Int. J. Nonlinear Anal. Appl. 15 (2024) 2, 87–94
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2023.30013.4313

Deep learning for big weather data analyzing and forecasting

Delveen L. Abd Al-Nabi[a,*], Shereen Sh. Ahmed[b]

[a]Department of Economics, College of Economics and Administration, Duhok University, Duhok, Kurdistan Region, Iraq

[b]Department of Computer Science, Faculty of Science, Zakho University, Duhok, Kurdistan Region, Iraq

(Communicated by Ehsan Kozehgar)

Abstract

Weather prediction is vital in daily life routines, for risk mitigation and resource management such as flood risk forecasting. Quantitative prediction of weather changes depends on different parameters such as rainfall time, temporal, barometric pressure, humidity, precipitation, solar radiation and wind. Therefore, a highly accurate system or a model to forecast the highly nonlinear changing happening in the climate is required. The focus of this research is direct prediction of forecasting from weather-changing parameters, the forecasts are performed using collected data values recorded in a big dataset (the dataset collects the weather parameter data of the Canary Islands (Las Palmas, Tenerife a Palma, Fuerteventura, La Gomera, Lanzarote and Hierro). The forecasting system is performed by proposing a deep learning approach (CNN). The research goal is predication the weather condition. The acquired classification accuracy for the climate condition using CNN (ShuffleNet) structure is 98%, and the recall and Precision results are 97.5 and 96.9 respectively.

Keywords: Deep learning, big data analysis, weather predication, machine learning
2020 MSC: 68T07, 68T09

1 Introduction

Weather changing prediction is the process of using science and modernization to understand the behavior of the air for certain range of area and time period. Climate estimates are performed by collecting considerable and recordable data (big data) about the current state of affairs changes of the circumference to expand how weather will alter. Climate information provide a wide range of good material to learn an accurate Artificial intelligent system or model [3, 2, 5].

Digital transformation has led to a massive explosion in the volume of data and the emergence of what is known as Big Data (BD), which is revolutionizing business as Company & McKinsey expects that global data by 2520 will be about 125 trillion gigabytes, and until this huge amount is exploited from the data, artificial intelligence (AI) is relied upon to analyze it, as it has become the dominant feature of many modern organizations using technology or dealing with companies technology-based in obtaining services, in order to run its business, process its data, and improve the quality of its services [4].

Big data analysis (BDA) and artificial intelligence (AI), where smart technologies and programs based on intelligence Artificial intelligence and its offshoots, especially machine learning (ML), extract huge amounts of data for better understanding analyzing and validation of big data [7, 10].

*Corresponding author
Email addresses: delveen.luqman@uod.ac (Delveen L. Abd Al-Nabi), shereen.ahmed@uoz.edu.krd (Shereen Sh. Ahmed)

Received: January 2023 *Accepted:* March 2023

The science of data analysis is one of the sciences that constitute the mainstay of computer science and technical engineering for the big data analysis process, it relies on advanced methods and techniques to take raw data, store it, and process it to extract value from them and reveal relationships and trends to help form ideas related to the primary goals of any activity, and to do with improvement initiatives.

Big data analytics is defined as "the use of techniques to examine and process Vast amounts of data to reveal trends and relationships that help to understand activities and functions, make predictions and make predictions on the results obtained.

Building Decisions, also can be defined as a process of group analysis big data. Big data analysis has many scales to work on it, it is used to discover patterns correlations, trends, preferences, and other valuable information, that could not previously be analyzed using traditional tools [11, 12].

Data Analytics refers to the advanced analytical methods, processes, and techniques used to understand processing large sets of data, generating and disclosing useful information, and drawing conclusions to support decision-making. The big data analyzing required many processes to acquired better understanding and data evaluation. These processing steps are:

- Data management: the process of acquisition, arranging, advocating and processing data to ensure accessibility.
- Data mining: the process that discover and establish the link, correlation and trend between different parameters.
- Predictive Analytics: Analytics that aim to predict the future outcomes of a set of data entries, including Statistical algorithms and machine learning techniques.
- Text analytics: analyze text data by combing through text from websites and, documents, s media and other text sources to revenue a useful awareness for the user. The text analysis process depends dramatically on machine learning and natural language processing technology.

There are several definitions of artificial intelligence as it is defined as "a computer system that has the ability to imitating human behavior and intelligence to perform tasks" (Bernard 2018, Bernard), and in another definition artificial intelligence "refers to machines that perform tasks that require a form of intelligence, which usually refers to learning, knowing, sensing, thinking, creating and achieving goals, and generating In artificial intelligence, it is based on techniques such as machine learning (ML) and deep learning.

AI algorithms can learn and manipulate massive amounts of data that will enhance and transform operations in various fields domains effectively during a given period of learning and understanding, AI technology can anticipate needs and make decisions Informed and relevant after data analysis as opposed to business intelligence technology which analyzes data and leaves out the decision-making parts. From the above, it can be said that artificial intelligence (AI) refers to computer systems or devices and machines that It simulates human intelligence to perform tasks, and has the ability to learn, think, understand, adapt, infer and make decisions quickly and efficiently [13, 15, 17].

Machine learning (MI) algorithms are part from Artificial intelligence science. Machine learning (ML) is the study of computer algorithms that automatically improve through experience [3]. Machine learning algorithms worked on building a mathematical model based on a data parameter values, known as "training data," so as to carry out predictions or decisions.

Machine learning is can be define as statistical commutating process, in other word making predictions using computers. The study of Mathematical Optimization produces theories, algorithms and application areas in the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. Machine learning in its application across business issues is also referred to as predictive analytics [5, 7].

The main contrition of this research analysis and understand big data (weather forecasting data) using the AI algorithms and machine learning specifically the deep learning algorithms. An automatic accurate weather forecasting system is produced.

2 Related work

Weather information is big data which contain information that continuously changed with time and area location. Therefore, climate data analyzing, understanding and forecasting is vital research area.

In spite of that, a few research worked on using the Artificial intelligence and machine learning algorithms to handle this type of data effectively.

[20] produced a model to automatic rainfall predication using Deep Belief Network (DBN), they conducted their work on rain measurement samples, a modified DBN with hidden layes (300-200-100-10) performed best with 4.59E-05, 0.0068 and 0.94 values of MSE, RMSE and R value respectively on testing samples. [6] he proposed model using convolutional generative adversarial network 4 years of ERA5 reanalysis data are trained to predict the associated meteorological fields in 2019. He used the video prediction deep learning (VPDL) algorithms for nowcasting predication. They used VPDL PredRNN++ and radar images to predict a sequence of reflectivity images for up to 1-h lead time for São Paulo, Brazil.

[8] produced Sub-seasonal climate forecasting (SSF) model, this model predicting climate variables changes such as temperature and rain in the 2-week to 2-month time scales. Different machine learning algorithms are tested to produce this system. Also, [18] used the neural network for weather predication, their model showed that additional shrouded neurons improved the system's capacity to rid highest request acquired information.

In this research an automatic weather forecasting system is produced using CNN deep learning algorithm, this algorithm used dramatically in many application [16], thus in this research the performance of its weather forecasting is tested.

3 The proposed model

The Fig. 1. illustrates the proposed model block diagram for big weather data predication, in the following section each block and algorithm are explained in details.

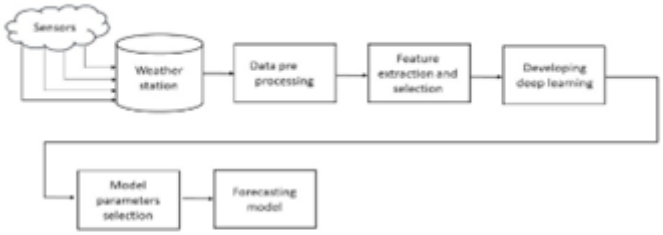

Figure 1: The weather predication proposed system

4 The dataset

The weather dataset collects using electronic sensors these sensors record the changes in weather parameters such as temperature, rain every day in parodic time. In this research a climate data is collected from seven Canary Islands Palmas, La Palma, Tenerife, Gomera, Fuerteventura, Hierro and Lanzarote: `https://www.kaggle.com/datasets/muthuj7/weather-dataset`.

This data contains different parameters such as wind speed, rain temperature and summary for the weather condition according to the parameters changes, in this work, the weather condition predication is the proposed system goal. Fig. 2. Shows some of data set parameters values.

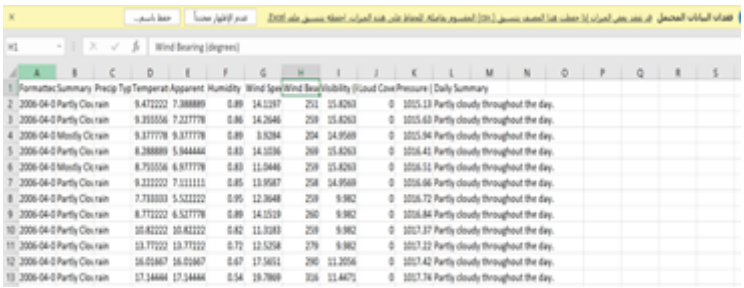

Figure 2: The weather data set

5 Data pre-processing

Since the data contain a number of textual data is a significant process for learning classification algorithms. Text pre-processing is the method that converts texts into an easy appropriate format for classification. Text pre-processing is very important because it can be used to decrease the feature space and computational process, and this, in turn, can positively affect the classification accuracy. Many pre-processing steps are performed to prepare texts for classification. For example, all unnecessary data such as, special characters, symbols, usernames, repeated and elongated letters, numbers, digits, and punctuations are removed. Unnecessary information that often reduces the efficiency and accuracy of the classification algorithm can also be removed. Tokenization is very important in many processing steps to split the text document into tokens based on space, commas, periods, semicolons, and quotes between words. These tokens may be symbols (words and /or numbers). Also, remove unimportant words such as stop words [21].

6 Feature extraction

Feature extraction operation is used to extract meaningful feature attributes from a textual and numerical parameter and make it ready for feeding to machine learning algorithms. This process is known as vectorization because the result of this process is numerical vectors from text attributes.The reason is that machine learning algorithms work on numerical vectors and cannot work directly on raw text data because it includes formats that are not acceptable by these algorithms. Methods of feature extraction are used in order to feed the extracted features into machine learning algorithms for learning patterns that can be implemented on new data points. There are different feature extraction methods such as term frequency inverse document frequency (TF-IDF) and Bag-of-Words (BoW). TF-IDF takes into account the term frequency and inverse document frequency when weighing each term [1, 19].

7 Deep learning

Convolution neural networks is a deep neural network Fig. 3. widely employed in image processing, classification, and recognition according to its great success in the field of computer vision [14, 24]. Recently, many applications have utilized deep learning algorithms, including CNN techniques, like handwritten recognition, automotive, medical and health care applications and telecommunication [9]. CNN transfers the input data, to the form of the multidimensional array. The CNN specifies weights for all neurons depended on the receptive fields' important role. Therefore, it can distinguish the neurons' significance from one another [23]. The basic operations in CNN are convolution and pooling. In convolution operation, many filters extract features from the input data, then construct a feature map that preserves their corresponding spatial information. The pooling operation reduces the feature maps' dimensionality generated from convolutional layers.

This work applied different CNNs models like AlexNet, VGG16, and ShuffleNet. However, both pre-trained VGG16 and AlexNet have not achieved promising results. Thus, ShuffleNet has been utilized as a deep feature extractor.

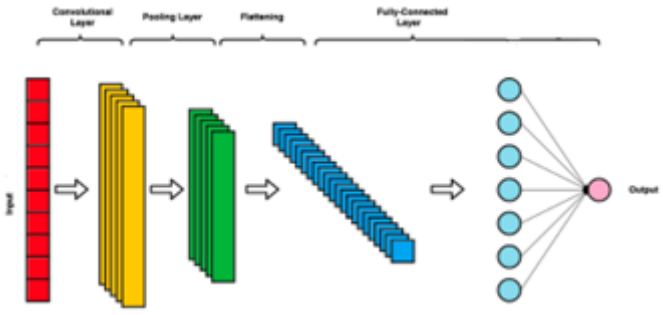

Figure 3: The CNN deep learning

8 Model parameters selection

ShuffleNet is a high-efficient CNN structure which decreases computation cost while sustaining high accuracy. It includes 172 layers and employs, in addition to a depth wise convolution unit, two new operations: channel shuffle and

pointwise group convolution. The pointwise group convolution decreases the computation complexity of convolution layers. The channel shuffle operation allows communicating between feature maps of different convolution groups. See Fig. 4. ShuffleNet unit is a residual block which includes the bottleneck structure. In its residual path, a channel shuffle operation follows 1x1 pointwise group convolution (GConv) to produce ShuffleNet unit. Then, 3X3 depthwidth convolution (DWConv) is built followed by another pointwise group convolution to improve the channel dimension to fit the shortcut branch. In case of applying stride to ShuffleNet unit, 3X3 average pooling (AVG Poo) is added on the shortcut path and a depth concatenation is used instead of element addition to increase channel dimension with slight additional computation cost. SuffleNet also contains batch normalization and nonlinearity, which is used in different CNNs (like ResNet and ResNeXt) to normalize the prior layers and increase the stability of the network during training. According to ShffleNet structure, it achieves higher speed than Alex Net and VGG and maintains comparable or better accuracy [24].

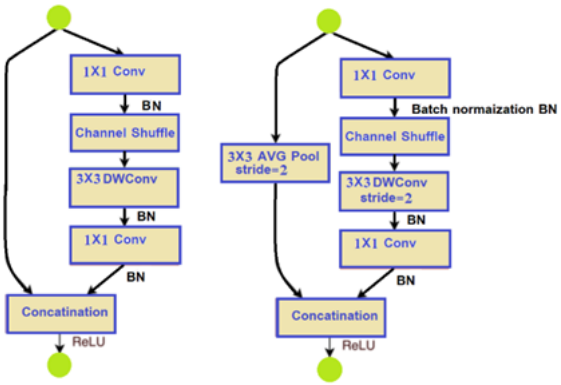

Figure 4: Convolution Neural network ShuffleNet unit basic structure

9 Forecasting model

The final model for weather condition forecasting using CNN (ShuffleNet) structure, after selecting the layers (filters) parameters is shown in Fig. 5.

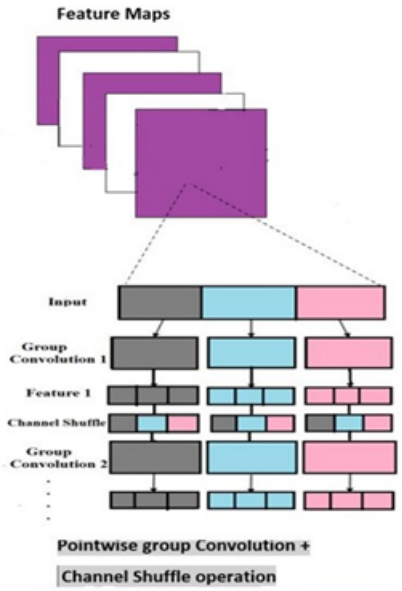

Figure 5: CNN (ShuffleNet) structure parameters

10 Experimental results

This research has utilized a confusion matrix to calculate the performance of the model. Each column in the matrix represents a class label, and each row represents the samples predicted by the classifier. The diagonal values represent the true positives of each class label, which is the number of samples correctly predicted (TP). The off-diagonal values represent misclassifications. The accuracy of the multi-class model can be calculated by equation 10.1.

$$Accuracy = \frac{the\ number\ of\ true\ predictions\ (the\ diagonal\ \ elements\ of\ matrix)}{the\ total\ predictions} \tag{10.1}$$

Other evaluation measures were computed, such as *precision* and *recall*, to calculate the individual class performance. The precision rate, also called positive predicted value, is the ratio of the correct predicted values of a specific class to the total predictions, as seen in equation 10.2. The precision of each class label is the rate of dividing the confusion matrix's diagonal value of that class by the sum of values of the corresponding row.

$$\boldsymbol{Precision - Class_i} = \frac{TP - Class\ i(number\ of\ samples\ in\ Class\ i\ predicted\ correctly)}{Total\ predicted\ as\ Class\ i(TP - Class\ i + FP - Class\ i)} \tag{10.2}$$

FP-Classi is the number of samples incorrectly predicted as Class*i*.

Recall, also called True Positive Rate (TPR) or sensitivity, is another measure to calculate the classifier's performance. It is the ratio of true positives of a specific class to the total samples of that one (sum of corresponding column's values); see equation 10.3.

$$\boldsymbol{Recall - Class_i} = \frac{TP - Class\ i(number\ of\ samples\ in\ Class\ i\ predicted\ correctly)}{Total\ elements\ of\ Class\ i(TP - Class\ i + FN - Class\ i)} \tag{10.3}$$

FN-Classi is the number of samples of Classi incorrectly predicted as elements of other classes [22].

The acquired classification accuracy for the weather condition is 98%, the recall and Precision results are 97.5 and 96.9 respectively.

11 Discussion

In this paper different CNN deep learning structures are tested (Alex Net, VGG16, and ShuffleNet). However, both pre-trained VGG16 and Alex Net have not achieved promising results. Thus, ShuffleNet has been utilized as a deep feature extractor. The goal of this research is predication the weather condition (Partly Cloudy, Mostly Cloudy, Overcast, Foggy Or clear). The proposed system handles the predication for these six states effectively (accuracy reached 98%). The data set parameters are numbers and text, this big data is called multiparameter structure data. It is more difficult to handle than the data which contain numbers only. Thus, in this work methods for text handling and waiting is implemented and this increase the predication accuracy dramatically from 69% without text preprocessing to 98%. Most the state of art research [20, 6] used the numerical data for forecasting certain weather parameters such as rain or temperature. Consequently, the proposed system in this research opens the door to handle complex structure data with different parameters types effectively.

12 Conclusion

The paper shows weather condition changes forecasting using deep learning algorithms the proposed system is conducted on weather data collected from a specific site in Canary Islands. Deep learning algorithms are expediting dramatically in nearly every area of science and replacing several other machine learning algorithms. Consequently, this gave us trend to further investigate these deep algorithms over time periods weather changes forecasting. Thus, the fundamental focus of this research was to perform forecasting for weather conditions. Big data with multiparameter (numbers and text) which its much more challenging than one type or numerical parameter data are utilized. During this research, it was observed that the parameters which are required to predict weather condition were extremely complex and tender for a short-term duration. Thus, variety CNN algorithm structures and statistical text features handling were investigated. In attention of the future works, we trust that results of this research can be used as principle for move forward forecasting of weather changes parameters with different climate conditions.

References

[1] A. Aichert, *Feature extraction techniques*, Camp. Med. Seminar Ws0708, 2008, pp. 1–8.

[2] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, *Disease prediction by machine learning over big data from healthcare communities*, IEEE Access **5** (2017), 8869–8879.

[3] X.-W. Chen and X. Lin, *Big data deep learning: Challenges and perspectives*, IEEE Access **2** (2014), 514–525.

[4] L. Deng and D. Yu, *Deep learning: Methods and applications*, Found. Trends Signal Process. **7** (2014), no. 3-4, 197–387.

[5] L. Deng, D. Yu, and J. Platt, *Scalable stacking and learning for building deep architectures*, IEEE Int. Conf. Acoustics Speech Signal Process. (ICASSP), IEEE, 2012, pp. 2133–2136.

[6] C.J. Devi, B.S. Prasad Reddy, K.V. Kumar, B.M. Reddy, and N.R. Nayak, *Ann approach for weather prediction using backpropagation*, Int. J. Engin. Trends Technol. **3** (2012), no. 1, 19–23.

[7] A. Efrati, *How "deep learning" works at Apple, beyond*, `https://www.theinformation.com/articles/how-deep-learning-works-at-apple-beyond`, 2017.

[8] M. Elhoseiny, S. Huang, and A. Elgammal, *Weather classification with deep convolutional neural networks*, IEEE Int. Conf. Image Process. (ICIP), IEEE, 2015, pp. 3349–3353.

[9] H. Gholamalinezhad and H. Khosravi, *Pooling methods in deep neural networks, a review*, ArXiv Preprint ArXiv:2009.07485 (2020).

[10] Á.B. Hernández, M.S. Perez, S. Gupta, and V. Muntés-Mulero, *Using machine learning to optimize parallelism in big data applications*, Future Gen. Comput. Syst. **86** (2018), 1076–1092.

[11] B. Hutchinson, L. Deng, and D. Yu, *Tensor deep stacking networks*, IEEE Trans. Pattern Anal. Machine Intell. **35** (2012), no. 8, 1944–1957.

[12] N. Jones, *The learning machines*, Nature **505** (2014), no. 7482, 146.

[13] A. Krizhevsky, I. Sutskever, and G.E. Hinton, *Imagenet classification with deep convolutional neural networks*, Advances in Neural Information Processing Systems, vol. 25, 2012.

[14] Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature **521** (2015), no. 7553, 436–444.

[15] C. Ma, H.H. Zhang, and X. Wang, *Machine learning for big data analytics in plants*, Trends Plant Sci. **19** (2014), no. 12, 798–808.

[16] A. Mathew, P. Amudha, and S. Sivakumari, *Deep learning techniques: an overview*, Int. Conf. Adv. Machine Learn. Technol. Appl., 2020, pp. 599–608.

[17] R. Raina, A. Madhavan, and A.Y. Ng, *Large-scale deep unsupervised learning using graphics processors*, Proc. 26th Ann. Int. Conf. Machine Learn., 2009, pp. 873–880.

[18] Sebastian Scher and Gabriele Messori, *Weather and climate forecasting with neural networks: Using general circulation models (GCMs) with different complexity as a study ground*, Geosci. Model Dev. **12** (2019), no. 7, 2797–2809.

[19] W. Scott, *TF-IDF for Document Ranking from scratch in python on real world dataset*, `https://towardsdatascience.com/tf-idf-for-document-ranking-from-scratch-in-python-on-real-world-dataset-796d339a4089`, 2020.

[20] A. Subashini, S.M. Thamarai, and T. Meyyappan, *Advanced weather forecasting prediction using deep learning*, Int. J. Res. Appl. Sci. Engin. Technol. **7** (2019), 939–945.

[21] F. Sun, A. Belatreche, S. Coleman, T.M. McGinnity, and Y. Li, *Pre-processing online financial text for sentiment classification: A natural language processing approach*, IEEE Conf. Comput. Intell. Financ. Engin. Econ., 2014, pp. 122–129.

[22] G. Varoquaux, P.R. Raamana, D.A. Engemann, A. Hoyos-Idrobo, Y. Schwartz, and B. Thirion, *Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines*, NeuroImage **145** (2017), 166–179.

[23] Rikiya Yamashita, Mizuho Nishio, Ron K. Do, and Kaori Togashi, *Convolutional neural networks: an overview and application in radiology*, Insights Imag. **9** (2018), no. 4, 611–629.

[24] X. Zhang, J. Zhao, and Y. LeCun, *Character-level convolutional networks for text classification*, Adv. Neural Inf. Process. Syst. **28** (2015).