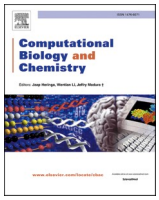
Computational Biology and Chemistry 118 (2025) 108448

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac



PLM-IL4: Enhancing IL-4-inducing peptide prediction with protein language model

Ruiqi Liu [a], Shankai Yan [a], Zilong Zhang [a], Junlin Xu [b], Yajie Meng [c], Qingchen Zhang [a], Leyi Wei [d,e], Quan Zou [f], Feifei Cui [a,*]

[a] School of Computer Science and Technology, Hainan University, Haikou 570228, China
[b] School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China
[c] School of Computer Science and Artificial Intelligence, Wuhan Textile University, Wuhan, Hubei 430200, China
[d] Centre for Artificial Intelligence driven Drug Discovery, Faculty of Applied Science, Macao Polytechnic University, Macao
[e] School of Informatics, Xiamen University, Xiamen, China
[f] Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China

ARTICLE INFO

Keywords:
IL-4 inducing peptides
ESM-2 model
Protein language model
SMOTE
ENN

ABSTRACT

Despite progress in developing antiviral drugs and vaccines, infections continue to be a significant challenge. Interleukin-4 (IL-4) is crucial for regulating immune responses and mediating allergic reactions. This research aims to improve the predictive accuracy of IL-4-inducing peptides by tackling data imbalance and enhancing feature extraction. Specifically, we introduce a new approach that utilizes SMOTE and ENN for balancing the dataset and applies a 30-layer ESM-2 model for extracting deep features. The extracted features are subsequently processed through a Gated Recurrent Unit (GRU) model, which is optimized through hyperparameter tuning. Our method achieves notable improvements, with an AUC of 0.98 and an accuracy of 93.1 %, highlighting its potential to support future immunotherapy and vaccine development efforts. The PLM-IL4 web server is freely accessible at http://www.bioai-lab.com/PLM-IL4, and the datasets used in this research are also available for download from the website.

1. Introduction

In the immune system, interleukin-4 (IL-4) is fundamental to several biological processes, especially in modulating allergic responses and regulating immune functions. IL-4 interacts with CD4 + T cells, facilitating the differentiation of helper T cells (Th cells) into Th2 cells. This differentiation, in turn, stimulates B cell proliferation and differentiation, culminating in antibody class switching and IgE production (Gause et al., 2013; Lambrecht and Hammad, 2015). Additionally, IL-4 is pivotal in the proliferation and differentiation of antigen-presenting cells, essential for controlling inflammatory responses and maintaining immune balance (Pulendran and Ahmed, 2011; Germain, 2002). Given the importance of IL-4 in diseases such as asthma, allergic inflammation, cancer, and autoimmune disorders, identifying peptides that induce IL-4 production has substantial scientific and clinical value (Chen, 2016; Zhu and Paul, 2010). The major histocompatibility complex (MHC) molecules play a key role in processing and presenting antigens on the cell surface, thereby facilitating immune responses against pathogens. Exogenous antigens undergo lysosomal processing before being presented by MHC class II molecules. The MHC class II-associated peptides subsequently bind to CD4 + T cells, initiating cytokine secretion that guides the differentiation of Th cells into Th1, Th2, Th17, or iTregs (Vyas et al., 2008; Blum et al., 2013; Li et al., 2019). Among these, IL-4 is the primary cytokine secreted by Th2 cells, counteracting Th1-triggered pro-inflammatory responses and significantly influencing various immune cells (Lee, 2006; Ma, 2012). IL-4 not only enhances cytotoxic T cell activity and promotes T cell proliferation in vitro but also supports CD8 + cell growth, regulates macrophage phenotypes, and promotes tissue repair and homeostasis (Agarwal and Rao, 1998; Zhu et al., 2009).

Although experimental validation is the most reliable method for identifying IL-4-inducing peptides, it is also time-consuming, laborious, and expensive (Barneh et al., 2024; Mehta et al., 2023). Given the large number of potential antigens and their fragments, developing systematic computational approaches for preliminary pathogen proteome screening is essential (Ali, 2019; Bhasin and Raghava, 2007). Over the past decades, significant efforts have been directed toward creating

* Corresponding author.
E-mail address: feifeicui@hainanu.edu.cn (F. Cui).

https://doi.org/10.1016/j.compbiolchem.2025.108448
Received 20 January 2025; Received in revised form 12 March 2025; Accepted 27 March 2025
Available online 4 April 2025

software programs, tools, online servers, and databases to assist researchers in designing and selecting antigens that activate different components of the host immune system, including cellular, innate, and humoral immunity (Jespersen, 2017; Vita, 2015; Dhanda, 2019). The advent of machine learning and artificial intelligence has driven remarkable advancements in medical imaging and bioinformatics, contributing to improvements in antigen prediction techniques (Ao et al., 2022; Rao et al., 2024; Tarca, 2007; Yan et al., 2024a,b). Notably, the Raghava team has developed several cytokine-specific prediction tools, such as IL-4-Pred, IL-10-Pred, and IL-6-Pred, which have shown enhanced predictive performance to some extent. Recent advancements in deep learning and natural language processing, especially with the advent of transformer models, have shown considerable potential in enhancing the predictive performance of bioinformatics applications (Dhanda et al., 2013a; Kumar et al., 2006; Xiao et al., 2024; Fu et al., 2024; Li et al., 2024). However, existing advanced methods, such as IL-4-Pred (Dhanda et al., 2013b), although providing acceptable prediction accuracy, still face challenges related to data imbalance and model complexity, which limit their applicability in immunotherapy. Meta-IL4 employed ensemble learning to predict IL-4-inducing peptides; however, its dependence on manual feature extraction reduced interpretability, and the overall performance of the machine learning models remained suboptimal (Hassan et al., 2023; Liu et al., 2023).

To improve the prediction accuracy and applicability of IL-4-

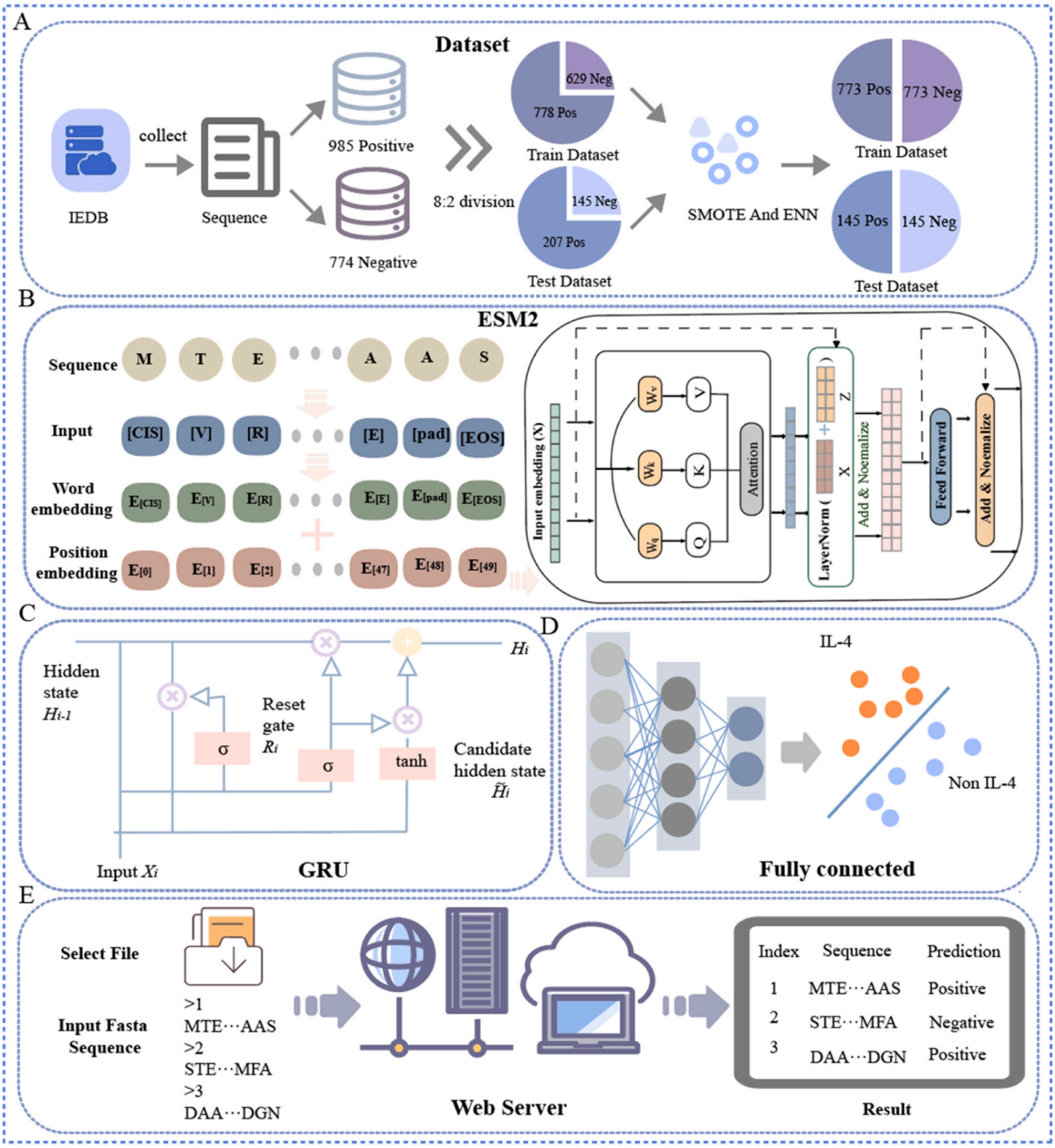


Fig. 1. Overview of the PLM-IL-4 Model Framework. (A) Illustrates the dataset preparation process, including data splitting in an 8:2 ratio and balancing using SMOTE and ENN. (B) Details the feature extraction using ESM-2 and the model construction process (C) Demonstrates the GRU's mechanism for capturing sequence dependencies and classifying peptides as IL-4 inducing or non-IL-4 inducing. (D) Shows the fully connected layer responsible for the final classification. (E) Depicts the deployment of the PLM-IL-4 web server, which provides predictions for the IL-4 inducing potential of input peptide sequences.

inducing peptides, we introduce an optimized method utilizing the ESM-2 protein language model (Geng et al., 2025). This method incorporates ENN (Edited Nearest Neighbors) and SMOTE (Synthetic Minority Over-sampling Technique) for data preprocessing to handle class imbalance and employs a 30-layer ESM-2 model for extracting deep features. The extracted features are subsequently input into a Gated Recurrent Unit (GRU) network, which includes multiple GRU layers integrated with dropout and batch normalization, followed by fully connected layers with dropout. The model is trained through hyperparameter tuning and learning rate scheduling to optimize its performance. Our objective is to develop a more efficient and robust prediction model to support future immunotherapy and vaccine development.

2. Materials and methods

Fig. 1 provides an overview of the IL4-ESM2 model framework for predicting IL-4 inducing peptides. In Fig. 1A, the dataset preparation process is illustrated, starting with the initial dataset obtained from IEDB, which includes 985 IL-4 inducing peptides (positive dataset) and 774 non-IL-4 inducing peptides (negative dataset). The dataset was split into training and testing sets in an 8:2 ratio, resulting in 773 positive and 629 negative sequences in the training set, and 207 positive and 145 negative sequences in the test set. To address class imbalance, SMOTE (Synthetic Minority Over-sampling Technique) was applied to generate synthetic samples, while ENN (Edited Nearest Neighbors) was used to remove unrepresentative instances. This process produced a balanced training set, containing 773 sequences in each class, and a test set with 145 sequences per class.

Fig. 1B details the model construction, beginning with the tokenization of peptide sequences, followed by generating word and positional embeddings. These embeddings are then fed into the 30-layer ESM-2 Transformer, which extracts deep contextual features. The extracted features undergo an attention mechanism and are processed by a feed-forward network before being passed into a Gated Recurrent Unit (GRU), as depicted in Fig. 1C. The GRU leverages reset and update gates to capture long-term dependencies within the sequential data, ultimately classifying sequences as IL-4 inducing or non-IL-4 inducing.

Fig. 1D presents the fully connected layer, which is responsible for final classification. Fig. 1E illustrates the deployment of the PLM-IL-4 web server, offering a user-friendly platform where researchers can input peptide sequences in FASTA format and receive predictions regarding their IL-4 inducing potential. This platform is designed to streamline peptide analysis, providing efficient and accessible tools for researchers and practitioners.

2.1. Dataset

The dataset for this study was obtained from Meta-IL4 and downloaded via the IEDB. It includes 985 IL-4 inducing peptides and 774 non-IL-4 inducing peptides. The dataset was split into training and testing sets in an 8:2 ratio, resulting in 773 positive and 629 negative sequences in the training set, and 207 positive and 145 negative sequences in the test set. To alleviate class imbalance in the training set, SMOTE was initially applied to oversample the minority class, thereby increasing its size through synthetic samples. Subsequently, ENN was utilized to further clean the data by removing unrepresentative or potentially noisy samples. Notably, the minority class was augmented significantly due to synthetic sample generation, whereas the majority class slightly decreased in number due to the removal of potentially noisy or anomalous instances, resulting in a cleaner and more representative dataset overall. This combined approach effectively addresses class imbalance, reduces negative impacts from noise samples, and enhances model accuracy and generalization capacity. Finally, random undersampling was applied to ensure strict class balance in both classes. To facilitate a balanced evaluation and enable a more intuitive comparison of model performance metrics, random undersampling was also employed on the test set to achieve equal representation of each class.

Fig. 2 illustrates the distribution of protein sequence lengths within the dataset analyzed in this study. The data shows that the majority of protein sequences fall within the range of 10–25 amino acids. Notably, sequences with lengths of 10, 15, and 20 amino acids are more frequently represented, indicating that these lengths are predominant in our dataset. This distribution reflects typical protein fragment sizes found in biological systems, which is advantageous for training deep learning models. Training on sequences of these prevalent lengths is expected to enhance the model's generalization ability, making it more applicable to real-world biological data.

2.2. ESM-2

Feature extraction was performed using a 30-layer ESM-2 model, a state-of-the-art deep learning model specifically designed for protein sequence analysis. The ESM-2 model, based on the transformer architecture, effectively captures complex patterns and relationships within biological sequences. By leveraging deep learning, the ESM-2 model generates rich feature representations that improve the accuracy of downstream tasks, including protein function prediction, structural analysis, and interaction studies (Lv et al., 2025).

The ESM-2 model utilized in this study was pretrained on a large dataset of protein sequences, enabling it to learn comprehensive representations that generalize well to new, unseen data. The 30-layer architecture allows the model to capture both local and global sequence features, providing a robust foundation for subsequent classification tasks. Feature extraction involves processing each protein sequence through the ESM-2 model to obtain a fixed-dimensional embedding, which is then used as input for the GRU-based classifier described in the subsequent sections.

2.3. Model Architecture

In this study, we developed a deep learning model based on GRU to predict IL-4 inducing peptides. GRU is a type of Recurrent Neural Network (RNN) (Qiao et al., 2023) particularly effective for sequential data, making it ideal for tasks like time series analysis and natural language processing (Yuan et al., 2024). To enhance model performance and generalization, we optimized hyperparameters and employed various regularization techniques.

The model begins with an input layer, where the input dimension corresponds to the feature vector length, derived from protein sequences using the ESM-2 model. The GRU layer forms the core of the model, utilizing gating mechanisms to effectively capture long-term dependencies in sequential data. Multiple GRU layers are employed to further improve representational capacity (Chung, 2014).

Key components of each GRU layer include the number of units, hyperparameter-tuned between 50 and 300, and a configuration to return sequences (except for the final layer) to retain sequential information across the network. L2 regularization (including kernel, recurrent, and bias regularization) is used to mitigate overfitting. Furthermore, recurrent dropout is applied to randomly drop neurons during the training process, thereby improving the generalization capability of the model.

After each GRU layer, we employ batch normalization to standardize the data, reduce internal covariate shifts, and accelerate training while enhancing model stability. The batch normalization is defined by:

$$\widehat{\boldsymbol{x}}^{(k)} = \frac{\boldsymbol{x}^{(k)} - {\mu_B}^{(k)}}{\sqrt{({\sigma_B}^{(k)})^2 + \epsilon}} \tag{1}$$

Where $\mu_B^{(k)}$ and $\sigma_B^{(k)}$ arepresent the mean and standard deviation of the k-th feature in mini-batch B, and ϵ is a small constant used to avoid division by zero.

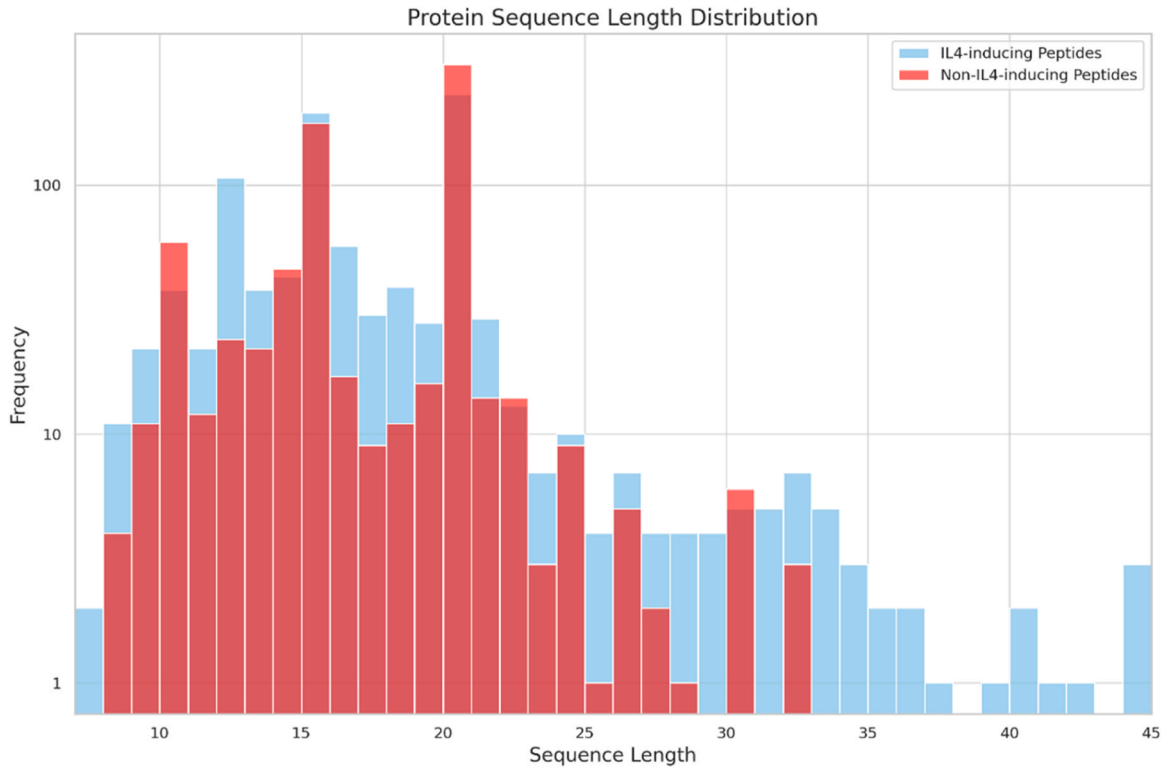


Fig. 2. Sequence lengths distribution of IL-4 inducing peptide and non-IL-4 inducing peptide.

Dropout layers were added after each GRU and dense layer, with dropout rates between 0.1 and 0.5, optimized through hyperparameter tuning (Srivastava, 2014). The purpose of dropout is to prevent overfitting by randomly deactivating a subset of neurons during training.

After the GRU layers, several dense layers were incorporated to extract features and carry out classification, with the number of neurons ranging from 32 to 256, determined through hyperparameter tuning. The output of a dense layer can be expressed as:

$$y^{(l)} = f\left(W^{(l)}x^{(l)}\right) + b^{l} \tag{2}$$

where f is the activation function, $W^{(l)}$ is the weight matrix, and b^{l} is the bias vector.

The output layer consists of a dense layer using a sigmoid activation function, which generates binary classification outcomes (0 or 1). The sigmoid function maps output values between 0 and 1, representing the predicted probability.

$$\hat{y} = \sigma(W_{out}h + b_{out}) \tag{3}$$

where σ is the sigmoid function, and W_{out} and b_{out} are the weight matrix and bias vector of the output layer.

During model compilation, both Adam and RMSprop optimizers were evaluated, and the optimal one was chosen through hyperparameter tuning. Binary cross-entropy was used as the loss function, with accuracy serving as the performance metric (Kingma et al., 2014). The binary cross-entropy loss is defined as:

$$loss = -\frac{1}{N}\sum_{i=1}^{N}(y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)) \tag{4}$$

where y_i represents the true labels, $\hat{y}_i$ represents the predicted probabilities, and N is the number of samples.

Hyperparameter tuning was conducted using the keras_tuner library, employing Bayesian optimization to identify the optimal parameter combinations, including the quantity of GRU layers, units per GRU layer, dropout rates, the count of dense layers, neurons per dense layer, and the type of optimizer used (Snoek et al., 2012; Jiao et al., 2024). To mitigate overfitting and enhance generalization during model training, several callback functions were utilized. EarlyStopping tracked validation loss and halted training when no further improvement was observed (Prechelt, 2002). LearningRateScheduler adjusted the learning rate adaptively based on the number of training epochs (Bengio, 2012). ReduceLROnPlateau reduced the learning rate when the validation loss stabilized, fostering better convergence (Zeiler, 2012).

2.4. Evaluation Metrics

To assess the model's performance, multiple widely used metrics were applied, including Precision (Pre), Recall (Rec) or Sensitivity (SE), F1 Score, Specificity (SP), Balanced Accuracy (BACC), Overall Accuracy (ACC), Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC) (Zhou et al., 2024; Wang et al., 2024). The definitions and corresponding formulas for these metrics are provided below:

$$ACC = \frac{TP + TN}{TP + FN + FP + TN} \tag{5}$$

$$SE = \frac{TP}{TP + FN} \tag{6}$$

$$SP = \frac{TN}{TN + FP} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$F1 = 2 \times \frac{Pre \times Rec}{Pre + Rec} \tag{9}$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{10}$$

In these equations, *TP* stands for true positives, *FP* for false positives, *TN* for true negatives, and *FN* for false negatives. The AUC evaluates the model's overall performance across various classification thresholds, yielding a single scalar value that encapsulates the classifier's effectiveness. These performance metrics provide a thorough assessment of the model's classification capability, aiding in identifying its strengths as well as areas needing further enhancement.

3. Results and discussion

3.1. Performance analysis

In Fig. 3, the bar chart presents a detailed comparison of the performance across six model architectures: LSTM+FC, GRU+FC, CNN+FC, and their respective variants integrated with attention mechanisms. The analysis reveals that the GRU+FC model consistently outperforms the others, demonstrating superior predictive accuracy and robustness, positioning it as the most favorable model in this study for predicting IL-4 inducing peptides.

The incorporation of attention mechanisms, particularly in the GRU+FC+Attention model, results in improvements in specific metrics such as Specificity (SP) and Precision. Despite these enhancements, the GRU+FC model maintains superior overall performance across most metrics, including Accuracy (ACC), Sensitivity (SN), Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC). This analysis underscores the effectiveness of the GRU+FC architecture, establishing it as the optimal model for predicting IL-4 inducing peptides and supporting future research in immunotherapy and vaccine development.

3.2. Comprehensive analysis of model uncertainty, stability, and performance across subsets

Using entropy as a measure, we analyzed the model's uncertainty (see Fig. 4A). The results indicate that most samples have an entropy value close to zero, suggesting high confidence in the model's predictions. A small subset of samples exhibits entropy values ranging from 0.1 to 0.7, indicating the presence of more complex features, which cause the model to display appropriate uncertainty during predictions (Gal and Ghahramani, 2016; Hendrycks and Gimpel, 2016).

Fig. 4B presents a comparison between the original sample and a perturbed sample. The perturbed sample was generated by adding random noise with a mean of 0 and a standard deviation of 0.1 to a test dataset sample, and the predictions for both the original and perturbed samples were recorded and compared. Despite the added noise, the overall trends remained consistent between the original and perturbed samples, demonstrating the model's robustness and stability in maintaining accurate predictions even with minor data variations (Lakshminarayanan et al., 2017).

Fig. 4C evaluates the performance of the trained GRU model across different test subsets using metrics such as Accuracy, Precision, Recall, F1 Score, Matthews Correlation Coefficient (MCC), and ROC-AUC. The analysis revealed some variations in performance across the subsets, with Subset 2 demonstrating superior performance in most metrics, particularly in Precision and Recall. Although MCC values were slightly lower, the model's performance remained stable, particularly in key metrics such as Accuracy and ROC-AUC.

Lastly, Fig. 4D presents a radar chart comparing the PLM-IL-4 model with the Meta-IL4 model across multiple performance metrics. The radar chart shows that the PLM-IL-4 model outperforms the Meta-IL4 model in most metrics, particularly in Specificity (SP), Matthews Correlation Coefficient (MCC), and AUC. These results indicate that the PLM-IL-4 model not only has an advantage in prediction accuracy but also exhibits greater robustness and reliability across critical performance metrics.

3.3. Performance evaluation of the PLM-IL-4 model against other predictors

The PLM-IL-4 model demonstrates significant improvements across several key performance metrics compared to the Meta-IL4 model, underscoring its superior predictive capabilities and robustness. Specifically, the accuracy of the PLM-IL-4 model increases by 2.7 percentage points, indicating stronger overall prediction performance and more reliable classification results when identifying IL-4 inducing peptides. Additionally, the specificity of the PLM-IL-4 model improves by 9.08 percentage points, highlighting its enhanced ability to accurately

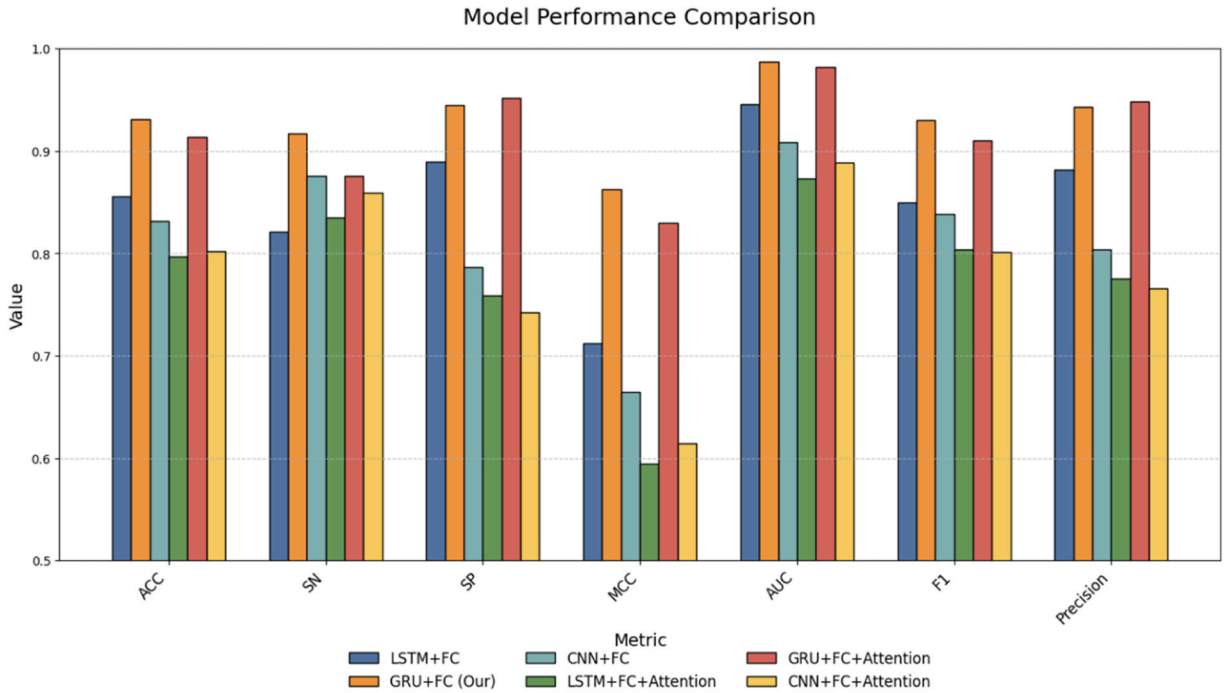


Fig. 3. The bar chart presents a comparison of six model architectures: LSTM+FC, GRU+FC, CNN+FC, and their corresponding variants with attention mechanisms.

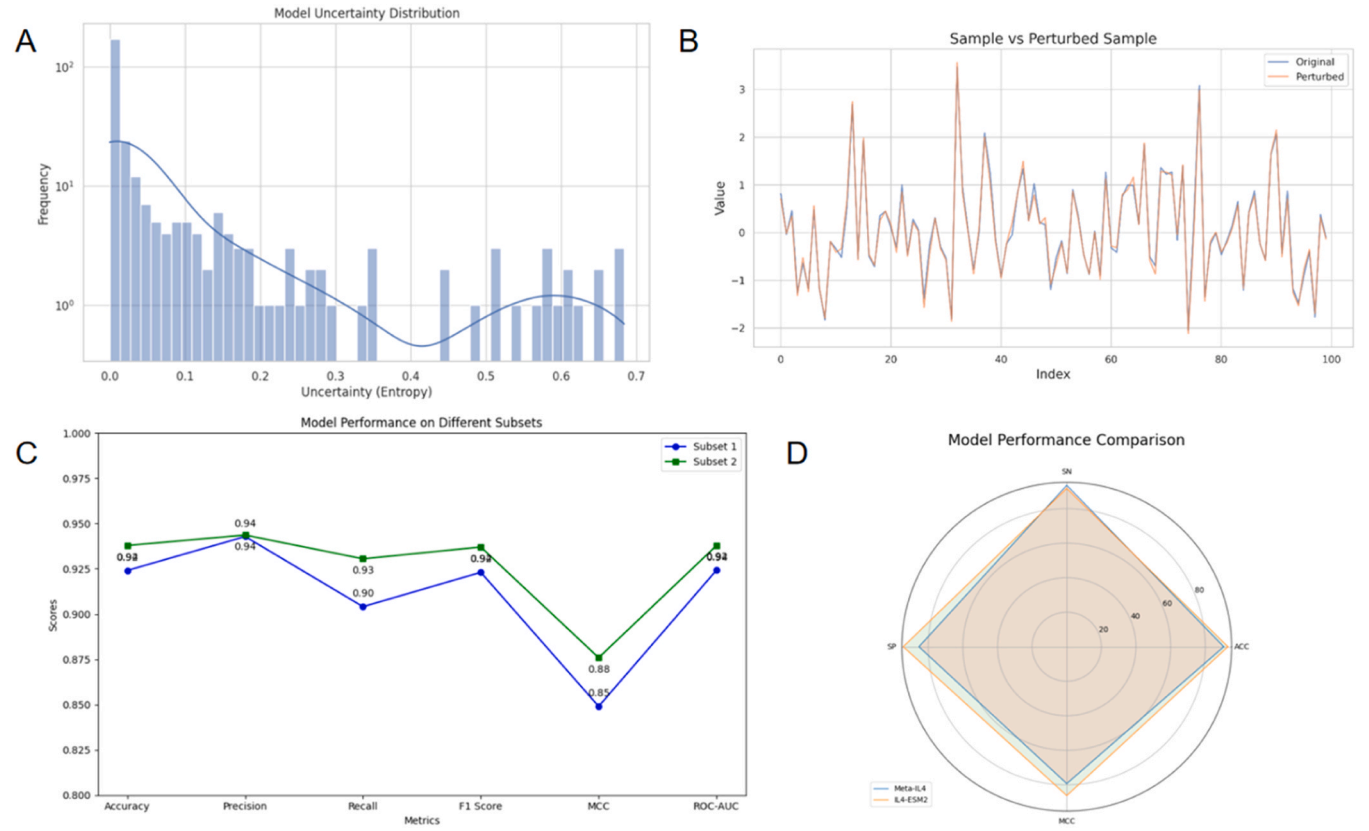


Fig. 4. Comprehensive analysis of model performance and uncertainty. (A) Model Uncertainty Distribution based on entropy values, indicating high confidence in most predictions with some samples showing appropriate uncertainty. (B) Comparison between the original and perturbed samples, demonstrating the model's stability in predictions despite the introduction of noise. (C) Performance evaluation of the GRU model on different test subsets, showing Subset 2's superior performance in most metrics, especially in Precision and Recall. (D) Radar chart comparing the PLM-IL-4 model with Meta-IL4 across multiple performance metrics, highlighting PLM-IL-4's overall superior performance, particularly in Specificity (SP), Matthews Correlation Coefficient (MCC), and AUC.

identify true negatives and significantly reduce false positives, which is crucial for improving reliability in distinguishing between different classes. Table 1

The model also shows substantial improvement in the MCC, which increases by 6.94 percentage points, indicating more balanced performance across true positives, true negatives, false positives, and false negatives. This further supports the PLM-IL-4 model's capacity to deliver more reliable and interpretable classification outcomes. Furthermore, the Area Under the Curve (AUC) score rises by 0.96 percentage points, demonstrating the model's superior ability to distinguish between classes across various decision thresholds.

The radar chart in Fig. 4D visually compares the performance of PLM-IL-4 and Meta-IL4 across several metrics. The PLM-IL-4 model shows clear superiority in terms of specificity, MCC, and AUC, suggesting that it not only provides more accurate predictions but also maintains robustness and balance in classification results. This comparison illustrates how PLM-IL-4 achieves a more stable and reliable prediction profile, enhancing its practical value in predicting IL-4 inducing peptides. In conclusion, these improvements establish the PLM-IL-4 model as a more reliable and robust tool, offering valuable insights for future research in immunotherapy and vaccine development.

Table 1
Performance comparison between the PLM-IL-4 and Meta-IL4 models.

	ACC	SN	SP	MCC	AUC	F1	precision
Meta-IL4	90.7	93.54	85.41	79.3	*	*	*
PLM-IL−4	93.1	91.72	94.48	86.24	98.67	93.01	94.33

"Note: Values marked with * indicate data that is not disclosed in this article."

3.4. Webserver construction

To facilitate practical applications, we developed a user-friendly web server for predicting IL-4 inducing peptides using the PLM-IL-4 model. The PLM-IL-4 web server, along with data resources and code, is freely accessible at http://www.bioai-lab.com/PLM-IL4. Users can input peptide sequences to receive predictions regarding their IL-4 inducing potential. The server features a simple interface that supports both single and batch sequence submissions. It efficiently processes requests and returns results, including probability scores and key metrics such as accuracy and specificity. This tool aims to assist researchers in conducting bioinformatics analyses, ultimately supporting immunotherapy and vaccine development.

4. Conclusion

In conclusion, the integration of the 30-layer ESM-2 model for feature extraction with the GRU-based architecture has significantly improved the prediction of IL-4 inducing peptides. The proposed model achieves superior performance, including an AUC of 0.9867 and an accuracy of 93.1 %, outperforming existing methods such as IL-4-Pred and Meta-IL4. These results demonstrate the effectiveness of leveraging advanced deep learning techniques for protein sequence analysis and underscore their potential to enhance immunotherapy and vaccine development.

The robustness and reliability of the model were validated through extensive experimentation and rigorous evaluation, confirming its suitability for practical applications in computational immunology. Future work will focus on further refining the model and exploring its

applicability to other cytokines and immune responses, paving the way for more comprehensive and accurate predictive tools in bioinformatics.

Funding

The work is supported by the National Natural Science Foundation of China (No. 62450002), the Science and Technology special fund of Hainan Province ZDYF2024GXJS018 and the Innovation Platform for "New Star of South China Sea" of Hainan Province under Grant No. NHXXRCXM202306.

CRediT authorship contribution statement

Wei Leyi: Supervision, Project administration, Formal analysis. **Zhang Qingchen:** Supervision, Resources, Data curation. **Meng Yajie:** Project administration, Methodology, Investigation. **Xu Junlin:** Investigation, Funding acquisition, Formal analysis. **Zhang Zilong:** Writing – review & editing, Project administration, Methodology, Funding acquisition. **Yan Shankai:** Writing – review & editing, Supervision, Funding acquisition. **Liu Ruiqi:** Writing – original draft, Visualization, Supervision, Investigation, Data curation. **Cui Feifei:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition. **Zou Quan:** Visualization, Methodology, Investigation.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

References

Agarwal, S., Rao, A.J.I., 1998. Modulation of chromatin structure regulates cytokine geneexpression during T cell differentiation. 9 (6), 765–775.
Ali M.et al., 2019. Induction of neoantigen-reactive T cells from healthy donors. 14 (6), 1926-1943.
Ao, C., Jiao, S., Wang, Y., Yu, L., Zou, Q., 2022. Biological Sequence Classification: A Review on Data and General Methods. Art no. 0011 Res., Rev. 2022, 0011. https://doi.org/10.34133/research.0011.
Barneh, F., Nazarian, A., Nadoshan, R.M., Bagheri, K.P., 2024. A Novel In silico Filtration Method for Discovery of Encrypted Antimicrobial Peptides, 2024 Curr. Bioinforma., Artic. 19 (5), 502–512. https://doi.org/10.2174/0115748936274103231114105340.
Bengio, Y., 2012. Practical recommendations for gradient-based training of deep architectures. Neural networks: Tricks of the trade: Second edition. Springer, pp. 437–478.
Bhasin, M., Raghava, G.J.J.o.b. , 2007. A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes. 32, 31–42.
Blum, J.S., Wearsch, P.A., Cresswell, 2013. P.J.A.r.o.i. Pathways of antigen processing 31 (1), 443–473.
Chen, L., et al., 2016. Genetic drivers of epigenetic and transcriptional variation in human immune cells. 167(5),1398-1414. e1324.
Chung, J., et al., 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling.
Dhanda, S.K., Vir, P. and Raghava, G.P.J.B.d., 2013a. Designing of interferon-gamma inducing MHC class-II binders. 8:1-15.
Dhanda, S.K., et al., 2013b. Prediction of IL4 inducing peptides. 2013(1):263952.
Fu, X., et al., 2024. Hyb_SEnc: An Antituberculosis Peptide Predictor Based on a Hybrid Feature Vector and Stacked Ensemble Learning. IEEE/ACM Trans. Comput. Biol. Bioinforma. 1–17. https://doi.org/10.1109/TCBB.2024.3425644.
Gal, Y., Ghahramani, Z., 2016. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. international conference on machine learning. PMLR, pp. 1050–1059.
Dhanda, S.K., et al., 2019. IEDB-AR: immune epitope database—analysis resource in 2019. 47(W1), W502-W506.
Gause, W.C., Wynn, T.A., Allen, J.E.J.N.R.I., 2013. Type 2 immunity and wound healing: evolutionary refinement of adaptive immunity by helminths. 13 (8), 607–614.
Geng, A., et al., 2025. ACP-CLB: An Anticancer Peptide Prediction Model Based on Multichannel Discriminative Processing and Integration of Large Pretrained Protein Language Models. J. Chem. Inf. Model. https://doi.org/10.1021/acs.jcim.4c02072.
Germain, R.N.J.N.r.i., 2002. T-cell development and the CD4–CD8 lineage decision. 2 (5), 309–322.
Hassan, M.T., Tayara, H., Chong, K.T.J.M., 2023. Meta-IL4: An ensemble learningapproach for IL-4-inducing peptide prediction. 217, 49–56.
Hendrycks, D. and Gimpel, K.J.a.p.a. 2016. A baseline for detecting misclassified and out-of-distribution examples in neural networks.
Jespersen, M.C., et al., 2017. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. 45 (W1), W24–W29.
Jiao, S., Ye, X., Sakurai, T., Zou, Q., Liu, R., 2024. Integrated convolution and self-attention for improving peptide toxicity prediction. Bioinformatics 40 (5). https://doi.org/10.1093/bioinformatics/btae297.
Kumar, M., Verma, R., Raghava, G.P.J.J.o.B.C., 2006. Prediction of mitochondrialproteins using support vector machine and hidden Markov model. 281 (9),5357–5363.
Lakshminarayanan, B., Pritzel, A., Blundell, 2017. C.J.A.i.n.i.p.s. Simple and scalable predictive uncertainty estimation using deep ensembles. 30.
Lambrecht, B.N., Hammad, H.J.N.i., 2015. The immunology of asthma. 16 (1), 45–56.
Kingma, D.P., Ba, Adam, J.J.a.p.a., 2014. A method for stochastic optimization.
Lee, G.R., et al., 2006. T helper cell differentiation: regulation by cis elements and. epigenetics. 24 (4), 369–379.
Li, Y., et al., 2024. msBERT-Promoter: a multi-scale ensemble predictor based on BERT pre-trained model for the two-stage prediction of DNA promoters and their strengths, 2024/05/30 BMC Biol. 22 (1), 126. https://doi.org/10.1186/s12915-024-01923-z.
Li, Y.J., Niu, M.T., Zou, Q., 2019. ELM-MHC: An Improved MHC Identification Method with Extreme Learning Machine Algorithm (Mar). J. Proteome Res. 18 (3), 1392–1401. https://doi.org/10.1021/acs.jproteome.9b00012.
Liu, R., Fu, X., Yan, S., Zhang, Z., Cui, F., 2023. AIPPT: Predicts anti-inflammatory peptides using the most characteristic subset of bases and sequences by stacking ensemble learning strategies, 5-8 Dec. 2023 2023 IEEE Int. Conf. Bioinforma. Biomed. (BIBM) 23–29. https://doi.org/10.1109/BIBM58861.2023.10385565.
Lv, J., et al., 2025. iBitter-GRE: A Novel Stacked Bitter Peptide Predictor with ESM-2 and Multi-View Features, 2025/04/15/ J. Mol. Biol. 437 (8), 169005. https://doi.org/10.1016/j.jmb.2025.169005.
Ma, C.S., et al., 2012. The origins, function, and regulation of T follicular helper cells. 209(7), 1241–1253.
Mehta, K., Vyas, P., Mujawar, S., Hazam, P.K., Vyas, A., 2023. Design and In-silico Screening of Short Antimicrobial Peptides (AMPs) as Anti-Tubercular Agents Targeting INHA, 2023 *Curr. Bioinforma.,* Artic. 18 (9), 715–736. https://doi.org/10.2174/1574893618666230419081901.
Prechelt, L., 2002. Early stopping-but when? Neural Networks: Tricks of the trade. Springer, pp. 55–69.
Pulendran, B., Ahmed, R.J.N.i., 2011. Immunological mechanisms of vaccination. 12 (6),509–517.
Qiao, B., Wu, Z., Ma, L., Zhou, Y., Sun, Y., 2023. Effective ensemble learning approach for SST field prediction using attention-based PredRNN. Art no. 171601 *Front. Comput. Sci.,* Artic. 17 (1), 171601. https://doi.org/10.1007/s11704-021-1080-7.
Rao, B., Han, B., Wei, L., Zhang, Z., Jiang, X., Manavalan, B., 2024. CFCN: An HLA-peptide Prediction Model based on Taylor Extension Theory and Multi-view Learning, 2024 *Curr. Bioinforma.,* Artic. 19 (10), 977–990. https://doi.org/10.2174/0115748936299044240202100019.
Snoek, J., Larochelle, H. and Adams, R.P.J.A.i.n.i.p.s. 2012. Practical bayesian optimization of machine learning algorithms. 25.
Vyas, J.M., Van der Veen, Ploegh, H.L.J.N.R.I., 2008. The known unknowns of antigenprocessing and presentation. 8 (8), 607–618.
Srivastava, N., et al., 2014. Dropout: a simple way to prevent neural networks from overfitting. 15 (1), 1929–1958.
Tarca, A.L., et al., 2007. Machine learning and its applications to biology. 3 (6), e116.
VitaR.et al., 2015. The immune epitope database (IEDB). 3.0 43 D1 D405 D412.
Wang, Y., Zhai, Y., Ding, Y., Zou, Q., 2024. SBSM-Pro: support bio-sequence machine for proteins, 2024/10/22 Sci. China Inf. Sci. 67 (11), 212106. https://doi.org/10.1007/s11432-024-4171-9.
Xiao, C., Zhou, Z., She, J., Yin, J., Cui, F., Zhang, Z., 2024. PEL-PVP: Application of plant vacuolar protein discriminator based on PEFT ESM-2 and bilayer LSTM in an unbalanced dataset, 2024/10/01/ Int. J. Biol. Macromol. 277, 134317. https://doi.org/10.1016/j.ijbiomac.2024.134317.
Yan, C., Geng, A., Pan, Z., Zhang, Z., Cui, F., 2024a. MultiFeatVotPIP: a voting-based ensemble learning framework for predicting proinflammatory peptides. Brief. Bioinforma. 25 (6). https://doi.org/10.1093/bib/bbae505.
Yan, K., Lv, H., Shao, J., Chen, S., Liu, B., 2024b. TPpred-SC: multi-functional therapeutic peptide prediction based on multi-label supervised contrastive learning. Art no. 212105 *Sci. China-Inf. Sci.*, Artic. 67 (11), 212105. https://doi.org/10.1007/s11432-024-4147-8.
Yuan, J., Wang, Z., Pan, Z., Li, A., Zhang, Z., Cui, F., 2024. DPNN-ac4C: a dual-path neural network with self-attention mechanism for identification of N4-acetylcytidine (ac4C) in mRNA. Bioinformatics 40 (11), btae625. https://doi.org/10.1093/bioinformatics/btae625.
Zeiler, M.D.J.a.p.a., 2012. Adadelta: an adaptive learning rate method.

Zhou, Z., et al., 2024. PSAC-6mA: 6mA site identifier using self-attention capsule network based on sequence-positioning, 2024/03/01/ Comput. Biol. Med. 171, 108129. https://doi.org/10.1016/j.compbiomed.2024.108129.

J. Zhu and W.E.J.C. r Paul, Heterogeneity and plasticity of T helper cells, vol. 20, no. 1, pp. 4-12, 2010.

J. Zhu, H. Yamane, and W.E.J.A. r o i Paul, Differentiation of effector CD4 T cell populations, vol. 28, no. 1, pp. 445-489, 2009.