A Motor Activity Analysis as a Depression Indicator: Predictive Approach Using A Hybrid CNN and LSTM Network.

*Note: Sub-titles are not captured in Xplore and should not be used

Vinsensius Odilo Sutedja
Dept. Computer Science
School of Computer Science, Bina Nusantara University
Jakarta, Indonesia
vinsensius.sutedja@binus.ac.id

Frans Andreas
Dept. Computer Science
School of Computer Science, Bina Nusantara University
Jakarta, Indonesia
frans.andreas@binus.ac.id

Ghinaa Zain Nabiilah
Dept. Computer Science
School of Computer Science, Bina Nusantara University
Jakarta, Indonesia
ghinaa.nabiilah@binus.ac.id

Jurike V. Moniaga
Dept. Computer Science
School of Computer Science, Bina Nusantara University
Jakarta, Indonesia
jurike@binus.edu

***Abstract*— Depression is a prevalent mental disorder affecting behavior, cognition, and physical activity. Early detection and intervention are crucial to mitigating its severe impacts, including suicidal ideation or suicide attempts. This study proposes a novel predictive model using a hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) network to analyze motor activity data as an indicator of depression. Utilizing the "Depresjon" dataset, which comprises activity levels from both depressed and non-depressed individuals, the model extracts features through CNN and processes these with LSTM to identify depression patterns. To address class imbalance in the dataset, Synthetic Minority Over-sampling Technique (SMOTE) was employed, enhancing the model's predictive accuracy. The final model achieved a high accuracy rate of 97.89%, significantly outperforming the other model created. This research demonstrates the potential of combining CNN and LSTM for effective depression prediction, providing a valuable tool for early diagnosis and intervention.**

Keywords— Depression, Motor Activity, CNN, LSTM, Predictive Model, SMOTE, Machine Learning

I. INTRODUCTION

Depression is one of the mental disorders that will affect how a person behaves and acts. The mindset and decision-making process will also deteriorate if the condition is left untreated or even unnoticed. It is unfortunately one of the most common mental disorders, and most sufferers will often lose concentration and focus, but it is exacerbated by difficulty sleeping and lack of sleep and some situations may wake up late. The worst effect that can appear in someone who has depression is when they start having suicidal thoughts, just thinking about something like that is a sign of a problem, but it does not rule out the worst possibility of doing it.

There are also other things that could increase the risk of becoming depressed, one example is from Pearce's paper[1], he conducted research on the association between physical activity and depression, and the result shows that there is a recommended level of activity for humans that if you do an activity more than the recommended level it can increase the risk of getting depressed. Another example of something that could increase the risk of becoming depressed is Han [2], he wrote a paper based on his research about the association between sleep time variance during weekends and weekdays with risk of depression. In the paper, it is discussed that if you have more sleep time on weekends above 20% difference then the risk of depression increases, it also applies if you sleep less on the weekend. There are many things that could increase the risk of being depressed, and for that reason, everyone could be subjected to depression without knowing it.

The impact of depression itself varies from small to large, and according to the World Health Organization, there are about 280 million people suffering from depression and 23 million of them are children and adolescents. For them to experience depression is tantamount to risking their future, for a small impact they may lose the focus needed to learn and in the worst case the suicide rate for children and adolescents may increase. According to Orsolini et.al. [3] MDD or Major Depressive Disorder has a 4.7% prevalence number globally and in Western countries, some of them even reach up to 8% prevalence in MDD. Orsolini et.al. [3] also stated that there is a relation between suicide risk and MDD, which explains that the risk of suicide could reach 15% for people who have MDD. Therefore, depression itself must be treated as soon as possible or at least recognized so that people can get help.

In this case, motion sensors that are already commonly used in sports might be an option to be developed into a tool that can predict whether someone has depression or not if combined with artificial intelligence. It is known that one of the effects of depression is difficulty sleeping, therefore people who suffer from depression should have a level of activity at night that is quite relatively high compared to normal people. This activity level pattern will be the benchmark in assessing whether someone has depression or not using this motion sensor. One of the examples of how this experiment works can be seen in Pedrelli et.al. paper [4]. The paper explained that detecting depression at an early stage by using a motor activity sensor is possible and in fact, could provide useful information on the severity.

In this experiment itself, this activity pattern will be the dataset used, and the dataset is called "depresjon" where there are 23 activity level data obtained from people who have depression and 32 activity level data obtained from people who do not have depression. This dataset has also been used several times, one of which is by Garcia-ceja et al.[5] where he first used this dataset to classify people with depression into unipolar and bipolar groups using the Support Vector Machine (SVM) method and got an accuracy of 0.72 although this paper is fairly old this paper is a pioneer for similar papers that use this dataset. The paper aims to provide an open dataset to understand the relationship between a person's motor activity level and depression itself. In the process of understanding the relationship, Garcia-ceja et al. created an AI model but the drawback written by them is that the amount of data is not balanced and it makes the classification process more likely to choose more data. Therefore, the method that will be used in this paper is to utilize both CNN and LSTM combining it into a hybrid method, CNN is going to be used to extract features from activity level data in the form of time series automatically then the features will be an input for the LSTM to learn. Then to solve the problem of the unbalanced amount of data, oversampling using SMOTE will be used.

II. Literature Review

In this 21st century, technology has developed very rapidly, but of all the technologies that exist so far there is only one purpose why technology continues to develop, namely to improve the quality of human life. Of all the fields that continue to develop every day, the field of artificial intelligence is one that is developing very rapidly. Starting in 1943 is the year when at least neural networks are a concept that has been studied and in 2021 neural networks can be used to simulate how the human brain works. Many AI models can be developed using this neural network, one of which is to detect depression or even classify it. There are many papers that discuss this topic with different datasets and different methods and in this section, we will discuss the results of some of these papers.

In the paper Jakobsen et al [6] where he created an AI model to compare the motor activity of patients with depression such as unipolar and bipolar with healthy patients. The experiment was carried out using a neural network and the SMOTE class balancing method, the machine learning approach also varied to get the most optimal results. From the results themselves, the most optimal is to use a deep neural network combined with the SMOTE class balancing method. The accuracy result is also fairly high at 0.84. This can be achieved because the amount of unbalanced data is made to be balanced first. This is evidenced by the accuracy result which is only 0.78 with a true positive number of 9, which means that many cases are misclassified due to this unbalanced amount of data.

Another paper by Julieta G. Rodriguez-Ruiz et al [7]. tried to compare day, night, and 24-hour motor activity in depressed and healthy patients. Their AI model will process the data through five stages namely, pre-processing, features extraction, features selection, classification, and validation. In pre-processing, their dataset will first be balanced in number and broken down into one-hour segmentations. From this segmentation, day data, night data, and daily data will be compiled. After cleaning the missing data the features extraction stage will be carried out using the fast fourier transform and continued with the features selection stage at this stage the data will be reduced in dimensions and then the classification stage is carried out by entering the data in the depressive and non-depressive groups and finally the validation stage where accuracy will be calculated. This paper uses a different approach because they extract the features from the dataset manually and create a model with the selected features. The accuracy obtained is also very high because it is at 0.98 and 0.99 where the best results are obtained from using the best 9 features from the night.

Another paper that produces high accuracy is the paper of M. Zakariah et al [8]. where their AI model combines neural network methods and UMAP as a dimension reduction. They did two approaches where one they will do the training directly after the data processing is done and the second approach they will extract the features and reduce the dimensions by using UMAP and then test and validation will be done. In the first approach where the features are not extracted, the accuracy value is 0.6340 and in the second approach where UMAP and neural network are used, the accuracy value is 0.991 which is a big increase in terms of accuracy. In this paper itself, it is written that the unbalanced number of datasets is a problem and the lack of further monitoring of patients is also a shortcoming in understanding the relationship between motor activity and depression.

Now we will discuss the use of hybrid models in other papers, an example of hybrid use is in Tsokov's paper [9], in his paper Tsokov tries to discuss the problem of air pollution, and methods to estimate it. In his paper he uses a hybrid model that combines 2D CNN and LSTM, the reason is 2D CNN will be used in capturing spatial values and LSTM will be used to process temporal values, making this model suitable for processing spatiotemporal data. The results obtained are also fairly good and consistent in predicting air pollution levels in the surrounding area, but these results are further strengthened by utilizing the k-best trained model.

Another paper that uses a hybrid model is Salur's paper [10], in his paper, he tries to classify sentiment using a hybrid model that combines CNN and LSTM. This paper is one of the references for this paper in how to utilize hybrid models because Salur utilizes CNN in extracting features and processing sequence data by LSTM. The results obtained also when compared with existing models show better results in terms of accuracy, precision, recall, and f1.

There is also a paper that uses this hybrid model to predict the morphological evolution of the beach nourishment project in the Netherlands, namely Li Y's paper [11]. This model combines CNN and LSTM also in monitoring data to predict changes in terrain effectively. This paper uses the same approach as Tsokov [9], where CNN is used to capture spatial values and LSTM is used to process data temporally. The results also show that this model can accurately predict the data of all the models used, which means that this model is an improvement of the existing models.

There is also a paper that uses this hybrid model for smart grid planning [12], it is used by forecasting the short term electric load so that the smart grid can be operated efficiently.

Although many models have been used to predict this, the author proposes this method to improve the performance of existing ones.

There is also a paper from [13], this paper tries to predict PM2.5 air pollutant which is the most impactful air pollutant for human health. The use of this model is to increase the accuracy of predicting PM2.5. The use is also still the same as how our model is used, which is by utilizing CNN in extracting features and LSTM in processing temporal data. The results obtained also state that this model is more accurate and efficient than other models.

In addition, there are also papers that utilize this hybrid model in the biological field [14]. This model is used to classify miRNA, it is said that this model is made to overcome the shortcomings of the previous model which still extracts features manually. After this model was trained and used, it showed better performance than the previous model, with an accuracy value of 0.943 and f1 of 0.925.

One more example from the use of CNN and LSTM for time series dataset can be seen in this paper [15]. This paper proposed a model for recognizing human activity using radar-based micro-Dopler signatures. The author decided to treat the radar spectrogram as a time series dataset, thus using CNN and LSTM to extract spatiotemporal data. The model itself shows high accuracy in recognizing human activity at 98.28% accuracy. This makes the model outperform the existing model since no other model pays attention to the temporal dependencies.

III. Methodology

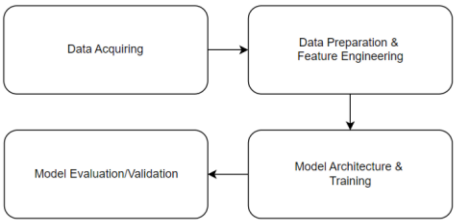

Fig 1. Methodology

In this section, the step-by-step of how the data is going to be processed will be explained, the step-by-step can be seen in Fig. 1 above it will start with obtaining the data, this dataset could be obtained from the Garcia-ceja paper [1]. The dataset itself will be processed first at the data preparation and feature engineering step, at this step the data will be normalized and split into sequences. After all that the model will be created and trained using the data that has been processed. Only after then the validation step begin to evaluate the model performance with accuracy as the metric.

A. Dataset

The dataset is in the form of a time series and each class is composed identically, it has three columns, one for the date of when the data was taken down to the minutes. The second column is the hour when data was taken. Since the data was obtained from a motion sensor that was attached to the patients it was programmed to take data at a certain interval which is every one minute. This data will be processed at the next step to prepare it as the input for the hybrid model of CNNtoLSTM.

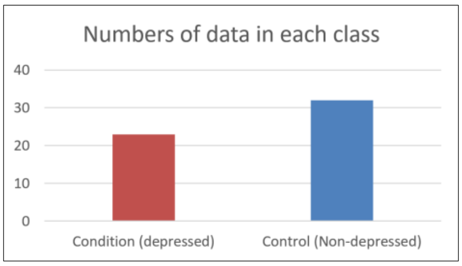

Fig 2. Comparison of the numbers of data in each class

B. Data Preparation and Feature Engineering

At this stage, the data will be prepared first by turning the data into a data frame that would fit the model later, this model itself would turn the normal sequence into a processed sequence by taking the hour part of the timestamp. So the processed sequence would consist of the hour when the data was taken and the level of activity at that time. After this step the sequence would be normalized by turning the original activity value into a value between 0 and 1, this is required since the neural network works best with this fuzzy value. Lastly, after getting the fuzzy value the sequence would be padded to make the missing value or turn it into a more uniform. At this point, we could already input the data frame into the model by creating the model first, but to avoid overfiting the model the data need to be oversampled first due to its unbalanced numbers of data between the two classes. In order to oversample it SMOTE also known as the Synthetic Minority Oversampling Technique will be utilized in the model, the simple reason behind this is that this technique focuses on increasing the number of data from the minority class to balance the overall number of data. The real reason for using SMOTE is that there is a paper dedicated to explaining SMOTE's benefit through small experiment [16].

The paper from Elreedy et.al. [16] explain about SMOTE in the opening of the paper, it is stated that to overcome the shortage of data SMOTE is one of the dominant methods that achieved great results in the literature. The way SMOTE generates new data is simply explained by generating new data from a line connecting a point with its K-nearest neighbor. The result achieved by SMOTE based on the paper also shows that SMOTE increased the accuracy of classification, but it varied from each other depending on the classifier used.

Another paper by Li et.al. [17] even make a better variant of SMOTE called NaNSMOTE, the difference between SMOTE and NaNSMOTE is that NaNSMOTE is more adaptive than original SMOTE in data selection and generation. This showed in the result that NaNSMOTE outperformed SMOTE in terms of data classification. So from similar paper suggestion to handle the Depresjon dataset with oversampling and the positive result from utilizing SMOTE, it became the real reason for choosing this method.

C. Model Architecture and Training

Since the data has been processed and prepared at this point, the model that would be used to classify the data can start to be created. The model itself would utilize both CNN and LSTM, the CNN would be used to automatically extract the features from the data, and then after that LSTM will learn the features from what the CNN has extracted. LSTM would study the pattern in order to identify and classify the data. This is the general flow of the model. The other detail of the model would consist of dimensional reduction by max-pooling layer to make it easier to be processed in the next layer, then there is also a dropout layer to avoid overfitting and next is flatten layer that would reduce the output from the previous layer into one dimension so that the next layer which is the dense layer with sigmoid activation could receive the output and classify it into the predicted class. All of the layers stated are what make the model and the model itself would be trained for 50 epochs before moving to the next step. For this experiment there would be three variants of the model will be created, the first is the very basic model, which has no oversampling or validation method, just pure data processing and inputting it into the model to be trained and evaluated. Then the second would be a model with an oversampling method, in this case, SMOTE will be used. Lastly model with oversampling method and validation method. Those three model would be created as a progress in order to make sure that the last model would achived the best result.

D. Model Validation

Validation is the step in which the model is evaluated, a good model would be able to generalize new unseen data that would be the input in the validation step. If the model could perform like when it is trained then it would be considered good. There are many validation techniques, but in this paper, only one will be used to evaluate the model, and that is the K-fold cross-validation technique. For K-fold cross-validation, it basically divided the dataset into k-subsets in which the model would be trained k times and evaluated using the k-number of subsets. Then the mean of the score would be the final score of the metric used which in this case is accuracy. That also becomes a reason why the K-fold Validation technique is applied, based on Asrol et.al. paper [18], the accuracy could be improved by using K-fold Validation, it is shown in the result of their experiment, when the accuracy gets boosted by using the K-fold technique. The same result can also be seen in Ling et.al. paper [19]. The model proposed in the paper also shows better accuracy results. This becomes a reason to apply k-fold cross-validation into the model for increasing accuracy in predicting depression.

IV. Result and Discussion

In this section, the result of the model will be shown and discussed. Previously it was said that there would be two models that would be made in order to make the final model. So in total, there would be three model created to be discussed in this section, but that does not include the modified version of the models. If there is some trouble encountered on spesific model there might be another version created before finalizing the model evaluation. The first model is the very basic model, it will just load the data, preprocess it, and input it into the model. After inputting the preprocessed data into the model, the evaluated result obviously still lacking at some part. In this case the accuracy still need to be improved, since it only achieved 0.7277 for this model. These accuracies are shown in form of a graph in Fig. 3 below.

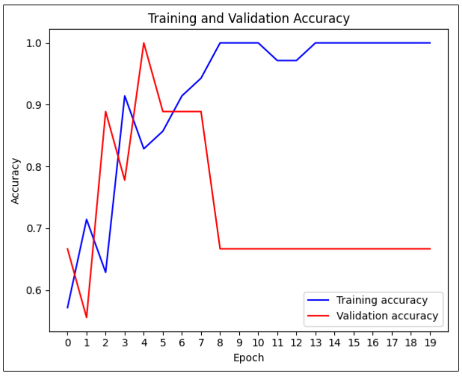

Fig. 3 Basic Model Training and Validation Accuracy

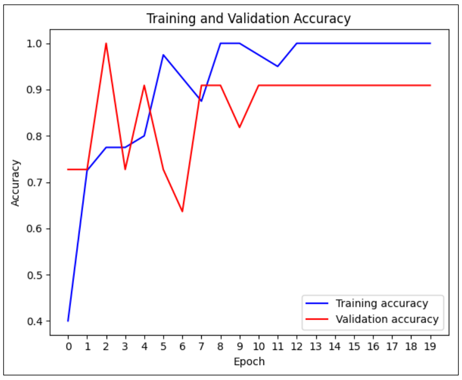

Fig. 4 Model with SMOTE Training and Validation Accuracy

The first model is very basic and easy to create so there is no problem encountered when creating or testing it, but there is some improvement needed from the result of the Basic Model. Then there is the second model. After learning that the first model has low accuracy the second model tried to improve it using the oversampling method SMOTE, discussed by Frogner J. [20] it increased the accuracy and performance of one of the models he created, there are also a few more cases for this dataset, where when SMOTE is used to balance the number of data, it ended with an improved result from the model than the one that trained without it. After the results come out after training, it is shown that the accuracy indeed increases from 0.7277 to 0.8462 this implies that imbalanced numbers of data in fact have an impact on how good the model is. The graph of accuracies for the second model can be seen in Fig. 4 above.

After implementing the oversampling method, there is only one thing left to implement, and that is the validation. For the validation method k-fold cross-validation is used, this method splits the data into k numbers of subsets that will be used to validate the model for k numbers of loop iteration. Other than k-fold cross-validation the final model itself has some additional technique and that is a regulation technique, it is for the model to prevent overfitting, and that is the training will end earlier from the epoch if the metric it monitored doesn't improve after five epochs in a row, in this case, the metric it monitored is validation loss, the reason other than from previous mistake is that validation loss is the metric to measure the model capability to work on the data it hasn't see during training. So due to this, the model that is considered best for testing is the one that has the lowest validation loss. By implementing these methods the model should be ready to be used for predicting depression, but the result that came out says otherwise. The accuracy of the model can be seen in Table I below.

TABLE I. ACCURACY OF FINAL MODEL AFTER APPLYING K-FOLD VALIDATION

Fold	*Accuracy*
1	0.7692
2	0.6154
3	0.6154
4	0.5385
5	0.7500
Mean	**0.6577**

TABLE II. ACCURACY OF FINAL MODEL AFTER FIXING AUGMENTATION DATA

Fold	*Accuracy*
1	0.7692
2	0.9231
3	1.0000
4	1.0000
5	0.9211
Mean	**0.9227**

TABLE III. ACCURACY OF FINAL MODEL AFTER OUTLIER DATA VALIDATION

Fold	*Accuracy*
1	1.0000
2	1.0000
3	1.0000
4	1.0000
5	0.8947
Mean	**0.9789**

The result of the FinalModel that has k-fold cross-validation applied shown in Table I, the final model not only added validation method but also an augmentation method for adding variant to the data, it is needed to prevent overfitting since the model will train at maximum 5 times than the previous model, but from the result above it has some problem that causes the accuracy to drop from 0.8462 to 0.6577, and it turns out that the data augmentation data has a logic problem, instead of adding variation to the data, it turns the normal data into a variant by adding noise and use the added noise data as the data that the model will learn from. This makes the model unable to learn the true feature from the data and hence the accuracy drops.

This problem is solved in the next model that and the accuracy shows improvement, from 0.8462 that gotten from SMOTE to 0.9227 that gotten from this model. It is shows that the final model is already performs well in predicting depression from the motor activity dataset, but there is some final validation suggested by adding validation for outliers data, this is perform by finding data that has extreme value and replacing it with median value, the reason this validation is added is to make the dataset more robust.

TABLE IV. EVALUATION COMPARISON BETWEEN MODELS

Evaluation of All the Models				
Models	***Accuracy***	***Precision***	***Recall***	***F1***
Basic Model	0.7277	0.7500	0.6000	0.6667
Model with SMOTE	0.8462	0.8571	0.8571	0.8571
Final Model	**0.9789**	**0.8571**	**1.0000**	**0.9231**

The result of the final model can be seen in Table II above it has a high accuracy in predicting depression from patients' motor activity data provided in the dataset, but just to make the dataset more robust outlier data validation is applied to the final model along with regulation technique to prevent overfitting. The final model was trained again and the result can be seen in Table III above. The final result shows that the the accuracy of the model improve both after fixing the augmatation problem and after adding outlier validation, it went from 0.8462 by using SMOTE, to 0.9227 just after adding variants to the data, and lastly 0.9789 after applying outlier validation. This concluded that the final model which was created using CNN and LSTM hybrid model and trained using augmented, oversampled, and validated data, can predict depression with relatively high accuracy at since the Basic Model and Model with SMOTE did not even reach 90 percent accuracy, other than accuracy the Final Model also exceed from all other aspect compared to the other two model. The comparison of accuracy, precision, recall, and f1 among the models itself could be seen in Table IV.

V. CONCLUSION

After training four models the result has been concluded that the model can be used to predict whether someone is depressed or not at a relatively early stage by analyzing their motor activity data with 97 percent confidence. Still, there are a lot of challenges throughout the process, first, the data itself is unbalanced, and that for people who would like to use this dataset have to make sure the data needs to be oversample or balanced before it is used for training, in order to avoid bias to the majority class. Then there is also an overfitting problem and low accuracy, this is solved in this paper by augmenting the data before oversampling it. This is done in order to add variation to the training data so that it can learn different patterns from the same data. Lastly in order to make the dataset more robust Outlier Validation is applied. Overall after solving all the problems one by one, the model has been improved to reached a good accuracy of 0.9789, and for futre works this model itself could still be improved by utilizing other classifiers or methods for oversampling and validation, with those the result achieved might be better than the one presented in this paper.

REFERENCES

[1] M. Pearce *et al.*, "Association between Physical Activity and Risk of Depression: A Systematic Review and Meta-analysis," *JAMA Psychiatry*, vol. 79, no. 6. American Medical Association, pp. 550–559, Jun. 01, 2022. doi: 10.1001/jamapsychiatry.2022.0609.

[2] K. T. Han and S. J. Kim, "Instability in daily life and depression: The impact of sleep variance between weekday and weekend in South Korean workers," *Health Soc Care Community*, vol. 28, no. 3, pp. 874–882, May 2020, doi: 10.1111/hsc.12918.

[3] L. Orsolini *et al.*, "Understanding the complex of suicide in depression: From research to clinics," *Psychiatry Investigation*, vol. 17, no. 3. Korean Neuropsychiatric Association, pp. 207–221, Mar. 01, 2020. doi: 10.30773/pi.2019.0171.

[4] P. Pedrelli *et al.*, "Monitoring Changes in Depression Severity Using Wearable and Mobile Sensors," *Front Psychiatry*, vol. 11, Dec. 2020, doi: 10.3389/fpsyt.2020.584711.

[5] E. Garcia-Ceja *et al.*, "Depresjon: A motor activity database of depression episodes in unipolar and bipolar patients," in *Proceedings of the 9th ACM Multimedia Systems Conference, MMSys 2018*, Association for Computing Machinery, Inc, Jun. 2018, pp. 472–477. doi: 10.1145/3204949.3208125.

[6] P. Jakobsen *et al.*, "Applying machine learning in motor activity time series of depressed bipolar and unipolar patients compared to healthy controls," *PLoS One*, vol. 15, no. 8, p. e0231995, Aug. 2020, doi: 10.1371/journal.pone.0231995.

[7] J. G. Rodríguez-Ruiz *et al.*, "Comparison of Night, Day and 24 h Motor Activity Data for the Classification of Depressive Episodes," *Diagnostics*, vol. 10, no. 3, p. 162, Mar. 2020, doi: 10.3390/diagnostics10030162.

[8] M. Zakariah and Y. A. Alotaibi, "Unipolar and Bipolar Depression Detection and Classification Based on Actigraphic Registration of Motor Activity Using Machine Learning and Uniform Manifold Approximation and Projection Methods," *Diagnostics*, vol. 13, no. 14, p. 2323, Jul. 2023, doi: 10.3390/diagnostics13142323.

[9] S. Tsokov, M. Lazarova, and A. Aleksieva-Petrova, "A Hybrid Spatiotemporal Deep Model Based on CNN and LSTM for Air Pollution Prediction," *Sustainability (Switzerland)*, vol. 14, no. 9, May 2022, doi: 10.3390/su14095104.

[10] M. U. Salur and I. Aydin, "A Novel Hybrid Deep Learning Model for Sentiment Classification," *IEEE Access*, vol. 8, pp. 58080–58093, 2020, doi: 10.1109/ACCESS.2020.2982538.

[11] Y. Li, P. Van Oosterom, Y. Ge, X. Zhang, and F. Baart, "A CNN LSTM method for the morphology evolution prediction of beach mega-nourishment," *IEEE Access*, vol. 8, pp. 184512–184523, 2020, doi: 10.1109/ACCESS.2020.3030119.

[12] M. Alhussein, K. Aurangzeb, and S. I. Haider, "Hybrid CNN-LSTM Model for Short-Term Individual Household Load Forecasting," *IEEE Access*, vol. 8, pp. 180544–180557, 2020, doi: 10.1109/ACCESS.2020.3028281.

[13] T. Li, M. Hua, and X. Wu, "A Hybrid CNN-LSTM Model for Forecasting Particulate Matter (PM2.5)," *IEEE Access*, vol. 8, pp. 26933–26940, 2020, doi: 10.1109/ACCESS.2020.2971348.

[14] A. Tasdelen and B. Sen, "A hybrid CNN-LSTM model for pre-miRNA classification," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-93656-0.

[15] J. Zhu, H. Chen, and W. Ye, "A Hybrid CNN-LSTM Network for the Classification of Human Activities Based on Micro-Doppler Radar," *IEEE Access*, vol. 8, pp. 24713–24720, 2020, doi: 10.1109/ACCESS.2020.2971064.

[16] D. Elreedy and A. F. Atiya, "A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance," *Inf Sci (N Y)*, vol. 505, pp. 32–64, Dec. 2019, doi: 10.1016/j.ins.2019.07.070.

[17] J. Li, Q. Zhu, Q. Wu, and Z. Fan, "A novel oversampling technique for class-imbalanced learning based on SMOTE and natural neighbors," *Inf Sci (N Y)*, vol. 565, pp. 438–455, Jul. 2021, doi: 10.1016/j.ins.2021.03.041.

[18] M. Asrol, P. Papilo, and F. E. Gunawan, "Support Vector Machine with K-fold Validation to Improve the Industry's Sustainability Performance Classification," in *Procedia Computer Science*, Elsevier B.V., 2021, pp. 854–862. doi: 10.1016/j.procs.2021.01.074.

[19] H. Ling, C. Qian, W. Kang, C. Liang, and H. Chen, "Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment," *Constr Build Mater*, vol. 206, pp. 355–363, May 2019, doi: 10.1016/j.conbuildmat.2019.02.071.

[20] J. I. Frogner *et al.*, "One-dimensional convolutional neural networks on motor activity measurements in detection of depression," in *HealthMedia 2019 - Proceedings of the 4th International Workshop on Multimedia for Personal Health and Health Care, co-located with MM 2019*, Association for Computing Machinery, Inc, Oct. 2019, pp. 9–15. doi: 10.1145/3347444.3356238.