. INTRODUCTION

Addressing targets, requirement and availability with consumer products is a tricky task, linearly proportional relation between manufacturing, supply chain and requirement can minimize losses and increase financial benefits. Frequently a wide gap exists between productivity and requirement, getting accurate feedback and addressing the gap effectively can minimize burden on manufacturing units. Rating, recommendations, reviews and feedback can provide quite insightful data about products in demand and can streamline proportionality between demand and supply. Plethora of diverse recommendation systems based on opinion mining techniques are available for resource optimization, however, there exists a wide gap in targeted approaches employing hybrid algorithms assisted by statistical techniques [1]. Currently, recommender systems (RS) are being utilized extensively in several industries, including online shopping, social media, electronic news channels, online music portals, travel planning, and entertainment. This is due to the growth and diversification of products and services [2], [3].

Traditional RS focus on accuracy and similarity, but they fail to incorporate item profitability, aspect-level preferences, or real utility constraints. Likewise, classical pattern-mining techniques such as apriori or frequent pattern (FP)-growth capture only frequency, not utility, leading to the omission of high-value but low-frequency patterns. High-utility itemset mining addresses this gap by modeling utility through quantitative factors such as profit, weight, and importance. Recent research shows that combining high-utility pattern mining (HUPM) with aspect-based opinion mining enables more meaningful feature-utility associations [4], [5], hybrid RS–HUPM models [6]–[8] therefore outperform earlier single-objective approaches by simultaneously capturing what users prefer, what yields value, and which item aspects drive priority, making them increasingly suitable for product selection, catalog optimization, and personalized decision support.

Manufacturing of consumer products has increased over the last decade, products in diversified packaging and formats are now available in abundance, sometimes even with less or no requirement resulting in losses. A wide gap exists between productivity and requirement, addressing the gap effectively can minimize losses and unnecessary burden on manufacturing units. Plethora of diverse recommendation systems based on opinion mining techniques are available for resource optimization, however, there exists a gap in targeted approaches employing hybrid algorithms.

Work proposed employs a hybrid approach of classification and high utility pattern to identify the dominant aspects of articles with high sale probability. The first step comprises of identification and classification of samples with context as prominent parameter, the second step takes output of first step as input, step comprises of opinion assisted mining for determining high utility pattern in articles, score rating is estimated for identifying samples with high utility from triples extracted from the labeled corpus. E-commerce items have been used to validate the proposed method [9], [10].

Paper is organized as section 2 presents opinion mining techniques along with identification of high utility patterns for developing targeted approach for resource optimization. Section 3 presents objectives identified from research gaps observed in literature survey. Section 4 discusses proposed methodology employed to achieve the objective. Section 5 presents the discussion and the results achieved from the proposed algorithm and finally section 6 presents the conclusion.

2. LITERATURE SURVEY

A targeted approach for investment in article manufacturing can optimize lot of resource allocation, effort and time. Proposed methodology is an effort to identify articles of high utility and hence focus production units on commodities of larger interest. A related study was conducted to investigate aspects of articles in high demand and further to narrow down the range.

2.1. Aspect-based sentiment analysis

The three main subtasks of aspect-based sentiment analysis (ABSA) include aspect-sentiment classification, aspect-term extraction, and aspect category detection. ABSA targets deep sentiment research within a specific entity [11]. Implicit elements can be difficult to identify; however, it is of extreme importance. Unfortunately, little consideration has been given in recent studies towards the extraction of implicit elements. From various research, the authors extracted the data and summarized the issues of the sentiment analysis (SA) classification techniques [12]. Research that views the impact of different product categories introduces an approach for fraudulent reviews identification using the ABSA method [13].

Additionally, talking-head attention and a model was proposed based on a local context focus mechanism, which improves SA results compared to existing baseline models [14]. A heterogeneous aspect graph neural network (HAGNN), which captures relationships between sentences and aspects to improve model performance, was proposed [15]. The performance of the unified framework is sub-optimal. A model called residual attention gating – three-channel graph convolutional network (RAG-TCGCN) for ABSA, which uses a three-channel graph convolutional network to learn important information from sentences [16]. In order to assess the sentiment polarity of certain aspects in text reviews, the study suggests using a dependency graph convolutional network model for ABSA [17].

2.2. High utility based on context identification

Context-based SA has been approached in a variety of ways, from the usage of contextualized lexicons to the hybridization of lexicons and context-aware machine learning algorithms. Different sort of context that is linguistic, user profile, social, geographical, and temporal. May be employed in a singleton as well as in combination to produce better outcomes [18]. In order to get context information for the sentiment classification problem, a unique neural network-based model was put forth [19].

According to Hari *et al.* [20], a novel feature fusion technique for multimodal SA and context modelling improved classification results by a margin of 1-2 percent for all combinations of modalities. A study that employed the Twitter rank algorithm to identify prominent users for a topic and extracted the emotion that was most influential for that topic [21], [22] focuses on recognizing feelings from implicit phrases that do not contain emotional terms. Algorithm is a multi-feature neural network model for sentiment recognition in implicit sentences that considers contextual, syntactic, and semantic information, among other factors. In order to determine the sentiment polarity of an aspect in a particular context, Zhu *et al.* [23] suggests using a context-enhanced attention network (CE-ATT) for aspect-based sentiment categorization.

High-utility pattern analysis following the initial introduction of high utility itemset mining (HUIM) in [4], a number of methods, including as UMining [5], incremental high utility pattern mining (IHUP) [7], and utility pattern-growth (UP-Growth) [8], were put forth. Since the downward-closure attribute [8] cannot be used to reduce the search space, the issue associated with HUIM is seen to be more challenging than the issue of frequent itemset mining (FIM). The utility of an itemset is neither monotonic nor anti-monotonic in HUIM, which is the reason for the downward closure property's need that it holds. The transaction-weighted downward closure (TWDC) feature was presented in [24] as a solution to this problem, and most of the subsequent research have made use of this property to reduce the search space and improve time efficiency. Table 1 presents the evolution of opinion mining techniques over the years.

Apart from designing, creating the prototype, getting raw material, fabricating, advertising, transportation and sales, and market need also plays a crucial role. With innovations in machine learning algorithms, several works with simple classifiers such as regression, support vector machine (SVM), and naive Bayes to advanced techniques like random forest, bagging, boosting to neural networks have been applied for enhanced results. However, a wide gap still exists in optimization of resources assisted by opinion mining and utility-based selection. Further the impact of utility pattern on optimization of resource allocation in product life cycle is still a potential area which is unexplored. Research gaps identified can be summarized as follows:

- Wide gap exists due to limited work in area of optimization of resources assisted by opinion mining.
- Impact of utility pattern on resource allocation in product life cycle is still a potential area which is unexplored.
- Hybrid recommenders that identify products with priority and high utility pattern can have major financial impact.

The work proposed is an effort to address gaps identified, use of utility patterns of high frequency products in tandem with classification for effective resource utilization and hence focus production units on commodities of larger interest.

Table 1. Summary of the evolution of opinion mining techniques

S.no	Category	References	Methodology
1	Hybrid model	[25]–[30]	Hybrid bidirectional gated recurrent unit (BGRU)/bidirectional long short-term memory (BiLSTM)-gated recurrent unit (GRU) models, BERT-GRU
2	High utility mining	[5]–[8], [23], [31]–[34]	UMining, IHUP, UP-Growth, TWDC, modified efficient frequent (MEF)-HUIM, incremental high average-utility itemset mining (IncHAUIM), EHNL, high average-utility itemset (HAUI)-Miner++, efficient high-utility itemset mining (EFIM) FP-tree projection
3	Hybrid nature inspired algorithms	[35]–[40]	Hybrid convolutional neural networks (CNNs)-long short-term memory (LSTM)-SVM, particle swarm optimization (PSO) SVM, naive Bayes-PSO, HUIM+PSO+CSA
4	Aspect sentiment classification	[15]–[17], [41]–[44]	HAGNN, RAG-TCGCN, dependency graph CNN, gated recurrent units, knowledge-aware dependency graph network, prompt knowledge tuning, retrieval contrastive learning
5	Context based	[19], [23], [45]–[47]	Neural network model, CE-ATT, context-sensitive multi-tier deep learning framework, SVM with grid search, transformer models RoBERT DistilBERT
6	Collaborative filtering with utility mining	[48]–[50]	Alternating least square (ALS) with attention mechanisms and BiLSTM, GloVe+BiLSTM, Bayesian hybrid collaborative filtering (BHCF)
7	High utility with negative unit profits	[51]–[53]	FHN, global high-utility mining (GHUM), top-k high utility itemset with positive or negative profits (TKN)

3. OBJECTIVES

Addressing the identified limitations, this work aims to design a hybrid algorithm that integrates

HUPM with context-aware classification to enable targeted investment decisions in manufacturing. Based on the research gaps, the objectives of the proposed work are defined as follows:

– To design an algorithm for targeted investment by combining utility mining and opinion-based classification, achieving greater than 10–15% reduction in false-positive product selections.
– Use utility-pattern analysis to guide product life-cycle resource allocation by enforcing minimum utility thresholds.
– Implement a hybrid recommender that integrates ratings, reviews, and behavioral data to achieve improvement in recommendation precision and stronger recall of high-utility products.

4. PROPOSED METHOD

Proposed work employs a hybrid approach of classification and high utility pattern to identify the dominant aspects of articles with high sale probability. Figure 1 depicts the flow graph design of three step proposed model, flow graph functions in three layers, first layer acquires samples and normalizes the dataset for optimal execution, second step comprises of identification and classification of samples with context as prominent parameter, vanilla feedforward neural network (VNN) with dense connections is employed for the classification and in third, articles identified from second step is subjected to utility pattern mining using triples for extracting commodities and aspect of high demand. E-commerce items have been used to validate the proposed method. The following steps are employed for the task.

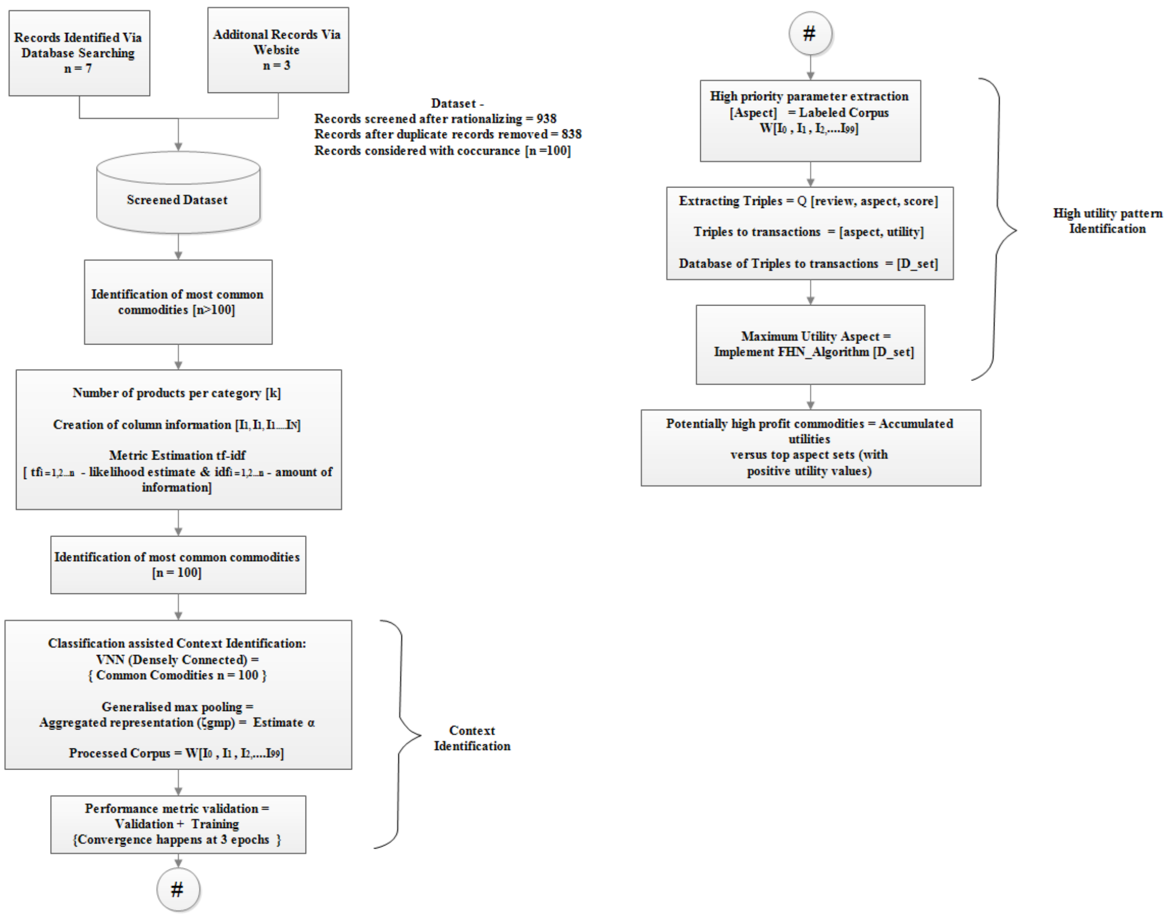

Figure 1. Flow graph design of two-step proposed model

4.1. Dataset

The product.json file, is used for the suggested work, obtained from GitHub. Sample acquired has the product names, descriptions, and categories. The product.json file contains a total of 51,646 products listed with dataset contains multiclass categories. The suggested technique can identify and forecast 1,802 classes since there are 1,802 non-overlapping classes. The dataset that is being studied is shown in

Figure 2, with the frequency of occurrence of items, category count, and a highlighted red line that separates all categories with a count of more than 500 from those without [54].

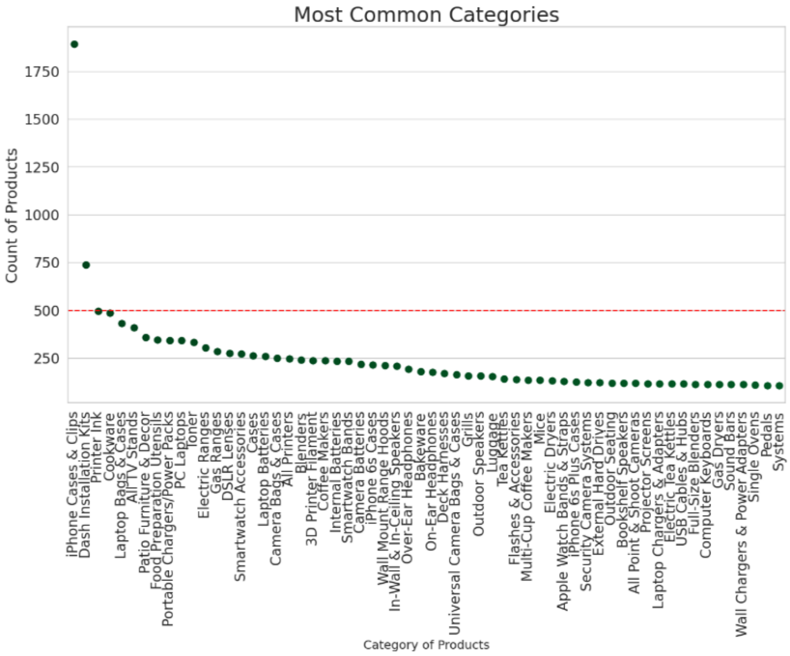

Figure 2. Frequency of occurrence of items

4.2. Deep neural network assisted priority context identification

To adapt machine algorithms for prediction, the dataset under investigation is divided, tokenized, preprocessed, and transformed into term frequency–inverse document frequency (TF-IDF). CNNs work on spatial pattern extraction and LSTMs captures sequential dependencies while VNNs acquire a latent distribution enabling better generalization, reduced overfitting, and improves interoperability through probabilistic outputs, supports generative capability and work better with smaller datasets allowing them to predict missing data, create synthetic samples, and discover deeper hidden patterns. Tasks which involve uncertainty modeling, noise tolerance, feature discovery, and probabilistic decision-making, VNNs outclass CNNs and LSTMs. Owing to reasons mentioned above VNN was used in the deep learning model, which is trained for 30 epochs and plots the losses and area under the curve (AUC). Figure 3 displays the VNN classification parameters for the proposed model.

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 500, 300)          11109300

 dense (Dense)               (None, 500, 300)          90300

 global_max_pooling1d (Glob  (None, 300)               0
 alMaxPooling1D)

 dense_1 (Dense)             (None, 69)                20769

=================================================================
Total params: 11220369 (42.80 MB)
Trainable params: 11220369 (42.80 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Figure 3. VNN classification parameters

VNN employed is with dense connections for classification [55]. A densely linked network employs feature reuse; it permits an output at one layer to be reused several times at succeeding levels [55], [56]. The quality of embedding pairs collected across the dataset is accumulated to reflect this aim in a loss function £ given by (1). Denoting the quality of a particular pair of embeddings with ij, the overall loss is defined as (1).

$$£ = \sum_{i\varepsilon I} \sum_{j\varepsilon N(i)} l_{ij} \quad (1)$$

Work proposed estimates by estimating algebraic sums for all descriptors, the algebraic sum may get escalated from errors generated due to uncorrelated descriptors that can influence the estimated similarity, even if error generated has a non-significant individual similarity. Weights is so taken that it equalizes the effect of every entity to the similarity estimates. The cumulative representation ζ can be represented in (2).

$$\zeta_{gmp} = \sum_{x\varepsilon X} \alpha(x) \beta(x) = \psi\alpha \quad (2)$$

The procedure to solve an optimization problem to estimate the weights α is known as generalized max pooling. Work proposed evaluates algebraic sums for all descriptors, the algebraic sum may get escalated from errors generated due to uncorrelated descriptors that can influence the estimated similarity, even if error generated has a non-significant individual similarity. Figure 4 depicts the losses incurred for the proposed model, which are approximately 0.04 for validation and less than 0.01 for training. Figure 5 depicts the AUC, the curve provides the precision with which the algorithm can identify diversified classes, it is approximately 0.97 for validation and 0.99 for training, results achieved show that corpus has been labelled with acceptable error ratio [57], [58].

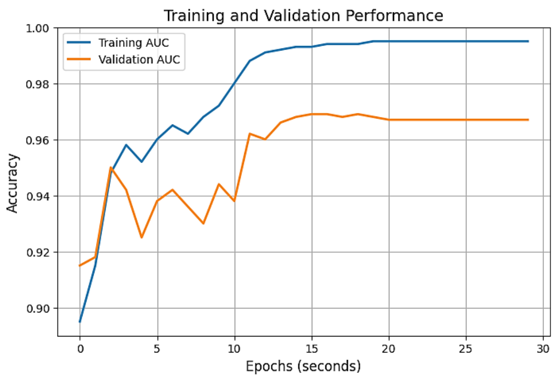

Figure 4. Losses incurred for the proposed model

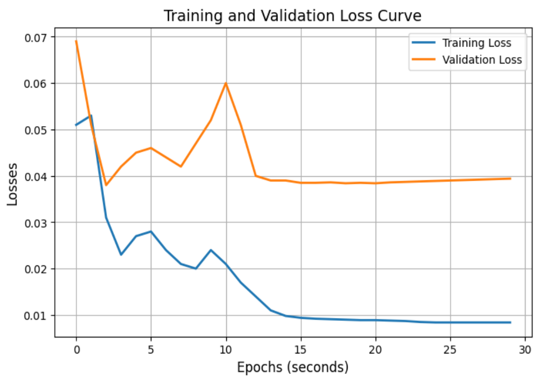

Figure 5. AUC represents the precision with which the algorithm can identify diversified classes

4.3. Pattern mining for high utility

VNN provides a processed dataset where commodities with context having high priority have been labelled, commodities extracted from stage one will be subjected to the mining process to identify articles of high utility. Work first establishes the primary entities of HUPM. Let I be finite class of unique items and D gives group of transactions in a database. It is distinguished by the transaction id for each transaction, Tk. Commonly, the former has an internal utility, $q(i, Tk)$, and the latter has exterior utility, $p(i)$. $u(X)$ represents the utility value of the article X across database D, which is estimated by summing together all the components of $u(X, Tk)$ so that T_k has X as member, given in (3).

$$u(i, Tk) = p(i) \times q(i, Tk) \quad (3)$$

High utility pattern extraction seeks to identify all high utility articles in the database D that have utility larger than or equal to minimum utility, given the database D and the minimal utility criterion. Aspect-based opinion mining constructs a set having triples: (ai, sj, sc). Here, sj refers to the associated sentiment for aspect ai and sc is the sentiment score. The procedure to extract triples (ai, sj, sc) is given in Algorithm 1.

Algorithm 1. High priority parameter extraction algorithm
Input: corpus
Output: high priority parameter extraction
Step 1: aspect assisted opinion mining.
Step 2: Q ← Corpus
Step 3: E ← β (Null Set)
Step 4: for τ ε Q
Step 5: Q; E ← Acquire Triples(τ)
Step 6: E ← E Ụ EQ # Review + aspect + score → E
Step 7: End
Step 8: CT ← ClusterTriples(E) # Cluster of triples = review + aspect
Step 9: D_set ← CT # Database of triples to transactions
Step 10: MUA ←ImplementFHN_Algorithm(D_set) # Maximum utility aspect

Sample reviews are included in Figure 6 from which triples are to be extracted. To extract triples in a particular review, the work proposed followed the procedure outlined in [13], [14], which is further explained as follows:

i) Corpus is divided down into sentences.
ii) NLP tokenizer is employed for the process as follows:
   – An individual aspect is labeled for every noun and noun phrase. In addition, every word in a noun corpus is converted to its lemmatized shape, and all lemmatized word forms are connected together so that inflectional variants of nouns and noun corpus can be combined in subsequent stages.
   – Sentiment lexicons are used to annotate words with emotions.
   – Sentiment words, annotated nouns, and noun phrases are arranged in the sentence's order of occurrence.
   – Every pair of nouns and sentiment words that co-occur makes a new triple.
iii) The triple sets union is acquired, which is formed for every review sentence.

Figure 7 provides corresponding triples that are retrieved using the previously described procedure. Extracted triples are converted into transactions with utility values in this step. Items have both internal and exterior usefulness in utility-based pattern mining.

	Reviews Text	Reviews Title
0	This keyboard is very easy to type on, but the...	Love the fingerprint reader
1	I love this. This small portable speaker has a...	Great sound
2	The product works as expected, but is really b...	Great product works well, too big
3	I have nothing to complain about this remote. ...	AV receiver it works great. The vibration feed...
4	BT audio on devices was very spotty. Gave up a...	Poor BT Connection
...	...	...

Figure 6. Sample reviews

	Review ID	Triples
0	1	[(home button, good, 1), (feature, excellent, 2)]
1	2	[(battery life, poor, -2), (voice quality, good, 1)]
2	3	[(keyboard, bad, -1), (shape, excellent, 2)]
3	4	[(mouse scroll, worst, -2), (size, small, -1)]
4	5	[(keyboard, amazing, 2)]

Figure 7. Corresponding triples

A review and a transaction are related, with elements from a review taken from the first step, matching things from a transaction in the third. For every object, the work proposes to give it an external utility. This estimate, which is domain-dependent and not included in the review data, may be a producer's choice because of cheap manufacturing costs, or it could be a customer's preference. Table 2 depicts the transformation model used for transformation of triples to transactions.

Table 2. Transformation model

Utility pattern mining	Aspect assisted opinion mining
Item (Ii)	Aspect
Transaction (Ti)	Review
Utility (Ui)	Sentiment product value X utility score
External (Ui)	Utility score
Internal (Ui)	Sentiment product review

Proposed model performs two procedures on the extracted aspect to complete the transformation. Triples are categorized by review id in the first one. A review depending upon issue addressed is likely to have more than one triple with the same aspect. The sentiment score of the elements that have several occurrences are resolved by adding up each sentiment score separately or calculating their average. In the former case, it is assumed that repeated mentions of an element increase its influence or preference (i.e., positive or negative). The array of triples with the identical review ID is arranged into a transaction during the next operation of the transformation procedure. An example transformation for the reviews mentioned in Figure 6 is shown in Figure 8 [59], [60].

```
   Sentiment Product Utility Pattern
0                 15     home button
1                 -3    battery life
2                  2        keyboard
3                 -3    mouse scroll
4                  8         monitor
```

Figure 8. Score estimates for the utility pattern

5. RESULTS AND DISCUSSION

The work considered only positive utility values for aspect groups having the highest positive ratings for sentiment, mapping articles with high sale impact only; estimates were calculated by extracting triples. The proposed work employed the faster high-utility itemset miner (FHN) algorithm [51], [52], using the FHN technique, high utility aspect (HUA) groups are retrieved from transaction database. Pattern employed is a collection of useful characteristics that are indicators of high profit features or aspects, with the highest positive sentiment ratings and, thus, the largest levels of satisfaction are indicated by the patterns with the highest positive utility. Figure 9 depicts comparison of accumulated utility for FIM and proposed technique under positive utility patterns.

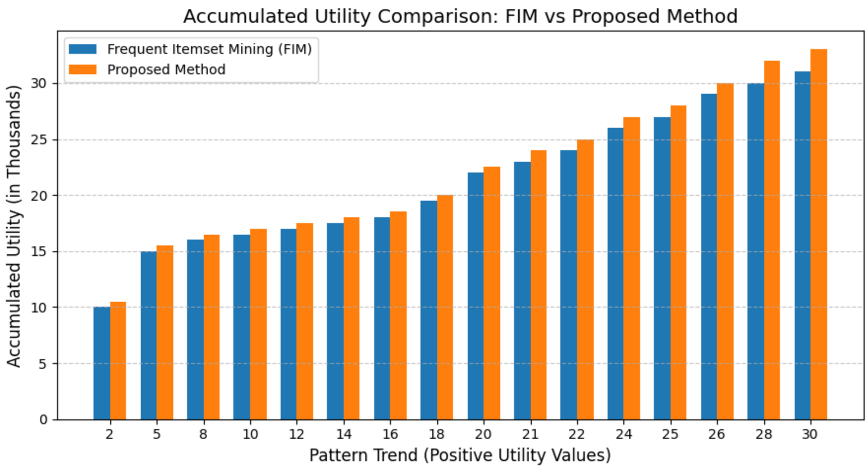

Figure 9. Comparison of accumulated utility for FIM and proposed technique under positive utility patterns

Hybrid VNN + FHN, CNN, LSTM, and transformer model were used for accessing the utility driven product analysis. CNN excels in extracting relevant features from short and structured texts, the LSTM network focuses on long-term context dependencies in product review texts, and Transformers shine with robust context embeddings that are insensitive to noises. Hybrid VNN + FHN combines deep context intelligence with efficient utilization-pattern discovery. By comparing the four algorithms, the most reliable one will be picked that balances the efficiency of product-pattern discovery, context intelligence, processing complexity, and noise robustness. The following metrics were used to assess the algorithms:

- Estimation of number of patterns (N).
- Utility of product and average utility u(pi).
- Runtime of an algorithm (T).

Table 3 depicts comparison metrics of the proposed hybrid algorithm with CNN, LSTM, and transformer-based models. It shows that the hybrid method demonstrated the strongest overall balance between utility, pattern extraction, and runtime. The weakest performance is of CNN for 143 patterns.

Table 3. Comparison metrics of proposed hybrid VNN+FHN with CNN, LSTM, and transformer model

Algorithm	Pattern count (N)	Average utility (units)	Run time (µs)
Hybrid (FHN + VNN)	158	162.88	70
CNN	143	155.46	60
LSTM	171	167.33	80
Transformer-based prediction model (TPM)	189	173.52	90

LSTM performs better with 171 patterns identified and higher utility of 167.33 with runtime of 80 µs. TPM delivers best deep-learning results, achieving the highest pattern count of 189 with utility value 173.52, however it has slowest runtime of 90 µs. The hybrid VNN + FHN technique balances these trade-offs by extracting 158 patterns with utility of 162.88 and a fast runtime of 70 µs, outperforming CNN in utility, LSTM and TPM in speed, and CNN in pattern count. In extended evaluation, hybrid VNN-FHN achieves highest overall utility of 182.4 units but required the longest runtime of 0.018 s, indicating increased computational cost with higher accuracy. CNN and LSTM produced moderate utilities of 150.6–161.2 units and occasionally misinterpreted short or noisy text, while the Transformer model maintained stable utility of 176.9 units but struggled with ambiguous or incomplete reviews. All models scaled well for 100–1,000 e-commerce products with runtimes under 0.02 s; CNN and LSTM showed the most stable runtimes of 0.010–0.013 s due to GPU-friendly architectures, whereas hybrid VNN-FHN slowed slightly as pattern size increased. In cross-domain scenarios finance, healthcare, and retail, all models can generalize with redefined utility functions, with hybrid VNN + FHN technique showing the best adaptability.

6. CONCLUSION

A targeted RS design helps in optimization of resources in terms of raw material, transportation, manufacturing cost, human resources, packaging, marketing, and sale. Proposed work employs a hybrid approach of classification and high utility pattern to identify the dominant aspects of articles with high sale probability. Proposed technique hybrid VNN + FHN was assessed and compared with CNN, LSTM, and Transformer model. LSTM performs better with 171 patterns identified and higher utility of 167.33 with runtime of 80 µs. TPM delivers best deep-learning results, achieving the highest pattern count of 189 with utility value 173.52, however it has slowest runtime of 90 µs. The hybrid VNN + FHN technique balances these trade-offs by extracting 158 patterns with utility of 162.88 and a fast runtime of 70 µs, outperforming CNN in utility, LSTM and TPM in speed, and CNN in pattern count. In cross-domain scenarios finance, healthcare, and retail, all models can generalize with redefined utility functions, with the hybrid VNN + FHN technique displaying the best adaptability.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Garima Srivastava	✓	✓	✓	✓	✓	✓		✓	✓	✓				
Vaishali Singh					✓	✓		✓		✓	✓	✓		
Sachin Kumar	✓	✓	✓	✓			✓			✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The dataset generated is available from the corresponding author, [GS], upon reasonable request.

REFERENCES

[1] H. J. Aleqabie, M. S. Sfoq, R. A. Albeer, and E. H. Abd, "A review of text mining techniques: trends, and applications in various domains," *Iraqi Journal for Computer Science and Mathematics*, vol. 5, no. 1, pp. 125–141, 2024, doi: 10.52866/ijcsm.2024.05.01.009.

[2] S. Li, B. Zhu, Y. Zhang, F. Liu, and Z. Yu, "A two-stage nonlinear user satisfaction decision model based on online review mining: considering non-compensatory and compensatory stages," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 19, no. 1, pp. 272–296, 2024, doi: 10.3390/jtaer19010015.

[3] J. Ding, M. Xu, Y. K. Tse, K. Y. Lin, and M. Zhang, "Customer opinions mining through social media: insights from sustainability fraud crisis - Volkswagen emissions scandal," *Enterprise Information Systems*, vol. 17, no. 8, 2023, doi: 10.1080/17517575.2022.2130012.

[4] H. Yao, H. J. Hamilton, and C. J. Butz, "A foundational approach to mining itemset utilities from databases," in *Proceedings of the 2004 SIAM International Conference on Data Mining*, Apr. 2004, pp. 482–486, doi: 10.1137/1.9781611972740.51.

[5] H. Yao, H. J. Hamilton, and L. Geng, "A unified framework for utility-based measures for mining itemsets," in *ACM SIGKDD 2nd Workshop on Utility Based Data Mining, Philadelphia,* 2006, pp. 28–37.

[6] V. Ramineni and L. Sudula, "Mining frequent high utility itemsets on dynamic databases without candidate generation," in *International Conference on Emerging Technologies in Engineering and Science: ICETES2023*, 2025, doi: 10.1063/5.0242321.

[7] C. F. Ahmed, S. K. Tanbeer, B. S. Jeong, and Y. K. Lee, "Efficient tree structures for high utility pattern mining in incremental databases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 12, pp. 1708–1721, 2009, doi: 10.1109/TKDE.2009.46.

[8] V. S. Tseng, B. E. Shie, C. W. Wu, and P. S. Yu, "Efficient algorithms for mining high utility itemsets from transactional databases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 8, pp. 1772–1786, 2013, doi: 10.1109/TKDE.2012.59.

[9] M. M. Aziz, A. A. Bakar, and M. R. Yaakub, "CoreNLP dependency parsing and pattern identification for enhanced opinion mining in aspect-based sentiment analysis," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 4, 2024, doi: 10.1016/j.jksuci.2024.102035.

[10] K. K. Yusuf, E. Ogbuju, T. Abiodun, and F. Oladipo, "A technical review of the state-of-the-art methods in aspect-based sentiment analysis," *Journal of Computing Theories and Applications*, vol. 1, no. 3, pp. 287–298, 2024, doi: 10.62411/jcta.9999.

[11] H. H. Do, P. Prasad, A. Maag, and A. Alsadoon, "Deep learning for aspect-based sentiment analysis: a comparative review," *Expert Systems with Applications*, vol. 118, pp. 272–299, Mar. 2019, doi: 10.1016/j.eswa.2018.10.003.

[12] A. Qazi, R. G. Raj, G. Hardaker, and C. Standing, "A systematic literature review on opinion types and sentiment analysis techniques," *Internet Research*, vol. 27, no. 3, pp. 608–630, Jun. 2017, doi: 10.1108/IntR-04-2016-0086.

[13] P. Hajek, L. Hikkerova, and J. M. Sahut, "Fake review detection in e-commerce platforms using aspect-based sentiment analysis," *Journal of Business Research*, vol. 167, 2023, doi: 10.1016/j.jbusres.2023.114143.

[14] Z. Lin and B. Li, "Aspect-based sentiment analysis based on local context focus mechanism and talking-head attention," in *2022 4th International Conference on Frontiers Technology of Information and Computer, ICFTIC 2022*, 2022, pp. 112–118, doi: 10.1109/ICFTIC57696.2022.10075321.

[15] W. An, F. Tian, P. Chen, and Q. Zheng, "Aspect-based sentiment analysis with heterogeneous graph neural network," *IEEE Transactions on Computational Social Systems*, vol. 10, no. 1, pp. 403–412, 2023, doi: 10.1109/TCSS.2022.3148866.

[16] H. Xu, S. Liu, W. Wang, and L. Deng, "RAG-TCGCN: aspect sentiment analysis based on residual attention gating and three-channel graph convolutional networks," *Applied Sciences*, vol. 12, no. 23, 2022, doi: 10.3390/app122312108.

[17] Y. Wang, C. Liu, J. Xie, S. Yang, Y. Jia, and H. Zan, "Aspect-based sentiment analysis with dependency relation graph convolutional network," in *2022 International Conference on Asian Language Processing, IALP 2022*, 2022, pp. 63–68. doi: 10.1109/IALP57159.2022.9961321.

[18] A. Kumar and G. Garg, "Systematic literature review on context-based sentiment analysis in social multimedia," *Multimedia Tools and Applications*, vol. 79, no. 21–22, pp. 15349–15380, 2020, doi: 10.1007/s11042-019-7346-5.

[19] H. Han, X. Bai, and P. Li, "Augmented sentiment representation by learning context information," *Neural Computing and Applications*, vol. 31, no. 12, pp. 8475–8482, 2019, doi: 10.1007/s00521-018-3698-4.

[20] S. S. Hari, S. Porkodi, R. Saranya, and N. Vijayakumar, "Intelligent model to improve the efficacy of healthcare content marketing by auto-tagging and exploring the veracity of content using opinion mining," *International Journal of Electronic Marketing and Retailing*, vol. 15, no. 2, pp. 240–260, 2024, doi: 10.1504/IJEMR.2024.136978.

[21] R. Alfrjani, T. Osman, and G. Cosma, "A hybrid semantic knowledgebase-machine learning approach for opinion mining," *Data & Knowledge Engineering*, vol. 121, pp. 88–108, 2019, doi: 10.1016/j.datak.2019.05.002.

[22] Y. Zhuang, Z. Liu, T.-T. Liu, C.-C. Hung, and Y.-J. Chai, "Implicit sentiment analysis based on multi-feature neural network model," *Soft Computing*, vol. 26, no. 2, pp. 635–644, Jan. 2022, doi: 10.1007/s00500-021-06486-7.

[23] Y. Zhu, W. Zheng, and J. Zheng, "A context enhanced attention network for aspect-based sentiment classification," in *2020 International Conference on Big Data and Artificial Intelligence and Software Engineering, ICBASE 2020*, 2020, pp. 108–112, doi: 10.1109/ICBASE51474.2020.00031.

[24] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases," in *VLDB '94: Proceedings of the 20th International Conference on Very Large Data Bases*, 1994, pp. 487–499, doi: 10.5555/645920.672836.

[25] M. V. Joseph, "A Bi-LSTM and GRU hybrid neural network with BERT feature extraction for amazon textual review analysis," *International Journal of Engineering Trends and Technology*, vol. 70, no. 5, pp. 131–144, 2022, doi: 10.14445/22315381/IJETT-V70I5P216.

[26] A. Hidri, S. A. Alsaif, M. Alahmari, E. AlShehri, and M. S. Hidri, "Opinion mining and analysis using hybrid deep neural networks," *Technologies*, vol. 13, no. 5, 2025, doi: 10.3390/technologies13050175.

[27] K. T. Shahriar and I. H. Sarker, "Exploring a hybrid deep learning framework to automatically discover topic and sentiment in COVID-19 Tweets," *Computing & AI Connect*, vol. 2, no. 1, 2025, doi: 10.69709/caic.2025.199112.

[28] S. Yao and J. Bi, "Sentiment analysis based on the BERT and bidirectional GRU Model," in *2024 IEEE 7th International Conference on Information Systems and Computer Aided Education, ICISCAE 2024*, 2024, pp. 290–295, doi: 10.1109/ICISCAE62304.2024.10761342.

[29] A. S. Talaat, "Sentiment analysis classification system using hybrid BERT models," *Journal of Big Data*, vol. 10, no. 1, Jun. 2023, doi: 10.1186/s40537-023-00781-w.

[30] B. Gupta, P. Prakasam, and T. Velmurugan, "Integrated BERT embeddings, BiLSTM-BiGRU and 1-D CNN model for binary sentiment classification analysis of movie reviews," *Multimedia Tools and Applications*, vol. 81, no. 23, pp. 33067–33086, 2022, doi: 10.1007/s11042-022-13155-w.

[31] J. Chen *et al.*, "Incremental high average-utility itemset mining: survey and challenges," *Scientific Reports*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-60279-0.

[32] K. Singh, A. Kumar, S. S. Singh, H. K. Shakya, and B. Biswas, "EHNL: an efficient algorithm for mining high utility itemsets with negative utility value and length constraints," *Information Sciences*, vol. 484, pp. 44–70, 2019, doi: 10.1016/j.ins.2019.01.056.

[33] K. K. Sethi and D. Ramesh, "High average-utility itemset mining with multiple minimum utility threshold: a generalized approach," *Engineering Applications of Artificial Intelligence*, vol. 96, 2020, doi: 10.1016/j.engappai.2020.103933.

[34] S. Zida, P. F.-Viger, J. C. W. Lin, C. W. Wu, and V. S. Tseng, "EFIM: a fast and memory efficient algorithm for high-utility itemset mining," *Knowledge and Information Systems*, vol. 51, no. 2, pp. 595–625, 2017, doi: 10.1007/s10115-016-0986-0.

[35] C. N. Dang, M. N. M.-García, and F. D. L. Prieta, "Hybrid deep learning models for sentiment analysis," *Complexity*, vol. 2021, pp. 1–16, 2021, doi: 10.1155/2021/9986920.

[36] R. Obiedat *et al.*, "Sentiment analysis of customers' reviews using a hybrid evolutionary SVM-based approach in an imbalanced data distribution," *IEEE Access*, vol. 10, pp. 22260–22273, 2022, doi: 10.1109/ACCESS.2022.3149482.

[37] S. Fahmi, L. Purnamawati, G. F. Shidik, M. Muljono, and A. Z. Fanani, "Sentiment analysis of student review in learning management system based on sastrawi stemmer and SVM-PSO," in *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, Semarang, Indonesia, 2020, pp. 643–648, doi: 10.1109/iSemantic50169.2020.9234291.

[38] V. Gupta and P. Rattan, "Improving Twitter sentiment analysis efficiency with SVM-PSO classification and EFWS heuristic," *Procedia Computer Science*, vol. 230, pp. 698–715, 2023, doi: 10.1016/j.procs.2023.12.125.

[39] K. Machová, M. Mikula, X. Gao, and M. Mach, "Lexicon-based sentiment analysis using particle swarm optimization," *Electronics*, vol. 9, no. 8, pp. 1–22, 2020, doi: 10.3390/electronics9081317.

[40] M. Ledmi, A. Ledmi, M. E. H. Souidi, A. H.-Cherif, T. M. Maarouk, and C. K. M. H.-Cherif, "High-utility itemsets mining integrating an improved crow search algorithm and particle search optimization," *Soft Computing*, vol. 28, no. 13–14, pp. 8471–8496, 2024, doi: 10.1007/s00500-024-09758-0.

[41] M. Al-Smadi, M. M. Hammad, S. A. Al-Zboon, S. Al-Tawalbeh, and E. Cambria, "Gated recurrent unit with multilingual universal sentence encoder for Arabic aspect-based sentiment analysis," *Knowledge-Based Systems*, vol. 261, 2023, doi: 10.1016/j.knosys.2021.107540.

[42] H. Wu, C. Huang, and S. Deng, "Improving aspect-based sentiment analysis with knowledge-aware dependency graph network," *Information Fusion*, vol. 92, pp. 289–299, Apr. 2023, doi: 10.1016/j.inffus.2022.12.004.

[43] X. Sun, K. Zhang, Q. Liu, M. Bao, and Y. Chen, "Harnessing domain insights: a prompt knowledge tuning method for aspect-based sentiment analysis," *Knowledge-Based Systems*, vol. 298, 2024, doi: 10.1016/j.knosys.2024.111975.

[44] Z. Jian, J. Li, Q. Wu, and J. Yao, "Retrieval contrastive learning for aspect-level sentiment classification," *Information Processing and Management*, vol. 61, no. 1, 2024, doi: 10.1016/j.ipm.2023.103539.

[45] G. K. P, A. A. V. S, J. P. V, A. Paul, and A. Nayyar, "A context-sensitive multi-tier deep learning framework for multimodal sentiment analysis," *Multimedia Tools and Applications*, vol. 83, no. 18, pp. 54249–54278, 2024, doi: 10.1007/s11042-023-17601-1.

[46] H. A. Kheruwala, M. S. Ahmad, J. P. Verma, S. Tanwar, and P. K. Singh, "Context-enriched machine learning-based approach for sentiment analysis," in *Recent Innovations in Computing*, Singapore: Springer, 2022, pp. 67–84, doi: 10.1007/978-981-16-8892-8_6.

[47] M. Fouad, W. Hussein, S. Rady, P. S. Yu, and T. Gharib, "A hybrid recommender system combining collaborative filtering with utility mining," *International Journal of Intelligent Computing and Information Sciences*, vol. 22, no. 4, pp. 13–24, 2022, doi: 10.21608/ijicis.2022.145103.1192.

[48] R. Alatrash and R. Priyadarshini, "Fine-grained sentiment-enhanced collaborative filtering-based hybrid recommender system," *Journal of Web Engineering*, vol. 22, no. 7, pp. 983–1035, 2023, doi: 10.13052/jwe1540-9589.2273.

[49] I. Karabila, N. Darraz, A. El-Ansari, N. Alami, and M. El Mallahi, "Enhancing collaborative filtering-based recommender system using sentiment analysis," *Future Internet*, vol. 15, no. 7, 2023, doi: 10.3390/fi15070235.

[50] Y. Zhang, K. Meng, W. Kong, Z. Y. Dong, and F. Qian, "Bayesian hybrid collaborative filtering-based residential electricity plan recommender system," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 8, pp. 4731–4741, 2019, doi: 10.1109/TII.2019.2917318.

[51] J. C.-W. Lin, P. F.-Viger, and W. Gan, "FHN: an efficient algorithm for mining high-utility itemsets with negative unit profits," *Knowledge-Based Systems*, vol. 111, pp. 283–298, Nov. 2016, doi: 10.1016/j.knosys.2016.08.022.

[52] D. Kim *et al.*, "Efficient mining of incremental high utility patterns with negative unit profits over all the accumulated stream data," *Knowledge-Based Systems*, vol. 325, 2025, doi: 10.1016/j.knosys.2025.113956.

[53] M. Ashraf, T. Abdelkader, S. Rady, and T. F. Gharib, "TKN: an efficient approach for discovering top-k high utility itemsets with positive or negative profits," *Information Sciences*, vol. 587, pp. 654–678, 2022, doi: 10.1016/j.ins.2021.12.024.

[54] E. Caron, "Open-data-set," *GitHub repository*, 2016. Accessed: Dec. 29, 2024. [Online]. Available: https://github.com/BestBuyAPIs/open-data-set

[55] H. Hashemi and H. Hashemi, "Sentiment analysis by hierarchical deep neural networks for audience opinion mining," *Asian Journal of Research in Computer Science*, vol. 17, no. 6, pp. 202–217, 2024, doi: 10.9734/ajrcos/2024/v17i6468.

[56] A. K. Shahade, K. H. Walse, V. M. Thakare, and M. Atique, "Multi-lingual opinion mining for social media discourses: an approach using deep learning based hybrid fine-tuned smith algorithm with adam optimizer," *International Journal of Information Management Data Insights*, vol. 3, no. 2, 2023, doi: 10.1016/j.jjimei.2023.100182.

[57] K. Saraswathi, N. T. Renukadevi, and S. Nandhinidevi, "Performance evaluation of deep learning and machine learning techniques for opinion mining," in *Intelligent Decision Making Through Bio-Inspired Optimization*, Hershey: IGI Global, 2024, pp. 1–17, doi: 10.4018/979-8-3693-2073-0.ch001.

[58] S. Bhowmik, S. Sultana, A. A. Sajid, S. Reno, and A. Manjrekar, "Robust multi-domain descriptive text classification leveraging conventional and hybrid deep learning models," *International Journal of Information Technology*, vol. 16, no. 5, pp. 3219–3231, 2024, doi: 10.1007/s41870-023-01559-2.

[59] F. Cauteruccio and G. Terracina, "Extended high-utility pattern mining: an answer set programming-based framework and applications," *Theory and Practice of Logic Programming*, vol. 24, no. 2, pp. 313–343, 2024, doi: 10.1017/S1471068423000066.

[60] Y. Li, Y. Dong, Y. Wang, and N. Zhang, "Product design opportunity identification through mining the critical minority of customer online reviews," *Electronic Commerce Research*, vol. 25, no. 1, pp. 211–239, 2025, doi: 10.1007/s10660-023-09683-8.

BIOGRAPHIES OF AUTHORS

Garima Srivastava is currently working as an assistant professor in the Department of Computer Science and Engineering, at Amity University, Lucknow. She holds an M.Tech. degree in Computer Science and Engineering from Guru Gobind Singh Indraprastha University, Delhi. She has around 17 years of teaching and research experience. Her research interest includes machine learning and NLP. She has published several research papers in national and international conferences and journals. She can be contacted at email: gsrivastava1@lko.amity.edu.

Dr. Vaishali Singh is currently working as an assistant professor in Department of Computer Science, Maharishi University of Information Technology, Lucknow. She has received her Ph.D. degree in the year 2017 from B.B. Ambedkar University, Lucknow, India. Her research interest includes information retrieval and question answering systems. She has published many research papers in reputed international conferences and journals. She can be contacted at email: singh.vaishali05@gmail.com.

Dr. Sachin Kumar a Ph.D. in Electronics from Dr. R M L University, M.Tech. and B.E in Electronics and Communication from Pune University, Presently with Amity University, Lucknow. He has almost 20 years of academic and research experience. His research interest includes applications of bio-medical signal processing, machine learning, artificial intelligence and applications of wireless sensor networks. He has published more than 80 research papers, book chapters, and articles in a number of reputed journals indexed in Scopus and SCI. He is an active member of editorial board/reviewer in various reputed journals. He can be contacted at email: skumar3@lko.amity.edu.