



Analysis And Prediction Of Increased Energy Consumption Using Machine Learning With Genetic Algorithm

Dibyahash Bordoloi Head of the Department, Department of Computer Science & Engineering, Graphic Era Hill University, Dehradun, Uttarakhand India, 248002 dibyahashbordoloi@geu.ac.in

Durgaprasad Gangodkar, Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 dgangodkar@yahoo.com

Devesh Pratap Singh Department of Computer Science & Engineering, Graphic Era Deemed to be University, Dehradun, Uttarakhand India, 248002 devesh.geu@gmail.com

Abstract- The machine learning (ML) is a kind of artificial intelligence (AI) that allows the machine to work based on the behaviours of human intelligence. This absorbs the predefined or the historical data as the input value to obtain the expected output values. The important example of the machine learning is the image processing techniques which enables by extracting the information from the images to obtain as digital information. Hence the machine learning with this advanced features helps in the prediction and analysis of energy consumptions with higher accuracy and precision values using genetic algorithm.

Keywords: Machine learning, artificial intelligence, historical data, energy consumption

I. Introduction

The increased energy consumption is due to the increase in population. This can be controlled and monitored by artificial intelligence. The another way of reducing the increased energy consumption is the usage of the renewable energy sources [1]. The machine learning with genetic algorithm is used to obtain higher precision and accuracy because it functions based on the functioning of the human intelligence. The several types of machine learning includes supervised learning, unsupervised learning and reinforcement learning. These learning helps to function the machine as the decision and task performed by the humans. They are programmed and functioned based on the input and output functions obtained by training and testing at the hidden layers [2]. The number of inputs are trained and tested to obtain the optimum desired output. This enables to allow the computer to feed the algorithm and obtain as single output source.

This leads to the automation in the overall systems, obtain wider range of applications, data acquisition, error prone output data and reduced computational cost [3]. Hence the machine learning is the subset in the artificial intelligence which learns the input data to give output through the future prediction. This is done through the genetic algorithm which used for solving both the constrained and

the unconstrained problems that are evolved with natural selection process. Genetic algorithms represent a powerful brute-force optimization method [4]. The approach is particularly useful in nonlinear or stochastic problems, problems represented by complex multidimensional surfaces, or problems with a large number of dependent variables. This is because unlike traditional optimization methods it does not rely on a local gradient calculation and therefore does not tend to be limited by local minima [5]. Hence the system helps to analyse and predict the higher energy consumption through real time implementation.

II. Proposed system

The main objective of the system is to obtain the analysis of energy consumption through the machine learning. This helps to reduce the increasing rate of energy consumption that tends to the shortage in future conditions. The machine learning is inculcated with the genetic algorithm for automatic prediction. This genetic algorithm works based on the training and testing with multiple inputs. The genetic algorithm is a high quality solutions to obtain the optimization with optimum output. This is used to solve the various complicated problems with possible accurate outcome in a shorter time for longer consequences.

III. Methodology

A database is created to obtain the necessary information and stores them for future references. This database is accompanied with the machine learning with genetic algorithm [6]. The genetic algorithm is incorporated with two major features such as fitness function and the crossover techniques. They help to solve more complex problems with lesser computational time. The basic principles of the functioning of the genetic algorithm includes the selection of the data, crossover and mutation techniques [7]. They are generated by the initiation of the data and evaluating their fitness functions with crossover and evaluating and merge of offsprings with the predefined values.

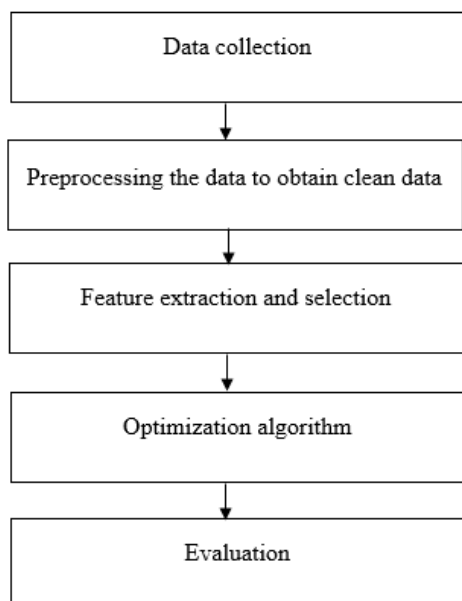


Fig 1: Machine learning stages

The figure 1 shows the various stages of the machine learning from data collection to incorporating the optimization algorithm used for the complete evaluation of the outcome. The optimization algorithm is an important tool to obtain the desired parameters with set of operating conditions.

Energy consumption is done by the prediction of the previous data that are proceeded to the preprocessing technique. The preprocessing involves the collected data to get rid of harmonics by eliminating them using filters. The clean dataset is essential to obtain the optimum output. Then the data are processed to future extraction techniques [8]. Feature extraction is defined as the classification of raw data into a processed numerical data in the original dataset. Then the optimization algorithm i.e. the genetic algorithm is used to obtain the accurate results. The figure 2 represents the analysis and prediction of the energy consumption.

```

In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import MinMaxScaler
from sklearn.cluster import KMeans
import pprint

from keras.models import Sequential
from keras.layers import Dense
from keras.layers import LSTM
from keras.layers import Dropout

%matplotlib inline
Using TensorFlow backend.

In [3]: df = pd.read_csv("AEP_hourly.csv")
print("-"*50)
print("First Five Rows ", "\n")
print(df.head(5), "\n")

print("-"*50)
print("Information About Dataset", "\n")
print(df.info(), "\n")

print("-"*50)
print("Describe the Dataset ", "\n")
print(df.describe(), "\n")

```

Fig 2: Analysis and prediction of energy consumption

The energy consumption is an increasing parameter which tends to rise gradually at every year. This is due to the increase in the functioning of the domestic loads. The increase in rise of energy consumption must be monitored to analyse and predict the need for the consumption of energy through the artificial intelligence. The rise of energy is depicted in figure 3.

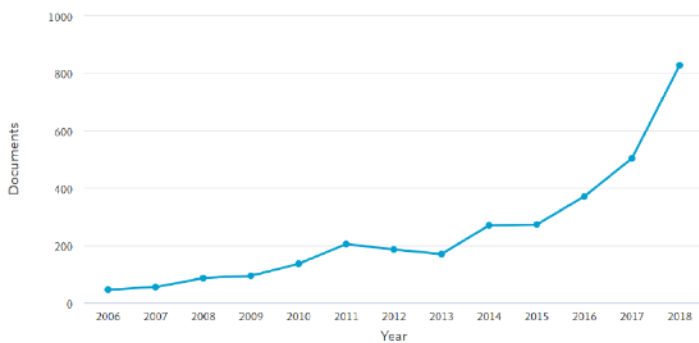


Fig 3: Rise in power consumption

The obtained energy must be evaluated by the genetic algorithm through time. The energy with respect to time gives the exact usage of energy at a particular intervals of time [9]. The evaluation of the energy can be classified as peak hours and the off peak hours utility with respect to time. The energy consumption with respect to time is represented in figure 4.



Fig 4: Energy consumption with respect to time

The another process involved in the energy consumption prediction is the resampling of the data. The resampling of data is the process of repeatedly including the subsets from the training and testing of the each sample in order to gain the outcome through the variability of the data [10]. The figure 5 represents the resampling of the data.

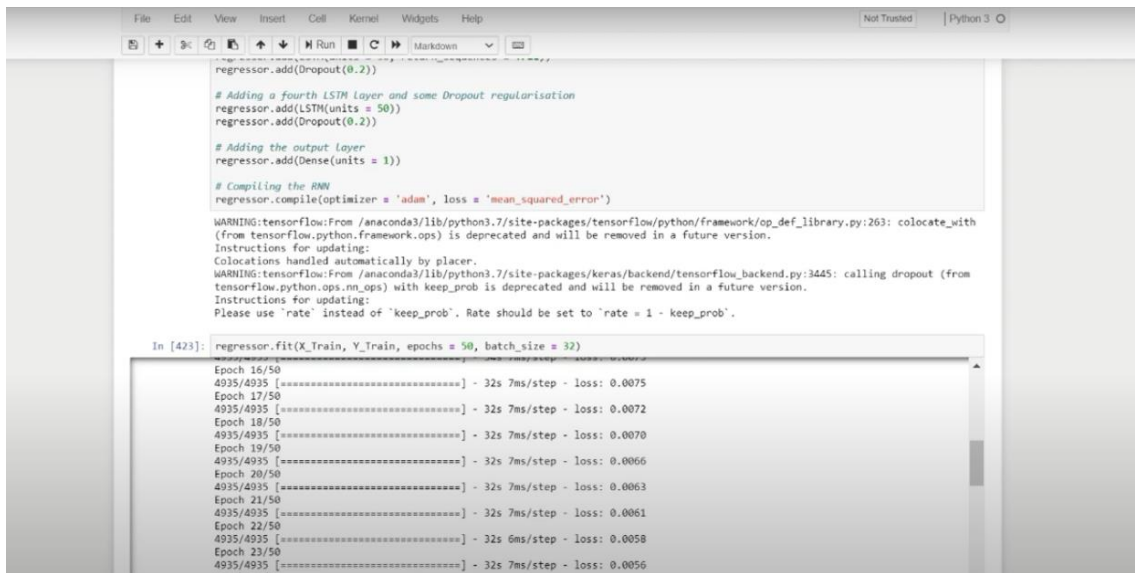


Fig 5: Resampling of data

The various types of resampling of data includes the normal resampling, permutation resampling and cross validation resampling [11]. It is done to improve the accuracy and enumerate the uncertainty of the system. This is done when the majority and minority class instances are out of balance condition.

By reducing the number of training data samples helps to reduce the run time and the storage problems [12].

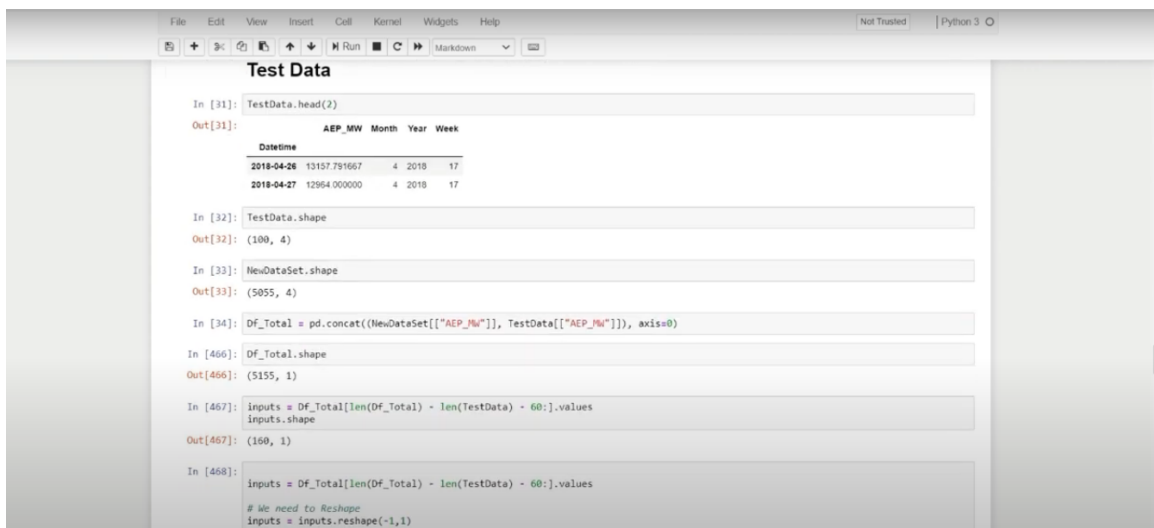
The training phase in the machine learning plays an important role to generate the optimum outcome. It is simply training a good values for all the weights and the bias [13]. This shows the accuracy of the model. The data are split into two parameters as training set and the testing set. The figure 6 represents the training phase.



```
regressor.add(Dropout(0.2))  
  
# Adding a fourth LSTM layer and some Dropout regularisation  
regressor.add(LSTM(units = 50))  
regressor.add(Dropout(0.2))  
  
# Adding the output layer  
regressor.add(Dense(units = 1))  
  
# Compiling the RNN  
regressor.compile(optimizer = 'adam', loss = 'mean_squared_error')  
  
WARNING:tensorflow:From /anaconda3/lib/python3.7/site-packages/tensorflow/python/framework/ops_def_library.py:263: colocate_with (from tensorflow.python.framework.ops) is deprecated and will be removed in a future version.  
Instructions for updating:  
Colocations handled automatically by placer.  
WARNING:tensorflow:From /anaconda3/lib/python3.7/site-packages/keras/backend/tensorflow_backend.py:3445: calling dropout (from tensorflow.python.ops.nn_ops) with keep_prob is deprecated and will be removed in a future version.  
Instructions for updating:  
Please use 'rate' instead of 'keep_prob'. Rate should be set to 'rate = 1 - keep_prob'.  
  
In [423]: regressor.fit(X_Train, Y_Train, epochs = 50, batch_size = 32)  
Epoch 16/50  
4935/4935 [=====] - 32s 7ms/step - loss: 0.0075  
Epoch 17/50  
4935/4935 [=====] - 32s 7ms/step - loss: 0.0072  
Epoch 18/50  
4935/4935 [=====] - 32s 7ms/step - loss: 0.0070  
Epoch 19/50  
4935/4935 [=====] - 32s 7ms/step - loss: 0.0066  
Epoch 20/50  
4935/4935 [=====] - 32s 7ms/step - loss: 0.0063  
Epoch 21/50  
4935/4935 [=====] - 32s 7ms/step - loss: 0.0061  
Epoch 22/50  
4935/4935 [=====] - 32s 6ms/step - loss: 0.0058  
Epoch 23/50  
4935/4935 [=====] - 32s 7ms/step - loss: 0.0056
```

Fig 6: Training of data

The training and testing of the data results in splitting the data as train and test data. The learning model is trained using the train data and performance model is measured using the test data. Thus the entire dataset is trained and tested to produce a given desired model [14].



```
Test Data  
  
In [31]: TestData.head(2)  
Out[31]:  
      AEP_MW  Month  Year  Week  
Datetime  
2018-04-26  13157.791667    4  2018   17  
2018-04-27  12964.000000    4  2018   17  
  
In [32]: TestData.shape  
Out[32]: (100, 4)  
  
In [33]: NewDataSet.shape  
Out[33]: (5055, 4)  
  
In [34]: Df_Total = pd.concat((NewDataSet[["AEP_MW"]], TestData[["AEP_MW"]]), axis=0)  
  
In [466]: Df_Total.shape  
Out[466]: (5155, 1)  
  
In [467]: inputs = Df_Total[len(Df_Total) - len(TestData) - 60:].values  
          inputs.shape  
Out[467]: (100, 1)  
  
In [468]: inputs = Df_Total[len(Df_Total) - len(TestData) - 60:].values  
          # We need to Reshape  
          inputs = inputs.reshape(-1,1)
```

Fig 7: Testing the data

The figure 7 represents the test data where the unseen data that tends to test the model. The training and testing helps to evaluate the performance and progress of the algorithms. This is a subset that gives the objective evaluation of a final model [15]. Thus the future predictions done by the predefined data are represented in figure 8.

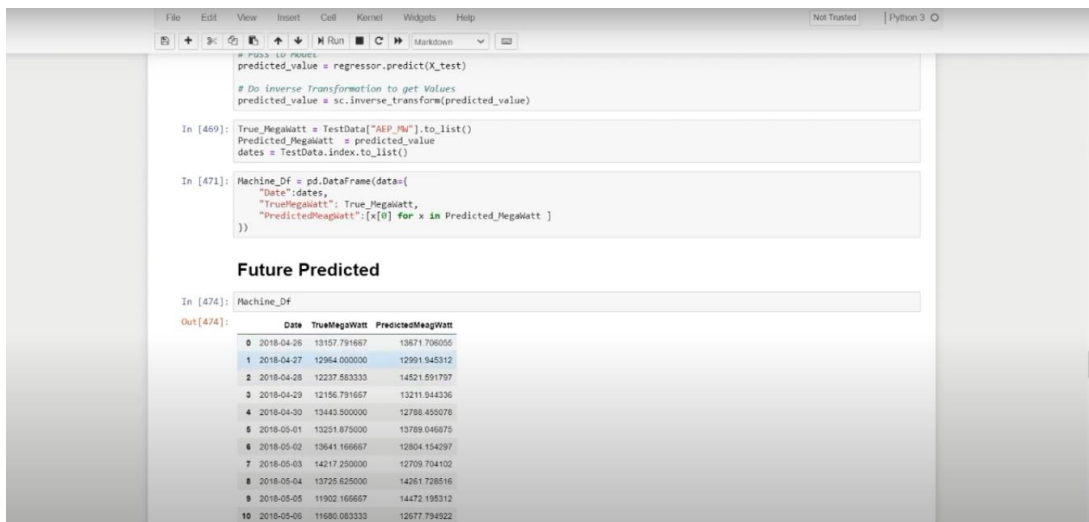


Fig 8: Future prediction

The future prediction in the machine learning is done through the past data collection. This is more important to make predictions that are necessary to match the realistic values. The grouping of three data set in the machine learning system are termed as clustering. The clustering tends to rely on the unsupervised machine learning. The clustering helps to identify the particular group of data with the necessary informations at a particular period of time respectively. The figure 9 represents the clustering of data based on the energy consumption over a particular intervals of time.

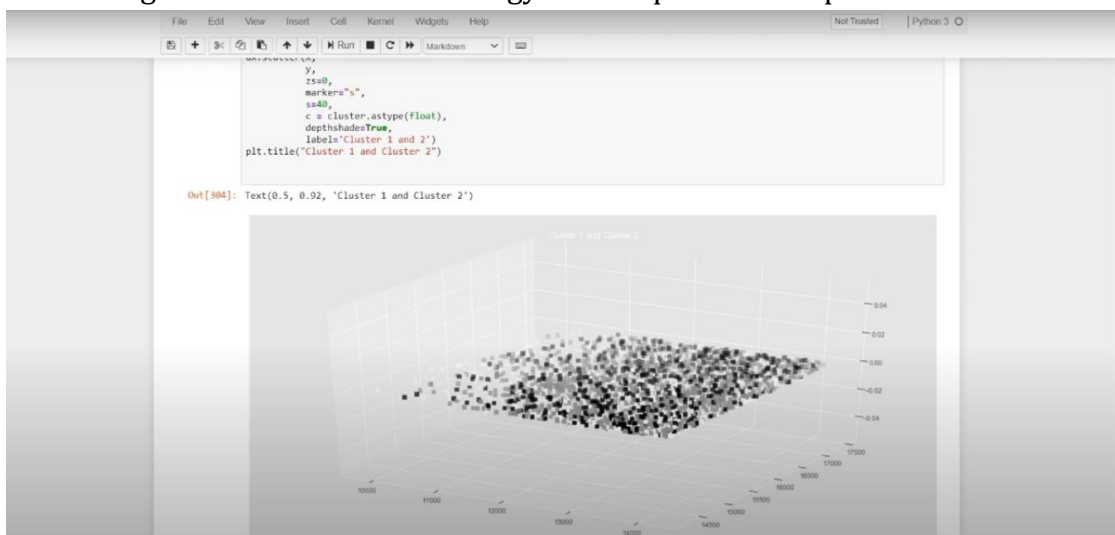


Fig 9: Clustering of data

IV.Simulation results

The analysis and prediction of the energy consumption is done by using the genetic algorithm is depicted as the future predicted waveforms as shown in figure 10. The predefined data are tested and trained to obtain the desired exact output.

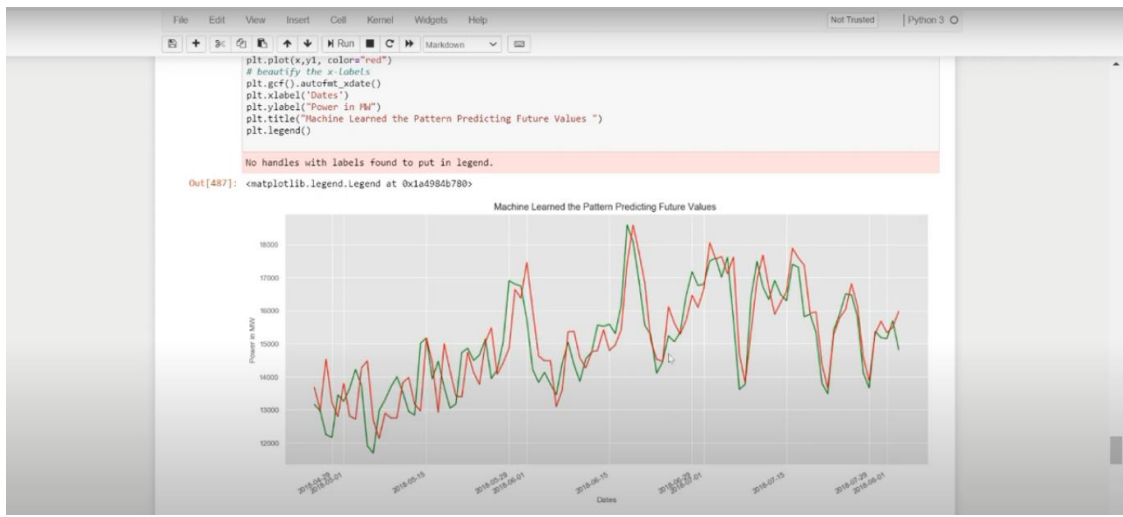


Fig 10: Output result

V.Conclusion

The important aspect of the proposed system is to obtain a energy consumption analysis and prediction in rise through automatically. The system is enabled with the machine learning incorporated with the genetic algorithm to obtain the exact results. The data are trained and tested to represent the optimum output values. Thus by using the machine learning model with genetic algorithm helps to obtain and predict th eincrease in th eenergy consumption with real time implementation.

References

[1] Ounifi, Hibat-Allah, et al. "A deep neural network based approach to energy efficiency analysis for cloud data center." 2019 IEEE 17th International Conference on Industrial Informatics (INDIN). Vol. 1. IEEE, 2019.

[2] Wang, Yixuan, Kun Wang, Huawei Huang, Toshiaki Miyazaki, and Song Guo. "Traffic and computation co-offloading with reinforcement learning in fog computing for industrial applications." IEEE Transactions on Industrial Informatics 15, no. 2 (2018): 976-986.

[3] Lu, Bingqian, et al. "Automating deep neural network model selection for edge inference." 2019 IEEE First International Conference on Cognitive Machine Intelligence (CogMI). IEEE, 2019.

[4] Nejad, Seyed Morteza Mirhoseini, Ghada Badawy, and Douglas G. Down. "EAWA: Energy-aware workload assignment in data centers." 2018 International Conference on High Performance Computing & Simulation (HPCS). IEEE, 2018.

[5] Li, Lianpeng, et al. "Sla-aware and energy-efficient vm consolidation in cloud data centers using host states naive bayesian prediction model." 2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social

Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom). IEEE, 2018.

[6] Chen, Jienan, et al. "iRAF: A deep reinforcement learning approach for collaborative mobile edge computing IoT networks." IEEE Internet of Things Journal 6.4 (2019): 7011-7024.

[7] Dartois, Jean-Emile, et al. "Investigating machine learning algorithms for modeling ssd i/o performance for container-based virtualization." IEEE transactions on cloud computing 9.3 (2019): 1103-1116.

[8] Duc, Thang Le, et al. "Machine learning methods for reliable resource provisioning in edge-cloud computing: A survey." ACM Computing Surveys (CSUR) 52.5 (2019): 1-39.

[9] Lin, Yun, Keping Zhou, and Jieli Li. "Application of cloud model in rock burst prediction and performance comparison with three machine learning algorithms." IEEE Access 6 (2018): 30958-30968

[10] Jiang, Chunxiao, et al. "Machine learning paradigms for next-generation wireless networks." IEEE Wireless Communications 24.2 (2016): 98-105..

[11] Gilda, Shlok. "Notice of violation of IEEE publication principles: Evaluating machine learning algorithms for fake news detection." 2017 IEEE 15th student conference on research and development (SCOReD). IEEE, 2017.

[12] Ye, Hao, et al. "Machine learning for vehicular networks: Recent advances and application examples." ieeve vehicular technology magazine 13.2 (2018): 94-101.

[13] Szott, Szymon, et al. "Wi-Fi Meets ML: A Survey on Improving IEEE 802.11 Performance With Machine Learning." IEEE Communications Surveys & Tutorials 24.3 (2022): 1843-1893.

[14] Chentharra, Shekha, et al. "Security and privacy-preserving challenges of e-health solutions in cloud computing." IEEE access 7 (2019): 74361-74382.

[15] Kobusińska, Anna, et al. "Emerging trends, issues and challenges in Internet of Things, Big Data and cloud computing." Future Generation computer systems 87 (2018): 416-419.