Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Optimizing hybrid neural networks for precise COVID-19 mRNA vaccine degradation prediction

Hwai Ing Soon [1, 2], Azian Azamimi Abdullah [1, 3, *], Hiromitsu Nishizaki [2], Mohd Yusoff Mashor [1], Latifah Munirah Kamarudin [1, 4], Zeti-Azura Mohamed-Hussein [5, 6], Zeehaida Mohamed [7], Wei Chern Ang [8, 9]

[1]*Faculty of Electronic Engineering and Technology, Universiti Malaysia Perlis (UniMAP), Arau, Perlis, Malaysia*
[2]*Integrated Graduate School of Medicine, Engineering and Agricultural Science, University of Yamanashi, Kofu, Yamanashi, Japan*
[3]*Medical Devices and Life Sciences Cluster, Sport Engineering Research Centre, Centre of Excellence (SERC), UniMAP, Arau, Perlis, Malaysia*
[4]*Advanced Sensor Technology, Centre of Excellence (CEASTech), Universiti Malaysia Perlis (UniMAP), Arau, Perlis, Malaysia*
[5]*Department of Applied Physics, Faculty of Science and Technology, Universiti Kebangsaan Malaysia (UKM), Bangi, Selangor, Malaysia*
[6]*UKM Medical Molecular Biology Institute, Universiti Kebangsaan Malaysia (UKM), Jalan Yaacob Latif, 56000 Cheras, Kuala Lumpur, Malaysia*
[7]*Department of Medical Microbiology Parasitology, School of Medical Sciences, Universiti Sains Malaysia (USM), 16150 Kubang Kerian Kelantan, Malaysia*
[8]*Clinical Research Centre (CRC), Hospital Tuanku Fauziah (HTF), Ministry of Health Malaysia, Kangar, 01000, Perlis, Malaysia*
[9]*Department of Pharmacy, Hospital Tuanku Fauziah (HTF), Ministry of Health Malaysia, Kangar, 01000, Perlis, Malaysia*

ARTICLE INFO

Article history:
Received 21 February 2024
Received in revised form
1 July 2024
Accepted 2 July 2024

Keywords:
Neural network
Hybridization
Hybridizing sequence
mRNA vaccines

ABSTRACT

Conventional hybrid models often miss an essential factor that can lead to less effective performance: intrinsic sequence dependence when combining various neural network (NN) architectures. This study addresses this issue by highlighting the importance of sequence hybridization in NN architecture integration, aiming to improve model effectiveness. It combines NN layers—dense, long short-term memory (LSTM), and gated recurrent unit (GRU)—using the Keras Sequential API for defining the architecture. To provide better context, bidirectional LSTM (BiLSTM) and bidirectional GRU (BiGRU) replace their unidirectional counterparts, enhancing the models through bidirectional structures. Out of 25 NN models tested, 18 four-layer hybrid NN models consist of one-quarter dense layer and the rest BiLSTM and BiGRU layers. These hybrid NN models undergo supervised learning regression analysis, with mean column-wise root mean square error (MCRMSE) as the performance metric. The results show that each hybrid NN model produces unique outcomes based on its specific hybrid sequence. The Hybrid_LGSS model performs better than existing three-layer BiLSTM networks in predictive accuracy and shows lower overfitting (MCRMSEs of 0.0749 and 0.0767 for training and validation, respectively). This indicates that the optimal hybridization sequence is crucial for achieving a balance between performance and simplicity. In summary, this research could help vaccinologists develop better mRNA vaccines and provide data analysts with new insights for improvement.

1. Introduction

Rather than relying on the rigid structure of traditional neural networks (NNs), recurrent neural networks (RNNs) augmented the traditional layered structure of NN to introduce specialized layers to process sequential data efficiently, offering a dynamic architecture ideal for capturing complex relationships within data (Kaur and Mohta, 2019; Mittal et al., 2023). The heart lies in recurrent layers, often implemented as long short-term memory (LSTM) or gated recurrent unit (GRU) units (Cahuantzi et al., 2023; Mittal et al., 2023). These layers possess an exceptional ability to retain the memory of past information, which is vital for capturing dependencies and patterns within sequential data that might go unnoticed by static

* Corresponding Author.
Email Address: azamimi@unimap.edu.my (A. A. Abdullah)
https://doi.org/10.21833/ijaas.2024.07.011
Corresponding author's ORCID profile:
https://orcid.org/0000-0001-5851-7705

models (Batur and Abualigah, 2021; Mittal et al., 2023). In this mRNA vaccine degradation prediction research, integrating dense layers alongside recurrent layers is practiced, granting hybrid NN models the capability to execute non-linear transformations. This strategic combination facilitates the adept capture of intricate relationships within the dataset, enhancing the models' capacity for nuanced pattern recognition and representation. Moreover, in the pursuit of optimal precision, this study explores the impact of hybridization sequences on the performance of hybrid models.

Recently, the predictive study of the degradation rate of mRNA vaccines has received a certain amount of attention in concomitance to the COVID-19 outbreak (Imran et al., 2020; Singhal, 2020; Ing et al., 2021a; 2022; 2021b; Muneer et al., 2021; Qaid et al., 2021; Wang, 2021). Data analysts have proposed using artificial intelligence (AI) approaches to reduce the complexity and duration of laboratory experiments. COVID-19 is a pandemic that killed millions in less than three years (Qaid et al., 2021; Cattaneo et al., 2022), leaving a famous record in the history of epidemiology and loimology. Antiviral drugs are being developed in parallel with the evolution of SARS-CoV-2. A few medications, i.e., tocilizumab and baricitinib, recently received official acceptance from the Food and Drug Administration (FDA) for clinical use (Banerjee et al., 2023). Nevertheless, there is more to be explored over time about their hidden side effects.

In this pandemic, vaccination is one of the most promising prophylactic measures. As with medicines, no vaccine can be guaranteed to be safe. Inflammation and fever are common post-vaccination symptoms. Viral vector vaccines are also likely to cause thrombosis (Okamura and Ebina, 2021; Cines and Greinacher, 2023), while live attenuated and inactivated vaccines are suspected of having the potential to restore their lively nature. In addition to that, there is no denying that some vaccines, such as DNA vaccines, are known to pose a risk of genetic mutation (Okamura and Ebina, 2021). Any vaccination carries the risk of side effects, and the degree of adversity can be a factor in the classification of vaccines. Of all vaccines, the mRNA vaccine has the lowest risk of side effects; therefore, it is widely used for vaccination.

mRNA vaccines are vaccines modified from mRNAs. mRNAs are innocuous, but as an advantage, they are self-adjuvanted (Xu et al., 2020; Okamura and Ebina, 2021) and have the most balanced and desirable immunological properties. In addition to that, the cellular and humoral immunities delivered by mRNA vaccines are not constrained by the restriction imposed by major histocompatibility complex (MHC) haplotyping, according to in situ protein expression (Okamura and Ebina, 2021; Olugbenga et al., 2022). The MHC is a group of genes that encode cell surface proteins. MHCs help an organism's immune system mount a defense by inducing immunological responses through the recognition of the invasion of foreign substances. The MHC is also recognized as a human leukocyte antigen (HLA) in the Homo sapiens species.

The mRNA is a potential vector that ensures biosecurity since it does not interact with the genome and merely serves as a transient information carrier (Okamura and Ebina, 2021; Xiao et al., 2022). Maximum flexibility can be offered by mRNA, as mRNA may express practically any protein without modifying it throughout the synthesis process. mRNA lends itself to the promotion of large-scale production of mRNA vaccines for the benefit of society, together with its excellent safety profile.

Compared to drugs and other vaccines, mRNA vaccines are the vaccine of choice for vaccinators owing to their safety profile and high tolerability. But unfortunately, history has shown that mRNA vaccines are susceptible to degradation, regardless of the era or the virus they target. mRNA is a subtype of RNA, and mRNA vaccines originate from mRNA. Therefore, the innate instability characteristics were inherited from the unstable RNA that carries an extra hydroxyl group in its structure, making mRNA vaccines more sensitive to environmental conditions than other vaccines. A sightless error or change in environmental conditions during the delivery and storage of mRNA vaccines can reduce their shelf life from months to hours. An approximate estimate of the RNA half-life and the shelf life of mRNA vaccines is provided in Table 1 and Table 2, respectively, for better comprehension.

Table 1: The estimated half-life of RNA under example-stimulated conditions (Wayment-Steele et al., 2021)

Condition	T (°C)[1]	pH	Mg^{2+} (mM)[2]	Half-life
Cold-chain	5	7.4	0	2.5 years
pKa shift	5	9.4	0	10 days
Room temperature[3]	25	7.4	0	5 days
Presence of Magnesium ion (Mg^{2+})	25	7.4	14	2.016 hours

[1]T = temperature, in degrees Celsius; [2]Magnesium ion concentration measured in millimolar; [3]Standard temperature defined by the International Union of Pure and Applied Chemistry (IUPAC) (Giełzak et al., 2023)

Table 2: Estimated shelf life of COVID-19 mRNA vaccine under example-stimulated conditions (Crommelin et al., 2021; Wang, 2021)

mRNA vaccine	pH	T (°C)[1]	Shelf life
CureVac	No data	≤ -60[1]	3 months
		2–8	3 months
		25	24 hours
Moderna	7–8	-20[1]	6 months
		2–8	30 days
		25	12 hours
Pfizer-BioNTech	7–8	-80[1] to -60[1]	6 months
		2–8	5 days
		25	2 hours

[1]Temperature of the frozen state of the respective vaccine

When stability is a concern, RNA's chemical structure is always a drawback, and vaccinations are no exception. There is a risk associated with the instability of mRNA vaccines. However, the risks posed by other vaccines were even more troubling than this drawback of mRNA vaccines. The instability of vaccines has been one of the inducing factors that result in alterations in the vaccines' potency, which could sacrifice millions (Dumpa et al., 2019). Despite vaccinologists' best efforts to extend the shelf life of

mRNA vaccines, several flaws persist. As things stand, learning the rates of vaccine degradation is currently the best course of action. In addition to keeping the fatality rate to a minimum, being knowledgeable about vaccination degradation rates can hasten the creation of vaccines with greater effectiveness, hence contributing to the purpose of this research in predicting the degradation rate of mRNA vaccines.

To clarify, the issues addressed by this research include the failure of practitioners to consider the order in which NN layers are hybridized. The architecture of neural network layers, also referred to as the arrangement and configuration of individual layers, significantly impacts the overall performance and complexity of the model. Different approaches to layer structuring, such as sequential stacking, wrapped stacking, wrapped hybridizing, or sequential hybridizing, offer varying advantages and trade-offs. However, practitioners hardly have this concern centralized while architecting NN layers. Therefore, it was raised and studied in this research.

The objectives of this study are as follows. Firstly, the strategic positioning of dense layers in the hybrid model structure is suggested to study their varying roles and impacts on a hybrid NN model's performance. Secondly, beyond optimizing the NN's specific layers, this research investigates whether varying the architectural approach while maintaining the same core layers can significantly impact overall performance. Last but not least, this research is put into practice to introduce a hybrid modeling approach that promotes lower prediction errors while reducing computational time and overfitting.

To achieve the objectives, 25 NN models were introduced: 20 models, comprising two stacked NN models and eighteen 4-layer hybrid NN models, followed a sequential architecture, while the remaining five models, including two stacked NN models and three 3-layer hybrid NN models, adopted a wrapped approach. NN layers, including dense, BiLSTM, and BiGRU, are engaged in the model development. Moreover, to study the impact of hybridizing NN layer sequences within a hybrid model on its predictive performance, the order of individual layers within the hybrid architecture is randomly shuffled, generating a diverse population of unique hybrid models. This randomized method allows systematic exploration of the potential benefits and drawbacks of different layer sequencing arrangements and gains valuable insights into their influence on prediction accuracy.

The results, in terms of mean column-wise root mean square error (MCRMSE), confirmed that integrating dense layers into a hybrid NN model not only serves to reduce complexity and mitigate overfitting but also strategically enhances predictive performance. In addition, the result demonstrated that the structuring approach will affect performance even if the underlying layers remain constant. Therefore, careful consideration should be given to the selection and configuration of layer structures, balancing desired performance gains with optimal model complexity and potential computational demands. The strong performance of structured models, particularly in terms of reduced prediction errors and increased computational efficiency, validated the effectiveness of the proposed approach for developing hybrid models, particularly Hybrid_LGSS, in mRNA vaccine degradation prediction. Briefly, the findings of this research contributed to:

- Dense layer incorporation: Results confirmed that strategically incorporating dense layers boosts feature extraction and overall performance, advocating their inclusion at pivotal intersections to unveil hidden patterns and relationships.
- Hybridizing sequence on hybrid model performance: Framework revealed that specific hybridizing sequences of NN layers demonstrably influenced prediction accuracy, highlighting the importance of sequence optimization. This finding promotes better understanding and awareness within the field, paving the way for improved model design.
- Enhanced precision of mRNA vaccine degradation rate: The findings underscored the significant impact of NN architecture on model performance. Models constructed using the proposed layer structuring approach consistently achieved superior results, demonstrating the importance of optimizing the order and composition of layers for optimal learning.

2. Related works

Over the past decade, the trend of leveraging hybrid models has continued to gain momentum. Studies leveraging hybrid models—a fusion of different machine-learning techniques—proliferated significantly to address complex challenges in diverse domains (Kurz et al., 2022; Ünal and Başçiftçi, 2022), proving that researchers and practitioners have recognized the potential of hybrid models in combining the strengths of different algorithms, thereby enhancing overall performance. As ML evolved, studies transitioned to more intricate hybrid models, incorporating advanced techniques like ensemble learning and DL architectures. The advent of ensemble methods, such as random forests (RF) and gradient boosting (GB) (Ünal and Başçiftçi, 2022; Buyrukoğlu and Savaş, 2023), allowed researchers to integrate predictions from multiple models, showcasing improved predictive capabilities. Concurrently, the emergence of deep learning paved the way for hybrid models that integrated convolutional neural networks (CNNs), RNNs, and other specialized architectures (Ünal and Başçiftçi, 2022). These contemporary methodologies underscore the versatility and adaptability inherent in hybrid models, thereby demonstrating their efficacy in addressing intricate challenges across diverse fields. In tandem with the continual progression of the machine learning domain, the

employment of hybrid models stands as a dynamic and evolving realm of research. Ongoing studies persistently delve into novel combinations and integrations, thereby extending the boundaries of achievable outcomes in hybrid models concerning predictive accuracy, interpretability, and applicability to real-world scenarios. The hybridizing model approach has gained considerable traction in the medical field (Schmidt and Hildebrandt, 2021; Ünal and Başçiftçi, 2022). Its application ranges from refining diagnostic accuracy, optimizing treatment strategies, and elucidating complex patterns in biomedical data, of which mRNA vaccine degradation rate is one of many.

Across the years, with the availability of datasets (Wayment-Steele et al., 2021) since 2020, the number of studies on the RNA degradation rate with AI has shown a significant increase. In comparison to other research topics, it is still in the minority, but it is steadily growing. Imran et al. (2020) and Singhal (2020) had each published an article on the topic. While Imran et al. (2020) proposed only the LSTM algorithm, Singhal (2020) proposed two additional algorithms, i.e., GRU and graph convolution networks (GCN), in addition to LSTM for training and testing. Imran et al. (2020) performed a 67:33 percentage split and optimized with the RMSprop optimizer; however, Singhal (2020) performed a 4-fold cross-validation and utilized the Adam optimizer. Comparing the results between Imran et al. (2020) and Singhal's (2020) models, it was noticed that the GRU model trained with parameters determined by Singhal (2020) could achieve a better score.

In 2021, five papers were published on this prediction topic. Qaid et al. (2021) and Wang (2021) published their experimental results before mid-year, while Ing et al. (2021a) and Muneer et al. (2021) published their observations in late 2021. All the papers before 2022 on this research topic were semi-supervised, except Qaid et al. (2021). Contrary to other authors (Imran et al., 2020; Singhal, 2020; Ing et al., 2021a; Ing et al., 2021b; Muneer et al., 2021; Wang, 2021; Qaid et al., 2021) considered the supervised learning paradigm. Qaid et al. (2021) suggested a bidirectional approach for both the LSTM and GRU algorithms, forming Bi-LSTM and Bi-GRU. Hybridizing the algorithms, Qaid et al. (2021) developed three three-layer hybrid models, i.e., 3-layer LSTM, 3-layer GRU, and a 3-layer hybrid model with a Bi-LSTM algorithm interposed between two Bi-GRU algorithms. With the dropout set at 0.5, embedding dimensions at 100, and hidden units at 512, the 3-layer LSTM model modified by Qaid et al. (2021) managed to lower the prediction error to 0.125 (MCRMSE).

GCN is once again considered for modifications by Wang (2021) and Muneer et al. (2021). Wang (2021) modeled GCN with a multi-head attention (MHA) mechanism and pseudo-labeling (PL) technique, then ensembled the modified GCN model with the WaveNet-GRU-LSTM model to produce a final model that achieved a value of 0.3489 MCRMSE in predicting the degradation rate of RNA. Muneer et al. (2021), on the other hand, hybridized GRU and convolutional neural networks (CNN) to GCN, forming two hybrid models, i.e., GCN_GRU and GCN_CNN. Both Wang (2021) and Muneer et al. (2021) performed a five-fold cross-validation data splitting method; however, the GCN_GRU model modeled with the technique performed by Muneer et al. (2021) managed to have an approximately 0.0074 MCRMSE lower than the model developed by Wang (2021) (0.3489).

Experiments were conducted by Ing et al. (2021a, 2021b) to explore the possibilities and potential of machine learning (ML) on the dataset (Ing et al., 2021a; 2021b). Similar to Singhal (2020), Wang (2021), and Muneer et al. (2021), Ing et al. (2021a, 2021b) experimented with the same data-splitting technique but with 10-fold cross-validation. With a learning rate of 0.01 and maximum binning of 512, the tree-based light gradient boosting machine (LGBM) outdid the other two ML algorithms, i.e., linear regression (LR) (Ing et al., 2021a) and random forest (RF) (Ing et al., 2021b), suggested by Ing et al. (2021a, 2021b) showing the potential of ML algorithms on this prediction topic.

Among the many papers available, it is observable that research with practitioners implementing the approach of hybridizing NN layers into hybrid NN models for mRNA vaccine degradation rate prediction is not in the minority. However, no single practitioner structures the hybrid models with consideration for the sequence of hybridization of NN layers. This observation prompts a recognition of an unexplored niche in the field, where opportunities for innovation and advancement lie dormant. In light of this, an exploration of the hybrid sequence is proposed, emphasizing the significance of order dependence within the NN layer sequence in determining the performance of a hybrid model. This exploration hints at the potential for unveiling novel insights and augmenting the predictive prowess of models, thereby establishing an intriguing frontier in the domain of mRNA degradation rate prediction. Furthermore, the contribution of this exploration to developing hybrid models is expected to extend beyond the scope of this research.

3. Methodology

Hybridization of ML models is the integration of two or more models into a single model; the resultant product is called a hybrid model. The main difference between hybridization and ensemble is that the ensemble method is a parallel incorporating technique, forming an ensemble model, while hybridization is achieved by incorporating models serially. The idea of integrating algorithms has become a commonplace approach for practitioners in search of better scoring.

Stacking, alongside ensemble and hybridization, emerged as another common approach to architecting neural network models for robust machine learning. These different approaches

combine multiple networks in unique ways to improve performance and reduce overfitting. While the specific choice of approach depends on the task and network architecture, all three offer valuable tools for building reliable and effective machine learning models. Collectively, these approaches mitigate variance and enhance generalization capabilities, synergistically reinforcing the model's resilience and predictive prowess.

Like hybridization, stacking constructs NN models sequentially, typically employing compatible layers to facilitate efficient feature extraction. However, the key distinction lies in the level of abstraction at which the combination takes place. Hybridization combines distinct network architectures or entire models as modular components, whereas stacking concentrates on integrating individual layers or sub-networks with analogous functionalities (He et al., 2016). This allows stacking to leverage specific expertise in designing individual components, while hybridization offers more flexibility in exploring diverse architectures.

3.1. Dataset

The NN models are implemented on the dataset released by Wayment-Steele et al. (2021), which is the same dataset used by Qaid et al. (2021) for supervised learning. The training data, comprised of 2400 samples of RNA sequences, is accompanied by a square-matrixed form of base-pairing probabilities (BPPs) feature data. The BPPs feature data will deliver the pairing probabilities of each nucleotide of every sample with adjoined nucleotides along the RNA sequence. Throughout the training phase, several authors (Ing et al., 2021a; 2022; 2021b; Qaid et al., 2021; Wang, 2021) additionally employed this feature data to promote the generation of more plausible predictions.

A good feature could be a reassuring and mighty ally that assists models to score higher, but an inferior one is merely a menace to their potencies. To allow models to extract useful information from the raw data, it has to be processed to ensure its presentability before being fed into models for training. The five aggregated numerical features—Exponential Weighted Average, Maximum, Normalize, Position Average, and Summation—engineered from the raw BPPs feature data are included as inputs for model construction, aiming to ensure comparability with the supervised learning of Qaid et al.'s (2021) proposed models. In addition to the training data sourced by Wayment-Steele et al. (2021), augmented data attached to Qaid et al. (2021) publications are considered to avoid any undesired fitting issues. The total number of samples in the training data doubled when augmentation samples were enabled. Cleaning the newly developed training data is the subsequent procedure. This leaves only 4192 clean samples in the training data after eliminating the 608 noisy samples with a signal-to-noise ratio greater than 1. In this research, a 90:10 percentage split is performed on the processed, cleaned training data to split the data into a training set of 3772 RNA-sequence samples and a validation set of 420 RNA-sequence samples. The dataset has eight main input fields (three categorical, five numerical) and five major output fields. Table 3 details the sample and concept for each field. The *sequence*, *structure*, and *predicted_loop_type* were initially introduced in string form and thus were referred to as categorical inputs. These three inputs are label-encoded into numerical representations to enable models to analyze the inputs, as indicated in Table 4.

Table 3: Input and output fields of the RNA data sets for COVID-19 mRNA vaccination production (Wayment-Steele et al., 2021)

Type	Field	Feature	Length[1,2]	Sample	Description
Categorical	Input	sequence	107	GGAAAAGCUC …	Sequence of RNA genetic codes
		structure	107	..(((.)…) …	Sequence of '(', ')', and '.' characters that serve as the pairing state of the RNA
		predicted_loop_type	107	EESSBSISSM …	Sequence of bpRNA that corresponds to RNA structure
Numerical	Input	bpps_ewa	107	0.0015, 0.0017, …	Sequence of exponentially weighted averaging probabilities
		bpps_max	107	0.0218, 0.0387, …	Sequence of maximum probabilities
		bpps_nrmlz	107	2.0660, 1.2272, …	Sequence of normalized probabilities
		bpps_pa	107	0.0031, 0.0031, …	Sequence of position-averaging probabilities
		bpps_sum	107	0.1984, 0.1837, …	Sequence of summed probabilities
	Output	reactivity	68	0.3297, 1.5693, …	Sequence of reactivity values
		deg_Mg_pH10	68	0.7556, 2.9830, …	Sequence of degradation values, under pH 10, in the presence of magnesium
		deg_pH10	68	2.3375, 3.5060, …	Sequence of degradation values, under pH 10
		deg_Mg_50C	68	0.3581, 2.9683, …	Sequence of degradation values, under 50°C, in the presence of magnesium
		deg_50C	68	0.6382, 3.4773, …	Sequence of degradation values, under 50°C

[1]'Public' sample: 107 bases (input) and 68 bases (output), 'private' sample: 130 bases (input) and 91 bases (output); [2]Only the first 68 bases are measured because: (1) the last 39 bases are sequencer-processing and oligonucleotide-binding bases (Kauffmann et al., 2022); and (2) technical constraints (Wayment-Steele et al., 2021)

3.2. Neural network layers

This study centered on architecting neural network models through sequential approaches, specifically hybridization and stacking with dense, BiLSTM, and BiGRU layers. Within the Keras framework, two alternatives for constructing single-layer bidirectional LSTMs and GRUs—(1) the wrapped manner (*layers.Bidirectional(layers)*) and (2) the sequential manner (*layers.Sequential (Bidirectional(layers))*)—were investigated. Table 5 below tabulated a comparison between the

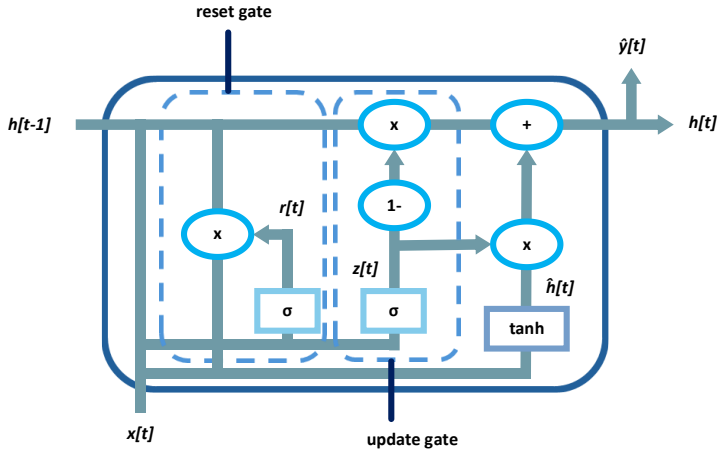
construction manner. LSTM and GRU are RNN algorithms that feature cyclical connections to process sequential data. Fig. 1 outlines the GRU and LSTM exemplary architectures for better visualization.

Table 4: Label encoding experienced by categorical input fields

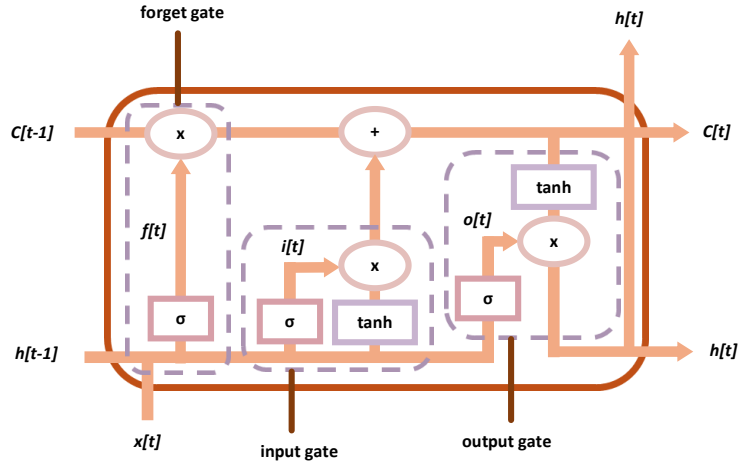
Input	Encoding method							
Structure	Character	(	)	.				
	Index	0	1	2				
RNA sequence	Character	A	U	G	C			
	Index	3	4	5	6			
Predicted_loop_type	Character	B	E	H	I	M	S	X
	Index	7	8	9	10	11	12	13

Table 5: Comparison of bidirectional layer construction methods in Keras

Manner	Wrapped layers.Bidirectional (layers)	Sequential layers.Sequential (Bidirectional(layers))
Sequence processing	Processes the sequence bidirectionally	Sequentially processes the sequence forward and backward
Output handling	Concatenates both-direction outputs into a single vector	Maintains separate outputs for each direction
Concatenation	Automatic concatenation that might be unnecessary	Offers control to prevent unnecessary concatenation
Information access	Gains simultaneous access to combined bidirectional information	Separate access to information from each direction
Computational efficiency	More efficient for tasks needing immediate access to combined information	Less efficient, necessitates storing and combining outputs later
Memory efficiency	Less efficient for large outputs	More efficient by avoiding concatenation
Potential benefits	Captures richer context and dependencies	Allows separate analysis of forward and backward outputs
Potential drawbacks	Computationally expensive and may overfit irrelevant patterns in non-sequential data	May yield less interpretable results due to separate outputs
Code snippets	layers.Bidirectional(layers.GRU) layers.Bidirectional(layers.LSTM)	layers.Sequential(Bidirectional(layers.GRU)) layers.Sequential(Bidirectional(layers.LSTM))



(a) GRU



(b) LSTM

Fig. 1: Sample architecture of RNN algorithms. (a) GRU; (b) LSTM (Vakharia et al., 2023)

As presented in Fig. 1, LSTMs consist of three gates (input, output, and forget gates), which control the extent of information to be retained, forgotten, or passed along. Contrastingly, GRUs simplify this architecture, utilizing only two gates (update and reset gates). Despite boasting higher accuracy through its intricate gating mechanisms, LSTM's increased complexity translates to higher computational costs, lengthening training and inference times compared to the simpler structured GRU (Mittal et al., 2023; Vakharia et al., 2023). The optimal choice hinges on the task's requirements and resource constraints. To counteract each other's deficiencies, it is proposed that both GRU and LSTM be involved in modeling a four-layer hybrid NN model. It is worth noting the use of *layers.Sequential()* is merely a structural decision that does not affect the configuration of the underlying bidirectional NNs; hence, both code snippets construct a bidirectional NN layer with 256 units in each direction, for a total of 512 hidden units. To potentially simplify the structure of the hybrid NN model, the NN layers are constructed sequentially using *layers.Sequential(Bidirectional(layers))*, which can contribute to overfitting prevention by facilitating the application of regularization techniques. This approach offers more control over layer arrangement and output flow than the wrapped (*layers.Bidirectional(layers)*) approach. The performances of the NN models structured through the sequential approach are subsequently compared with the results obtained by the replicated reference NN models developed in a wrapped manner suggested by Qaid et al. (2021).

In addition to the BiLSTM and BiGRU layers, a fully connected (Dense) layer has been integrated into the four-layer hybrid NN architecture in further processing the extracted features. As is standard practice in Keras, this dense layer was constructed sequentially, as indicated by the code snippet *layers.Sequential(layers.Dense())*. Dense layers play a pivotal role in neural network architectures, contributing significantly to the model's overall functionality. Primarily, these layers apply non-linear transformations to incoming inputs, enabling the model to uncover intricate patterns and dependencies that might elude linear methods. This capability underpins their versatility across diverse machine-learning tasks. Dense layers allow the model to capture complex non-linear relationships and generate higher-level features, often expressed as combinations or abstractions of the original input features, thereby enhancing overall model performance (Rao and Reimherr, 2023).

Furthermore, dense layers excel in the integration of information from different layers, contributing to comprehensive decision-making processes (Huang et al., 2017; Rao and Reimherr, 2023). Beyond feature learning, dense layers can also serve as dimensionality reduction tools. By reducing the dimensionality of the input data, they can mitigate overfitting and improve computational efficiency (Dileep et al., 2020; Hasan et al., 2023). Their adaptable nature and flexibility are underscored by their applicability to diverse tasks such as classification, regression, feature extraction, and dimensionality reduction, making them a cornerstone in neural network architecture design.

3.3. Model architecture

The order of the fully connected (Dense) layer within a hybrid NN architecture can critically impact the model's information flow and expressive power (Dileep et al., 2020). While later inclusion encourages higher-level feature abstraction, final output generation, and decision-making, early placement fosters low-level feature extraction, intermediate representation shaping, complex feature interactions, and non-linear transformations. The general potential effects of the inclusion order of dense NN layer in a hybrid NN model are demonstrated in Table 6.

Table 6: Effects of dense layer inclusion order on hybrid NN model performance

Position of dense layer	Potential strengths	Potential risks
Beginning	Early feature extraction Learn key feature combinations from input Reduces dimensionality Aiding efficiency	Discard useful features if without proper config. Difficulty capturing sequential relationships
Intermediate	Refine features extracted Learn features from intermediates Capture more generalizable patterns Introduces non-linearities	Overfitting if without proper regularization Disrupt temporal flow if not carefully integrated
End	Integrate and condense information Serves as a crucial decision-making layer	Less effective for learning new features Limit nns' direct influence on final predictions

Order refers to the proper sequence of arrangements that encourages the avoidance of disruptions during processing by offering a comforting framework that should be followed to prevent errors. Ensemble learning, stacking, and hybridization are algorithmic integration approaches, as stated, although ensemble learning is performed in parallel and the other two in series. Since it performs the parallel approach, the effect of the order of NN layers in an ensemble model on prediction performance is negligible. However, this does not apply to hybridization, which performs serial algorithmic integration. Therefore, experiments are conducted by considering the hybridization order during four-layer hybrid modeling to predict the degradation rate of RNA-sequence samples for mRNA vaccine development. While stacking, akin to hybridization, involves serial

algorithmic integration, it is worth highlighting that, in this research, the NN layers employed for stacking do not encompass the integration of distinct NN architectures within a stacked NN model. Consequently, the consideration of ordering or sequencing is not centered on the stacking approach of this research. Given these considerations, this study engineered hybrid NN models to predict the degradation rate of mRNA vaccines. The investigation delves into the impact of hybridizing sequences and the structural configuration of NN layers within the model. The composition of NN layers in each of the constructed NN models is outlined in Table 7, and Fig. 2 shows a sample of the architecture of the NN models for better insight.

According to the reference, Qaid et al. (2021) engaged in supervised learning. Fortunately, Singhal (2020) and Imran et al. (2020) disclosed the predictive performance of their models in the training and validation phases, allowing a comparison between the models' potencies. According to the results, Qaid et al.'s (2021) best model scored a value of 0.109 and 0.125 for training and validation losses, respectively, which is much lower than the best models of Imran *et al.* (training loss = 0.4904, validation loss = 0.5165) and Singhal (2020) (training loss = 0.1143, validation loss = 0.1787). In addition, the difference between validation and training losses evidenced that the model proposed by Qaid et al. (2021) exhibited a lesser propensity to overfit. Hence, the modeling concept put forward by Qaid et al. (2021) is worth serving as a paradigm for improvements.

Table 7: Hybridizing sequences of neural network layers for neural network model development

No. [1]	Model [2]	Sequence			
		Layer 1	Layer 2	Layer 3	Layer 4
1	Stack_GGG	BiGRU	BiGRU	BiGRU	nil
2	Stack_SSS	BiLSTM	BiLSTM	BiLSTM	nil
3	Hybrid_LGSS	DENSE	BiGRU	BiLSTM	BiLSTM
4	Hybrid_LSGS	DENSE	BiLSTM	BiGRU	BiLSTM
5	Hybrid_LSSG	DENSE	BiLSTM	BiLSTM	BiGRU
6	Hybrid_GLSS	BiGRU	DENSE	BiLSTM	BiLSTM
7	Hybrid_GSLS	BiGRU	BiLSTM	DENSE	BiLSTM
8	Hybrid_GSSL	BiGRU	BiLSTM	BiLSTM	DENSE
9	Hybrid_SSLG	BiLSTM	BiLSTM	DENSE	BiGRU
10	Hybrid_SLSG	BiLSTM	DENSE	BiLSTM	BiGRU
11	Hybrid_SLGS	BiLSTM	DENSE	BiGRU	BiLSTM
12	Hybrid_LSGG	DENSE	BiLSTM	BiGRU	BiGRU
13	Hybrid_LGSG	DENSE	BiGRU	BiLSTM	BiGRU
14	Hybrid_LGGS	DENSE	BiGRU	BiGRU	BiLSTM
15	Hybrid_SLGG	BiLSTM	DENSE	BiGRU	BiGRU
16	Hybrid_SGLG	BiLSTM	BiGRU	DENSE	BiGRU
17	Hybrid_SGGL	BiLSTM	BiGRU	BiGRU	DENSE
18	Hybrid_GGLS	BiGRU	BiGRU	DENSE	BiLSTM
19	Hybrid_GLGS	BiGRU	DENSE	BiGRU	BiLSTM
20	Hybrid_GLSG	BiGRU	DENSE	BiLSTM	BiGRU
21	GRU [3]	BiGRU	BiGRU	BiGRU	nil
22	Hybrid_1	BiLSTM	BiGRU	BiGRU	nil
23	Hybrid_2 [3]	BiGRU	BiLSTM	BiGRU	nil
24	Hybrid_3	BiGRU	BiGRU	BiLSTM	nil
25	LSTM [3]	BiLSTM	BiLSTM	BiLSTM	nil

[1]Models No. 1 to 20 are developed using sequential manner, while Models No. 21 to 25 are replicated (No. 21, 23, and 25) and developed (No. 22 and No. 24) using wrapped manner, in line with the idea proposed by Qaid et al. (2021). Differentiating with a blue partition line; [2]G = BiGRU, S = BiLSTM, L = Fully Connected (Dense); [3]Models proposed by Qaid et al. (2021)

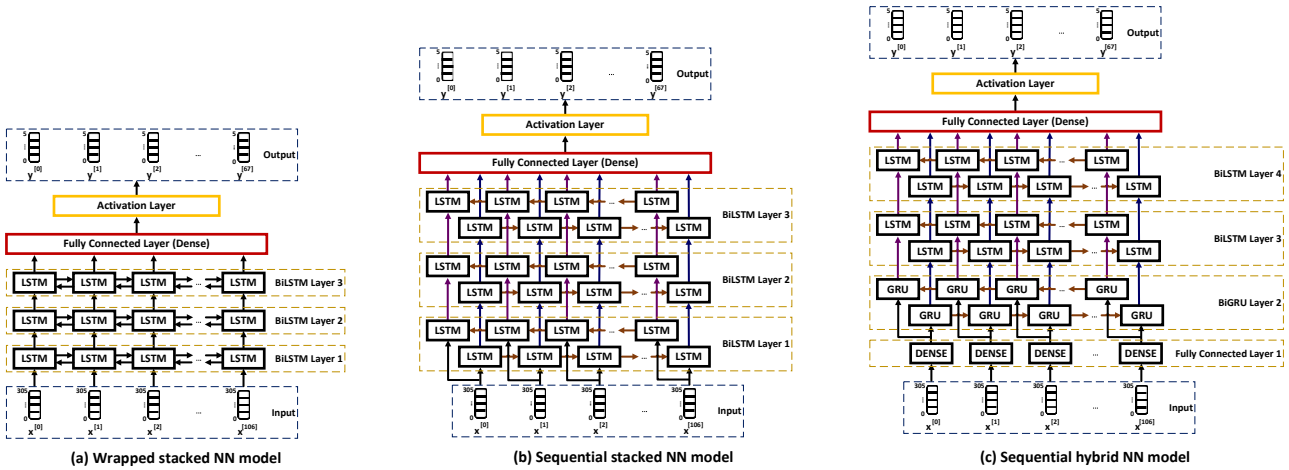


Fig. 2: Sample architecture of NN model. (a) Wrapped approach 3-layer stacked BiLSTM model; (b) Sequential approach stacked BiLSTM model; (c) 4-layer Sequential approach hybrid LGSS model

3.4. Hyperparameter initialization

The primary focus of this research lies in a comprehensive comparison of the structuring manner, dense layer impact, and hybridization sequences of NN layers in neural networks. The model proposed for this prediction task focuses specifically on a sequential model, as opposed to the NN models modified by Qaid et al. (2021) through the wrapped approach. Contrary to stack models, which can suffer from complexity problems depending on the NN layers involved, a hybrid approach is adopted. This strategic choice presents a valuable alternative to excessively complex or overly simplistic NN architectures, achieving a harmonious balance between complexity, accuracy, and simplicity. This is accomplished by fostering the integration of diverse NN layers, aligning with the overarching goals and vision of the model. Notably, to ensure equity for a fair comparison, including those replicated from NN models proposed by Qaid et al. (2021), uniformity is maintained by utilizing identical hyperparameters to all NN layers across each model of this experiment.

A linear activation function was employed to address the regression nature of this research. After evaluating a range of dropout values (0 to 0.5), a dropout rate of 0.125 was meticulously chosen

based on the average performance across every engaged NN model. To further mitigate potential overfitting, orthogonal kernel initialization, a hyperparameter selection that promotes model generalization was selected. Each character of the categorical features (*sequence*, *structure*, and *predicted_loop_type*) was embedded in a 100-dimensional space, concatenated into a total of 300 features from the initial 14 encoded categorical features. This dimensional expansion enriches the model's information intake. Values assigned to parameters and hyperparameters are detailed more comprehensively in Table 8.

Table 8: NN model configuration parameters and hyperparameters

Parameters/Hyperparameters	Value
Activation Function	Linear
Dropout	0.125
Embedding Layer	100
Hidden Units	256
Kernel Initializer	Orthogonal
Layers [1]	Bidirectional
Optimizer	Adam
Return Sequence[1]	True
Test Size	0.1

[1]Applied only to BiGRU and BiLSTM

3.5. MCRMSE

Performance metrics are sometimes referred to as error metrics for regression or simply as evaluation metrics in general. As the name suggests, the primary role of a performance metric is to evaluate models' performance on given data. As opposed to classification, which measures accuracy (Fern et al., 2022), an error is a metric to determine the potency of a regressor. Along with this research topic, practitioners are introducing a new performance metric named MCRMSE. As the name implies, MCRMSE is the mean of root mean square errors (RMSE). An RMSE equation can be defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(predicted_i - actual_i)^2} \quad (1)$$

therefore, the equation of MCRMSE is defined as:

$$MCRMSE = \frac{1}{N_t}\sum_{j=1}^{N_t}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(actual_{ij} - predicted_{ij})^2}. \quad (2)$$

The MCRMSE performance metric is introduced to address the numerous outputs of this research topic. Instead of receiving numerous prediction errors, one for each output, which might be confusing and misleading, a single-value prediction error that can indicate the model's overall performance is preferred. For better understanding, taking this prediction topic as an example, *reactivity*, *deg_Mg_pH10*, *deg_Mg_50C*, *deg_pH10*, and *deg_50C* are the outputs. At the end of the evaluation phase, each will display its corresponding RMSE result, resulting in five RMSEs to be obtained. The RMSE that results from averaging these five RMSEs over the number of output fields—in this example, five—is the same as the single-value MCRMSE. In short, the MCRMSE equation eliminates the need for practitioners to manually average the prediction errors when more than one output field is studied in an experiment. The concept of the overall process could be summarized in the workflow as shown in Fig. 3.

4. Results and discussions

The focus of this article is to present the idea of developing a hybrid model while considering the hybridizing sequence, impacts of dense layers, and the manner of structuring NN layers in neural networks while ensuring the performance of the improved model outperforms the off-the-shelf models. To ensure the achievement of utter comparability, the experimental setups and the suggested parameters and hyperparameters were applied to our models (No. 1 to 20), which are architected with code snippet (sequential manner): *layers.Sequential(Bidirectional(layers))*, and to the models replicated according to the ideas proposed by Qaid et al. (2021) (No. 21 to 25), which are architected with the code snippet (wrapped manner): *layers.Bidirectional(layers)*.

The code is coded in Python on a Jupyter environment with a high-caliber NVIDIA GeForce RTX 3060 Graphics Processing Unit (GPU) that boasts excellent Ampere architecture and compute capability (8.6) (Padilla-Perez et al., 2022) as the NVIDIA GeForce RTX 3090 (Jiménez-Ruiz et al., 2023). In addition to that, the TensorFlow library, with a version of 2.6.0, is optimized throughout the experiment. The prediction results of each of the structured models after evaluation with the MCRMSE performance metric under these experimental setups are tabulated in Table 9.

As declared, MCRMSE is derived from RMSE. Thus, MCRMSE is also a negative orientation scoring technique. A model that has minimal MCRMSE results is more optimal. However, a low prediction error is not the only standard to be met. Observation on whether the model is experiencing any fitting issues is required. This assessment is necessary to determine the potential of a model. The difference between the training loss (Loss $_{Training}$) and validation loss (Loss $_{Validation}$) can be used to observe the fitting problems. The smaller the difference between the losses in terms of error, the lesser the tendency of a model to overfit. On the other hand, if a negative value is obtained from the difference, the model is underfitting. In a word, a regressor can only be considered top-notch if and only if it can show minimal prediction errors while at the same time neither underfitting nor overfitting the data.

According to the results presented in Table 9, models developed using the proposed approach (in a sequential manner) have shown better performance in terms of timing and predictability. It is noticeable that, although all the NN layers engaged in the NN models and the sequence of hybridization were mirroring, e.g., the Stack_GGG model (model No. 1)

compared to the GRU model (model No. 21) and the Stack_SSS model (model No. 2) compared to the LSTM model (model No. 25), the models developed with our proposed approach (in a sequential manner) still dominated the models constructed by Qaid et al. (2021) (in a wrapped manner) in terms of time for this research.

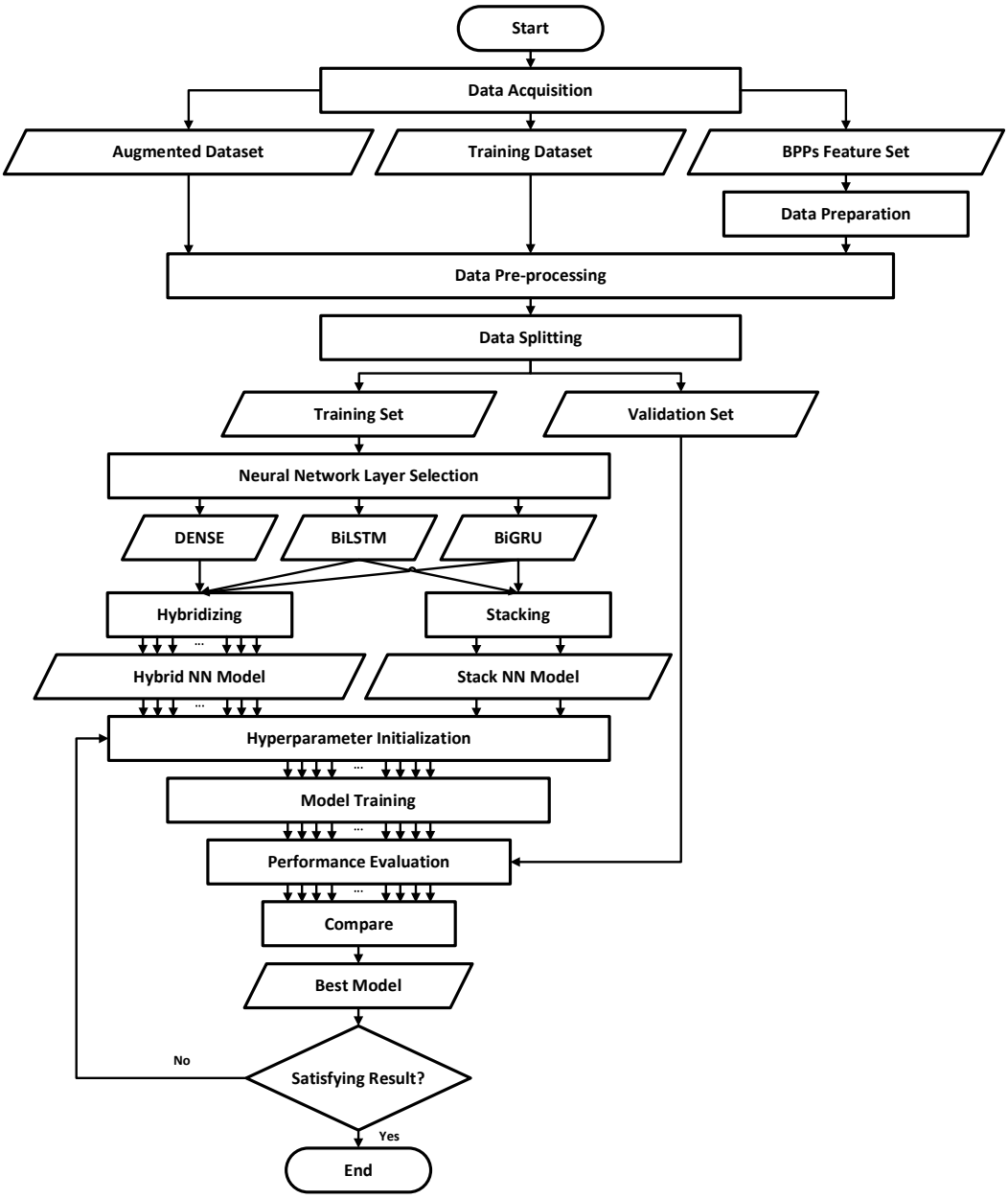


Fig. 3: Methodology workflow

Moreover, it was found that, with comparable results, not only in terms of computational time that reflects the complexity of respective NN models, sequential-manner-developed stacked NN models have been more resistant to overfitting issues than wrapped-manner-developed stacked NN models.

Across all the models, model No. 3, Hybrid_LGSS, which completes the whole process in 12.2771 minutes, is the best among all the structured models, considering every aspect. Not only does Hybrid_LGSS manage to have the best precision (Loss $_{Training}$: 0.0749, Loss $_{Validation}$: 0.0767), but it also shows no sign of overfitting nor underfitting with its minimal difference (0.0018) compared to the LSTM model (model No. 25, best among model No. 21 to 25) structured with Qaid et al.'s (2021) idea that takes 14.6032 minutes to score MCRMSEs of 0.0563 and 0.0776 for training and validation losses, respectively. The difference between the losses of Model No. 25 (the LSTM model) is 0.0213, which is 0.0195 higher than that of Hybrid_LGSS, showing the LSTM model is more prone to overfitting. To summarize, despite being a 4-layer model, our model, particularly Hybrid_LGSS, took less time, performed better, and is less prone to overfitting than models structured with the ideas suggested by Qaid et al. (2021), demonstrating the achievement of the development of improved models with lower complexity but advanced functionality. Therefore, according to the experimental results derived from the hybrid NN models structured in this study, it is evident that the ability to learn from context and

sequence is significantly enhanced when integrated with dense layers. These dense layers exhibit proficiency in capturing non-linear relationships within the data, empowering RNNs to extract intricate features and dependencies. This integration holds the potential to yield more precise predictions of vaccine degradation over time.

Table 9: MCRMSEs determined for the developed models

No.	Model	Loss $_{Training}$[1]	Loss $_{Validation}$[1]	Difference[2]	Time (minute)
1	Stack_GGG	0.0857	0.0890	0.0033	11.3136
2	Stack_SSS	0.0696	0.0785	0.0089	13.6362
3	Hybrid_LGSS	0.0749	0.0767	0.0018	12.2771
4	Hybrid_LSGS	0.0792	0.0808	0.0016	11.3495
5	Hybrid_LSSG	0.0718	0.0778	0.0060	12.4918
6	Hybrid_GLSS	0.0839	0.0868	0.0029	12.0642
7	Hybrid_GSLS	0.0834	0.0884	0.0050	12.0382
8	Hybrid_GSSL	0.0789	0.0815	0.0026	13.2822
9	Hybrid_SSLG	0.0783	0.0897	0.0114	12.0673
10	Hybrid_SLSG	0.0827	0.0909	0.0082	11.2852
11	Hybrid_SLGS	0.0863	0.0881	0.0018	11.7262
12	Hybrid_LSGG	0.0816	0.0846	0.0030	10.1630
13	Hybrid_LGSG	0.0758	0.0785	0.0027	12.4274
14	Hybrid_LGGS	0.0824	0.0825	0.0001	11.0270
15	Hybrid_SLGG	0.0918	0.0959	0.0041	11.8359
16	Hybrid_SGLG	0.0887	0.0972	0.0085	11.1258
17	Hybrid_SGGL	0.0853	0.0889	0.0034	11.8305
18	Hybrid_GGLS	0.0926	0.0956	0.0030	11.1290
19	Hybrid_GLGS	0.0878	0.0881	0.0003	11.3620
20	Hybrid_GLSG	0.0874	0.0911	0.0037	10.9147
21	GRU	0.0737	0.0900	0.0163	12.9262
22	Hybrid_1	0.0696	0.0864	0.0168	12.1602
23	Hybrid_2	0.0668	0.0834	0.0166	12.9176
24	Hybrid_3	0.0676	0.0848	0.0172	12.9482
25	LSTM	0.0563	0.0776	0.0213	14.6032

[1]Losses, also known as prediction errors. The lower, the better; [2]Difference = Loss Training – Loss Validation, the smaller the difference, the lesser the overfitting

In addition, it is clear from the results that the slightest change in the order of the hybridization sequence will result in a different ranking of the performance of the hybrid models. This is despite the fact that the NN layers used were the same. By way of illustration, taking Hybrid_LGSS (model No. 3) and Hybrid_GLSS (model No. 6) as examples, the results are poles apart with a simple conversion in the sequence of a dense layer and the BiGRU layer, demonstrating the impact of the order of the hybridization sequence on the performance of a hybrid model.

5. Conclusion and future directions

To conclude, the experiment's primary goal of developing a better hybrid model with a less sophisticated structure yet superior performance was accomplished. The experimental results have also proven that the proposed idea for hybrid modeling is applicable and advisable. Additionally, the reduction in complexity of a hybrid NN model is achievable not only with the structuring approach but also with the inclusion of a dense layer. The dense layer has proven its potential to improve the stability and efficiency of the learning process of subsequent layers by preconditioning the input features through scaling or normalizing their distributions. By introducing a non-linear transformation, the dense layer might also guide the information flow through the network, allowing the subsequent NN layers to focus on specific aspects of the data that are most relevant for degradation prediction. Moreover, the displayed results leave a cautionary note for practitioners to never neglect the order of the hybridization sequence during hybrid modeling. In addition, the results reaffirmed the argument that LSTM is a better choice than GRU when dealing with a larger data set and when accuracy is a primary concern.

mRNA has been promoted since 1989 as a therapeutic and received public advocacy in vaccinology a few years later. The benefits of the biosafety of mRNA are vast and well documented; thus, although the genetic construct of mRNA is tiny, this does not restrict it from improving the health and quality of life of every being. mRNAs, including their vaccines, have no interaction with the genome, and therefore, the possibility of detrimental genomic interaction is a negative score, thus clearing the concerns. In short, mRNA vaccines are much less pernicious than other types of vaccines. Therefore, in spite of the fact that mRNA vaccines may be susceptible to degradation, they have been in use for years without abandonment. The less detrimental characteristics of mRNA have scientists willingly committing to investing effort and time into improving them against not just COVID-19 but also other infectious diseases.

To this end, considering various factors, predicting the mRNA vaccine's degradation rate with AI is accompanied by multiple advantages, including reducing the time required by the overall procedure compared to laboratory experiments and, most importantly, improving the overall efficiency of mRNA vaccine development. In addition, since this experiment requires only the genetic codes of RNA sequences in an mRNA vaccine rather than the sample vaccines, it reduces the need for and waste of resources, thus gracefully eliminating further complications to the problem of limited vaccine supply.

The application of a hybrid NN model in this manuscript has led to significant advancements in predicting mRNA vaccine degradation rates. The research of this manuscript highlights the crucial role of the sequence of layer hybridization within NNs. By strategically layering different NN layers, the model's efficacy and predictive accuracy are enhanced with utilized NN layers. In particular, feature extraction and data interpretation capabilities have been markedly improved by integrating dense layers into RNN configurations.

Through rigorous testing, it has been demonstrated that the sequential structuring of a model can reduce its susceptibility to overfitting compared to a wrapped configuration. Furthermore, varying the NN layer sequence significantly affects the outcomes. In addition to revealing different aspects of mRNA stability, each configuration demonstrated the dramatic effect that subtle changes in layer configuration could have on prediction performance. This variability underscores

the importance of customizing the layer sequence to align with specific data characteristics, revealing unique aspects of the data that inform vaccine stability.

In the mRNA vaccine development realm, the proposed hybrid model's ability to adapt the sequence of layers provides a powerful tool for researchers and practitioners. This flexibility allows model fine-tuning predictions based on the unique properties of each mRNA sequence and vaccine formulation, leading to more accurate predictions. Bridging the gap between advanced computational techniques and tangible health benefits, this manuscript detailed the practical applications and identified the most effective layer configurations. The proposed model significantly reduces the time and resources required for empirical testing, thereby accelerating the vaccine development cycle—crucial during urgent health crises.

In practical terms, the proposed NN model can be integrated into several stages of the vaccine development pipeline. First, at the formulation phase, it can predict the stability of various mRNA sequences under different conditions, guiding the selection of the most stable candidates for further development. This application reduces the reliance on extensive empirical testing. Later, in the production phase, the model predicts mRNA degradation behaviors based on environmental factors such as temperature and pH, which are critical for scaling up manufacturing processes effectively.

By highlighting the importance of layer sequencing in improving prediction accuracy and reducing errors, the implications of the findings extend beyond conventional modeling techniques and impact the broader ML field. Practitioners are encouraged to account for the sequence of layers when structuring models within NN frameworks and across all ML applications. This strategic sequencing represents more than a mere technical refinement but a crucial practice that significantly improves the utility and effectiveness of a hybrid NN model. As this field evolves, this manuscript serves as a testament to the intersection of computational power and biological insight, urging ongoing collaboration between computational and biological scientists to refine prediction models, ensuring models meet the practical requirements of vaccine development and are adaptable across various research settings.

The experiment as a whole is significant; however, Wayment-Steele et al. (2021) claimed that due to technical restrictions in the data sample development process, compared to the actual laboratory RNA sequence length for the production of mRNA vaccines, the number of bases that make up a sequence sample in this dataset has not even reached 3% of the actual length (Zhang et al., 2019). A more comprehensive and expanded database is strongly encouraged in light of this. At the same time, the extension of the genetic code sequences for improvement could help increase the reliability and validity of the obtained results that reflect the predictive performance. For future directions, practitioners can begin by polishing the experimentation techniques. Experiment with a hybrid approach that combines various alternative algorithms or engages in fine-tuning hyperparameters to achieve a dual objective: reducing overall computational time and enhancing performance.

Nowadays, the public is encouraged to get involved in achieving the Sustainable Development Goals (SDGs) of the 2030 Agenda to address global challenges. Vaccines are frequently used botanically to protect against plant infections, in addition to being beneficial to Animalia. Therefore, advancing this experiment undoubtedly contributes to achieving the third goal of the SDGs 2030: Good Health and Well-being. Overall, this study is not only adding to and offering perspective to the field of data analysis, but it is also a vital study that supports the vaccination community in playing a significant role in balancing biology and ecology. Therefore, it is worthy of attention, exploration, promotion, and advocacy.

Acknowledgment

The authors gratefully acknowledge funding from the Fundamental Research Grant Scheme (FRGS) funded by the Ministry of Higher Education (MOHE), Malaysia. FRGS/1/2021/TKO/UNIMAP/02/65.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Banerjee S, Banerjee D, Singh A, Kumar S, Pooja D, Ram V, Kulhari H, and Saharan VA (2023). A clinical insight on new discovered molecules and repurposed drugs for the treatment of COVID-19. Vaccines, 11(2): 332. https://doi.org/10.3390/vaccines11020332 **PMid:36851211 PMCid:PMC9967525**

Batur ŞC and Laith A (2021). A novel deep learning-based feature selection model for improving the static analysis of vulnerability detection. Neural Computing and Applications, 33(20): 14049-14067. https://doi.org/10.1007/s00521-021-06047-x

Buyrukoğlu S and Savaş S (2023). Stacked-based ensemble machine learning model for positioning footballer. Arabian Journal for Science and Engineering, 48(2): 1371-1383. https://doi.org/10.1007/s13369-022-06857-8

Cahuantzi R, Chen X, and Güttel S (2023). A comparison of LSTM and GRU networks for learning symbolic sequences. In the Science and Information Conference, Springer Nature, Cham, Switzerland: 771-785. https://doi.org/10.1007/978-3-031-37963-5_53

Cattaneo A, Vitali A, Mazzoleni M, and Previdi F (2022). An agent-based model to assess large-scale COVID-19 vaccination campaigns for the Italian territory: The case study of

Lombardy region. Computer Methods and Programs in Biomedicine, 224: 107029. https://doi.org/10.1016/j.cmpb.2022.107029 **PMid:35908330 PMCid:PMC9287580**

Cines DB and Greinacher A (2023). Vaccine-induced immune thrombotic thrombocytopenia. Blood, 141(14): 1659-1665. https://doi.org/10.1182/blood.2022017696 **PMid:36669155 PMCid:PMC9870607**

Crommelin DJ, Anchordoquy TJ, Volkin DB, Jiskoot W, and Mastrobattista E (2021). Addressing the cold reality of mRNA vaccine stability. Journal of Pharmaceutical Sciences, 110(3): 997-1001. https://doi.org/10.1016/j.xphs.2020.12.006 **PMid:33321139 PMCid:PMC7834447**

Dileep P, Das D, and Bora PK (2020). Dense layer dropout based CNN architecture for automatic modulation classification. In the National Conference on Communications, IEEE, Kharagpur, India: 1-5. https://doi.org/10.1109/NCC48643.2020.9055989

Dumpa N, Goel K, Guo Y, McFall H, Pillai AR, Shukla A, Repka MA, and Murthy SN (2019). Stability of vaccines. AAPS PharmSciTech, 20: 42. https://doi.org/10.1208/s12249-018-1254-2 **PMid:30610415**

Fern SH, Amir A, and Azemi SN (2022). Multi-class imbalanced classification problems in network attack detections. In the Proceedings of the 6th International Conference on Electrical, Control and Computer Engineering, Springer Singapore, Kuantan, Pahang, Malaysia: 1057-1069. https://doi.org/10.1007/978-981-16-8690-0_92

Giełżak J, Dejak B, Sokołowski J, and Bociong K (2023). Changes in strength parameters of composite cements as affected by storage temperature—A review of the literature. Coatings, 13(2): 244. https://doi.org/10.3390/coatings13020244

Hasan MS, Alam R, and Adnan MA (2023). Compressed neural architecture utilizing dimensionality reduction and quantization. Applied Intelligence, 53(2): 1271-1286. https://doi.org/10.1007/s10489-022-03221-z

He K, Zhang X, Ren S, and Sun J (2016). Deep residual learning for image recognition. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Las Vegas, USA: 770-778. https://doi.org/10.1109/CVPR.2016.90 **PMid:26180094**

Huang G, Liu Z, Van Der Maaten L, and Weinberger KQ (2017). Densely connected convolutional networks. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Honolulu, USA: 4700-4708. https://doi.org/10.1109/CVPR.2017.243 **PMCid:PMC5598342**

Imran SA, Islam MT, Shahnaz C, Islam MT, Imam OT, and Haque M (2020). COVID-19 mRNA vaccine degradation prediction using regularized LSTM model. In the IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), IEEE, Bhubaneswar, India: 328-331. https://doi.org/10.1109/WIECON-ECE52138.2020.9398044

Ing SH, Abdullah AA, and Kanaya S (2021a). Development of COVID-19 mRNA vaccine degradation prediction system. In the International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, IEEE, Zallaq, Bahrain: 449-454. https://doi.org/10.1109/3ICT53449.2021.9582052

Ing SH, Abdullah AA, Harun NH, and Kanaya S (2021b). COVID-19 mRNA vaccine degradation prediction using LR and LGBM algorithms. Journal of Physics: Conference Series, 1997(1): 012005. https://doi.org/10.1088/1742-6596/1997/1/012005

Ing SH, Abdullah AA, Mashor MY, Mohamed-Hussein ZA, Mohamed Z, and Ang WC (2022). Exploration of hybrid deep learning algorithms for COVID-19 mRNA vaccine degradation prediction system. International Journal of Advances in Intelligent Informatics, 8(3): 404-416. https://doi.org/10.26555/ijain.v8i3.950

Jiménez-Ruiz A, Fernández-Escribano G, Cañas-Carretón M, and Sánchez JL (2023). Using graphics processing units for the efficient dynamic simulation of wind farms. Computers and Electrical Engineering, 112: 109018. https://doi.org/10.1016/j.compeleceng.2023.109018

Kauffmann AD, Kennedy SD, Moss WN, Kierzek E, Kierzek R, and Turner DH (2022). Nuclear magnetic resonance reveals a two hairpin equilibrium near the 3′-splice site of influenza A segment 7 mRNA that can be shifted by oligonucleotides. RNA, 28(4): 508-522. https://doi.org/10.1261/rna.078951.121 **PMid:34983822 PMCid:PMC8925974**

Kaur M and Mohta A (2019). A review of deep learning with recurrent neural network. In the International Conference on Smart Systems and Inventive Technology, IEEE, Tirunelveli, India: 460-465. https://doi.org/10.1109/ICSSIT46314.2019.8987837 **PMid:30649504 PMCid:PMC6432867**

Kurz S, De Gersem H, Galetzka A, Klaedtke A, Liebsch M, Loukrezis D, Russenschuck S, and Schmidt M (2022). Hybrid modeling: Towards the next level of scientific computing in engineering. Journal of Mathematics in Industry, 12: 8. https://doi.org/10.1186/s13362-022-00123-0

Mittal S, Chawla T, Jajoo J, Bansal M, Parashar H, and Ruby D (2023). Performance comparison of recurrent neural networks for metamorphic malware. In the International Conference on Data Science and Network Security, IEEE, Tiptur, India: 1-6. https://doi.org/10.1109/ICDSNS58469.2023.10245978

Muneer A, Fati SM, Akbar NA, Agustriawan D, and Wahyudi ST (2022). iVaccine-Deep: Prediction of COVID-19 mRNA vaccine degradation using deep learning. Journal of King Saud University-Computer and Information Sciences, 34(9): 7419-7432. https://doi.org/10.1016/j.jksuci.2021.10.001 **PMid:38620874 PMCid:PMC8513509**

Okamura S and Ebina H (2021). Could live attenuated vaccines better control COVID-19? Vaccine, 39(39): 5719-5726. https://doi.org/10.1016/j.vaccine.2021.08.018 **PMid:34426024 PMCid:PMC8354792**

Olugbenga O, Oladipo EK, Kolawole OM, Oloke JK, Adelusi TI, Irewolede BA, Dairo EO, Ayeni AE, Kolapo KT, Akindiya OE, and Oluwasegun JA (2022). Bioinformatics, computational informatics, and modeling approaches to the design of mRNA COVID-19 vaccine candidates. Computation, 10(7). https://doi.org/10.3390/computation10070117

Padilla-Perez D, Medina-Sanchez I, Hernández J, and Couder-Castañeda C (2022). Accelerating electromagnetic field simulations based on memory-optimized CPML-FDTD with OpenACC. Applied Sciences, 12(22): 11430. https://doi.org/10.3390/app122211430

Qaid TS, Mazaar H, Alqahtani MS, Raweh AA, and Alakwaa W (2021). Deep sequence modelling for predicting COVID-19 mRNA vaccine degradation. PeerJ Computer Science, 7: e597. https://doi.org/10.7717/peerj-cs.597 **PMid:34239977 PMCid:PMC8237341**

Rao AR and Reimherr M (2023). Modern non-linear function-on-function regression. Statistics and Computing, 33(6): 130. https://doi.org/10.1007/s11222-023-10299-z

Schmidt B and Hildebrandt A (2021). Deep learning in next-generation sequencing. Drug Discovery Today, 26(1): 173-180. https://doi.org/10.1016/j.drudis.2020.10.002 **PMid:33059075 PMCid:PMC7550123**

Singhal A (2020). Predicting hydroxyl mediated nucleophilic degradation and molecular stability of RNA sequences through the application of deep learning methods. Arxiv

Preprint Arxiv:2011.05136.
https://doi.org/10.48550/arXiv.2011.05136

Ünal HT and Başçiftçi F (2022). Evolutionary design of neural network architectures: A review of three decades of research. Artificial Intelligence Review, 55(3): 1723-1802.
https://doi.org/10.1007/s10462-021-10049-5

Vakharia V, Shah M, Nair P, Borade H, Sahlot P, and Wankhede V (2023). Estimation of lithium-ion battery discharge capacity by integrating optimized explainable-AI and stacked LSTM model. Batteries, 9(2): 125.
https://doi.org/10.3390/batteries9020125

Wang Y (2021). Predicting the degradation of COVID-19 mRNA vaccine with graph convolutional networks. In the Proceedings of the 2021 6th International Conference on Machine Learning Technologies, ACM, New York, USA: 111–116. https://doi.org/10.1145/3468891.3468907

Wayment-Steele HK, Kim DS, Choe CA, Nicol JJ, Wellington-Oguri R, Watkins AM, and Das R (2021). Theoretical basis for stabilizing messenger RNA through secondary structure design. Nucleic Acids Research, 49(18): 10604-10617.
https://doi.org/10.1093/nar/gkab764
PMid:34520542 PMCid:PMC8499941

Xiao Y, Tang Z, Huang X, Chen W, Zhou J, Liu H, Liu C, Kong N, and Tao W (2022). Emerging mRNA technologies: Delivery strategies and biomedical applications. Chemical Society Reviews, 51(10): 3828-3845.
https://doi.org/10.1039/D1CS00617G **PMid:35437544**

Xu S, Yang K, Li R, and Zhang L (2020). mRNA vaccine era—Mechanisms, drug platform and clinical prospection. International Journal of Molecular Sciences, 21(18): 6582.
https://doi.org/10.3390/ijms21186582
PMid:32916818 PMCid:PMC7554980

Zhang C, Maruggi G, Shan H, and Li J (2019). Advances in mRNA vaccines for infectious diseases. Frontiers in Immunology, 10: 429065.
https://doi.org/10.3389/fimmu.2019.00594
PMid:30972078 PMCid:PMC6446947