UtilitasMathematica
ISSN 0315-3681 Volume 120, 2023

Effective of Various Vaccines on Antibody Response and Genetic Immune Using Deep Learning Method

Nuha Hisham Mohammed[1], Sura Zaki Al-Rashid[2]
[1]Department of Information Technology, University of Babylon, Babylon, Iraq,
nuhahm.sw.msc@student.uobabylon.edu.iq
[2]Department of Information Technology, University of Babylon, Babylon, Iraq,
sura_os@itnet.uobabylon.edu.iq

Abstract

The COVID-19 epidemic has affected daily life on a global scale. Many research teams from major pharmaceutical companies and university institutions around the world have been developing vaccines since the beginning of the pandemic. The effectiveness, acceptability, and results of vaccinations are influenced by gender. The SARS-CoV-2 mRNA vaccines were released on the market in reaction to the Covid-19 public health emergencies. There is no history of using mRNA vaccines to treat infectious diseases. The numerous modifications of the vaccine's mRNA work to protect it from cellular defenses, lengthen its biological half-life and increase the creation of spike protein. In this paper, we propose a novel model to predict the antibody response based on deep learning by proposing a Convolutional Neural Network (CNN) model. The proposed system consists of several stages, where the GSE201533 dataset which used, containing 26,370 features is first split into two sets, then the missing value and normalization were applied as a preprocessing stage, then the best features were selected using three techniques (Mutual information, Chi-square, and Analysis of Variance (ANOVA)). Then the selected features were classified using the proposed CNN model. The proposed CNN contributed to raising the accuracy of the model and reducing the time required for prediction. The experimental results indicate an accuracy rate of 100% in all cases.

Keywords: COVID-19, Convolutional Neural Network (CNN), mRNA, Analysis of Variance (ANOVA), Chi-square, Mutual Information (MI).

I. INTRODUCTION

The COVID-19 pandemic has had a profound global impact, with over 260 million cases and 5 million fatalities reported by the World Health Organization (WHO) [1]. The COVID-19 pandemic is recognized as the most severe of this century and has had a substantial influence on public life [2]. Vaccination is still the most efficient public health approach to contain the COVID-19 virus and reduce the growth in the rates of concurrent infectious disease death, morbidity, and disability [3]. The World Health Organization (hereafter, WHO) estimates that immunization prevents approximately 3 million deaths and 75,000 cases of disability each year. As a result, vaccination is being heavily considered a potent containment technique, especially in light of its prior accomplishments [4]. The goal of vaccination is to imitate the immune reaction to a natural infection using non-pathogenic

material, hence granting immunity in the case of pathogen exposure. The use of attenuated viral vaccines as well as whole organisms has been the main method used to achieve this goal [5]. It has proven more technically difficult to employ viral fragments or their protein derivatives, sometimes known as "subunit vaccinations" [6]. In any case, the deployment of any vaccination program is predicated on the implicit premise that the vaccine delivers the impacts of a "benign infection", thus avoiding the detrimental effects of a true infection, boosting the immune system versus future exposure [7]. The majority of the COVID-19-related material implies that the immune reaction to mRNA-depended vaccination is comparable to that of natural infection [8]. Same antigen-specific antibodies in diverse donations from the same genetic components and typically as an outcome of comparable mechanisms of antigen recognition are referred to as a common or public antibody reaction. Understanding how people react to specific antigens is crucial for identifying the molecular characteristics of recurrent antibodies among the wide antibody repertoire at the population level, as well as for the creation of successful vaccinations [9]. The objective of machine learning and deep learning is to find groups of characteristics (i.e., biological elements) that should forecast a result. The result indicates the vaccine "labels" that had to be established during the training procedure. The purpose of classification models is to forecast a discrete class of labels as shown in "Fig. 1" [10]. These days, a variety of computational areas apply machine learning (ML) and deep learning (DL) algorithms because of their potent performance [11]. The benefit of the machine and deep learning in healthcare is its capacity to generate enormous datasets that are beyond the capabilities of humans, and then reliably translate analysis of those datasets into clinical insights that assist physicians in planning and delivering care, ultimately leading to better outcomes, lower healthcare costs, and higher patient satisfaction [12].

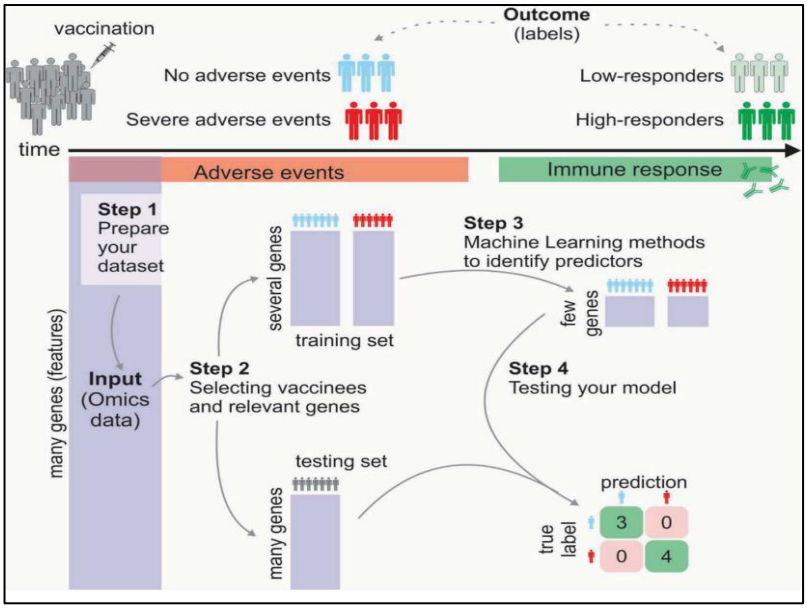

Fig. 1. The essential four processes for determining discriminatory markers for reactogenicity and vaccine-induced immunity [10].

The remainder of the paper is structured as follows; section 2 discusses related works of proposed genetic immune and antibody response analyses. Section 3 describes the proposed methodology, while Section 4 analyzes the experimental results. Section 5 contains the conclusion portion.

II. RELATED WORKS

Since the emergence of the Corona epidemic in the year 2019, many researchers have been conducted related to vaccines and their effect on humans. This section reviews a group of studies that were conducted based on deep learning to predict the body's response to these vaccines. J. Chen et al. [13] demonstrated that a significant number of the 462 mutations on the receptor-binding domain (RBD) will disrupt the binding of S protein and antibodies and affect the feasibility and reliability of antibody treatments and vaccines by combining genetics, biophysics, deep learning, and algebraic topology. There is a list of 31 mutants that disrupt antibodies, and many other disruptive mutations are described as well. By comparing their predictions to those of more than 1400 deep mutations on the S protein RBD, they were found to be accurate. S. Seneff et al. [14] provided proof that vaccination results in a serious disruption of sort I interferon signaling, which has a variety of harmful effects on human health. Spike protein-containing exosomes in large numbers and essential microRNAs are released into the bloodstream by immune cells that have ingested vaccine nanoparticles, and these exosomes trigger a signaling reaction in recipient cells at remote places. V. Rustagi et al. [15] developed models to predict the number of people that pass away with COVID infection depending on doses of vaccine (partial or complete vaccination). Karl Pearson's coefficient was used to analyze the dataset's estimated variance. A further quartic polynomial regression model was developed for COVID-19 cases and immunization dosages. Depending on the number of vaccination doses garnered, this predictor model aids in estimating the cases of COVID-19-related deaths and determining the susceptibility to COVID-19 infection. M. Kurano et al. [16] investigated whether antibody testing may help forecast the maximal severity of the illness in COVID-19 individuals as well as the SARS-CoV-2 antibody responses that are associated with the COVID-19 infection's peak severity in its early stages. In accordance with the severity of the condition, they divided the patients into four groups. According to the findings, physicians may be able to detect non-immunized COVID-19 patients who require admission to a critical care unit by using antibody testing. S. Haggenburg et al. [17] Examined whether a third dosage of mRNA-1273 vaccination results in higher neutralizing antibody levels in immunosuppressed patients with hematologic malignancies that are comparable to those attained in healthy people following the recommended 2-dose mRNA-1273 immunization regimen. E. Wang and A. Chakraborty [18] merged deep mutational scanning data and assessments of coronavirus spike sequences and structures to create SARS-CoV-2 variant antigens. These antigens bind to ACE2 and are stable, making them feasible alternatives to vaccination or infection with the wild-type virus. Using a mathematical model of affinity maturation, the immune reaction to immunization with different permutations of the hypothesized antigens was studied (AM). According to the research, a combination of proteins is more probable to encourage the development of higher antibody titer of antibodies that are specifically directed against SARS-coV-2 variants. Y. Wang et al. [19] showed that the typical (public) reactions to the various domains of the spike protein were very diverse by examining immunoglobulin D and V gene usages, area H3 sequences that determine complementarity, and somatic hypermutations. Additionally, they utilized these sequences to train a deep learning system to precisely differentiate both human antibodies to the SARS-CoV-2 spike protein as well as

the influenza hemagglutinin protein. S. Shan et al. [20] put forth a geometric deep-learning method that effectively boosts antibody affinity to produce broader and more effective neutralizing activity versus SARS-CoV-2 variants. Our improved antibodies may be made into antibody drug candidates for both existing and new variations. These outcomes demonstrate the effectiveness of our deep learning strategy in antibody optimization and suggest future directions for its use in protein engineering. H. Shon et al. [21] developed a classification method depending on deep learning and indicate its application to gene expression data gathered from patients with stomach cancer. Data from 60,483 genes from 334 stomach cancer patients were evaluated using principal component analysis, heat maps, and the convolutional neural network (CNN) method from The Cancer Genome Atlas. They combined clinical data with RNA-seq gene expression data, searched for important genes, and analyzed them using CNN deep learning. They attained an accuracy of 50.51% for vital status and 95.96% for sample type. J. Zrimec et al. [22] use deep learning on more than 20,000 mRNA datasets to study the genetic regulatory network governing the amount of mRNA in 7 model species, ranging from bacteria to humans. With up to 82% of the variance of transcript levels recorded in the gene regulatory structure, we can estimate mRNA abundance in all species simply from DNA sequence.

III. METHODOLOGY

Deep learning methods are used in this study to analyze genetic immune and antibody responses. "Fig. 2" depicts the proposed system. The proposed system is separated into three stages, which we will go over in detail later. This is the dataset description. The following step is to extract features using (mutual information, chi-square, and ANOVA). Finally, CNN-depended intelligent classification algorithms are utilized to predict the genetic immune and antibody responses.

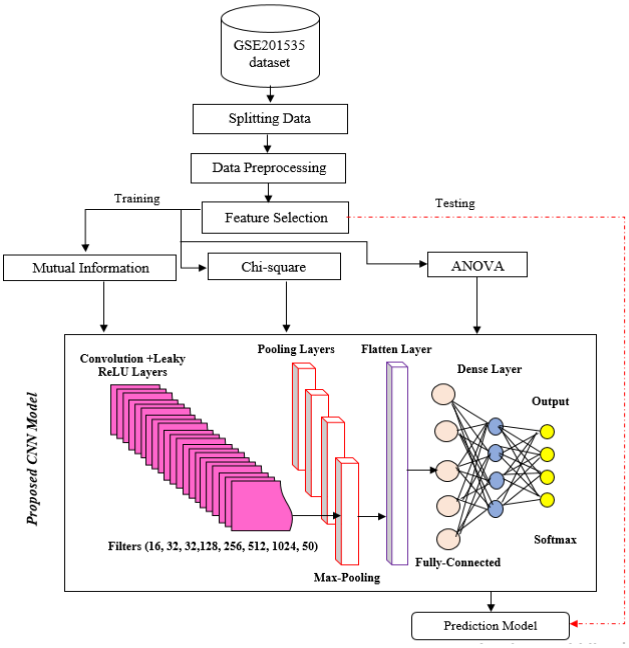

Fig. 2. An overall diagram of the proposed schema.

A. Dataset Description

We have used a GSE201533 dataset available on: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE201535, which consists of 161 samples (any row) and 26370 columns (feature). The genes of each blood sample are stored separately in a text file compressed in .zip extension for each dataset. These data should be extracted and stored properly in a tabular structure. Each text file is a record of a blood sample. Which consists of two columns. The first column is for the gene name. And the second column is for the level of the biomarker of that gene in the blood sample, as shown in "Fig. 3".

```
A1BG      32
A1BG-AS1          42
A1CF      3
A2M       79
A2M-AS1   35
A2ML1     8
A2MP1     44
A3GALT2   14
A4GALT    0
A4GNT     4
AA06      0
AAAS      1485
AACS      455
AACSP1    0
AADAC     0
AADACL2   0
AADACL2-AS1       0
```

Fig. 3. A sample of the original gene data in .txt file

Each file name holds important information of the serial number of the experiment (the session of blood drawn), the vaccine name if the vaccine dose is the first or second dose, the person's ID, and the day the blood is drawn. The generated dataset file consists of representing the blood samples drawn, and columns representing the features of each blood sample.

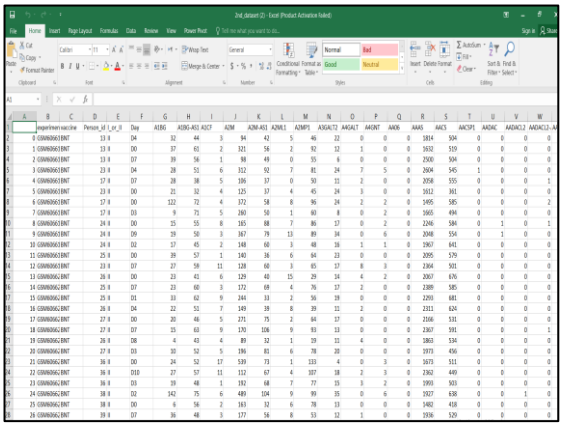

Fig. 4. Dataset after the reading process

B. Data Splitting Stage

The supplied data is divided into two halves, typically for the purpose of cross-validation. The first part of the dataset is utilized to generate a prediction model, whereas the second part of the dataset is utilized to assess the model's effectiveness. Cross-validation procedures are one way for ensuring

proper generalization and preventing overtraining. To obtain the final model outcomes, the dataset must be divided into two subsets: one for training (70%) and the other for testing (30%). To produce a reliable and consistent approximation of the model's outcomes, cross-major validation is used [23].

C. Preprocessing Stage

This stage includes some of the processes as follows:

1) Missing Value: In gene expression datasets, they may include missing values for some features [24]. There are several ways to solve the missing values problem such as guessing the missing values, ignoring the missing values, and removing data objects [25].

2) Normalization: One of the initial processing steps for processing the dataset that is applied before the data is used, such as increasing or decreasing the range of values. Normalization is convenient and useful in dataset problems that depend on classification, the conversion of feature values for a specific and small range such as 0 to 1. There are many ways to normalize, such as Z-normalization and Min-Max Normalization. Min-Max Normalization is a linear transformation technique used in processors in which preserving the relationship between the original dataset is important. In addition, it is considered one of the simple techniques that are suitable for the dataset within predefined limits [26][27]. Normalization is done according to "Eq. (1)":

$$x\ new = \frac{x - \min(x)}{\max(x) - \min(x)} \qquad (1)$$

Where x new represent normalized x.

D. Feature Selection Stage

For gene immune and antibody response analysis, feature selection and rating are essential. Finding the best "k" features out of all the feasible characteristics involves the process of extracting their scores. A feature's frequency is counted in training positive and negative class samples separately, and a function of both is then obtained. There are numerous parameters that need to be tracked in order to study gene immune and antibody response, some of which will be helpful while others may not be. Eliminating unnecessary features improves accuracy and reduces computing time, resulting in improved performance. There are three techniques of feature selection were used in the proposed model.

1) Select features based on Mutual Information: Mutual information is frequently used in feature selection because it can identify non-linear interactions between numerous variables. The majority of mutual information-based feature selection methods use the two formulas "Eq. 2" and "Eq. 3", which compute the relevance of a feature subset and redundancy, respectively.

$$\text{Red} = \text{MI}\ (s_1 . s_2 . \ldots . s_m) \qquad (2)$$

$$\text{Rel} = \text{MI}\ (S . C) \qquad (3)$$

where C is the class label, S is the feature set, which contains m features $s_1 . \ldots . s_m$. By eliminating all unnecessary and superfluous characteristics, feature selection aims to provide an

ideal feature subset. Hence, the best feature subset minimizes the quality measure specified in "Eq. 4".

$$F = -\alpha * \text{Rel} + (1 - \alpha) * \text{Red} \quad (4)$$

where is used to regulate how much relevance and redundancy contribute to the fitness score [28].

Select features based on Chi-square: The chi-squared test, which assesses deviation from the predicted distribution when the feature event is independent of the class value, is a numerical analysis. True positives (TP), false positives (FP), true negatives (TN), false negatives (FN), probability of a no. of positive instances (P_pos), and probability of the no. of negative cases (P_neg) are some of the metrics used to compute the chi-square value [29].

$$chi - square = t\left(t_p(t_p + f_p)p_{pos}\right) + t\left(f_n(f_n + t_n)p_{pos}\right) + t\left(f_p(t_p + f_p)p_{neg}\right) + t\left(t_n(f_n + t_n)p_{neg}\right) \quad (5)$$

3) Select features based on ANOVA: For each continuous predictor, it was intended to run an ANOVA F-test to determine whether or not all of the various classes of Y have the same mean as X. Applying the following notation:

$$s_j^2 = \sum_{i=1}^{N_j}\left(X_{ij} - \bar{X}_j\right)^2/\left(N_j - 1\right) \quad (6)$$

$\bar{x}$: The grand mean of predictor X:

$$\bar{x} = \sum_{j=1}^{J} N_j\bar{X}_j/N \quad (7)$$

The aforementioned notations are depending on pairs of (X, Y). Then, using the F statistic, the p-value is determined by p-value = Prob {F (J-1, N-J) >F}: where,

$$F = \frac{\sum_{j=1}^{J} N_j(\bar{X}_j - \bar{X})/(J-1)}{\sum_{j=1}^{J}(N_j - 1)s_j^2/(N-1)} \quad (8)$$

A random variable called F (J-1, N-1) has degrees of freedom J-1 and N-J and follows the F distribution. A predictor's p-value should be set to zero if the denominator is zero. By ranking the predictor in ascending order depending on the p-value, the predictor is ranked. If there are ties, sort by F in order of importance first, then if there are still ties, sort by N from lowest to highest [30].

Prediction stage using the proposed CNN model

Due to its extensive network depth, CNN is categorized as a deep neural network. With each layer, CNN gains the ability to recognize a wider variety of things. Images of various resolutions are processed via image processing, with the output from each image processed and used as input to the next layer [31]. Feature selection techniques (Mutual information, Chi-square, and Analysis of Variance (ANOVA)) and classification are two distinct parts of the proposed CNN architecture. The convolutional, Leaky Rectified Linear Unit (LeakyReLU), and pooling are the three operations following the feature selection portion [32].

Convolutional layer: Among filters and data as input. To build a feature map, the filter will travel about in the data and connect the input and filter value with a "dot" operation. Optimization of the convolutional layer using modified filter size, stride, and zero padding [33].

LeakyReLU: It is an effort to address the fading ReLU issue. When x<0, instead of being zero, a leaky ReLU will have a slight negative slope (of 0.01, or so). The equation in Eq. (2), where α is a tiny constant, is computed using the function [34].

$$f(x) = 1(x < 0)(\alpha x) + 1(x >= 0)(x) \quad (9)$$

3) Pooling Layer: The convolutional layer's output is received by the pooling layer, which also minimizes the amount of data on this layer. Filters of various sizes and strides make up the pooling layer, which travels over the feature map area. Max pooling and mean pooling are the pooling layers that are frequently used in their application [35].

4) Fully Connected Layer: Processing of the feature maps from the feature selection layers is done by the fully connected layer. In contrast to the convolutional layer, which links neurons just to certain parts of the input, the fully connected layer connects all neurons to enable linear data classification, and the fully connected layer flattens multidimensional feature map arrays into dimensional arrays [36].

5) Softmax Activation: A different kind of Logistic Regression called Softmax Classifier may classify more than two classes. The output of the last layer may be transformed to its underlying probability distribution utilizing Softmax. Softmax has the advantage that the output probability can be between 0 and 1 and that the sum of the probabilities is 1 [37].

The suggested CNN model comprised 27 layers as follows:

- The nine-layer Convolutional Neural Network (CNN).
- 8 Leaky ReLU layers.
- 6 Max-pooling layers.
- 1 Flatten layer.
- 3 Dense layer.

These layers are described in further depth in "Table 1".

TABLE I. THE SUGGESTED CNN LAYERS.

NO.	Layer Type	Filters	Size/ stride	Activation function
1	Convolutional	16	3/1	--
2	Leaky ReLU	--	--	--
3	Max Pooling	--	1/1	--
4	Convolutional	32	3/1	--
5	Leaky ReLU	--	--	--
6	Max Pooling	--	1/1	--

7	Convolutional	32	3/1	--
8	Leaky ReLU	--	--	--
9	Max Pooling	--	1/1	--
10	Convolutional	128	3/1	--
11	Leaky ReLU	--	--	--
12	Leaky ReLU	--	--	--
13	Max Pooling	--	1/1	--
14	Dense	128	--	Linear
15	Convolutional	256	3/1	--
16	Leaky ReLU	--	--	--
17	Max Pooling	--	1/1	--
18	Convolutional	512	3/1	--
19	Leaky ReLU	--	--	--
20	Max Pooling	--	1/1	--
21	Convolutional	1024	3/1	--
22	Leaky ReLU	--	--	--
23	Max Pooling	--	1/1	--
24	Dense	1024	--	Linear
25	Convolutional	50	3/1	--
26	Flatten	--	--	--
27	Dense	2	--	Softmax

IV. RESULTS AND DISCUSSION

The metrics used to assess the performance includes of the classifiers will be described in this part, followed by a comparison of the outcomes of the suggested approaches. Four evaluation criteria were used accuracy, precision, recall, and f-measure [38].

A. Comparison Results of the Proposed Models

Three models have been proposed to analyze and predict gene immune and antibody response. These models are:

1) The proposed CNN with Mutual Information (MI): The first method was using the structure of the Convolutional Neural Network (CNN) with mutual information feature selection technique with more than one value as shown in "Table 2".

TABLE II. RESULTS OF THE PROPOSED CNN + MUTUAL INFORMATION.

Technique	**Accuracy %**	**Precision %**	**Recall %**	**F-measure %**	**Time in sec.**
MI-20	100	100	100	100	51
MI-40	100	100	100	100	98
MI-60	100	100	100	100	135

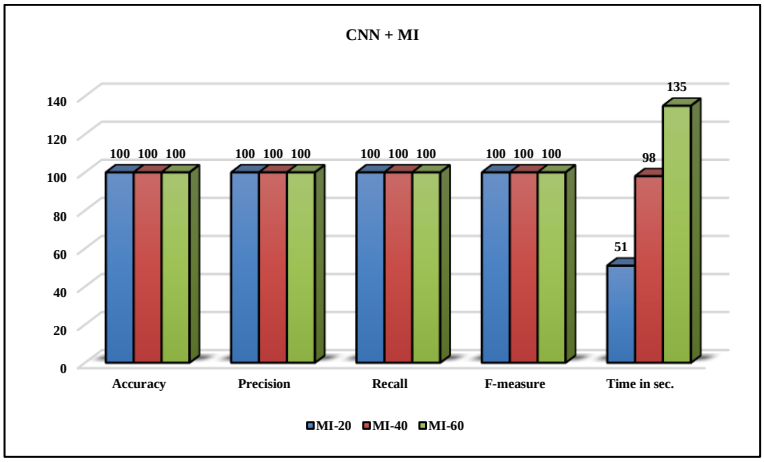

Fig. 5. Chart of the proposed CNN+MI model

2) The proposed CNN with Chi-square: The second method was using the proposed Convolutional Neural Network (CNN) architecture with a Chi-square feature selection technique with more than one value as shown in "Table 3".

TABLE III. RESULTS OF THE PROPOSED CNN + CHI-SQUARE

Technique	Accuracy%	Precision %	Recall %	F-measure %	Time in sec.
Chi-20	100	100	100	100	51
Chi-40	100	100	100	100	98
Chi-60	100	100	100	100	135

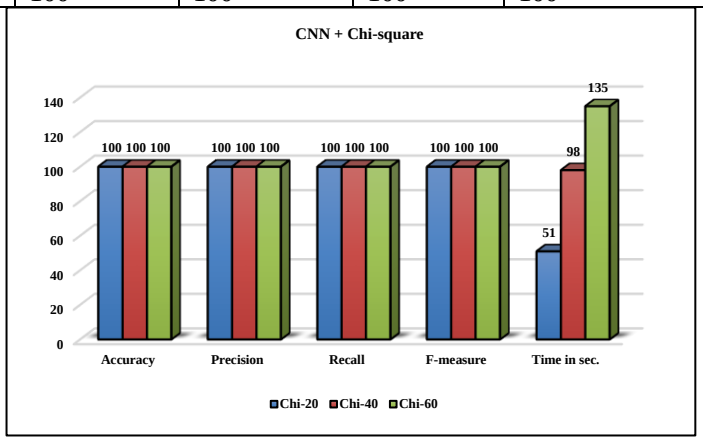

Fig. 6. Chart of the proposed CNN + Chi-square model.

3) The proposed CNN with ANOVA: The second method was using the proposed Convolutional Neural Network (CNN) architecture with an ANOVA feature selection technique with more than one value as shown in "Table 4".

TABLE IV. RESULTS OF THE PROPOSED CNN + ANOVA

Technique	Accuracy %	Precision %	Recall %	F-measure %	Time in sec.
ANOVA-20	100	100	100	100	51
ANOVA -40	100	100	100	100	98
ANOVA -60	100	100	100	100	135

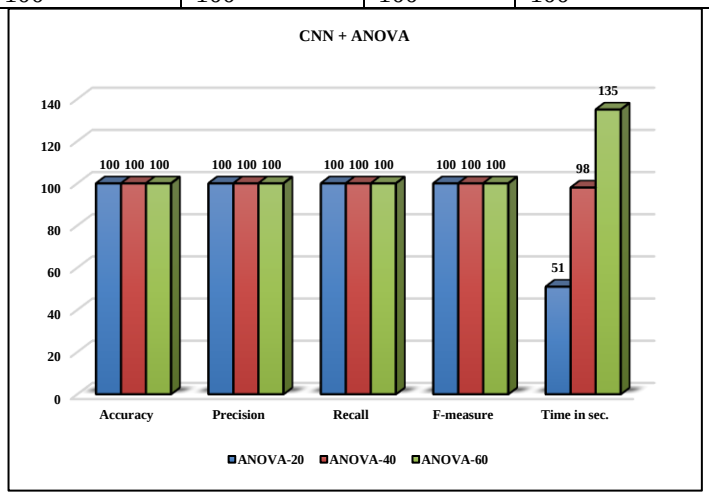

Fig. 7. Chart of the proposed CNN + ANOVA model.

From the above results, it is clear that the results of the three feature selection techniques, which were used with the convolutional neural network model, gave similar results in terms of prediction accuracy. The accuracy reached 100% with very few times not exceeding seconds.

B) comparison with related studies

In this section, we compare the results obtained by applying the proposed model in three cases each time using a different feature selection technique with the results of previous studies that also worked on gene expression.

TABLE V. ACCURACY COMPARISON RESULTS

Technique	Accuracy %
[19]	95.96
[20]	82
Proposed CNN+ MI	100
Proposed CNN+ Chi	100
Proposed CNN + ANOVA	100

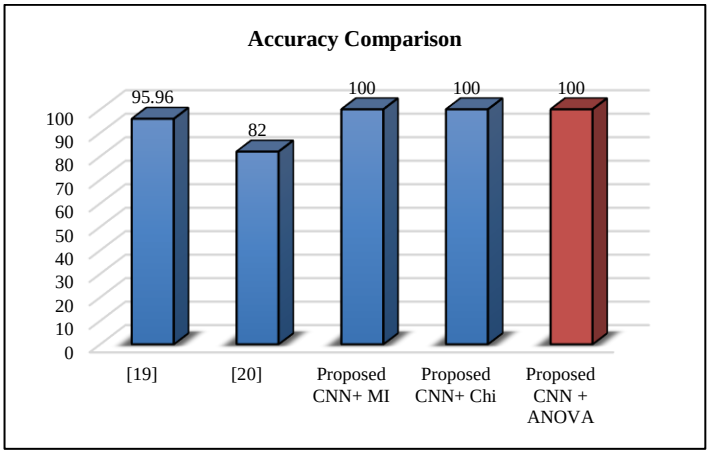

Fig. 8. Chart of an accuracy comparison with previous studies.

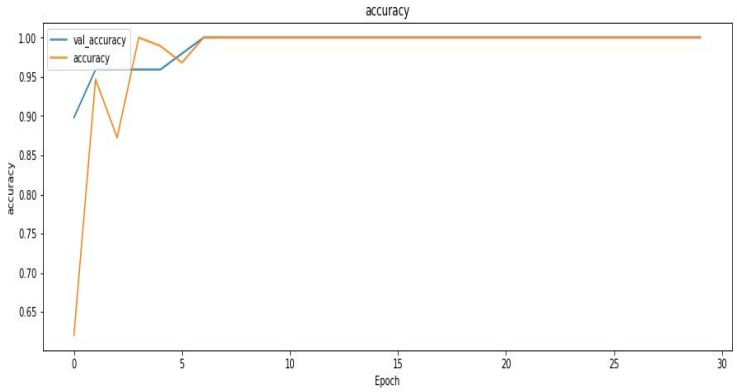

Fig. 9. Valid accuracy vs accuracy

V. CONCLUSIONS

In this study, we propose a novel Convolutional Neural Network (CNN) model to predict the antibody response based on deep learning. The proposed CNN architecture and the feature selection techniques including (Mutual information, Chi-square, and ANOVA) that were employed both improved the model's accuracy and decreased the amount of time needed for prediction. The GSE201535 dataset was used to execute the suggested model, and accuracy results always achieved 100%. The proposed system reaches a stage of equilibrium and can be applied to any data using any of the feature selection techniques. There is no problem with accuracy such as the problem of overfitting because the valid accuracy is equal to the testing accuracy as shown in fig (9) above, the percentage of loss function decreases as the number of epochs increases, and the loss function value does not reach zero, and this means that the accuracy of the system is correct. The proposed system is superior to previous studies in determining the effect of gene expression in determining some diseases or determining the response of bodies to vaccines, as in the current study.

References

[1] J. Zhou, S. Yang, C. Xiao, and F. Chen, “Examination of Community Sentiment Dynamics Due to COVID-19 Pandemic: A Case Study from A State in Australia.” SN Computer Science, Vol. 2, No. 3, pp. 1–11, 2021.
[2] V. Galasso, V. Pons, P. Profeta, and M. Becher, “Gender Differences in COVID-19 Attitudes and Behavior: Panel Evidence from Eight Countries.” Proc Natl Acad Sci, Vol. 117, No. 44, pp. 27285–91, 2020.
[3] E. Mathieu, H. Ritchie, and E. Ortiz-Ospina, “A Global Database of COVID-19 Vaccinations,” Nature Human Behaviour, Vol. 5, No. 7, pp. 947–953, 2021.
[4] D. Trapani and G. Curigliano, “COVID-19 Vaccines in Patients with Cancer,” The Lancet Oncology, Vol. 22, No. 6, pp. 738-739, 2021.
[5] S. Pullan and M. Dey, “Vaccine Hesitancy and Anti-Vaccination in The Time of COVID-19: A Google Trends Analysis,” Vaccine, Vol. 39, No. 14, pp. 1877–1881, 2021.
[6] D. Amanatidis, I. Mylona, and I. Kamenidou, “Mining textual and imagery Instagram data during the COVID-19 pandemic,” Applied Sciences, vol. 11, no. 9, p. 4281, 2021.
[7] M. Awal, “A novel Bayesian optimization based machine learning framework for COVID-19 detection from inpatient facility data,” IEEE Access, vol. 9, pp. 10263– 10281, 2021.
[8] K. Bairagi, “Controlling the outbreak of COVID-19: a noncooperative game perspective,” IEEE Access, vol. 8, pp. 215570–215581, 2020.
[9] G. Troiano and A. Nardi, “Vaccine hesitancy in the era of COVID-19,” Public Health, vol. 194, pp. 245–251, 2021.
[10] P. Gonzalez-Dias, E. Lee, and S. Sorgi, "Methods for predicting vaccine immunogenicity and reactogenicity". HUMAN VACCINES & IMMUNOTHERAPEUTICS, VOL. 16, NO. 2, 269–276, 2020.
[11] I. Ibrahim and A. Abdulazeez, "The Role of Machine Learning Algorithms for Diagnosing Diseases," Journal of Applied Science and Technology Trends, Vol. 2, No. 1, pp. 10-19, 2021.
[12] G. Muhammad, M. Alhamid, and X. Long, "Computing and processing on the edge: Smart pathology detection for connected healthcare", IEEE Network, Vol. 33, No. 6, pp. 44-49, 2019.
[13] J. Chen, K. Gao, R. Wang, and G. Wei, "Prediction and Mitigation of Mutation Threats to COVID-19 Vaccines and Antibody Therapies". Chem. Science, Vol. 12, pp. 6929–6948, 2021.
[14] S. Seneff, G. Nigh, A. Kyriakopoulos, and P. McCullough, "Innate immune suppression by SARS-CoV-2 mRNA vaccinations: The role of G-quadruplexes, exosomes, and MicroRNAs". Food and Chemical Toxicology, Vol. 164, 2022.
[15] V. Rustagi et al., "Analyzing the Effect of Vaccination Over COVID Cases and Deaths in Asian Countries Using Machine Learning Models". Sec. Clinical Microbiology, Vol. 11. 2021.
[16] M. Kurano, "Measurement of SARS-CoV-2 Antibody Titers Improves the Prediction Accuracy of COVID-19 Maximum Severity by Machine Learning in Non-Vaccinated Patients". Sec. Viral Immunology, Vol. 13, 2022.

[17] S. Haggenburg, "Antibody Response in Immunocompromised Patients with Hematologic CancersWho Received a 3-DosemRNA-1273 Vaccination Schedule for COVID-19". American Medical Association, Vol. 8, No. 10, 2022.
[18] E. Wang and A. Chakraborty, "Design of Immunogens for Eliciting Antibody Responses That May Protect Against SARS-Cov-2 Variants." PLoS Comput Biol, Vol. 18, No. 9, 2022.
[19] Y. Wang, M. Yuan, H. Lv, J. Peng, I. Wilson, N. Wu, "A Large-Scale Systematic Survey Reveals Recurring Molecular Features of Public Antibody Responses to SARS-CoV-2". CellPress, Immunit, Vol. 55, pp. 1105–1117, 2022.
[20] S. Shan, S. Luo, Z. Yang, and J. Peng, "Deep Learning Guided Optimization of Human Antibody Against SARS-Cov-2 Variants with Broad Neutralization". PNAS, Vol. 119 No. 11, 2022.
[21] H. Shon, Y. YearnGui, O. Kyoung, E. Cha, K. Kim, "Classification of stomach cancer gene expression data using CNN algorithm of deep learning," J Biomed Transl Res, Vol. 20, No. 1, pp. 15-20, 2019.
[22] J. Zrimec, C. Börlin, F. Buric, et al., "Deep learning suggests that gene expression is encoded in all parts of a co-evolving interacting gene regulatory structure. Nat Commun, vol. 11, No. 6141, 2020.
[23] Q. Nguyen, H. Ly, L. Ho, and N. Al-Ansari, "Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil," Mathematical Problems in Engineering, 2021.
[24] C. Sen Seah et al., "An effective pre-processing phase for gene expression classification," Indonesian Journal of Electrical Engineering and Computer Science, 2018.
[25] A. K. Al-Mashanji and S. Z. AL-Rashid, "Predicting with the quantify intensities of transcription factor-target genes binding using random forest technique," Int. J. Nonlinear Anal. Appl., vol. 12, no. 2, pp. 145–161, 2021, doi: 10.22075/ijnaa.2021.5026.
[26] S. G. K. Patro and K. Kumar, "Normalization : A Preprocessing Stage," no. March 2015, doi: 10.17148/IARJSET.2015.2305.
[27] C. Science and S. Publications, "Data Mining : A Preprocessing Engine Luai Al Shalabi, Zyad Shaaban and Basel Kasasbeh Applied Science University, Amman, Jordan," vol. 2, no. 9, pp. 735–739, 2006.
[28] F. Macedo, R. Valadas, E. Carrasquinha, and M. Oliveira, "Feature selection using Decomposed Mutual Information Maximization". Neurocomputing, Vol. 513, pp. 215-232, 2022.
[29] S. Ikram and A. Cherukuri, "Intrusion detection model using the fusion of chi-square feature selection and multi-class SVM," Journal of King Saud University –Computer and Information Sciences, 2016.
[30] N. Elssied, O. Ibrahim and A. Osman, "A Novel Feature Selection Based on One-Way ANOVA F-Test for E-Mail Spam Classification", Research Journal of Applied Sciences, Engineering, and Technology, Vol. 7, No. 3, pp. 625-638, 2014.
[31] T. Bezdan and N. Džakula, "Convolutional Neural Network Layers and Architectures," International Scientific Conference On Information Technology and Data Related Research, 2019.
[32] S. AlRashid, M. Dosh, and A. Obaid, "Classification of the Senescence-Accelerated Mouse (SAM) Strains with Its Behaviour Using Deep Learning," 1Vol. 8, No. 2, 2022.

[33] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, “Activation Functions: Comparison of Trends in Practice and Research for Deep Learning,” arXiv:1811.03378v1 [cs.LG] 8 Nov 2018.
[34] Y. Bai, “RELU-Function and Derived Function Review,” SHS Web of Conferences, vol. 144, no. 25, 2022.
[35] H. Al-Bermany and S. Al-Rashid, "Microarray Gene Expression Data for Detection Alzheimer’s Disease Using K-means and Deep Learning". Proceedings of the 7th International Engineering Conference "Research and Innovation Amid Global Pandemic", No. 9476128, pp. 13-19, 2021.
[36] M. Yani, B. Irawan, and C. Setiningsih, “Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry’s Nail,” International Conference on Electronics Representation and Algorithm (ICERA 2019).
[37] M. Valizadeh and S. Wolff, “Convolutional Neural Network Applications in Additive Manufacturing: A review,” vol. 4, 2022.
[38] S. Jadon, “COVID-19 detection from scarce chest x-ray image data using few-shot deep learning approach,” Proc. SPIE 11601, Medical Imaging 2021: Imaging Informatics for Healthcare, Research, and Applications, 2021.