International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue03, 2021

Deep Learning Based Neural Network Model for Anomaly Detection in Automobile Industry

Ganagavalli K[1], Dr.Santhi.V[2] and Krishnamoorthy V[3]

[1] Bannari Amman Institute of Technology, Sathyamangalam, Erode
[2] PSG College of Technology, Coimbatore
[3]Bannari Amman Institute of Technology, Sathyamangalam, Erode

Email: [1]kganaga@gmail.com

Abstract.* *Because of the recent advancement in the technologies like Internet of things and Artificial Intelligence, most of the industries have adapted into automation. In our routine life most of the devices have become automated one with the help of more connectivity and flawless integration of information technology. Most of the modern vehicles are equipped with smart technologies with security concepts. To identify the abnormality in vehicle network, an anomaly detection mechanism is proposed in this paper. By using this anomaly detection system we are able distinguish the vehicular attacks in the networks and also able to detect the manufacturing defects for ensuring the quality assurance. Due his paper describes a deep learning based neural network model that will explore the abnormalities in the vehicular network data. Various steps in the neural network implementation part have been illustrated. It also evaluates the experimental results by applying the deep learning method on the given real-time data. Due to the immediate responsiveness of the system, it has been gaining much more attention in the automobile industry.
Keywords: in-vehicle networks, anomaly detection, abnormality.

1. INTRODUCTION

Internet of things is a technology where devices will be connected via Internet and they can exchange information with each other without any human interventions. This type of applications can be used for automating the daily routine of people and also automating processes in private industries and government organizations.

Recently most of the vehicles have been equipped with different type of sensors and actuators. Each of them will be having nearly 40-80 electrical control units which are able to connect and communicate to the vehicle network. The vehicle network is used to transmit the data coming from the sensors and actuators which will form a complex network transmission[3]. So for identifying the faults in the vehicle during the manufacture and also to assess the behavior vehicle the data from sensors and actuators will be recorded and will be analyzed using some technology[4]. The recording will be done by the automobile manufacturer during the test drive done before starting the production of any vehicle model.

With the advancement of big data analytics in various industries, algorithms can be designed for identifying the abnormal behavior of vehicles and data[5]. This can be reported to the concerned management so that those types of faults can be either avoided or rectified during the manufacturing phase of the particular model. To deal with this type of problem this paper proposes a deep learning based model which will detect the abnormal activity

2. RELATED WORK

In [1] the author proposes an anomaly detection system based using one class support vector machine which works on the multivariate time series data[6]. The methodology is used to identify the faults in the recorded data from the vehicle networks.

In [2] the author proposes a technique for anomaly detection using intrusion detection technique in the recorded data. The basic objective of this technique is by using the entropy assigned to the data transmitted i9n-vehicle network[7][8]. Each data will be having some entropy which is the randomness. By observing the entropy of healthy data, the system can able to distinguish the abnormal data which will be generated during attacks by the intruders in the network.

3. PROPOSED WORK

a. System overview

The proposed work is an automated system for identifying faults and abnormal behavior in the recorded data of newly designed automobiles using in-vehicle network. The system employs a LSTM neural network model for classifying the given data and labeling the data into normal and abnormal data. In this work healthy and broken data have been taken for training the model and based on the trained data, the abnormality will be identified whenever it runs the test on the real-time data.
This system consists of three parts:
a) Gathering input and preprocessing of data
b) Designing Sequential model
c) Training of test data.

b. Gathering Input and Data Preprocessing

In the proposed work LSTM neural network model has been applied. The system takes the recorded inputs from the in-vehicle network. The input data contains the vibration data and acceleration data which was recorded during the test drives. The input data contains both normal data and abnormal data which are defined as healthy and broken data.

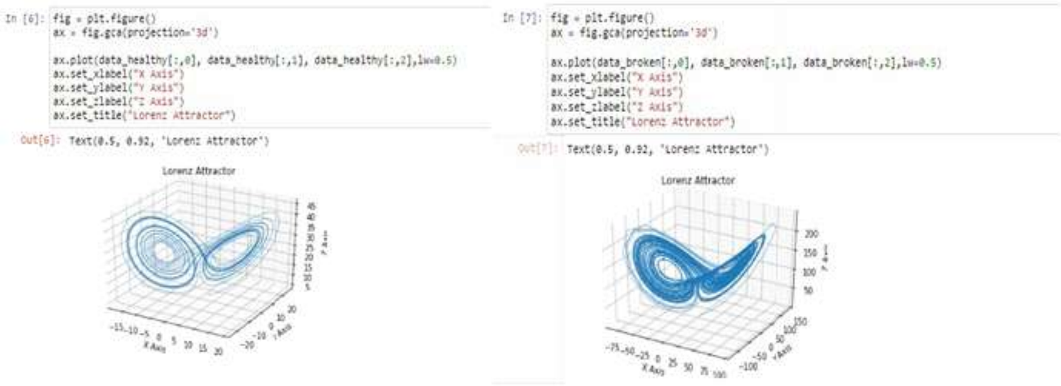

Fig. 1. Healthy and broken data

Before training the data, the input data should be preprocessed by removing unwanted data and outliers. In this system FFT (Fast Fourier Transform) is used for preprocess the data. The following figure Fig.2. Shows the preprocessed data.

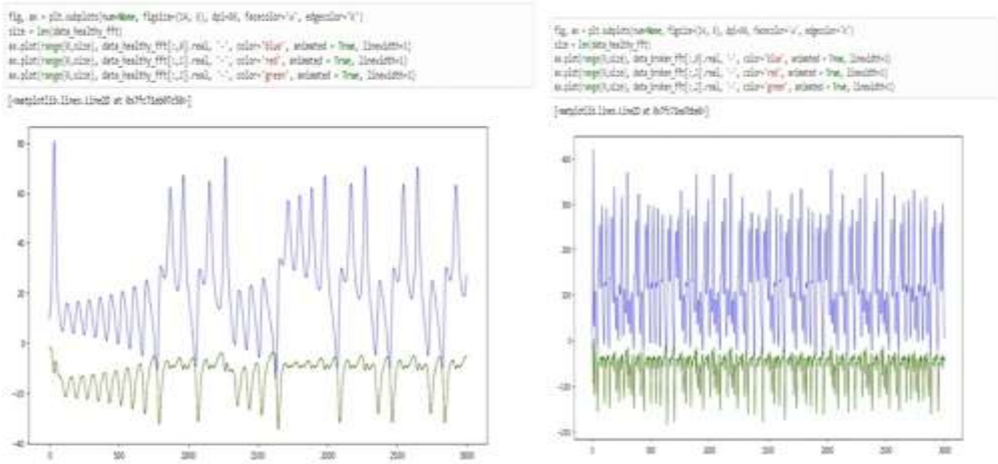

Fig. 2. Preprocessed input data (Healthy and Broken)

c. Sequential Model

The sequential model has been designed for classification data by assigning labels. A sequential model will be having a number of hidden layers where the input will be processed and given to the next layer in the sequence. Most of prediction and classification problems can be implemented as a sequence model. Here the LSTM neural network is defined a three layer sequence model for classification. The model is defined as follows in the diagram Fig.3.

```
number_of_neurons_layer1 = 3
number_of_neurons_layer2 = 3
number_of_neurons_layer3 = 1
number_of_epochs = 20
```

```
[24]: # design network
from tensorflow.keras import optimizers
sgd = optimizers.SGD(lr=0.01, clipnorm=1.)

model = Sequential()
model.add(Dense(number_of_neurons_layer1,input_shape=(dim, ), activation='relu'))
model.add(Dense(number_of_neurons_layer2, activation='relu'))
model.add(Dense(number_of_neurons_layer3, activation='sigmoid'))
model.compile(loss='binary_crossentropy', optimizer=sgd)

def train(data,label):
  model.fit(label, data, epochs=number_of_epochs, batch_size=72, validation_data=(data, data), verbose=0, shuffle=False,callbacks=[lr])
def score(data):
    return model.predict(data)
```

Fig. 3. LSTM sequential model

In the LSTM sequential model three layers have been defined. Each layer will be having user define number of neurons based on the complexity of problem. Here the dimension has been given as 3000 samples and activation function is "relu:. Once we have designed the sequential model the system is ready for training the data.

d. Training data

Once the model is designed, we can feed the input data for training. The sequential model classifies the data by assigning labels. Here the model assigns "1" for broken data and "0" for healthy data. The labeled data can viewed as follows:

```
In [28]: pd.DataFrame(train_both)
Out[28]:
```

	0	1	2	3	4	5	6	7	8	9	...
0	0.250133	0.255146	0.260753	0.267016	0.274003	0.281791	0.290462	0.300109	0.310834	0.322747	...
1	1.000000	0.997476	0.995146	0.992915	0.990694	0.988398	0.985940	0.983234	0.980193	0.976722	...
2	1.000000	0.997476	0.995146	0.992915	0.990694	0.988398	0.985940	0.983234	0.980193	0.976722	...
3	0.151988	0.159267	0.170598	0.188207	0.215687	0.258880	0.327175	0.434534	0.596514	0.810387	...
4	0.957909	0.954314	0.949721	0.942610	0.930689	0.910010	0.873657	0.810554	0.708197	0.572845	...
5	0.957909	0.954314	0.949721	0.942610	0.930689	0.910010	0.873657	0.810554	0.708197	0.572845	...

6 rows × 3001 columns

Fig. 4. Classification of data by training

Here we have totally taken 3000 sample data which was collected from in-vehicle network. The data has been trained so that the model can able to identify the unhealthy data whenever it encounters a new set of recordings.

e. Experimental Results

Once the system is trained we can able to feed new recorded data so that it can easily classify the data whether it is healthy or unhealthy. By using this system, high rate of accuracy is achieved which is comparatively more than the other techniques. The implementation results have been given as follows:

```
In [81]:  score(data_healthy_scaled)

Out[81]:  array([[0.47900042],
                 [0.45667896],
                 [0.45667896]], dtype=float32)

In [82]:  score(data_broken_scaled)

Out[82]:  array([[0.49421665],
                 [0.4567853 ],
                 [0.4567853 ]], dtype=float32)
```

Fig. 5. SVM Prediction rate

```
In [39]:  score(data_healthy_scaled)

Out[39]:  array([[0.6107939],
                 [0.612749 ],
                 [0.612749 ]], dtype=float32)

In [40]:  score(data_broken_scaled)

Out[40]:  array([[0.52274895],
                 [0.6218389 ],
                 [0.6218389 ]], dtype=float32)
```

Fig. 6. LSTM prediction rate

From this we can see that the rate of accuracy (score) is achieved more than 60% which is much better than the other algorithms.

4. CONCLUSION

This paper proposes a classification technique which is based on LSTM sequential model that is used for detecting anomaly activity in the in-vehicle network. Based on the experimental results we can able to see the rate of accuracy achieved. With this rate of accuracy, this system can be more reliable in detecting abnormalities and unhealthy data which may be transmitted by attackers or caused by system defects. By using this system, the automobile industries can able to predict the defects during the test drive itself and may be helpful in identifying any type of attacks caused in the networked data.

5. REFERENCES

[1] Andreas Theissler, “Anomaly detection in recordings from in-vehicle networks”, Big Data Applications And Principles, First International Workshop, BIGDAP 2014,pp.1-10.
[2] M. Müter and N. Asaj, "Entropy-based anomaly detection for in-vehicle networks," 2011 IEEE Intelligent Vehicles Symposium (IV), Baden-Baden, 2011, pp. 1110-1115, doi: 10.1109/IVS.2011.5940552.
[3] S.Woo, H.J.Jo and D.H.Lee, “A practical wireless attack on the connected car and security protocol for in-vehicle CAN”, IEEE Trans.Intell.Transp.System,vol.16, No.2, pp.993-1006, Apr,2015.
[4] J.H.Kim, S.H.Seo, N.T.Hai, “Gateway framework for in-vehicle networks based on CAN, FlexRay and Ethernet”, IEEE Trans Veh.Technology, Vol.64, No.10, pp.4472-4486,Oct,2015.
[5] C.Miller and C.Valasek “A survey of remote automotive attacks surfaces” in Proc.Scribd, vol.21, Washington,DC,USA,2014, pp,34-90.
[6] Mao, J., Sun, Q., Wang, X., Muthu, B., & Sujatha Krishnamoorthy, S. (2020). The importance of public support in the implementation of green transportation in the smart cities. Journal of Computational Intelligence. Wiley publications .https://doi.org/10.1111/coin.12326. 26th April 2020
[7] Yasoda, K., Ponmagal, R.S., Bhuvaneshwari, K.S. K Venkatachalam, “ Automatic detection and classification of EEG artifacts using fuzzy kernel SVM and wavelet ICA (WICA)” Soft Computing Journal (2020).
[8] K. Venkatachalam, A. Devipriya, J. Maniraj, M. Sivaram, A. Ambikapathy, and S. A. Iraj, "A novel method of motor imagery classification using eeg signal," *Artificial intelligence in medicine,* vol. 103, p. 101787, 2020.