

An intelligent oil accident predicting and classifying system using deep learning techniques

Yasmen Samhan Abd ElWahab¹, Mona Mohamed Nasr², Fahad Kamal Al Sheref¹

¹Information Systems Department, Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef, Egypt

²Information Systems Department, Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt

Article Info

Article history:

Received Sep 1, 2022

Revised Sep 19, 2022

Accepted Oct 3, 2022

Keywords:

3W database

Adam

Deep learning classification

Optimizers

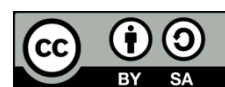
Petrol

Tanh

ABSTRACT

This study discusses the problem of oil and gas faults that lead to spills or explosions that lead to a lot of losses in human life, oil field extraction, and costs. Petrol is an important field in our lives because it controls all aspects of human life and their way of life, so our research focused on petrol and its problems in order to introduce a better way of life. The data used in this research was taken from the 3w database that was prepared by Petrobras, the Brazilian oil holding. The 9 classes classified in that work include the normal state that indicates the factors that will not lead to a problem. Deep learning classification techniques were used in this study. 99% accuracy was obtained in that model, and it refers to a successful prediction and classification of each class. Different results were observed when different hidden layers, optimizers, neurons, epochs, and activation functions were used. 99% was achieved when using Adam's optimizer and Tanh's activation function.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Yasmen Samhan

Department of Information System, faculty of Computers and Artificial Intelligence, Beni-Suef University

Beni Suef Governorate 2722165, Egypt

Email: yasmen-samhan@yahoo.com

1. INTRODUCTION

Accidents are happening all the time. Dealing with accidents is so different from trying to avoid the accident itself and is more important as knowing the source of the accident is more important than handling the accident itself. If we handle the accident, it will not solve the problem from the root, and with the same causes, there will be a high prediction for the accident happening again. Our study focuses on causes that lead to other causes that will cause accidents, and by occurring the same first causes, a prediction class will be discovered. Oil is one of the most important and dangerous fields in our lives as one mistake not only leads to money loss but also leads to a lot of other life losses such as marine life, birds' life, sea creatures, and the most sensitive and costly human life as explained in [1], [2]. Each one of them is so important and affects the world in all its lives, such as the economy and reputation of those who fear working in the oil fields; a field without people to work on is no field; oil controls the world; most of the machines are working with oil and gas, and all the kings of oil extraction and transforming. Oil field safety is first and companies should grant that safety to their employees with each degree, so this study focuses on reasons that can lead to any problem that can lead to any explosions or spilling in that field [3].

Oil problems are one of the most important problems we need to speak about and discuss in detail. First, we are going to learn about different important issues such as: 1) the importance of oil usage in our lives, 2) how the process of extraction and transportation happened, 3) how problems and accidents affect our lives generally and our economy especially, 4) how to fix and treat these problems, and 5) some of the

accidents and countries that have been affected and faced a lot of losses because of these problems as discussed in [4].

Bejarano *et al.* [5] discussed, companies can extract oil from onshore and offshore. Onshore refers to all extraction from all locations on the ground, not water, such as green and desert lands around the world, while offshore refers to all locations in water places such as seas or oceans. This research and the dataset are about the offshore problem. Why offshore when marine life is so important for the world and all these sea creatures that enhance human life are the ones closest to danger when a problem occurs? Also, onshore oil accounts for 70% of the world's oil extraction rate, whereas offshore oil accounts for 30% of the world's oil extraction rate, though this percentage may differ from country to country. Devold [6] discussed, some of the reasons for the difference in oil extraction are that oil in the sea requires special and expensive tools to extract it, moving through the sea is difficult and requires ships, a plane takes off to the location for quick arrival, and the transportation process requires massive ships to simply transport the production to the land. Not only the sea creatures that will be affected by that problem, but also all the people who live in the location of the accident and also the people that live near to that location. All birds that drink from the oil-infected water will perish. What a loss in all these lives. This study is interesting in trying to prevent that danger as much as possible and asking for more research in such a field.

The deep neural network (DNN) was used in this model with four fully connected layers with tanh activation function and an output layer with softmax activation function. The adam optimizer gave the highest performance of the model. A powerful technique such as deep learning created a powerful model that predicts and classifies problems in the oil sector, solving the problem of overfitting of the data. DNN was used in a lot of great research that solved many problems in different fields, and it was better than other methodologies [7].

2. RELATED WORK

Problems with oil are many, and the role of new science is to present solutions to all problems that face the world in all fields. As mentioned before, the importance of oil fields and its effect on both the economy and real life means lots of research has been done in this field. Here are some papers that handle specific issues in that field.

A pressure-balance-based approach was proposed in this study [8] to capture two types of active circular pressure behaviour in a gas well with sustained case pressure (SCP) produced by pipe leakage at different ideal depths and gas-lift faucet failure during operation. This method employs models that simulate the effects of tubing and annulus fluid temperature and pressure spreading to determine tubing spill spots and evaluate the average annulus pressure at the rig under various conditions. This study put the diagnosis of offshore gas leaks to the test.

In May 2015, a multi-frequency radar system was utilised to picture monitored hazardous and noxious substance (HNS) discharges above the Mediterranean Sea [9]. The primary goal of this study was to develop a process for gathering evidence of illegal oceanic contamination by harmful liquid chemicals utilising airborne radar sensors, also known as a flying radar sensor. They demonstrate the ability of radar imaging to detect, describe, and discriminate compounds discovered in the water. They used a regulated polarisation variance parameter to list both the impacts of the substance discharged on the ocean surface and the relative concentration of the material in the spill, demonstrating that radar imaging can take into account HNS information. One can tell the difference between a product that forms a film on the sea's surface and one that combines with saltwater, which is crucial information for effective cleanup methods.

For the vulnerability valuation of the urban buried gas pipeline network, an equivalent structure for the novel application of supporting vector machine (SVM) and artificial neural network (ANN) techniques was proposed, and an evaluation was conducted on different phases between these two techniques. Model development sample data was simulated and split into two data sets for model training and confirmation. As a result, the training outcomes of ANN are more sophisticated than those of SVM, causing ANN to be overfitting. The SVM model's training results are better matched with the required outputs than the ANN model's, and it can achieve satisfactory production values and is more accurate than the ANN model when applied to unidentified situations, as demonstrated in [10].

A common methodology, CNN, employed a basic discrimination analysis to characterise the selected characteristics subgroup to categorise unstructured data in [11]. Entropy, alpha, and single-bounce eigenvalue relative difference are among the properties investigated in the C-band polarimetric mode (SERD). We also propose a new technique for distinguishing oil spills and lookalikes in SAR images. The accuracy of the categorization attained using 900 test data samples is 91.33 percent. As a consequence, we can see that not only can the suggested technique reliably identify black spots on SAR images, but also that the suggested algorithm can categorise unstructured characteristics. The finding also demonstrates that if the difference in sea conditions (such as climatic, geographical, sea temperature, and environmental factors) between test and training data is

too great, classification accuracy will suffer greatly. To lessen the effects of these variables on future tests, more oil spill data from different marine locations will be incorporated in the creation of the CNN model, and the strength of the proposed technique's training structure will be improved. The classification precision would be improved if the prior possibilities of oil spills and lookalikes were taken into consideration for the classification result.

Park *et al.* [12] is an example. They found oil spill regions using high-resolution optical photos using an artificial neural network (ANN) and the operational conquering of austere sunshine outcomes. To do so, a directional median filter (DMF) was designed, and its performance was compared to that of a traditional low-pass filter. A KOMPSAT2 image created as a result of the oil leak catastrophes in the Gulf of Mexico in 2010 was used in a presentation test. The proposed strategy had two primary phases: (1) the DMF was used to adjust the sunlight effects induced by the ocean waves, and (2) the ANN method was utilised to find the location of the oil spill. The following are the outcomes: (a) In controlling the effect of sunlight in a high-resolution optical image, the created DMF, which considers the size and angle of ocean waves, has been great and intelligent; and (b) oil spill regions have been efficiently found utilising the recommended ANN approach. With precision of around 98.12 percent and 89.56 percent, the oil spill region was identified independently using the receiver operating characteristic (ROC) curve and detection probability (POD) values. These findings reveal that, when compared to the usual detection algorithm, the suggested technique improves precision by roughly 9%.

Yan *et al.* [13] is an example. They suggested an oceanic phenomena exploration technique in SAR pictures focusing on coevolutionary neural networks (CNN). The strategy then employs ResNet-50 to eliminate multilevel functionality. Second, it employs the dreadful spatial pyramid pooling (ASPP) module to extract multiscale functionality. Finally, it includes multilevel and multiscale aspects to investigate oceanic processes. The Sentinel-1 satellite's SAR photos are used to create a dataset for studying maritime phenomena. The proposed methodology can be used on the dataset with 91 percent precision.

The investigation in [14] compares the structure of the typical sound event detection (SED) pipeline constituents to previously documented responses to irregular SED. On an isolated and multi-class basis, the system's performance was estimated. It was revealed that performance in a single class is unlikely to be identical and that techniques diverged. Performing the little aberration in a lone class would be a multifunctional strategy that isn't biased toward a particular outcome. As a result, the technique is thought to be better for a wider spectrum of SED problems. The most practicable way is a combination of creating attributes utilised as log-Mel energies in a convolutional recurrent neural network (CRNN) with long short-term memory (LSTM) with a thresholding confirmation phase, which has found 93.1% F1 score and 0.1307 ER.

Using the software Weka, mining techniques were employed in this study report [15] to analyse the leading reasons for underground and surface accidents in Spain. Between 2005 and 2015, data from the Spanish Ministry of Employment and Social Security was used. Physical effort or overexertion of body movement was found to be the most significant immediate cause. Because Weka is unable to handle massive data and emit memory messages in large databases, this work used a bad programme like Weka with limited data sets.

A random forest machine-learning method was used in [16] to anticipate and classify eight different types of oil-field defects. Although a satisfactory result was attained, three classes were excluded from this study because their transient phase was not described in the 3W database. Furthermore, on the enormous dataset, machine learning approaches did not perform well. This study paper employed a good technique, but their classification only yielded 94 percent, whereas our methodology yielded 99 percent by employing a more effective deep learning technology.

In general, most of the above research papers focused on monitoring the location of oil production and, after an accident, by radar, they can discover that there is an accident. Other research focused on how to deal with an accident after it happens and, by advanced technologies, they will determine which tools will be used in handling that problem. Handling the accident after it happens is good work, but what about trying to prevent it from happening as in [15], [16]. They focused on the causes of the accident to predict and classify accidents before they happen by studying their causes, but they used minig and machine learning techniques that face problems in large datasets, and deep learning is a more powerful technique than them.

3. THE PROPOSED METHOD

3.1. Steps that developed for building the model

Building a model for predicting oil problems was complicated and had different steps with different comparisons for deep learning techniques, So, in order to complete the task of creating all of the models and determining which one is best for running and classifying this study, precise and organised steps were taken. 1. Our model handled and obtained a perfect comparison and the result shows that good steps with clear monitoring for this result are a perfect way to obtain the best model, and the following steps were taken to build our model as illustrated in Figure 1.

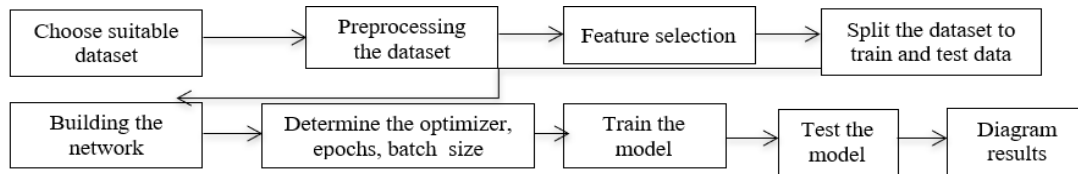


Figure 1. Steps of building the proposed deep learning model

Step 1: Deep learning techniques were deployed in our model with different results, and these different results depended on all the different techniques used in our models, such as relu and tanh activation functions, optimizers, number of hidden layers, and number of neurons, so building the network used in our model was the most important step in building our model. Many networks were built, but the one with tanh activation function, Adam optimizer, four hidden layers, and eight neurons gave the best performance for the model.

Step 2: Determining the capability of software for the model is one of the most important steps to take in building any model. Our model was written in Python code and many libraries were used, especially tensorflow_keras, which is considered one of the best libraries for building deep learning models. Google Colab was used in building our model. It's perfect for building huge and complicated models as it provides good space for both DICK and RAM and needs only a good internet connection to build any model, regardless of the capabilities of the computer or laptop used.

Step 3: The dataset used in this framework was derived from the 3w database created by Petrobras, the Brazilian oil holding company. It's about a lot of files; each class has tens of files in a separate file, but our research depends on how much data there is, and it's hard to obtain data in such a field [17]. This data with its format is too small and not ready for the model to give a good result for the deep neural network. As for organising this data, many files from each class have merged to handle all classes, and second, the huge amount of data means the good result will be obtained from that model. During building the model, some problems were discovered in the dataset that were handled in different ways, such as 1) missing values in the dataset that were handled by replacing each missing value with a zero value, as zero will not affect the nature of the dataset and make the result with a real value, and 2) scaling was done to the dataset used in this model to have huge data for the model to obtain the best result, and that's handled by using StandardScaler from the sklearn library.

Step 4: The dataset was divided into 75% training and 25% testing, which is considered a good split for building the model to handle all possible cases in the training and testing process. And Step 5: After preprocessing the dataset and dealing with all problems in it, as well as drawing our network that we thought was neither too complex nor too simple but was suitable for building the model, the model was built and a different comparison was performed to obtain a model with high accuracy and a low loss function. The comparison was in the number of neurons, tanh and relu activation functions, number of epochs, and different types of optimizers.

3.2. Description of the dataset

The dataset used in this model was not in one file, but a merge of some files was done to obtain a good and useful dataset for our model. This dataset contains features that by preprocessing it we can predict and classify the classes of oil problems such as: Abrupt Increase of basic sediment and water (BSW), Spurious Closure of the downhole safety valve (DHSV), Severe Slugging, Flow Instability, Rapid Productivity Loss, Quick Restriction in the production choke (PCK), Scaling in the production choke (PCK), and Hydrate in the Production Line. These eight classes mean there's a problem that will happen that can lead to a big disaster like an explosion or spill that can also sometimes lead to an explosion. In addition, in these eight classes, there's a normal class that indicates all cases or factors that will not lead to any problem, which means there's a safety mode and monitoring. What bad cases can happen in that sensitive field where any small problem in it will lead to huge problems? Knowing the source of the problem is more important than handling the problem itself, but handling and fixing these resources is so useful for avoiding and preventing repeating the same problem soon. Now, the role of defining these resources will be known below, as well as the features in our model, such as P-PDG: Pressure at permanent downhole gauge (PDG) Pa, P-TPT: Pressure at temperature/pressure transducer (TPT) Pa, T-TPT: Temperature at temperature/pressure transducer (TPT) °C, P-MON-CKP: Pressure upstream of production choke (CKP) Pa, T-JUS-CKP: Temperature downstream of production choke (CKP) and there's three features in the dataset that had not used as they have a lot of missing values that will effect on model accuracy such as P-JUS-CKGL: Pressure downstream of gas lift choke (CKGL) Pa, T-JUS-CKGL: Temperature downstream of gas lift choke (CKGL)

°C, and QGL: Gas lift flow rate m³/s. Pa refers to pascal, C refers to Celsius, and m³/s refers to meters cubed per second. It's a powerful dataset that, by using it in our model, a successful model has been built to predict and classify each type of problem that leads to an accident.

3.3. Explaining steps of running the model

As mentioned, the step of building a network is the most important step in our model, as any detail of the network will affect the model's performance. Each output layer is an input to the next layer, so the input layer $a^{(L-1)}$ and the output layer $a^{(L)}$. L refers to layer, and each layer has weights $w^{(L)}$, activation function (L), and bias $b^{(L)}$. The model has forward and back propagation [18], [7]. Forward propagation is calculated according to (1) and (2):

$$z^L = W^{[L]}a^{[L-1]} + b^L \quad (1)$$

$$a^L = g^L(z^L) \quad (2)$$

g^L refers to the activation function as in our model tanh and relu.

After calculating the forward propagation will calculate the difference between predicted and real value and a backpropagation will be calculated by calculate the differences between predicted value and real value then reduce the cost function and updated the weights and bias and repeat the two process till obtaining the perfect performance for the model and if the difference value between predicted and real value is small, a little updates for weights and bias will happen and if big a many updates will happen. Backpropagation is calculated by calculated Derivation of the cost function as refers with (3), (4), (5), (6), (7): first define $da^{(L)}$ input, $da^{(L-1)}$ output, $dw^{(L)}$ weight, and $db^{(L)}$ bias

$$dz^{[L]} = da^{[L]} * g^{[L]}(z^{[L]}) \quad (3)$$

$$dw^{[L]} = dz^{[L]} * a^{[L-1]} \quad (4)$$

$$db^{[L]} = dz^{[L]} \quad (5)$$

$$da^{[L-1]} = w^{[L]} * dz^{[L]} \quad (6)$$

$$dz^{[L]} = w^{[L-1]} dz^{[L-1]} * g^{[L]}(z^{[L]}) \quad (7)$$

dw ... derivate the cost function for weights

db ... derivate the cost function for bias

The best accuracy and loss function will be obtained after running that network with specific batch size and epochs.

4. RESEARCH METHOD

4.1. Deep learning

Deep learning is a new field of machine learning research that was introduced to move machine learning closer to one of its original and preliminary goals, such as artificial intelligence. It is about the multi-stage education of representation and abstraction that helps to understand data such as sound, such as images and text as identified in [18]. Deep learning helps us teach machines how to complete and achieve challenging tasks without programming them. We're entering the age of machine learning and artificial intelligence. Deep learning methods are those that can overcome a variety of problems through deep learning.

Li *et al.* [10] discussed, (ANN) is a non-parametric, experimental threat minimization-based modelling technique with the ability to approximate any non-linear function to arbitrary and tyrannical precision, depending on the training of the network neurons to adjust the weights connecting these neurons, and is typically used for device security and protection estimation. Many fields have successfully applied ANN to deficiency estimation, but there are many drawbacks, such as non-convex objective functions, difficulty in measuring the number of hidden layers, and overfitting, which arise due to the huge number of parameters in the model and are a major and significant weakness of the theory of experimental hazard minimization.

As mentioned before, the importance of that field in our lives for all the world was so great that a powerful technique was used to build the model. It's a complicated technique that's difficult to learn but more

powerful to apply. [19] Deep learning has ushered in a new era that has the potential to make the world more intelligent and interesting. It's a complicated technique that's difficult to learn but more powerful to apply. It needs a lot of data to let the model learn well, and providing this data in some fields is not available. If more data can be used, more accuracy will be gained. Deep learning models depend on a lot of different technologies that can be used for building models. Figure 2 shows the network of the proposed model. Normal means that there are no problems, it refers to class 0.

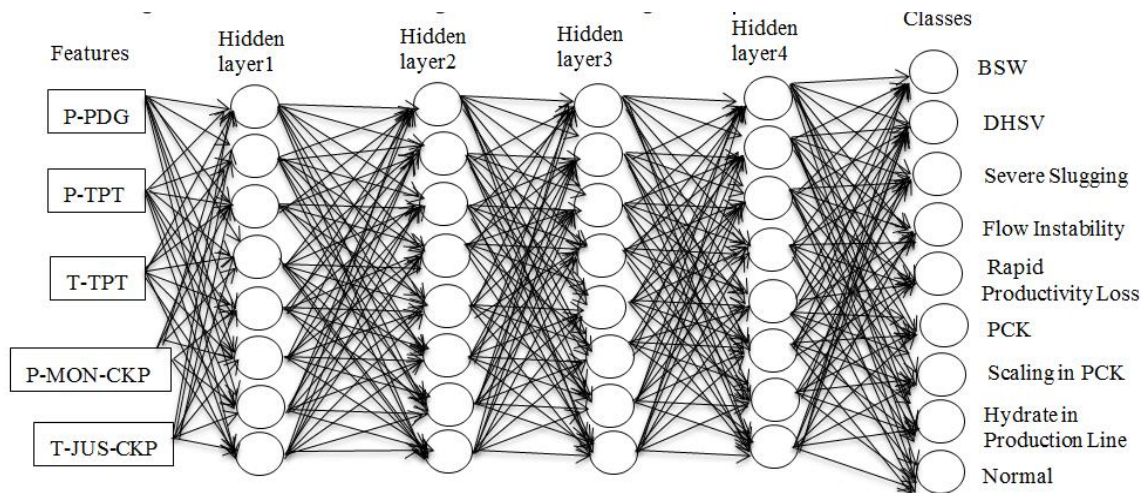


Figure 2. The architecture of the proposed Deep Learning model

4.2. Techniques of deep learning

4.2.1. The number of hidden layers and neurons

The number of hidden layers used in the model depends on the different types and sizes of datasets used in the model; the more complex the dataset, the more hidden layers are used, especially in image and video datasets. However, the dataset used in that model is a CSV file, so only four hidden layers were used in this framework.

As for hidden layers, it depends on the dataset and its complexity. Four to eight neurons were used for each hidden layer, and a result was obtained. That's in addition to the input layer that has five neurons as the number of features is, so there are three layers named input, hidden, and output layer.

4.2.2. The activation functions

There are many activation functions, such as sigmoid, Tanh, SoftMax, and ReLU. Each activation function is better when used in a different dataset. After multiplying the inputs and weights for each feature, an activation function will be deployed on the output of that equation to get the predicted value that will be compared with the real value to determine the loss value, and according to this value, it will either stop training in that step. Examples of these activation functions include.

4.2.2.1 Hyperbolic tangent function (Tanh) activation function

In neural networks, the steps begin with inputs multiplied by weights that play a significant role in the training outcome [19] plus bias, and end with an output that some activation function deploys on that input [20]. Refer with: Figure 3. Here in this step, which activation function will be better than others and why? Tanh is different from the sigmoid function as Tanh is a triangle function and it ranges from 1 to -1 and then its average is zero, which will be more comfortable in calculations. Using Tanh in the hidden layer is better than sigmoid and Relu as proved in [20] as it ranges from 1 to 0 and its average is .5. The model that has been built in this study proves all these facts about Tanh activation function. As a result, the Tanh function's output refers to the (8).

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{8}$$

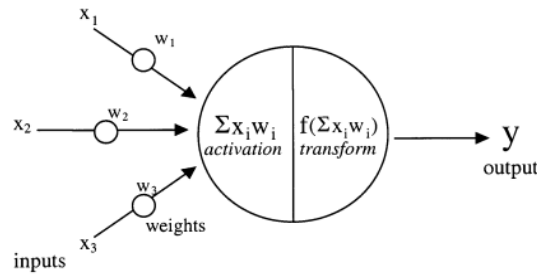


Figure 3. A neuron model for artificial neural network [21]

4.2.2.2 Rectified Linear Unit (RELU)

Relu is another type of deep learning activation function and is the most widely used in deep learning because it's so fast in learning as explained in [22], [23] and by using it, it proved that it's the most successful function [24] and gives better performance and accuracy than Tanh and sigmoid in some situations as shown in [25], [26]. The Relu activation function applies a threshold operation to each input element, setting any values less than zero to zero refers with (9), resulting in the RELU (refer to Figure 4 for visual representation).

$$f(x) = \max(0, x) = \begin{cases} xi, & \text{if } xi \geq 0 \\ 0, & \text{if } xi < 0 \end{cases} \quad (9)$$

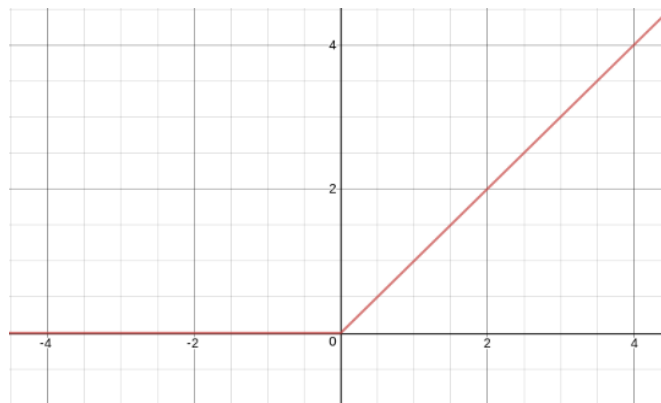


Figure 4. The rectified linear unit (ReLU) activation function produces 0 as an output when $x < 0$, and then produces a linear with slope of 1 when $x > 0$ [27]

4.2.2.3 SoftMax activation function

The SoftMax activation function is another type of deep neural network function that is used for the process of processing. It's been used when multiple classifications will be made [28], as happened in this framework. In the layer of softmax activation function, the number of neurons is equal to the number of the classes of predicted value. In the softmax activation function, if the model accuracy is 100%, it puts 1 for the predicted class and zero for all other classes, and if the model accuracy is 95%, it puts .95 for the predicted class and the .05 will be distributed to all other classes according to their probability of achieving the current features, as the sum of all probabilities should be 1 in all situations submission.

4.2.3. The number of epochs

Detecting the suitable number of epochs needed to achieve the best performance of the model depends on monitoring and noticing the model's performance when it is running and determining if the quality increases in high epochs or not, and if it decreases at some stage of the run to know when the model can stop running and give the highest performance. In the building model in this study, 2000 epochs were used to compare results for each building network and different optimizers and determine which parameters are best for that model.

4.2.4. The optimizer

Before speaking about the optimization process, which includes many techniques to discuss, First, let’s know the mini batch gradient descent. That means the dataset will be split into many pieces, or here called a batch, as in our research and testing, as with each batch, the two processes of forward and backpropagation will be deployed and a new weight will be created for the next batch and so on, so by the end of the data set, thousands of updates of weights will have happened and that process will speed the training process with perfect accuracy. But before the mini batch gradient descent, I have to make a shuffle before the mini_batch gradient descent to make the train data in all different data, not only on the first records that may belong to the only city or case [29], [30].

After knowing the idea of mini-batch gradient descent, you will know the idea of stochastic gradient descent (SGD). Stochastic Gradient Descent will be so clear as this idea means that the number of the batch is equal to the number of records in the data set. That means if there are one million records for training only, it means that there are one million batches, which means there are also one million values for weight and bias [31], [32]. The second optimizer is adaptive moment optimization (Adam). It’s another optimizer that is used for obtaining good training as it is fast and with practise it gives high accuracy, but it depends on many parameters that take memory space [33].

5. RESULTS AND DISCUSSION

As mentioned in the previous section, the dataset and the details of the technique were used in our model to build a network that consists of four hidden layers, eight neuorns, and two types of activation functions and optimizers such as tanh, relu, SGD, and Adam. As a comparison between all the results, the Adam optimizer with tanh activation function had given 99% of accuracy, but the SGD optimizer with RELU activation function had given 93% of accuracy. It’s a good performance for a deep learning model that needs a lot of data to give good accuracy. The number of epochs has played a big role in the result, as with this data, 2000 epochs were needed to obtain that result. Also, 10000 batch-sized was used in this model as demonstrated in Figures 5, 6, 7, 8, 9, 10, 11, 12, 13, and 14 show different results for our comparison of deep learning techniques. The SoftMax activation function was necessary to use in the last layer as our model predicts and classifies nine classes, and it can’t predict multi-classification without that activation function. Our model with this high accuracy of 99%, as shown in Figures 15 and 16, will predict and classify problems that can happen in the oil section, and by using this model, a lot of accidents will be prevented in the future, and put a great strategy to handle all problem types.

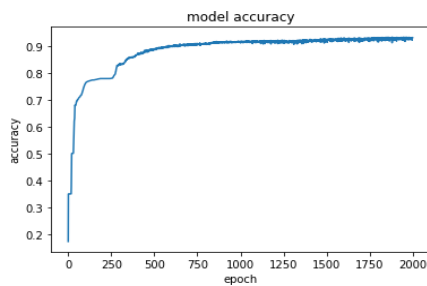


Figure 5. Model accuracy in state of 8 neurons, sgd optimizer, and tanh activation function

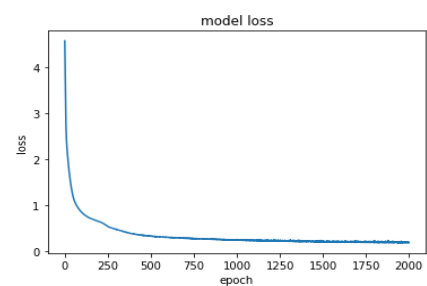


Figure 6. Model loss function in status of 8 neurons, sgd optimizer, and tanh activation function

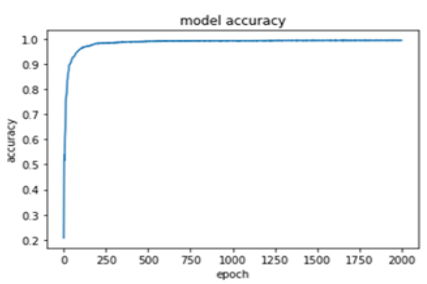


Figure 7. Model accuracy in state of 8 neurons, adam optimizer, and tanh activation function

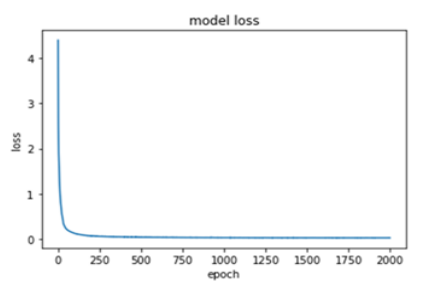


Figure 8. Model loss function in status of 8 neurons, adam optimizer, and tanh activation function

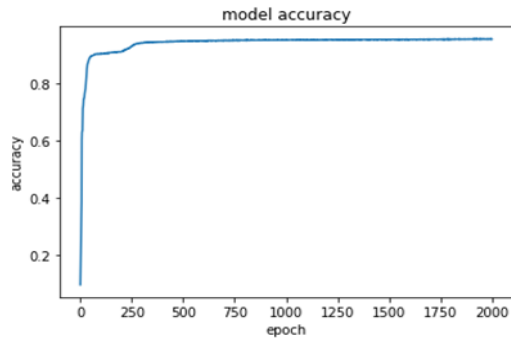


Figure 9. Model accuracy in state of 4 neurons, adam optimizer, and relu activation function

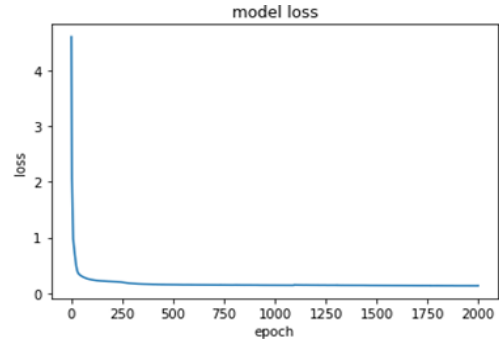


Figure 10. Model loss function in status of 4 neurons, adam optimizer, and relu activation function

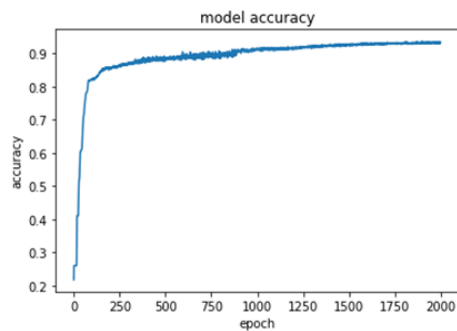


Figure 11. Model accuracy in state of 8 neurons, sgd optimizer, and relu activation function

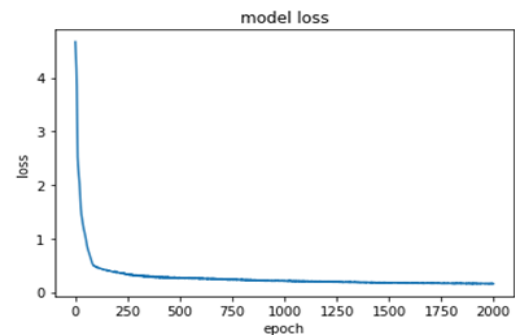


Figure 12. Model loss function in status of 8 neurons, sgd optimizer, and relu activation function

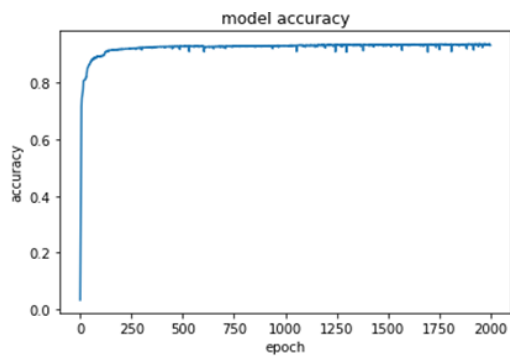


Figure 13. Model accuracy in state of 8 neurons, adam, and relu activation function

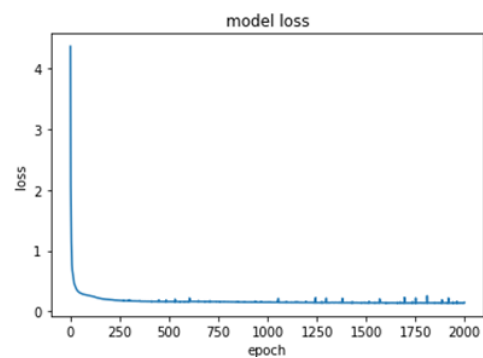


Figure 14. Model loss function in status of 8 neurons, adam optimizer, and relu activation function

```
Epoch 2000/2000
74/74 [=====] - 4s 47ms/step - loss: 0.0211 - accuracy: 0.9937
```

Figure 15. The result of model training in epoch 2000

```
val_loss, val_acc = model.evaluate(X_test, y_test)
7704/7704 [=====] - 10s 1ms/step - loss: 0.0216 - accuracy: 0.9937
```

Figure 16. The result of model testing

From the drawings above, it is noted that the higher the number of epochs, the higher the quality of the model, and that's so important to gain the best performance for the model. This model predicts and classifies the oil problems that face production under some conditions, and by predicting these problems, companies can prevent these problems from happening and handle them before the disaster happens. These figures indicate the comparison between different deep learning techniques, and the next tables from table 1 to table 5 will show the results with numbers for each one, table 6 will show the comparison between our research methodology and others, and table 7 will show the values of parameters used in the proposed model:

Table 1. Model accuracy and loss function in status of 8 neurons, sgd optimizer, and tanh activation function

Epoch	Loss function	Model accuracy
10	2.5082	0.3511
100	0.8379	0.7608
1000	0.2352	0.9172
2000	0.1784	0.9320

Table 2. Model accuracy and loss function in status of 8 neurons, adam optimizer, and tanh activation function

Epoch	Loss function	Model accuracy
10	1.2542	0.6342
100	0.1214	0.9608
1000	0.0262	0.9922
2000	0.0211	0.9937

Table 3. Model accuracy and loss function in status of 4 neurons, adam optimizer, and relu activation function

Epoch	Loss function	Model accuracy
10	0.9813	0.6344
100	0.2436	0.9030
1000	0.1458	0.9527
2000	0.1361	0.9550

Table 4. Model accuracy and loss function in status of 8 neurons, sgd optimizer, and relu activation function

Epoch	Loss function	Model accuracy
10	3.0529	0.2599
100	0.4739	0.8211
1000	0.2108	0.9128
2000	0.1585	0.9342

Table 3. Model accuracy and loss function in status of 4 neurons, adam optimizer, and relu activation function

Epoch	Loss function	Model accuracy
10	0.9813	0.6344
100	0.2436	0.9030
1000	0.1458	0.9527
2000	0.1361	0.9550

Table 4. Model accuracy and loss function in status of 8 neurons, sgd optimizer, and relu activation function

Epoch	Loss function	Model accuracy
10	3.0529	0.2599
100	0.4739	0.8211
1000	0.2108	0.9128
2000	0.1585	0.9342

Table 5. Model accuracy and loss function in status of 8 neurons, adam, and relu activation function

Epoch	Loss function	Model accuracy
10	0.7131	0.7502
100	0.2598	0.8934
1000	0.1522	0.9301
2000	0.1485	0.9319

Table 6. Comparison between deep learning methodology and other methodologies

Dataset	Technique	Accuracy
data from the Spanish Ministry of Employment and Social Safety between 2005 and 2015 [15]	Mining techniques	77.4% of cases were properly assigned for underground mines, and 70.6% for surface mines
3w database that prepared by Petrobras, the Brazilian oil holding [16]	random forest classification	94%
3w database that prepared by Petrobras, the Brazilian oil holding [16]	Deep learning techniques	99%

Table 7. Values of parameters used in the proposed model

Parameter	Value
Number of hidden layers	4
Number of neurons	8
Number of epochs	2000
Size of batch_size	10000
Type of activation functions	Tanh, softmax
Type of optimizer	Adam
Type of loss function	sparse_categorical_crossentropy

6. CONCLUSION

Model classification accuracy for oil problems that can lead to spilling or explosion was calculated based on the probability score of the labelled data by using accuracy metrics in training and model evaluation in testing. In this paper, we indicate how to use deep learning techniques such as optimizers, activation functions, hidden layers, number of neurons, and number of epochs in the field of oil. We also describe how deep learning was useful in this study, and a comparison between deep learning techniques was done to produce a more accurate model, which proved that the activation function, the Adam optimizer, four hidden layers, eight neurons, and 2,000 epochs gave the best accuracy with 99%. Related work has been discussed with different tools and datasets. Our dataset is new for use and organised for working on. The reasons for the accident or what they can lead to have been discussed, as well as the importance of that field. It's an important field for researchers and the whole world to be interested in, so there are a lot of other ideas that can be applied and discovered in that field, and researchers need to handle all sides of that field soon. After predicting the accident, similarity algorithms or similar advanced techniques can provide the best solution for dealing with that accident based on its type.




REFERENCES

- [1] D. U. Johnbosco, D. A. Bibiana.C, and S. N. Richard.E, "Impact of used motor oil on the soil qualities of orji mechanic village owerri, nigeria," *Int. J. Eng. Technol. Manag. Res.*, vol. 7, no. 2, pp. 1–12, Mar. 2020, doi: 10.29121/ijetmr.v7.i2.2020.524.
- [2] P. P. Falciglia *et al.*, "Remediation of petrol hydrocarbon-contaminated marine sediments by thermal desorption," *Chemosphere*, vol. 260, p. 127576, Dec. 2020, doi: 10.1016/j.chemosphere.2020.127576.
- [3] C. Mani Kumar, P. Sumithabhashini, B. Sumanya, P., Sravanthi, and G. Susmitha, "Advanced IOT based LPG LeakageSensing and AlertingSystem over GSM and WIFI," vol. 11, no. 2, pp. 792–798, 2020.
- [4] M. Mohamed Nasr, F. Kamal Kamel, and Y. S. Abd Elwahab, "A survey on predicting oil spills by studying its causes using deep learning techniques," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 1, p. 580, Apr. 2021, doi: 10.11591/ijeecs.v22.i1.pp580-589.
- [5] A. C. Bejarano, E. Levine, and A. J. Mearns, "Effectiveness and potential ecological effects of offshore surface dispersant use during the Deepwater Horizon oil spill: a retrospective analysis of monitoring data," *Environ. Monit. Assess.*, vol. 185, no. 12, pp. 10281–10295, Dec. 2013, doi: 10.1007/s10661-013-3332-y.
- [6] H. Devold, *Oil and Gas Production Handbook - An introduction to oil and gas production, transport, refining and petrochemical industry*. 2013.
- [7] B. Bhandari, A. Alsadoon, P. W. C. Prasad, S. Abdullah, and S. Haddad, "Deep learning neural network for texture feature extraction in oral cancer: enhanced loss function," *Multimed. Tools Appl.*, vol. 79, no. 37–38, pp. 27867–27890, Oct. 2020, doi: 10.1007/s11042-020-09384-6.
- [8] S. Wu, L. Zhang, J. Fan, X. Zhang, D. Liu, and D. Wang, "A leakage diagnosis testing model for gas wells with sustained casing pressure from offshore platform," *J. Nat. Gas Sci. Eng.*, vol. 55, pp. 276–287, Jul. 2018, doi: 10.1016/j.jngse.2018.05.006.
- [9] S. Angelliaume, B. Minchew, S. Chataing, P. Martineau, and V. Miegbielle, "Multifrequency radar imagery and characterization of hazardous and noxious substances at sea," in *Remote Detection and Maritime Pollution*, Wiley, 2021, pp. 17–39. doi: 10.1002/9781119801849.ch2.
- [10] F. Li, W. Wang, J. Xu, J. Yi, and Q. Wang, "Comparative study on vulnerability assessment for urban buried gas pipeline network based on SVM and ANN methods," *Process Saf. Environ. Prot.*, vol. 122, pp. 23–32, Feb. 2019, doi: 10.1016/j.psep.2018.11.014.
- [11] H. Guo, D. Wu, and J. An, "Discrimination of oil slicks and lookalikes in polarimetric SAR images using CNN," *Sensors*, vol. 17, no. 8, p. 1837, Aug. 2017, doi: 10.3390/s17081837.
- [12] S.-H. Park, H.-S. Jung, and M.-J. Lee, "Oil spill mapping from kompsat-2 high-resolution image using directional median filtering and artificial neural network," *Remote Sens.*, vol. 12, no. 2, p. 253, Jan. 2020, doi: 10.3390/rs12020253.
- [13] Z. Yan, J. Chong, Y. Zhao, K. Sun, Y. Wang, and Y. Li, "Multifeature fusion neural network for oceanic phenomena detection in SAR images," *Sensors*, vol. 20, no. 1, p. 210, Dec. 2019, doi: 10.3390/s20010210.
- [14] A. S. M. Affendi and M. Yusoff, "Review of anomalous sound event detection approaches," *IAES Int. J. Artif. Intell.*, vol. 8, no. 3, p. 264, Dec. 2019, doi: 10.11591/ijai.v8.i3.pp264-269.
- [15] L. Sanmiquel, M. Bascompta, J. Rossell, H. Anticoi, and E. Guash, "Analysis of occupational accidents in underground and surface mining in spain using data-mining techniques," *Int. J. Environ. Res. Public Health*, vol. 15, no. 3, p. 462, Mar. 2018, doi: 10.3390/ijerph15030462.
- [16] M. A. Marins *et al.*, "Fault detection and classification in oil wells and production/service lines using random forest," *J. Pet. Sci. Eng.*, vol. 197, p. 107879, Feb. 2021, doi: 10.1016/j.petrol.2020.107879.
- [17] R. E. V. Vargas *et al.*, "A realistic and public dataset with rare undesirable real events in oil wells," *J. Pet. Sci. Eng.*, vol. 181, p. 106223, Oct. 2019, doi: 10.1016/j.petrol.2019.106223.
- [18] Theano Development Team, "Deep Learning Tutorial, Release 0.1." LISA lab, University of Montreal, 2015. [Online]. Available: <https://www.cs.virginia.edu/yanjun/teach/2014f/lecture/L20-handout-deeplearningTutorial.pdf>
- [19] S. Hayou, A. Doucet, and J. Rousseau, "On the impact of the activation function on deep neural networks training," in *36th International Conference on Machine Learning, ICML 2019*, 2019, vol. 2019-June, pp. 4746–4754.
- [20] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *Int. J. Artif. Intell. Expert Syst.*, vol. 1, no. 4, pp. 111–122, 2015.
- [21] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *J. Pharm. Biomed. Anal.*, vol. 22, no. 5, pp. 717–727, 2000, doi: 10.1016/S0731-7085(99)00272-1.
- [22] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *ReLU Improve RBMs*, 2010, vol. 33, no. 5, pp. 384–387.
- [23] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [24] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for activation functions," 2017.




- [25] M. D. Zeiler *et al.*, "On rectified linear units for speech processing," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 3517–3521. doi: 10.1109/ICASSP.2013.6638312.
- [26] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 8609–8613. doi: 10.1109/ICASSP.2013.6639346.
- [27] A. F. Agarap, "Deep learning using rectified linear units (relu)," 2018. doi: <https://doi.org/10.48550/arXiv.1803.08375>.
- [28] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: comparison of trends in practice and research for deep learning," 2018. [Online]. Available: <http://arxiv.org/abs/1811.03378>
- [29] A. Botev, H. Ritter, and D. Barber, "Practical Gauss-Newton optimisation for deep learning," in *34th International Conference on Machine Learning, ICML 2017*, 2017, vol. 2, pp. 860–880.
- [30] F. Schneider, L. Balles, and P. Hennig, "Deepobs: A deep learning optimizer benchmark suite," *7th Int. Conf. Learn. Represent. ICLR 2019*, 2019.
- [31] S. Hadjis, C. Zhang, I. Mitliagkas, D. Iter, and C. Ré, "Omnivore: An optimizer for multi-device deep learning on CPUs and GPUs," 2016. [Online]. Available: <http://arxiv.org/abs/1606.04487>
- [32] Y. Wu, M. Ren, R. Liao, and R. B. Grosse, "Understanding short-horizon bias in stochastic meta-optimization," 2018.
- [33] J. L. Ba and D. P. Kingma, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015, pp. 1–15.

BIOGRAPHIES OF AUTHORS






Yasmen Samhan Abd Elwahab    received the B.Sc. degree from the Faculty of Computers and Artificial Intelligence, Beni-suef University, Beni suef, Egypt, in 2016. She is currently M.Sc student in Faculty of Computers and Artificial Intelligence, Helwan University. She is currently an Teaching Assistant in Information systems Department, Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni Suef, Egypt. Machine learning, Deep learning, and game developer. She can be contacted at email: yasmen.drengg@gmail.com.



Prof. Mona Mohamed Nasr    is a Professor at Information Systems Department, Faculty of Computers and Artificial Intelligence, Helwan University HU, Prof. Mona has held many positions such as (Chief Information Officer (CIO) for HU, Scientific Computing Center Manager at HU, Head of Information Systems Department at Faculty of Computers and Artificial Intelligence HU, Vice Dean for Community Service and Environmental Development at FCAI-HU). She received the M.Sc. degree in information systems from Faculty of Computers and Artificial Intelligence, Helwan University, Egypt, 2000. Ph.D. degree in Information Systems from Faculty of Computers and Artificial Intelligence, Helwan University, Egypt, 2006. Prof. Nasr is a well-known speaker with a charismatic presence at academic conferences and seminars, university, and public lectures and radio and television programs in his area of expertise. She published many articles in international journals and conferences and she is an active reviewer for numerous international journals. Prof. Nasr Attended Teaching the Teacher TTT Program in Artificial Intelligence from EPITA, School of Engineering and Computer Science (EPITA-Paris, France) Nov- Dec 2020, she is a member in many other societies such as (IEDRC, IACSIT, MJC, IAOE, TWOWS/OWSD, CSTA, ArabWIC ABI, etc). She can be contacted at email: drmona.m.nasr@gmail.com.



Fahad Kamal ALsheref    was born in Sohag, Egypt, in 1983. He received the B.Sc. degree from the Faculty of Computers and Information, Assuit University, Assuit, Egypt, in 2005, M.Sc. and Ph. D degrees from the Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt in 2011 and 2012. He is currently an associate professor with Information systems Department, Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni Suef, Egypt. He has authored or co-authored over 16 researches publications in peer-reviewed reputed journals. His research interests include social informatics, Health information systems, machine learning, and data mining. Dr. Fahad has se a reviewer for a number of international journals. He can be contacted at email: Drfahad@fcis.bsu.edu.eg.