Prediction of Seizure Onset with Machine Learning: A Treatment Option for Landau Kleffner Syndrome

Anuprabha Dandapat[1] and Yin Kwong John Lee[#]

[1]Lexington High School
[#]Advisor

ABSTRACT

Landau Kleffner Syndrome (LKS) is a rare genetic disorder that presents in the form of auditory verbal agnosia and aphasia (loss of ability to interpret and express language) as well as electroencephalographic abnormalities, which present in children from the ages of 2 to 8. The disorder is classified as developmental/epileptic encephalopathy with spike wave activation on sleep (DEE-SWAS), with 70% of affected patients having epileptic seizures. The disorder is diagnosed based on language regression in addition to severe EEG abnormalities during non-REM sleep. Sometimes patients with epileptiform disorders can sense when a seizure is imminent and alert caretakers of the situation so they can be safely situated when the seizure occurs. This period is called the preictal period or prodrome in which patients can sometimes detect changes in mood and behavior. However, in disorders such as LKS, the loss of verbal expression proves difficult for expression of needs, and could lead to dangerous situations along with a constant need for such patients to be monitored. In this study, machine learning techniques are utilized to analyze several hours of ictal and preictal EEG data from 23 patients in order to predict the onset of a seizure. This study aims to promote patient safety and reduce need for constant monitoring by predicting when a seizure will occur. Several machine learning models have been constructed and their accuracies have been analyzed and compared to those of other studies to create an optimized program which can accurately interpret EEG data.

Introduction

Children affected by Landau Kleffner Syndrome (LKS) have trouble recognising speech and speaking (aphasia and agnosia). Their inability to recognise voices and interpret speech is often mistaken for hearing loss, but hearing loss is not associated with LKS. As the disease progresses, children may lose the ability to recognise non-language sounds like a horn beeping, clapping, etc. Eventually, the child may lose their ability to speak and may have difficulty reading and writing. Behavioral hyperactivity and attention deficit are also commonly associated with LKS. These symptoms interfere with the child's ability to learn in a classroom setting and form social relationships with peers. Landau Kleffner Syndrome (LKS) is a rare genetic syndrome in childhood that manifests in the form of loss of verbal and auditory function and epileptic activity. The disease occurs in children from the age of 2 to 8 and is twice as prevalent in male patients. The underlying causes of this disorder remain unknown, but it has been linked to genetic abnormalities. The rare pediatric disease is considered to be a subtype of developmental/epileptic encephalopathy. It is generally followed by progressive neuropsychological impairment and is characterized by the appearance of paroxysmal electroencephalograph (EEG). Additionally, epileptic seizures develop in 70% of patients with LKS.

During a seizure, loss of motor control and sudden jerking movements can pose a significant threat to the patients with epileptic disorders. In the first stages of a seizure, a patient can usually detect the impending event, and alert caretakers, or position themselves such that their surroundings are conducive to their safety for the duration of the seizure. However, often it is not possible for patients with LKS to communicate their needs, especially because the disease presents in young children, and due to language regression. Patients with LKS often are subject to frequent hospital visits due to injuries from seizure events. Consequently, caretakers often need to constantly monitor the patient.

In order to prevent this, machine learning can be harnessed to predict seizure onset in combination with at-home EEG devices. Machine learning is a branch of artificial intelligence in which a computer is able to learn based on the mistakes it makes on a certain amount of training data, without being explicitly programmed. The resulting model can be used to predict other data with high accuracy, even when variables are changed.

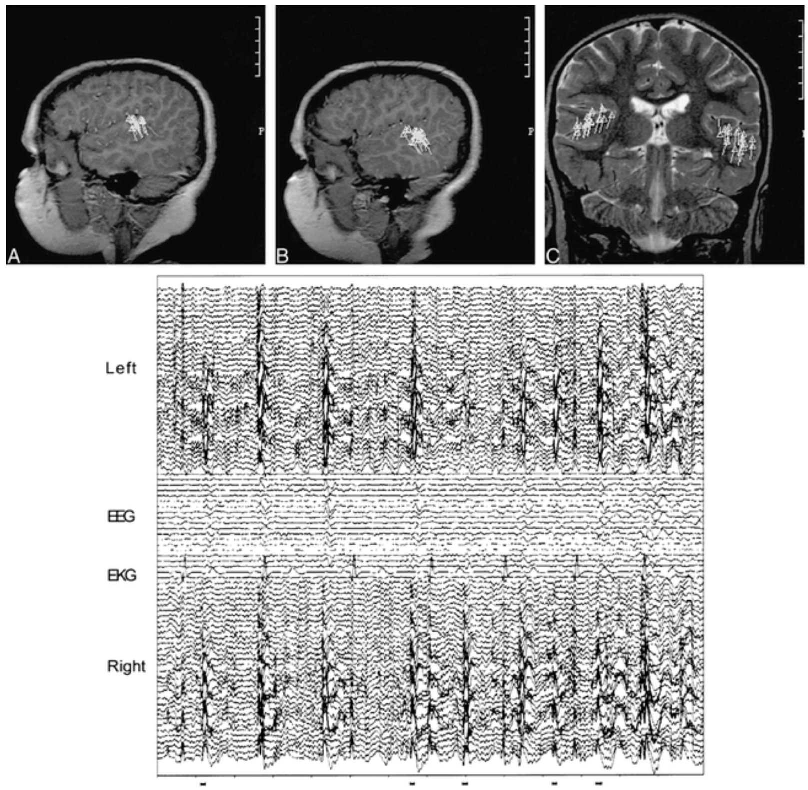

Figure 1. Images from the EEG of a 6-year-old male patient. In figure 1a, 1b, and 1c sagittal T1-weighted images from the right and left hemispheres shows perisylvian clustering of spike activity in Wernicke's area (speech comprehension area. 1d shows MEG wave forms from the left hemisphere, right hemisphere, and concurrent EEG. (Image taken from Sobel et al. AJNR Am J Neuroradiol 2000, 21 (2) 301-307)

LKS is diagnosed by performing an EEG. Along with significant loss of language comprehension, presence of severe epileptic discontinuities is required for diagnosis. Additionally, a magnetoencephalogram can be used for detection of epileptic activity. Magnetic resonance imaging (MRI) can be used to make sure that symptoms are not due to other underlying causes. Other testing includes behavioral and/or brainstem evoked audiometry and standardized psychometric and speech/language testing to exclude hearing loss and provide the basis for therapies to aid in recovery. The cause of LKS remains unknown, but LKS along with other epileptic disorders have been linked with mutations in the GRIN2A gene. It has been speculated that the disease could also be linked with the RELN, BSN, EPHB2 and NID2 genes. Due to some patients responding well to immunosuppressive medication, autoimmune mechanisms have also been proposed. Treatments usually include a combination of antiepileptic drugs and speech therapy. Due to deteriorating communication skills, patients with this disorder may be unable to communicate their needs, and could get injured during a seizure if not in a safe environment. In this study, we intend to compare different machine learning models to predict the onset of seizures in these patients, in order to promote patient safety.

HIGH SCHOOL EDITION
Journal of Student Research

Seizures

Seizures are uncontrolled bursts of electrical activity in the brain which cause a variety of symptoms including loss of consciousness and uncontrollable muscle jerking. Patients with LKS generally develop focal motor seizures and tonic-clonic seizures, and often have to go to the hospital because of resulting injuries. During a seizure, the patient might bite the inside of their mouth and draw blood, or have trouble breathing due to stiffness of chest muscles. A seizure is composed of four phases, during which the patient experiences different symptoms. The prodromal phase is often not considered part of the seizure, but occurs hours or days before the seizure occurs. In this phase, patients may feel a sense of irritability or confusion. The period directly before the seizure is called the preictal phase or the aura, in which a variety of symptoms are experienced, including dizziness, nausea, and a bitter taste in the mouth. The preictal stage is the first time when the patient can sense that a seizure is imminent. Not all patients experience this phase, but activity can be detected on the EEG. Finally, the ictal and post ictal phases are those during and after the seizure respectively.

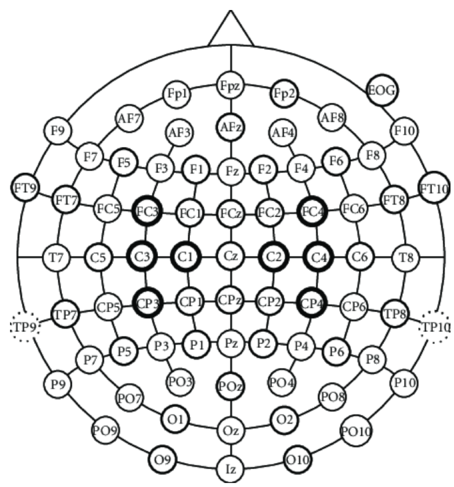

Figure 2. EEG electrode placement. The diagram shows placement of the electrodes on the scalp for a total of 39 electrodes. Eight electrodes over or close to the motor cortex are shown in bold circles (C1, C2, C3, C4, FC3, FC4, CP3, and CP4). (Picture taken from Schroder et al. EURASIP Journal on Applied Signal Processing 2005:19, 3103–3112)

EEGs are a noninvasive technique used to monitor electrophysiological activity in the brain. It can be used during surgery to monitor the amount of anesthesia required. For those with epilepsy, rapid spiking motion is visible during the ictal period of the seizure. A number of electrodes are placed on the scalp to monitor electrical activity. The EEG measures polarity between different electrodes to generate a voltage-time graph. Machine learning can be used to interpret the series of amplitudes in the graph at any given time to determine whether the patient is in the preictal stage, so that caretakers can be alerted.

Related Literature

The intention of this study is to use data from the ictal and preictal stages of seizure from various patients through use of machine learning models. The resulting models will be able to be generalized to various epileptic disorders, including LKS. Proper preprocessing is required for optimal accuracy of the generated models, and to improve on findings of other studies.

Landau Kleffner Syndrome

A study by Pearl et al. describes the symptoms, EEG characteristics, and treatments for the disorder. The paper states that epileptiform discharges are triggered by sleep onset and occur throughout non-REM sleep. Pearl et al. further explores the connection between Autism Spectrum Disorder (ASD) and LKS, stating epileptiform tendencies in patients with ASD. The study found that 19% of the 894 patients with ASD had epileptiform potential. Due to high potential for relapse, patients with LKS need extensive care from parents, speech/language therapists, neuropsychologists, and neurologists.

EEG Data Analysis with Machine Learning

Classification analysis of EEG signals with machine learning techniques has been demonstrated by several studies. SVM and KNN models were shown to perform proficiently in most studies. Usman et al. developed an SVM algorithm with 92.23% sensitivity using the CHB-MIT dataset to predict onset of seizures. The 23 channels were converted into a single signal. In addition, empirical mode decomposition was performed to increase the signal to noise ratio.
Yoo et al. used similar EEG data to predict cognitive load using adaptive boost and gradient boosting algorithms. These algorithms had accuracies and F1 scores ranging from 60-70%. Their highest performing model was the bi-LSTM model with 87.10% accuracy.

In addition, Zandi et al. and Teixeira et al. have also used similar techniques for prediction of seizure onset. Zandi et al. have proposed a SVM model using 18 EEG channels for predicting seizures using scalp EEG signals and a Bayesian Gaussian mixture model. The proposed model had 91.11% accuracy.

Teixeira et al. have proposed a model for prediction of seizures by using only six EEG channels, and have extracted 22 linear univariate features for each channel. The overall feature space expands to 132 dimensions. The use of fewer EEG channels is in order to minimize discomfort for the patient, so that only 6 electrodes have to be attached to the patient's scalp. The accuracy of the model was 73.5%.

Aljalal et al. have used EEG signals in detection of Parkinson's Disease, discrete wavelet transform, different entropy measures, and machine learning techniques. Features are extracted from wavelet packet-derived reconstructed signals using log energy entropy, Shannon entropy, threshold entropy, sure entropy, and norm entropy. The accuracy of the DWT + TShEn and KNN classifier was 99.89%, using a small number of EEG channels. Limitations for all papers include a lack of publicly available EEG data for reproducibility of results.

Methods

Dataset

The CHB-MIT Scalp EEG Database contained data from a period of 4096 seconds of both ictal and preictal data from 24 patients. In the full preprocessed data, there are 23 channels of EEG data with the last column being the outcome column, with '0' indicating that the data was taken from the preictal stage of a seizure, and '1' indicating that the measured EEG signal was from the ictal stage of the seizure. The initial dimensions of the data were 2097150 x 24, where there was exactly the same amount of ictal and preictal seizure data. The 2D dataset was extracted from the raw '.edf' files to give the '.csv' classification data.

HIGH SCHOOL EDITION
Journal of Student Research

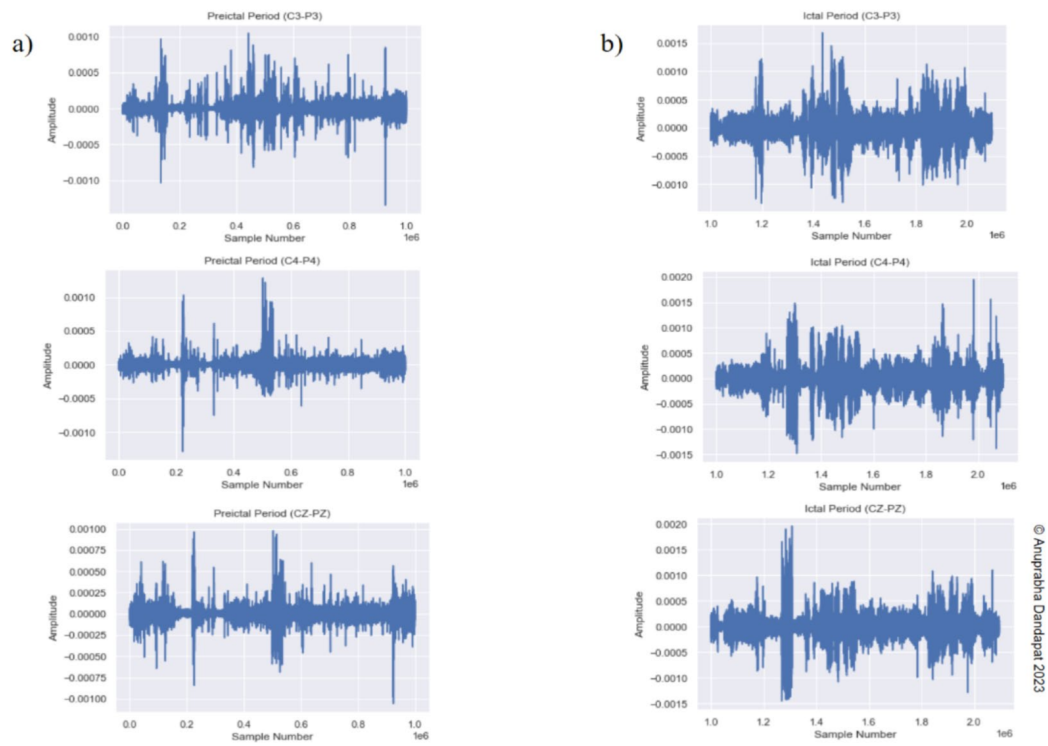

Figure 3. Ictal and preictal EEG from CHB-MIT dataset. Figure 3a shows the preictal EEG graph for 3 EEG Channels (C3-P3, C4-P4, CZ-PZ). Figure 3b shows the ictal EEG graph for the 3 given channels. It can be observed that in the ictal period there is a rapid spiking motion, which distinguishes the seizure phase from other phases.

Exploratory Data Analysis And Data Preprocessing

Exploratory Data Analysis

Exploratory data analysis (EDA) was used to visualize the data and identify data distribution. In order to do this, several plots were generated to assess data correlation and distribution. From the correlation matrix it was noted that there was a lack of correlation between features in general, but columns ‘F3-C3’ and ‘F7-T7’ had moderately high positive correlation with ‘FZ-CZ.’ From histogram plots, scatterplots, and boxplots, it was found that the data was clustered around 0, and formed a symmetric bell shaped curve for all features. In order to remove noise from the data, any values that lied more than 3 standard deviations from the mean were removed.

HIGH SCHOOL EDITION
Journal of Student Research

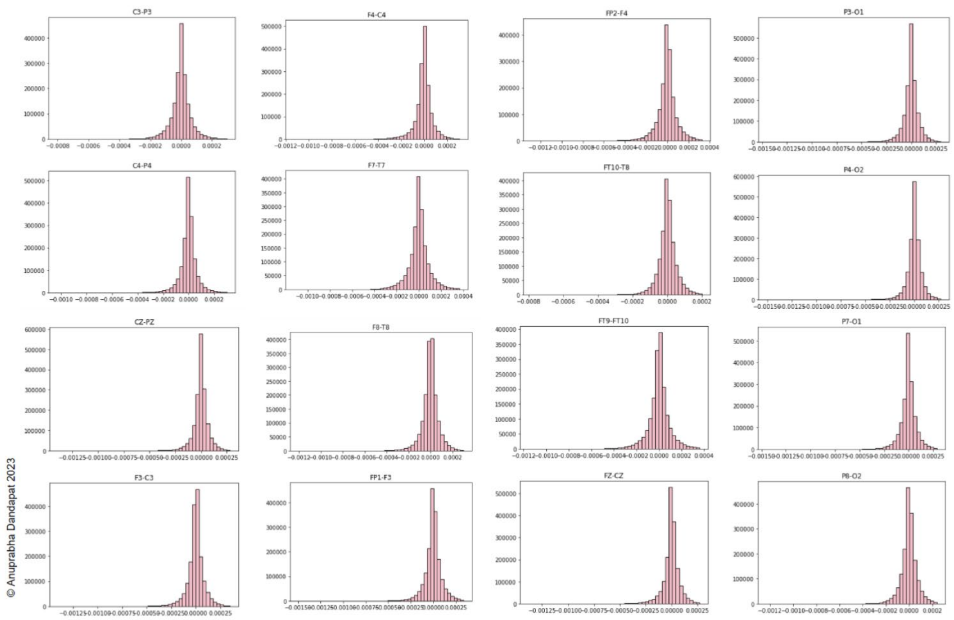

Figure 4. Histograms representing data distribution from 16 EEG channels. It can be observed that data cluster around 0 and are symmetric.

Data Preprocessing

Data preprocessing is an extremely crucial step needed to remove any values or features which could negatively impact the accuracy of the model. In this stage, null values, outliers, correlated data, etc. are removed from the dataset in order to reduce chance of overfitting or underfitting. Feature importance scores were calculated using scikit learn and all columns with importance scores of less than 0.03 were removed from the dataset.

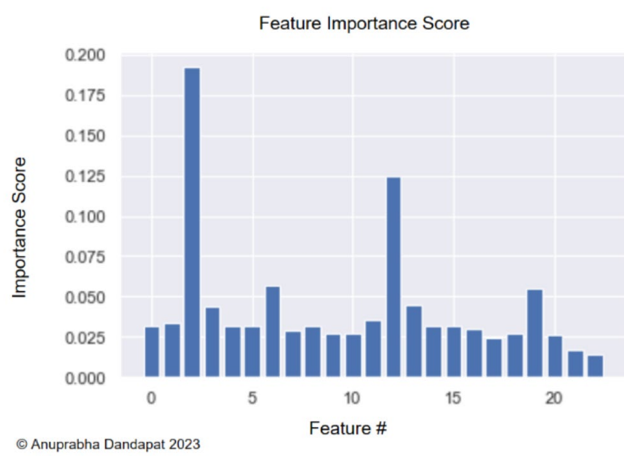

Figure 5. Feature Importance Scores. The features were analyzed to determine their importance in determining the outcome. As a result, labels with feature importance scores of less than 0.03 were dropped from the dataset.

Additionally, multicollinearity was assessed using the variance inflation factor. Multicollinearity occurs when there are multiple independent variables with strong correlation, which may interfere with model performance. The formula for the variance inflation factor is shown in Equation 1, where R^2 is the coefficient of determination.

Equation 1: Variance inflation factor used for testing multicollinearity:

$$VIF = \frac{1}{1 - R^2}$$

All columns with variance inflation factor greater than 5 were removed from the dataset, as they indicated high correlation with other independent variables.

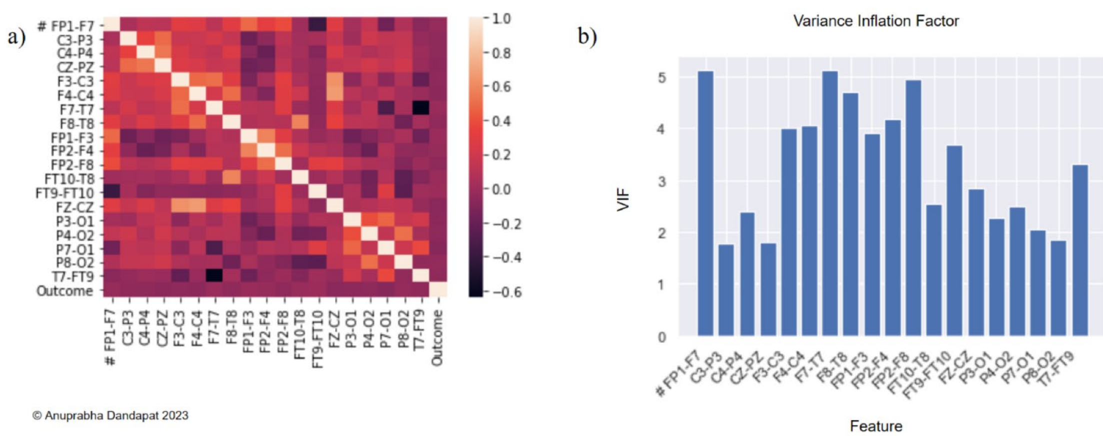

Figure 6. Collinearity Analysis. Correlation of independent variables is shown above. In figure 6a, the heatmap shows the correlation of the independent variables toward each other. Figure 6b shows the graphed variance inflation factor for each parameter. Features with VIF higher than 5 were removed.

Using scikit learn's train_test_split function, the data was randomly split into training and testing data, where 80% of the data was used as training data, and the remaining 20% was used as testing data. As a result of preprocessing, 17 features were retained to be used in classification models.

Model fitting

Samples were classified as precital or interictal using a variety of models. Models classified a sample as preictal or interictal based on the sequence of amplitudes over a period of 2 seconds. 7 models were created to analyze the data. Metrics measured for all models included accuracy, F1 score, and AUC score, where F1 score is the accuracy after being penalized for false positives and negatives. AUC score is the area under the Receiver Operating Characteristics (ROC) curve, which indicates the ability of the model to distinguish between classes. Confusion matrices were also generated for all models.

K Nearest Neighbors

K Nearest Neighbors (kNN) is a nonlinear supervised learning model that can be utilized for both classification and regression. Each datapoint in the model, as observed in Figure 7 is analyzed by determining its nearest values and using them to predict the outcome of the datapoint.

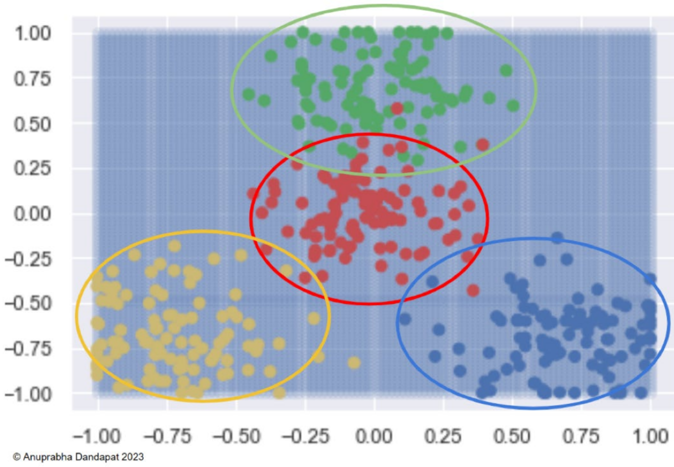

Figure 7. KNN model functionality. Class labels for each data point are determined based on the n neighbors closest to them.

The length of the vector between data points is computed by using the Euclidean formula (Equation 2), in which the distance between two points is computed by taking the square root of the sum of squared differences in each dimension.

Equation 2: Euclidean formula used to compute norm of distance vectors:

$$d(a,b) \ = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \ldots + (a_n - b_n)^2}$$

In this study, the model was fitted to the data and used the 3 nearest neighbors to predict outcomes. The accuracy was assessed by using the root mean square error (rmse).

Support Vector Machine

Support Vector Machine (SVM) algorithms are supervised machine learning models used for classification and regression. The model utilizes discriminative classification to find a curve that separates outcomes. The curve has a margin of a certain width and is drawn up to the nearest point to separate the different classes, as shown in Figure 8. The curve that allows for the greatest margin is used in the model to guide predictions. The points that lie on the margins are known as the support vectors and essentially dictate the structure and function of the model.

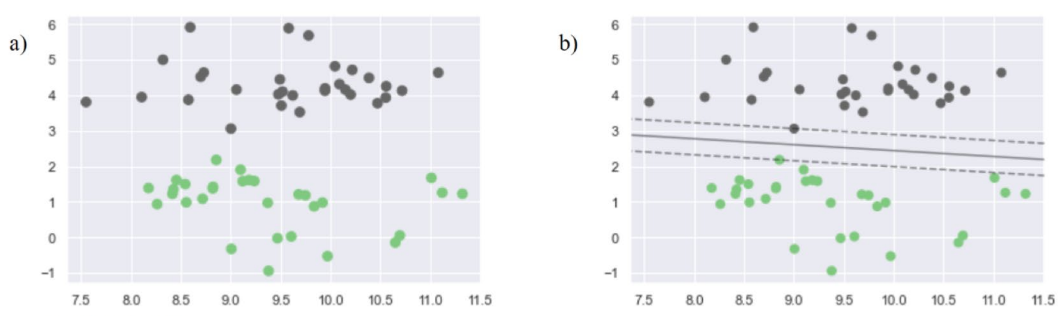

Figure 8. SVM Model Functionality. The SVM model functions through discriminative classification in which boundaries between labels are drawn with a margin, and the model is optimized by maximizing the number of data points that lie on the margin. As seen in figure 8a, the points can be classified into two categories by several curves. In figure 8b, the optimized model is shown with maximum margin size and with the support vectors lying on the margin.

HIGH SCHOOL EDITION
Journal of Student Research

Decision Tree

Decision tree classifiers are flowcharts that are composed of several nodes to output a final classification. There are three types of nodes: root node, internal node, and leaf node, where the leaf node contains the overall outcome. Samples go down the decision tree cascade and are assigned a label based on the leaf node they arrive at. The tree is built by assessing which features are best for outcome prediction. Figure 9 shows the general structure of a decision tree model.

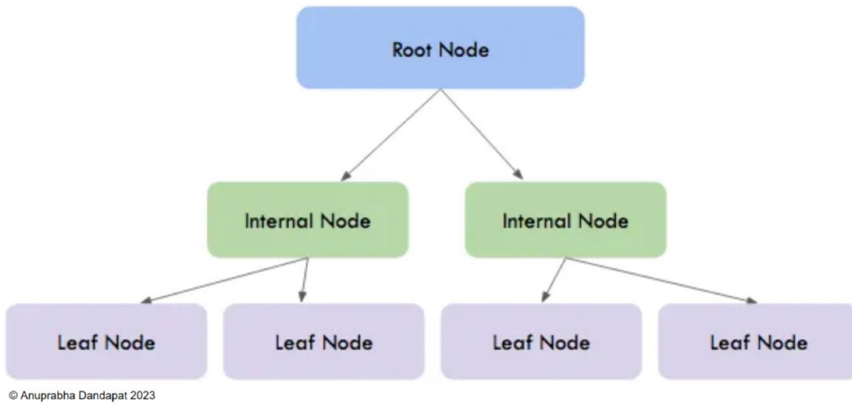

Figure 9. A basic decision tree with one root node, two internal nodes, and four leaf nodes.

Determining the importance of features is done by calculating the Gini impurity of each sample, which is a number between 0 and 0.5 that determines the probability of misclassifying a random datapoint. As a result, the root node is the feature with the lowest impurity score, and the leaf nodes are the ones with the highest impurity scores in order to optimize the classification cascade. Equation 3 shows how the Gini impurity is calculated, where K is the number of class labels.

Equation 3: Gini Impurity formula:

$$\text{Gini Impurity}(df) \; = \; 1 - \sum_{i=0}^{k} (p_i)^2$$

Random Forest

Random forest classifiers use the same principle as decision trees, but instead of one tree, the classifier takes the input of multiple different trees to generate an output. The classifier takes datapoints at random to create a bootstrapped dataset, and generate a decision tree. This process is repeated to make hundreds of decision trees, and is called bagging. The result of using multiple decision trees is a far more accurate classification tool than a singular decision tree.

AdaBoost

AdaBoost is another decision tree based algorithm in which several weak learners are combined to make a stronger classifier. Each tree in the AdaBoost model consists of only a root node and two leaf nodes (a stump), and has a different amount of influence on the final classification. As shown in Figure 10, the outcomes of the stumps depend on the previous stumps. Each stump is made based on the errors that the previous stump made in classifying the sample. If a certain sample is incorrectly classified, its weight is increased in order to emphasize the need to properly classify it in the next stump.

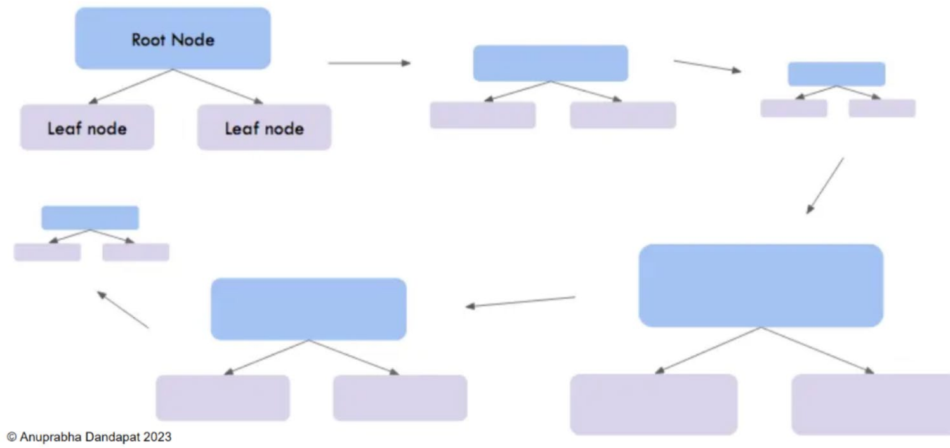

Figure 10. A random forest with AdaBoost. As shown, each stump takes input from the previous stumps to optimize classification.

Sample weights determine the amount of say each stump has on the final classification. All sample weights are initially equal to the reciprocal of the number of total samples, and adjusted as the algorithm learns. Based on the total error, each sample has a different amount of say, in accordance with Equation 4, where a is the amount of say, and x is the total error.

Equation 4: Amount of say for each stump:

$$a = \frac{1}{2} \cdot log(\frac{1-x}{x})$$

Following the above equation, the amount of say is zero when the total error is ½, as it would be no different from flipping a coin. Above a total error of ½, the amount of say becomes increasingly negative, and below ½ the amount of say is increasingly positive. In addition, sample weights for each stump are adjusted by multiplying the initial sample weight by e^{a} or e^{-a} depending on whether the stump classified the sample correctly or incorrectly.

Gradient Boosted Regression Tree

Gradient boosted regression trees (GBRTs) are similar to AdaBoost, but use trees with larger depth instead of stumps. This algorithm generates regression trees instead of classification trees, and predicts residuals instead of class labels. The split for each tree is determined by finding the Mean Squared Error (MSE), and using the one with the lowest MSE as calculated by Equation 5.

Equation 5: Mean Squared Error:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Observed_i - Predicted_i)^2$$

After the performance of the tree is evaluated, the process is repeated to generate the next tree.

Multilayer Perceptron

The Multilayer Perceptron (MLP) Classifier is a deep learning model that optimizes the log-loss function using stochastic gradient descent. The model trains using backpropagation, and is able to fit non linear data. The constructed model had 2 hidden layers, and used cross entropy to measure loss over 100 iterations. Figure 11 shows the MLP model that was constructed in this study with the appropriate number of neurons and layers.

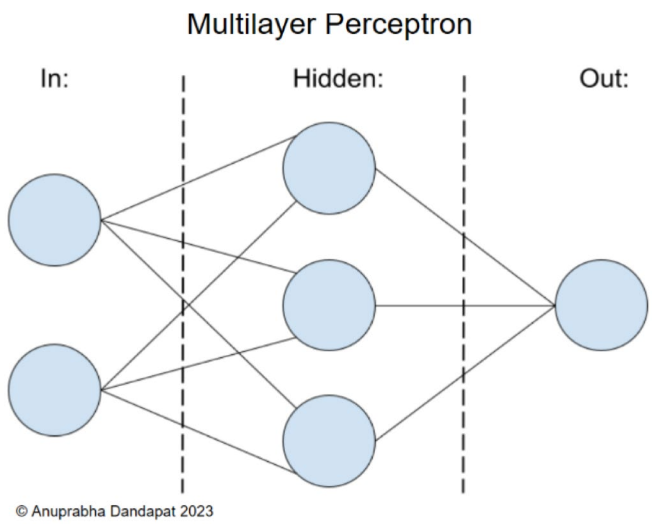

Figure 11. Multilayer Perceptron Classifier with 1 hidden layer.

Neural Network (Sequential Model)

A neural network was created in order to accurately predict new values using multiple different parameters. Neural networks consist of many different layers which continue in a chain-like manner until an output is produced. Depending on how important a given parameter is to the final result, each node is assigned weights and biases. The result from each node influences the result in the subsequent nodes, and is thus well suited for dealing with non-linearities. Cycles of back and forward propagation allow for the model to be trained effectively. The model was constructed using keras, and consisted of 5 dense layers (Figure 12).

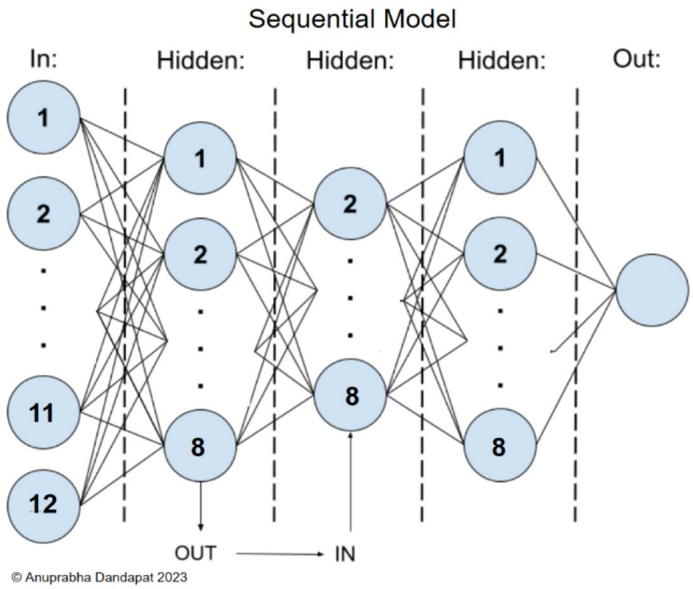

Figure 12. Artificial Neural Network with 3 hidden layers. As shown in the figure, the input layer has 12 neurons, each of the hidden layers has 8 neurons, and the output layer consists of 1 neuron.

The model was optimized with the adam optimizer, and used ReLU and sigmoid activation and accuracy was assessed using binary cross entropy. Training was observed over 200 epochs.

Results

In order to identify a suitable model, 8 different machine learning models have been compared. After analysis of the 8 different models, the following results were obtained. The Artificial Neural Network was selected due to its high accuracy of 97.76%. The accuracy and loss for the first 50 epochs are plotted in Figure 13, and indicate the learning rate of the model. The first 7 models took 5-10 minutes to be fitted, while the neural network took over 250 minutes to compile.

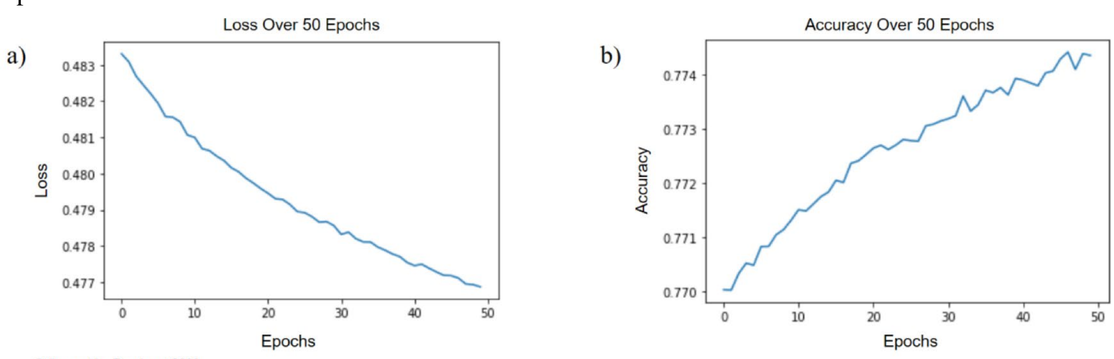

Figure 13. Loss and accuracy over the first 50 epochs of training for the ANN model. Figure 13a shows the decrease in loss over the training period. Figure 13b shows the increase in accuracy over the training period.

The two highest performing models (KNN and ANN) have shown higher accuracy than those produced by Usman et al., Zandi et al., and Teixeira et al. Comparison between the proposed models is shown in Table 1.

Table 1. Comparison of accuracy between proposed models

Model	Dataset	EEG Signal Type	Number of EEG channels	Number of Subjects	Accuracy %
Teixeira et al.	EPILEPSIAE	Scalp EEG	6	227	73.5
Zandi et al.	VGH	Scalp EEG	18	17	91.11
	CHB-MIT	Scalp EEG	23	3	83.81
Usman et al.	CHB-MIT	Scalp EEG	23	24	92.23
Proposed Model	CHB-MIT	Scalp EEG	17	23	97.76

The accuracies of the remaining 7 models were assessed. It was found that the KNN model also performed well (94.53%), while all tree classification models performed with around 80% accuracy, with the basic decision tree performing the best. The F1 scores were calculated to be very close to the accuracy scores, which indicates that there were not many false positives and negatives. Confusion matrices were plotted (Figure 14) to assess the true/false positive/negative rate for all four tree models. This showed that the basic decision tree model had the lowest false negative and positive rate, which is also reflected in the F1 and AUC score metrics.

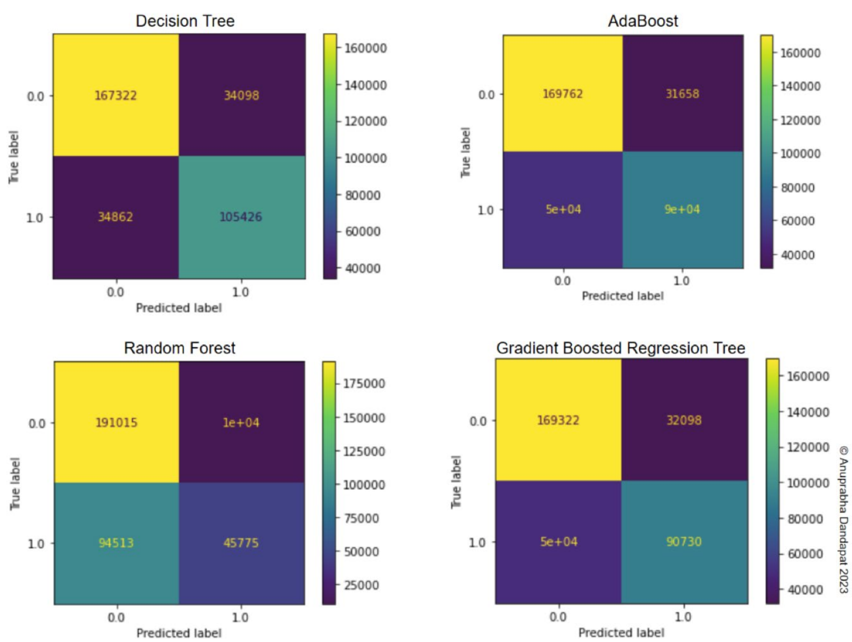

Figure 14. Confusion matrices for 4 decision tree based models. The basic decision tree produced the highest accuracy and thus has a comparatively low false negative/positive rate. The top left and bottom right represent the number of true positives and true negatives respectively, while the top right and bottom left represent the false positives and false negatives respectively. False negatives are the most dangerous in the context of seizure prediction, because this means the patient will not be alerted even though they are about to have a seizure.

Table 2 shows all calculated metrics for the 8 models, and Figure 15 plots the accuracies of all models. Although the MLP classifier utilized dense learning techniques, the accuracy was lower than any other model, as it was only trained for 12 epochs, and had only one hidden layer. Furthermore, the F1 and AUC scores were observed, and the results indicate that there was high sensitivity due to low number of false positives and negatives, and the model was able to effectively distinguish between ictal and preictal states, as shown in the AUC scores.

Table 2. Accuracy scores and other metrics for 8 constructed models.

Model Name	Accuracy (%)	F1 Score	AUC score
K Nearest Neighbors (KNN)	94.53	0.95	0.974
Support Vector Machine (SVM)	77.81	0.78	0.760
Decision Tree Classifier	79.82	0.80	0.791
Random Forest Classifier	69.29	0.69	0.813
AdaBoost Classifier	75.97	0.76	0.836
Gradient Boosted Regression Tree (GBRT)	76.10	0.76	0.837
MLP Classifier	58.95	0.59	0.500

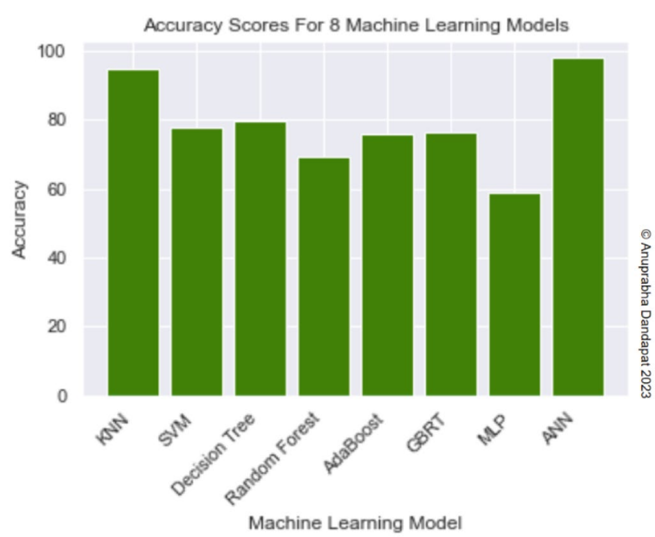

Figure 15. Comparison of accuracy for 8 constructed models. As shown, the KNN and ANN algorithms performed the best out of the constructed models with 94.53% and 97.76% accuracies respectively.

It is evident that the proposed model(s) have higher accuracy than currently proposed models and thus could potentially be used as a tool for patients with LKS and other epileptic disorders.

Discussion

The study was performed to generate 8 machine learning models for the prediction of epileptic seizures. The intended use of these models is to use them in diagnosis of LKS and other epileptic disorders. The models performed with suitable accuracy, and can be generalized to patients with other epileptic disorders. In comparison to models from other studies who performed similar classification analysis on EEG data, the proposed model performed better in terms of accuracy. The two selected models for the most optimal epileptic seizure prediction in this study were the KNN (94.53%) and ANN (97.76% over 200 epochs) models. Zandi et al. and Usman et al. both produced SVM models with highest accuracy out of their other constructed models. It was found that the SVM model produced in this study did not have comparable accuracy to those of other studies (58.95% over 100 epochs), but the KNN and ANN models showed higher accuracy than other models for epileptic seizure prediction. Use of 17 channels proved to perform better than models using a smaller number of channels. Usman et al. similarly used input from several channels, and performed empirical mode decomposition to combine input from 23 channels. Furthermore, the tree models including the decision tree, random forest, AdaBoost model, and Gradient Boosted Regression tree model performed similarly to those of Yoo et al (70-80%).

Conclusion

In this study, we have used ictal and preictal data from 23 patients from the CHB-MIT dataset in order to predict seizure onset. The intended use of the constructed models is to create a tool for LKS patients that can alert caretakers about seizures prior to the event. The data was rigorously preprocessed, and 8 models were constructed to analyze the EEG data and determine whether the patient was about to have a seizure. Of these 8 models, the MLP Classifier consisting of 2 dense layers performed the worst (58.95% over 100 epochs), the 4 tree algorithms performed with acceptable accuracy (70-80%), and the KNN (94.53%) and Artificial Neural Network with 5 dense layers (97.76% over 200 epochs) performed the best, and surpassed accuracy scores of other existing studies. These models can be used in combination with available at-home EEG technologies for patients with diseases like LKS in order to ensure patient safety. In the future, we hope to construct a similarly accurate model which can take into account more factors to generate predictions. As the neural network was shown to perform the best, it may be useful to explore other types of neural networks or to add layers to the existing model to enhance accuracy. In the future this model can be fitted with data from more patients, and those with different types of seizures in order to produce more useful and accurate outputs.

Limitations

In order to enhance the models, a larger amount of data from various individuals is required. Also to be improved is the time complexity of the programs, which greatly hinders efficiency in the training stage.

References

1.17. neural network models (supervised). scikit. (n.d.). https://scikit-learn.org/stable/modules/neural_networks_supervised.html#:~:text=Starting%20from%20initial%20random%20weights,meant%20to%20decrease%20the%20loss.

A. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont. (2013). Predicting epileptic seizures in scalp EEG based on a variational bayesian gaussian mixture model of zero-crossing intervals," IEEE Transactions on Biomedical Engineering, vol. 60, no. 5, pp. 1401–1413.

C. A. Teixeira, B. Direito, M. Bandarabadi. (2014). Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients, Computer Methods and Programs in Biomedicine, vol. 114, no. 3, pp. 324–336.

David F. Sobel, Maung Aung, Hiroshi Otsubo and Michael C. Smith. (2000). Magnetoencephalography in Children with Landau-Kleffner Syndrome and Acquired Epileptic Aphasia. AJNR Am J Neuroradiol, 21 (2) 301-307

Dobilas, S. (2022, February 5). *Gradient boosted trees for classification-one of the Best Machine Learning Algorithms*. Medium. https://towardsdatascience.com/gradient-boosted-trees-for-classification-one-of-the-best-machine-learning-algorithms-35245dab03f2

Landau Kleffner syndrome - symptoms, causes, treatment: Nord. National Organization for Rare Disorders. (2023, May 1). https://rarediseases.org/rare-diseases/landau-kleffner-syndrome/#diagnosis

Landau-Kleffner Syndrome in children: Children's health. Landau-Kleffner Syndrome. (n.d.). https://www.childrens.com/specialties-services/conditions/landau-kleffner-syndrome

Light, G. A., Williams, L. E., Minow, F., Sprock, J., Rissling, A., Sharp, R., Swerdlow, N. R., & Braff, D. L. (2010, July). *Electroencephalography (EEG) and event-related potentials (erps) with human participants*. Current protocols in neuroscience. CHAPTER: Unit–6.2524. doi:10.1002/0471142301.ns0625s52.

Majid Aljalal, Saeed A. Aldosari, Marta Molinas, Khalil AlSharabi, Fahd A. Alturki. (2022). Detection of Parkinson's disease from EEG signals using discrete wavelet transform, different entropy measures, and machine learning techniques. Scientific Reports 12, 22547. https://doi.org/10.1038/s41598-022-26644-7

Michael Schroder, Thomas Navin Lal, Thilo Hinterberger, Martin Bogdan, N.JeremyHill, NielsBirbaumer, WolfgangRosenstiel, Bernhard Scholkopf. (2005). Robust EEG Channel Selection across Subjects for Brain-Computer Interfaces. EURASIP Journal on Applied Signal Processing:19, 3103–3112.

Pearl PL, Carrazana EJ, Holmes GL. The Landau-Kleffner Syndrome. Epilepsy Currents. 2001;1(2):39-45. doi:10.1111/j.1469-5812.2005.00134.x-i1

Sobel, D. F., Aung, M., Otsubo, H., & Smith, M. C. (2000, February 1). Magnetoencephalography in children with Landau-Kleffner syndrome and acquired epileptic aphasia. American Journal of Neuroradiology. Vol. 21, Issue 2

Tf.keras.sequential : tensorflow V2.13.0. TensorFlow. (n.d.). https://www.tensorflow.org/api_docs/python/tf/keras/Sequential

U.S. Department of Health and Human Services. (n.d.). *Landau-Kleffner syndrome*. National Institute of Neurological Disorders and Stroke. https://www.ninds.nih.gov/health-information/disorders/landau-kleffner-syndrome

Usman, S. M., Usman, M., & Fong, S. (2017). *Epileptic seizures prediction using machine learning methods*. Computational and Mathematical Methods in Medicine. | Article ID 9074759 | https://doi.org/10.1155/2017/9074759.

WebMD. (n.d.). *Seizures: What happens to your body*. WebMD. https://www.webmd.com/epilepsy/stages-seizures

Yoo G, Kim H, Hong S. Prediction of Cognitive Load from Electroencephalography Signals Using Long Short-Term Memory Network. Bioengineering (Basel). 2023 Mar 15;10(3):361. doi: 10.3390/bioengineering10030361. PMID: 36978752; PMCID: PMC10044910.