SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Integrated Rnn-Based Rvm Model to Detect Bipolar Disorder

Dr. Lingeswari Sivagnanam[1], Dr. S. Malathi[2], Dr. S. Kanageswari[3], Dr. Angel Nancy A[4], Dr. Stephy Joy D[5], Dr Priyanka M[6]

[1]Assistant Professor, Department of Computer Science w/s in AI & ML, SRM Institute of Science and Technology, Ramapuram, Chennai 600089. E-Mail: tejumitu1279@gmail.com, Orcid: http://orcid.org/0009-0005-9149-1304
[2]Assistant Professor, Department of Computer Science, ST. THOMAS COLLEGE OF ARTS AND SCIENCE, KOYAMBEDU, CHENNAI -600107, E-Mail: drmalathisambandan@gmail.com, ORCID: 0009-002-9091-7201
[3]Assistant professor, Department of Computer science, Loyola College, Chennai. E-Mail: kanageswari@loyolacollege.edu, ORCID: 0000-0001-8780-5757
[4]Assistant Professor in Computer Science, Women's Christian College, Chennai, University of Madras
Mail: angelnancy@wcc.edu.in, ORCID: 0000-0003-2407-8964
[5]Assistant Professor, Department of computer science with AI, Women's Christian College, Nungambakkam, Chennai.
E-Mail: stephy@wcc.edu.in, ORCID: 0000-0002-5929-343X
[6]Assistant Professor, Department of IT, Women's Christian College, Nungambakkam, Chennai. E-mail: priyanka@wcc.edu.in

Abstract
This research introduces a novel integrated model that combines Recurrent Neural Networks (RNN) with the Relevance Vector Machine (RVM) to resolve key challenges in mental health diagnosis. The proposed RNN-based RVM (RR) model advances traditional methods by enhancing the accuracy of bipolar disorder detection, thereby supporting early intervention and improved clinical outcomes. The objectives are to leverage RNN for large-scale data analysis and RVM for robust classification, to validate performance across diverse datasets. By applying the integrated model to three different types of datasets, the study demonstrates its potential to enable earlier diagnosis, more effective treatments and reduced diagnostic burden in Bipolar Disorder (BD) detection. The findings are expected to provide valuable insights into the role of advanced machine learning techniques in enhancing BD diagnostics.

Keywords: Recurrent Neural Networks (RNN), Relevance Vector Machine (RVM), RNN-based RVM (RR), Integrated model and Bipolar Disorder (BD).

1. INTRODUCTION

The core component of the Knowledge Discovery in Databases (KDD) process is data mining, which involves the extraction of meaningful patterns, trends and insights from large volumes of highly variable data using techniques in statistics, machine learning and artificial intelligence methods (Fayyad et al. 1996; Han et al. 2011). Machine Learning (ML) is considered as an advancement in Data Mining. They are deeply interconnected, machine learning can provide algorithms and techniques that enable advanced data mining processes and can automate the discovery of patterns, relationships and insights from data (Kotsiantis, Sotiris B. 2007). Bipolar disorder is a kind of mental illness, The diagnosis of BD is primarily clinical and is based on identifying typical patterns of manic, hypomanic and depressive episodes. The detection of bipolar disorder is performed by the integration of RVM and RNN model as RR model using three different types of datasets. The datasets used in this research are the WESAD data that consist of the psychological signal-based data, multimodal data and the real-world clinical data.

2. LITERATURE REVVIEW

Drougkas et al. (2024) used CNN and RNN models to examine multimodal ML (speech, text, physiological signals) and found AUC of 0.87 and precision of 0.85 for identifying disorder markers. The study claims that multimodal fusion provides increased performance over unimodal methods but adds more complexity and timing issues across the different modalities. Li et al. (2024) developed a hybrid RNN-RVM model that combined longitudinal EEG and clinical features to detect bipolar disorder in earlier stages. They achieved an overall accuracy of 92%, while interpretability improved, highlighting the value of integrating temporal neural data with the probabilistic nature of RVM modeling. Calesella et al. (2024) constructed a multimodal ML pipeline that integrated structural MRI with clinical data to discriminate between BD and MDD. Their model was trained on 1,200 neuroimaging scans, based on SVM and gradient boosting. The model achieved AUC of 0.91, accuracy

of 88% and the authors concluded that their model outperformed individual modality models, highlighting the diagnostic power of multimodal fusion of neuroimaging modalities and clinical data to discriminate mood disorders.

3. DATASET DESCRIPTION

The research work utilizes the Wearable Stress and Affect Detection (WESAD) dataset (Schmidt et al., 2018), a widely used multimodal dataset collected from 15 healthy individuals. WESAD comprises physiological signals recorded from both wrist-worn and chest-worn devices, including Electrodermal Activity (EDA), Heart Rate (HR) via Blood Volume Pulse (BVP), Accelerometer (ACC), Temperature (TEMP) and Respiration (RESP), along with corresponding affective state labels such as amusement, baseline and stress conditions.

The Multimodal Dataset contains psychotherapy interview transcriptions processed through the Speech-to-text transcription was performed using Google Cloud Speech API platform. Turkish Audio-Visual Bipolar Disorder Corpus (AVEC 2018 BDS), existing dataset utilizes audio and visual information to investigate human behaviours and emotional states. the features of the multimodal dataset which includes 46 speaking subjects with 352 videos, in the audio format of .wav and video format of .mp4, gender, etc.,

Real-World Clinical Dataset for this research has been collected from Mental Health Centre. The data was retrieved from the patient's data sheet. It contains both the patient's history as well as treatment details. Data relevant to the research has been analysed with the help of a Clinical Specialist associated with mental health care. The data retrieved from the excel sheet and stored in .CSV format. Around twenty-three features such as patient ID, age, gender, family history, total duration, psychiatric illness, substance use etc., were collected from the patient's clinical datasheet.

4. METHODS

4.1. RECURRENT NEURAL NETWORK MODEL

Neural Networks are computational models inspired by the human brain, consisting of nodes (neurons) that are interconnected. RNN is a type of neural network that processes sequential information (Goodfellow et al. 2016). Recurrent Neural Network (RNN) represent an exciting analytic strategy for studying and diagnosing bipolar disorder due to their ability to model the temporal and sequential patterns in physiological and behavioural data.

Moreover, RNN-based methods facilitate the development of automated, context-aware diagnostic systems that allow for ongoing monitoring and personalized interventions. RNN will be able to analyze long-term dependencies in time-series data, rather than relying solely on static features as traditional models. Empirical studies support the role of RNN in advancing data-driven psychiatry by facilitating early detection, prognosis, and longitudinal monitoring of bipolar disorder. These findings are further supported by other empirical work (PubMed, 2024; OpenMETU, 2023; PNR Journal, 2022).

The structure of RNN is shown in Figure 1 it consists of an input, recurrent hidden units and an output, all connected through weight matrices that are updated during the training process and continues to be a central tool in sequence modeling tasks.

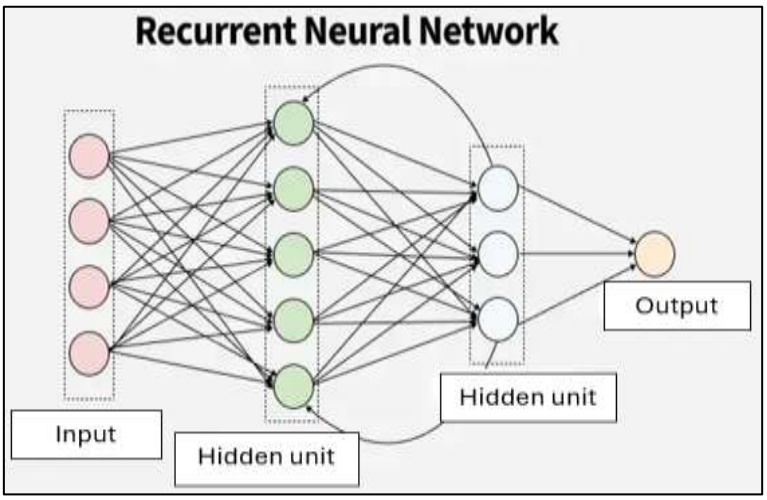

Figure 1. RNN Structure

The pictorial representation of the RNN model is depicted in Figure 2. Accordingly, an RNN adds recurrence, meaning that the hidden state at time step t depends on both the current input $(x_1, x_2, \ldots\ldots, x_T)$ and the hidden states $(h_1, h_2, \ldots\ldots, h_T)$ from the previous time step, in contrast to

other neural network models that process fixed-size vectors independently. This produces the final output at y from the last hidden state h_T.

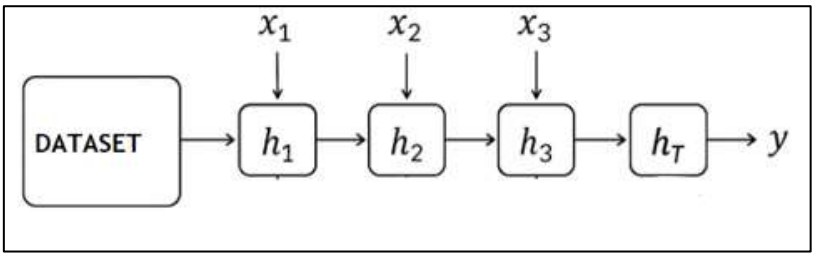

Figure 2. Representation of RNN Model

For Dataset D with input sequence, $x = (x_1, x_2, \ldots \ldots, x_T)$, the RNN updates its hidden state h_t at each time step t as follows:

$$h_t = f(W_{xh}x_t + W_{hh}h_{t-1} + b_h) \qquad equation\ (1)$$

where x_t is the input at the current time step and T denotes the sequence length (number of time steps). W_{xh} is the weight matrix mapping input to the hidden state, W_{hh} is the recurrent weight matrix mapping previous hidden state to current hidden state, h_{t-1} is the hidden state from the previous time step, b_h is the bias term and f is a nonlinear activation function. The function f can be replaced with ReLU,

$$ReLU(z) = max(0, z) \qquad equation\ (2)$$

The RNN hidden state equation with a function f will be defined as follows:

$$h_t = max(0, W_{xh}x_t + W_{hh}h_{t-1} + b_h) \qquad equation(1)$$

RNN face challenges due to gradients, which refer to how much each weight contributed to the error. As the input sequences grow longer, they suffered from the vanishing gradient problem, where gradients become too small to update parameters effectively and the exploding gradient problem, as gradients grew uncontrollably.

Traditional RNN used sigmoid or tanh activations, which compress values into a narrow range and often cause gradients to vanish when backpropagated through many time steps. RNN with ReLU activations become more stable when modeling complex sequences. In practice, ReLU and its variants (Leaky ReLU, Parametric ReLU) became popular choices for more complex recurrent architectures because of their efficiency and ability to accelerate the training process.

After the inputs have been completely processed, the output is generally computed by means of the Softmax activation function in classification methods as follows:

$$y_t = g(W_{hy}h_t + b_y) \qquad equation\ (4)$$

where g refers to the Softmax activation function that converts the raw scores (logits) into probabilities across all classes for output, y_t is the RNN's output (class probabilities or predictions) at time step t, W_{hy} is the weight matrix connecting hidden states to the output layer and b_y is the bias vector for the output layer.

Recurrent Neural Network (RNN) is modelled after the human brain's sequential processing mechanism, where previous events shape the associated experiences that are perceived in a current form. Their function is an external representation of the human cognitive ability to remember, engage with and learn through temporal patterns. The main features of inspiration are listed below:

- Temporal Memory: RNN duplicate the human ability to remember temporal information in the past to evaluate choices in the present or future.
- Sequential processing: RNN process input based on time, as humans naturally think about and interpret events in terms of time.
- Context awareness: RNN develop understanding of previous state and current state and as a result, they learn the possibility.
- Constantly Learning: The model continuously rewrites its hidden states as new data points occur and therefore, it structures data around time.
- Information storage: RNN have recurrent connection to hold memory across time steps to model sequences.
- Shared Parameters in Time: A single set of weights learns across all time steps, and therefore learns in sequences without the complexity of the processes.
- Mimicking human brain function: RNN emulate some aspects of short-term memory and sequential reasoning that is common among neural activity of the brain.

Ultimately, RNN can provide a feasible method for processing, analyzing and possibly predicting outcomes from sequential or time-track data.

Despite the above advantages of RNN, there are still several issues that limit the efficiency and scalability of Recurrent Neural Network (RNN) such as vanishing and exploding gradient problems,

limitations in learning long-term dependencies, slow sequential training, high computational cost associated with the long sequence processing and limited memory for higher sequences.

4.2. RELEVANCE VECTOR MACHINE LEARNING MODEL

The Relevance Vector Machine (RVM), introduced by Tipping (2001), is a sparse Bayesian learning model that provides an elegant probabilistic alternative to the Support Vector Machine (SVM). While SVM often rely on a large number of support vectors to define the decision boundary, RVM employs a Bayesian framework to automatically prune irrelevant training samples, leaving only a small subset called relevance vectors that are sufficient to represent the model (Tipping, 2001; Bishop, 2006). This sparsity results in simpler and computationally efficient classifiers, making RVM particularly attractive in various domains where interpretability and efficiency are crucial. The RVM produces probabilistic output, which can be particularly useful for estimating uncertainty in medical diagnoses.

In the research landscape of bipolar disorder, RVM has been useful for examining various sources of data associated with diagnosing patients, as well as identifying relevant biomarkers that characterize the difference among healthy controls, individuals diagnosed with bipolar disorder and various psychiatric groups. RVM has inherently employs a sparse learning approach that highlights the most salient and informative aspects of the features that contribute most to the model, in order to aid in interpretability and diagnostic confidence. Tipping (2001) has established the robustness and accuracy of RVM in predicting and understanding bipolar disorder (Costafreda, 2011).

Figure 3 shows the representation of Relevance Vector Machine (RVM) that begins with input features and contains Bayesian priors to identify important structure. In the next step, a kernel expansion layer builds a model for underlying nonlinear relationships. At the end of this process, RVM outputs probabilistic predictions that articulate classification and decision confidence based on posterior probability.

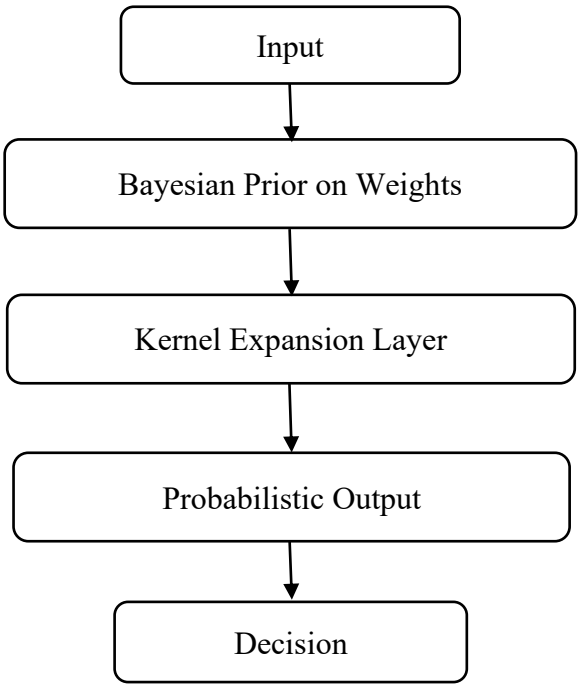

Figure 3. Representation of RVM Model

The overall method produces a sparse, interpretable and reliable model, justifying the usefulness of RVM in complex classification tasks such as bipolar disorder inferences using various data. This combination of sparsity, probabilistic reasoning and interpretability makes RVM a powerful tool for classification tasks with real-world constraints (Tipping, 2001; Bishop, 2006).

Formally, given a training dataset $D = \{(x_i, y_i)\}_{i=1}^{N}$ where x_i are the input features and $y \in \{-1, +1\}$ are the class labels, the decision function of RVM is expressed as follows:

$$y(x) = \sum_{i=1}^{N} w_i k(x, x_i) + b \qquad equation\ (5)$$

Here, $k(x, x_i)$ is a kernel function (commonly Gaussian RBF), x is the new input, w_i is the weight assigned to the training point x_i, b is the bias and the resulting score $y(x)$ is a real-valued logit that can be passed through a sigmoid activation function to yield a probabilistic interpretation. $\sum_{i=1}^{N}$ is the summation over training samples. The model combines contributions from all training samples, N is the total number of training samples. However, in practice, only a few samples survive as relevance vectors, because most weights $w_i \approx 0$. RVM places a zero-mean Gaussian prior over each weight.

$$p(w_i \mid \alpha_i) = \mathcal{N}(w_i \mid 0, \alpha_i^{-1}) \qquad equation\ (6)$$

where α_i is a hyperparameter controlling the precision of each weight. During training, most α_i values tend to infinity, effectively driving the corresponding w_i to zero. This results in automatic sparsity, since only a few non-zero weights remain corresponding to the most informative training samples, the relevance vectors. $\mathcal{N}$ is the Gaussian distribution, used to probabilistically model the weights, enforce regularization and enable automatic sparsity through Bayesian learning. (Tipping, 2001; Faul & Tipping, 2002).

The predictive distribution of RVM is probabilistic, which distinguishes it from other classifiers. The raw score of $y(x)$ is passed through a sigmoid function to compute the probability of belonging to a class as follows:

$$P(c = 1|x) = \sigma\big(y(x)\big) = \frac{1}{1+e^{-y(x)}} \qquad equation\ (7)$$

The sigmoid function σ(·) maps values from −∞ to +∞ into a probability in the interval (0,1). P is the predictive probability that the new input x belongs to class 1 (the positive class), c represents the class label and e is Euler's number, a mathematical constant. This probabilistic output allows RVM to express not just the predicted class but also the model's confidence in its decision. Such interpretability is especially valuable in sensitive areas like bipolar disorder detection, where confidence estimation is critical for supporting clinical decision-making (Bishop, 2006; Tax & Duin, 2004).

Unlike many traditional classifiers, the integrated probabilistic reasoning with kernel-based learning captures complicated nonlinear data and yet still remains straightforward to use. Some of the key characteristic features of Relevance Vector Machine are as follows:

- Bayesian Basis: Underlying Bayesian-based probabilistic learning framework provides uncertainty in all predictions.
- Sparsity: Uses only a few relevance vectors for a compact and computationally efficient model.
- Automatic Relevance Determination (ARD): Automatically learns to discard, and keep only the most relevant features for the prediction task.
- Probabilistic Output: Provides posterior probabilities is compared with hard classifications, thus adding interpretability to the RVM's utilization for classification.
- Strong generalization capability: Reduces occurrences of overfitting through Bayesian regularization and therefore performs better on seen or unseen data.
- Kernel Flexibility: Can incorporate various kernel functions, thus modelling for complex nonlinear relationships in the data.
- Interpretability: Useful for informing a practitioner about what data and features were influential (or not) for driving the decision boundary for prediction.
- Efficiency: Achieves comparable or higher accuracy with fewer parameters with many more parameters such as Support vector machine and therefore cost less computationally.
- Noise robustness: More suitable when the data is very noisy and uncertain, therefore more reliable in real world situations.
- Field applicability: Particularly well used in neuropsychiatry and medicinal research. For example, for bipolar disorder classification.

Taken together, the RVM's simple use with very few parameters and very few vectors lead to strong generalization capability that it can achieve high accuracy and prediction power for classification tasks of bipolar disorder detection.

Although the RVM is highly effective it is not free from limitations. The computational time is relatively high and the complicated Bayesian optimization means that the RVM procedure could experience some convergence issues while training the model. Furthermore, RVM performance can be sensitive to initialization, and the RVM procedure can become very complicated when using large datasets since it involves significant matrix operations. The software support is also limited, as is handling multi-class problems. Noisy or irrelevant data can also affect its performance.

5. PROPOSED METHODOLOGY

The proposed system follows a structured pipeline, ensuring optimal training, testing and classification using the RNN-based RVM (RR) model. Recurrent Neural Network (RNN) and Relevance Vector Machine (RVM) integration is based on both models. RNN is well suited to learn temporal dynamics over sequences, as it can retain memory across time steps. RVM sparse Bayesian models generate probabilistic predictions and eliminate irrelevant data points through automatic piecewise, relevance decisions.

The novel integrated modeling approach relies on sequential data input where the RNN captures meaningful temporal feature representations through feature extraction and processes them through the RVM. The RNN encodes temporal features to extract the most salient input representations. The salient input features are transformed into a kernel matrix that captures similarities among the data points. The computed kernel matrix is provided as input into the RVM to perform Bayesian learning on the received kernel and it identifies a small set of relevance vectors. Thus, it generates a decision boundary in the proposed space for classification among classes. For each new input sequence, the

RNN first extracts relevant temporal features, which are then processed by the RVM to perform classification and to generate probabilistic predictions corresponding to that sequence.
Through this integrated model, there can be a balance between RNN's temporal representation power and the RVM's interpretable and confidence-based decision-making, enhancing diagnostic precision for bipolar disorder.
The updating hidden state takes place with the help of equation (1). Once all the time steps of a sequence are processed, the last hidden state h_T at the end of the sequence is taken out, and a fixed-length feature representation of that sequence is stored. This is done to each sequence of the dataset in order to get a set of final hidden states. There is no explicit limitation or restriction in the memory of hidden states, as it depends upon the dataset and its features during the training process. Figure 4 presents the workflow of RR model.

Algorithm 1. RNN-based RVM (RR) Model
Input:
- Sequential dataset $D = \{(x^i, y^i)\}_{i=1}^{N}$
- RNN parameters: $W_{xh}, W_{hh}, W_{hy}, b_h, b_y$
- RVM parameters: kernel function k, weights w_i, bias b, pruning threshold ϵ

Output:
- Trained RNN-RVM (RR) model
- Sparse set of relevance vectors with probabilistic predictions

Procedure:
Step 1: Initialization
Initialize RNN (W_{xh}, W_{hh}, b_h) weights and RVM (w_i, α_i, b) parameters

Step 2: Sequence Encoding with RNN
(i) For each training sequence $x^{(i)} = \left(x_1^{(i)}, x_2^{(i)}, \ldots \ldots x_T^{(i)}\right)$
(ii) Finding hidden states according to equation (3.1), respectively.
(iii) Extract the final hidden state $h_T^{(i)}$ as the feature vector.

Step 3: Construction of Feature Set
(i) Form the new dataset of hidden representations:
(ii) $H = \{\left(h_T^{(i)}, y^{(i)}\right)\}_{i=1}^{N}$

Step 4: Kernel Mapping (RVM)
Compute the kernel matrix, $\Phi_{ij} = k(h_T^{(i)}, h_T^{j})$

Step 5: Sparse Bayesian Training (RVM)
(i) Estimate posterior distribution over weights according to equation (6)
(ii) Update hyperparameters α_i
(iii) If $\alpha_i \to \infty$, prune the corresponding feature vector $h_T^{(i)}$

Step 6: Identify Relevance Vectors
Retain only vectors with finite α_i as the relevance vectors (RV).

Step 7: Prediction for New Sequences
(i) For an unseen sequence x_{new},
(ii) Encode using the RNN to obtain
h_{new}
(iii) Compute the RVM decision function

$$y(h_{\text{new}}) = \sum_{i \in \text{RV}} w_i \, K(h_{\text{new}}, h^{(i)}) + b$$

(iv) Estimate class probability

$$P(\, c = 1 \mid h_{\text{new}}\,) = \sigma\big(y(h_{\text{new}})\big)$$

(v) Predict class label

$$\hat{c} = \arg\max_{c} P(c \mid h_{\text{new}})$$

Step 8: End of algorithm
Such RNN-based embeddings are fed to the probabilistic learning of RVM system. Embeddings are also initialized with a weight w_i and hyperparameter α_i. A kernel matrix $\Phi_{ij} = k(h_T^{(i)}, h_T^{j})$ is obtained with a predefined kernel $h^{(i)}$ and it is a measure of similarity between the hidden state.

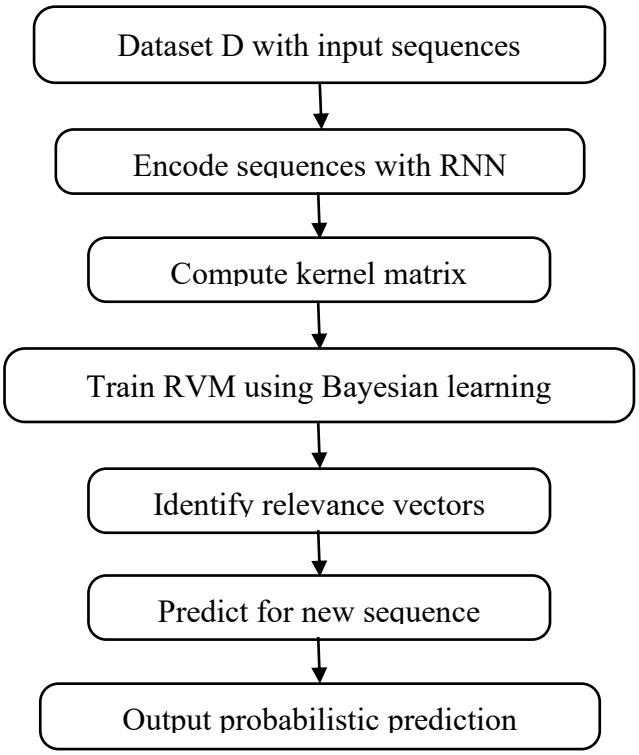

Figure 4. Workflow of RR Model

In training, when α_i happens to be very large, the corresponding feature vector $h^{(i)}$ becomes irrelevant to the model and can be removed. Such RNN-based embeddings are fed to the probabilistic learning of RVM system. Embeddings are also initialized with a weight w_i and hyperparameter α_i. A kernel matrix $\Phi_{ij} = k(h_T^{(i)}, h_T^j)$ is obtained with a predefined kernel $h^{(i)}$ and it is a measure of similarity between the hidden state. In training, when α_i happens to be very large, the corresponding feature vector $h^{(i)}$ becomes irrelevant to the model and can be removed.

After training, the rest vectors create a sparse subset called the Relevance Vectors (RVs), which are the main determinants of the resultant output. In case of a new sequence, the RNN structure performs the same calculation with the final hidden state, h_{new}. This is then sent to the RVM, as per the prediction calculation in the RR model architecture. The notation c refers to the class labelled as 0 or 1, to identify the result as non-bipolar or bipolar. The result is a probabilistic forecast that indicates the patterns learned in the sequence data as well as the sparsity-based selection of pertinent instances. This coupled RNN-RVM model, therefore, offers a unified model for temporal feature extraction and sparse Bayesian classification in detecting bipolar disorder.

The ultimate output is a probabilistic classification output that provides predicted class labels and posterior probabilities for each unseen input sequence. This allows the model to provide accurate decisions and interpretable confidence estimates, which support uncertainty-aware prediction for real-world applications.

6. RESULTS AND DISCUSSION

The consistent performance of the proposed RR model is evaluated using three different datasets as WESAD data, multimodal data (voice, audio and video) and real-world clinical data. The results are obtained by executing the algorithms for different runs. From the repeated runs, the outcome with the lowest error rate was selected, and its standard deviation is reported as the final result. Several parameters to be fixed were chosen based on the literature and heuristic methods are also followed to explore various combinations of these parameters. The values for the parameters applied in the RR model are given in Table 1.

Table 1. Parameters of RR Model

Parameter	Description	Value
Hidden state dimension (H)	The number of hidden units in the RNN controls feature representation capacity	128

Sequence length (T)	Number of time steps processed in each input sequence	Depends on the dataset
Relevance threshold (ϵ)	Cut-off value for pruning irrelevant vectors	Typically, 10^{-3} to 10^{-6}
Hyperparameters (α_i)	Bayesian priors controlling the sparsity of RVM weights	Initialized large, updated iteratively
Maximum iterations (max)	Limit for RNN training and RVM Bayesian optimization	150
Batch size	Number of sequences processed per training step	32

The online datasets are benchmarked or pre-classified whereas the real-world dataset was pre-processed using appropriate preprocessing techniques and their labels were compared with the results obtained. The comparison of objective function values across datasets such as WESAD Data, Multimodal Data and Real-World Clinical Data indicate the RR model's improved performance compared to RNN and RVM models. Table 2 represents the comparative analysis of objective function values for the RR Model.

Table 2. Comparative Analysis of Objective Function Values for the RR Model

Dataset	Values	RNN	RVM	RR Model
WESAD Data	Minimum Loss	0.184	0.165	0.091
	Max Loss	0.211	0.192	0.107
	Average Loss	0.196	0.178	0.099
Multimodal Data	Minimum Loss	0.152	0.138	0.079
	Maximum Loss	0.187	0.166	0.092
	Average Loss	0.169	0.152	0.084
Real-World Clinical Data	Minimum Loss	0.165	0.142	0.089
	Maximum Loss	0.193	0.171	0.101
	Average Loss	0.178	0.156	0.094

The RR model produced the lowest loss values (Minimum = 0.091, maximum = 0.107, average = 0.099) for the WESAD Data suggesting that the RR model can effectively maximize temporal feature representation while minimizing overfitting. The RR model shows a marked decrease in average loss (0.084) recorded for the multimodal data which is compared to both RNN (0.169) and RVM (0.152) models, indicating the RR framework's ability to manage complexity and multimodal noise by adaptive regularization. In case of real-world clinical data, containing the highest level of variability and noise, the RR model shows the lowest objective loss value recorded across all datasets (average = 0.094), further demonstrating strong generalization ability in realistic data conditions on the input data.

In general, the consistent decrease in loss value across all datasets suggests the RR framework is successful in determining relevance versus regularization for the model, resulting in consistent learning stability, convergence and predictive reliability to handle real-world data. These findings offer support for the utility of the RR model in clinical and behavioural prediction tasks, especially within scenarios like the identification of bipolar disorder when interpretation and heterogeneity of the data are of primary concern.

The empirical evaluation is conducted on various data sources with varying levels of data complexity and class-label distributions related to BD classification. The proposed RR model is compared with existing algorithms which include RNN and RVM, using performance evaluation metrics such as accuracy, precision, recall, F1-score and AUC. The models have been trained on 75% of the dataset and 25% is used for testing true and false predictions were analyzed to evaluate the model's performance. The data is transmitted through the models, so models learned patterns from the data to provide reliable and accurate results.

Table 3 shows the accuracy-based comparison of the proposed RNN-based RVM (RR) model. It shows that the RR model consistently performs better across datasets than the RNN or RVM models alone.

Accuracy is about 95.73, 93.35 and 91.72 for the WESAD, multimodal and real-world clinical datasets, respectively, which is significantly better than individual baseline models.

Table 3. Comparative Analysis of Accuracy

Dataset	RNN	RVM	RR Model
WESAD Data	83.04	85.63	95.73
Multimodal Data	87.21	85.06	93.35
Real-World Clinical Data	81.02	84.23	91.72

Table 4, Table 5 and Table 6 represent the precision, recall and F1-score and further validate the robustness of the RR model in accurately classifying bipolar states while minimizing false alarm rates. The proposed model achieves a precision of 94.73 and a recall of 95.32 in the WESAD dataset, highlighting its high sensitivity to changing neural signal features. In the multimodal dataset, the recall and F1-score (95.00 and 94.04, respectively) confirm the model's ability to manage text, audio and behavioural data features. The real-world clinical dataset shows a low but stable performance (accuracy of 91.72 and AUC of 90.74).

Table 4. Comparative Analysis of Precision

Dataset	RNN	RVM	RR Model
WESAD Data	83.04	85.63	95.73
Multimodal Data	87.21	85.06	93.35
Real-World Clinical Data	81.02	84.23	91.72

Table 5. Comparative Analysis of Recall

Dataset	RNN	RVM	RR Model
WESAD Data	83.01	84.27	95.32
Multimodal Data	89.14	87.90	95.00
Real-World Clinical Data	85.98	86.78	94.87

Table 6. Comparative Analysis of F1-score

Dataset	RNN	RVM	RR Model
WESAD Data	82.89	84.99	95.02
Multimodal Data	87.14	85.79	94.04
Real-World Clinical Data	84.02	85.06	93.76

Table 7 shows the AUC results that supports the discriminative ability of the RR model. The model achieves AUC of 94.32, 92.15 and 90.74 for the WESAD, multimodal and real-world clinical datasets, respectively. Overall, this further supports RR model ability to distinguish feature spaces from structured signal features to unstructured textual and behavioural data while being clinically interpretable and diagnostically reliable.

Table 7. Comparative Analysis of AUC

Dataset	RNN	RVM	RR Model
WESAD Data	80.05	85.12	94.32
Multimodal Data	86.74	85.92	92.15

Real-World Clinical Data	83.87	85.10	90.74

Figures 5-9 below are the graphical representation of the comparative analysis tables with respect to the performance evaluation metrics.

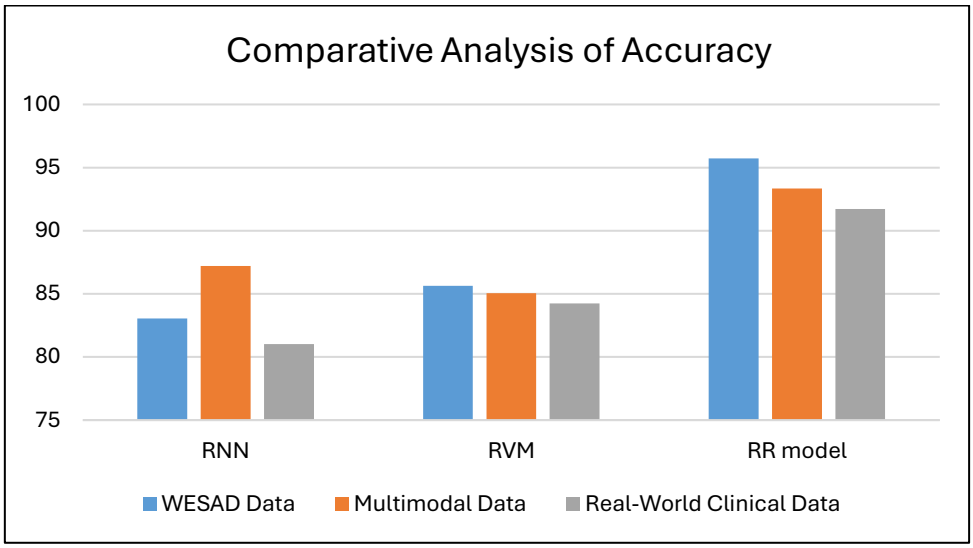

Figure 5. Comparative Analysis of Accuracy

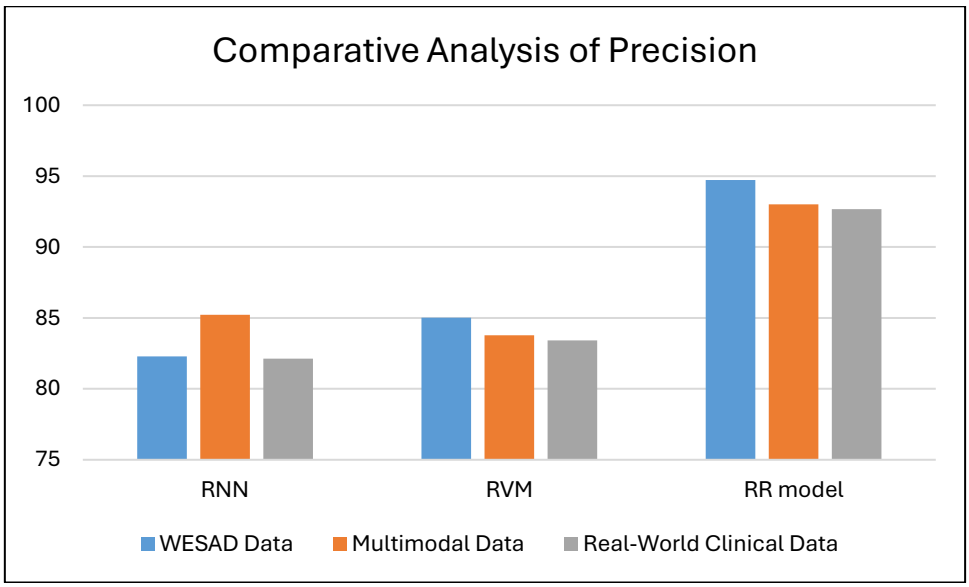

Figure 6. Comparative Analysis of Precision

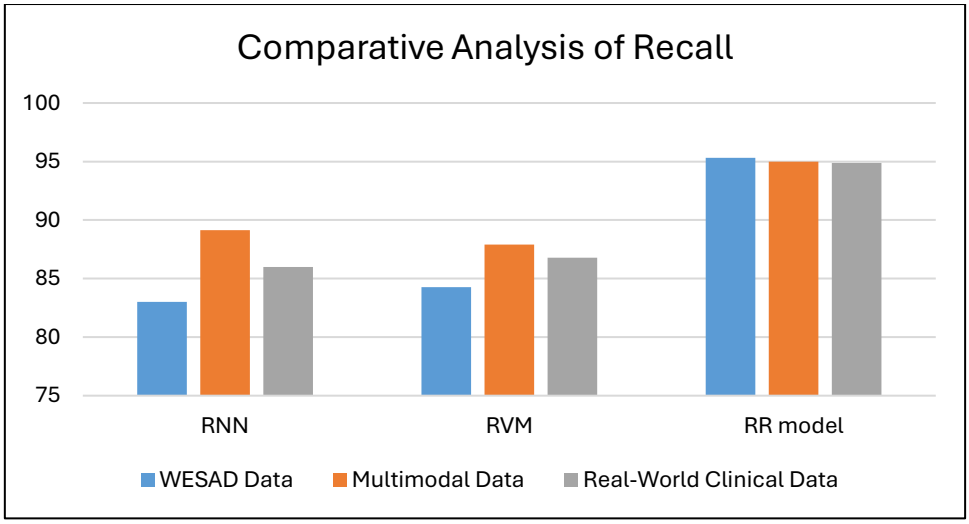

Figure 7. Comparative Analysis of Recall

The performance of the RNN, RVM and proposed RR model is delineated in the Figures across metrics such as Accuracy, Precision, Recall, F1-score and AUC. Each performance evaluation metric represents a separate aspect of model classification ability and reliability. Overall, it is evident that the RR model outperforms RNN and RVM in all metrics.

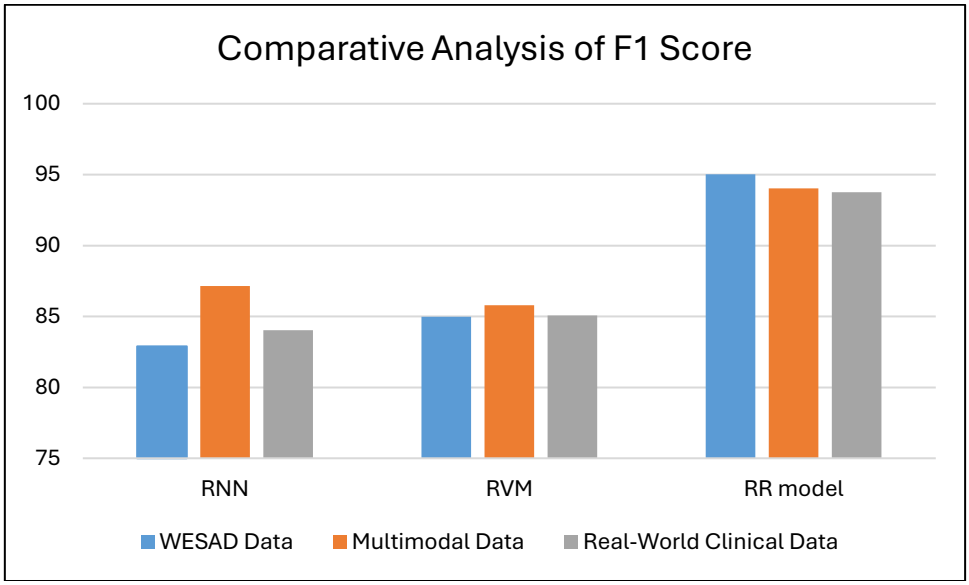

Figure 8. Comparative Analysis of F1-score

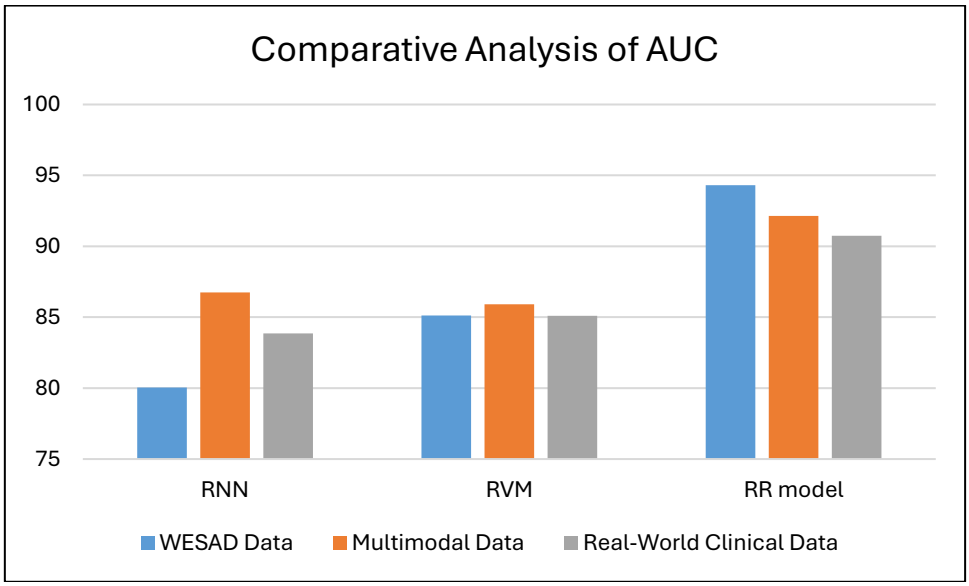

Figure 9. Comparative Analysis of AUC

From the above graphs, it is clearly shown that the scores are slightly lower on the real-world clinical data compared to the other two datasets. The proposed RR model still surpassed the existing methods, achieving an accuracy of 91.72, a precision of 92.68, a recall of 94.87, F1-score of 93.76 and an AUC of 90.74. The RVM model performs better than RNN on WESAD and multimodal datasets. Meanwhile, RNN showed stronger results on real-world data. Such differences may arise in the models that depend on the dataset characteristics, but the RR model consistently outperforms on all the datasets as analysed with other models in all five performance evaluation metrics.

7. CONCLUSION

It is evident from the study that the integration of RNN with RVM provides a powerful methodology for handling sequential data while ensuring sparse and probabilistic predictions. It proves to be the state-of-the-art integrated framework in selecting various parameters needed for classification. Results clearly suggest that the RNN based RVM model offers superior generalization, sparsity and probabilistic decision-making, but it is still facing challenges in its computational complexity, Bayesian optimization cost and scalability challenges.

References

1) Bishop, C. M. (2006). *Pattern recognition and machine learning*. pp. 1–58. Springer.
2) Tax, D. M. J., & Duin, R. P. W. (2004). *Cost-based classifier evaluation for imbalanced problems*. In Proceedings of the Joint IAPR International Workshops on Structural, Syntactic, and Statistical Pattern Recognition (SSPR) (pp. 762–770). Springer.
2. Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1, 211–244.
3. Faul, A. C., & Tipping, M. E. (2002). Analysis of sparse Bayesian learning. In *Advances in Neural Information Processing Systems (NIPS 14)* (pp. 383–389). MIT Press.
4. Costafreda, S. G. (2011). Pattern recognition of brain activation: Towards biomarkers in neuropsychiatry. *The British Journal of Psychiatry*, 199(1), 4–5.
5. Dunster, G. P., Swendsen, J., & Merikangas, K. R. (2021). Real-time mobile monitoring of bipolar disorder: A review of evidence and future directions. *Neuropsychopharmacology*, 46, 197–208.
6. Chien, Y.-L., et al. (2025). Using wearable device and machine learning to predict mood symptoms in bipolar disorder: Development and usability study. *JMIR Medical Informatics*.
7. Pfaffenseller, B., et al. (2025). Self-assessment and rest-activity rhythm monitoring for effective bipolar disorder management: A longitudinal actigraphy study. *International Journal of Bipolar Disorders*.
8. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
9. Ringeval, F., Schuller, B., Valstar, M., Cowie, R., Kaya, H., & Pantic, M. (2018).
10. *AVEC 2018 Workshop and Challenge: Bipolar Disorder and Cross-Cultural Affect Recognition*. Proceedings of the 2018 Audio/Visual Emotion Challenge and Workshop (AVEC 2018), 3–13. ACM.
11. Schmidt, P., Reiss, A., Dürichen, R., & Van Laerhoven, K. (2018).
12. Introducing WESAD, a multimodal dataset for wearable stress and affect detection. *Proceedings of the 20th ACM International Conference on Multimodal Interaction (ICMI 2018)*, 400–408.
13. Calesella, F., Colombo, F., Bravi, B., Fortaner-Uyà, L., Monopoli, C., Poletti, S., … Vai, B. (2024). *A machine learning pipeline for efficient differentiation between bipolar and major depressive disorder based on multimodal structural neuroimaging*. Neuroscience Applied, 3, 103931.
14. Li, H., Cui, L., Cao, L., et al. (2020). Identification of bipolar disorder using a combination of multimodality magnetic resonance imaging and machine learning techniques. *BMC Psychiatry, 20*(1), 488.
15. Drougkas, A., et al. (2024). Multimodal machine learning for mental health assessment. *Pattern Recognition, 146*.
16. Fayyad, U. M., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM, 39*(11), 27–34.
17. Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
18. Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica, 31*(3), 249–268.