TASK Quarterly 26 (4) 2022

Comparative analysis of machine learning algorithms based on an air pollution prediction model

Aneta Wiktorzak
Lomza State University of Applied Sciences
1 Akademicka St., 18-400 Lomza, Poland
Date: 08 January 2024

Bartosz Kaczorowski
Lomza State University of Applied Sciences
1 Akademicka St., 18-400 Lomza, Poland
Date: 08 January 2024

https://doi.org/10.34808/dp3a-5n10

Abstract

In this paper it has been assumed that the use of artificial intelligence algorithms to predict the level of air quality gives good results. Our goal was to perform a comparative analysis of machine learning algorithms based on an air pollution prediction model. By repeatedly performing tests on a number of models, it was possible to establish both the positive and negative influence of the parameters on the result generated by the ANN model. The research was based on some selected both current and historical data of the air pollution concentration altitude and weather data. The research was carried out with the help of the Python 3 programming language, along with the necessary libraries such as TensorFlow and Jupyter Notebook. The analysis of the results showed that the optimal solution was to use the Long Stort Term Memory LSTM algorithm in smog prediction. It is a recursive model of an artificial neural network that is ideally suited for prediction tasks.

Further research on the models may develop in various directions, ranging from increasing the number of trials which would be linked to more reliable data, ending with increasing the number of types of algorithms studied. Developing the models by testing other types of activation and optimization functions would also be able to improve the understanding of how they affect the data presented. A very interesting developmental task may be to focus on a self-learning artificial intelligence algorithm, so that the algorithm can learn on a regular basis, not only on historical data. These studies would contribute significantly to the amount of data collected, its analysis and prediction quality in the future.

Keywords:

ANN, Python, LSTM

1. Introduction

Along with the constantly developing urbanization rate, the problem of smog becomes an increasingly important issue, which leads to negative changes in the natural environment. Air pollution monitoring has become a very important and complex challenge. By analyzing the degree of air pollution, it is possible to determine the trend lines towards which the pollutant is heading before it occurs. This allows for the introduction of effective countermeasures that will reduce the pollution generated in the natural environment.

In this paper it has been assumed that the use of artificial intelligence algorithms to predict the level of air quality gives good results. The aim of the research was to conduct a comparative analysis of the application of machine learning algorithms to predict the occurrence of smog. The research was based on some selected both current and historical data of the air pollution concentration altitude and weather data. The task was to develop an artificial neural network model to predict the state of air pollution and to calculate the average level of PM10 in the atmospheric air 24 hours in advance, while using the 24-hour range of historical concentration values.

2. Air pollution monitoring

Monitoring the state of air pollution, especially in countries with a high urbanization rate and the number of cars per person, is very significant. The most common, and at the same time, the most frequently researched area when it comes to the impact, effects and causes of air pollution, is the phenomenon of smog [1] which is created by combining fog with smoke and exhaust fumes, produced, inter alia, by cars and factories. It is noteworthy that location is a big factor in the process of smog formation. However, it is not always that smog is caused mainly by human activity, it is also possible to appear in the natural environment, but these are marginal cases in relation to the overall situation. Nevertheless, areas with a high level of population density, factory districts and the vicinity of various types of mines are the most vulnerable [2]. In Poland, it is visible, inter alia, in Silesia, an area densely occupied by coal mines and heavy industry, where the effects of suspended dust are most visible, ranging from the ubiquitous dust to an increased incidence of respiratory diseases [3]. It may be stated that this type of air pollution is directly responsible for the major climatic and atmospheric changes, both those felt through human senses and imperceptible, albeit fatal while in long exposure. It has been observed that irreversible changes, extinction of species and mass threats to the human life and health may occur with longer exposure of the natural environment to the impact of smog.

There is no one universal method to fight and protect oneself against the negative effects of smog, nevertheless there are some preventive measures such as filter masks or air purifiers that prove helpful. It is also recommended that the air quality should be monitored with the use of more and less specialized devices. For example, thanks to various types of smog systems and meters, an ordinary citizen is able to check the level of air pollution and decide for himself/herself whether it is worth going for an evening walk, or if he/she should resign from the activity due to the bad quality of air. This type of awareness is one of the best possible ways to protect oneself from excessive smog exposure and it is highly recommended in large urban agglomerations, where such threats are usually much greater than in suburban areas or villages [4].

Nowadays, smog sensors are widely available. That means that they can be easily installed on private properties without incurring high costs. In this way, in turn, the air quality in the neighborhood can be checked and monitored on a regular basis, via the use of popular websites or applications. It is believed that frequent air checks contribute to making the public more aware of this problem. Increasing people's awareness of the importance of air quality monitoring and of the effects which smog has on human life and health, may lead to their taking appropriate pro-ecological decisions on a micro and macro scale.

To meet these needs, many companies and institutions have decided to collect air pollution data for research purposes and also promote pro-ecological activities. The activities of several air quality monitoring companies in Poland are described below:

- Airly, the company has air pollution sensors around the world, committed to supporting the community by providing up-to-date and reliable data. Thanks to the available APIs, it is possible to view and use data in various programs [5];
- The Chief Inspectorate for Environmental Protection deals with many aspects of the ecosystem, one of which is the monitoring of air pollution levels in Poland. The organization also provides free APIs [6];
- Syngeos, a company that monitors air pollution in Europe, is characterized by professionalism and ease of use of the application and the website. The data provided by this company is used, inter alia, by schools and public institutions [7];
- Eko Patrol GIG, a company whose goal is to build a dense measurement network that provides systemic support for anti-smog activities, useful both for local authorities and residents. The program consists of stationary and portable monitoring devices, mobile measurement laboratories and individual dust meters [8];

- InConTech, the company focuses on the production of universal software and hardware devices that allow building the entire infrastructure of the Internet of Things [9].

Monitoring the state of air pollution is undoubtedly a very complex problem faced by companies and institutions, private or state-owned, trying to study atmospheric phenomena. In addition to data such as air pollution levels, smog sensors can also collect a range of weather and atmospheric data such as temperature, humidity, wind speed and direction.

3. Prediction of air pollution

It is commonly accepted that continuous monitoring is required to better understand the changes taking place in the natural environment. By analyzing the levels of suspended pollen, it is possible to determine the trend lines to which a given pollution is heading before it occurs, it allows for the introduction of effective countermeasures that will lower the concentration, and thus, reduce the pollution generated in the natural environment.

In the prediction of air pollution it is important to describe the appropriate features assigned to the measurement data. The measured values must be closely related to each other, so that the algorithm used could search for dependencies and thus be able to make predictions with a good degree of accuracy. Unrelated or logically incorrect features may lead to a disturbance of the results of the algorithm and adversely affect other values, which may be considered less important inside the algorithm. Fig. 1 shows the dependence of PM10 pollution on several features in the period from August 2021 to February 2022 in Lomza, Poland.

Therefore, an important element of preparation for air pollution measurements is the selection of appropriate attributes, such as temperature, pressure, humidity, wind direction and speed. The temperature allows introducing certain seasonality into the data. The wind data is also significant for the operation of the prediction algorithm, at higher wind speeds the air pollution will be correspondingly lower due to the "blowing off" of solid particles and other substances in the air, their constant movement causes the sensors not to be able to read the data, but also the human body is not able to absorb such amounts as there would be in stagnant air. As a parameter, humidity accurately indicates the relationship between, inter alia, rainfall and harmful substances in the air, after the rainfall, the air is much cleaner due to the dissolution of pollutants in the precipitation. Dry air promotes the absorption of solid particles, the pharyngeal mucosa begins to dry out, which can lead to irritation. All the above-described weather factors, such as temperature, wind, humidity, have a very large impact on the accuracy of the air pollution forecast.

Other compounds in the air, as well as the feature to be predicted, are relevant to the operation of the algorithm. Increasing the measured values may result in the fact that the concentration of particulate matter in the future may be increased, which is an important argument for a machine learning model that can discover this relationship and update its forecasts by analyzing it. The location of the sensor itself can also be a significant prediction feature, i.e. sensors in densely populated areas may show different readings than those found outside the city. Furthermore, the exact location of a sensor on a structure, or a building, must be considered. For example, the height at which the sensor is placed, or the distance between particular sensors, may influence the collection of data. Therefore, for the prediction model to be trained and then implemented, one must consider various factors which influence the air quality in a particular region.

4. Aritificial intelligence and prediction of air pollution

Artificial intelligence (AI) can be defined as the intelligence displayed by devices, as opposed to natural intelligence. It is used to create models and programs that apply, at least in part, behaviors that could be defined as intelligent [10].

One of the areas closely related to artificial intelligence is Machine Learning (ML). ML deals with algorithms devoted to self-learning, i.e. those that improve automatically through the experience they have gained during operation and exposure to data. These algorithms build a mathematical model on the basis of the entered data, called training [11]. The advantage of such algorithms is that they are not directly programmed by a human, and the model itself selects appropriate weights for individual input data in such a way that they are as result-dependent as possible, which is expected especially in decision-making or forecasting tasks. They are used in places where conventional algorithms might turn out to be impossible or very complicated.

Machine Learning is the next step in the development of artificial intelligence, its practical application, used in modern technologies, which are increasingly based on it [11]. These algorithms are to ensure increased effectiveness, efficiency, cost reduction and uptime. In order for such assumptions to be realistic, the model must be very well learned and repeatedly tested so that it provides the most accurate data. Evaluating whether an ML model is well-trained can depend on the specific

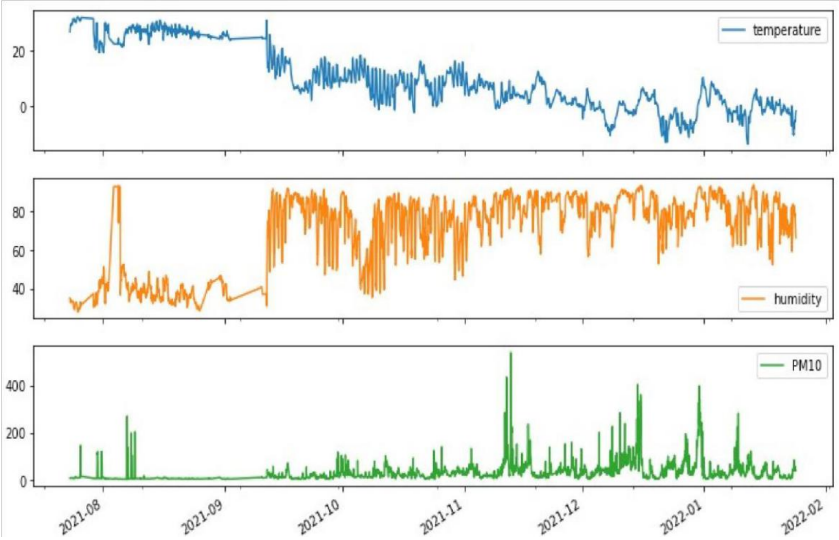

Figure 1: Features dependent on PM10. Source: Own elaboration based on collected data from InConTech sensors in Lomza

type of task and analyzed data. There are various error metrics that can be used for this purpose, and the choice depends on the context. Here are several popular error metrics and their values that can be applied in different cases:

- Mean Squared Error (MSE): MSE measures the average square of the difference between model predictions and actual values. Lower MSE values indicate better performance. A well-trained model will have a low MSE on test data.
- Root Mean Squared Error (RMSE): RMSE is the square root of MSE and provides an error measure in the same units as the data. A lower RMSE indicates better model quality.
- Mean Absolute Error (MAE): MAE measures the average absolute difference between predictions and actual data. Lower MAE values indicate more precise predictions.
- Coefficient of Determination (R-squared): R-squared evaluates how much of the data variability is explained by the model. A value close to 1 indicates a good predictive ability.
- Accuracy: In classification problems, accuracy measures the percentage of correct model classifications. A higher accuracy indicates a better model.

It is worth noting that what constitutes a "good" error value can vary depending on the context and task requirements. Therefore, to assess whether a model is well-trained, it is valuable to compare its results with those of other models or establish specific threshold values that are acceptable in a given context. ML can also be used in forecasting tasks, after providing properly formatted input data it is possible to quite accurately reflect the reality for a selected number of steps into the future. Prediction algorithms are specialized pieces of a program code that are used to make predictions of single values or their groups by analyzing the input data [11]. They are very useful for large data sets, so-called "big data", which a human would not be able to analyze in a timely manner. Systems developed for meteorological or atmospheric phenomena, such as predicting air pollution, are very popular in the public due to their impact on the natural environment.

The most popular form of creating prediction algorithms is machine learning employing deep learning (DL) models. These methods allow, in a much simpler way, generating functions whose task is to predict values. In these methods, the algorithms are created by themselves, the weights of specific data fed to the inputs are determined by the algorithm automatically on the basis of the resulting outputs. The calculation of the weights itself is extremely difficult, but in the final step it can turn out to be very accurate [11].

Unlike classical machine learning algorithms, deep learning has the ability to use more data resources and is not so limited in terms of learning opportunities. Deep Learning (DL) is a subgroup of machine learning based on Artificial Neural Networks (ANN) [12]. The name comes from the layout of the network structure, it consists of many input, output and hidden layers. The layer units transform the input information so that further layers are able to perform a prediction task. One of the more important aspects is that Deep Learning does not require data preprocessing, as it does in machine learning, these algorithms are able to independently extract interesting unstructured data and create functions based on it. DL is, in its assumptions, much closer to the human mind than machine learning, but it is associated with a greater demand for computing power. Since people began to notice the benefits of using Deep Learning, it has appeared in many industries, be it programming, medical, agricultural

and financial. The most popular deep learning algorithms and those that usually bring the best results in the time planned for their learning are [13]:

- Long Short Term Memory (LSTM);
- Convolutional Neural Networks (CNN);
- Recurrent Neural Networks (RNN).

5. Experimental research

Coming back to the thesis that was put forward in the Introduction, DL algorithms were analyzed for their use in the smog prediction task. The method of training and the parameters of the model itself are key to finding the right correlations between the input and output data. During the research, many models were built and learned, which were then checked on test and validation data. The course of each test was recorded, and the collected data was analyzed in terms of correctness of operation and the dependence of individual parameters on each other.

The analysis of the results has shown that the optimal solution is to use the Long Short Term Memory LSTM algorithm in smog prediction, it is a recursive model of an artificial neural network that is ideally suited for prediction tasks. Unlike other models, it has feedback loops that enable better communication between individual layers and neurons. An individual LSTM called a unit consists of a cell, an input gate, an output gate, and a forget gate. The cell remembers the values at any time interval, and three gates regulate the flow of information to and from the cell. Due to the fact of being recursive, the function is very well suited for the tasks of predicting time series and classification.

Since events can occur at different time intervals in time series data sets, the LSTM has been developed to deal with the vanishing gradient problem that occurs when training ordinary recursive neural networks [13]. The error of the disappearing gradient consists in the fact that the weights on the lines connecting individual neurons are constantly updated in the model learning phase, in the case of very low updating values, some kind of stagnation occurs and no changes are made, the model errs in anticipation of a larger excitation signal. At worst, this can stop the neural network from training any further. A visualization of a recursive neural network such as the LSTM is shown in Fig. 2.

When working with any algorithm of artificial intelligence and machine learning, it is important to enter the appropriate parameters by changing them, it is possible to easily notice the changes which they introduce. Such parameters include the number of epochs, the number of units, and the layer complexity.

As part of the analysis of the correctness of prediction of machine learning algorithms and artificial neural networks, appropriate metrics measuring errors and correctness, i.e. a unified system of assessments, should be used. It is very easy to estimate the results of the model prediction, even on greatly large data sets. Consideration should also be given to what metrics are used to evaluate models trained for specific activities. The metrics that evaluate the classification tasks are completely different from those used for prediction.

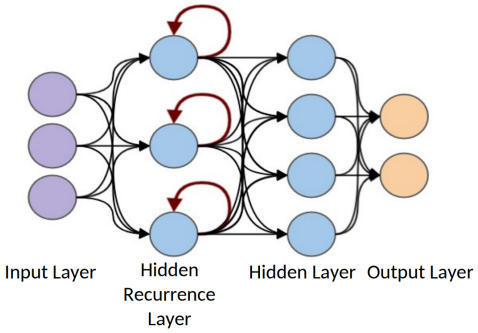

Figure 2: Recursive neural network Source: Own elaboration based on austingwalters.com

An example of this is the accuracy metric which tells about how often the result of the model is equal to the expected result. When working with classification, such a metric can bring very good results. The model classifies a given item to one of the strictly defined labels. In the case of predicting e.g. air pollution, such metric would probably almost always bring zero, due to the fact that it is very difficult for the model to indicate the ideal prediction result consistent with the expected one, along with the usual decimals.

During the research, the purpose of which was to choose the best machine learning algorithm to solve the task of predicting the state of air pollution, it was decided to carry out multiple attempts to train the neural network using the LSTM model. For this purpose, its parameters such as the number of units, the activation function, the optimizer were changed, but also the number of network layers was changed. Thanks to the use of various permutations of the size and types of neural network parameters, it was possible to determine the optimal characteristics of the LSTM model for smog prediction. The data on which the research was conducted was made available by InConTech, which has a number of sensors located in Lomza and its vicinity, which was a great convenience in working on a comparative analysis of machine learning algorithms based on the air pollution prediction model. Measurement data contain: concentrations of airborne substances and dusts, PM10, PM2.5, O3, SO2, and atmospheric data such as temperature, pressure, humidity, and wind speed and direction. This is the basic necessary data required to conduct air pollution monitoring tests, each measurement also includes the device identification number, the altitude

and latitude of the measurement point (sensor) and the exact date of the test. Examples of measurements stored in the database are shown in Fig. 3.

In the case of operations on parameters such as pressure, humidity, PM10, PM2.5, O3, SO2, the missing values were supplemented by assigning the average value from the entire set. By using this method, the possibility of generating data significantly changing the overall picture of the set is reduced, assigning an average value will not affect the newly determined aggregated value after the change operation.

6. Testing

The research process can be divided into three stages, with the first two already discussed, involving data collection and operations on it, and the last one focusing on modeling a real machine learning algorithm. This stage involves creating a function, adding appropriate layers with possible parameters to it, then compiling the model, taking into account the estimated metrics, and saving the entire history of the model's progress to a file. A total of 150 training iterations were conducted, divided into 30 models, with 5 trials for each parameter change within the predictive function. The cumulative training time for all models was nearly 16 hours, with an average training time of 7 and 1/2 minutes per process. The complexity of the model's layers and the number of units in each layer contributed the most to increasing this value. Another significant factor were the specifications of the computer on which the research was conducted.

During the experiments, it was observed that increasing the number of LSTM layers resulted in relatively better results compared to models created with a single-layer structure. This trend was noticed in most models, regardless of the selected parameters. However, increasing the number of layers in the model also came with a significant increase in the training time and computational complexity, which depended on the computational power of the computer used. More complex algorithms consumed a substantial amount of machine resources during the training, both due to the recursive nature of the layers and the memory footprint associated with the process. Table 1 presents the discussed data, and it can be observed that despite generally low-quality results, the best activation function turned out to be SIGMOID, achieving the best results for the R-squared metric.

Data analysis has revealed that the optimal solution is to use the ADAMAX optimizer. Experimental studies have shown that using the other two optimizers, ADAM and SGD, results in worse performance. When analyzing the training data, it can also be observed that the ADAM algorithm achieved an average score of around 50% correctly predicted values. The data collected during the training process is presented in Table 2.

The results obtained in the process of learning and testing the artificial neural network were analyzed in terms of correctness and dependence on the model parameters. During the training, a number of metrics were used to assess the results. These include MSE, RMSE, MAE and R-squared, all of which show the full picture of how the model works. In the Keras library, which is a part of TensorFlow, the metric of checking the R-squared value in specific learning phases is not explicitly declared. For this purpose an original version of the function was developed that allowed recording the value in question [13].

After analyzing the results of the correct operation of the models, it was found that model 7 achieved the best results and was characterized by the highest correctness of operation during the network learning process. Nevertheless, it should be noted that this model behaved unstably in the final stages. In such cases, a hybrid solution is very often used, i.e. for the prediction of results, a neural network is used in the initial phase of operation, and in the final phase, a stable, simple function is used to determine the objective function optimum.

The Model 7 parameters are:

- one LSTM input layer with 200 units;
- one DENSE output layer which is responsible for the output data format from the model;
- SIGMOID activation function;
- ADAMAX optimizer.

Graphs of the network training process in Model 7, with different metrics, can be seen below.

Model 7 for the R-squared metric, although it showed no stability in the final stages, returned results close to 70% correct, which is a very satisfactory result with a relatively small set on which the model was learned. The rest of the metrics have values close to zero, which is the result sought in these categories.

The image of the correctness of the model can be best observed by analyzing the predictions in relation to the measured values, Fig. 8 shows the discussed results for the best model number 7, one of the 24-hour windows that were generated by the model is presented. The 'x' markers stand for the predicted values, and the blue line with dots are the measured values. There is a large correlation between the individual values. The analysis of the results shows a good quality/accuracy of the predictions.

The conducted research has allowed identifying the best machine learning models that can be used in air quality forecasting. The models described by the selected parameters, changing in each study, generated very different

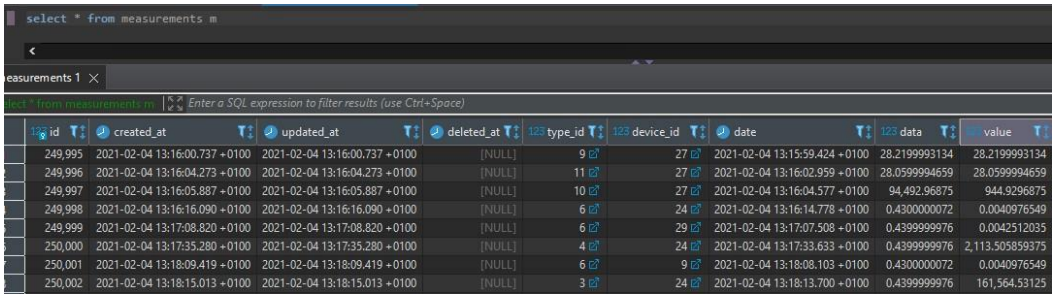

Figure 3: Screenshot of measurements stored in the database. Source: Own study based on the InConTech database

Table 1: Averaged evaluation results for single-layer models

Activation	val_perf				test_perf			
	MSE	RMSE	MAE	R^2	MSE	RMSE	MAE	R^2
SOFTMAX	3.29	1.80	1.06	0.08	0.99	0.99	0.65	0.04
RELU	203.48	8.26	4.67	-50.00	42.68	3.95	2.67	-40.28
SIGMOID	2.89	1.70	0.95	0.20	0.82	0.90	0.55	0.20
Optimizer	val_perf				test_perf			
	MSE	RMSE	MAE	R^2	MSE	RMSE	MAE	R^2
ADAM	2.87	1.69	1.00	0.20	0.94	0.97	0.63	0.09
SGD	3.98	1.97	1.13	-0.10	1.02	1.00	0.62	0.01
ADAMAX	3.07	1.75	0.94	0.15	0.78	0.88	0.55	0.25

Table 2: Averaged evaluation results for two-layer models

Activation	val_perf				test_perf			
	MSE	RMSE	MAE	R^2	MSE	RMSE	MAE	R^2
2xSOFTMAX	3.56	1.88	1.03	-0.02	1.04	1.00	0.64	0.08
2xSIGMOID	2.84	1.68	0.92	0.20	0.97	0.98	0.58	0.11
SIGMOID/SOFTMAX	2.81	1.68	0.82	0.19	1.11	1.05	0.58	0.18
SOFTMAX/SIGMOID	3.17	1.77	0.87	0.04	1.45	1.19	0.64	-0.14
Optimizer	val_perf				test_perf			
	MSE	RMSE	MAE	R^2	MSE	RMSE	MAE	R^2
ADAM	3.07	1.75	0.93	0.10	1.21	1.09	0.63	-0.05
SGD	3.46	1.84	0.93	-0.06	1.42	1.18	0.68	-0.05
ADAMAX	3.06	1.75	0.91	0.14	0.96	0.98	0.57	0.20

results. The best models were those with 200 units in the LSTM layer(s). Their excessive increase or reduction had a negative impact on the learning process, too many units, extended the learning time without any major changes in the quality of the model prediction, reducing the number of units resulted in a very large reduction of the learning time, but it significantly worsened the results. Optimization functions such as ADAM and ADAMAX brought the best results with reference to the tests carried out and described, it may clearly result from the fact that both algorithms are very similar to each other in terms of structure, and ADAMAX itself can be treated as an extension of the original algorithm. In the case of the tested data, the use of the SGD function significantly reduced the quality of prediction. This function did not work well in the tested model, with the data prepared in such a way.

In the conducted research, it was possible to observe a significant influence of the activation function on the result of the operation of the entire model. The SIGMOID function proved to be much better than the others and made it possible to obtain satisfactory results of the prediction of the state of air pollution. The best optimal model learned had this activation function. When the SIGMOID activation function was used, the model was stable, there were no anomalies, and the artificial neural network generated correct results. It can be concluded

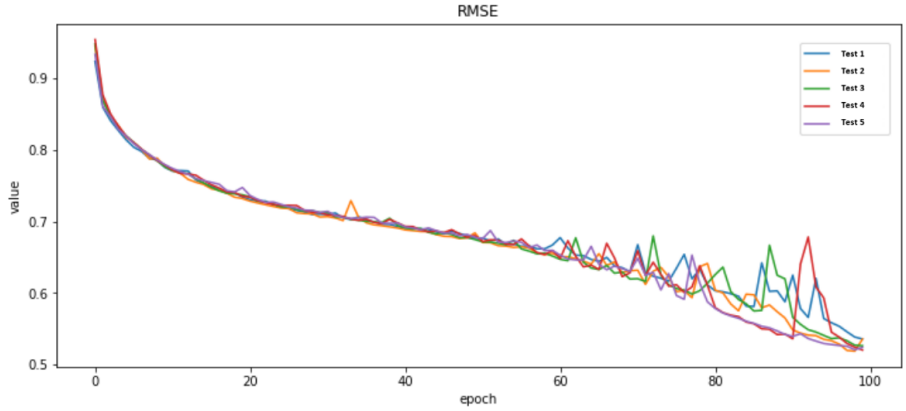

Figure 4: Model 7, RMSE metric Source: Own study based on collected data

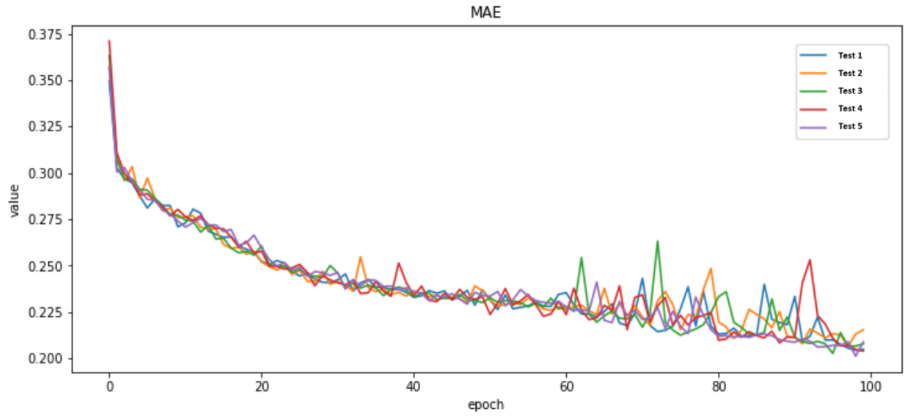

Figure 5: Model 7, MAE metric. Source: Own study based on collected data

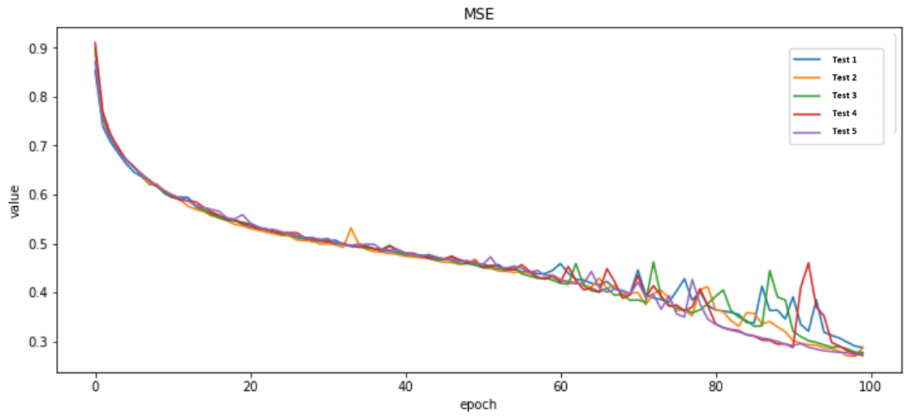

Figure 6: Model 7, MSE metric. Source: Own study based on collected data

that the use of the sigmoidal activation function gives good results in smog prediction tasks.

7. Conclusions

As we wrote in the Introduction our goal was to perform a comparative analysis of machine learning algo-

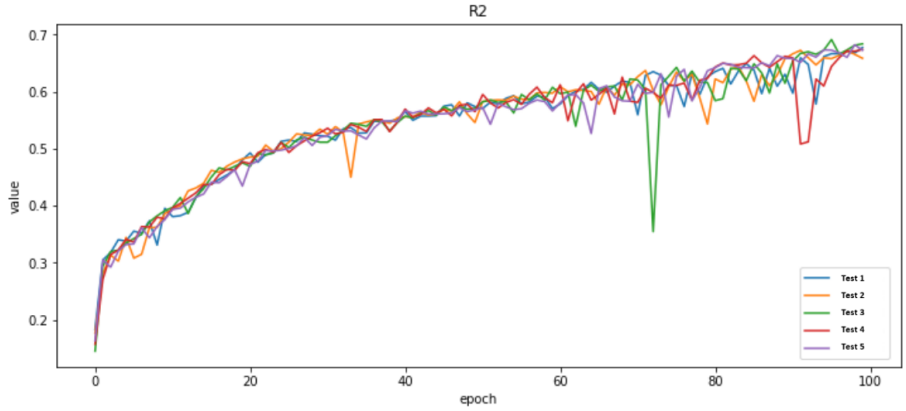

Figure 7: Model 7, R-squared metric. Source: Own study based on collected data

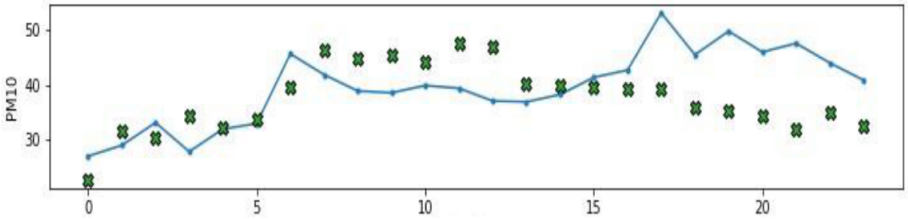

Figure 8: Values predicted for the measured values. Source: Own study based on collected data

rithms based on an air pollution prediction model. Using the data provided by InConTech, including atmospheric and partly meteorological data, research was carried out to determine the influence of model parameters on their final results and the quality of predictions. The analysis of the results showed that the optimal solution was to use the Long Stort Term Memory LSTM algorithm in smog prediction. It is a recursive model of an artificial neural network that is ideally suited for prediction tasks. Variable parameters such as: optimization function, activation function, number of network layers, number of units in each layer were subjected to the study. It was possible to establish both the positive and negative influence of the parameters on the result generated by the ANN model by repeatedly performing tests on a number of models. The research was carried out with the help of the Python 3 programming language, along with the necessary libraries such as TensorFlow and Jupyter Notebook.

Further research on the models may develop in various directions, ranging from increasing the number of trials, which would be linked to more reliable data, ending with increasing the number of types of algorithms studied. Developing the models by testing other types of activation and optimization functions would also be able to improve understanding of how they affect the data presented. One way to develop research is to check the application of more layers, perhaps even combining them with other types of algorithms. A very interesting developmental task may be to focus on a self-learning artificial intelligence algorithm, so that the algorithm can learn on a regular basis, not only on historical data. These studies would contribute significantly to the amount of data collected, its analysis and prediction quality in the future.

References

[1] P. Kleczkowski, *Smog w Polsce.* Warszawa: Wydawnictwo naukowe PWN, 2020.

[2] J. Chełmiński, *SMOG. Diesle, kopciuchy, kominy, czyli dlaczego w Polsce nie da się oddychać?* Poznań: Wydawnictwo Poznańskie, 2019.

[3] H. Mazurek and A. Badyda, *Smog. Konsekwencje zdrowotne zanieczyszczeń powietrza.* Warszawa: PZWL Wydawnictwo Lekarskie, 2021.

[4] K. Górka, *Ocena skuteczności polityki antysmogowej Państwa.* Wrocław: Prace naukowe uniwersytetu ekonomicznego we Wrocławiu, 2018.

[5] Airly, "Airly." https://airly.org, 2022. dostęp: 16.05.2022.

[6] GIOŚ, "GioŚ." https://www.gios.gov.pl/pl/, 2022. dostęp: 16.05.2022.

[7] Syngeos, "Syngeos." https://syngeos.pl/, 2022. dostęp: 16.05.2022.

[8] E. P. GIG, "Eko patrol gig." https://monitoringjakoscipowietrza.pl/, 2022. dostęp: 16.05.2022.

[9] InConTech, "Incontech." http://incontech.eu/, 2022. dostęp: 16.05.2022.

[10] M. Kasperski, *Sztuczna inteligencja.* Warszawa: Wydawnictwo Helion, 2003.

[11] S. Raschaka and V. Mirjalili, *Python. Machine learning i deep learning.* Gliwice: Wydawnictwo Helion, 2021. Biblioteki scikitlearn i TensorFlow 2.

[12] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning. Systemy uczące się.* Warszawa: Wydawnictwo PWN, 2018.

[13] Z. Valentino, *Deep Learning. Uczenie głębokie z językiem Python. Sztuczna inteligencja i sieci neuronowe.* Gliwice: Wydawnictwo Helion, 2021.

[14] D. Qin, J. Yu, G. Zou, R. Yong, Q. Zhao, and B. Zhang, "A novel combined prediction scheme based on cnn and lstm for urban pm2.5 concentration," *IEEE Access*, vol. 7, pp. 172052–172061, 2019.

[15] T.-C. Bui, V.-D. Le, and S.-K. Cha, "A deep learning approach for forecasting air pollution in south korea using lstm," *arXiv*, p. 1801.05746, 2018.

[16] M. Zeinalnezhad, A. Gholamzadeh, and J. Klemeš, "Air pollution prediction using semi-experimental regression model and adaptive neuro-fuzzy inference system," *Journal of Cleaner Production*, vol. 261, p. 121218, 2020.

[17] D.-R. Liu, S.-J. Lee, Y. Huang, and C.-J. Chiu, "Air pollution forecasting based on attention-based lstm neural network and ensemble learning," *Expert Systems with Applications*, vol. 160, p. 113726, 2020.

[18] A. Heydari, M. Majidi Nezhad, D. Astiaso Garcia, A. H. Gandomi, H. Karami, and A. H. Alavi, "Air pollution forecasting application based on deep learning model and optimization algorithm," *Clean Technologies and Environmental Policy*, vol. 24, p. 607–621, 2022.