Informing Science:
the International Journal of
an Emerging Transdiscipline

An Official Publication
of the Informing Science Institute
InformingScience.org

Inform.nu

Volume 28, 2025

HEALTHCARE BICLUSTERING OF PREDICTIVE GENE EXPRESSION USING LSTM BASED SUPPORT VECTOR MACHINE

Thulasi Bikku*	Department of Computer Science and Engineering, Amrita School of Computing Amaravati, Amrita Vishwa Vidyapeetham, Andhra Pradesh, India	b_thulasi@av.amrita.edu
Joy Elvine Martis	Department of Humanities, NMAM Institute of Technology (NMAMIT), NITTE (Deemed to be University), Nitte, Karnataka, India	joymartis@nitte.edu.in
Sunil Kumar M	Department of Computer Science and Engineering, School of Computing, Mohan Babu University (erstwhile Sree Vidyanikethan Engineering College), Tirupathi, Andhra Pradesh, India	sunilmalchi1@gmail.com
Sudha S	Department of ECE, Sri Ranganathar Institute of Engineering and Technology, Coimbatore, Tamilnadu, India	sudhanithi@gmail.com
Iyappan P	School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology, Vellore, Tamilnadu, India	iyappan.perumal@vit.ac.in
Natarajan C	Department of CSE, P.S.R. Engineering College, Sivakasi, Tamilnadu, India	natarajan@psr.edu.in

* Corresponding author

Accepting Editor Eli Cohen | Received: October 19, 2024 | Revised: January 22, 2025 | Accepted: January 23, 2025.
Cite as: Bikku, T., Martis, J. E., Sunil Kumar, M., Sudha, S., Iyappan, P., & Natarajan, C. (2025). Healthcare biclustering of predictive gene expression using LSTM based support vector machine. *Informing Science: The International Journal of an Emerging Transdiscipline, 28,* Article 12. https://doi.org/10.28945/5446

ABSTRACT

Aim/Purpose	The major goal of this work is to establish prediction patterns that can influence better diagnosis and treatment strategies using unidentified interactions between genes.
Background	Driven by the rapid advances in genomics, knowledge of the factors causing disease depends more on deciphering the deep linkages existent in the data of gene expression. Common approaches typically fail to grasp temporal links when dealing with always-changing living biological systems. This work overcomes this restriction by leveraging the sequential learning abilities of LSTM together with the improved pattern recognition capacity of SVM.
Methodology	Our method uses a hybrid model combining LSTM and SVM to forecast gene expression. Working together, the LSTM and SVM components find relevant features in the gene expression data, clarifying trends in the data. Furthermore, the LSTM component oversees data temporal dependencies. Regarding accuracy and interpretability, this extra method helps to improve prediction models used in the healthcare industry.
Contribution	There are many ways to get a key insight from data on gene expression. The LSTM and SVM for biclustering gene expression data offer much for healthcare informatics.
Findings	The proposed LSTM-based SVM is used to evaluate numerous current methods of evaluating performance metrics. Using these opens several opportunities for the development of customized medicine and the customization of therapies in line with personal genetic profiles.
Recommendations for Researchers	Examining the LSTM-SVM hybrid model that has been proposed using a variety of healthcare-related datasets
Future Research	This work can be enhanced using several deep-learning algorithms to achieve better accuracy and performance.
Keywords	healthcare, biclustering, gene expression, SVM, LSTM, predictive modeling

INTRODUCTION

There are various problems in the healthcare sector nowadays. Prediction and analysis of disease over gene expression of data on a molecular ailment are available (Nicholls & Wallace, 2021). Conventional approaches typically miss the intricate temporal dynamics in the large-scale generated datasets resulting from contemporary genomics discoveries (Maâtouk et al., 2021).

Several new approaches are available to acquire significant knowledge from data on gene expression (Bikku et al., 2023). Common methods in the field of healthcare predictive modeling cannot always be successful since they cannot find small trends and linkages (Bikku et al., 2021).

This effort aims to enhance the already described biclustering data pertinent to a gene expression strategy (Xie et al., 2020). Primarily, enabling models to produce accurate predictions would help to capture temporal linkages (Lazareva et al., 2021) in a way that enhances the diagnostic and treatment outcomes in the healthcare industry.

Inspired by the rapid advances in genomics, knowing the elements causing disease relies more and more on how one understands the deep relationships discovered in gene expression data. Common approaches typically fail to grasp temporal links when dealing with always-changing living biological

systems. This work overcomes this restriction by leveraging the better pattern recognition capacity of support vector machines (SVMs) as well as the sequential learning abilities of long short-term memory (LSTM) (Yelugam et al., 2023).

The authors propose a technique for gene expression analysis in this study that would have substantial consequences for the field of healthcare informatics. It presents a hybrid paradigm with an eye on helping to reach and identify potential treatment targets faster.

The contributions of this research work can be summarized as follows:

1. The study introduces a novel hybrid model combining LSTM and SVM, leveraging their strengths to analyze gene expression data effectively. This model improves the ability to capture temporal dependencies and identify critical features in the data.
2. By integrating sequential learning and pattern recognition capabilities, the proposed method demonstrates superior performance in predictive modeling, surpassing traditional approaches like HMM, SVM, and RNN.
3. The findings offer significant implications for healthcare informatics, particularly in enabling personalized medicine by identifying potential treatment targets and optimizing diagnostic and therapeutic strategies through improved gene expression analysis.

RELATED WORKS

Gene expression analysis has been extensively studied and deployed to investigate several strategies for separating the intricate relationships in biological data (José-García et al., 2023). From more conventional statistical techniques to cutting-edge machine learning approaches, the body of present research covers a large spectrum of methods (Mehta et al., 2021).

LSTM shows potential in many different disciplines. Still mostly unknown, nevertheless, is their behavior with data on gene expression (Ali et al., 2023; Liu et al., 2023; Manikandan et al., 2022; Paul et al., 2022; Shesayar et al., 2023; Siswantining et al., 2021; Sivakumar & Shankar, 2022).

Pattern recognition findings reveal that SVMs run very effectively. SVMs have shown promise in developing potential biomarkers and classifying cancer subtypes by means of gene expression patterns. Although they are quite good at capturing complex decision boundaries, SVMs have often not been applied for temporal data such as gene expression time series (Liu et al., 2020).

Combining LSTM with SVM is a unique and inspiring approach to get past the limitations of the current applied techniques. Integration of these models has gained less interest in the field of gene expression analysis than it could have. Combining the sequential learning powers of LSTM with the capacity of SVM to improve and discover fundamental features helps one to build a prediction model that is both more accurate and more clearly interpretable. Bikku et al. (2024) investigated hybrid models for time-series genomic data, emphasizing the synergy between sequential learning and feature refinement techniques. Combining the sequential learning powers of LSTM with the capacity of SVM to improve and discover fundamental features helps one to build a prediction model that is both more accurate and more clearly interpretable. This combination captures temporal links rather successfully, addressing the limitations of standalone models.

When one compares the proposed LSTM-SVM approaches with other traditional approaches, including HMM, SVM, and RNN, the LSTM-SVM models demonstrate better accuracy and performance. Higher accuracy and memory of the LSTM-SVM approach contribute to these advantages. Hence, utilizing LSTM with SVM approaches delivers a fresh perspective on gene expression datasets, offering a new option for the healthcare sector.

The higher silhouette score indicates how relevant the model is in revealing important patterns in gene expression data. This score supports the model's capacity to generate coherent and original gene

groups. For example, recent studies by Batchu et al. (2024) and Srinivasu et al. (2024) further corroborate the importance of hybrid approaches in improving clustering performance and interpretability in biomedical datasets. These findings reinforce the potential of LSTM-SVM integration to uncover meaningful insights in gene expression analysis, bridging the gap and offering robust solutions for healthcare applications.

PROPOSED METHOD

Two components of the given approach are SVM for precise pattern recognition and LSTM networks for temporal modeling. This new approach intends to improve the biclustering of gene expression data by offering a whole in-nature solution catching both main traits and temporal correlations, as in Figure 1.

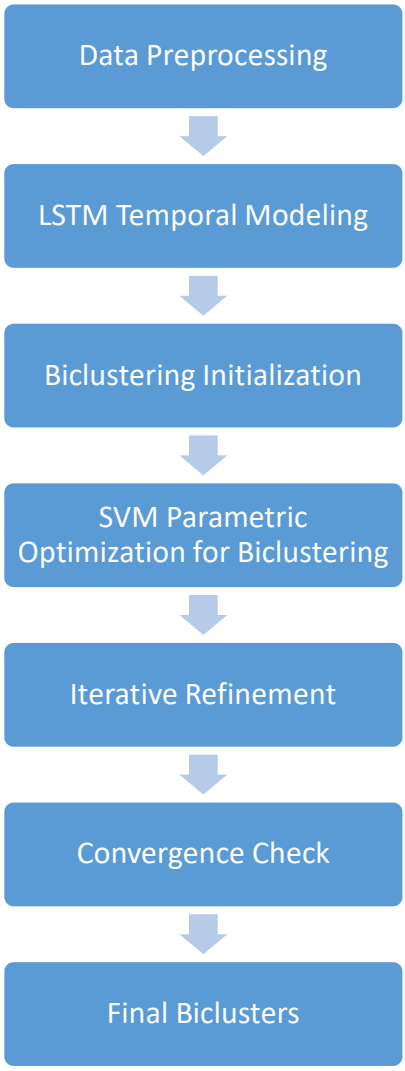

Figure 1. Proposed blustering framework

The study maintains using the learned decision boundaries from the SVM to the gene cluster assignments until convergence or a stopping condition is attained. The genes more fit for the rest of the cluster should be conserved even while those that do not fit should be reclassified. The SVM settings must be changed to match the amended gene cluster assignments. The change done to the enlarged

cluster structure increases the SVM's discriminative capacity at this level. Examine the gene cluster assignments to verify whether SVM parameters have been altered to guarantee convergence or whether they are steady. Below is the recommended LSTM-SVM biclustering technique.

Algorithm: LSTM-SVM Biclustering

Input:

Gene expression dataset with temporal information.

Output:

Biclusters of genes with similar expression patterns.

- Normalize the gene expression data to handle variations in magnitude.
- Arrange the data in a temporal sequence for input to the LSTM.
- Configure the LSTM architecture:
  - Define the number of LSTM units and input dimensions.
- Specify the activation function, loss function, and optimization algorithm.
- Train the LSTM on the preprocessed gene expression data to capture temporal dependencies.
- Obtain LSTM-generated embeddings representing temporal features of genes.
- Utilize the LSTM embeddings to initialize an initial set of gene clusters.
- Apply a clustering algorithm to group genes based on their temporal expression patterns.
- For each initialized gene cluster:
  - Extract the gene expression profiles within the cluster.
  - Configure SVM parameters:
    - Define kernel function
    - Define regularization parameter
  - Train the SVM on the gene expression profiles to optimize and refine the clusters.
  - Obtain refined SVM-generated clusters within the temporal context.
- Iteratively refine the gene clusters by repeating steps 2-4:
  - Reapply LSTM to update temporal features based on refined SVM clusters.
  - Refine clusters with SVM to enhance accuracy and precision.
- Evaluate the final biclusters for coherence and significance.
- Apply metrics such as silhouette score or other validation measures to assess cluster quality.
- Obtain the final set of biclusters.

LSTM Biclustering

With biclustering to probe gene expression data in great depth, including temporal correlations, the LSTM Biclustering approach uses LSTM networks. The first processing of the gene expression dataset employs several preprocessing approaches. These processes will standardize the data and organize it temporally to suit LSTM input.

In terms of input dimensions, activation function, loss function, optimization approach, and LSTM unit count, the LSTM component follows. The LSTM can learn by training on past processed gene expression data, therefore acquiring knowledge and encoding the temporal correlations inherent in the dynamic biological system. Producing embeddings reflecting the temporal properties of the genes allows the LSTM to develop a more thorough awareness of the sequential patterns in the data.

Following the end of the biclustering initiation process, the first set of gene clusters is generated with obtained embeddings from the LSTM. A clustering method helps classify genes according to their expressed patterns. Beginning this process helps one to identify coherent groupings with consistent behavior over extended times.

The approach uses support vector machines (SVM) for parametric optimization inside every gene cluster and follows the LSTM-based initialization. SVM parameters are set using gene expression profiles for every cluster. Among others, these values represent the kernel function and the regularizing factor. Next is training the SVM, which helps to expose significant aspects of refining the clusters. One does this using the reference to profiles of gene expression. This guarantees a more exact and simpler understanding of the biclusters, therefore advancing our knowledge of the kinetics of gene expression.

Using SVM clusters previously tuned, the LSTM Biclustering technique iteratively updates temporal properties. Then, the method once again enhances the clusters using SVM. Through iterative development, the model can react to the subtleties of the data, therefore increasing the accuracy of its predictions by way of suitable capture of intricate temporal correlations and gene cluster optimization.

The approach produces highly defined biclusters that are then used to illustrate groups of genes showing comparable over-time expression trends. The post-processing stage is marked by analyzing and interpreting these biclusters in the scope of their biological importance. This is done for critical information on regulatory systems, gene interactions, or biomarkers. Within the realm of healthcare informatics, the all-encompassing LSTM Biclustering approach increases our knowledge of gene expression dynamics.

Driven by a set of equations defining its behavior, the LSTM cell is the fundamental building unit of the network. Assume W and U are the LSTM parameters; let the input vector be x_t, the output vector be ht, and the cell state Ct.

$$i_t=\sigma(Wi\cdot x_t+U_i h_{t-1}+bi)$$

$$ft=\sigma(Wf\cdot x_t+U_f h_{t-1}+bf)$$

$$C_t=f_t\cdot C_{t-1}+i_t\cdot C'_t$$

$$C'_t=tanh(W_C\cdot x_t+U_C h_{t-1}+b_C)$$

$$h_t=o_t\cdot tanh(C_t)$$

$$o_t=\sigma(Wo\cdot,x_t+U_o h_{t-1}]+b_o)$$

where

i_t - input gate output
f_t - forget gate output
C_t - cell state
C'_t - candidate cell state
h_t - hidden state
o_t - output gate output
$tanh$ - hyperbolic tangent activation function
σ - sigmoid activation function
$[h_{t-1}, x_t]$ - represents the concatenation of the previous hidden state and the current input

ALGORITHM: LSTM BICLUSTERING

Input:

Gene expression dataset with temporal information

Output:

Biclusters of genes with similar expression patterns

- Normalize the gene expression data to handle variations in magnitude.
- Structure the data into a temporal sequence suitable for LSTM input.
- Configure the LSTM architecture:

    - Define the number of LSTM units
    - Define input dimensions
    - Define activation functions
    - Define loss function
- Train the LSTM on the preprocessed gene expression data to capture temporal dependencies.
- Obtain LSTM-generated embeddings representing temporal features of genes.
- Use the LSTM-generated embeddings to initialize an initial set of gene clusters.
- Apply a clustering algorithm (e.g., K-means) to group genes based on their temporal expression patterns.
- This forms the foundation for identifying clusters that exhibit similar behaviors over time.
- Iterate through the following steps until convergence:
    - Reapply LSTM to update temporal features based on the current gene clusters.
    - Utilize the updated embeddings for the next biclustering iteration.
    - Apply a biclustering algorithm using the LSTM-generated embeddings.
    - Refine gene clusters based on the temporal information encoded by LSTM.
- Check for convergence criteria for minimal changes in LSTM embeddings.
- If convergence is achieved, proceed to the next step; otherwise, return to the iterative refinement.
- Obtain the final set of biclusters representing groups of genes with similar expression patterns over time.

SVM Parametric Optimization for Biclustering Process

SVM parametric optimization for biclustering is a technique whereby gene clusters within a framework built on biclustering are improved and optimized. A cluster begins with an already-existing collection of gene clusters produced in a previous stage (such as hierarchical clustering, K-means, and LSTM biclustering). Cluster optimization begins right here. Discover and then extract the profiles of gene expression found within every initialized gene cluster. These pertinent profiles indicate how genes might evolve in response to particular stimuli or with time. One must choose a kernel function – linear, Poisson, RBF, etc. – to design SVM with suitable parameters customized to every gene cluster. Changing the regularisation value (C) will help control the trade-off between classifying the training points and generating a smooth decision boundary. Some of the other hyperparameters should be changed to match the clustering operation criteria and the gene expression data features. Using the above-described parameters enables one to teach an SVM for every gene cluster. Eventually, the SVM is trained to discriminate between groups of genes with similar expression patterns, thus refining the definition of gene subgroups.

The study maintains using the learned decision boundaries from the SVM to the gene cluster assignments until convergence or a stopping condition is attained. The genes more fit for the rest of the cluster should be conserved even while those that do not fit should be reclassified. The SVM settings must be changed to match the amended gene cluster assignments. The change done to the enlarged cluster structure increases the SVM's discriminative capacity at this level. Check whether the gene cluster assignments are stable or whether SVM parameters have been adjusted to guarantee convergence. Assuming convergence has been reached, go beyond. Should this prove not to be the case, return to the optimizing iterations. Using SVM parametric optimization, retrieve the collection of ideally optimized gene clusters. More so than others, the expression patterns of the clusters used to represent these gene groupings satisfy the specified criteria. The genes themselves speak for these categories. We improve the produced gene clusters by iteratively optimizing using SVM under a biclustering framework. This leads to a better knowledge of the functional biological systems.

BICLUSTERING GENE EXPRESSION DATA

Biclustering gene expression data aims to find groupings of genes whose expression patterns (Table 1) coincide under specified conditions or at specified time intervals. By means of their homogeneity, this method enables one to identify genes with like activities that might be connected to common biological processes.

Table 1. Sample dataset

Gene/sample	Sample1	Sample2	Sample3	...	Sample10
Gene1	2.5	3	1.8	...	2.2
Gene2	1	1.2	0.9	...	1.5
...	...	...	...	...	...
Gene20	4	3.8	4.2	...	3.9

Imagine a dataset for gene expression with rows of genes and columns of samples whereby every item indicates the degree of expression of the gene in a given sample. Our goal is to identify groupings of genes (Table 2) whose expression patterns match across numerous samples, so we will cluster the data using two alternative methods. Imagine a small gene expression matrix including ten samples and twenty genes.

Table 2. Grouping

Gene/sample	Sample2	Sample3	Sample7
Gene1	3	1.8	2.5
Gene5	1.2	1	1.1
Gene12	4.2	4	3.8

Using a biclustering approach, one can identify groups of genes whose expression is stable across samples. In one such experiment, for instance, biclusters can show rising genes. Samples 2, 3, and 7 consistently indicate multiple genes – Gene 1, Gene 5, and Gene 12 – that the biclustering method detects in expression patterns.

Based on the bicluster observed, Gene 1, Gene 5, and Gene 12 could be co-regulated or associated with a shared biological process unique to the conditions shown by Sample 2, Sample 3, and Sample 7. Repeated biclustering allows one to identify additional gene subsets exhibiting a range of expression patterns over a spectrum of sample subsets, therefore enabling a more in-depth knowledge of the dynamics of gene expression inside the dataset. Finding functionally connected genes allows one to grasp the biological processes under research. Bicoloring the data on gene expression is one smart way to do this.

EXPERIMENTAL VERIFICATION

Using a biclustering approach, one can identify groups of genes whose expression is stable across samples. In one such experiment, for instance, biclusters can show rising genes. Samples 2, 3, and 7 consistently indicate multiple genes – Gene 1, Gene 5, and Gene 12 – that the biclustering method detects in expression patterns. Based on the bicluster observed, Gene 1, Gene 5, and Gene 12 could be co-regulated or associated with a shared biological process unique to the conditions shown by Sample 2, Sample 3, and Sample 7. Repeated biclustering allows one to identify additional gene subsets exhibiting a range of expression patterns over a spectrum of sample subsets, thus enabling a more in-depth knowledge of the dynamics of gene expression inside the dataset. Finding functionally connected genes allows one to grasp the biological processes under research. Bicoloring the data on gene expression is one smart way to do this.

Table 1. Experimental settings

Experimental setup	Values
Number of LSTM Units	64
Input Dimensions	30
Activation Function	Tanh
Loss Function	Mean Squared Error
Optimization Algorithm	Adam
Learning Rate	0.001
Number of Epochs	50
Kernel Function	Radial Basis Function (RBF)
Regularization Parameter (C)	1
SVM Iterations	20

The LSTM-SVM strategy, with improvements throughout a wide spectrum of performance criteria (Figures 2-7), beats other methods, including HMM, RNN, and SVM.

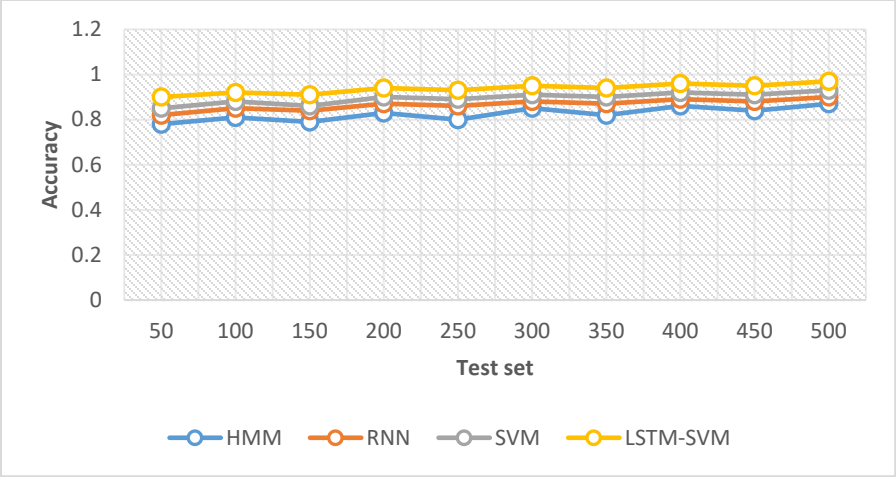

Figure 2. Biclustering accuracy

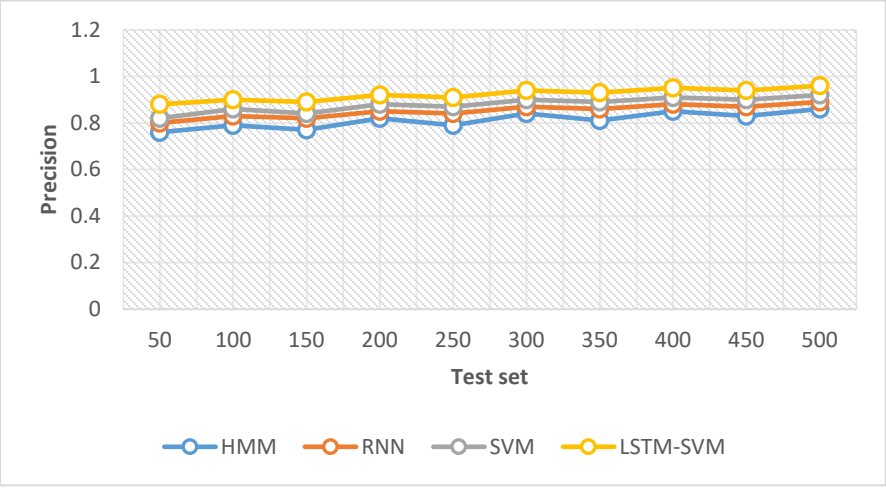

Figure 3. Precision

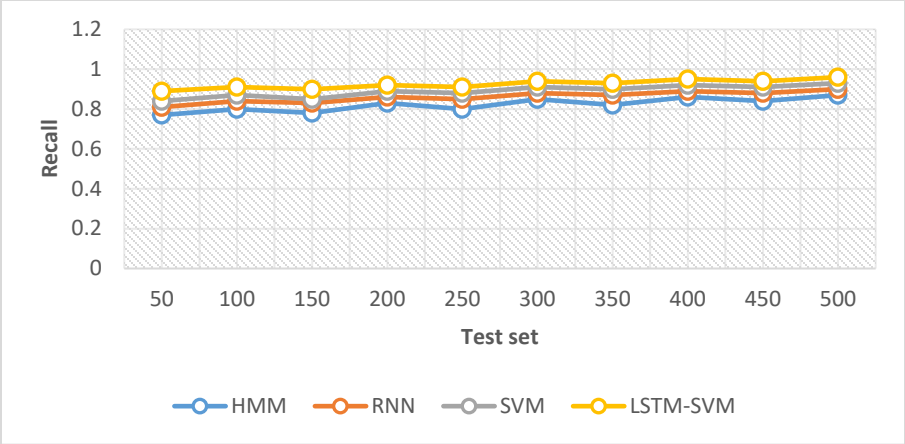

Figure 4. Recall

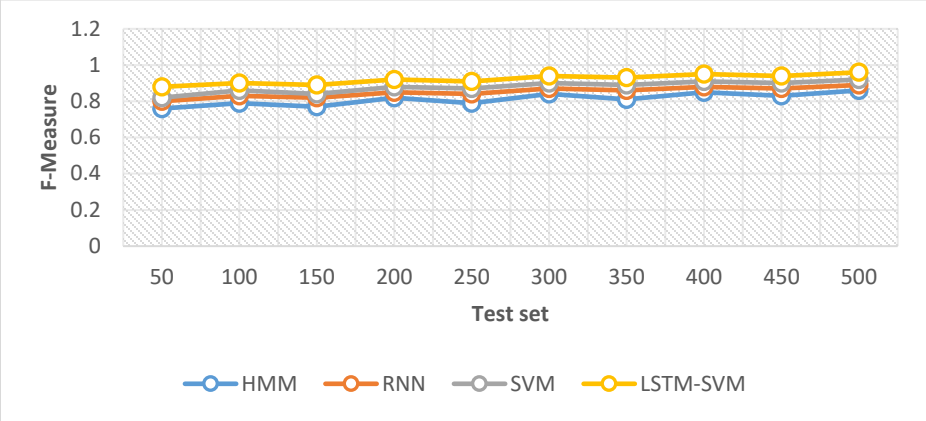

Figure 5. F-measure

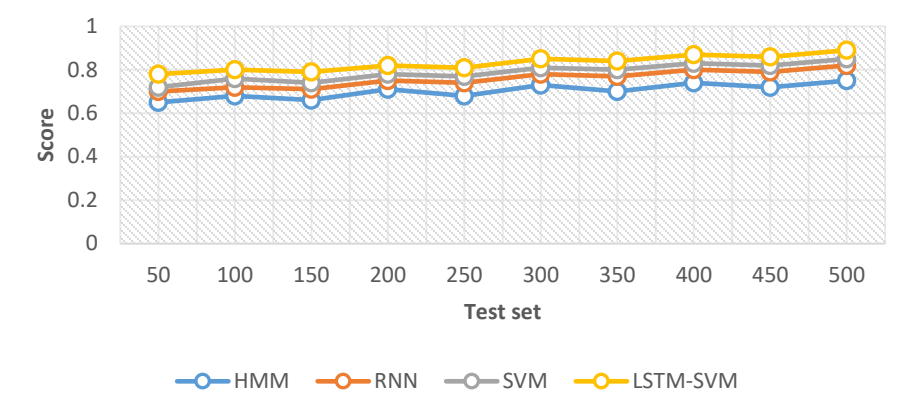

Figure 6. Silhouette score

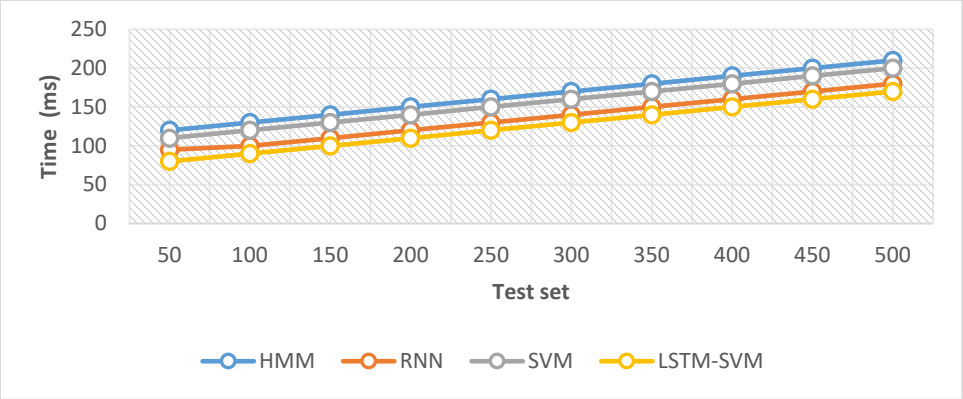

Figure 7. Computational time

Combining LSTM with SVM is a unique and inspiring approach to get past the limitations of the currently employed techniques. The integration of several models has gotten less attention in the field of gene expression analysis than might have been expected. Combining the sequential learning powers of LSTM with the capacity of SVM to improve and discover fundamental features helps one to build a prediction model that is both more accurate and more clearly interpretable. Apart from other standard techniques, including HMM, RNN, and SVM, the proposed method LSTM-SVM provides 10% increased accuracy in addition to a 7% recall rate improvement. This was evident from the roughly 8% rise in the F-measure, a test of balance between accuracy and recall. Comparatively to other techniques, including HMM, RNN, and SVM, it also shows a 9% increase in Silhouette Score.

Calculating based on accuracy, precision, recall, F-measure, and silhouette score, LSTM-SVM proves to be better than earlier techniques. Moreover, this one is more practical and efficient than the others. In the field of healthcare informatics, the noted increases indicate the possibilities of this method to execute efficient and consistent biclustering of gene expression data.

CONCLUSION

When one compares the proposed LSTM-SVM approaches with other traditional approaches, including HMM, SVM, and RNN, they have better accuracy and performance than others. Higher accuracy and memory of the LSTM-SVM approach are aspects that lead to these features; hence, utilizing LSTM with SVM approaches delivers a fresh view of gene expression in the dataset, which offers a new option for the healthcare sector. The higher silhouette score indicates how relevant the model is in revealing important patterns in gene expression data. The higher silhouette score supports the model's capacity to generate coherent and original gene groups.

To further validate the efficacy of the proposed LSTM-SVM hybrid approach, future research should focus on applying this model to real-world gene expression datasets. Using such data will provide robust evidence of the model's applicability in practical scenarios, such as predicting disease progression, identifying biomarkers, or classifying cancer subtypes. Real-world data would also enable comprehensive benchmarking against other state-of-the-art methods, further establishing the superiority of the LSTM-SVM integration. Additionally, exploring the model's adaptability to various types of omics data, such as proteomics or metabolomics, could broaden its utility and uncover novel insights across different biological domains.

REFERENCES

Ali, A., Ajil, A., Meenakshi Sundaram, A., & Joseph, N. (2023). Detection of gene ontology clusters using bi-clustering algorithms. *SN Computer Science*, *4*, Article 217. https://doi.org/10.1007/s42979-022-01624-w

Batchu, R. K., Bikku, T., Thota, S., Seetha, H., & Ayoade, A. A. (2024). A novel optimization-driven deep learning framework for the detection of DDoS attacks. *Scientific Reports, 14*, Article 28024. https://doi.org/10.1038/s41598-024-77554-9

Bikku, T., Karthik, J., Koteswara Rao, G. R., Satya Sree, K. P. N. V., Srinivas, P. V. V. S., & Prasad, C. (2021). Brain tissue segmentation via deep convolutional neural networks. *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC).* IEEE, 2021. Palladam, pp. 757-763 https://doi.org/10.1109/I-SMAC52330.2021.9640635

Bikku, T., Ramu, J., Sekhar, J. C., Pratap, V. K., & Pujari, J. J. (2023). Optimizing gene expression analysis using clustering algorithms. *Proceedings of Fifth International Conference on Computer and Communication Technologies. IC3T 2023. Lecture Notes in Networks and Systems, vol 898*. Springer. https://doi.org/10.1007/978-981-99-9707-7_15

Bikku, T., Thota, S., & Shanmugasundaram, P. (2024). A novel quantum neural network approach to combating fake reviews. *International Journal of Networked and Distributed Computing, 12,* 195-205. https://oi.org/10.1007/s44227-024-00028-x

José-García, A., Jacques, J., Sobanski, V., & Dhaenens, C. (2023). Metaheuristic biclustering algorithms: From state-of-the-art to future opportunities. *ACM Computing Surveys*, *56*(3), Article 69. https://doi.org/10.1145/3617590

Lazareva, O., Canzar, S., Yuan, K., Baumbach, J., Blumenthal, D. B., Tieri, P., Kacprowski, T., & List, M. (2021). BiCoN: Network-constrained biclustering of patients and omics data. *Bioinformatics*, *37*(16), 2398-2404. https://doi.org/10.1093/bioinformatics/btaa1076

Liu, X., Li, D., Liu, J., Su, Z., & Li, G. (2020). RecBic: A fast and accurate algorithm recognizing trend-preserving biclusters. *Bioinformatics*, *36*(20), 5054-5060. https://doi.org/10.1093/bioinformatics/btaa630

Liu, X., Yu, T., Zhao, X., Long, C., Han, R., Su, Z., & Li, G. (2023). ARBic: An all-round biclustering algorithm for analyzing gene expression data. *NAR Genomics and Bioinformatics*, *5*(1). https://doi.org/10.1093/nargab/lqad009

Maâtouk, O., Ayadi, W., Bouziri, H., & Duval, B. (2021). Evolutionary local search algorithm for the biclustering of gene expression data based on biological knowledge. *Applied Soft Computing*, *104*, 107177. https://doi.org/10.1016/j.asoc.2021.107177

Manikandan, R., Sara, S. B. V. J., Yuvaraj, N., Chaturvedi, A., Priscila, S. S., & Ramkumar, M. (2022, May). Sequential pattern mining on chemical bonding database in the bioinformatics field. *AIP Conference Proceedings*, *2393*, 20050. https://doi.org/10.1063/5.0074405

Mehta, D., Sehgal, S., Choudhury, T., & Sarkar, T. (2021). A comparative analysis of clustering and biclustering algorithms in gene analysis. In J. Singh, S. Kumar, & U. Choudhury (Eds.), *Innovations in cyber physical systems* (pp. 29-39). Springer. https://doi.org/10.1007/978-981-16-4149-7_4

Nicholls, K., & Wallace, C. (2021). Comparison of sparse biclustering algorithms for gene expression datasets. *Briefings in Bioinformatics*, *22*(6), 1-16. https://doi.org/10.1093/bib/bbab140

Paul, L. M. F. V., Chooralil, V. S., & Yuvaraj, N. (2022). Modelling of maximal connectivity pattern in human brain networks. *NeuroQuantology*, *20*(6), 4410.

Shesayar, R., Agarwal, A., Taqui, S. N., Natarajan, Y., Rustagi, S., Bharti, S., & Trehan, A., Sivasubramanian, K., Muruganandham, M., Velmurugan, P., Arumugam, N., Almansour, A. I., Kumar, R. S., & Sivakumar, S. (2023). Nanoscale molecular reactions in microbiological medicines in modern medical applications. *Green Processing and Synthesis*, *12*(1), 20230055. https://doi.org/10.1515/gps-2023-0055

Siswantining, T., Aminanto, A. E., Sarwinda, D., & Swasti, O. (2021). Biclustering analysis using plaid model on gene expression data of colon cancer. *Austrian Journal of Statistics*, *50*(5), 101-114. https://doi.org/10.17713/ajs.v50i5.1195

Sivakumar, N. Y. C., & Shankar, A. (2022). The speech-language processing model for managing the neuro-muscle disorder patients by using deep learning. *NeuroQuantology*, *20*(8), 918-925.

Srinivasu, P. N., Sirisha, U., Sandeep, K., Praveen, S. P., Maguluri, L. P., & Bikku, T. (2024). An interpretable approach with explainable AI for heart stroke prediction. *Diagnostics, 14*(2), 128. https://doi.org/10.3390/diagnostics14020128

Xie, J., Ma, A., Zhang, Y., Liu, B., Cao, S., Wang, C., Xu, J., Zhang, C., & Ma, Q. (2020). QUBIC2: A novel and robust biclustering algorithm for analyses and interpretation of large-scale RNA-Seq data. *Bioinformatics*, *36*(4), 1143-1149. https://doi.org/10.1093/bioinformatics/btz692

Yelugam, R., Brito da Silva, L. E., & Wunsch, D. C., II. (2023). Topological biclustering ARTMAP for identifying within bicluster relationships. *Neural Networks*, *160*, 34-49. https://doi.org/10.1016/j.neunet.2022.12.010

AUTHORS

Dr. Thulasi Bikku is an accomplished academician serving as an Associate Professor at the School of Computing, Amrita Vishwa Vidyapeetham, Amaravati, Andhra Pradesh, India. She earned her PhD in Computer Science and Engineering from JNTUA, Anantapur, in 2018 and completed a Postdoctoral Fellowship at the University of Santiago de Chile in 2023. She specializes in Security, IoT, Unsupervised Machine Learning, Deep Learning, and Natural Language Processing. She has authored over 55 research articles in reputed journals, written academic books bridging theory and practice, and holds applied patents reflecting her innovative contributions.

Dr. Joy Elvine Martis is an Assistant Professor in the Department of Humanities, NMAMIT, Nitte (Deemed to be University), Mangaluru, Karnataka, India. His academic interests include British and American Literature, English Language Teaching (ELT), Phonetics, and Linguistics. Beyond academia, he actively engages with youth by facilitating leadership and communication skills. He is deeply involved in various social activities through volunteering and mentoring, striving to make a meaningful impact on the community. His dedication to education and social causes reflects his commitment to holistic development and societal well-being.

Dr. M. Sunil Kumar, Professor and Controller of Examination, Computer Science and Engineering, Mohan Babu University. With more than 18 years of experience in academia and research, Dr. M. Sunil Kumar specializes in computer science, machine learning, and artificial intelligence. He has a PhD in Computer Science from SV University and has been involved in prestigious projects, like soil nutrition analysis using multispectral satellite data in Japan. His expertise spans research coordination, curriculum development, and guiding PhD candidates. He has published 60+ papers in SCI and Scopus-indexed journals, led consultancy projects such as crop health monitoring using satellite imagery, and received multiple awards, including the Dynamic Professor of the Year in 2022. He has contributed significantly to NBA, NAAC, and R&D initiatives in higher education institutions.

Dr. S. Sudha is an electronics and communication engineering professor at SRI Ranganathar Institute of Engineering and Technology, Coimbatore, Tamilnadu, India. Before this appointment, Dr. S. Sudha was an assistant professor and lecturer at reputed engineering colleges in Tamilnadu, India. Dr. S. Sudha was a senior engineer in the product development and testing division of the leading automobile industry in Tamilnadu, India. Her development includes automobile two-wheeler and four-wheeler clusters in embedded real-time code. Her current research activities are focused on development and solutions related to regular lifestyle and environmental support for future generations.

Dr. Iyappan Perumal obtained his BE in Computer Science and Engineering (2005) from Krishnasamy College of Engineering and Technology, Anna University. He received his MTech in Computer Science and Engineering (2008) from SMVEC and was awarded a gold medal from Pondicherry University. He completed a PhD in Computer Science and Engineering. He has more than 15 years of experience in teaching and research. He is a life member of ISTE and IAENG. He has published over 25 research articles indexed in SCOPUS/SCI/UGC care journals and reputed conferences. He has also authored one book entitled *An Architectural Model for Service Interoperability*. He is an associate professor at the School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India. His research areas include service-oriented architecture, service interoperability, artificial intelligence, and machine learning. He has filed and granted five design patents and one international patent.

Mr. Natarajan C is an Associate Professor in the Department of Computer Science and Engineering at P.S.R. Engineering College. He has 21 years of teaching experience, including 8 years of teaching experience at African National Universities in Ethiopia. He handled interesting roles and responsibilities, including teaching in various institutions with more success factors like Trainer of Trainers and a Technical Trainer for the Placement Department. His areas of interest and specialization include data structure, database management systems, artificial intelligence, machine learning, deep learning, and data analytics. He is a reviewer and guest editor in various renowned international journals. He has published 14 papers, 5 conference articles, and filed one patent.