

REACTION AND MATERIAL PREDICTION USING ARTIFICIAL NEURAL NETWORKS (ANNs) ALGORITHMS FOR CHEMICAL PREDICTIVE MODEL

V. Kanchana,^{1*} and Shanmuga Sundaram Anandan,²

¹Associate Professor, Department of Chemistry, AMET University, Chennai, Tamilnadu, India.

²Professor, Department of Mechanical Engineering, Sree Sastha Institute of Engineering and Technology, Chennai, Tamilnadu, India.

*Corresponding Author: kanrishi04@gmail.com

Abstract: Chemical reactions and the material prediction are crucial processes, which is currently applied in several industries like materials science, pharmaceuticals, and environmental engineering. The accurate prediction of the reaction and material process is very essential to the development of new compounds. Artificial Neural Networks (ANNs) have become powerful tools for this purpose due to their ability to manage the complex relationships in the chemical prediction process. Our research introduces a new method for predicting chemical reactions and material prediction using ANNs. We use different ANN structures, such as MLP, RNN, and GNN to model various aspects of chemical reactions and material properties. The reaction and prediction accuracy and loss factors are calculated for the methods like RMP-MLP, RMP-RNN and RMP-GNN. For the process of comparative analysis, the earlier research methods which are considered are SVM, RF and ANN. The parameters which are taken into concentration for the parameter analysis are MAE, MSE, RMSE and R^2 .

Index Terms: Reaction and Material Prediction, Artificial Neural Networks (ANNs), Multilayer Perception (MLP), Recurrent Neural Network (RNN), and Graph Neural Network (GNN)

Introduction:

Sustainable Chemistry is a concept of designing compound goods and growth to minimize or defect a utilize and creation about dangerous material [1]. In 1998, Anastas and Warner introduced the twelve principles of Green Chemistry, which provided a roadmap for creating sustainable chemical processes. These principles emphasize the importance of finding alternatives to harmful solvents. The fifth principle specifically recommends using harmless solvents when avoiding them altogether is not possible [2]. Despite an estimated global consumption of 30 million metric tons per year, there is still a concerning increase in the use of organic solvents in scientific and technological activities. Moreover, certain highly dangerous solvents continue to be widely used [3]. Therefore, implementing solvent-free processes aligns with the principles of Green Chemistry. While some progress has been made in this area, such as solvent less sample preparation methods and greener reactions without solvents, there are still an lot to do previously these strategies remains widely utensil. Another approach is to minimize solvent usage or replace hazardous solvents with cleaner alternatives, which is strongly recommended to reduce associated risks [4]. To aid in selecting safer options, various guides have been developed in literature. However, a lack of important data, including material and environmental properties, may hinder their practical application. To address this issue, default values or substitutions according to similar chemicals can be used when data is missing for a particular solvent class [5].

The utilization of ML might be employed into detect model and rage in extensive amounts about input through competent pattern such can make predictions for new instances. An demand of ML in a realm about chemistry is gaining prominence in research [6]. This resurgence can be attributed to various factors, including progressive as to graphics processing units, heavier pools of data, also novel code [7]. The growth is not evenly distributed, and they hypothesize that the reason for the faster advancements in fields like analytical chemistry comparison for alternative such as green combination reaction is an accessibility of vast guidance data pool in traditionally data-heavy fields. For instance, as to logical hermetic, ML algorithms were utilized toward identify substance type with focus under a typical detection restrict discovering hidden system within noise gesture. Nonetheless, significant evolution relic created as regards employing ML techniques in organosulfur chemistry as well [8].

2 Related Works:

In [9], they developed LASSO and RF-ML patterns to predict cyanobacteria blooms in aquaculture ponds, with LASSO showing superior performance ($R^2 = 0.918$). The findings highlight the importance of monitoring organic carbon and reducing phosphorus in feed for effective cyanobacteria prevention and pond management. In [10], the review highlights the capability of ML in processing design and appeal poly phenol-based relevant, offering solutions to the complexity and diversity of these compounds for various farmland like biomedical, ecological, and power sectors. In [11], an ANN, is specifically a multi-layer perceptron, to predict precise state-to-state rate continual for atom-diatom collisions by using a combination of few predict and many low-accuracy rates, demonstrating significant improvement over previous methods. In [12], the potential of energetic ML in chemical engineering, highlighting its efficiency and cost-effectiveness through merger ML with design of experiments. In [13], a NIRS, and chemo metric methods like PCA to assess Yongchuan Xiuya tea quality. The model combining NIRS with BP-ANN demonstrates rapid and accurate assessment, showcasing potential for evaluating other teas' quality digitally. In [14], the use of NIR spectroscopy

with ANN modeling showed strong correlations, proving effective in predicting total polyphony, antioxidant activity, and extraction yield from NIR spectra of the extracts.

In [15], in Sohag, Egypt, utilizes DRIFT-FTIR spectroscopy coupled with machine learning patterns like PLSR to estimate SOC, showing PLSR as the most effective with an R^2 of 0.82. The potential for this approach in predicting SOC but recommends further studies to enhance model accuracy and compare with traditional soil analysis methods. In [16], employs CPANN to develop accurate patterns predicting the binding of compounds to ARs and ERs, offering a vital tool for screening EDCs and prioritizing safety assessments. In [17], the "Anti-Dengue" algorithm combines QSAR and machine learning techniques to predict dengue virus inhibitors, achieving a PCC of 0.81 on an independent validation dataset. It identifies potential repurposed drugs like goserelin, gonadorelin, and nafarelin, offering a promising tool for accelerating antiviral drug development against dengue outbreaks. In [18], ANNs to portend mechanical properties of stainless steel placed on alchemist formula and gelid depth decrease, offering valuable insights for materials engineering. The patterns serve as supportive appliance meant for plotting original inox class and optimizing cold shaping action. In [19], forecast elasticity of cold shaping inox according to bio chemistry and gelid depth decrease, aiding in the development of new grades and optimization of cold-forming processes in materials engineering.

In [20], classical logistic regression with AI techniques like ANNs, SVMs, DTs for bankruptcy prediction in the Slovak chemical industry, finding similar outcomes across methods and highlighting the ongoing debate on the potential advantages of advanced AI patterns in financial analysis. In [21], showcasing the efficiency of ANN patterns along r^2 until 1.000 and SVM patterns with r^2 stretch out 0.981. Sensitivity inspection highlights the importance of dehydrated peach and osmotic pretreatment, leading to optimal formulation with 15% lyophilized peach for cookie quality enhancement. In [22], he production of kombucha fresh cheese with desired standard character, demonstrating high r^2 values. The model identifies optimal trial, temperature fluid extract of sage, showcasing specific favorable process boundary for increase product development in the dairy industry. In [23], a SOQASP utilizing AIoT sensors and deep learning patterns like ANN, CNN, and LSTM, achieving high accuracy in assessing sesame oil quality. The platform offers a user-friendly and efficient solution for digitizing oil quality measurement, supply observation, and scalability in the industry. In [24], a MR-PSO is used to accurately predict mechanical belongings of insulation paper in power transformers less than thermal aging. It provides valuable insights for maintenance planners, simulating nonlinear behavior with acceptable margins of error according to time, temperature, and chemical markers.

In [25], the effectiveness of ANN patterns coupled with NIR spectroscopy for predicting material and combustion about oil-seed crop educe mixture with high accuracy. Various NIR spectral preprocessing techniques increase efficiency of ANN model for quantitative forecast. In [26], a CS-R-P for MB removal, optimized using an ANN coupled with IWGO creed, accomplish a high chance-half correlation along $R^2 = 0.99$. The aerogram demonstrates effective MB adsorption through exchange and hydrogen binding, with the optimized conditions leading to a 94.6% removal efficiency. The summary of the earlier approaches are described in table 1.

Table 1: Earlier Research Summary

Ref. No	Algorithm	Methodology	Advantages	Disadvantages	Performance	Efficiency	Accuracy	Features Used	Measurements
[9]	LASSO Regression	ML patterns (LASSO regression & random forest) to predict cyanobacteria abundance in aquaculture ponds	Feasible prediction of cyanobacteria abundance, operational solution	Limited monitoring data might affect prediction accuracy	R2 = 0.918	High	High	Chemical Oxygen Demand, TOC	Cyanobacteria abundance
[10]	Random Forest	ML application for poly phenol-based materials research and development	Proficiency in predicting, optimizing, designing materials	Complexity and diversity of poly phenol-based materials	NA	NA	NA	Polyphonic compounds, chemical structures	Properties and applications of poly phenol-based materials
[11]	Artificial Neural Network	ANN for predicting precise state-to-state rate firm for atom-diatom collisions	Accurate predictions, improved rate constants	Need for accurate rates as input, complexity	Prediction accuracy from low-accuracy rates	High	High	High	Atom-diatom collision rates
[12]	Active Machine Learning	Combining machine learning with design of experiments for efficient research in chemical engineering	More efficient than traditional methods, investigations across scales	Challenges in convincing researchers, data creation flexibility	NA	High	NA	Chemical engineering processes	Process investigations and optimization
[13]	PLSR, BP-ANN	Attribute appraise pattern for Yongchuan Xiuya tea using NIRS along with machine learning methods	Rapid, accurate, digital assessment of tea quality	Model comparison needed for best results, potential model bias	RP2 = 0.9315	High	High	NIR Spectroscopy, PLSR, BP-ANN	Tea quality attributes, chemical components
[14]	Artificial Neural Network	Prophecy of material and combustion on grape skin extract using NIR spectroscopy and ANN patterns	Valuable tool for predicting properties, correlation with NIR spectra	Data preprocessing required, model calibration	NA	High	High	NIR Spectroscopy, ANN	material and combustion of grape skin extract
[15]	PLSR, ANN, SVR, RF	Estimating SOC in Sohag, Egypt service DRIFT-FTIR spectroscopy and ML patterns	Potential for SOC estimation, various ML patterns tested	Variable performance among ML patterns, need for further improvement	R2 = 0.82	Moderate	Moderate	DRIFT-FTIR Spectroscopy, PLSR, ANN, SVR, RF	Soil Organic Carbon (SOC) in Sohag, Egypt
[16]	Counter-Propagation Artificial Neural Networks	Screening and prioritizing EDCs using ANN patterns for binding to nuclear receptors	Excellent prediction accuracy, safety prioritization tool	Need for diverse compound data-set, potential model bias	Prediction accuracy for binding affinity to nuclear receptors	High	High	High	Endocrine-disrupting chemicals (EDCs), ANN
[17]	SVM, ANN, kNN, DT	Developing "Anti-Dengue" algorithm for predicting dengue virus inhibitors using QSAR and ML patterns	Effective in identifying potential drug candidates, QSAR applications	Potential for over fitting, model validation challenges	PCC = 0.81	High	High	QSAR, SVM, ANN, kNN, RF	Dengue virus inhibitor compounds
[18]	ANN, MLR	Predicting strain-hardening properties of stainless-steel using ML patterns	Support for materials engineering, prediction of mechanical properties	Potential for reduced data quality, comparison challenges	NA	High	NA	Chemical composition, Cold thickness reduction	Mechanical properties of stainless steel
[19]	ANN, MLR	Predicting strain-hardening properties of stainless-steel using ML patterns	Support for materials engineering, prediction of mechanical properties	Potential for reduced data quality, comparison challenges	NA	High	NA	Chemical composition, Cold thickness reduction	Mechanical properties of stainless steel
[20]	LR, ANN, SVM, DT	Bankruptcy prophecy patterns for manufacture in Slovakia using classical and AI methods	Effectiveness in bankruptcy prediction, comparison of methods	Similar outcomes across methods, input data quality concerns	NA	High	High	Financial indicators, LR, ANN, SVM, DT	Bankruptcy prediction in chemical industry
[21]	SVM, ANN	Quality prediction and optimization of cookies enriched with dehydrated peach using SVM and ANN patterns	Efficient prediction of quality parameters, multi-objective optimization	NA	r2 up to 1.000	High	High	Dehydrated peach presence, Treatment methods	Quality parameters of cookies
[22]	Artificial Neural Networks	Producing kombucha fresh cheese with desired quality parameters using ANN patterns	Accurate prediction of quality, robust generalization capabilities	NA	r2 = 0.993	High	High	Chemical composition, Antioxidant potential,	Quality parameters of kombucha fresh cheese

								Microbiological profile	
[23]	ANN, CNN, LSTM	Sesame Oil Quality Assessment Service Platform using AIoT sensors, ANN, CNN, LSTM	Rapid assessment of oil quality, effective AI methods	Data preprocessing complexities, model calibration	LDA = 95.13	High	High	AIoT sensors, ANN, CNN, LSTM	Quality assessment of sesame oil
[24]	ANN, PSO	Predicting mechanical properties of insulation paper in power transformers using ANN patterns	Accurate simulation of insulation behavior, optimized model	Need for optimization, experimental data comparison	NA	High	NA	Tensile strength, Degree of polymerization (DPV), Time, Temperature	Mechanical properties of insulation paper in transformers
[25]	ANN	Characterization of oil-in-plant extort immiscible through NIR spectra and ANN patterns	Quantitative prediction of properties, efficient modeling	Spectral preprocessing effects, model tuning	R2 > 0.9	High	High	NIR Spectra, ANN	material and combustion of emulsions
[26]	ANN, IWGO	Adsorption of ethylene blue using chitosan-based aerosols modeled and optimized with ANN and IWGO algorithms	Efficient optimization, robust modeling	Model complexity, optimization parameters	Correlation coefficient R2 = 0.99	High	High	Chitosan-resole-pectin concentration, Thermal treatment	Adsorption properties of chitosan-based aerogels

3. Fundamentals:

3.1 Fundamentals of MLP:

The MLP is a kind of FF-ANN that number of multiple layers. This associated layer is referred to in different ways in the field of ANN, including ambiguous, strict, and loose terminology. The MLP is comprised of three specific layers: hidden, output, input layers. Each node within the network is linked to one of these layers. The Hidden and Output layers make use of nonlinear activation methods, while the Input layer utilizes linear activation techniques. The MLP also employs a supervised learning approach known as back propagation. This method has various applications such as speech recognition and image recognition. The ANN is a popular method for solving complex issues like clustering, classification, pattern recognition, and prediction by using data-based modeling. The structure includes three tiers: input, concealed, and output. The axon in each tier are linked to those in the subsequent tier via weights that range from 1 to -1. Every neuron in MLP-NN conducts two mathematical computations - addition and multiplication - on the inputs, weights, and biases.

$$S_j = \sum_{i=1}^n w_{ij} I_i + \beta_i \quad (1)$$

This formula uses n to represent number of inputs, I_i as the input variable, β_i as bias meant to jth axon, and w_{ij} as weight connecting the input to the neuron. To continue, an activation function must be applied to correlation 6's output. Several types of activation functions can be used in an MLP-NN, with some of the most common options listed. The network must then be trained by adjusting and updating weights and biases using different guidance algorithms. Popular options include LM, CG, NM, GD, and QN. The choice of a guidance algorithm should depend on the specific problem. However, LM has been found to have faster computational speeds compared to other methods, but it requires more memory. In contrast, GD has slower computational speeds but requires less memory. Therefore, once agreements amount of data, using GD may be more advantageous in terms of memory usage.

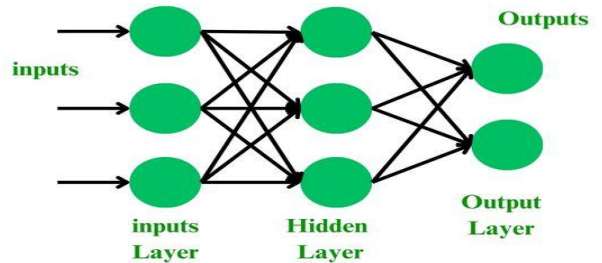


Figure 1: A Structure of Multi-layer Perception

Fundamentals of RNN:

RNNs have a similar structure to feed forward neural networks, but with one key difference: they have a recurrent secret province that relies on the previous time's activation. This makes them part of a diverse group of patterns capable of performing unpredictable computations. In 1991, Siegelman and Sontag demonstrated that by using sigmoidal activation functions, a finite-sized RNN can replicate a universal Turing machine. Due to their ability to handle temporal dependencies, RNNs are well-suited for tasks involving sequences of interdependent data points. The fundamental architecture of an RNN. In contrast, feed forward networks simply take inputs and produce outputs, making them useful for classification tasks such as identifying images as apples or oranges. However, unlike RNNs, there is no connection between present and past predictions in feed forward networks. For example, if an image is classified as an apple in the first moment, this will not affect the classification of the next image. Even if the next image is an orange, it will still be undisclosed as such. This is not the case with RNNs, as their predictions at each moment are influenced by both previous outputs and current inputs.

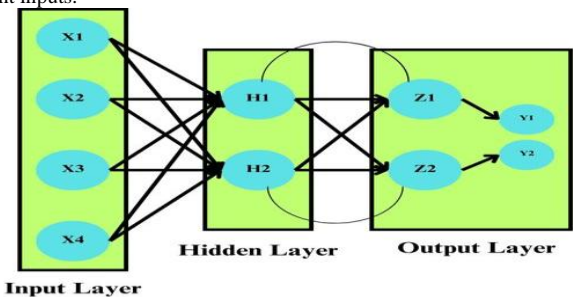


Figure 2: Architecture of RNN

The FFN has a surface input elements of instance, input size, while recurrent neural networks have a three-dimensional structure of instance, input size, time series length. The FNN follows the sequence of Input - Hidden - Output in its three phases of operation. This same order is also observed as regards

RNN as Input & Previous Hidden Output. An uncoil RNN takes the first input X_1 to calculate an secret province H_1 and output Y_1 . For subsequent time steps H_2 , secret province H_1 is computed using previous secret province also input, followed by calculating output Y_2 . This can be expressed using the RNN formula, which states that the current secret province is determined by a function f that takes into account previous secret province and current input, amidst as its parameter. A RNN typically learns to use $h^{(t)}$ as a summary of important information from past inputs $h^{(t-1)}$ up to time $x^{(t)}$. To calculate defeat conducive to t inputs and related outputs, the loss function sums up absence everywhere time steps. As this consider, RNN are utilized to forecast rainfall patterns in Bihar from 1901 to 1997 and evaluated on data from 1998 to 2017. The findings suggest that past rainfall data can impact future rainfall.

$$L(\{x^{(1)}, \dots, x^{(t)}\}) = \sum_t L^{(t)} = \sum_t -\log y_{y^{(t)}}^{(t)} \quad (2)$$

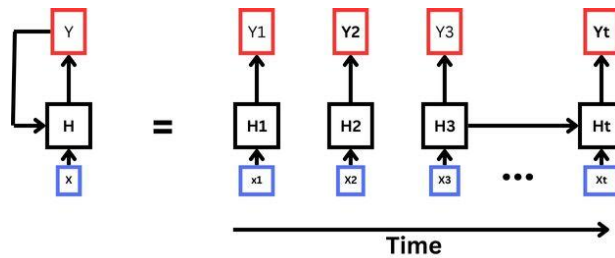


Figure 3: Unrolled RNN

Fundamentals of GNN:
 A fore mentioned, tag distinguishing special of GNNs compared to one more popular types of neural networks is that their internal structure is not predetermined. Instead, it adapts to the input graph's specific layout. As regarding GNN, mutually factor within the graph illustrate by a secret province, which is initially a angle containing predetermined features. Through a process called pass a message, these secret provinces are associate according to linked in the graph. This occurs through successive repetition, during which each node's secret province is updated using information from its neighboring nodes. This allows for gradual dissemination of information throughout the entire graph over multiple repetition. Each node has its own unique representation according to its distinct perceptive, as be in receipt of its covenant. It should be noted that once a amount services performing a simple equals graph's equator, information has potentially spread throughout the whole graph. The GNN architecture including members: pass a message and recite. The pass a message component conducts insistent one way transfer operation more elements in graph and able either run for a set number of repetition (T) or to a criteria is met. Overall, each repetition of pass a message can be described using these equations:

$$M_{ij}^* = m(h_i^t, h_j^t, e_{ij}) \quad (3)$$

$$M_i^{t+1} = aggr(M_{ij}^*) \quad (4)$$

$$h_i^{t+1} = u(h_i^t, M_i^{t+1}) \quad (5)$$

To put it differently, the item responsible for passing messages runs three distinct functions that are carried out on all elements of the graph:

1. The message function $[m_{ij}(h_i, h_j, e_{ij})]$ encodes information about the connection between two nodes $(h_i \text{ and } h_j)$ in a graph. Way of in secret provinces of both lumps along an optional angle describing their connection

(e_{ij}) properties, and outputs a message containing relevant relationship (M_{ij}^*) information.

2. The accumulation province $[aggr(M_{ij}^*)]$: combines all messages from neighboring nodes to create a certain extent angle (M_i) representing the resident area of a burl. Its main purpose is to handle graphs with varying node degrees by producing a consistent output representation. This is usually achieved through element-wise total, although other methods such as mean, min, max, variance, or unions can also be used.

3. The modernize function $[u(h_i^t, (M_i^{t+1}))]$: then combines secret province of each node (h_i^t) with its accrued communication to create an updated representation (h_i^{t+1}) . This updated state is then used as input for the next repetition of pass a message, repeating the process outlined in equations (1)-(3). A message (m) and update (u) functions are utensil using separate NN, with their inside specification learned during guidance. Once

pass a message stage is complete, resulting secret provinces (h_i) go with recite item where they are transformed into output values $[y = r(|v \in V)]$ or labels for the GNN. It should be noted that GNNs have two types of outputs: node-level along with universal graph-level labels.

Anterior is obtained by individually forwarding recite function (r_n) into every another secret province, while the latter involves totaling all secret provinces and then applying a global recite function (r_G) to the resulting angle. Similar to the pass a message item, accumulation province allows for champion of graphs of any size by producing a fixed-size angle after all of

input graph size. A recite functions $(r_n \text{ or } r_G)$ are also approximated using neural networks, while accumulation province usually embrace element wise totality.

A typical GNN is composed of four distinct functions $(m, aggrandu, r)$, which are duplicated and combined depending on the input graph. This implies that same Message, Aggregation, Update, and recite functions are utilized total components within a graph. All through the guidance process, these functions are gradually approximated using a standard back propagation algorithm throughout the entire GNN structure. This takes into account both output labels of the GNN and those of guidance data-set. The crucial aspect here is that all NN parameters are collectively optimized for each guidance sample across all instances in GNN model. This leads to general note, Update, and recite functions iterate each node's opinions in a graph. Function declared with one can then referred new graphs on condition serve their intended purpose, such as predicting path delays in our scenario.

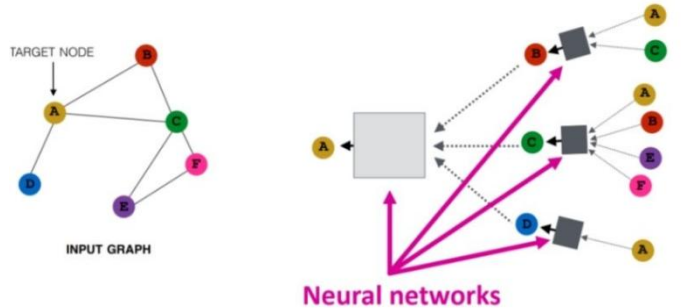


Figure 3: Architecture of GNN

PROPOSED RMP-ANN MODEL:

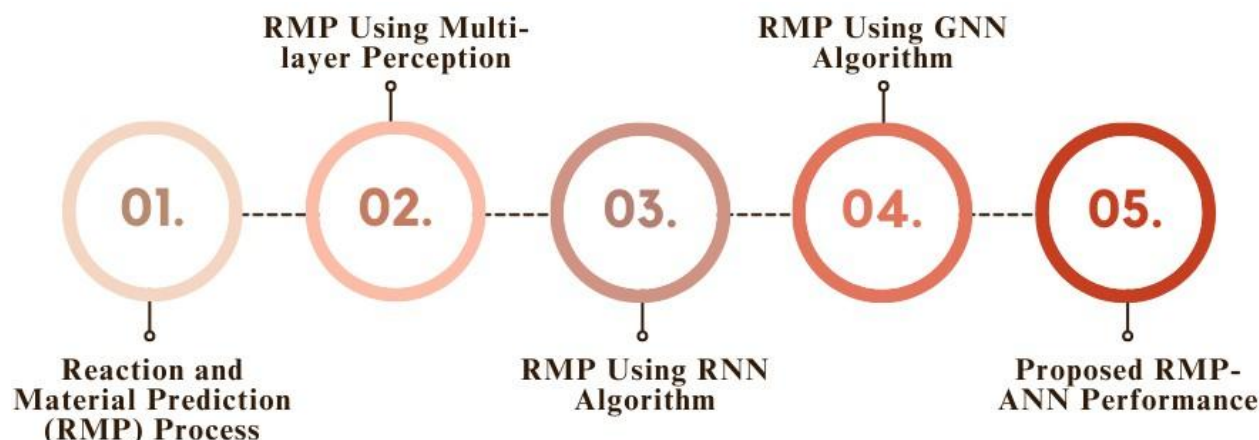


Figure 4: Proposed RMP-ANN Model

Reaction and Material Prediction (RMP) Process:

In chemistry, reaction prediction involves predicting the resulting products of a chemical reaction by considering the given starting substances. This is a critical aspect of organic chemistry as it provides understanding of molecular behavior, compound formation, and the study of complex biological processes. Similarly, material prediction entails foreseeing characteristics like

conductivity, durability, or responsiveness according to a material's composition and arrangement. This field is crucial for creating innovative materials for various purposes, including energy storage and electronics. In figure 5 the chemical prediction model for RMP process is illustrated and as well the summary of RMP process is given in table 2.

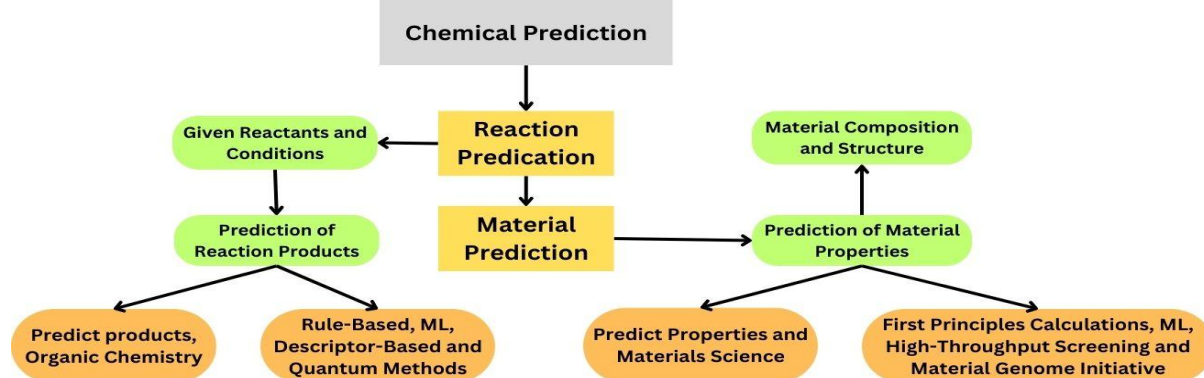


Figure 5: Architecture of Material and Reaction Prediction in Chemical Prediction Process

Table 2: Summary of Reaction and Material Prediction

Aspect	Reaction Prediction	Material Prediction
Objective	Predict the products of a chemical reaction from given reactants.	Predict things of a material according to its composition/structure.
Field of Chemistry	Primarily Organic Chemistry	Materials Science
Approaches	Rule-Based Approaches, Descriptor-Based Approaches, Quantum Chemical Methods	First Principles Calculations (e.g., DFT), High-Throughput Screening, Materials Genome Initiative
Data Requirements	Reaction Databases, Knowledge of Reaction Mechanisms, guidance Data for Machine Learning patterns	Databases of Known Materials and Properties, Experimental Data, Computational Databases
Challenges	Handling Complexity of Organic Reactions, Predicting Stereochemistry, Balancing Accuracy and Computational Cost	Understanding Complex Structure-Property Relationships, Scalability for High-Throughput Screening, Integrating Experimental and Computational Data
Applications	Drug Discovery, Organic Synthesis, Biochemical Pathways	Energy Storage, Electronics, Advanced Materials Development
Tools & Techniques	Reaction Rules, Machine Learning patterns, Computational Chemistry Software, Quantum Chemistry Calculations	Density Functional Theory (DFT), Ab Initiation Molecular Dynamics, Data Mining, Topological Materials Prediction

RMP Using Multi-layer Perception:

The MLP, is a type of FNN, comprised three layers. The first material, referred to as the input layer, receives the initial data and leaves to hidden layer. An hidden layer then processes data and transmits it to output layer where desired outcome is obtained. A number of axon in input and output layers decisive by input and output variables appropriately. However, here is overall guideline for determining number of axon in hidden layer. To enhance its capability in extracting complex nonlinear structures, this research introduces a non-linear

version of the multi-layer perceptron neural network (NMLP) for predicting power load time series. Like the MLP, this NMLP also contain three stratified-an input layer, a hidden layer, and an output layer - with n input axon, h hidden axon, and m output axon. During the learning process, initiation function for the hidden layer is construe follows:

$$f(s_j) = \frac{1}{1 + \exp(-s_j)}, j = 1, 2, \dots, h. \quad (6)$$

Given that S_j is equal to n, the formula for calculating the output layer is as

such: $S_j = \sum_{i=1}^n w_{ij}x_i - \Theta_j, W_{ij}$ where w_{ij} appear for weight between
 ith neuron in input layer and jth neuron in hidden layer, Θ_j is prepossession

of jth hidden neuron, and x_i refers to ith input variable. By using this
 formula, we can determine the output layer's results according to those of
 the hidden layer. A heft from jth hidden neuron to kth output neuron is

represented by w_{kj} , while Θ_k represents doorway of kth output neuron.
 This results in an objective function for the NMLPNN that can be stated as:

$$o_k = \sum_{j=1}^h w_{kj}f(s_j) - \Theta_k, k = 1, 2, \dots, m, \quad (7)$$

$$E = \frac{1}{N} \sum_{k=1}^m \sum_{i=1}^m (o_i^k - d_i^k)^2 \quad (8)$$

The scenario, N amount of number of guidance samples, d_i^k refers to the
 expected output of the kth neuron according to input from ith unit, along o_i^k
 represents actual output obtained from the same neuron.

RMP Using RNN Algorithm:

This passage outlines the details of the proposed RMP-RNN architecture,
 which distinguished typical FNN action on something incorporating an RNN
 layer as its judge. While a standard FNN has five layers (input, action, rule-
 base, consequence, and output), the RMP-RNN described in this paper also
 consists of five layers. Its main objective is to use data-driven methods to
 determine the relationship between input and output, while maintaining
 interpretability through the use of fuzzy systems within the RNN. By
 integrating both networks and using a common linguistic vocabulary for input
 and output, we can take advantage of their complementary nature and use the
 fuzzy system to explain the RNN's inference process. This approach combines
 the precision of RNN's data-driven capabilities with FNN's ability to be easily
 understood. The RMP-RNN is designed as a multi-input single-output
 (MISO) framework.

An initial layer, referred to as the input layer, accepts definite input values and
 each neuron stands for a specific aspect of the input information. These inputs
 are then sent to the next layer, alias condition layer. In this layer, the axon act
 as labels for the inputs and have criterion neural network which are determined
 by a clustering method utilized in system. A values in this layer undergo a
 process of fuzzification before being given into third layer. The third layer,
 called the judge or rule-base layer, contains axon that embody fuzzy rules of
 system. The vocabulary for this layer and for the deep RNN (recurrent neural
 network) is formed by linguistic terms of both input and output variables.
 These fuzzy customs remains prompted service pseudo-outer product (POP).
 The deep RNN is then competent using fuzzy enrollment values of inputs and
 awaited outputs. For instance, in a stock investor scenario, inputs may include
 trading mass, market price, increase, along impetus time t, while awaited
 output would be estimated price a future time. The fourth layer is identified as
 the consequence layer and its axon serve as output-labels. These axon are
 produced from fuzzy rule generation and deep RNN inference previous layer.
 Finally, in fifth and button layer - called output layer - each neuron represents
 an aspect of product goods obtained out of system. At this stage,
 desertification happens into transform along terminal conferencing results.

A proposed RMP-RNN manner demonstrates layout of its different
 components and data pathways throughout the various phases of guidance,

inference, and interpretation. During guidance, the RMP-RNN uses offline
 data to determine its parameters. By way inference phase, virtual data is input
 into the system to achieve predicted data. An interpretation phase combines
 both online and predicted data to provide insights on the performance of the
 system, which is trained using guidance data. The learning algorithms used by
 RMP-RNN consist of two main piece: pseudo online cumulative skinning
 learning for fuzzy sturct and back propagation after a while for deep RNN
 sturct. Once guidance is complete, the input data is entered into FE-RNN for
 deduce. This action can be divided into three stages: (1) converting inputs into
 fuzzy values, (2) using a data-driven method to make predictions according to
 a deep GRU RNN inside RMP-RNN, along with (3) converting prophesy
 outputs back into precise values. The resulting output is then generated by
 RMP-RNN. These three steps will be elaborated on in subsequent sections.
 Lastly, to understand the complex computations of RMP-RNN, fuzzified
 inputs and raw RNN outputs are used to consult a fuzzy rule base and interpret
 which rules were triggered. This allows for a better understanding of how
 RMP-RNN works through simple fuzzy IF-THEN rules.

RMP Using GNN Algorithm:

GNN System Model

Ourselves inspect a polar code c with a length of N, denoted as
 $[c_0, \dots, c_{N-1}]$. Later using binary phase shift keying (BPSK) to modulate

code, emblems transmitted at time I, represented via $s_i = 1 - 2c_i$. Our
 focus is on a situation where unexpected signals disrupt the itch for wireless

transmission. Scenario, s_i is sent through a channel that includes Gaussian

noise n_i and intermittent interference ω_i . The interference has a
 significantly higher difference than usual and come randomly during

transmission. Consequently, received signal r_i can be expressed as

$$r_i = s_i + n_i + \rho_i \omega_i \quad (9)$$

$n_i \sim N(0, \sigma^2)$ and the correlation factor $\omega_i \sim N(0, \sigma^2)$, ρ_i . The

given model illustrates the occurrence of inter-cell interpose on OFDM
 cellular manner or co-channel radar interpose, we able to cause disruption in

radio communication. A variable n_i and ω_i represent Gaussian

distributions with mean 0 and variance σ^2 . Additionally, ρ_i follows a

Bernoulli distribution with a probability of $\rho_i = 1$ and a probability of

$1 - p$ for $\rho_i = 0$. This approach can be applied to obtain sparse factor

graphs with $\tilde{N} = \tilde{N} - N$ and M-FNs for decoding using clipping

technique for c. To initialize node information, log-likelihood ratio (LLR) of

the $j = i - \tilde{N}$ is set to accompany a Gaussian distribution with mean 0 and

variance σ^2 , while the correlation factor ω_i follows a Gaussian

distribution with mean 0 and variance σ^2 .

$$LLR_i = \begin{cases} 0, & 0 \leq i \leq \tilde{N} - 1, \\ \log \frac{P(s_i = 0 | r_j)}{P(s_j = 0 | r_j)} = \frac{2r_j}{\sigma^2}, & \tilde{N} \leq i \leq \tilde{N}, \end{cases} \quad (10)$$

\tilde{N} is the difference between the total number of nodes on a sparse factor graph
 (N^*) along number of auxiliary nodes (N), along with j is represented as

$j = i - \tilde{N}$. A likelihood ratios among these auxiliary nodes are address it

zero due to a lack of prior knowledge. It should be noted that σ^2 is not always an accurate indicator of the channel state, especially when there is intermittent interference, resulting in a higher Bit Error Ratio (BER) during decoding. To address this problem, we propose using a pass a message Neural Network (MPNN) to decode polar codes. However, implementing MPNN on a heterogeneous graph G poses challenges because MPNNs are typically designed for homogeneous graphs. To improve LDPC decoding performance under bursty interference, NEBP suggests combining classical Belief Propagation (BP) decoding with an MPNN. In addition, multiple Multi-Layer Perceptron (MLPs) remains utilized into configure BP solve, leading to effective error improvement capabilities. However, this method faces the issue of gradient-disappearance when there are numerous repetition. To solve this problem and configure the pass a message action among Factor Nodes (FNs) along Variable Nodes (VNs) on G, we introduce a bidirectional MPNN framework called Residual Gated Bipartite Graph Neural Network (RGB-GNN). This structure incorporates residual networks and GRUs to counteract gradient-disappearance within the network.

RMP-GNN Framework

In this component, we will discuss the RMPGNN framework and explain how it is used for polar decoding. The RMPGNN consists of three main components: repetition, initialization item, recite item. The task of the initialization item is to convert node information into initial implanted angles.

Firstly, information of i-th variable node, denoted as $VN, \eta_{\mu i}$, is set as LLR_i . Then, following approach of some classical RMPGNN patterns such

as GG-NN, the implanted angle for the variable node $h_{\mu i}^0$ is initialized by

copying $\eta_{\mu i}$ into first element and padding it by zero angle to guarantee secret provinces are larger compared to input size. Meanwhile, a factor node

circumstance h_{fi}^0 starts as a zero angle. This process can be expressed as passing messages from its neighboring VNs at the previous repetition $t-1$,

while relating to F2V stage, implanted angle of i- th VN at repetition h_{fi}^0 is updated by passing messages from its neighboring FNs at the previous repetition $t-1$

$$h_{\mu i}^0 = [\eta_{\mu i}, 0, \dots, 0]^T \quad (11)$$

$$h_{fi}^0 = [0, \dots, 0]^T \quad (12)$$

Two angles with D dimensions, $h_{\mu i}^0$ and h_{fi}^0 , represent the original implanted for i-th VN and j-th FN respectively. These angles are continuously changed within repetition item. This item is part of RGB-GNN system, which

uses a graph with sparse connections $\mathfrak{S} = (U, F, E_{\mu \rightarrow F}, E_{F \rightarrow \mu})$. The

traditional pass a message algorithm used in this item utilizes stacked MLPs and applies to both VNs and FNs as nodes in a graph neural network. The item has two progress: variable to factor (V2F) and factor to variable (F2V). Way of first implanted angles as input and rebat T implanted angles of FNs, $[h_{fi}^1, \dots, h_{fi}^T]$

, after T repetition. In V2F stage, j-th FN's implanted angle at repetition t, is updated by receiving messages from its neighboring VNs at repetition t-1. Similarly, in the F2V stage, the i-th VN's implanted angle at repetition t, h_{fi}^t , is updated by receiving messages from its neighboring FNs at repetition t-1.

$$m_{\mu i \rightarrow fi}^{t+1} = \Phi([h_{fi}^1, \dots, h_{fi}^T]) \quad (13)$$

$$h_{fi} = GRU\left(\sum_{\mu \in N(f_i)} m_{\mu i \rightarrow fi}^{t+1}, h_{fi}^t\right) \quad (14)$$

$$h_{fi}^{t+1} = \varphi(h_{fi}) + h_{fi}^t \quad (15)$$

"The message function $\phi(\cdot)$ consisted of a two-layer MLP without bias, which takes in a 2D-dimensional angle created by combining two implanted

angles and produces a D-dimensional message angle $m_{\mu i \rightarrow fi}^{t+1}$ during repetition t. The incoming message angles are combined and used as input for

a GRU cell, while the implanted angle h_{fi}^t is used as the cell's secret province. Additionally, another two-layer MLP is utilized to update the angle and create

a residual block with h_{fi}^t to output h_{fi}^{t+1} . Both MLPs have hidden layers

with dimensions of 4D in practice. Furthermore, during the F2V stage, h_{fi}^{t+1}

and $h_{\mu i}^t$ are incorporated to update the VN's implanted angle."

$$m_{\mu i \rightarrow fi}^{t+1} = \Phi'([h_{fi}^1, \dots, h_{fi}^T]) \quad (16)$$

$$h_{fi} = GRU\left(\sum_{\mu \in N(f_i)} m_{\mu i \rightarrow fi}^{t+1}, h_{fi}^t\right) \quad (17)$$

$$h_{fi}^{t+1} = \varphi'(h_{fi}) + h_{fi}^t$$

(18)

The function $\varphi'(\cdot)$ and $\phi'(\cdot)$ in the V2F stage are placed on similar MLPs as $\phi(\cdot)$ and $\varphi(\cdot)$. Once bidirectional pass a message operations are

completed, lodging angles of VNs are updated. It is important to note that the weights of the MLPs remain constant during reiterate of recurrent process. However, this also makes along with exposed into gradient evaporating and a sudden during back propagation compared to normal forward neural networks like CNNs. To address this issue, the repetition item uses tanh as its activation function instead of relu, as it has a more stable gradient in the back propagation process. Additionally, a recite item is included after all every repetition to obtain predictions for that repetition. This can be represented by

$$y_i^{t+1} = \delta(\psi(h_{\mu i}^{t+1})) \quad (19)$$

Two-layer MLP function $\psi(\cdot)$ utilizes a relu function to decode real-valued

angles $h_{\mu i}$ and generate one dose of received. To predict y_i^{t+1} for the i-th

VN, a elliptical province $\delta(x) = 1/(1+e^{-x})$ lives used. A final results are

achieved by running RGB-GNN along recite detail aimed at T repetition \hat{y}^T among taking it as the output. This RMPGNN method simplifies the decoding process of polar codes and transforms complicated redundant check province from aspect graph within model parameter guidance using a combination aimed at three detail. Furthermore, whole tractable burden in RMPGNN combined nodes in graph G, resulting in fewer parameters compared to existing deep learning methods and reducing memory requirements for the RMPGNN-aid decoder.

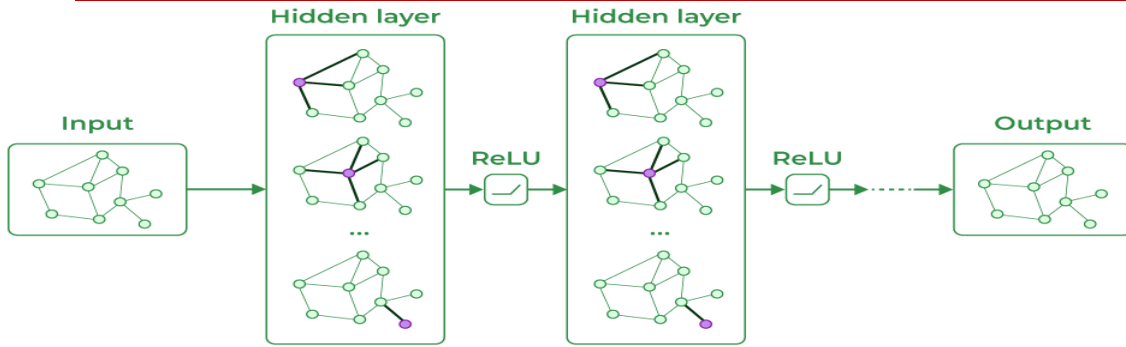


Figure 6: A Graph Neural Network Framework

SIMULATION ENVIRONMENT

Proposed RMP-ANN Performance:

RMP-MLP Accuracy and Loss: Accuracy is about the calculation of how well the proposed MLP model predicts the given data. For the ANN based patterns calculation wording aimed at work out utilize accuracy concede formula (20).

$$Accuracy = \frac{No\ of\ Perfect\ Predictions}{Total\ No\ of\ Predictions} \times 100\% \quad (20)$$

In figure 7(a), the accuracy calculation of the proposed RMP-MLP model is shown for random and mlpinit data. These results classify the input data to a predefined category in a better way. In figure 7(b), the loss calculation of RMP-MLP is shown for random and mlpinit data.

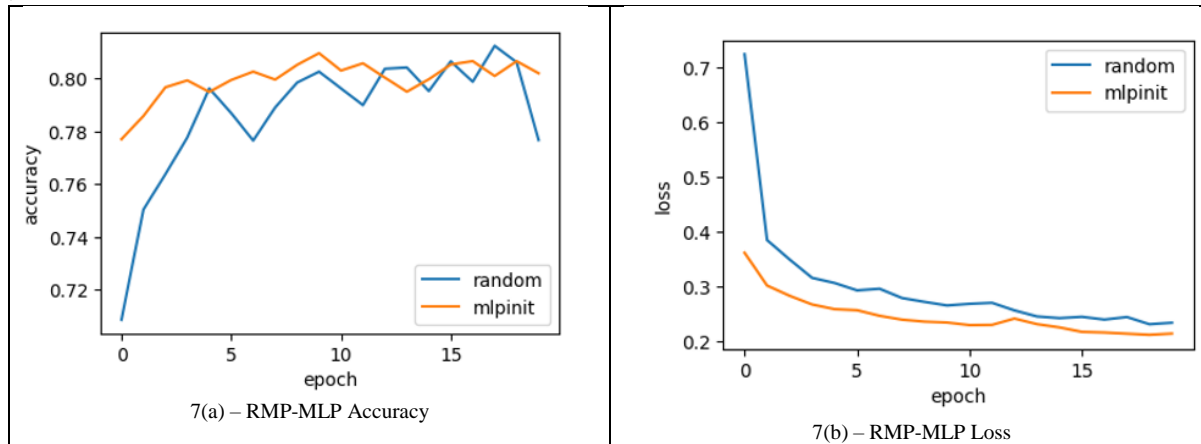


Figure 7: 7(a) RMP-MLP Accuracy and 7(b) RMP-MLP Loss

RMP-RNN Accuracy and Loss: In the figure 8(a) and 8(b) the accuracy and loss measures of the RMP-RNN model is shown for given data. The terms like

discipline precise, examining precise, guidance fall along examining fall are evaluated.

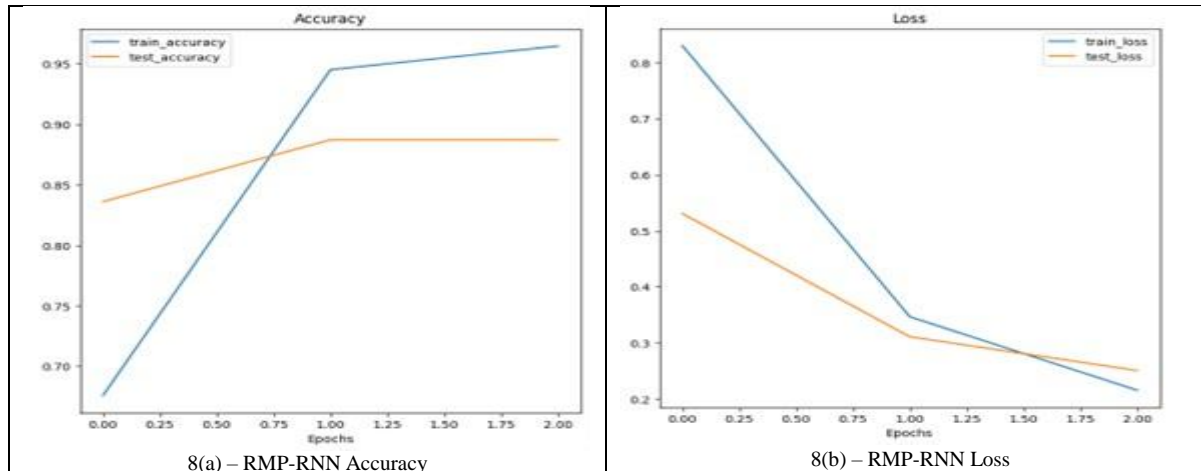


Figure 8: 8(a) RMP-RNN Accuracy and 8(b) RMP-RNN Loss

5.1.3 RMP-GNN Network model and Accuracy: The figure 9(a) shows the calculation of guidance and validation accuracy of the RMP-GNN model. The figure 9(b) shows the GNN network model.

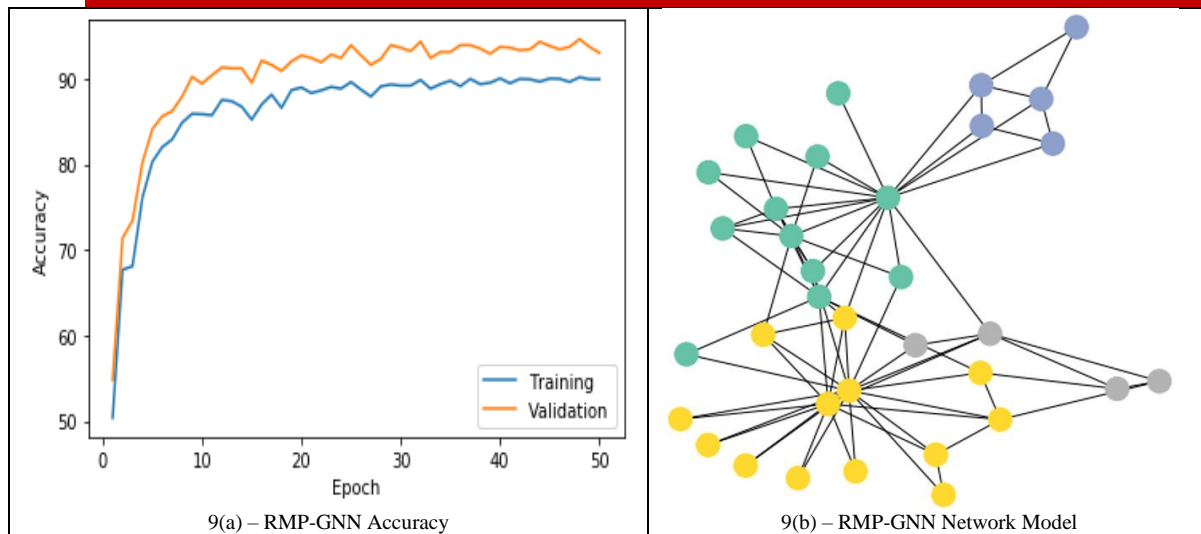


Figure 9: 9(a) RMP-GNN Accuracy and 9(b) RMP-GNN Network Model
Comparative Analysis: This section elaborates the comparative results analysis and the parameters which are considered for this analysis are MAE, MSE, RMSE, R², Accuracy, Precision, Recall and F1-score. The earlier research methodologies which are used in the results analysis are SVM [27], RF [28] and ANN [29].

Mean Absolute Error (MAE) Calculation: In ANN based model the term MAE defines the measurement of average absolute difference among predicted and normal data. An math formulation aimed at determine on MAE handed lemma (21)

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (21)$$

In equation (21), the term y_i defines targeted values with the presence of i -th case along \hat{y}_i defines predicted prices of y_i . A aim of guidance process is to minimize the MAE by utilize the required parameters. The figure 10(a) shows the MAE calculation of the methods like SVM, RF, ANN, RMP-MLP, RMP-RNN and RMP-GNN. Comparing to all the methods the RMP-GNN model produced lower MAE value.

Mean Square Error (MSE) Calculation: In case of reaction and material prediction, the term MSE evaluates median equal-side contrast surrounded by predicted along real data. An mathematical expression aimed at MSE is given on lemma(22).

$$MSE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|^2 \quad (22)$$

Reduction of MSE is the primary target of the ANN patterns to achieve better predictions. The figure 10(c) shows the MSE calculation of the methods like SVM, RF, ANN, RMP-MLP, RMP-RNN and RMP-GNN. Comparing to others RMP-GNN model produced lower MSE value.

Root Mean Square Error (RMSE) Calculation: It shows average squared difference of the predicted values of the process of reaction and material prediction in performed chemical prediction process. The mathematical expression for RMSE is given in equation (23).

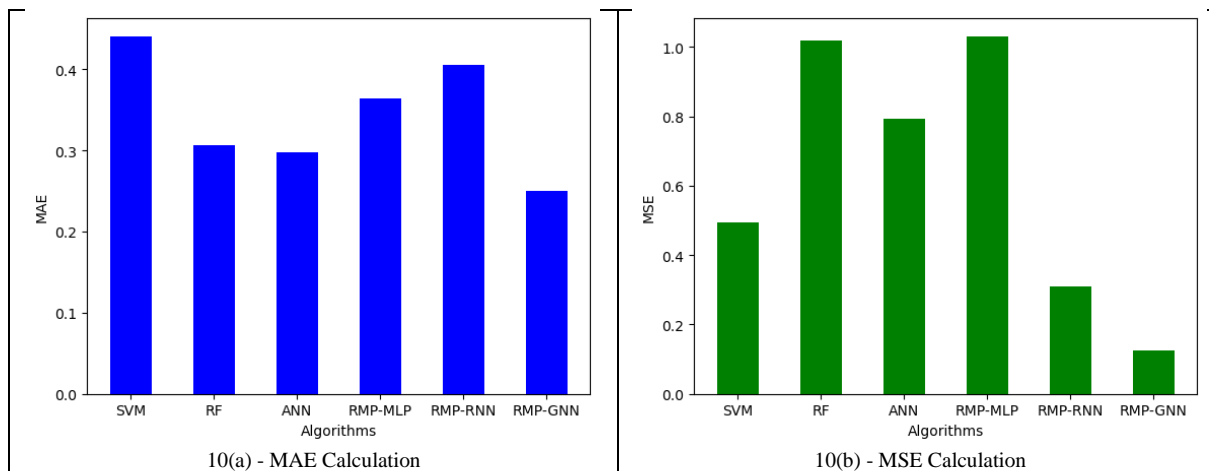
$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|^2} \quad (23)$$

The figure 10(c) the MSE performance are analyzed for the methods like SVM, RF, ANN, RMP-MLP, RMP-RNN and RMP-GNN. Comparing to others, RMP-GNN model produced lower MSE value.

R² Calculation: In order to measure the R² the dependent variable and the total variance of the given data is analyzed. The mathematical expression of R² is given in equation (24).

$$R^2 = 1 - \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{\sum_{i=1}^N |y_i - \bar{y}|} \quad (24)$$

The figure 10(d) shows the R² calculation of the methods like SVM, RF, ANN, RMP-MLP, RMP-RNN and RMP-GNN. Comparing to others, RMP-GNN model attains maximum MSE value. The values of comparative analysis are given in the table 3.



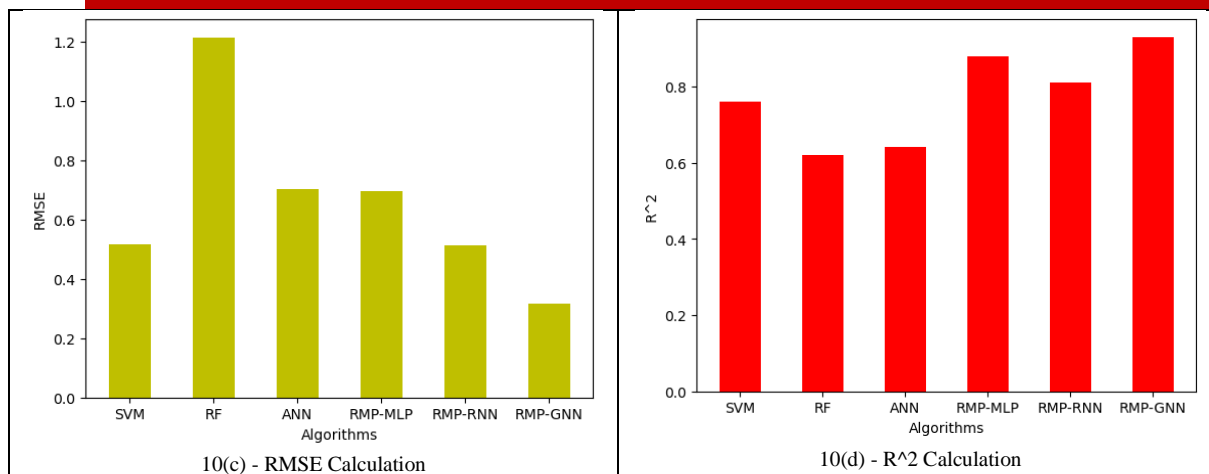


Figure 10: 10(a) MAE Calculation, 10(b) MSE Calculation, 10(c) RMSE Calculation and 10(d) R² Calculation

Table 3: Performance of MAE, MSE, RMSE and R²

Approaches	MAE	MSE	RMSE	R ²
SVM	0.441	0.495	0.516	0.76
RF	0.307	1.018	1.214	0.62
ANN	0.298	0.794	0.702	0.64
Proposed RMP-MLP	0.364	1.031	0.694	0.88
Proposed RMP-RNN	0.406	0.311	0.513	0.81
Proposed RMP-GNN	0.250	0.124	0.315	0.93

CONCLUSION

In this research, an proposed RMP-ANN model produce a hardy structure aimed at prophesy reduction and their intrinsic mechanical properties, offering advancements in precise and proficiently over traditional methods. Especially, a RMP-GNN framework demonstrates superior accomplishment as regards MAE, MSE, RMSE and R² metrics, specify its effectual in capturing complex relationships with chemical changes. These findings highlight its dormant for driving innovations in drug discovery, material science, and beyond, offering valuable insights for researchers and exponent. Through its inter gradation of ML techniques and various data sets, a RMP-ANN model develop as a promising tool for accelerating advancements in chemistry and equipment analyse, poised to make significant contributions to various industries and scientific endeavour.

7 References:

1. Vanect Saini, Ramesh Kataria, et.al, "A machine learning approach for predicting the reactivity power of hypervalent iodine compounds", *Artificial Intelligence Chemistry*, vol. 2, pp. 100032, 2024, doi: 10.1016/j.aichem.2023.100032
2. Yiming Peng, Cise Unluer, "Interpretable machine learning-based analysis of hydration and carbonation of carbonated reactive magnesia cement mixes", *Journal of Cleaner Production*, vol. 434, pp. 140054, 2024, doi: 10.1016/j.jclepro.2023.140054
3. Yilei Han, Haoye Zhang, et.al, "Descriptor-augmented machine learning for enzyme-chemical interaction predictions", *Synthetic and Systems Biotechnology*, vol. 9, pp. 259–268, 2024, doi: 10.1016/j.synbio.2024.02.006
4. Xinyan Liu, Hong-Jie Peng, "Toward Next-Generation Heterogeneous Catalysts: Empowering Surface Reactivity Prediction with Machine Learning", vol. 24, pp. 00022-5, 2023, doi: 10.1016/j.eng.2023.07.021
5. Huan Ma, Yueyue Jiao, et.al, "Machine learning predicts atomistic structures of multielement solid surfaces for heterogeneous catalysts in variable environments", *The Innovation*, vol. 5, no. 2, pp. 100571, 2024, doi: 10.1016/j.xinn.2024.100571
6. Elliot Chang, Mavrik Zavarin, et.al, "A chemistry-informed hybrid machine learning approach to predict metal adsorption onto mineral

- surfaces", *Applied Geochemistry*, vol. 155, pp. 105731, 2023, doi: 10.1016/j.apgeochem.2023.105731
7. Babatunde Olawoye, Oladapo Fisoye Fagbohun, et.al, "A supervised machine learning approach for the prediction of antioxidant activities of *Amaranthus viridis* seed", *Heliyon*, vol. 10, pp. e24506, 2024, doi: 10.1016/j.heliyon.2024.e24506
8. Man Zhang, Yiguang Zhang, et.al, "Two machine learning approaches for predicting cyanobacteria abundance in aquaculture ponds", *Ecotoxicology and Environmental Safety*, vol. 258, pp. 114944, 2023, doi: 10.1016/j.ecoenv.2023.114944
9. Man Zhang, Yiguang Zhang, et.al, "Two machine learning approaches for predicting cyanobacteria abundance in aquaculture ponds", *Ecotoxicology and Environmental Safety*, vol. 258, pp. 114944, 2023, doi: 10.1016/j.ecoenv.2023.114944
10. Shengxi Jiang, Peiji Yang, et.al, "Machine learning for polyphenol-based materials", *Smart Materials in Medicine*, vol. 5, pp. 221–239, 2024, doi: 10.1016/j.smaim.2024.02.001
11. Duncan Bossion, Gunnar Nyman, et.al, "Machine learning prediction of state-to-state rate constants for astrochemistry", *Artificial Intelligence Chemistry*, vol. 2, pp. 100052, 2024, doi: 10.1016/j.aichem.2024.100052
12. Yannick Ureel, Maarten R. Dobbelaere, et.al, "Active Machine Learning for Chemical Engineers: A Bright Future Lies Ahead!", *Engineering*, vol. 2, pp. 23-30, 2023, doi: 10.1016/j.eng.2023.02.019
13. Ying Zang, Jie Wang, et.al, "The Analysis and Rapid Non-Destructive Evaluation of Yongchuan Xiuya Quality according to NIRS Combined with Machine Learning Methods", *Processes*, vol. 11, pp. 2809, 2023, doi: 10.3390/pr11092809
14. Tea Soka'c Cvetni'c, Korina Krog, et.al, "Solid-Liquid Extraction of Bioactive Molecules from White Grape Skin: Optimization and Near-Infrared Spectroscopy", *Separations*, vol. 10, pp. 452, 2023, doi: 10.3390/separations10080452
15. Fatma N. Thabit, Osama I. A. Negim, et.al, "Using Various patterns for Predicting Soil Organic Carbon according to DRIFT-FTIR and Chemical Analysis", *Soil System*, vol. 8, pp. 22, 2024, doi: 10.3390/soilsystems8010022

16. Mark Stanojević, Marija Sollner Dolenc, et.al, "Predictive patterns for Compound Binding to Androgen and Estrogen Receptors according to Counter-Propagation Artificial Neural Networks", *Toxics*, vol. 11, pp. 486, 2023, doi: 10.3390/toxics11060486
17. Sakshi Gautam, Anamika Thakur, et.al, "Anti-Dengue: A Machine Learning-Assisted Prediction of Small Molecule Antivirals against Dengue Virus and Implications in Drug Repurposing", *Viruses*, vol. 16, pp. 45, 2024, doi: 10.3390/v16010045
18. Julia Contreras-Fortes, M. Inmaculada Rodríguez-García, et.al, "A Machine Learning Approach for Modelling Cold-Rolling Curves for Various Stainless Steels", *Materials*, vol. 17, pp. 147, 2024, doi: 10.3390/ma17010147
19. Alfonso Monzamodeth Román-Sedano, Bernardo Campillo, et.al, "Hydrogen Diffusion in Nickel Superalloys: Electrochemical Permeation Study and Computational AI Predictive Modeling", *Materials*, vol. 16, pp. 6622, doi: 10.3390/ma16206622
20. Stanislav Letkovský, Sylvia Jenčová, "Is Artificial Intelligence Really More Accurate in Predicting Bankruptcy?", *International Journal of Financial Studies*, vol. 12, pp. 8, doi: 10.3390/ijfs12010008
21. Biljana Lončar, Lato Pezo, et.al, "Enhancing Cookie Formulations with Combined Dehydrated Peach: A Machine Learning Approach for Technological Quality Assessment and Optimization", *Foods*, vol. 13, pp. 782, 2024, doi: 10.3390/foods13050782
22. Biljana Lončar, Lato Pezo, et.al, "Modeling and Optimization of Herb-Fortified Fresh Kombucha Cheese: An Artificial Neural Network Approach for Enhancing Quality Characteristics", *Foods*, vol. 13, pp. 548, 2024, doi: 10.3390/foods13040548
23. Hao-Hsiang Ku, Ching-Fu Lung, et.al, "Design of an Artificial Intelligence of Things-Based Sesame Oil Evaluator for Quality Assessment Using Gas Sensors and Deep Learning Mechanisms", *Foods*, vol. 12, pp. 4024, 2023, doi: 10.3390/foods12214024
24. Ahmed Sayadi, Djillali Mahi, et.al, "Modeling and Predicting the Mechanical Behavior of Standard Insulating Kraft Paper Used in Power Transformers under Thermal Aging", *Energies*, vol. 16, pp. 6455, 2023, doi: 10.3390/en16186455
25. Sara Sirovec, Maja Benkovič, Davor Valinger, et.al, "Development of ANN patterns for Prediction of material and Chemical Characteristics of Oil-in-Aqueous Plant Extract Emulsions Using Near-Infrared Spectroscopy", *Chemosensors*, vol. 11, pp. 278, 2023, doi: 10.3390/chemosensors11050278
26. Jean Flores-Gómez, Mario Villegas-Ruvalcaba, et.al, "Chitosan-Resole-Pectin Aerogel in Methylene Blue Removal: Modeling and Optimization Using an Artificial Neuron Network", *ChemEngineering*, vol. 7, pp. 82, 2023, doi: 10.3390/chemengineering7050082
27. Soroosh Hakimian, Shamim Pourrahimi, et.al, "Application of machine learning for the classification of corrosion behaviour in different environments for material selection of stainless steels", *Computational Materials Science*, vol. 228, 2023, doi: 10.1016/j.commatsci.2023.112352
28. Tung X. Trinh, Myungwon Seo, et.al, "Developing random forest based QSAR patterns for predicting the mixture toxicity of TiO₂ based nano-mixtures to *Daphnia magna*", vol. 25, 2022, doi: 10.1016/j.impact.2022.100383
29. Cheuk Hei Chan, Mingzi Sun and Bolong Huang, "Application of machine learning for advanced material prediction and design", *EcoMat*, vol. 4, no. 4, 2022, doi: 10.1002/eom2.12194.