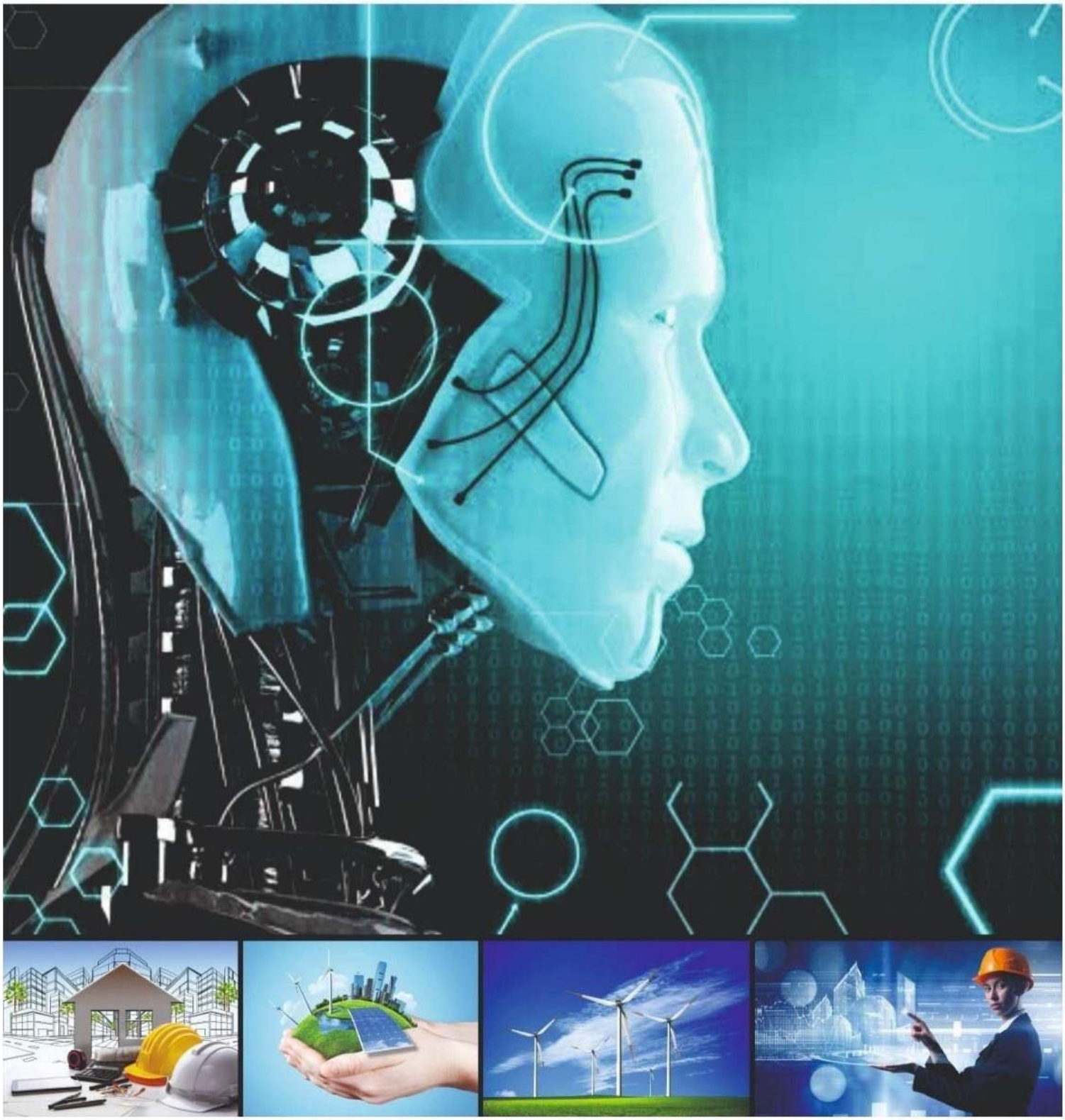
MALATYA TURGUT ÖZAL UNIVERSITY

NATURENGS

MTU Journal of Engineering and Natural Sciences

Volume: 3 Issue: 2 - December 2022



https://dergipark.org.tr/tr/pub/naturengs

www.naturengs.com

MALATYA TURGUT ÖZAL UNIVERSITY

NATURENGS

MTU Journal of Engineering and Natural Sciences

Volume: 3 / Issue: 2 / December - 2022

We are delighted to present the second volume of the journal NATURENGS owned by Malatya Turgut Özal University. MTU Journal of Engineering and Natural Sciences – NATURENGS is a double- blind peer-reviewed, open-access international journal which will publish electronically two times in a year by the Malatya Turgut Özal University from June 2020.

We set out with the desire to create an environment where scientific and/or technological studies carried out in universities, industry and other research institutions will be shared. We aim to advance by giving priority to studies involving scientific and / or technological originality.

Manuscripts submitted for publication are analyzed in terms of scientific quality, ethics and research methods in terms of its compliance by the Editorial Board representatives of the relevant areas. Then, the abstracts of the appropriate articles are sent to at least two different referees with a well-known in scientific area. If the referees agree to review the article, full text in the framework of the privacy protocol is sent. By the decisions of referees, either directly or corrected article is published or rejected. Confidential reports of the referees in the journal archive will be retained for ten years. All post-evaluation process is done electronically on the internet.

In the journal's publication policy, we would like to state that we will not compromise on quality. In this process, we know that we have undertaken important tasks, especially the selection of referees and monitoring of evaluations. Our journal is indexed in *Index Copernicus, CiteFactor, Google Scholar, Scientific Indexing Services, ASOS and ESJI* international databases. We will work with the devotion to get our journal into the TR-Index and then the Science Citation Index database as soon as possible.

We would like to thanks all the Authors and Referees who contributed to this issue.

Assoc. Prof. Aydan AKSOĞAN KORKMAZ
On behalf of the Editorial Board

MALATYA TURGUT ÖZAL UNIVERSITY

NATURENGS

MTU Journal of Engineering and Natural Sciences

Volume: 3 / Issue: 2 / December- 2022

ISSN: 2717-8013

Owner / Publisher

Prof. Dr. Recep BENTLİ for Malatya Turgut Özal University

Chef in Editor

Assoc. Prof. Aydan AKSOĞAN KORKMAZ

Malatya Turgut Özal University, 44210 Battalgazi/Malatya,

TURKEY Phone: +90-422-846 12 55 Fax: +90-422-846 12 25

e-mail: aydan.aksogan@ozal.edu.tr

Editor

Assoc. Prof. Harun KAYA

Malatya Turgut Özal University, 44210 Battalgazi/Malatya,

TURKEY Phone: +90-422-846 12 55 Fax: +90-422-846 12 25

e-mail: harun.kaya@ozal.edu.tr

Contact Information

MTU Journal of Engineering and Natural Sciences – NATURENGS,

Malatya Turgut Özal University, 44210, Battalgazi/Malatya, TURKEY

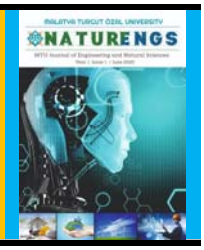
Phone: +90-422 846 12 55, Fax: +90-422 846 12 25,

e-mail: naturengs@ozal.edu.tr

web: https://dergipark.org.tr/tr/pub/naturengs

CONTENTS

NATURENGS
MTU Journal of Engineering and Natural Sciences
https://dergipark.org.tr/tr/pub/naturengs
DOI: 10.46572/naturengs.1182766



Research Article

Deep Learning and Machine Learning-Based Sentiment Analysis on BitCoin (BTC) Price Prediction

Ayşenur Sarıkaya[1*], Serpil Aslan [2]

[1]Department of Informatics, Malatya Turgut Ozal University, Malatya, Turkey.
[2] Department of Software Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Ozal University, Malatya, Turkey.

(**Received:** 30.09.2022; **Accepted:** 16.11.2022)

ABSTRACT: Emotions form an essential and fundamental aspect of our lives. What we do and say reflects some of our feelings in some way, though not directly. We must examine these feelings using emotional data, also known as affect data, to comprehend a person's basic behavior. Text, voice, facial expressions, and other data types can be included. Since social networking websites have become so popular, many individuals have started reading the material on these numerous sites. Twitter is one of these social networking sites. People's feelings and thoughts about a subject reveal positive, negative, and neutral emotional values. Doing sentiment analysis on Twitter is a critical and challenging task. In this study, we aim to investigate the sentiments of Bitcoin and provide an overview of its effect on the value of Bitcoin by utilizing the power of deep learning architectures and machine learning methods. The study collected tweets in English shared on Twitter between December 12, 2021, and March 13, 2022. First, people's feelings about Bitcoin were assessed using TextBlob, a natural language processing (NLP) tool. Then, it was done using basic machine learning algorithms for sentiment classification and Convolutional neural network (CNN), Long short-term memory (LSTM), and Bidirected Long short-term memory (BiLSTM) deep learning architectures that we modeled. However, deep learning models were tested separately with the TF-IDF and Glove word embedding approaches. Experimental results prove the success of deep learning architectures using the Glove word embedding approach.

Keywords: BitCoin, Sentiment Analysis, Machine Learning, Deep Learning, Glove, TF-IDF, CNN, LSTM, BiLSTM.

1. INTRODUCTION

The general purpose web has increased with the Internet's rising popularity. As a result, many individuals communicate their views and opinions through various online resources. It will be used for users, who will be used automatically, to use the public's constant looking and use. The popularity of social networking sites has led to the emergence of several fields devoted to extracting vital information from social networks and their content.

Sentiment analysis is the process of determining a text's emotional meaning from its content. Given the long and well-known use of public opinion in decision-making, sentiment analysis is an area of NLP, and there must have been much early research on this topic. However, sentiment analysis is still being developed in the new millennium. Scholars, corporations, governments, and organizations have adopted sentiment analysis in recent years [1]. Sentiment analysis has grown in popularity among the research community in recent years. The phrase "Idea Analysis or Idea Mining"

*Corresponding Author: aysnr.zrn@gmail.com
ORCID number of authors: [1] 0000-0003-3696-3645, [2] 0000-0001-8009-063X

is another name for sentiment analysis. The sentiment analysis work has lately expanded. Twitter remains the best indicator of the world's broader pulse and what is happening within it. The data provided by Twitter and the insights we can gather from it can truly change the world in more ways than most people realize. Twitter data is now included in most stock analyst systems, including Bloomberg's.

In this study, the shares of Twitter users about BitCoin were initially automatically captured with the Twitter API application we designed [2]. In this process, the extracted data is not clean as it is a raw data set. For this reason, pre-processing steps were applied to the text. Many unnecessary special characters, expressions, links, tags, emojis, etc., are cleaned in this context. Thus, unnecessary characters are deleted. After the pre-processing step, meaningful information was extracted from the obtained quality dataset using the Glove data embedding approach. In the Glove-based BiLSTM approach we designed for this study, the data in the created corpus were classified into three classes positive, negative and neutral. The determination of the emotional tendencies in social media is directly proportional to the preferred topic and the popularity of the topic on social media. Therefore, detecting missing, incorrect, and false information about BitCoin will help control the sudden rise and fall of cryptocurrencies.

2. RELATED WORKS

Sentiment Analysis is analyzing data and classifying it according to the research needs. Sentiment analysis is a method that automatically scans natural language utterances, picks out important claims or opinions, and organizes them into groups according to emotional attitudes. These emotions can be used to gain a better understanding of various events and the effects they cause. Analyzing these emotions allows us to determine what people like, want, and their primary concerns. Sentiment analysis methods are explained under dictionary-based, Machine learning, and hybrid approaches.

Dictionaries are a set of tokens, each with a predetermined score reflecting the neutrality, positivity, or negativity of the text to which it is allocated [3]. The Dictionary-Based Approach adds the positive, negative, and neutral evaluations separately for each token to do a full-text analysis. The highest value of the individual scores is used to determine the text's general polarity in the final step. As a result, the text is first separated into symbols made up of single words, after which each symbol's polarity is determined and added. The dictionary-based approach is suitable for feature- and sentence-level sentiment analysis. Since no training data is needed, it can be referred to as an unsupervised technique.

Systems that use machine learning can classify emotions. The machine learning method takes on the problem of sentiment classification, which is a typical text classification problem, by using syntactic and/or linguistic components. The categorization model links the attributes of the underlying record to one of the class labels. The model then predicts the class label for a particular instance of an unknown class. We are given a difficult categorization task when a sample is given just one title. A soft classification problem is one in which the probabilistic value of the labels is assigned to a sample. Systems can learn new skills using machine learning without explicitly coding them. Algorithms for sentiment analysis, for instance, can be taught to read beyond simple definitions to recognize sarcasm, context, and phrases that aren't supposed to be used in that way. The hybrid method combines dictionary-based and machine-learning techniques. The term "hybrid" describes the fusion of sentiment analysis, machine learning, and dictionary-based methods. Emotional dictionaries play a crucial part in most systems in the modern hybrid approach, which mixes the two. Sensitivity analysis is a hybrid method for polarity recognition that combines statistical and knowledge-based techniques [4]. He suggested a hybrid approach to machine learning that combined two feature selection

methods, “Multithreaded Optimizer and Relief Algorithms with SVM” [5]. A hybrid strategy using machine learning, incorporating RF and SVM, was developed for the sentiment analysis problem [6]. The hybrid model, which combines both methods, had a more accurate performance in the dataset of product reviews provided by amazon.com, which was close to 84 percent. Many researchers have proposed a hybrid architecture that blends dictionary-based and machine-learning techniques to improve the results. However, many studies are still being done on this topic, which is still quite popular.

3. CRYPTO CURRENCY, TWITTER AND SENTIMENT ANALYSIS

Bitcoin: A Peer-to-Peer Electronic Cash System, written by "Satoshi Nakamoto" in 2009, describes a peer-to-peer payment system employing electronic cash (cryptocurrencies) that may be delivered directly from someone [7]. Bitcoin is the first cryptocurrency, which came to life using cryptography, that is, encryption technology, using the name "Satoshi Nakamoto" and was invented in 2009 by an unknown person or group. However, peer-to-peer payment systems are one use case for blockchain technology and many more. They are suitable for IoT applications, distributed storage systems, healthcare, and more since they give the other side security, anonymity, and a distributed ledger without using a third party to validate the transaction [8].

Information gathered through social media networks is referred to as social data. There are social media platforms where people interact with each other on many issues and share their feelings on certain issues. One of these platforms is Twitter. The tweets individuals write about specific topics enable Twitter, a social media tool, to be used as a dynamic data source. The tweets written about the topics on the agenda can give information about the direction of the comments on that topic. Researchers can collect tweets on hot topics and reveal positive, negative, or neutral feelings about those topics [3] and use Twitter data for training to reveal the success of hybrid models in sentiment analysis. The Twitter API was used to retrieve up to 6900 tweets for educational purposes. According to the data, their model outperformed most models, which indicated a 96% reduction in features. Additionally, they highlighted hybrid models' advantages and concluded that they might surpass all others with exemplary architecture and careful hyperparameter selection [5].

One of the hot topics is the digital currency Bitcoin. The thirteen years since Bitcoin launched have seen much growth and controversy. Recently, cryptocurrency has gained much attention, partly because of its destructive potential and claims of unheard-of profits [9] [10]. Additionally, scholars are becoming more aware of Twitter's ability to predict various events, particularly those related to financial markets. This subject is widely discussed in tweets and has grown in popularity, especially in recent years. Bitcoin and altcoins are the currencies that are invested in and evaluated as financial resources. People, especially investors, are effective in encouraging investments in those coins by sharing tweets about the coins they invest in. This study collected tweets about Bitcoin and the most popular coins daily. In our work with the collected tweets, we will analyze the effect of Twitter on bitcoin with a deep learning-based model. The forecasting capability of Twitter sentiment in various cryptocurrency ecosystems is the topic of this study.

Twitter Sentiment Analysis [11] has become a trendy research topic for researchers working on NLP and Sentiment Analysis (SA). However, the diversity and size of the data in social media make it

impossible for people to conduct sentiment analysis. This situation necessitated an automated sentiment analysis system.

NLP, a set of techniques allowing computers to analyze and understand text, has emerged as a subject of study or development due to the proliferation of unstructured data [12]. This study used a set of NLP tools called "sensitivity analysis". Whitelaw et al. [13] defined sentiment analysis as “labeling documents as positive or negative according to the target”. Sentiment analysis approaches generally classify emotions as positive, negative, or neutral. Sentiment analysis is not a new field of research; many researchers have put forward many studies on sentiment analysis over the years.

4. MATERIALS AND METHODS

4.1. Data Collection

Twitter, the world's largest social media platform, was chosen as the data source in the study. The data were collected from tweet data shared publicly on Twitter. For this data set, only tweets shared in English were collected in tweets shared publicly by area. The study used the MAXQDA [14] qualitative data analysis tool to collect Twitter data. First, after the "Twitter API Key" entries were made using the Twitter Developer Account, 256502 Tweets about Bitcoin and sub-coins published between December 12, 2021, and March 13, 2022, were collected in English. Then, to reach the target tweets, the search terms "BTC OR ETH OR XRP OR BCH OR EOS OR LTC OR ADA OR XLM OR TRX" were used after identifying the most tweeted hashtags about Bitcoin on Twitter. Finally, the collected data was converted to a CSV file in the pre-processing and emotion classification steps. As a result, a dataset of 256502 English tweets was collected between December 12 and March 13, 2022. To ensure the accuracy of this data set, the pre-processing stage, whose details are given in the next stage, was carried out. After the pre-processing steps, the size of the dataset consisting of English tweets was reduced to 152398 tweets.

4.2. Data Cleaning and Pre-processing

The unstructured nature and high noise levels of Twitter data are well recognized. Therefore, to be effective in sentiment analysis, the gathered Twitter data must undergo a lengthy preparation process. Text cleaning is one of the text mining operations to clean up words or other components that are difficult to deduce or analyze the meaning of the text. During the analysis phase, many unnecessary special characters, expressions, links, tags, emojis, etc. These characters are challenging to analyze for sentiment because they don't provide a lot of information. These characters are challenging to analyze for sentiment because they don't offer much information. That may adversely affect the experimental results includes. At this stage, pre-processing was performed on the data set by applying the following procedures:

- Duplicate tweets are removed from the dataset.
- Numbers are removed.
- Punctuation marks are removed.
- Twitter RT, @, and links in sentences are cleared.
- The text is converted to lowercase, so words like “bitcoin” and “Bitcoin” are considered the same for analysis.

- Special characters and facial expressions (emojis) used to express emotions are removed after being detected in the text using their special codes.
- Tokenization is applied for each tweet. Tokenization is separating the words in the text according to the spaces.
- Then, stopwords are removed from the tweets.
- The words in each tweet are separated into their roots. Lemmatization uses morphological analysis of words.
- Finally, the sentiment of tweets was evaluated using TextBlob [15], an NLP tool.

The purpose of pre-processing is to reduce the number of words in the text without disturbing the basic meaning of the text. In the raw data set collected, some unnecessary words and expressions will not be used in the sentiment analysis phase. Therefore, pre-processing must be performed before performing any data analysis.

4.3. Word Representation Approaches

TF-IDF

The usefulness of a word to a document in a collection of documents is evaluated using a statistical method called TF-IDF [12]. This is accomplished by multiplying the frequency of a word within a document and the reverse document frequency over a set of documents. The most significant benefit is that it has a variety of uses for automated text analysis, including word scoring in machine learning algorithms for NLP.

TF-IDF was created for document search and information retrieval. It works by increasing a term's frequency in a document, but this is offset by the number of papers in which it appears. Because they are not particularly pertinent to that document, words like this, what, and if, often used throughout all documents, rank poorly. A word vector is a set of numbers for each potential sentence word in a document. By taking a document's text and turning it into one of these vectors, the text's content is somehow represented by the vectors' numbers. We can use TF-IDF to link each word in a document to a number representing its importance to that content. We then use a machine learning technique to hunt for documents with similar vectors and similar, related terms.

GLOVE

Spherical Vectors for Word Representations is the abbreviation for GloVe. The GloVe [16] is an unsupervised learning technique for discovering word placement for various words in vector representation. The global word-to-word co-occurrence matrix, which tracks how frequently words occur together in a given whole, provides nonzero inputs that the GloVe model is trained with. One pass across the full corpus obtains the information needed to fill this matrix. Large businesses may find this shift computationally expensive, but there is only one upfront expense. Subsequent training iterations are completed much quicker since nonzero matrix entries often make up a smaller proportion of the corpus than words.

4.4. Machine Learning Models

SVM

The Support Vector Machine (SVM) [17] uses classification methods to address two-group classification issues. After providing tagged training data sets to an SVM model for each category, they can classify the new text.

Compared to more contemporary algorithms like neural networks, it offers two major advantages: faster processing and better performance with fewer data (in thousands). Therefore, the method is ideally suited for text classification problems where access to a dataset with up to several thousand labeled samples is common.

Karasu et al., in the study conducted by [18], LR and SVM, two machine learning algorithms, were used to predict the price of bitcoin using a time series of daily closing prices from 2012 to 2018. It has been seen that the proposed SVM model gives more successful results than the LR model.

RF

Machine learning models called Random Forest (RF) models [19] combine the output predictions of some regression decision trees. A random vector generated from the input data is used to create each tree independently, ensuring that all the trees in the forest have the same distribution. Bootstrapping and random feature selection are used to average predictions from forests. RF models are reliable predictors for high-dimensional and small-sample data [20]. Wimala Gunaratne et al. [21] used ANN, SVM, RF, and NB algorithms for cryptocurrency price prediction. The success percentage of studies conducted on Twitter heavily depends on the correctness of the data.

A classifier that develops from decision trees is called a random forest. There are numerous decision trees in it. Each decision tree classifies a fresh sample by assigning a classification to the incoming data. Each tree receives sampled data from the original dataset as input. After gathering the categories, the random forest chooses the forecast with the highest votes.

Additionally, to develop the tree at each node, a random selection of features from the optional characteristics is selected. Finally, no tree gets pruned as it grows. The random forest technique enables a strong classifier to be formed from multiple weak or weakly connected classifiers.

LR

The supervised learning classification procedure known as logistic regression (LR) is utilized to calculate the likelihood of a target variable. Due to the binary nature of the target or dependent variable, only two viable classes exist. Data in a logistic regression model are recorded as 1 (representing success/yes) or 0 (meaning failure/no), depending on the binary nature of the dependent variable. A logistic regression model makes mathematical predictions about P(Y=1) as a function of X. Detecting spam, predicting diabetes, finding cancer, etc. It is one of the most straightforward ML techniques that may be applied to many categorization issues.

In the study of Chen et al. [16], statistical methods for daily Bitcoin price prediction include Linear Regression (LR) and Linear Discriminant Analysis (LDA); machine learning algorithms utilized include Random Forest (RF), XGBoost (XGB), Second Order Discriminatory Analysis (QDA), SVM, and Long-Short-Term Memory Networks (LSTM). For daily Bitcoin price prediction, statistical methods had an accuracy of 66%, while machine learning algorithms had a success rate of 65.3%. Even the highest success rate was inadequate compared to previous research because of the study's numerous flaws. Other machine learning algorithms weren't employed either for the study's execution.

4.5. Deep Learning Models

CNN

This method, also referred to as CNN [22], allows us to analyze an image briefly. It is a deep learning algorithm that helps us distinguish various objects in the image from each other. As can be seen from Fig. 1, similar to a typical multilayer neural network, CNN has one or more convolutional layers, pooling layers, and one or more linked layers. With the same number of hidden units, CNNs have the advantage of requiring less training and fewer parameters than fully linked networks. CNN is an example of a feedforward neural network. CNN is ideal for image processing applications but can also be used in audio and NLP applications. A CNN model comprises three layers: the convolutional layer, the pooling layer, and the fully connected layer. Convolutional, pooling, and fully connected layers are included in the feature extraction phase; however, only the fully connected layer is included in the classification step. The feature map is extracted from the convolutional layer using filters. The input parameters are reduced in dimension by the pooling layer, which comes after the convolutional layer. One of the average pooling or maximum pooling approaches is typically selected for this technique. The filter moves along the input in the maximum pooling method and chooses the maximum values that will make up the output. In the average pooling method, the moving filter creates the output values by averaging the input values in the area. The flattened layer prepares the incoming input for the fully connected layer by converting it into a one-dimensional array. The layer where the qualities are categorized is the completely connected layer. Each neuron in a layer is linked to every neuron in the layer above it in a completely connected layer.

Roy and Ojha [23] stated that Twitter is a big gold mine where people share their instant feelings and thoughts, and based on this, they conducted a sentiment analysis on Twitter. Three deep-learning models were created and compared for sentiment analysis. Google BERT, LSTM, and CNN algorithms were used, and it was determined that the BERT model outperformed the others. The language of the study was English, which was an increasing factor in the accuracy rate.

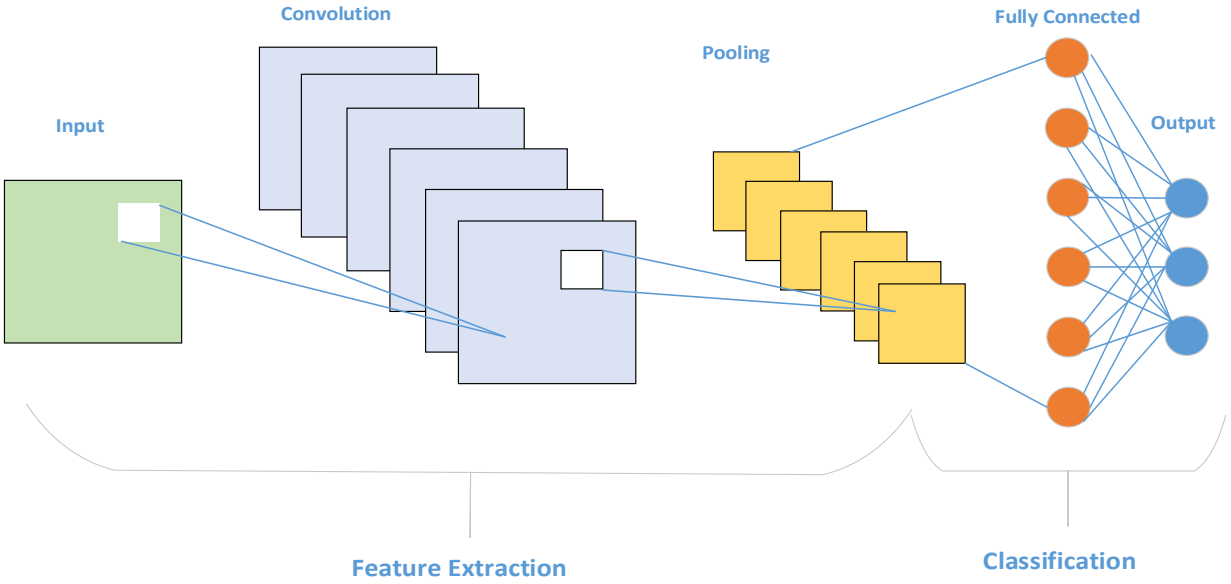


Figure 1. Basic Structure of the CNN Model

LSTM

LSTM is the most commonly used and successful RNN method in sentiment analysis. The method introduced by Hochreiter and Schmidhuber in 1997 is designed to avoid the problem of long-term addiction [24]. As seen from Fig. 2, an LSTM cell consists of an information flow line called the cell state and three interactive gates that form the decision mechanism. These gates are called forget gate (forget gate), the input gate (input gate), and the output gate (output gate). The LSTM operation first decides which information to keep or not in the cell state. The information obtained from the previously hidden layer and the information now being input are then transmitted to the forget gate. The sigmoid function is then used to construct the outcome. The result of the sigmoid function ranges from 0 to 1. 0 means forget information, 1 means retain information. The second stage involves selecting the new information that will be kept in the cell state. The input gate then selects which data to update using the sigmoid function. The cell state is then updated after evaluating the results and creating value vectors using the Tahn function. Which information from the cell state will be sent as a result is decided in the last phase. The Tahn layer organizes the information from the cell state, the output port completes the decision step, and the outcome is generated.

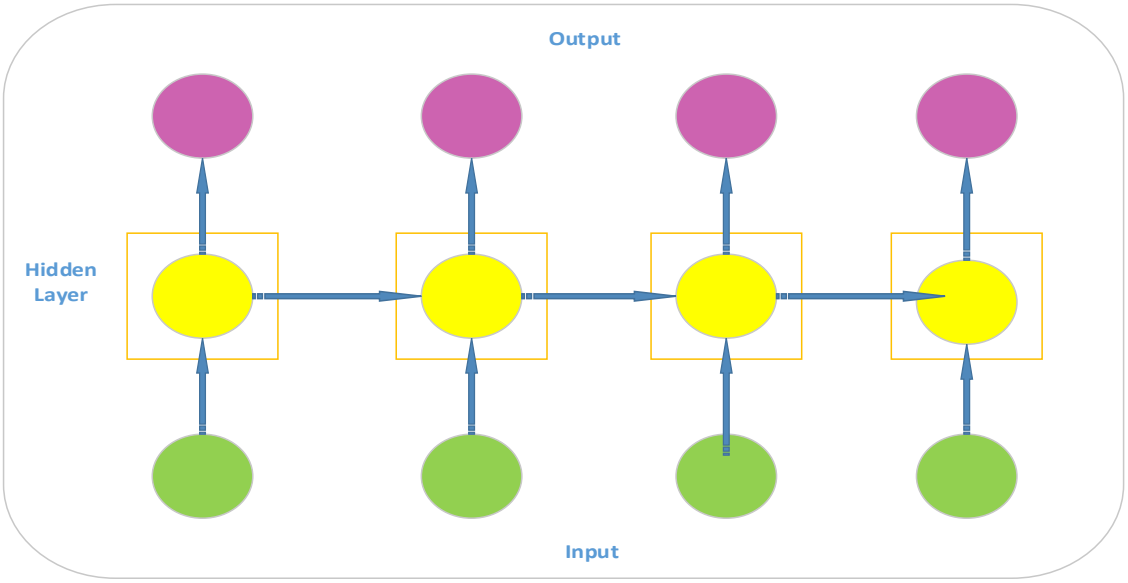


Figure 2. Basic Structure of the LSTM Model

BLSTM

A group of processing models called Bidirectional LSTM (BLSTM) [25] consists of two LSTMsThey have the ability to store data at any moment, including information from the past and the future. While the other receives input flowing backward, the former does the opposite. The network can access more data using BiLSTMs, which helps the context of the algorithm. You can run your inputs in two directions—from the present to the future and from the present to the future—by using bidirectional. This method combines two hidden states and is different from the unidirectional in that it uses backward working LSTM to safeguard information from the future.

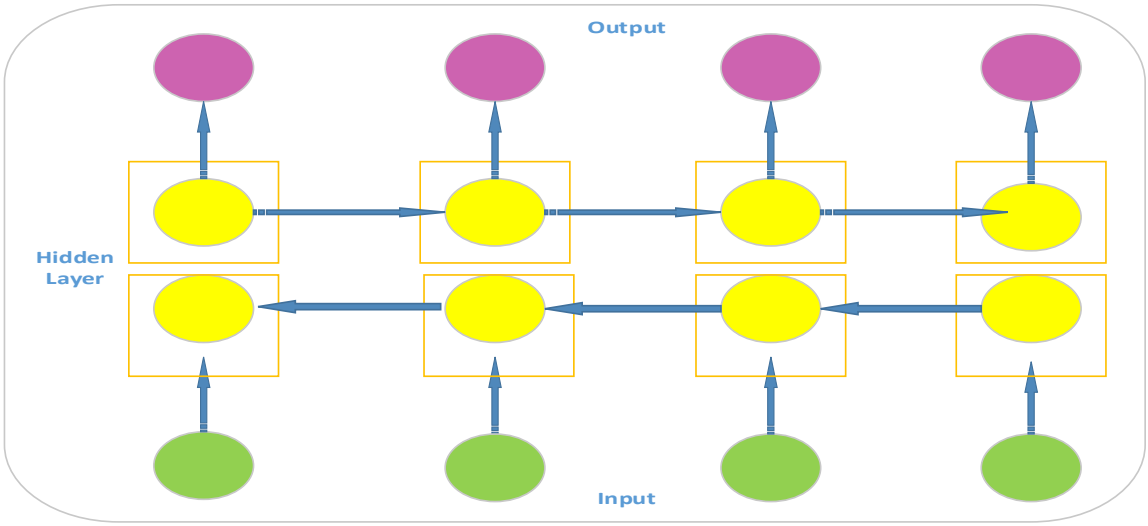


Figure 3. Basic Structure of the BiLSTM Model

4.6. Evaluation Metrics

We used four performance scales to evaluate the performance of deep learning and machine learning models. This study uses "Accuracy, Precision, Recall, and F1-score" [26], which are among the commonly used evaluation criteria.

Precision is called the precision of a classifier and indicates what percentage of all clusters are positively labeled and positive. It is calculated as follows:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

Recall is often called a measure of completeness and indicates the percentage of true positive predictions labeled correctly. It is calculated as follows:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

Accuracy assessment may not be a good metric because there is no balanced data set. In such cases, the F1-score is used. Because the F1-score provides the results according to each target class, it is a statistical classification analysis criterion that considers both the precision of the classifier and the recall criteria. It is calculated as follows:

$$F1 - score = 2\frac{Precision\ x\ Recall}{Precision + Recall} \quad (3)$$

5. EXPERIMENTAL RESULTS

This section presents the results obtained with the machine and deep learning classifiers on the dataset we created by collecting tweets shared from Twitter about Bitcoin and altcoins through MAXQDA qualitative data analysis. Experiments were tested using the Python programming language on the Google Collaborate Pro platform. Pandas, Keras, Numpy, spaCy, and Sklearn python programming libraries were used in the experiments. All experiments were tested on a PC with an Intel Core i7 processor, 16 GB RAM, and Windows 10 operating system. The parameters of the deep models that were compared are shown in detail in Table 1.

CNN is good at capturing local invariant features but not at capturing word order information in a sentence. This greatly reduces a sentence's semantic representation. Therefore, the experimental results of the CNN model have lower performance for both word embedding approaches compared to other models.

For all the deep learning models compared, it is seen that the Glove word embedding approach outperforms the trained variations based on the TF-IDF word approach. TF-IDF identifies words that frequently occur in the given text and phrases that are not common in the remaining dataset. TF-IDF ignores the order of words and returns the mxn matrix (or mxn depending on the application, where n is the number of words in the vocabulary and m is the number of documents. TF-IDF cannot capture

the order of words and semantic and syntactic information of words in a document because it is created with the word bag logic. On the other hand, Glove assigns each word a distinct vector based on the words that surround it. While TF-IDF is based on a sparse vector representation, Glove belongs to dense vector representations. For all these reasons, Glove was preferred as the word embedding approach in our proposed model. Glove and TF-IDF word embedding approaches were tested separately using CNN, LSTM, and BiLSTM deep learning architectures. As can be seen from Fig. 4 and Fig. 5, GloVe has achieved the best performance than TF-IDF for all compared deep models. Fig. 6, Fig. 7 and Fig. 8 represent the performance comparison of CNN, LSTM and BiLSTM using a different word embedding vectors, respectively. As can be seen from Fig. 6-7, GloVe has achieved the best performance than TF-IDF for all compared deep models. The glove method is computationally expensive, but it is a one-time upfront cost. Compared to TF-IDF, it performs better when training on datasets of similar or larger size. This is because it produces its best guess by learning the entire co-occurrence dataset once. It differs from the TF-IDF method, where additional sentences can be given for training and prediction.

Table 1. Hyperparameter setting.

	Hyperparameter	
Embedding and Input Layer	Optimizer	Adam
	Loss Function	Binary Cross Entropy
	Learning Rate	0.0001
	SpatialDropout1D	0,2
	Maximum Length	60
	Embedding size	200
	Batchsize	32
	Epoch	20
CNN	CNN filter size	100
	CNN kernel size	3
	CNN Activation	ReLu
	CNN Maxpooling size	2
	Dropout	0.25
LSTM	LSTM Node	128
	LSTM Activation	tanh
	LSTM Dropout	0.25
BiLSTM	BiLSTM Node	128
	BiLSTM Activation	ReLu
	BiLSTM Dropout	0.25
FC Layer	Dense 1	50
	Dense 1 Activation	ReLu
	Dense 2	1
	Dense 2 Activation	softmax

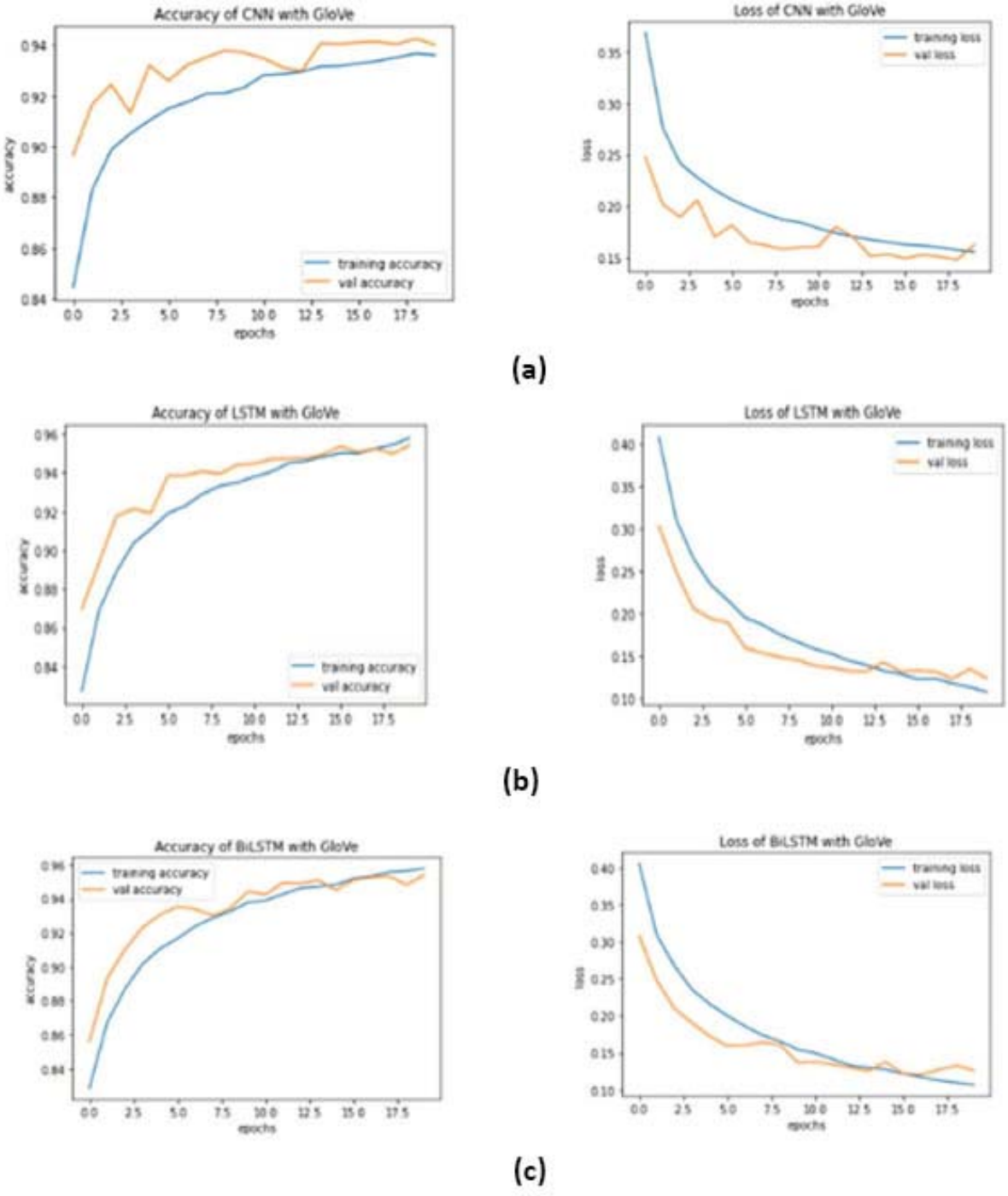


Figure 4. Comparison of (a) Glove-CNN, (b) Glove-LSTM, (c) Glove-BiLSTM Accuracy and Loss Curves

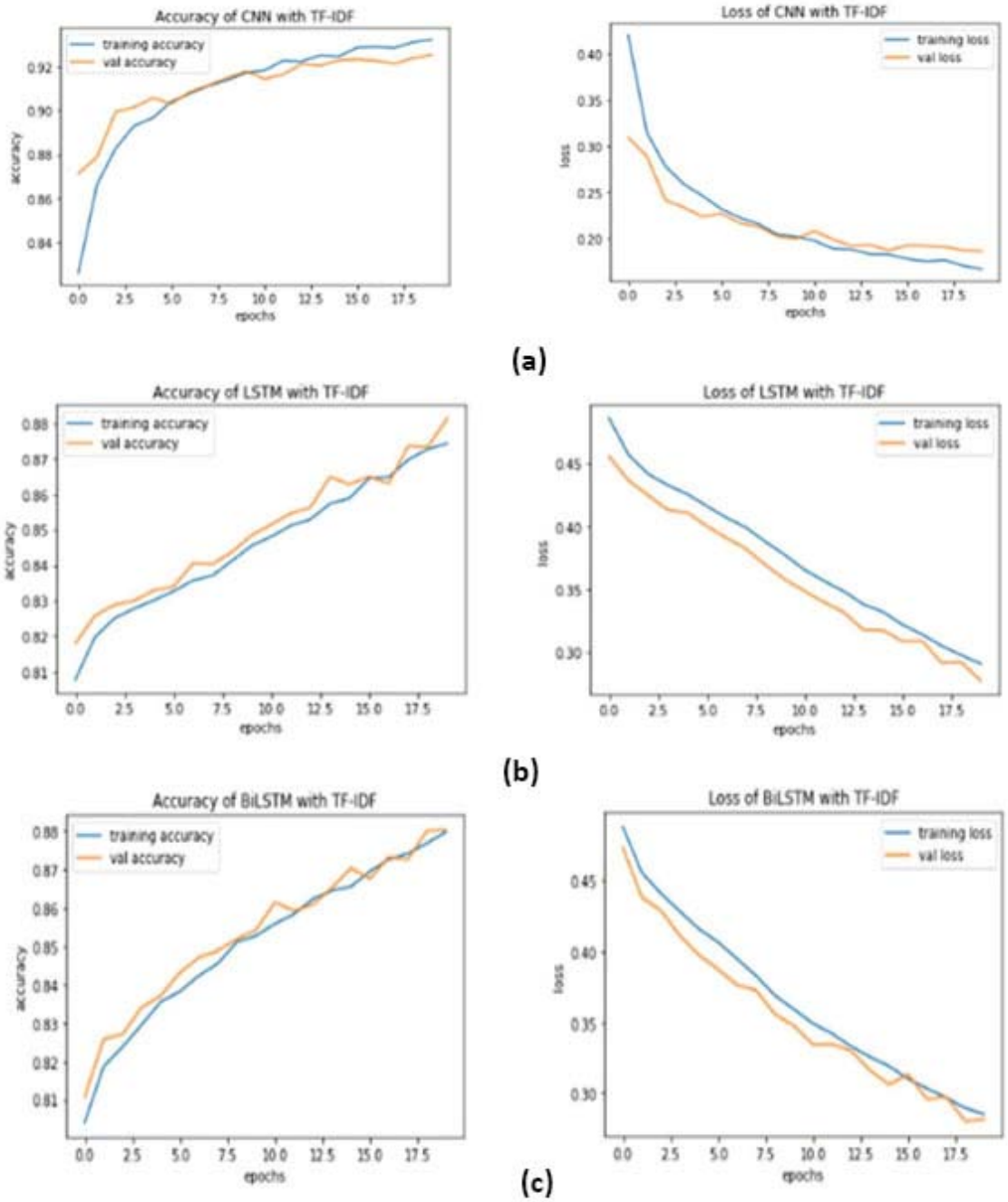


Figure 5. Comparison of (a) TF-IDF-CNN, (b) TF-IDF -LSTM, (c) TF-IDF -BiLSTM Accuracy and Loss Curves

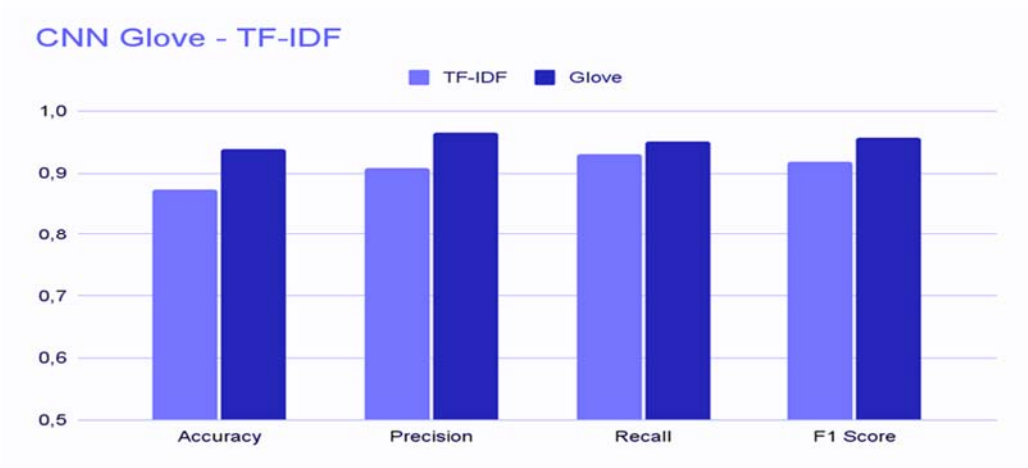


Figure 6. Performance Comparison of CNN using Different Word Embedding Vectors

Figure 7. Performance Comparison of LSTM using Different Word Embedding Vectors

Figure 8. Performance Comparison of BiLSTM using Different Word Embedding Vectors

Fig. 9 shows the performance comparison of the RNN-based BiLSTM model with the basic machine learning algorithms, which achieved the best performance. However, as seen in Fig. 9, machine learning algorithms have a severe performance weakness compared to deep learning-based algorithms. The most important reason for this is that in machine learning methods, the learning process is divided into small steps, and the results of each step are combined into a single output. In contrast, in deep learning methods, the learning process is processed from end to end.

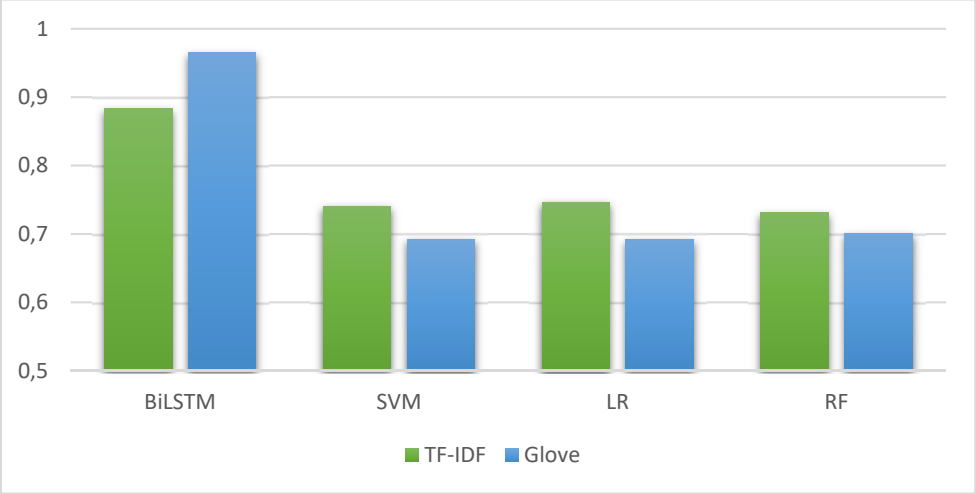


Figure 9. Comparison of the BiLSTM Deep Learning Model's Accuracy Performance with Other Machine Learning Models Using Different Word Embedding Vectors

6. CONCLUSION

Our study evaluated their feelings about Bitcoin using TextBlob, an NLP tool. Then, it was done using basic machine learning algorithms for emotion classification and CNN, LSTM, BiLSTM deep learning architectures that we modeled. Extensive experiments prove that the proposed model performs better. However, after the deep learning models were tested separately with the TF-IDF and Glove word embedding approaches, the experimental results prove the success of deep learning architectures using the Glove word embedding approach.

This study thoroughly investigated the English-language tweets that were shared between December 12, 2021, and March 13, 2022. For this analysis, 152398 shared tweets about Bitcoin and other currencies were used. The sentiment analysis results revealed that the overall sentiment polarity was positive and that there were around twice as many positive tweets as negative ones. The experimental findings of this study will make it easier to spot unreliable, inaccurate, and incomplete information on Bitcoin and other cryptocurrencies. In this way, individuals who experience Bitcoin investment indecision will be able to make the right choices in investment. Those who follow market investments can apply it efficiently in determining the valuation of Bitcoin investment. In the future, it is planned to develop an optimal deep learning-based system to detect fake accounts and false information and intervene in social media based on current work.

Acknowledgments

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declaration of Competing Interest
There is no conflict of interest.

Author Contribution
All authors contributed equally to every step of the article.

REFERENCES

[1] Sánchez-Rada, J. F., Iglesias, C. A., Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison. Information Fusion, 2019. 52: p. 344-356.

[2] McNally, S., Roche, J., Caton, S., Predicting the price of bitcoin using machine learning. In 2018 26th euromicro international conference on parallel, distributed and network-based processing (PDP). IEEE, 2018. p. 339-343

[3] Kiritchenko, S., Zhu, X., Mohammad, S. M., Sentiment analysis of short informal texts. Journal of Artificial Intelligence Research, 2014. 50: p. 723-762.

[4] Hassonah, M. A., Al-Sayyed, R., Rodan, A., Ala'M, A. Z., Aljarah, I., Faris, H., An efficient hybrid filter and evolutionary wrapper approach for sentiment analysis of various topics on Twitter. Knowledge-Based Systems, 2020. 192: p. 105353.

[5] Jin, C., Kong, X., Chang, J., Cheng, H., Liu, X., Internal crack detection of castings: a study based on relief algorithm and Adaboost-SVM. The International Journal of Advanced Manufacturing Technology, 2020. 108(9): p. 3313-3322.

[6] Al Amrani, Y., Lazaar, M., El Kadiri, K. E., Random forest and support vector machine based hybrid approach to sentiment analysis. Procedia Computer Science, 2018. 127: p. 511-520.

[7] Nakamoto, S., Bitcoin: A peer-to-peer electronic cash system Bitcoin: A Peer-to-Peer Electronic Cash System. 2009. Bitcoin. org. Disponible en https://bitcoin. org/en/bitcoin-paper.

[8] Miraz, M. H., Ali, M., Applications of blockchain technology beyond cryptocurrency. 2018. arXiv preprint arXiv:1801.03528.

[9] Kilimci, Z. H., Sentiment analysis based direction prediction in bitcoin using deep learning algorithms and word embedding models. International Journal of Intelligent Systems and Applications in Engineering, 2020. 8(2): p. 60-65.

[10] Pant, D. R., Neupane, P., Poudel, A., Pokhrel, A. K., Lama, B. K, Recurrent neural network based bitcoin price prediction by twitter sentiment analysis. 2018. In 2018 IEEE 3rd International Conference on Computing, Communication and Security.

[11] Spencer, J., Uchyigit, G., Sentimentor: Sentiment analysis of twitter data. 2012. In SDAD@ ECML/PKDD p. 56-66.

[12] Kamyab, M., Liu, G., Adjeisah, M., Attention-Based CNN and Bi-LSTM Model Based on TF-IDF and GloVe Word Embedding for Sentiment Analysis. Applied Sciences, 2021. 11(23): p. 11255.

[13] Whitelaw, C., Garg, N., Argamon, S., Using appraisal groups for sentiment analysis. 2005. In Proceedings of the 14th ACM international conference on information and knowledge management, p. 625-631.

[14] We used MAXQDA 2020 (VERBI Software, 2019) for data analysis.

[15] Loria, S., Textblob Documentation. Release. 2018. 0.15, 2, p. 269.

[16] Pennington, J., Socher, R., Manning, C. D., Glove: Global vectors for word representation. 2014. In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), p. 1532-1543.

[17] Jakkula, V., Tutorial on support vector machine (svm). School of EECS, Washington State University, 2006. 37(2.5): p. 3.

[18] Karasu, S., Altan, A., Saraç, Z., Hacioğlu, R., Prediction of Bitcoin prices with machine learning methods using time series data. 2018. In 2018 26th signal processing and communications applications conference (SIU), IEEE, p. 1-4.

[19] Al Amrani, Y., Lazaar, M., El Kadiri, K. E., Random forest and support vector machine based hybrid approach to sentiment analysis. Procedia Computer Science, 2018. 127: p. 511-520.

[20] Biau, G., Scornet, E., A random forest guided tour. Test, 2016. 25(2): p. 197-227.

[21] Wimalagunaratne, M., Poravi, G., A predictive model for the global cryptocurrency market: A holistic approach to predicting cryptocurrency prices. 2018. In 2018 8th International Conference on Intelligent Systems, Modelling and Simulation (ISMS).

[22] Priyadarshini, I., Puri, V., A convolutional neural network (CNN) based ensemble model for exoplanet detection. Earth Science Informatics, 2021. 14(2): p. 735-747.

[23] Ramadhani, A. M., Goo, H. S., Twitter sentiment analysis using deep learning methods. 2017. In 2017 7th International annual engineering seminar (InAES), IEEE, p. 1-4.

[24] Gers, F. A., Schmidhuber, J., Cummins, F., Learning to forget: Continual prediction with LSTM. Neural computation, 2000. 12(10): p. 2451-2471.

[25] Graves, A., Fernández, S., Schmidhuber, J., Bidirectional LSTM networks for improved phoneme classification and recognition. 2005. In International conference on artificial neural networks, p. 799-804.

[26] Aslan, S., Kaya, M., Topic recommendation for authors as a link prediction problem. Future Generation Computer Systems, 2018. 89: p. 249-264.

NATURENGS
MTU Journal of Engineering and Natural Sciences
https://dergipark.org.tr/tr/pub/naturengs
DOI: 10.46572/naturengs.1143413

Review Article

Wild Plants Consumed as Food and Medicine in Elazıg and Malatya

Ayse Biçer*
Department of BioEngineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Özal University, Malatya, Turkey.

(**Received:** 12.07.2022; **Accepted:** 12.10.2022)

ABSTRACT: This study investigated the vitamin contents and health benefits of nine wild plants consumed as food, medicine, or animal feed in and around Elazığ and Malatya, Turkey. The study introduced *Rheum ribes, Gundelia tournefortii, Tragopogon reticulatus, Urtica dioica, Asphodelus aestivus, Medicago sativa, Nasturtium officinale, Mentha pulegium,* and *Capsella pastoris*. The results showed that the nine wild plants were common ingredients in Turkish cuisine and had numerous health benefits

Keywords: Wild plants, Turkish cuisine, Food, Vitamins, Medicine

1. INTRODUCTION

Nature offers self-grown medicinal plants. Wild plants grow by themselves wherever conditions permit. People use wild plants for cooking vegetable dishes [1]. Environmental conditions are constantly changing, and the world's population is overgrowing. Therefore, more and more people turn to wild plants to meet their nutritional needs. Dishes made from edible wild plants have become a significant part of Turkish culinary culture. Turkish people consume edible wild plants, either raw or cooked. They sometimes bring wild plants to a boil and then sieve and add bulghur or rice to them. Some people eat wild plants with eggs, yogurt with garlic, or as stuffed dishes [2].

There are about 800,000 plant species worldwide and more than 9,000 in Turkey [3]. Throughout history, humans have been consuming plants for nutritional and medicinal purposes and using them to make dyes, resin, gum, or soft drinks. Plants are also used in the cosmetic and ornamental plants industries [4, 5]. Plants have been used to treat diseases for thousands of years. Humans have passed down that knowledge over successive generations [6]. Today, more and more people turn to natural products for various reasons. First, people have become more and more concerned with their health. Second, some synthetic drugs fail to be effective or show serious side effects. Many medicines are made of natural herbs and plant extractions [7, 8].

People living in the Eastern Anatolia Region of Turkey commonly consume wild plants as food [6]. Particularly those living in Van, Hakkari, Ağrı, and Bitlis widely use wild plants as food [9]. The local people of Elazig and Malatya use wild plants as raw or cooked vegetables in salads and jams or as spices when fruits and vegetables are expensive or scarce [10]. Wild plants are rich in vitamins, minerals, fiber, protein, and antioxidants [2].

*Corresponding Author: ayse.bicer@ozal.edu.tr
ORCID number of authors: 0000 -0003-4514-5644

Although wild plants have been a part of our local food system for centuries, urbanization has led people to adopt more sedentary lifestyles and consume more processed foods. We must register plant species to preserve our cultural heritage and pass it on to future generations. This study focused on nine wild plants [*ışkın* (*Rheum ribes*), *kenger* (*Gundelia tournefortii*), *yemlik* (*Tragopogon reticulatus), ısırgan (Urtica dioica), çiriş (Asphodelus aestivus l), yonca (Medicago sativa), su teresi - acice (Nasturtium officinale), su yarpuzu (Mentha pulegium), kuşekmeği (Capsella pastoris*)] consumed by the local people of Elazig and Malatya. The study investigated how and why they consumed those plants.

2. WILD PLANTS AS FOOD

This section introduced the nine wild plants consumed by the local people of Elazig and Malatya.

*Işkın (*Syrian rhubarb; *Rheum ribes)*

Işkın (Syrian rhubarb*; Rheum ribes*) is a perennial herbaceous plant that grows in mountainous regions. It grows by itself in the spring in the area [11]. It is a member of the family of *Polygonaceae* [12]. It is the only *rheum* species that grows between 1800 and 2800 altitude in rocky meadows in Turkey. It can grow as tall as 40 cm. It has leaves attached and parallels the ground in the lower parts. It has a leafless stem extending as a middle rod [13]. At the top of it are spike-shaped yellow-green flowers. There are hairs on the outer surface of the stems (Fig 1).

Fig. 1 *Işkın (*Syrian rhubarb*; rheum ribes)*

Işkın contains vitamins A, B1, B2, C, D, E, and K. It is also used as a laxative in rural areas because it contains fiber [14]. It tastes bitter and sour. Experts recommend that it not be consumed too much because it is rich in oxalic acid [15]. It should be collected before the bud stage. People peel the hairy bark of its stems and shoots, salt it, consume it raw, or cook it with minced meat or eggs. It is put in cakes and muffins in some countries (Germany, England, Sweden, etc.). It is also used to make compote and jam [16]. It contains high levels of vitamin C. When cooked and eaten, it prevents the spread of infections to other tissues and helps enlarged tonsils get smaller. It is suitable for runny nose, headache, and fatigue. It also increases the body's resistance to winter diseases.

Kenger (*Gundelia tournefortii)*

Kenger (*Gundelia tournefortii)* belongs to the family of *Asteraceae*. It is found in the temperate regions of Western Asia [17]. It is also used for medicinal purposes in Cyprus, Egypt, Iran, Israel, Turkey, Azerbaijan, and Turkmenistan [18]. It grows in many regions of Turkey. It grows by itself in the mountainous areas and high latitudes in April and May [13].

Kenger is perennial, hairy, milky, prickly, and herbaceous. It has a few branches in its trunk. It is short and thick. It has leathery, veined, whitish leaves and purplish-red flowers (Fig 2). Its stem is cut to obtain milk, which is used to make *kenger* gum, which is very expensive. Therefore, it is a precious plant. People bring it to a boil and make a salad from it. It is a delicious plant with numerous health benefits. The heads ripening in the Mediterranean region are roasted and ground to make *kenger* coffee. Its leaves, stem, roots, and seed are consumed as food. Its seed is rich in crude oil, protein, and fiber [17].

Fig. 2 *Kenger* (*Gundelia tournefortii)*

Many people consume *kenger* because it possesses therapeutic dynamics against hypertension and diabetes and its other health benefits. It is used to prevent toothaches and gum ailments and to whiten teeth. It helps pass a kidney stone. It is suitable for fatigue and anemia. One of its biggest benefits is allowing the body to produce antibodies, which results in cell regeneration. It has antioxidants that help the body flush out harmful toxins.

Yemlik (*Tragopogon reticulatus*)

Yemlik (*Tragopogon reticulatus*) belongs to the family of *Asteraceae*. It is a perennial herbaceous plant [16]. It grows in forage meadows, plateaus, and mountains that develop with spring rains (Fig 3). It starts to leaf out as of April and can be consumed afterward. However, it must be collected and consumed before flowering. It is rich in vitamins (A, B2, B6, C, and E) and minerals (iron, calcium, etc.) [19]. It is collected by hand or by cutting with a knife. It is consumed in salads or with bulgur or rice [16].

Yemlik is good for anemia as it is rich in iron. It regulates the digestive system, increases appetite, and prevents weight gain. It also helps strengthen the bones due to its calcium and vitamin A. It is also suitable for eye problems (nyctalopia, etc.) as it is rich in vitamin A.

Fig. 3 *Yemlik* (*Tragopogon reticulatus*)

Isırgan (Common Nettle; *Urtica dioica*)

Isırgan (common nettle; *Urtica dioica*) grows in the mountains, by the walls, and in ruins. It is also known as *dancak otu*, *gidişgen*, *sırgan*, etc. [3]. It has dark green leaves. It is a petiolate plant with an inflammatory effect on the skin (stinging and burning sensation). Its leaves are rich in minerals, chlorophyll, amino acids, lecithin, carotenoids, flavonoids, sterols, tannins, and vitamins. It is consumed locally in meals, soups, and salads (Fig 4).

Fig. 4 *Isırgan* (Common Nettle; Urtica dioica)

It has many health benefits ranging from the digestive to the immune system. Isırgan increases blood circulation strengthens the immune system, protects against diabetes, and cleans the blood. It is rich in vitamins A and C. It prevents the formation of free radicals as it contains antioxidants. It is therapeutic for seasonal allergies. It helps slow down the cancer-free growth of the prostate gland in men by affecting hormone levels or interacting with cells in the prostate. It nourishes the scalp and makes the hair look shinier and healthier because it contains silica and sulfur. It ensures that calcium, which is vital for bone development and health, remains in the body. It can prevent or slow down the onset of osteoporosis.

Çiriş (Asphodelus aestivus L).

Çiriş (*Asphodelus aestivus L.*) grows in Turkey's mountains of East, Southeast, and Central Anatolia. It is widely consumed in those regions [20]. Its leaves are similar to leek leaves but much smaller than a leek. It is also known as *yabani pırasa* (wild leek), *güllük*, *yeling otu,* and *sarı zambak* (yellow lily) [21]. It is widely consumed as it is rich in vitamins, antioxidants, fibers, and minerals and has nutritional properties [22]. Its root, flowering stem, and seeds are

consumed. Its leaves are also used in dishes and canned goods (Fig 5). People consume it in stews, soups, salads, roasts, pilafs, and pastries. It is dried and ground and used in powder form. The most popular dishes made with it are rice with *çiriş* and fried egg with *çiriş*.

Fig. 5 *Çiriş Otu* (*Asphodelus aestivus L.*)

Çiriş contains vitamins A, C, B1, B2, B3, B6, and B9 [23, 24]. It has numerous health benefits. It treats hemorrhoids, rheumatism, eczema, acne, and ringworm. It is a diuretic plant that also helps with menstruation and milk production. Its roots treat jaundice, liver disorders, stomach irritation, acne, and bone fractures [25].

Çiriş is used as an adhesive in bookbinding and shoemaking. It is also used to give hardness and shine to *ehram* fabric in the Erzurum region.

Yonca (*Medicago sativa)*

Yonca (Medicago sativa) belongs to the family of *Leguminosae*. It is rich in proteins, minerals (calcium, etc.), and vitamins B, C, D, E, and K. It is consumed as fresh or dried. It has numerous health benefits. However, it is mainly used as animal feed (Fig 6). It is one of the most nutritious forage plants.

Fig. 6 *Yonca (Medicago sativa)*

Experts recommend that *yonca* be fed to young animals because it plays a vital role in bone development.

Yonca is also used for medicinal purposes. It is an appetite-stimulating plant that gives strength and energy and helps the stomach work regularly. It calms the nerves, relieves severe headaches, reduces rheumatic pains, stops diarrhea, and stimulates breast milk production. It is also good for anemia. Local people also argue that it prevents Parkinson's disease.

Su Teresi - Acice **(Watercress; *Nasturtium officinale)***

Su teresi - Acice (watercress; *Nasturtium officinale) belongs to the family of Cruciferae, although it looks like a flowery plant* (Fig 7). *Su teresi* is a kind of mint that grows in wet environments. Besides giving flavor to salads, it is perfect for health. *Su teresi*, a member of the Brussels sprout family, has a slightly spicy flavor.

Fig. 7 *Su Teresi - Acice* (watercress; *Nasturtium officinale)*

Su teresi boosts immunity as it is a potent antioxidant rich in vitamin C and folic acid. It prevents the risk of diabetes, cancer, and heart disease. Its most important feature is that it is low in calories. It is very healthy for people with high blood pressure. It helps lower bad cholesterol. It also regulates the liver, increases sexual potency, regulates the menstrual cycle, and prevents bleeding and cardiovascular and eye diseases (cataracts, etc.). It improves brain functions and prevents Alzheimer's-like diseases. It makes the teeth healthier and stimulates bone development as it contains calcium. It also improves vision and cognitive function.

Su yarpuzu **(Pennyroyal; *Mentha pulegium*)**

Su yarpuzu (pennyroyal; *Mentha pulegium*) belongs to the family of *Lamiaceae*. It is also known as *filiskin*, *narpuz*, *pülüskün*, *yabani nane* (wild mint), or *su nanesi* (water mint) (Fig 8). It is a refreshing plant like mint. Not only is it tasty, but it also has health benefits. Therefore, it is widely used (dried or fresh) in teas, dishes, salads, and soups. In some regions, people add it to *tarhana*.

Fig. 8 *Su yarpuzu* (Pennyroyal; *Mentha pulegium*)

Su yarpuzu is an antispasmodic and expectorant plant that prevents infections and regulates menstruation. It helps aid breathing and acts as a decongestant. It is good for the heart and stomach. It is also used to treat common colds, sinusitis, cholera, food poisoning, bronchitis, and tuberculosis [26].

***Kuşekmeği* (shepherd's purse; *Capsella pastoris*)**

Kuşekmeği (shepherd's purse; *Capsella pastoris*) belongs to the family of *Cruciferae*. It is a short plant almost stuck to the ground in the pre-bloom period. It is known for its serrated leaves. Its leaves are edible (Fig 9). As the season progresses, those leaves turn yellow, and a small ball of tiny white flowers appears on the branch that emerges from the middle. It is at most half a meter long. It has heart-shaped mini leaves. It grows in uncultivated fields, gardens, and meadows. It is also known as *kuşkuş otu*, *guşguş*, *acıbici*, *dağ marulu* (mountain lettuce), *kuş mancası*, or *çobançantası* (shepherd's purse).

Fig. 9 *Kuşekmeği* (shepherd's purse; *Capsella pastoris*)

Kuşekmeği can be eaten raw at very young periods but is usually cooked. It is cooked with minced meat, tomato paste, and bulgur. It is also fried in oil. It is included in mixed herb roasts. It can be mixed with vegetables and consumed as a winter soup. It is a kind of gözleme (Turkish stuffed pancake). It is cut into small pieces, placed between thin phyllo dough, and cooked on sheet metal. It is brushed with butter or oil. It is a good source of vitamins.

Tea is made by brewing its fresh and dry leaves and flowers. *Kuşekmeği* tea helps prevent bleeding and drop blood pressure. It is used for heart and circulatory problems and nervous heart complaints. It is good for urinary tract infections and stomach, kidney, and uterus-related ailments. In addition, tea made from its seeds and leaves stops diarrhea and nosebleeds and treats hemorrhoids and dysentery.

3. RESULTS AND DISCUSSION

This study focused on nine wild plants consumed by the local people of Elazig and Malatya for nutritional or medicinal purposes. We conducted a literature review and interviews with the local people of Elazig and Malatya to understand how and why they consumed those plants.

Işkın is rich in vitamins A, B1, B2, C, D, E, and K. It can be eaten raw or cooked with minced meat or eggs. It treats leukemia, cough, antipyretic, runny nose, headache, and fatigue. However, experts recommend that it not be consumed too much because it is rich in oxalic acid (which may be mildly irritating to the whole body).

Kenger is widely consumed for its health benefits and therapeutic properties. It is good for blood pressure and diabetes. It helps pass a kidney stone. It is used to prevent toothaches and

gum ailments and to whiten teeth. It is rich in vitamins A, E, and C. It is consumed raw in salads or cooked. It is also used to produce chewing gum.

Yemlik is rich in vitamins (A, C, E, B1, B2, B3, and B6) and minerals (iron, calcium, etc.). It is consumed raw in salads or cooked with bulgur and rice. It is good for anemia, nyctalopia, and the digestive system. It is an appetite-stimulating plant that helps strengthen bones. However, it must be harvested and consumed before blooming due to its pubescence and hardening.

Isırgan is rich in vitamins A and C. It is used in dishes, soups, and salads. It prevents the formation of free radicals and boosts immunity as it contains antioxidants. Therefore, it reduces the risk of cancer, heart attack, and stroke. It helps reduce anemia, allergies, prostate disorders, and kidney and gallbladder stones.

Çiriş is consumed raw or cooked. It contains vitamins A, C, B1, B2, B3, B6, and B9. It is good for hemorrhoids, rheumatism, ringworm, eczema, jaundice, liver disorders, stomach irritation, acne, and bone fractures. It is used to regulate menstruation. It is also a diuretic.

Yonca is rich in protein, minerals (calcium), and vitamins (B, C, D, E, and K). It can be consumed raw in salads or cooked with bulgur and rice. It can also be boiled to make tea. It is one of the most nutritious forage plants. It is an appetite-stimulating plant that helps relieve stomach problems and increase breast milk production. It is good for headaches, rheumatic pain, diarrhea, and anemia.

Su teresi can be consumed raw or cooked. It is rich in vitamins A, B6, C, E, and K. It contains antioxidants. Therefore, it reduces the risk of diabetes, cancer, high blood pressure, and heart disease. It prevents bleeding and cataracts. It helps regulate the liver and improve cognitive function.

Yarpuz is widely used (dried or fresh) in dishes, salads, and soups. It can also be boiled to make tea. It is rich in vitamins A, B1, B2, B3, B6, C, and E. It is good for common colds, sinusitis, cholera, food poisoning, bronchitis, and tuberculosis. It is an antispasmodic and expectorant plant that prevents infections and regulates menstruation. It helps aid breathing and acts as a decongestant.

Kuşekmeği can be boiled to make soup or cooked with bulgur. It is rich in vitamins B and C. It helps regulate blood pressure and menstruation. It is good for constipation, hemorrhoids, and tooth and throat aches. It is also a diuretic.

4. CONCLUSIONS

The following are the conclusions based on the results:

✓ All wild plants in question are traditional plants that have been consumed (raw or cooked) by Turkish people, especially those living in and around Elazig and Malatya, for hundreds of years.

✓ Although wild plants have health benefits, they should not be consumed too much.

✓ All wild plants in question are rich in vitamins B and C. All wild plants but *yonca* and *kuşekmeği* contain vitamin A. All wild plants but *çiriş* and *kuşekmeği* contain vitamin E. *Işkın*, *ısırgan*, *yonca*, and *su teresi* are rich in vitamin K.

✓ *Isırgan*, *ışkın*, and *su teresi* boost immunity and help prevent diabetes, cancer, and heart disease.

✓ *Yonca*, *yemlik*, and *kuşekmeği* help regulate the digestive system and relieve headaches, toothaches, and rheumatic pain.

✓ *Işkın*, *su yarpuzu*, and *ısırgan* are good for lung and respiratory ailments.

In conclusion, the local people of the Elazig-Malatya region have been consuming wild plants for nutritional and medicinal purposes for hundreds of years.

Acknowledgments

This work was not supported by any institution or organization.

Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could influence the work reported in this paper.

Author Contribution

Ayşe Biçer contributed 100% at every stage of the article.

REFERENCES

[1] Tuzlacı, E., Türkiye'nin yabani besin bitkileri ve ot yemekleri. 2011. Alfa yayınları**,** Document Type: Book, ISBN: 9786051063485, Istanbul.

[2] Yücel, E., Şengün, İ., Çoban, Z., Afyonkarahisar çevresinde gıda olarak tüketilen yabani otlar ve tüketim biçimleri. Biological Diversity and Conservation, 2012. 5(2): p. 95-105.

[3] Baytop, T., Türkiye'de bitkiler ile tedavi (Geçmişte ve Bugün). 1984. İstanbul Üniv. Yayınları No: 3255, Eczacılık Fak. Yayınları No: 40, İstanbul.

[4] Erdogrul, O.T., Ates, A., Antimicrobial activities of various medicinal and commercial plant extracts, Turkish Journal of Biology, 2003. 27: p. 157–163.

[5] Faydaoğlu, E., Surucuoglu, M.S., Geçmişten günümüze tıbbi ve aromatik bitkilerin kullanılması ve ekonomik önemi. Kastamonu Üniversitesi Orman Fakültesi Dergisi, 2011. 11(1): p. 52-67.

[6] Ozturk, M., Ozcelik, H., Useful plants of East Anatolia. 1991. Siirt İlim, Spor, Kültür ve Arastırma Vakfı, Document Type: Book, Ankara.

[7] Yücel, E., Tapırdamaz, A., Yücel Şengün, İ., Yılmaz, G., Ak, A., Determining the usage ways and nutrient contents of some wild plants around Kisecik Town (Karaman/Turkey). Biological Diversity and Conservation, 2011. 4(3): p. 71-82.

[8] Tosun, F., Kızılay, A., Sener, Ç., Vural, B., Palittapongarnpim, M., Antimycobacterial activity of some Turkish plants. Pharmaceutical Biology, 2004. 42: p. 39- 43.

[9] Okcu, Z., Kaplan, B., Doğu Anadolu Bölgesinde gıda olarak kullanılan yabani bitkiler. Türk Tarım-Gıda Bilim ve Teknolojisi Dergisi, 2018. 6(3): p. 260-265.

[10] Siyamoglu, B., Ege Bölgesinde insan beslenmesinde kullanılan bazı yabani otlar (silcan, karakan, pirzola kekiği ve kudret narı) üzerinde araştırmalar. Ege Üniversitesi Ziraat Fakültesi Dergisi, 1984. 21(3): p. 75–88.

[11] Munzuroglu, O., Karatas, F., Gür, N., Isgın (Rheum ribes L.) bitkisindeki A, E ve C vitaminleri ile selenyum düzeylerinin araştırılması. *Turk J Biol* 24, 2000. p. 397–404, TÜBITAK.

[12] Shokravi, A., Agha, N.K., Synthesis of 1,2,3,4,5,6,7,8-octahydro-9-ethoxy 10-hydroxy-1-anthracenone (OEHA), Iranian Journal of Chemistry and Chemical Engineering, 1997. 16: p. 10–15.

[13] Yıldız, S., Yukarı Fırat Havzasında yetişen kenger (gundelia tournefortii L.), güllük (eremurus spectabilis m. bieb.) ve ışkın (rheum ribes L.) bitkilerindeki polifenollerin ve bazı metallerin tayini, Yüksek Lisans Tezi, 2014. Fırat Üniversitesi, Fen Bilimleri Enstitüsü.

[14] Meral, N., The effect of different temperatures on antioxidant activity and phenolic profile of the rheum ribes, Yüzüncü Yıl Üniversitesi Tarım Bilimleri Dergisi (YYU J AGR SCI), 2017. 27(1): p. 88-94.

[15] Atasoy, N., Van bölgesinde yetişen endemik bitkilerde pro-vitamin A(B-karoten) tayini, Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 2010. 15(2): p. 134-142.

[16] Doğan, S., Gevaş (Van) ilçesinde yöresel olarak taze tüketilen bazı yabani bitkiler ve besin değerlerinin belirlenmesi, Yüksek Lisans Tezi, 2016. Yüzüncü Yıl Üniversitesi Fen Bilimleri Enstitüsü Tarla Bitkileri Anabilim Dalı.

[17] Nasrollahzadeh, M., Mahamb, M.S., Mohammad, S., Green synthesis of CuO nanoparticles by aqueous extract of Gundelia tournefortii and evaluation of their catalytic activity for the synthesis of N-monosubstituted ureas and reduction of 4-nitrophenol. Journal of Colloid and Interface Science, 2015. 455: p. 245–253.

[18] Haghi, G., Hatami, A., Arshi, G., Distribution of caffeic acid derivatives in gundelia tournefortii L, Food Chemistry, 2011. 124: p. 1029–1035.

[19] Çöteli, E., Erden, Y., Karataş, F., Investigation of amounts of malondialdehyde, glutathione and vitamins with total antioxidan capacity in plant mentha polegium L., Suleyman Demirel University Journal of Natural and Applied Science, 2013. 17(2): 4-10.

[20] Oskay, M., Aktas, K., Sari, D., Azeri, C., A comparative study of antimicrobial activity using well and disk diffusion method on asphodelus aestivus (liliaceae*)*, Ekoloji, 2007. 16(62): p. 62-65.

[21] Gülçin, İ., Oktay, M., Kireçci, E., Küfrevioğlu, Ö.İ., Screening of antioxidant and antimicrobial activities of anise (pimpinella anisum L.) seed extracts, Food Chemistry, 2003. 83: p. 371-382.

[22] Badayman, M., Dinçel, E., Alçay, A.Ü., Ciris herb and its use in Turkish cuisine, Aydın Gastronomy, 2018. 2(1): p. 51-55.

[23] Karataş, F., Bektaş, İ., Birişik, A., Aydın, Z., Kurtul, A., Çiriş otu'nda (*asphodelus aestivus* l.) suda çözünen bazı bileşiklerin araştırılması, SDU Journal of Science, 2011. 6(1): p. 35-39.

[24] Peksel, A., Imamoglu, S., Antioxidative properties of extracts from Asphodelus aestivus brot *(Liliaceae), Annals of Nutrition and Metabolism*, 2009. 55: p. 596 -596.

[25] Pourfarzad, A., Najafi, M.B.H., Khodaparast, M.H.H., Khayyat, M.H., Malekpour, A., Fractionation of eremurus spectabilis fructans byethanol: Box– Behnken design and principal component analysis, *Carbohydrate Polymers*, 2014. 106: p. 374-383.

[26] Çöteli, E., Karataş, F., Investigation of amounts of glutathione and vitamins with total antioxidant capacity in leaves of plant tragopogon reticulatus, GÜFBED/GUSTIJ, 2015. 5(2): p. 78-86

NATURENGS
MTU Journal of Engineering and Natural Sciences
https://dergipark.org.tr/tr/pub/naturengs
DOI: 10.46572/naturengs.950220

Research Article

A Simple Expression for the Refractive Index of Distilled Water

Ramazan Emre ODUNCUOGLU[1*], Murat ODUNCUOGLU [2]

[1]Department of Computer Engineering, Faculty of Engineering, Alanya Alaaddin Keykubat University, Antalya, Turkey.
[2]Physics Department, Sciences and Arts Faculty, Yıldız Technical University, Istanbul, Turkey.

(**Received:** 09.06.2021; **Accepted:** 19.09.2022)

ABSTRACT: Water is conceivably the most important material in the universe and essential to the functioning of all the known life forms. The refractive index is an important physical parameter that can be used to characterize and understand the properties of water. Therefore, accurate knowledge of the refractive index of water is of great importance. A simple expression for the real part of the refractive index of water was investigated and a new equation was proposed as a function of the temperature of distilled water between 0 and 100 ºC and wavelengths in the range of 200 to 1100 nm. Water is transparent in visible light and has a complex optical absorption property in the infrared and ultraviolet spectra. The refractive index highly depends on wavelength and temperature. The expression for the refractive index is helpful for different applications in biomedical optics. The proposed formula derived by using genetic programming methods has accurate terms, has a good agreement, and demonstrates increased performance with experimental measurements for calculations of knowledge of the refractive index of water at given ranges.

Keywords: Water, Temperature, Wavelength, Modelling, Refractive index.

1. INTRODUCTION

Water is an essential biological liquid. The liquid component of blood is about 90% water in plasma form. Human blood consists of approximately 54.3% of plasma. The water is more than half of the total blood volume. Water is presented in free and bound states with biomolecules in biological tissues. The water molecules can be combined with blood contents such as proteins, different concentrations of glucose, and other components. The substances were dissolved and transmitted into and out of a cell. Blood sample analysis is very important for the diagnosis of various diseases. The observed changes in the optical properties of water have advantages in medical and optical diagnostic techniques. The change in liquid refractive index, n, can be calculated by different experimental methods [1]. An expression overcomes the difficulties of determining the refractive index in less time for the given ranges.

The working principle, performance, and optimization of the sensors and new devices were determined and characterized by physics and responses in selected conditions of the elementary material. The knowledge of the refractive index of biological solutions and clinically relevant materials facilitates the design and optimization of new sensors [2]. In addition, determining

*Corresponding Author: 180254023@ogr.alanya.edu.tr
ORCID number of authors: [1]0000-0001-7014-0453, [2]0000-0002-3130-5646

the optical properties of water is an essential prerequisite to understanding the physical and chemical properties of aqueous environments [3]. In a laboratory, the refractive index was measured at a constant temperature. For real practical applications, the refractive index measurement as a function of temperature and wavelength simultaneously must fulfill some important requirements of the sensor research. The properties of different phases of water have been extensively investigated by scientists and engineers as a function of the wavelength (λ) or temperature (T) for several decades [3-7]. Measurement of the optical refractive index as an inherent characteristic of the substance is a formal investigation in research [8]. The refractive index represents the response of the electronic charge distribution to the disturbance brought about by the electric field component of incident electromagnetic radiation [9]. Water molecules are polar and if exposed to an electromagnetic wave, they will swing around and try to stay lined up with the alternating electric field. The properties of water play a role in understanding the electron charge transfer rates and behavior of aqueous solutions and determining the oxidation states of minerals and rocks used throughout science and industry [10]. These two parameters are the most known factors affecting the refractive index by modifications in the band gap energy and dielectric constant of water.

The structural and transport properties of the most abundant molecule on the earth's surface have been investigated to understand the material properties. In this article, the expression for the refractive index of distilled water formula as a function of wavelength and temperature has been studied using genetic programming in the given ranges. The genetic programming provided experimental data to calculate the desired parameter of given problems [11]. This is a powerful modeling technique for the simple expression of the water refractive index. The method adapted from the Darwinian evolution theory of natural selection has chromosomes and gene expression trees [12, 13]. The iterative process is used to generate an expression with a given accuracy. Genetic programming is a tool for predicting parameters in medicine, social sciences, and other disciplines. Genetic programming was explained in books and papers [13-18]. It aimed to be in the best fitness zone, minimize errors, simplify, and find the best predicted results [19]. A combination of these was investigated to increase the algorithm's performance to predict output parameters. The accuracy of the model was evaluated using statistical parameters such as root mean square error (RMSE), which is the best model with the lowest value or mean absolute error (MAE). The coefficient of correlation and explanations of these parameters have been provided in detail Ref. [19]. The correlation coefficient (R) was used to calculate the model performance. The calculated results of this study with these parameters were in good agreement with the experimental results.

2. MATERIAL AND METHODS

Common computing techniques were used to propose new expressions and solve real-life problems. The latest and high-performance computers and advanced techniques could solve problems and determine parameters for scientists and computer programmers. The genetic programming technique created the water refractive index expression in this study. The variables consist of temperature and wavelength as input and refractive index for output. The function sets used were +, −, *, /, $\sqrt{}$, exp(x), x^2, x^2, $x^{1/3}$, with seven head sizes and three genes were used for the proposed simple expression of calculating the n(T, λ). The number of training samples was 70% of the total samples (155), and 25% for validation, with 5% for testing. The training data were used to implement and build a model for the sake of the remastered version and the validation data was used to validate the results with the renewed model. 5% of the data points in all the data are excluded from the dataset and are used to test the proposed expressions. Validation and test data were chosen randomly and equally spaced in these data sets.

Table 1. The genetic programming parameters used for the proposed model.

Statistical Coefficient	Training	Validation
R^2	0.93510	0.94291
Mean Squared Error (MSE)	$3.1\ 10^{-5}$	$3.4\ 10^{-5}$
Mean Absolute Error (MAE)	0.00531	0.00483
Relative Absolute Error (RAE)	0.25895	0.28174
Root Mean Squared Error (RMSE)	0.00617	0.00584
Correlation Coefficient	0.96700	0.97103

The correlation coefficient, R^2, showed the flawless fit of the data and how it is greater than 0.93 in Table 1. These values demonstrated that the proposed model's performance is highly accurate. The correct number of training and test data depends on the task, the desired performance, the input features, the noise in the training data, the noise itself, the complexity of the model, and so on. The bigger the dataset, it is that much more preferred to train it. The employed computational method has the potential to save costs while broadening accessibility.

3. RESULTS AND DISCUSSION

The change in refractive index in water is not attributable to temperature because it is not linear [9, 10]. The expression for the refractive index was collected from the literature [9, 20-22]. The performance of genetic programming by comparing the experimental and the predicted values are given in Fig. 1. It is observed that the critical trends in formulation proposed by the genetic programming formula for the refractive index in overall performance are successful. There are different formulations for determining the water refractive index but no direct, explicit formulations related to temperature and wavelength.

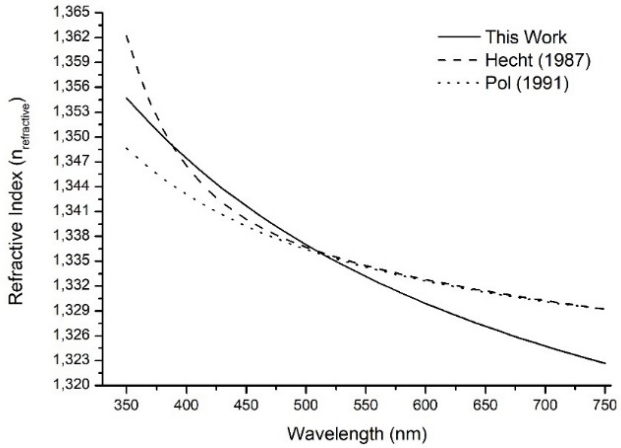


Figure 1. The comparison of equations for the refractive index of water.

The wavelength dependence of the water refractive index in the visible region was used to compare the results of the proposed expression given in the literature [23-25] as

$$n(\lambda) = A + \frac{B}{\lambda^2} - \frac{C}{\lambda^4} + \frac{D}{\lambda^6} \quad (1)$$

$$n(\lambda) = 1.31848 + \frac{6.662}{\lambda - 129.2} \quad (2)$$

where λ is a wavelength, the coefficients are A = 1.3199, B = $6.788\ 10^3$, C = $1.132\ 10^9$ and D = $1.11\ 10^{14}$ in Eqn. 1. The water was transparent at visible wavelengths. In Fig. 1, we

compare the refractive index of water as a function of wavelengths in the visible range of the spectrum at room temperature. The refractive index values are decreased with increasing wavelengths and our proposed formula's results show good agreement. The results of these equations have been obtained for a room temperature environment and the temperature was not included in Eqn. 1 and Eqn. 2. Water refractive index's temperature and wavelength dependence are critical for biomedical optical applications. The goal of this study is to obtain the simple expression for water refractive index as a function of temperature and wavelength

$$n_{refractive}(\lambda,T) = 1.28900 + \frac{16.45262\,\lambda + \lambda^{4/3} + 11.86573\,\lambda - 108.85351}{\lambda\,(\lambda + T - 9.173768)} \tag{3}$$

where λ is a wavelength, and T is the temperature. This expression is valid in the range from 200 nm to 1100 nm, and the temperature ranges from 0 °C to 100 °C.

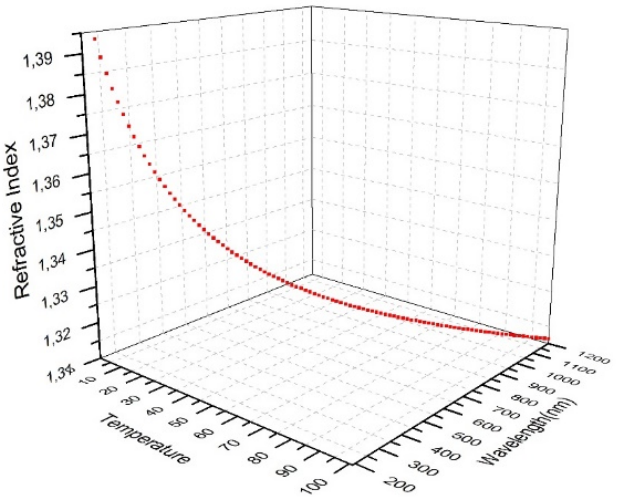


Figure 2. The synergistic effects of each parameter in refractive index, n(λ, T).

The data shows that the water refractive index decreases with the increase in both wavelength and temperature in Fig 2. The water refractive index was directly calculated and we overcame this limitation by using our proposed expression, which agreed with our experimental and predicted results. The Cauchy formula was used to determine the water refractive index dependency on temperature and wavelength. The coefficients in Eqn.1 were temperature-dependent and in the form of A(T), B(T), C(T) and D(T). There were variable coefficients for each temperature value and the variables were determined by curve fitting methods in Ref. [22].

In addition, the electronic band gap of water has been inferred from using the formula of its refractive index, *n*, which can be obtained by genetic programming. According to the single-oscillator model, the Penn model, the refractive and the electronic gap E_g are correlated as

$$n^2 = 1 + \frac{(\hbar\omega)^2}{{E_g}^2} \tag{4}$$

where ω is the frequency [26], this correlation between *n* and E_g holds for materials. Measurements of the band gap of water by light are not practicable under some conditions; therefore, establishing the correlation between the band gap and the refractive index is a crucial step in inferring band gaps from measured dielectric constants [27]. The image below compares the experimental and predicted values of Fig. 3 and Fig. 4.

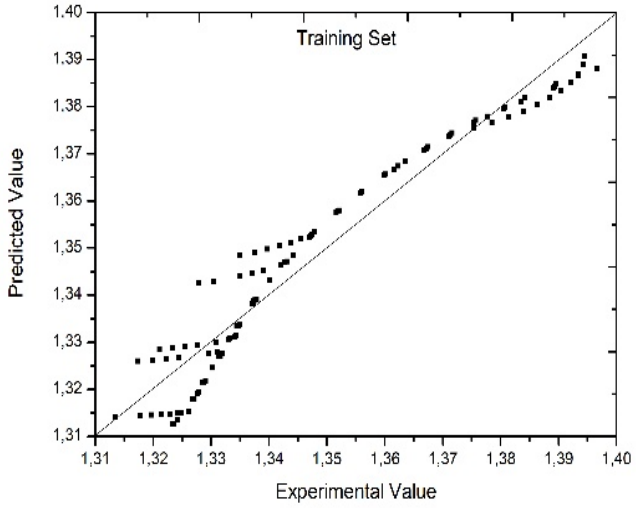


Figure 3. The values of water refractive index for the training dataset.

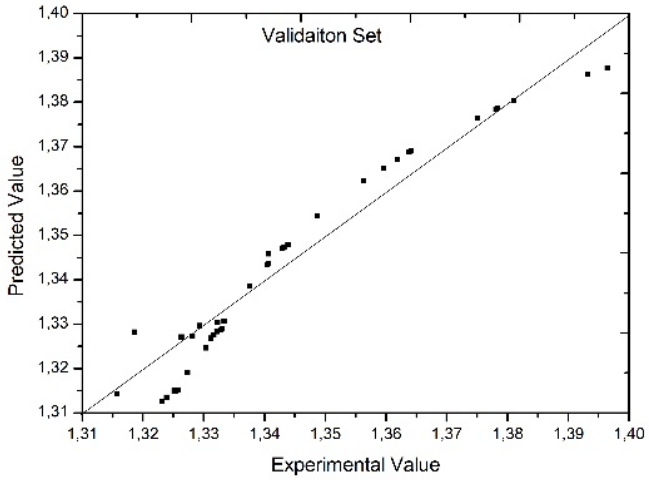


Figure 4. The values of water refractive index for validation dataset.

The refractive index of water decreasing with the temperature and wavelength can be easily shown in the proposed formula. Therefore, comparing the difference between test and reference substance is preferable to achieve reliable changes in the result. The water content of substances generally characterizes the refractive indices and the entire spectra exhibit additional strong features due to solvents in its contents. The effect of temperature on the refractive index is much smaller than that of solution or solvent alone. The index difference corresponds to distilled water and different liquids or water with various soluble additives dissolved inside. The sources of differences at low and high values on experimental and predicted uncertainties in the wavelength and temperature data caused values. It can be noted from Fig. 3 and Fig 4 the good agreement between our theoretical predictions with the experimental results. The results shown in such a figure suggest that our predicted formula and our results were almost ideal.

4. CONCLUSIONS

In this study, we proposed a simple expression for water refractive index as a function of wavelength and temperature. The genetic programming works well and provides optimized parameters of real parts of the refractive index. For optimizing and adjusting the parameters of new methods and techniques at given temperatures, the refractive index of water can be used

as a reference. It allows us to compare all the parameters with one another in calibration and measurement. The predictive capacity of the program is 93.51% of the training and 94.29% of the validation results. The proposed equations are user-friendly. The proposed formulation's statistical parameters (R2, MAPE, and RMS) show high potential for predicting the refractive index and using the results for diagnostic purposes. This study also developed and analyzed a simple linear refractive index formula of distilled water as a reference liquid for parameters. The results have potential in the future of biomedical device applications. The water-related science has a broad area, and these types of studies and results are essential for researchers. The optical properties of materials are necessary to determine and improve the prospects of device performance. The knowledge of such quantities presented here is expected for designing optical devices and their applications.

Acknowledgments
This work was not supported by any institution or organization.

Conflict of Interest
All authors declare that they have no conflicts of interest.

Author Contribution
REO. and MO. verified the analytical methods and REO. performed the computations. REO wrote the manuscript with support from MO. The authors contributed to the final version of the manuscript.

REFERENCES

[1] Li, G., et al., Fabry-Perot cavity enhanced prism for highly sensitive refractive index measurement of water. Optik, 2021. 245: p. 167688.

[2] Rowe, D.J., D. Smith, and J.S. Wilkinson, Complex refractive index spectra of whole blood and aqueous solutions of anticoagulants, analgesics and buffers in the mid-infrared. Scientific Reports, 2017. 7(1): p. 7356.

[3] Pan, D., Q. Wan, and G. Galli, The refractive index and electronic gap of water and ice increase with increasing pressure. Nature Communications, 2014. 5(1): p. 3919.

[4] Bartels-Rausch, T., et al., Ice structures, patterns, and processes: A view across the icefields. Reviews of Modern Physics, 2012. 84(2): p. 885.

[5] French, M. and R. Redmer, Optical properties of water at high temperature. Physics of Plasmas, 2011. 18(4): p. 043301.

[6] Fang, C., et al., The accurate calculation of the band gap of liquid water by means of GW corrections applied to plane-wave density functional theory molecular dynamics simulations. Physical Chemistry Chemical Physics, 2015. 17(1): p. 365-375.

[7] Zha, C.-S., et al., Optical study of H 2 O ice to 120 GPa: Dielectric function, molecular polarizability, and equation of state. The Journal of chemical physics, 2007. 126(7): p. 074506.

[8] Domanski, A.W., M. Roszko, and M. Swillo. Compact optical fiber refractive index differential sensor for salinity measurements. in IEEE Instrumentation and Measurement Technology Conference Sensing, Processing, Networking. IMTC Proceedings. 1997.

[9] Aly, K.M. and E. Esmail, Refractive index of salt water: effect of temperature. Optical Materials, 1993. 2(3): p. 195-199.

[10] Tilton, L. and J. Taylor, Refractive index and dispersion of distilled water for visible radiation, at temperatures 0 to 60 C. J. Res. Natl. Bur. Stand., 1938. 20: p. 419-477.

[11] Oduncuoğlu, M., Refractive Index Formula of Blood as a Function of Temperature and Concentration. Anais da Academia Brasileira de Ciências, 2021. 93.

[12] Koza, J., On the programming of computers by means of natural selection. Genetic programming, 1992.

[13] Koza, J.R., Genetic programming II: automatic discovery of reusable programs. 1994: MIT press.

[14] Koza, J.R., What is genetic programming (GP). How Genetic Programming Works, 2007.

[15] Booker, L.B., D.E. Goldberg, and J.H. Holland, Classifier systems and genetic algorithms. Artificial intelligence, 1989. 40(1-3): p. 235-282.

[16] Goldberg, D.E., Genetic algorithms. 2006: Pearson Education India.

[17] Pouraliakbar, H., et al., Predicting the ultimate grain size of aluminum sheets undergone constrained groove pressing. The International Journal of Advanced Manufacturing Technology, 2016. 86(5): p. 1639-1658.

[18] Kurt, H.I. and M. Oduncuoğlu, Formulation of the effect of different alloying elements on the tensile strength of the in situ Al-Mg2Si composites. Metals, 2015. 5(1): p. 371-382.

[19] Kurt, H.I., M. Oduncuoğlu, and R. Asmatulu, Wear behavior of aluminum matrix hybrid composites fabricated through friction stir welding process. Journal of Iron and Steel Research International, 2016. 23(10): p. 1119-1126.

[20] Lide, D.R., CRC handbook of chemistry and physics. Vol. 85. 2004: CRC press.

[21] Bashkatov, A.N. and E.A. Genina. Water refractive index in dependence on temperature and wavelength: a simple approximation. in Saratov Fall Meeting 2002: Optical Technologies in Biophysics and Medicine IV. 2003. International Society for Optics and Photonics.

[22] Kohl, M., M. Essenpreis, and M. Cope, The influence of glucose concentration upon the transport of light in tissue-simulating phantoms. Physics in Medicine & Biology, 1995. 40(7): p. 1267.

[23] Rol, P.O., Optics for transscleral laser applications. 1991, ETH Zurich.

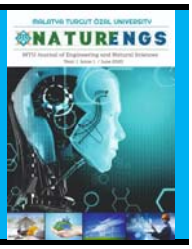
[24] Hecht, E., Optics. 2017: Pearson Education.

[25] Daimon, M. and A. Masumura, Measurement of the refractive index of distilled water from the near-infrared region to the ultraviolet region. Applied Optics, 2007. 46(18): p. 3811-3820.

[26] Penn, D.R., Wave-number-dependent dielectric function of semiconductors. Physical Review, 1962. 128(5): p. 2093.

[27] Oduncuoğlu, M., Optical Properties of Dilute Bismuth of Semiconductor Alloys. El-Cezeri Journal of Science and Engineering, 2020. 7(3): p. 1355-1361.

NATURENGS
MTU Journal of Engineering and Natural Sciences
https://dergipark.org.tr/tr/pub/naturengs
DOI: 10.46572/naturengs.1120218



Research Article

Equilibrium Studies for Dye Adsorption onto Red Clay

Muhammed ONAY[1*], Çiğdem SARICI ÖZDEMİR [2]

[1]Department of Engineering Basic Science, Faculty of Engineering and Natural Sciences, Malatya Turgut Özal University, Malatya, Turkey.
[2]Department of Chemical Engineering, Faculty of Engineering, Inonu University, Malatya, Turkey.

(**Received:** 23.05.2022; **Accepted:** 20.12.2022)

ABSTRACT: In this study, it was aimed to remove malachite green, a cationic dye, from aqueous solutions by adsorption method under various experimental conditions by using red clay. Red clay was used because it is abundant in nature and easily accessible in our region. In addition, it has been preferred as an adsorbent because it is used economically without any pre-treatment. The effects of initial malachite green concentration, temperature, adsorbent amount, and contact time on adsorption were evaluated. To examine the percent dye removal effect of the initial malachite green concentration, five different amounts of 25-150 ppm were examined. Five different parts, in the range of 0.5-2 grams, were reviewed to investigate the effect of red clay on the percentage of dye removal. Studies were conducted at four different values, 30-120 minutes, to find the optimum time for the adsorption process. In the adsorption process of red clay and malachite green, experiments were carried out at three different degrees, 298, 313 and 333 K. After reaching equilibrium in the adsorption process, the data obtained were analyzed and studies were carried out by applying them to isotherm models. The results obtained from the adsorption process were compared with Langmuir, Freundlich, Dubinin-Radushkevich and Temkin isotherm models. The experimental studies have been determined to be more compatible with the Dubinin-Radushkevich model. According to the results, it is incompatible with the Langmuir model. Accordingly, it can be said that the adsorption takes place in a multilayer and heterogeneous form.

Keywords: Dye removal, Red clay, Cationic dye, Isotherm.

1. INTRODUCTION

Recently, important events such as low carbon footprint, climate change and water scarcity have been on the agenda more and more. One of the main reasons that cause these negative situations is environmental pollution. Harmful gases and wastewater emitted into the atmosphere during production in industrial facilities are one of the factors that cause environmental pollution. In addition, it is known that water resources have decreased significantly today. In our country and region, where textile production is intense, the need for water is relatively high. Considering all these situations, the purification and reuse of the water used will significantly benefit both nature and the economy.

Industrial wastewater treatment processes to remove water pollutants such as dyes and heavy metals remain challenging for countries today. Paints are used in food, textile paper, cosmetics, etc. It has been used as a colorant in different industries. Even releasing minimal amounts of this paint into the environment poses a serious problem. Dye removal methods from textile

*Corresponding Author: muhammed.onay@ozal.edu.tr
ORCID number of authors: [1] 0000-0002-4276-0425, [2] 0000-0003-2129-3044

wastewater; chemical oxidation, foam flotation, adsorption, coagulation, electrodialysis. Among these methods, the adsorption method is known to have a better potential than any other listed process. Among these styles, adsorption seems more advantageous than different styles due to its low cost, easy design and installation, easy availability of the adsorbents used, and the ability to process dyes even at high concentrations. [1-5].

When the studies carried out using activated carbon are examined, it has been proven that it is very effective in removing dyestuffs by the adsorption method. However, since activated carbon is still expensive to produce or procure, there is a trend toward using low-cost adsorbents. Agricultural and industrial by-product wastes can be considered in the class of low-cost adsorbents. In addition, adsorbents such as soil and clay, which are used in their current form without any treatment in nature, are preferred because of their low cost, abundance and ability to remove dyestuffs and metal ions from aqueous solutions. [6, 7].

Clay is a soft and very fine-grained (fine than sand) material. The atoms in the clay material are either in the form of a lattice or a chain arrangement. The main ingredient of clay is aluminum silicate hydrate; depending on its type, compounds of other elements such as sodium, potassium, calcium, magnesium and iron can be found. Depending on the mineral content and chemical composition of the minerals, the color of the clays can be in various shades of white, gray, green, red and brown. Clays can be used as natural adsorbents [8].

This work investigated the adsorption of malachite green dye in an aqueous medium using red clay as an adsorbent in the batch system. Issues such as the adsorption efficiency of parameters such as initial dye concentration, amount of adsorbent, temperature and total contact time were discussed. The results of the studies on the adsorption equilibrium and the obtained findings were compared with the isotherm models.

2. MATERIAL AND METHODS

2.1. Materials

In this study, red clay was naturally obtained from the district of Hekimhan in Malatya province. In addition, red clay is abundant in many parts of Turkey and the world. Red clay was used as an adsorbent to separate malachite green from an aqueous solution by adsorption. Before the experimental adsorption process, the red clay was kept in an oven at 100 °C for about one day to reduce the humidity value. The dried red clay was sieved with less than 200 meshes particle size.

In this study, malachite green (MG) was used as the dyestuff separated from the aqueous solution, the adsorbate phase. MG is in the class of cationic dyes. MG, purchased in a dry form ready to be used in studies, has a molecular formula of $C_{23}H_{25}ClN_2$ and a molecular weight of 364.93 (g mol^{-1}). A stock solution was prepared at 1000 ppm to be used in the experiments. New solutions were prepared at the concentrations preferred in the experiments by mixing the stock solution with distilled water at the determined ratios. Filtered samples after adsorption were measured at 601 nm in a UV spectrophotometer. Five-milliliter cuvettes were used for measurements.

2.2. Adsorption Experiments

The sample containers in which the adsorption experiments were carried out, that is, the adsorbent and adsorbate phases came into contact, were made of plastic material with a volume of 50 ml, with a lid. Five different solutions, 25, 50, 100, 125 and 150 ppm, were prepared to examine the effect of initial dyestuff concentration on percent dye removal. Experiments were carried out with five different amounts of red clay 0.5, 1, 1.5 and 2 grams to examine the effect of the adsorbent amount on percent dye removal. Experiments were carried out at four different times, 30, 60, 90 and 120 minutes, to examine the effect of contact time on percent dye removal. Experiments were carried out at three different degrees: 298, 313 and 333 kelvin to investigate the effect of temperature on percent dye removal.
The percentage of dye removal [9] calculated according to the experimental results can be shown with the mathematical expression given below. Eq. (1):

$$\text{Percentage dye removal} = \frac{(C_0 - C_e)}{C_0} \times 100 \quad (1)$$

The initial dye concentration is represented by Co (mg L^{-1}), Ce indicates the concentration value at equilibrium. The amount of dyestuff adsorbed on the surface per unit weight of the adsorbent is expressed as q_e (mg g-1) and is calculated with (Equation 2) [10]:

$$q_e = (C_0 - C_e) \times \frac{V}{m} \quad (2)$$

What is represented by the symbol V (L) in the equation is the volume of the dye solution. The adsorbent mass is shown with m (g). All experiments carried out in adsorption processes were repeated at least twice. The arithmetic mean values obtained from these values were used in the following calculations.

3. RESULTS AND DISCUSSION

3.1. Adsorbent Characterization

Scanning electron microscopy is one of the important analysis methods used to characterize the structure of adsorbents. Here the porosity of the surface and secondary structures can be detected. SEM images of the adsorbent are given in (Figure 1). When the image is examined, it is seen that the red clay has a fibrous structure and various layers.

The surface properties of all adsorbents, whether natural, synthetic, or waste, are significant for the course of the process. Whether the structure is crystalline or amorphous is determined by XRD analysis. The XRD diagram of the red clay is given in (Figure 2). When the peaks appearing in the graph are compared with the literature, it is thought that the adsorbent shows an amorphous structure. There are 2 Theta D 20.72 views.

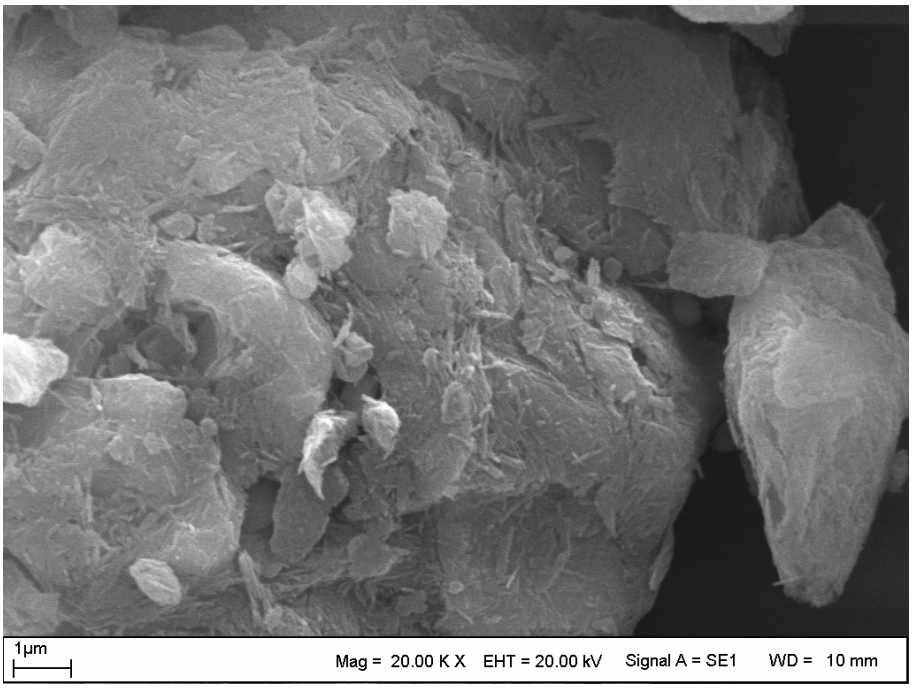


Figure 1. Scanning electron microscope image for red clay.

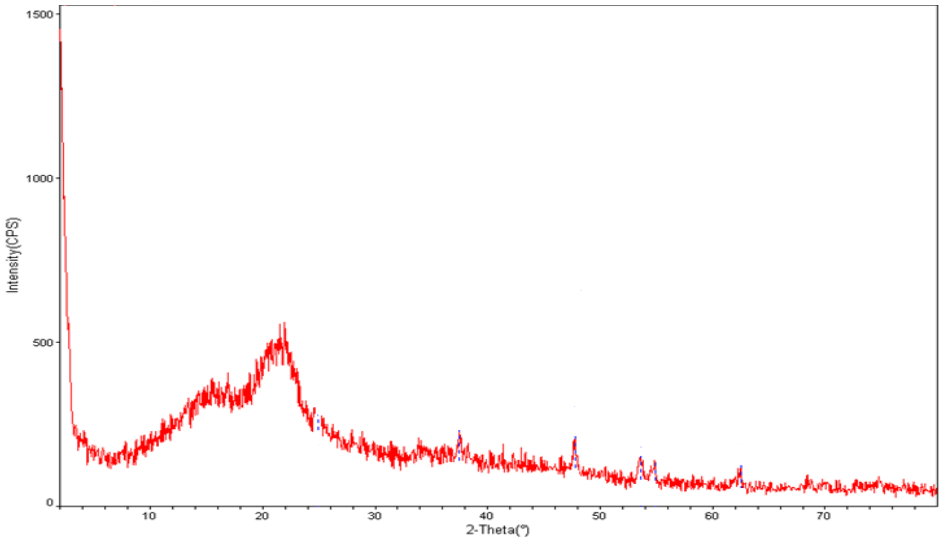


Figure 2. The peaks formed as a result of X-ray diffraction analysis of the adsorbent.

3.2. Effect of Initial Concentration

The effect of the initial concentration for malachite green adsorption is shown in (Figure 3). In this experimental set, where the contact time is one hour and the amount of red clay is 1 gram, it is seen that the dye removal rate increases as the initial concentration increases. However, the yield decreases at values higher than 100 ppm. Experiments were carried out at 298 kelvins and 400 rpm stirring speed.

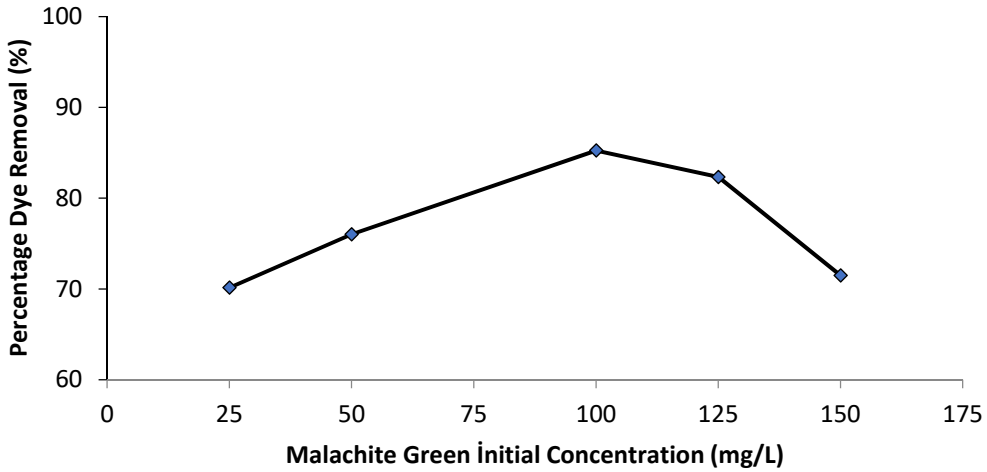


Figure 3. The effect of initial adsorbate concentration (t = 1 h, T= 298 K, 400 rpm, m = 1 g)

3.3. Effect of Adsorbent Dosage

To examine the positive or negative effects of the amount of adsorbent on dye removal, studies were carried out at an initial dye concentration of 100 ppm and a temperature of 298 K at a stirring speed of 400 rpm for 1 hour. Looking at the graph (Figure 4), it can be said that the dye removal efficiency increases as the amount of red clay increases. This is technically expected. However, the removal rates of 1, 1.5 and 2 grams are close. This is an indication that the saturation value has been reached.

3.4. Effect of Contact Time

The results of the adsorption time for malachite green adsorption are described in (Figure 5). When the adsorbing time increases, the removal of dyestuff increases when the adsorbent and adsorbate phases come into contact for a longer time. However, the increase occurs in the first 60 minutes and stabilizes. Times more extended than 60 minutes do not affect dye removal. To see the results of the adsorption time, experiments were carried out at an initial concentration of 100 ppm, using 1 gram of red clay at a stirring speed of 400 rpm and a temperature of 298 K.

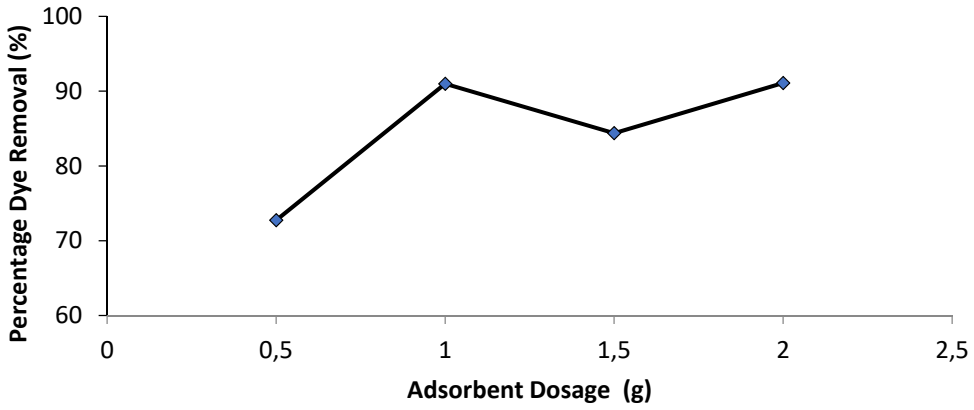


Figure 4. Effect of adsorbent dosage (t = 1 h, T = 298 K, 400 rpm, C_0 = 100 ppm)

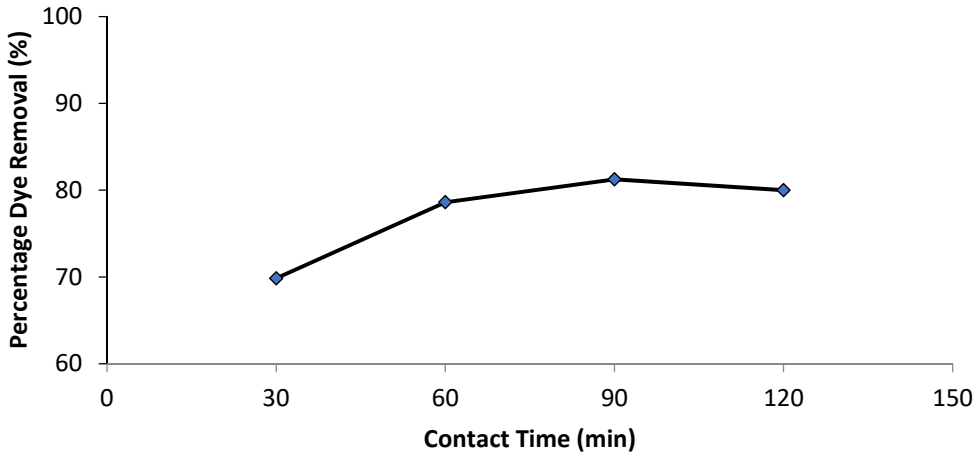


Figure 5. Effect of contact time (m = 1 g, T = 298 K, 400 rpm, C_0 = 100 ppm).

3.5. Effect of Temperature

To determine the relationship between process temperature change and adsorption success rate, it was carried out at three different degrees, 298, 313 and 333 Kelvin, at an initial concentration of 100 ppm, at a mixing speed of 400 rpm for 1 hour and the amount of 1 gram of red clay. When (Figure 6) is examined, as the temperature value is increased, the adsorption efficiency increases. As the temperature increases, it can be thought that the malachite green molecules move more. In this way, it can be believed that an adsorption event is endothermic.

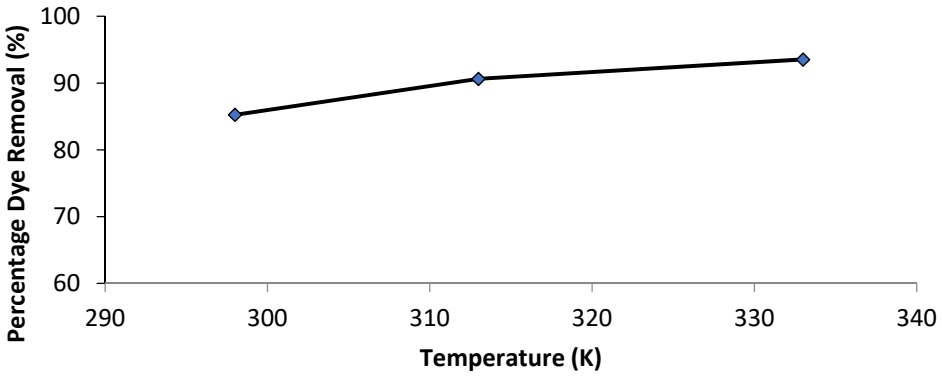


Figure 6. Effect of temperature (m = 1 g, t = 1 h, 400 rpm, C_0 = 100 ppm).

3.6. Adsorption Equilibrium

Isotherms are used to determine the adsorption equilibrium parameters and the physical and/or chemical properties of adsorption. They are used to predict the optimum amount of adsorbent. It is also used to make an assessment about the nature of the interactions that occur between the adsorbent and the adsorbate. The adsorption equilibrium graph is shown in (Figure 7). Four adsorption isotherms commonly used in the literature: Langmuir, Freundlich, Dubinin-Radushkevich, and Temkin are used in the present study.

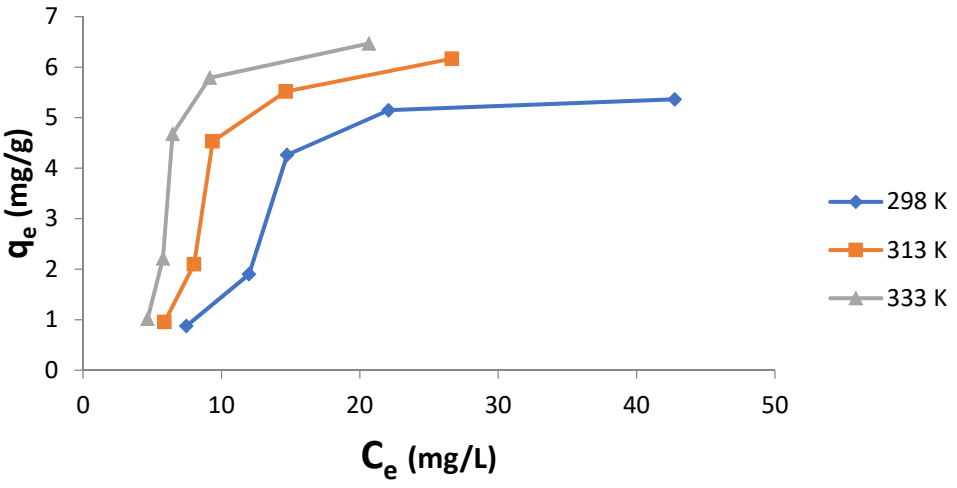


Figure 7. Adsorption equilibrium graph

3.6.1. The Langmuir İsotherm Model

With the Langmuir isotherm [11], an idea can be obtained about the number and structure of active sites on the adsorbent surface. When all active sites are filled, when maximum adsorption is achieved, it provides information about the saturation of these sites. It is possible to write the linear equation derived from the Langmuir model according to (Equation 3):

$$\frac{C_e}{q_e} = \frac{1}{Q_0\ b} + \frac{C_e}{Q_0} \tag{3}$$

Equilibrium concentration Ce (mg/L), q_e is the amount retained on the adsorbent surface in equilibrium (mg/g), Q_0 is the Langmuir constant (mg.g-1) that represents the system thermodynamically, and b shows the direct adsorption energy (L/mg).
In addition, the following results are the separation factor R_L obtained by the Langmuir isotherm can give an idea about the properties of adsorption [11]. R_L values are calculated from the following Eq. (4):

$$R_L = \frac{1}{1+bC_0} \tag{4}$$

As it is known, C_0 is the initial malachite green concentration (mg L^{-1}).

3.6.2. The Freundlich İsotherm Model

The assumptions of the Freundlich isotherm model [12] are as follows; After the active centers of the adsorbent are filled, an exponential decrease in adsorption energy occurs. The mathematically linear form of the Freundlich model is expressed by Equation (5):

$$\ln q_e = \ln K_f + \frac{1}{n} \ln C_e \tag{5}$$

Kf is a fixed number for the adsorption system associated with the bond energy. The value of 1/n gives information about the distribution of high-energy regions on the adsorbent surface compared to the other areas. It depends on whether the adsorption process is physical or chemical and, therefore, thermodynamically energized [12].

3.6.3. The Dubinin-Radushkevich (DR) İsotherm Model

The D-R model, based on the type, number, etc. of interactions between adsorbate particles in the aqueous solution, gives information about its features. It also allows calculating the kind of adsorption reaction and the reaction activation energy. The mathematical expression of my model is shown in the following equation Eq. (6).

$$q_e = q_m \exp\ (K\ \varepsilon^2\ ) \tag{6}$$

ε equals the empirical value RT ln(1 + 1/Ce). q_e can be considered the amount of dye adsorbed to the red clay surface per unit (mol/L). qm is the numerical amount of saturation capacity of the layer (mol/g), as this model is assumed to be a single layer. Ce is the value of the density of the solution at equilibrium. (mol/L). K' (mol^2/kJ^2) is an indicator of the thermodynamic energy of adsorption and is a constant value. R is the ideal gas constant (kJ/mol*K) and T is the temperature (K). The linear form of the Dubinin-Radushkevich isotherm model [13] is given in (Equation 7);

$$\ln q_e = \ln q_m - \ K\ \varepsilon^2 \tag{7}$$

The K value is represented by the average adsorption energy (E, kJ.mol-1) arising from the adsorption process and Eq. (8):

$$\mathrm{E} = \frac{1}{\sqrt{2K}} \tag{8}$$

3.6.4. The Temkin İsoterm Model

According to the Temkin model assumption, all molecules in the layers are considered to occupy the active centers without interacting due to the reduction of the reactive effect between the adsorbent and the adsorbent. Therefore, it is assumed that the heat of adsorption decreases in parallel with the coating of the surface. The linear mathematical expression of the Temkin isotherm [14] can be defined by the following Equation (9):

$$q_e = \frac{R_T}{b_T} \ln A_T + \frac{R_T}{b_T} \ln C_e \tag{9}$$

AT is a constant number expressed by the binding energy and represents the maximum value. The constant coefficients of the equations representing the adsorption isotherm models were calculated through error analysis using linearized versions of the isotherm mono equations. Obtained constant values and coupling coefficients R^2 are shown in (Table 1). When the table is examined, it can be said that it does not fit the Langmuir model since the R^2 value is very low. This is due to the layered structure of red clay. Since the R^2 value is higher in the D-R isotherm than the others, it can be considered to fit this isotherm. According to this model, it can be accepted that the clay has a porous structure and a heterogeneous structure. In the

Freundlich isotherm, on the other hand, it is seen that the K_f value increases as the temperature increases. This is because the adsorption capacity changes in heterogeneous multilayer adsorption with temperature change.

Table 1. Isotherm constants

Isotherm	298 K	313 K	333 K
Langmuir			
Q_0 (mg /g)	36.9	52.356	84.561
b (L/mg)	0.048	0.046	0.025
R^2	0.629	0.596	0.606
Freundlich			
K_f (L/g)	1.582	2.030	3.863
n	0.968	0.879	0.969
R^2	0.751	0.719	0.597
D-R			
q_m (mol. g^{-1}).10^3	1.59	1.78	1.93
K ($mol^2.kJ^{-2}$)	0.003	0.003	0.003
E ($kJ.mol^{-1}$)	12.91	12.91	12.91
R^2	0.931	0.942	0.883
Temkin			
A_T	1.589	1.817	1.452
b_T	90.37	75.19	81.42
R^2	0.805	0.825	0.722

4. CONCLUSIONS

Below are the results of malachite green adsorption on red clay:

- The red clay obtained from the Hekimhan region can be used as an adsorbent for the transport of malachite green, a cationic dye, from its aqueous solution by adsorption method.
- The effects of many parameters, such as initial dyestuff concentration, amount of adsorbent, adsorption time and temperature on the percentage of dye removal, were analyzed. The adsorbent amount and contact time determined optimum levels. It was determined that the yield increased as the temperature increased.
- The test results were evaluated regarding four different isotherm models. It was determined that the model with the lowest correlation coefficient values was the Langmuir isotherm model, while the highest values were observed in the Dubinin-Radushkevich isotherm model. It can be thought that the correlation coefficient is in the range of 0-1 and the results closer to 1 are more compatible with the applied model.
- The fact that the experimental data do not fit the Langmuir isotherm can indicate that the red clay has a multilayered structure. Compliance with the Dubinin-Radushkevich isotherm shows that the adsorbent has a heterogeneous and porous structure.

Acknowledgments

The datasets generated during and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declaration of Competing Interest

There is no conflict of interest.

Author Contribution

All authors contributed equally to every step of the article.

REFERENCES

[1] Hamzezadeh, A., Rashtbari, Y., Afshin, S., Morovati, M., Vosoughi, M., Application of low-cost material for adsorption of dye from aqueous solution. International Journal of Environmental Analytical Chemistry, 2022. 102(1): p. 254-269.

[2] Sultana, S., Islam, K., Hasan, M. A., Khan, H. J., Khan, M. A. R., Deb, A., Rahman, M. W., Adsorption of crystal violet dye by coconut husk powder: isotherm, kinetics and thermodynamics perspectives. Environmental Nanotechnology, Monitoring & Management, 2022. 17: p. 100651.

[3] Sarıcı Özdemir, Ç., Equilibrium, kinetic, diffusion and thermodynamic applications for dye adsorption with pine cone. Separation Science and Technology, 2019. 54(18): p. 3046-3054.

[4] Yagub, M. T., Sen, T. K., Afroze, S., Ang, H. M., Dye and its removal from aqueous solution by adsorption: a review. Advances in colloid and interface science, 2014. 209: p. 172-184.

[5] Salleh, M. A. M., Mahmoud, D. K., Karim, W. A. W. A., Idris, A., Cationic and anionic dye adsorption by agricultural solid wastes: a comprehensive review. Desalination, 2011. 280(1-3): p. 1-13.

[6] Ahmed, M. J., Application of agricultural based activated carbons by microwave and conventional activations for basic dye adsorption. Journal of Environmental Chemical Engineering, 2016. 4(1): p. 89-99.

[7] Jawad, A. H., Abdulhameed, A. S., Wilson, L. D., Syed-Hassan, S. S. A., ALOthman, Z. A., Khan, M. R., High surface area and mesoporous activated carbon from KOH-activated dragon fruit peels for methylene blue dye adsorption: Optimization and mechanism study. Chinese Journal of Chemical Engineering, 2021. 32: p. 281-290.

[8] Kausar, A., Iqbal, M., Javed, A., Aftab, K., Bhatti, H. N., Nouren, S., Dyes adsorption using clay and modified clay: a review. Journal of Molecular Liquids, 2018. 256: p. 395-407.

[9] Gadekar, M. R., Ahammed, M. M., Modelling dye removal by adsorption onto water treatment residuals using combined response surface methodology-artificial neural network approach. Journal of environmental management, 2019. 231: p. 241-248.

[10] Kearns, J. P., Wellborn, L. S., Summers, R. S., Knappe, D. R. U., 4-D adsorption to biochars: Effect of preparation conditions on equilibrium adsorption capacity and comparison with commercial activated carbon literature data. Water research, 2014. 62: p. 20-28.

[11] Langmuir, I., The adsorption of gases on plane surfaces of glass, mica and platinum. Journal of the American Chemical society, 1918. 40(9): p. 1361-1403.

[12] Freundlich, H., Über die adsorption in lösungen. Zeitschrift für physikalische Chemie, 1907. 57(1): p. 385-470.

[13] Dubinin, M. M.; Radushkevich, L. V., The Equation of Thecharacteristic Curve of Activated Charcoal. Proceedings of the Academy of Sciences. Phys. Chem. Sect., 1947. 55: p. 331–337.

[14] Temkin, M. J.; Pyzhev, V., Recent Modifications to Langmuir Isotherms. Acta Phys. Chem., 1940. 12: p. 271–279.

NATURENGS
MTU Journal of Engineering and Natural Sciences
https://dergipark.org.tr/tr/pub/naturengs
DOI: 10.46572/naturengs.1171676

Research Article

Identities of Generalized Pell and Pell-Lucas Sequences

Yashwant K Panwar*

Department of Mathematics, Government Model College, Jhabua, India.

(**Received:**06.09.2022; **Accepted:** 20.12.2022)

ABSTRACT: This paper presents sums of generalized Pell and Pell-Lucas sequences. Tulay Yagmur introduced these sequences in 2019. We have used their Generating function, Binet's formula and Induction method to derive the identities. We establish some connection formulae of involving them. Also, we present its two cross two matrix representation.

Keywords: Generalized Pell sequence, Generalized Pell-Lucas sequence, Binet's formula and Generating function.

1. INTRODUCTION

The Fibonacci sequence, Lucas sequence, Pell sequence, Pell-Lucas sequence, Jacobsthal sequence and Jacobsthal-Lucas sequence are most prominent examples of recursive sequences.

The sequence of Fibonacci numbers [6], F_n is defined by

$$F_n = F_{n-1} + F_{n-2},\ n \geq 2 \ with\ F_0 = 0,\ F_1 = 1 \tag{1}$$

The sequence of Lucas numbers [6], L_n is defined by

$$L_n = L_{n-1} + L_{n-2},\ n \geq 2 \ with\ L_0 = 2,\ L_1 = 1 \tag{2}$$

The sequence of Pell numbers [7], P_n is defined by

$$P_n = 2P_{n-1} + P_{n-2},\ n \geq 2 \ with\ P_0 = 0,\ P_1 = 1 \tag{3}$$

The sequence of Pell-Lucas numbers [7], Q_n is defined by

$$Q_n = 2Q_{n-1} + Q_{n-2},\ n \geq 2 \ with\ Q_0 = 2,\ Q_1 = 2 \tag{4}$$

Goksal Bilgici [1], defined new generalizations of Fibonacci and Lucas sequences for any real nonzero numbers a and b,

$$f_k = 2af_{k-1} + (b - a^2) f_{k-2},\ k \geq 2 \ \ with \ \ f_0 = 0,\ f_1 = 1 \tag{5}$$

$$l_k = 2al_{k-1} + (b - a^2) l_{k-2},\ k \geq 2 \ \ with \ \ l_0 = 2,\ l_1 = 2a \tag{6}$$

*Corresponding Author: yashwantpanwar@gmail.com
ORCID number of authors: 0000-0002-7429-4043

Tulay Yagmur [9], defined generalizations of Pell and Pell-Lucas sequences

$p_k = 2ap_{k-1} + (b-a^2)p_{k-2}$, $k \geq 2$ *with* $p_0 = 0,\ p_1 = 1$

$q_k = 2aq_{k-1} + (b-a^2)q_{k-2}$, $k \geq 2$ *with* $q_0 = 2,\ q_1 = 2a$

The main objective of this study is to give some Explicit sums of generalized Pell and Pell-Lucas sequences. Moreover, we introduce the special sums of the generalized Pell and Pell-Lucas sequences and prove them using Binet's formula.

2. PRELIMINARIES

In this section, we review basic definitions and introduce relevant facts.

For $n \geq 2$, The generalized Pell sequence [9], is defined by

$$p_k = 2ap_{k-1} + (b-a^2)p_{k-2}\ ,\ k \geq 2\ \ with\ \ p_0 = 0,\ p_1 = 1 \tag{7}$$

First few generalized Pell numbers are

$$\{p_n\} = \{0, 1, 2a, 3a^2 + b, 4a^3 + 4ab, 5a^4 + 10a^2b + b^2, ...\}$$

For $n \geq 2$, The generalized Pell-Lucas sequence [9], is defined by

$$q_k = 2aq_{k-1} + (b-a^2)q_{k-2}\ ,\ k \geq 2\ \ with\ \ q_0 = 2,\ q_1 = 2a \tag{8}$$

First few generalized Jacobsthal-Lucas numbers are

$$\{q_n\} = \{2, 2a, 2a^2 + 2b, 2a^3 + 6ab, 2a^4 + 12a^2b + 2b^2, 2a^5 + 20a^3b + 10ab^2, ...\}$$

Any nonzero real numbers are in (7) and (8).

If $a = 1$ & $b = 2$, then we obtained classical Pell sequence and Pell-Lucas sequences,

If $a = \frac{1}{2}$ & $b = \frac{9}{4}$, then we obtained classical Jacobsthal and Jacobsthal-Lucas sequences,

If $a = \frac{1}{2}$ & $b = \frac{5}{4}$, then we obtained classical Fibonacci and Lucas sequences,

If $a = \frac{3}{2}$ & $b = \frac{1}{4}$, then we obtained classical Mersenne and Fermat sequences.

For any positive integer *k*,

If $a = 1$ & $b = (1+k)$, then we obtained *k*-Pell and *k*-Pell-Lucas sequences,

If $a = \frac{k}{2}$ & $b = \left(\frac{4+k^2}{4}\right)$, then we obtained *k*-Fibonacci and *k*-Lucas sequences,

If $a = \frac{k}{2}$ & $b = \left(\frac{8+k^2}{4}\right)$, then we obtained *k*- Jacobsthal and *k*- Jacobsthal-Lucas sequences.

Generating function for generalized Pell and Pell-Lucas numbers are

$$\sum_{k=0}^{\infty} p_k x^k = \frac{x}{1-2ax-(b-a^2)x^2} \tag{9}$$

$$\sum_{k=0}^{\infty} q_k x^k = \frac{2-2ax}{1-2ax-(b-a^2)x^2} \tag{10}$$

The Binet's formula for generalized Pell and Pell-Lucas numbers are

$$p_k = \frac{\Re_1^k - \Re_2^k}{\Re_1 - \Re_2} \tag{11}$$

$$q_k = \Re_1^k + \Re_2^k \tag{12}$$

where $\Re_1$ and $\Re_2$ are the roots of the characteristic equation,

$$x^2 - 2ax - (b-a^2) = 0 \tag{13}$$

with $\Re_1 = a+\sqrt{b}$, $\Re_2 = a-\sqrt{b}$; $\Re_1 + \Re_2 = 2a$, $\Re_1 - \Re_2 = 2\sqrt{b}$, $\Re_1\Re_2 = a^2 - b$.

Also $p_{-k} = \frac{-1}{(a^2-b)^k} p_k$ *and* $q_{-k} = \frac{1}{(a^2-b)^k} q_k$.

3. IDENTITIES OF THE GENERALIZED PELL AND PELL-LUCAS SEQUENCE

This section introduces and proves some interesting identities of generalized Pell and Pell-Lucas sequences.

3.1. Explicit Sums of generalized Pell and Pell-Lucas Sequence

This section studies the sums of generalized Pell and Pell-Lucas sequences. This enables us to give in a straightforward way several formulas for the sums of such numbers.

Theorem 1: Explicit sum formula for generalized Pell sequence

$$p_k = \sum_{i=0}^{\left\lceil \frac{k-1}{2} \right\rceil} \binom{k-i-1}{i} (2a)^{k-2i-1} (b-a^2)^i \tag{14}$$

Proof: The proof is clear from the generating function of generalized Pell sequence.

Theorem 2: Explicit sum Formula generalized Pell-Lucas sequence

$$q_k = 2\sum_{i=0}^{\left\lceil \frac{k}{2} \right\rceil} \binom{k-i}{i} (2a)^{k-2i} (b-a^2)^i - \sum_{i=0}^{\left\lceil \frac{k-1}{2} \right\rceil} \binom{k-i-1}{i} (2a)^{k-2i} (b-a^2)^i \tag{15}$$

Proof: The proof is clear from the generating function of generalized Pell-Lucas sequence.

Lemma 3: For every s *and* t, the following equality holds

$$p_{s(n+2)+t} = q_s\ p_{s(n+1)+t} - (a^2-b)^s\ p_{sn+t} \tag{16}$$

Proof: From the Binet's formula of generalized Pell and Pell-Lucas sequence,

$$\begin{aligned} q_s\ p_{s(n+1)+t} &= \left(\Re_1^s + \Re_2^s\right)\left(\frac{\Re_1^{s(n+1)+t} - \Re_2^{s(n+1)+t}}{\Re_1 - \Re_2}\right) \\ &= \frac{1}{\Re_1 - \Re_2}\left[\Re_1^{s(n+2)+t} + (a^2-b)^s \Re_1^{sn+t} - (a^2-b)^s \Re_2^{sn+t} - \Re_2^{s(n+2)+t}\right] \\ &= \frac{1}{\Re_1 - \Re_2}\left[\left\{\Re_1^{s(n+2)+t} - \Re_2^{s(n+2)+t}\right\} + (a^2-b)^s\left(\Re_1^{sn+t} - \Re_2^{sn+t}\right)\right] \\ &= p_{s(n+2)+t} + (a^2-b)^s\ p_{sn+t} \end{aligned}$$

then, equality becomes,

$$p_{s(n+2)+t} = q_s\ p_{s(n+1)+t} - (a^2-b)^s\ p_{sn+t}$$

Theorem 4: For fixed integers s, t *with* $0 \le t \le s-1$, the following equality holds

$$\sum_{i=0}^{n} p_{si+t} = \frac{p_{s(n+1)+t} - (a^2-b)^t\ p_{s-t} - p_t - (a^2-b)^s\ p_{sn+t}}{q_s - (a^2-b)^s - 1} \tag{17}$$

Proof: From the Binet's formula of generalized Pell sequence,

$$\begin{aligned} \sum_{i=0}^{n} p_{si+t} &= \sum_{i=0}^{n} \frac{\Re_1^{si+t} - \Re_2^{si+t}}{\Re_1 - \Re_2} \\ &= \frac{1}{\Re_1 - \Re_2}\left[\sum_{i=0}^{n} \Re_1^{si+t} - \sum_{i=0}^{n} \Re_2^{si+t}\right] \\ &= \frac{1}{\Re_1 - \Re_2}\left[\frac{\Re_1^{sn+t+s} - \Re_1^t}{\Re_1^s - 1} - \frac{\Re_2^{sn+t+s} - \Re_2^t}{\Re_2^s - 1}\right] \\ &= \frac{1}{(a^2-b)^s - q_s + 1}\left[(a^2-b)^s\ p_{sn+t} - p_{s(n+1)+t} + p_t + (a^2-b)^t\ p_{s-t}\right] \\ &= \frac{p_{s(n+1)+t} - (a^2-b)^t\ p_{s-t} - p_t - (a^2-b)^s\ p_{sn+t}}{q_s - (a^2-b)^s - 1} \end{aligned}$$

This completes the proof.

Corollary 5: Sum of odd new generalized Pell sequence,

If $s = 2m+1$ then Eq. (17) is

$$\sum_{i=0}^{n} p_{(2m+1)i+t} = \frac{p_{(2m+1)(n+1)+t} - (a^2-b)^t p_{2m+1-t} - p_t - (a^2-b)^{(2m+1)} p_{(2m+1)n+t}}{q_{(2m+1)} - (a^2-b)^{(2m+1)} - 1} \tag{18}$$

For example

(1) If $m = 0$ then $s = 1$

$$\sum_{i=0}^{n} p_{i+t} = \frac{p_{n+t+1} - (a^2-b)^t p_{1-t} - p_t - (a^2-b) p_{n+t}}{2a - (a^2-b) - 1} \tag{19}$$

(i) For $t = 0$: $\sum_{i=0}^{n} p_i = \frac{p_{n+1} - 1 - (a^2-b) p_n}{2a - (a^2-b) - 1}$

(2) If $m = 1$ then $s = 3$

$$\sum_{i=0}^{n} p_{3i+t} = \frac{p_{3n+t+3} - (a^2-b)^t p_{3-t} - p_t - (a^2-b)^3 p_{3n+t}}{a^3(2-a^3) + 3ab(2-ab) + b(3a^4+b^2) - 1} \tag{20}$$

(i) For $t = 0$: $\sum_{i=0}^{n} p_{3i} = \frac{p_{3n+3} - (3a^2+b) - (a^2-b)^3 p_{3n}}{a^3(2-a^3) + 3ab(2-ab) + b(3a^4+b^2) - 1}$

(ii) For $t = 1$: $\sum_{i=0}^{n} p_{3i+1} = \frac{p_{3n+4} - 2a(a^2+b) - 1 - (a^2-b)^3 p_{3n+1}}{a^3(2-a^3) + 3ab(2-ab) + b(3a^4+b^2) - 1}$

(iii) For $t = 2$: $\sum_{i=0}^{n} p_{3i+2} = \frac{p_{3n+5} - (a^2-b)^2 - 2a - (a^2-b)^3 p_{3n+2}}{a^3(2-a^3) + 3ab(2-ab) + b(3a^4+b^2) - 1}$

(3) If $m = 2$ then $s = 5$

$$\sum_{i=0}^{n} p_{5i+t} = \frac{p_{5n+t+5} - (a^2-b)^t p_{5-t} - p_t - (a^2-b)^5 p_{5n+t}}{q_5 - (a^2-b)^5 - 1} \tag{21}$$

(i) For $t = 0$: $\sum_{i=0}^{n} p_{5i} = \frac{p_{5n+5} - p_5 - (a^2-b)^5 p_{5n}}{q_5 - (a^2-b)^5 - 1}$

(ii) For $t = 1$: $\sum_{i=0}^{n} p_{5i+1} = \frac{p_{5n+6} - (a^2-b) p_4 - 1 - (a^2-b)^5 p_{5n+1}}{q_5 - (a^2-b)^5 - 1}$

(iii) For $t = 2$: $\sum_{i=0}^{n} p_{5i+2} = \frac{p_{5n+7} - (a^2-b)^2 p_3 - 2a - (a^2-b)^5 p_{5n+2}}{q_5 - (a^2-b)^5 - 1}$

(iv) For $t = 3$: $\sum_{i=0}^{n} p_{5i+3} = \frac{p_{5n+8} - (a^2-b)^3 2a - (3a^2+b) - (a^2-b)^5 p_{5n+3}}{q_5 - (a^2-b)^5 - 1}$

(v) For $t = 4$: $\sum_{i=0}^{n} p_{5i+4} = \frac{p_{5n+9} - (a^2-b)^4 - 4a(a^2+b) - (a^2-b)^5 p_{5n+4}}{q_5 - (a^2-b)^5 - 1}$

(vi) For $t = 5$: $\sum_{i=0}^{n} p_{5i+5} = \frac{p_{5n+10} - p_5 - (a^2-b)^5 p_{5n+5}}{q_5 - (a^2-b)^5 - 1}$

Corollary 6: Sum of even generalized Pell sequence, if $s = 2m$ then Eq. (17) is

$$\sum_{i=0}^{n} p_{2mi+t} = \frac{p_{2m(n+1)+t} - (a^2-b)^t p_{2m-t} - p_t - (a^2-b)^{2m} p_{2mn+t}}{q_{2m} - (a^2-b)^{2m} - 1} \tag{22}$$

For example

(1) If $m = 1$ then $s = 2$

$$\sum_{i=0}^{n} p_{2i+t} = \frac{p_{2n+2+t} - (a^2-b)^t p_{2-t} - p_t - (a^2-b)^2 p_{2n+t}}{q_2 - (a^2-b)^2 - 1} \tag{23}$$

(i) For $t = 0$: $\sum_{i=0}^{n} p_{2i} = \frac{p_{2n+2} - 2a - (a^2-b)^2 p_{2n}}{(2a^2+2b) - (a^2-b)^2 - 1}$

(ii) For $t = 1$: $\sum_{i=0}^{n} p_{2i+1} = \frac{p_{2n+3} - (a^2-b) - 1 - (a^2-b)^2 p_{2n+1}}{(2a^2+2b) - (a^2-b)^2 - 1}$

(iii) For $t = 2$: $\sum_{i=0}^{n} p_{2i+2} = \frac{p_{2n+4} - 2a - (a^2-b)^2 p_{2n+2}}{(2a^2+2b) - (a^2-b)^2 - 1}$

(2) If $m = 2$ then $s = 4$

$$\sum_{i=0}^{n} p_{4i+t} = \frac{p_{4n+4+t} - (a^2-b)^t p_{4-t} - p_t - (a^2-b)^4 p_{4n+t}}{q_4 - (a^2-b)^4 - 1} \tag{24}$$

(i) For $t = 0$: $\sum_{i=0}^{n} p_{4i} = \frac{p_{4n+4} - p_4 - (a^2-b)^4 p_{4n}}{q_4 - (a^2-b)^4 - 1}$

(ii) For $t = 1$: $\sum_{i=0}^{n} p_{4i+1} = \frac{p_{4n+5} - (a^2-b)p_3 - 1 - (a^2-b)^4 p_{4n+1}}{q_4 - (a^2-b)^4 - 1}$

(iii) For $t = 2$: $\sum_{i=0}^{n} p_{4i+2} = \frac{p_{4n+6} - (a^2-b)^2 2a - 2a - (a^2-b)^4 p_{4n+2}}{q_4 - (a^2-b)^4 - 1}$

(iv) For $t = 3$: $\sum_{i=0}^{n} p_{4i+3} = \frac{p_{4n+7} - (a^2-b)^3 - p_3 - (a^2-b)^4 p_{4n+3}}{q_4 - (a^2-b)^4 - 1}$

(v) For $t = 4$: $\sum_{i=0}^{n} p_{4i+4} = \frac{p_{4n+8} - p_4 - (a^2-b)^4 p_{4n+4}}{q_4 - (a^2-b)^4 - 1}$

(3) If $m = 3$ then $s = 6$

$$\sum_{i=0}^{n} p_{6i+t} = \frac{p_{6n+6+t} - (a^2-b)^t p_{6-t} - p_t - (a^2-b)^6 p_{6n+t}}{q_6 - (a^2-b)^6 - 1} \tag{25}$$

(i) For $t = 0$: $\sum_{i=0}^{n} p_{6i} = \frac{p_{6n+6} - p_6 - (a^2-b)^6 p_{6n}}{q_6 - (a^2-b)^6 - 1}$

(ii) For $t = 1$: $\sum_{i=0}^{n} p_{6i+1} = \frac{p_{6n+7} - (a^2-b)p_5 - 1 - (a^2-b)^6 p_{6n+1}}{q_6 - (a^2-b)^6 - 1}$

(iii) For $t=2$: $\sum_{i=0}^{n} p_{6i+2} = \dfrac{p_{6n+8} - (a^2-b)^2 p_4 - 2a - (a^2-b)^6 p_{6n+2}}{q_6 - (a^2-b)^6 - 1}$

(iv) For $t=3$: $\sum_{i=0}^{n} p_{6i+3} = \dfrac{p_{6n+9} - (a^2-b)^2 p_3 - p_3 - (a^2-b)^6 p_{6n+3}}{q_6 - (a^2-b)^6 - 1}$

Theorem 7: For fixed integers s, t *with* $0 \le t \le s-1$, the following equality holds

$$\sum_{i=0}^{n} (-1)^i p_{si+t} = \frac{(-1)^n p_{s(n+1)+t} + (-1)^n (a^2-b)^s p_{sn+t} - (a^2-b)^t p_{s-t} + p_t}{q_s + (a^2-b)^s + 1} \tag{26}$$

For different values of s and t:

(i) $$\sum_{i=0}^{n} (-1)^i p_i = \frac{(-1)^n p_{n+1} + (-1)^n (a^2-b) p_n - 1}{2a + a^2 - b + 1}$$

(ii) $$\sum_{i=0}^{n} (-1)^i p_{2i} = \frac{(-1)^n p_{2n+2} + (-1)^n (a^2-b)^2 p_{2n} - 2a}{(2a^2+2b) + (a^2-b)^2 + 1}$$

(iii) $$\sum_{i=0}^{n} (-1)^i p_{2i+1} = \frac{(-1)^n p_{2n+3} + (-1)^n (a^2-b)^2 p_{2n+1} - (a^2-b) + 1}{(2a^2+2b) + (a^2-b)^2 + 1}$$

(iv) $$\sum_{i=0}^{n} (-1)^i p_{4i} = \frac{(-1)^n p_{4n+4} + (-1)^n (a^2-b)^4 p_{4n} - p_4}{q_4 + (a^2-b)^4 + 1}$$

(v) $$\sum_{i=0}^{n} (-1)^i p_{4i+1} = \frac{(-1)^n p_{4n+5} + (-1)^n (a^2-b)^4 p_{4n+1} - (a^2-b) p_3 + 1}{q_4 + (a^2-b)^4 + 1}$$

(vi) $$\sum_{i=0}^{n} (-1)^i p_{4i+2} = \frac{(-1)^n p_{4n+6} + (-1)^n (a^2-b)^4 p_{4n+2} - (a^2-b)^2 2a + 2a}{q_4 + (a^2-b)^4 + 1}$$

(vii) $$\sum_{i=0}^{n} (-1)^i p_{4i+3} = \frac{(-1)^n p_{4n+7} + (-1)^n (a^2-b)^4 p_{4n+3} - (a^2-b)^3 + p_3}{q_4 + (a^2-b)^4 + 1}$$

3.2. Product of Generalized Pell and Pell-Lucas Sequences

In this section, we present identities involving the product of generalized Pell and Pell-Lucas numbers and related identities consisting of even and odd terms.

Theorem 8: If p_k *and* q_k are generalized Pell and Pell-Lucas numbers, then holds for every k *and* s,

I. $$p_{2k+s} q_{2k+1} = p_{4k+s+1} + (a^2-b)^{2k+1} p_{s-1} \tag{27}$$

II. $$p_{2k+s} q_{2k+2} = p_{4k+s+2} + (a^2-b)^{2k+2} p_{s-2} \tag{28}$$

III. $$p_{2k+s} q_{2k} = p_{4k+s} + (a^2-b)^{2k} p_s \tag{29}$$

Theorem 9:

I. $p_{2k-s}q_{2k+1} = p_{4k-s+1} + (a^2-b)^{2k+1} p_{-s-1}$ (30)

II. $p_{2k-s}q_{2k-1} = p_{4k-s-1} + (a^2-b)^{2k-1} p_{1-s}$ (31)

III. $p_{2k-s}q_{2k} = p_{4k-s} + (a^2-b)^{2k} p_{-s}$ (32)

Theorem 10:

I. $p_{2k}q_{2k+s} = p_{4k+s} - (a^2-b)^{2k} p_s$ (33)

II. $4bp_{2k}p_{2k+s} = q_{4k+s} - (a^2-b)^{2k} q_s$ (34)

III. $q_{2k}q_{2k+s} = q_{4k+s} + (a^2-b)^{2k} q_s$ (35)

The proof is clear by the Binet's formula of generalized Pell and Pell-Lucas numbers.

3.2. Sum and difference of squares Generalized Pell and Pell-Lucas Sequences

In this section, the sum and difference of generalized Pell and Pell-Lucas numbers are treated in the following theorem.

Theorem 11: $4b\left(p_{n+1}^2 + p_{n-1}^2\right) = q_{2n+2} + q_{2n-2} - 2(a^2-b)^{n-1}\{(a^2-b)^2+1\}$ (36)

Theorem 12: $4b\left(p_{n+1}^2 - p_{n-1}^2\right) = q_{2n+2} - q_{2n-2} - 2(a^2-b)^{n-1}\{(a^2-b)^2-1\}$ (37)

The proof is clear by the Binet's formula of generalized Pell and Pell-Lucas numbers.

3.3. Matrix Representation of Generalized Pell and Pell-Lucas Sequences

In this section, we present two crosses and two matrices for generalized Pell and Pell-Lucas sequences are given by $\mathrm{A} = \begin{bmatrix} 2a & 1 \\ (b-a^2) & 0 \end{bmatrix}$.

Theorem 13: For $n \in \square$, we have $\begin{bmatrix} p_{n+1} \\ (b-a^2)p_n \end{bmatrix} = \mathrm{A}\begin{bmatrix} p_n \\ (b-a^2)p_{n-1} \end{bmatrix}$ (38)

Proof: To prove the result we will use induction on n. (6.1) is valid for n = 1. Suppose (38) is good for n, we get

$$\begin{bmatrix} p_{n+2} \\ (b-a^2)p_{n+1} \end{bmatrix} = \begin{bmatrix} 2ap_{n+1} + (b-a^2)p_n \\ (b-a^2)p_{n+1} \end{bmatrix}$$

$$= \begin{bmatrix} 2a & 1 \\ (b-a^2) & 0 \end{bmatrix}\begin{bmatrix} p_{n+1} \\ (b-a^2)p_n \end{bmatrix}$$

$$= \begin{bmatrix} 2a & 1 \\ (b-a^2) & 0 \end{bmatrix}\begin{bmatrix} 2a & 1 \\ (b-a^2) & 0 \end{bmatrix}\begin{bmatrix} p_n \\ (b-a^2)p_{n-1} \end{bmatrix}$$

$$
\begin{aligned}
&= \begin{bmatrix} 2a & 1 \\ (b-a^2) & 0 \end{bmatrix} \begin{bmatrix} 2ap_n + (b-a^2)p_{n-1} \\ (b-a^2)p_n \end{bmatrix} \\
&= \begin{bmatrix} 2a & 1 \\ (b-a^2) & 0 \end{bmatrix} \begin{bmatrix} p_{n+1} \\ (b-a^2)p_n \end{bmatrix} \\
&= \mathrm{A} \begin{bmatrix} p_{n+1} \\ (b-a^2)p_n \end{bmatrix}
\end{aligned}
$$

Theorem 14: For $n \in \square$, we have $\begin{bmatrix} q_{n+1} \\ (b-a^2)q_n \end{bmatrix} = \mathrm{A} \begin{bmatrix} q_n \\ (b-a^2)q_{n-1} \end{bmatrix}$ (39)

Theorem 15: For , $n \in \square$ we have $\begin{bmatrix} p_{n+1} \\ (b-a^2)p_n \end{bmatrix} = \mathrm{A}^n \begin{bmatrix} p_1 \\ (b-a^2)p_0 \end{bmatrix}$ (40)

Theorem 16: For $n \in \square$, we have $\begin{bmatrix} q_{n+1} \\ (b-a^2)q_n \end{bmatrix} = \mathrm{A}^n \begin{bmatrix} q_1 \\ (b-a^2)q_0 \end{bmatrix}$ (41)

4. CONCLUSIONS

This study describes explicit sums of generalized Pell and Pell-Lucas sequences. This enables us to give in a straightforward way several formulas for the sums of such generalized numbers. We describe some generalized identities involving the product of generalized Pell and Pell-Lucas sequences. Also, we present identities related to their sum and difference of squares involving them and its two cross two matrices and find exciting properties such as the nth power of the matrix.

Acknowledgments

The author thanks the editor and anonymous referees for helpful suggestions and comments.

Declaration of Competing Interest

The author declares that they have no known competing financial interests or personal relationships that could influence the work reported in this paper.

Author Contribution

Pell and Pell-Lucas sequences were introduced by Tulay Yagmur in 2019. Yashwant K Panwar contributed 100% at every stage of the article.

REFERENCES

[1] Bilgici, G., New Generalizations of Fibonacci and Lucas Sequences. Applied Mathematical Sciences, 2014. 8(29): p. 1429-1437.

[2] Dasdemir, A., On the Pell, Pell-Lucas and modified Pell numbers by matrix method. Appl. Math. Sci., 2011. 5(64): p. 3173–3181.

[3] Falcon,, S., Plaza, A., On the *k*-Fibonacci Numbers. Chaos,Solitons and Fractals, 2007. 32(5): p. 1615-1624.

[4] Gulec, H.H., Taskara N., On the (s,t)-Pell and (s,t)-Pell-Lucas sequences and their matrix representations. Appl. Math. Lett., 2012. 25: p. 1554–1559.

[5] Horadam, A.F., Pell Identities. *The Fibonacci Quarterly*, 1971 9(3): p. 245-263.

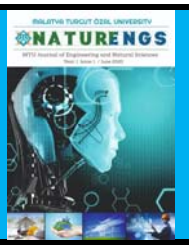
[6] Koshy T., Fibonacci and Lucas numbers with applications. 2001. New York, Wiley- Interscience.

[7] Koshy T., Pell and Pell-Lucas numbers with applications. 2014. Springer, Berlin.

[8] Taşyurdu Y, Cobanoğlu N, Dilmen Z., On the New Family of k-Fibonacci Numbers. Erzincan University Journal of Science and Technology, 2016. 9(1): p. 95-101.

[9] Yagmur T., New Approach to Pell and Pell-Lucas Sequence. Kyungpook Mathematical Journal, 2019. 59(1): p. 23-34.

NATURENGS
MTU Journal of Engineering and Natural Sciences
https://dergipark.org.tr/tr/pub/naturengs
DOI: 10.46572/naturengs.1205919



Research Article

Detection of Tonsillopharyngitis with Grad-Cam and Optimization-Based Model Using Oropharyngeal Images

Muhammed YILDIRIM[1*], Orkun EROGLU [2]

[1] Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Ozal University, Malatya, Turkey.
[2] Department of Otorhinolaryngology, School of Medicine, Firat University, Elazig, Turkey.

(**Received:** 17.11.2022; **Accepted:** 09.12.2022)

ABSTRACT: Tonsillopharyngitis is a sudden onset contagious infection of the pharynx and tonsils. Patients experience a rapid general condition and loss of workforce. In addition to affecting patients, it spreads and affects other individuals. In addition, it causes severe complications and increases hospital costs. Therefore, early and accurate diagnosis is essential. In this study, a hybrid model is developed to diagnose tonsillopharyngitis. First, the heat maps of the images in the original data set by applying the Gradient-weighted Class Activation Mapping (Grad-Cam) method. In the proposed model, feature maps are obtained from the original and heatmap datasets using the Darknet53 architecture as the base. It aims to increase the proposed model's performance by bringing together different features of the same image. After the feature map obtained after the feature fusion step is optimized with the Relief method, classification is carried out using an SVM shallow classifier.

Keywords: Artificial Intelligence, CNN, Grad-Cam, Relief, Tonsillopharyngitis.

1. INTRODUCTION

Tonsillopharyngitis is a contagious infection and inflammation of the pharynx and tonsils. It is one of the most common infections in children and adults [1]. Bacteria and viruses can cause infection. It is one of the common reasons for hospital admissions, especially in childhood. Transmission is more common in communal living areas. Tonsillopharyngitis caused by microorganisms called group A beta-hemolytic streptococcus is most common in school-age children. The disease often increases in the winter season. The most common complaints of the patients are sudden onset of sore throat, fever, and inability to swallow and feed. Physical examination, throat culture, serological tests, and rapid antigen tests are used for diagnosis. Physical examination may reveal swelling of the tonsils, painful lymphadenopathy in the anterior part of the neck, and rash on the soft palate. Early diagnosis of the disease prevents the development of complications in patients as early treatment will be initiated. In addition, early diagnosis prevents the transmission of the disease to other people [2-4].

Tonsillopharyngitis is a common life-threatening infection when complications develop. It is a significant health problem because it causes loss in the workforce and can easily be transmitted in family and communal living spaces. Detection of this problem at an early stage is of great

*Corresponding Author: muhammed.yildirim@ozal.edu.tr
ORCID number of authors: [1] 0000-0003-1866-4721, [2] 0000-0001-9392-5755

importance for treating the patient [5]. In recent years, the use of artificial intelligence methods in the medical field has been the focus of the attention of researchers [6-8]. This study aims to recognize tonsillopharyngitis images by developing an artificial intelligence-based hybrid model.

The novelties and contributions of the study are as follows.

- In this study, heat maps of oropharynx images in the data set were obtained with Grad-Cam technology.
- Feature maps were extracted from the original and heatmap datasets using the Darknet53 architecture. In this step, different features of the same image are obtained.
- It aims to increase the model's performance by combining the features obtained from both the original data set and the data set created by obtaining heat maps.
- The Relief method is used to reduce the size of the feature map obtained after the feature merging step. In this step, it is aimed that the model will run faster and produce more successful results since the unnecessary features are eliminated and the size reduction process is performed.
- High performance is achieved by classifying the optimized feature map in the SVM shallow classifier.
- While the highest accuracy value obtained in eight different pre-trained models is 84.71%, the accuracy value obtained in the proposed model is 89.7%.

Not many studies were found when the literature on the subject was searched. Yoo et al. study created a data set consisting of 131 throats and 208 normal throat images. Since the images in the created data set were few, the researchers made data multiplexing. GAN structures were used for data multiplexing. In the study, researchers used ResNet50, InceptionV3 and MobileNetV2 architectures. The accuracy value obtained by using 4-fold cross-validation in the ResNet50 model was 95.3%. The researchers stated that the proposed system could be used in smartphone-based applications [9].

The article's organization is as follows: The material and method part is examined in the second part. In the third part, the results of the application are presented in detail. The fourth part of the study, discussion, and conclusions in the fifth part are included.

2. MATERIAL AND METHODS

This section examines the data set used in the study, pre-trained models, classifiers, Grad-Cam technology, and the proposed model. The proposed model aims to obtain high success rates in classifying tonsillopharyngitis images.

2.1. Data set

Tonsillopharyngitis is a common, acute-onset infection of the pharynx and tonsils. Patients have sore throats and difficulty swallowing. Physical examination is essential in diagnosis. Physical study reveals redness in the oropharynx, swelling of the tonsils, and exudate. The infection may

spread to the surrounding soft tissues in the neck and cause serious complications such as peritonsillar abscesses and mild neck tissue infections. The data set used in the study consists of oropharyngeal images with tonsillopharyngitis and normal oropharyngeal images. While creating this data set we used, the researchers applied the knowledge of two experts. In addition, ambiguous images were eliminated by experts. While collecting the data set used in the study, images up to April 2020 were collected [9, 10]. The related data set contains 131 images of the oropharynx with tonsillopharyngitis and 208 images of the normal oropharynx. Image examples from the original dataset are presented in Figure 1.

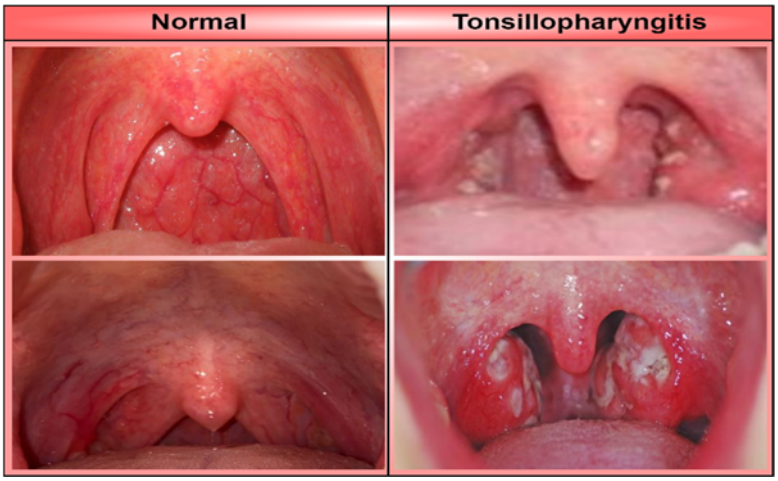


Figure 1. Examples from oropharynx images

2.2. Grad-Cam technology

Grad-Cam technology is a technique used to obtain heat maps of images. Grad-Cam uses gradients of any target concept. Grad-Cam can be applied to various CNN architectures [11]. In this study, heat maps of existing images were obtained using Grad-Cam technology. Then, the feature maps of these images were extracted and combined with the feature maps of the original images. This way, the model's performance is increased by using different features of the same image. The image samples obtained by applying Grad-Cam to the original images are given in Figure 2.

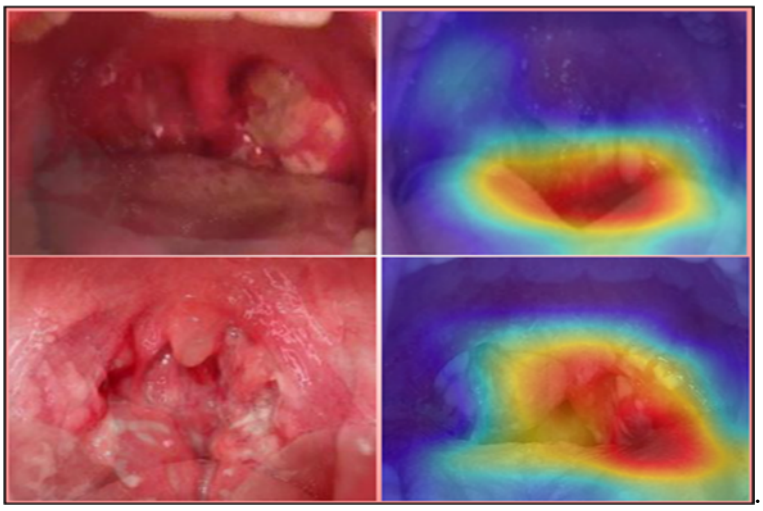


Figure 2. Examples from Grad-Cam images

2.3. Pre-trained architectures, Relief, and classifiers

Artificial intelligence-based methods have been used frequently in recent years, especially in the biomedical field. Especially after AlexNet architecture won the ImageNet competition in 2012, deep learning, which had a stagnant period, started to become popular again. Eight different pre-trained architectures were used in this study to diagnose tonsillopharyngitis using oropharyngeal images. The results obtained in these architectures used in the literature were compared with the proposed hybrid model. In this study, AlexNet [12], DarkNet53 [13], DenseNet201 [14], EfficientNetb0 [15], GoogleNet [16], MobileNetV2 [17], ResNet101 [18], and ShuffleNet [19] architectures were used. In these CNN architectures used in the study, feature extraction is done automatically. In these architectures, there is no need for expert knowledge in the feature extraction stage.

After combining the feature maps obtained with different techniques in the proposed hybrid model, the Relief method optimized the obtained feature map [20]. After eliminating unnecessary features, the optimized feature map obtained was classified into different machine learning classifiers. The classifiers used in the study are Support Vector Machine (SVM)[21], k-nearest neighbor (KNN)[22], Ensemble Subspace (ES)[23], Decision Tree(DT)[24], Discriminant Analysis (DA)[25], Naïve Bayes (NB)[26] and Logistic Regression(LR)[27].

2.4. Developed model

In the model developed for the classification of oropharyngeal images, first of all, heat maps of the pictures of the original data set were obtained with Grad-Cam technology. Then, feature maps were extracted from both the original dataset and the dataset created using Grad-Cam technology using DarkNet53. After these extracted feature maps were combined, the feature map was optimized using Relief. After the Relief eliminated the unnecessary features, these features were classified into different classifiers. In the proposed model, the highest success rates were obtained in SVM.

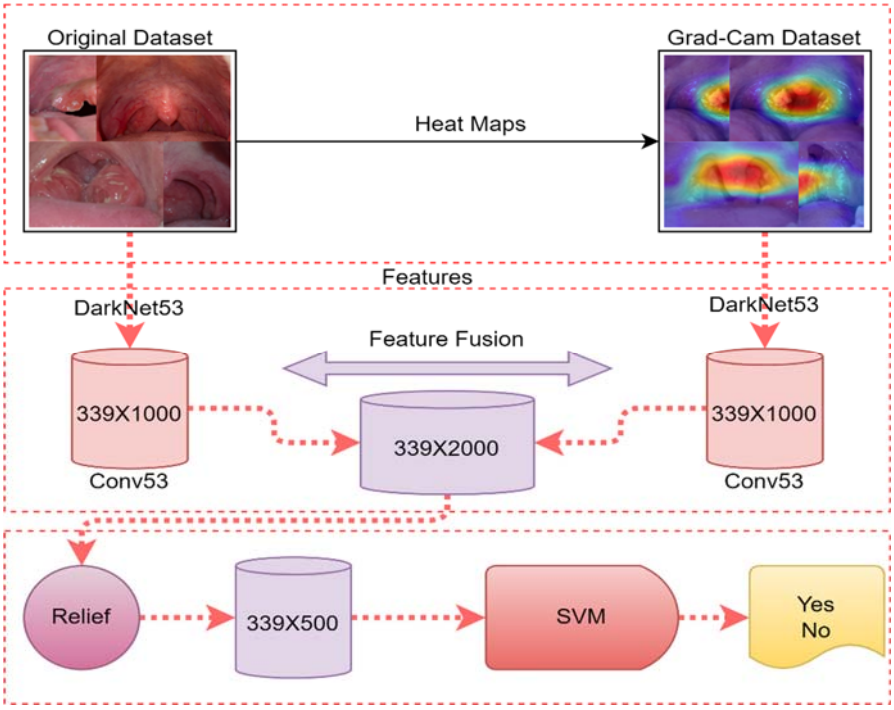


Figure 3. Proposed model for the classification of oropharyngeal images

As seen in Figure 3, feature extraction was done with DarkNet53 architecture. Features are taken from the Conv53 layer of the DarkNet53 architecture. 1000 features were taken for each image from the original data set and the data set created with Grad-Cam technology. The dimensions of the feature maps obtained from the original dataset and with Grad-Cam technology are 339x1000. After the merge, the size of the new feature map is 339x2000. For the model developed for classifying oropharyngeal images to work faster and more efficiently, the features in the size of 339x2000 were optimized with the Relief method. The size of the feature map optimized by the Relief method was 339x500. The optimized feature map in the last step is classified in the SVM classifier. The results were obtained by using the pre-trained models in the literature to compare the performances of the developed model.

3. RESULTS

In this study, which was carried out to diagnose tonsillopharyngitis using oropharyngeal images, the results of the application were obtained in a Matlab environment. To compare the performance of the proposed hybrid model, results were obtained with 8 pre-trained models accepted in the literature using the original data set. While training the pre-trained models, 75% of the images in the original dataset were used. The remaining 25% of the dataset is reserved for testing the models. In addition, the epoch value of 5, the InitialLearnRate value of 1e-4, and the MiniBatchSize value of 16 were selected while training the pre-trained models.

The hybrid model developed for classifying oropharyngeal images was also used as the basis of the DarkNet53 architecture, and feature extraction was made from both the original data set and the data set obtained using Grad-Cam technology. Then, the obtained features were combined and optimized and classified in different classifiers. In this way, each of the models used in the study was tested with a further 20% data set each time. Default training parameters and 5-fold cross-validation were used in shallow classifiers.

Different evaluation metrics were used to measure the performance of the models used in the study to diagnose tonsillopharyngitis disease using oropharyngeal images. Accuracy, Sensitivity, Specificity, Negative Predictive Value(NPV), False Positive Rate(FPR), False Negative Rate(FNR), False Discovery Rate(FDR), F1 score, and Matthews Correlation Coefficient(MCC) are the leading performance measurement metrics used in the study [28-30].

3.1. Results of pre-trained models

This study obtained results primarily in pre-trained models for diagnosing tonsillopharyngitis using oropharyngeal images. The results obtained in these models were compared with the developed model. The confusion matrix of the pre-trained models is in Figure 4.

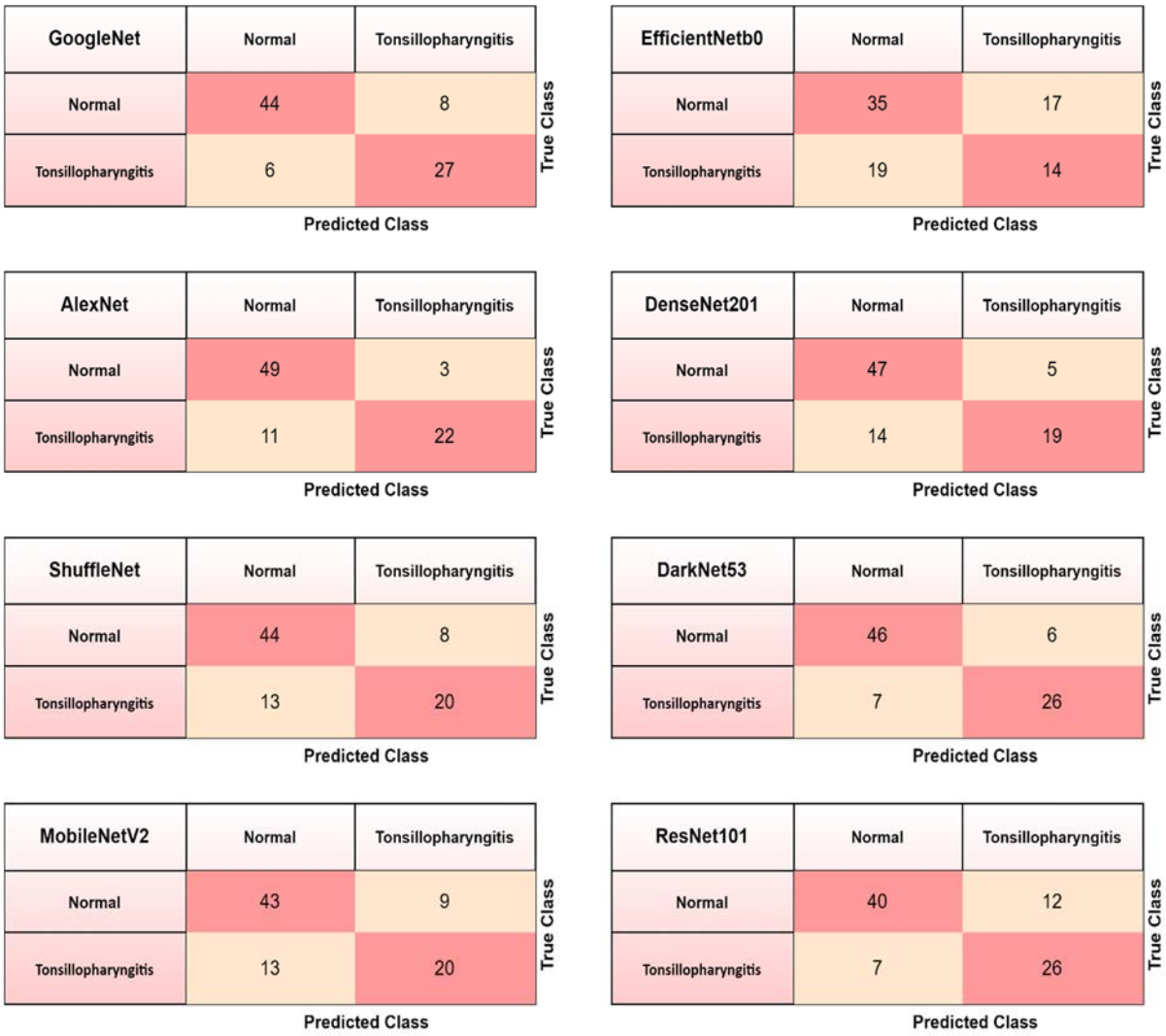


Figure 4. Confusion matrix of the pre-trained models

The EfficientNetb0 architecture predicted 49 correctly and 36 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images was 57.64%. The second model used to diagnose tonsillopharyngitis using oropharyngeal images is MobileNetV2. The MobileNetV2 architecture predicted 63 correctly and 22 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images was 74.11%. The third model used to diagnose tonsillopharyngitis using oropharyngeal images is ShuffleNet. This architecture predicted 64 correctly and 21 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 75.29%. The fourth model used to diagnose tonsillopharyngitis using oropharyngeal images is the ResNet101. The ResNet101 predicted 66 correctly and 19 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 77.64%. The fifth model used to diagnose tonsillopharyngitis using oropharyngeal images is DenseNet201. The DenseNet201 predicted 66 correctly and 19 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 77.64%. The sixth model used to diagnose tonsillopharyngitis using oropharyngeal images is the AlexNet. The AlexNet predicted 71 correctly and 14 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 83.52%. The seventh model used to diagnose tonsillopharyngitis using oropharyngeal images is GoogleNet. The GoogleNet predicted 71

correctly and 14 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 83.52%. The last model used to study is the DarkNet53. The DarkNet53 predicted 72 correctly and 13 incorrectly of the 85 images used for testing. The success of this model in classifying oropharyngeal images is 84.71%. The results of the pre-trained models used in classifying oropharyngeal images are in Table 1.

Table 1. Accuracy of pre-trained models

Model	Accuracy
EfficientNetb0	57.64%
MobileNetV2	74.11%
ShuffleNet	75.29%
ResNet101	77.64%
DenseNet201	77.64%
AlexNet	83.52%
GoogleNet	83.52%
DarkNet53	84.71%

When the results obtained in 8 different pre-trained models on oropharyngeal images for the diagnosis of Tonsilopharyngitis disease are examined in Table 1, it is seen that the highest accuracy rate is obtained in the DarkNet53 architecture. DarkNet53 is the most accurate model among the pre-trained models for classifying oropharyngeal images.

3.2. The results of the CNN architectures used as the base in the proposed model

In this section, DarkNet53 architectures are used as the basis and the results obtained in 7 different classifiers are examined. First, a new data set was created by applying the Grad-Cam method to the original data set, for which heat maps were obtained. In the next step, DarkNet53 architectures were used as the base, and feature maps were extracted from the two existing datasets. After these extracted feature maps were combined, unnecessary features were eliminated by the Relief method. After the Relief method, features decreased from 2000 to 500. In the last step, the feature map optimized with the Relief method was classified into 7 different shallow classifiers. The obtained accuracy values are presented in Table 2.

Table 2. Accuracy Rate of Proposed Models at Different Classifiers (%)

	DT	DA	LR	SVM	NB	KNN	SE
Proposed Model	72.0	72.9	55.2	**89.7**	79.1	87	86.1

In the last step of the proposed model, accuracy values were obtained using different classifiers. Among the 7 different classifiers used in the study, the highest accuracy value was obtained in the SVM classifier at 89.7%. In contrast, the lowest accuracy value was obtained in the LR

classifier at 55.2%. The confusion matrix obtained proposed model for classifying oropharyngeal images is presented in Figure 5.

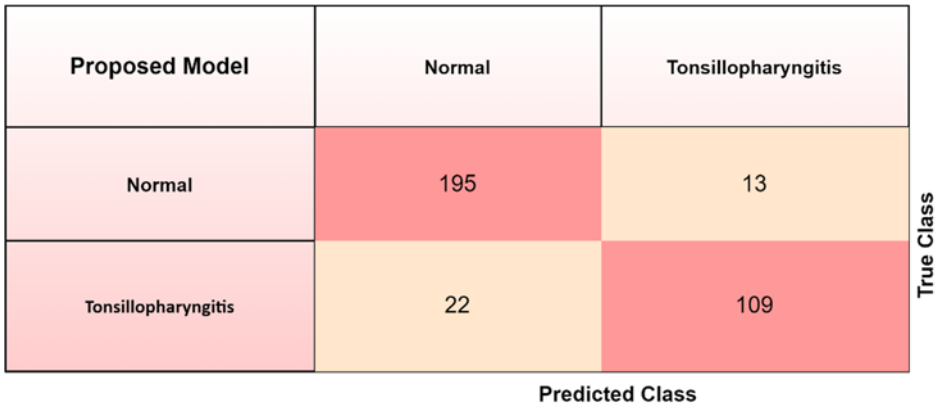


Figure 5. Confusion matrix of the proposed model for classification of oropharyngeal images

When Figure 5 is examined, it is seen that the proposed model for the classification of oropharyngeal images predicts 304 of 339 oropharyngeal images correctly and 35 of them incorrectly. It is seen that the proposed model predicts 195 of 208 Normal oropharyngeal images correctly and 13 of them as Tonsilopharyngitis. Similarly, out of 131 oropharyngeal images belonging to the Tonsilopharyngitis class, the proposed model predicted 109 correctly and 22 incorrectly. The success rate of the proposed model in classifying oropharyngeal images is 89.7%. The success metrics of the proposed model in classifying oropharyngeal images are in Table 3.

Table 3. Performance metrics of the proposed model for the classification of oropharyngeal images

Accuracy	**Sensitivity**	**Specificity**
89.7%	89.9%	89.3%
NPV	**FPR**	**FNR**
83.2%	0.1066	0.1014
FDR	**F1 score**	**MCC**
0.0625	91.8%	78.1%

When the performance evaluation metrics given in Table 3 are examined, it is seen that the proposed model for the classification of oropharyngeal images achieves high performance in the category of oropharyngeal images.

4. DISCUSSION

Tonsillopharyngitis is a common life-threatening infection when complications develop. It is an infection of the tonsil and pharynx tissues caused by bacteria or viruses. Chronic infection of the tonsils and obstructive hypertrophic tonsils are among the most common diseases in the pediatric age group [31]. It causes many complications, such as mouth breathing, nasal speech, maxillofacial developmental disorders, and psychosocial, neurocognitive, and developmental disorders in the following years [31]. In addition, serious complications such as peritonsillar

abscess, deep neck soft tissue infection, and arthritis may be seen in patients whose treatment is delayed [32].

Tonsillopharyngitis causes a sudden onset of addiction, resulting in loss of workforce. In addition, it is a significant health problem since it can be easily transmitted within the family and in public places. For this reason, early and accurate diagnosis is essential because the disease can spread rapidly and cause serious complications with high mortality if the treatment is delayed. In this study, tonsillopharyngitis was diagnosed using artificial intelligence methods in the data set consisting of tonsillopharyngitis and normal oropharyngeal images. High success rates have been achieved in the hybrid model developed for this purpose. The developed hybrid model was also compared with the pre-trained models in the literature. When the results obtained in the developed model are compared with the results obtained in the pre-trained models, it is seen that there is a significant increase.

The accuracy values of the pre-trained models and the proposed model used in the study for the classification of oropharyngeal images are compared in Figure 6.

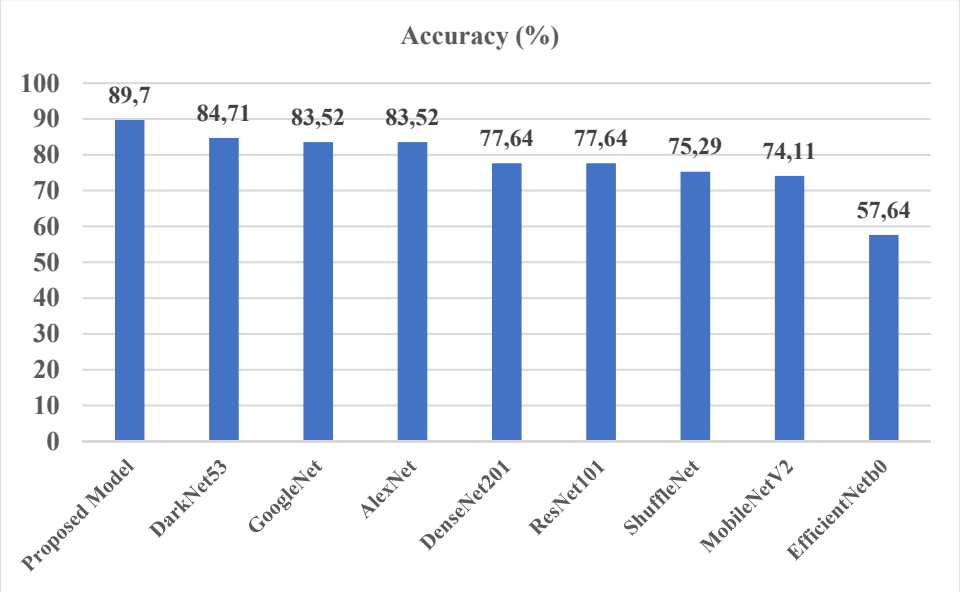


Figure 6. Compare of model's accuracy rate

When Figure 6 is examined, the proposed model's highest accuracy value is 89.7%. The closest value to this value was obtained in DarkNet53 with 84.71%.

This study used the CNN-based DarkNet53 model as the basis for feature extraction. Heat maps of the images in the data set were obtained using Grad-Cam technology. Features were extracted from the data sets created using Grad-Cam and interpolation method and these features were combined. The Relief dimension reduction method performed feature reduction after the feature fusion process. The number of features has been reduced from 2000 to 500. In this way, it is aimed to run the model faster. The optimized features in the last step of the model are classified in the SVM classifier.

The study's biggest limitation is that the data is obtained from a single center. Testing the model with data obtained from different centers would be more appropriate. It is among our aims to get data from various centers and work with more experts.

5. CONCLUSIONS

Tonsillopharyngitis is a disease with high morbidity and mortality. Detection of this disease by computer-aided systems is of great importance. Computer-assisted artificial intelligence-based systems have been used frequently in recent years, especially in the biomedical field, and successful results have been obtained. In this study, tonsillopharyngitis was diagnosed with an accuracy rate of 89.7% in the data set consisting of oropharyngeal images. This value shows that the proposed computer-assisted artificial intelligence-based model can be used to diagnose Tonsillopharyngitis.

Acknowledgments
The authors thank the owners of the database for sharing their data.

Declaration of Competing Interest
The authors declare that there is no conflict of interest in the study.

Author Contribution
The authors contributed equally to the article.

REFERENCES

[1] Bochner, R.E., Gangar, M., Belamarich, P.F., A clinical approach to tonsillitis, tonsillar hypertrophy, and peritonsillar and retropharyngeal abscesses. Pediatrics in Review, 2017. 38(2): p. 81-92.

[2] Vicedomini, D., et al., Diagnosis and management of acute pharyngotonsillitis in the primary care pediatrician's office. Minerva Pediatrica, 2014. 66(1): p. 69-76.

[3] Demir, N., Bayar Muluk, N., Chua, D., Acute Tonsillopharyngitis in Children, in Pediatric ENT Infections. 2022, Springer. p. 515-523.

[4] Osiejewska, A., et al., Acute tonsillopharyngitis-a review. Journal of Education, Health and Sport, 2022. 12(7): p. 873-882.

[5] Amiraraghi, N., et al., Intramural oesophageal abscess: an unusual complication of tonsillitis. BMJ Case Reports CP, 2019. 12(2): p. bcr-2018-226010.

[6] Bingol, H., NCA-based hybrid convolutional neural network model for classification of cervical cancer on gauss-enhanced pap-smear images. International Journal of Imaging Systems and Technology, 2022.

[7] Toğaçar, M., B. Ergen, Tümen, V., Use of dominant activations obtained by processing OCT images with the CNNs and slime mold method in retinal disease detection. Biocybernetics and Biomedical Engineering, 2022.

[8] Kiziloluk, S. and Sert, E., COVID-CCD-Net: COVID-19 and colon cancer diagnosis system with optimized CNN hyperparameters using gradient-based optimizer. Medical and Biological Engineering & Computing, 2022. 60(6): p. 1595-1612.

[9] Yoo, T.K., et al., Toward automated severe pharyngitis detection with smartphone camera using deep learning networks. Computers in biology and medicine, 2020. 125: p. 103980.

[10] https://data.mendeley.com/datasets/8ynyhnj2kz/1.

[11] Selvaraju, R.R., et al., Grad-cam: Visual explanations from deep networks via gradient-based localization. in Proceedings of the IEEE international conference on computer vision. 2017.

[12] Krizhevsky, A., Sutskever, I., Hinton, G.E., Imagenet classification with deep convolutional neural networks. Communications of the ACM, 2017. 60(6): p. 84-90.

[13] Redmon, J., Farhadi, A., Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.

[14] Huang, G., et al. Densely connected convolutional networks. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[15] Tan, M., Le, Q., Efficientnet: Rethinking model scaling for convolutional neural networks. in International conference on machine learning. 2019. PMLR.

[16] Szegedy, C., et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[17] Sandler, M., et al. Mobilenetv2: Inverted residuals and linear bottlenecks. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[18] He, K., et al. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[19] Zhang, X., et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[20] Yu, L., Liu, H., Feature selection for high-dimensional data: A fast correlation-based filter solution. in Proceedings of the 20th international conference on machine learning (ICML-03). 2003.

[21] Joachims, T., 11 making large-scale support vector machine learning practical, in Advances in kernel methods: support vector learning. 1999, MIT press. p. 169.

[22] Guo, G., et al. KNN model-based approach in classification. in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". 2003. Springer.

[23] Banfield, R.E., et al. A comparison of ensemble creation techniques. in International Workshop on Multiple Classifier Systems. 2004. Springer.

[24] Cramer, G., Ford, R., Hall, R., Estimation of toxic hazard—a decision tree approach. Food and cosmetics toxicology, 1976. 16(3): p. 255-276.

[25] Klecka, W.R., Iversen, G.R., Klecka, W.R., Discriminant analysis. Vol. 19. 1980: Sage.

[26] Lewis, D.D. Naive (Bayes) at forty: The independence assumption in information retrieval. in European conference on machine learning. 1998. Springer.

[27] Kleinbaum, D.G., et al., Logistic regression. 2002: Springer.

[28] Yildirim, M., Automatic classification and diagnosis of heart valve diseases using heart sounds with MFCC and proposed deep model. Concurrency and Computation: Practice and Experience, 2022: p. e7232.

[29] Tümen, V., SpiCoNET: A Hybrid Deep Learning Model to Diagnose COVID-19 and Pneumonia Using Chest X-Ray Images. Traitement du Signal, 2022. 39(4).

[30] Özbay, E., An active deep learning method for diabetic retinopathy detection in segmented fundus images using artificial bee colony algorithm. Artificial Intelligence Review, 2022. 1-28.

[31] Eroglu O, Keles E., Karlidag T, Kaygusuz I, Turker C, Yalcin S., Review of Our Tonsillectomy Indications Firat Med J, 2018. 23 (4): 178-183.

[32] Mazur, E., et al., Concurrent peritonsillar abscess and poststreptococcal reactive arthritis complicating acute streptococcal tonsillitis in a young healthy adult: a case report. BMC Infectious Diseases, 2015. 15(1): p. 1-5.