Hybrid Deep Learning Algorithms for Big Output Spaces: A Comparative Analysis

Dr. Jitendra Sheetlani [1], Dr. J. P. Patra[2], Bhimsen Moharana [3]

[1]*Research Guide, Dept. of Computer* Science

Sri Satya Sai University of Technology and Medical Sciences,

Sehore Bhopal-Indore Road, Madhya Pradesh, India.

[2]*Research Co-Guide, Director. Dept. of* Computer Science *& Engineering*

Dept. of CSE, CSVTU University, Chhattisgarh

[3]*Research Scholar, Dept. of* Computer Science

Sri Satya Sai University of Technology and Medical Sciences,

Sehore Bhopal-Indore Road, Madhya Pradesh, India

Abstract

Classification of imbalanced data is an important research problem as lots of real-world data sets have skewed class distributions in which the majority of data instances (examples) belong to one class and far fewer instances belong to others. While in many applications, the minority instances represent the concept of interest (e.g., fraud in banking operations, abnormal cell in medical data, etc.), a classifier induced from an imbalanced data set is more likely to be biased towards the majority class and show very poor classification accuracy on the minority class. Despite extensive research efforts, imbalanced data classification remains one of the most challenging problems in data mining and machine learning, especially for multimedia data. To tackle this challenge, in this paper, we propose an extended deep learning approach to achieve promising performance in classifying skewed multimedia data sets. Specifically, we investigate the integration of bootstrapping methods and a state-of-the-art deep learning approach, Convolutional Neural Networks (CNNs), with extensive empirical studies. Because deep learning approaches such as CNNs are usually computationally expensive, we propose to feed low-level features to CNNs and prove its feasibility in achieving promising performance while saving a lot of training time

Keywords: Deep Learning Algorithms , Big Output Spaces , supervised and unsupervised learning.

Introduction

"Big Data" and "Deep Learning" are two things that data scientists pay a lot of attention to in the fast-changing digital world of today. The challenge of managing and analysing large amounts of unstructured digital raw data is called "Big Data," and its catchy acronym sums up what it is all about. When putting together large amounts of data, the needs of the business must be taken into account. This change is caused by the fast growth of digital data in all of its many shapes, sizes, and forms. Most of the world's largest technology companies, many of which have collected huge amounts of data, have storage systems that can hold exabytes. Due to the fact that billions of people use social media sites like YouTube, Twitter, and Facebook, a huge amount of data is created. On the other hand, traditional technologies can't handle such a huge amount of information. [Needs citation] Several companies have made products that use this technology, and the goal of this section is to show some of the many business uses that are now possible because big data analytics is available. The main goal of large-scale data analysis is to find important patterns in the huge amounts of data that have been collected. On the other hand, Big Data Analytics faces new problems with data analysis and machine learning. For example, [1] [2], these problems include the need for many different types and sizes of data input, as well as the need for a fast and accurate data stream, accurate data processing, and high-quality data. Also, different types and sizes of data must be entered in different ways. Big Data makes it hard to collect and store valuable information because it can affect almost every part of modern life. To keep up with the fast growth of hidden information in non-traditional data, we need new ways of doing things. To be successful at this, you will need to put together a group of experts from many different fields. The goals of big data analytics can only be reached with the help of machine learning and the power of computers. When it came to machine learning, the biggest problems were how to code the data it got and how to use the patterns it learned in new situations. Both of these were problems that needed to be fixed right away. The results that machine learning algorithms come up with depend a lot on how the data are represented. Getting the data to a machine learner in the right format can help it do its job better. On the other hand, if the machine learner is complicated and hard to use, it might not work right if the data are not shown in the right way. The process of feature engineering, which is an important part of machine learning, is used to build features and describe raw input data. Adding new features takes a lot of time and work, and the target market is usually the one who decides which new features should be added. Machine learning can be used in a wide range of business fields, such as medicine, the Internet of Things (IoT), search, and many others. Deep learning is an important part of machine learning that is needed to do big data analytics on very large datasets [4]. Traditional teaching methods focus on learning that is only superficially structured and don't make use of more in-depth experiences. Deep learning, on the other hand, uses both supervised and unsupervised machine learning techniques to automatically classify features by building hierarchical data representations. This is not the case with traditional ways of teaching, which focus more on structures that seem to be set up in a certain way. Deep learning is becoming more and more popular in many fields, such as health, computer vision, and speech recognition. The brain is a great tool, especially when it comes to figuring out what signals mean. When it comes to making artificial intelligence, the human brain is a great place to look for ideas. To put it another way, Facebook, Apple, and Google are the best at getting and analysing large amounts of digital data. For example, the iPhone's operating system has a digital version of Siri, the personal assistant software made by Apple. Siri uses a technology that Apple has patented called "deep learning" to do tasks and get information from users. You can also use it to send money, check the news and weather, set an alarm, and even change the lighting in the room with the touch of a button. As you use it more, it will learn more about your preferences and be able to meet your needs before you even ask for them. Deep learning algorithms are at the heart of many popular services, like Google Translate, searches for images and videos, and Android's speech recognition. Deep learning is used by companies like Microsoft and IBM in their day-to-day business operations.

Related work

Classes of algorithms or models for classifying data that are imbalanced can be separated from classes of data manipulation. As a result of these studies, deep learning algorithms for imbalanced data categorization can be developed.

When it comes to categorization of unbalanced data, algorithmic and model-based approaches spent much of their time investigating and perfecting their training algorithms. Cost-sensitive learning techniques, for example, aim to improve classification performance by maximising the loss functions associated with a data set. This approach of learning was prompted by the realisation that most real-world applications do not have consistent prices for misclassifications. For these approaches to work, it is necessary to build a cost matrix based on data and then apply it to learning to determine the actual costs associated with each type of error. When it comes to altering the bias of a machine to favour minorities, the idea of cost-conscious learners comes to mind. Many studies have shown their ability to improve classification accuracy when faced with unbalanced data, but none of those studies were systematic or comprehensive. Over the past few years, deep learning, which is based on research on artificial neural networks, has garnered considerable research efforts in a wide range of domains. Some examples of these domains are signal and information processing, machine learning, and artificial intelligence. There have been numerous approaches to deep learning investigated as a result of this You can find examples of these in the Deep Belief Network (DBN) and Boltzmann Machine (BM) categories, as well as the Restricted Boltzmann Machines (RBM) and Deep Boltzmann Machines (DBM) and Deep Neural Networks (DNN). The article provides a more in-depth look at the most recent research in deep learning. For example, a discriminative deep architecture known as the convolutional neural network (CNN), which falls within the DNN category and has achieved the state of the art in computer vision and image recognition, is one of these neural networks. Two layers make up the convolutional and pooling layers of each CNN module. These modules are often stacked one on top of the other in order to form a deeper model. The pooling layer reduces the data rate from the layer below by subsampling the output of the convolutional layer, which shares many of its weights. It is still unclear how well CNNs perform on data sets that are very unbalanced despite their promising results in a variety of applications.

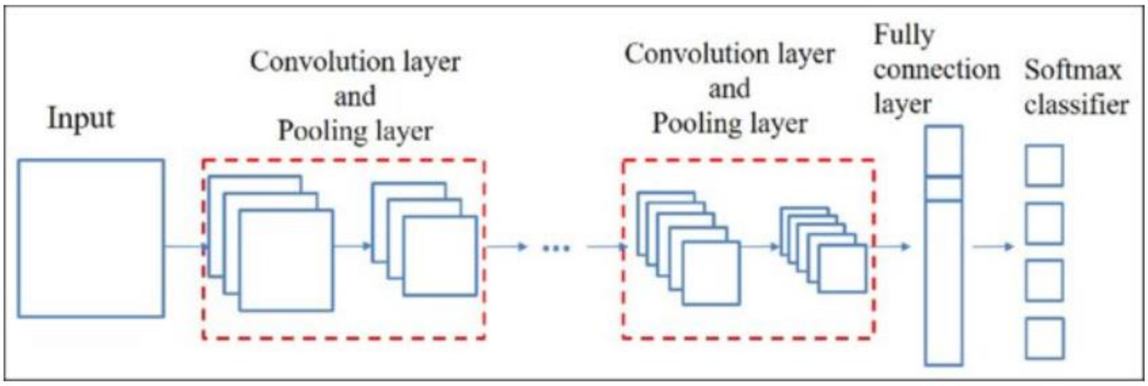

Figure 1: CNN Model

As a result, the goal of this research is to find out how well CNNs classify imbalanced data and, more importantly, how to expand them in order to increase their overall performance.. To be more explicit, we propose that CNNs be enhanced by using a bootstrapping sample method that is specifically tuned to meet the unique characteristics of CNNs.. To boost CNN's performance on multimedia data classification, we've developed a bootstrap sampling strategy that involves oversampling and decision fusion. However, the negative bootstrap method, which combines random sampling and adaptive selection to iteratively uncover relevant negatives, is in contrast to this approach The negative bootstrap method, on the other hand, uses an iterative process to uncover relevant negatives.

He, Lu et al[1] Researchers in the fields of multimedia computing and machine vision have been stymied for a long time by the challenge of action recognition. The most recent developments in deep learning, together with layered convolutional Independent Subspace Analysis, have been applied to a great number of data sets that are freely available to the public (ISA). Unfortunately, one of the most significant challenges associated with the adoption of this innovative deep learning-based strategy on a broad scale is the excessive latency that is involved in training with high-dimension data. In this work, we present a brand new hardware accelerator that may be used for action identification that is based on deep learning. We propose to speed up the convolutional stacking ISA approach, which is one of the essential components of deep learning-based action recognition systems. [Citation needed] [Citation needed]

C. -H. Huang et al[2] The goal of this research is to make hardware deployment more flexible and to study the dependability of hardware utilisation for virtual machines (VM). The Apriori algorithm and deep learning technologies are at the core of the new method described in this research. With this method, dependable hardware resources can be used more efficiently.

S. Nandy et al.[3] Develop a deep learning model that combines a bidirectional LSTM layer with a CatBoost Algorithm layer in order to recognise tweets and recognise signs of depression. Propose this model. The results of this research show that the proposed model works better than existing approaches to classification using machine learning, and they also show that the prevalence of depression has dramatically grown since the beginning of the epidemic. The most important contributions of the study are the deep learning models and their capacity to detect depression.

J. Zhang et al.,[4] Develop a deep learning model that combines a bidirectional LSTM layer with a CatBoost Algorithm layer in order to recognise tweets and recognise signs of depression. Propose this model. The results of this research show that the proposed model works better than existing approaches to classification using machine learning, and they also show that the prevalence of depression has dramatically grown since the beginning of the epidemic. The most important contributions of the study are the deep learning models and their capacity to detect depression.

M. N. Al-Andoli, et al[5] As a first step toward devising workable solutions to the challenges posed by DL-based CD techniques, this paper examines relevant studies on DL that combine parallel and

MH approaches. There have been many challenges in CD research, and we offer some proposals for future research avenues as well.

Proposed methodology

One kind of neural network is a convolutional neural network. These deep learning models, also called convolutional neural networks (CNNs), are adaptations of multilayer perceptions that use the least amount of preprocessing and are based on two ideas. First, each hidden unit only has connections for a small number of input devices. This keeps the total number of hidden devices low. This makes sure that there aren't too many hidden devices (called feature maps in CNNs). The idea for these confined networks came from the discovery that neurons in the visual cortex have receptive fields that are limited to a certain area. Another option is to use the fact that natural images don't move as a way to reduce the amount of work needed to process them. To do this, you would have to use the fact that images of nature don't move. This means that the numbers are the same no matter where you are in relation to the picture or where you choose to look at it. We can combine the features that were learned over small patches that are randomly chosen from the larger image to get a different feature activation value for each place in the image. This will help us get a better idea of what the image looks like. Once the process of convolution is done, we can either use the features directly or combine their statistics to classify the data. It is not uncommon for aggregated statistics to not only improve the results but also have a lower dimensionality (compared to using all of the features that were collected) (less over-fitting). There are three different layers in a convolutional network: the convolutional layer, the pooling layer, which is the second most common, and the full connectivity layer. Most of the time, the convolutional layer is used (the third).

1) Convolutional layer: A convolutional layer is constructed out of feature maps, which serve as its fundamental elements. According to Equation (1), the lth layer's feature map is created by feeding the feature maps of the preceding layer into an activation function f that has learnable kernels and additive bias. This process results in the creation of the feature map for the lth layer. In this particular scenario, the first layer is comprised of the input data, and the activation function f is often the logistic (sigmoid) function. A collection of input maps makes up the second layer.

$$X_j^l = f\left(\sum_{i \in M_j} X_i^{l-1} * K_{ij}^l + b_j^l\right), l \geq 2;$$
$$X_j^l = f\left(\beta_j^l pool(X_j^{l-1}) + b_j^l\right), l \geq 2.$$

2) Pooling layer: According to the aforementioned equation, a pooling layer is obligated to generate downsampled versions of the feature maps that it is provided with as input (2). The notation represents the multiplicative bias, denotes the additive bias, and pool(.) denotes a pooling operation that, in general, computes the aggregated statistics of the input maps, such as their mean or maximum values. This particular scenario represents the multiplicative bias, and denotes the additive bias. To put it another way, in this context, symbolises the additive bias, while represents the multiplicative bias. The layer may be referred to as either mean pooling or max pooling, or it may be referred to as something entirely else, depending on the technique of pooling that was used. However, it may also be referred to as something else entirely. The vast majority of the time, this layer is added after each and every convolutional layer that was already present in the chain that led up to it.

3) Fully-connected layer: The neural network advances to the fully connected layers after a series of convolutional and pooling layers. These layers are responsible for higher-level reasoning. Every single neuron in a fully connected layer is connected to every single neuron in the layer below it — regardless of whether those neurons are fully connected, pooling, or convolutional.

Results Analysis

This experiment makes use of the IACC. data set that was provided by the TRECVID 2011 benchmark. The goal of the experiment is to recognise the semantic notion that is contained within a video shot. Such a technology may be essential for video retrieval, categorization, and other video exploitations. High-level semantic concepts are all referred to by the words "car," "road," and "tree," respectively. It struggles with problems such as data imbalance, inability to scale, and semantic mismatch, to name a few of them. Figure 3 displays three keyframes, each of which has a notion labelled next to it. When applying traditional deep learning approaches like CNNs, the TRECVID dataset's poor performance can be attributed to a number of problems, including underfitting, a huge diversity of features, noisy and insufficient data annotation, and so on. Because different testing concepts have varying degrees of data imbalance, it is possible that a constant batch size is not the best choice for all of the ideas represented in the TRECVID data set. As a consequence of this, the batch size is dynamically determined by the number of positive training instances contained inside the training set. In the course of our experiment, we made use of a batch size that was twice as large as the number of positive training samples.

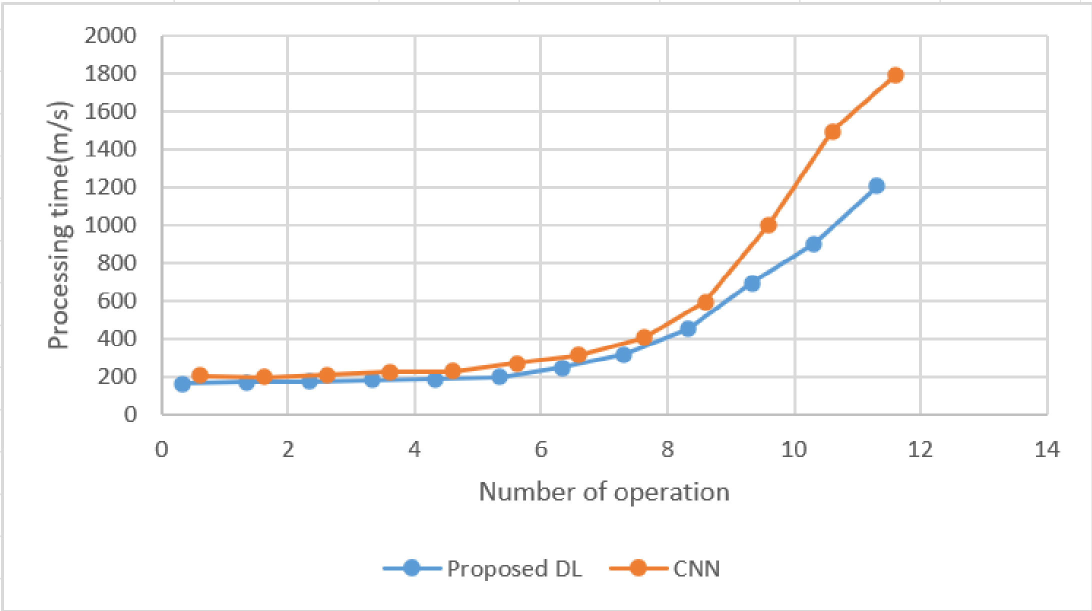

Figure 2: Hybrid Deep Learning Algorithms for Big Output Spaces

In the TRECVID data set, there are 262,911 instances of bicycles, trees, and politics that are used for training, and 113,046 instances of politics that are utilised for testing. In total, the data set has a total of 333,045 cases. In the context of unbalanced data classification, the importance of recall in comparison to precision is discussed in and the F-score illustrates the trade-off between the two concepts. Our framework's recall and F-score values are compared which was the highest performer in the TRECVID 2011 semantic indexing challenge. As can be seen from the table, our F-scores and recall values are much better than those obtained by t proposed concepts. This is due to the fact that our data is less noisy and incomplete, and as a direct consequence, our F-scores and recall values are far higher. It is also important to point out that the TiTech group can only locate zero or one true positive for each of the fifty concepts, but our framework has an average recall value of approximately 0.3. In a previous experiment using TRECVID's dataset, convolutional neural networks (CNNs) were shown to perform worse than all other classifiers. This finding clearly demonstrates the importance of use bootstrapping to incorporate CNNs into our system for analyzing imbalanced multimedia data.

Conclusion

In the paper, we talk about how to improve the performance of CNNs, a deep learning method, by combining it with a bootstrapping mechanism. During the bootstrapping process, a set of "pseudo-balanced" training batches are made based on how the data set is set up. Then, these batches are put into the CNN so that it can classify them. The results of our experiments show that our proposed framework works well to find multimedia data with a very skewed data distribution. With the help of the TRECVID data set, we were able to get these results. It has also been shown that our extended CNN framework can work well with low-level features, which cuts the training time needed for deep learning by a lot. This is different from most of the deep learning research that has been done so far, which uses raw media data as input.

Reference

[1]. He, Lu; Luo, Yan; Cao, Yu (2014). *[IEEE 2014 IEEE 22nd Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM) - Boston, MA, USA (2014.5.11-2014.5.13)] 2014 IEEE 22nd Annual International Symposium on Field-Programmable Custom Computing Machines - Accelerator of Stacked Convolutional Independent Subspace Analysis for Deep Learning-Based Action Recognition. , (), 104–104.* doi:10.1109/fccm.2014.37 .

[2]. C. -H. Huang, B. -H. Huang and T. -Y. Wu, "Hardware Resource Reliability Analysis based on Deep Learning for Virtual Machine Deployment Optimization," 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), 2020, pp. 726-727, doi: 10.1109/GCCE50665.2020.9292051.

[3]. S. Nandy and V. Kumar, "My Mind is a Prison: A Boosted Deep Learning approach to detect the rise in depression since COVID-19 using a stacked bi-LSTM CatBoost model," 2021 IEEE International Conference on Big Data (Big Data), 2021, pp. 4396-4400, doi: 10.1109/BigData52589.2021.9671482.

[4]. J. Zhang et al., "Robust CAPTCHAs Towards Malicious OCR," in IEEE Transactions on Multimedia, vol. 23, pp. 2575-2587, 2021, doi: 10.1109/TMM.2020.3013376.

[5]. M. N. Al-Andoli, S. C. Tan, W. P. Cheah and S. Y. Tan, "A Review on Community Detection in Large Complex Networks from Conventional to Deep Learning Methods: A Call for the Use of Parallel Meta-Heuristic Algorithms," in IEEE Access, vol. 9, pp. 96501-96527, 2021, doi: 10.1109/ACCESS.2021.3095335.

[6]. W. Liu and T. Zhang, "Multimedia Hashing and Networking," in IEEE MultiMedia, vol. 23, no. 3, pp. 75-79, July-Sept. 2016, doi: 10.1109/MMUL.2016.39.

[7]. Y. Yan, M. Chen, M. -L. Shyu and S. -C. Chen, "Deep Learning for Imbalanced Multimedia Data Classification," 2015 IEEE International Symposium on Multimedia (ISM), 2015, pp. 483-488, doi: 10.1109/ISM.2015.126.

[8]. L. Yuwei, Y. Linyu, L. Xiangyu and C. Long, "Summary of Data Mining and Analysis of Launch Vehicle Test Based on Small Sample Learning Method," 2021 International Conference on

Artificial Intelligence, Big Data and Algorithms (CAIBDA), 2021, pp. 158-161, doi: 10.1109/CAIBDA53561.2021.00040.

[9]. Y. Wang, "Visualization Analysis of Meteorological Big Data through Deep Learning and Network Model," 2021 IEEE 3rd International Conference on Civil Aviation Safety and Information Technology (ICCASIT), 2021, pp. 601-604, doi: 10.1109/ICCASIT53235.2021.9633473.

[10]. K. R. Swetha, N. M, A. M. P and M. Y. M, "Prediction of Pneumonia Using Big Data, Deep Learning and Machine Learning Techniques," 2021 6th International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 1697-1700, doi: 10.1109/ICCES51350.2021.9489188.

[11]. L. Liu, "Deception, Robustness and Trust in Big Data Fueled Deep Learning Systems," 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 3-3, doi: 10.1109/BigData47090.2019.9005597.

[12]. C. Seo, J. Kim and S. Kwon, "A study on modeling using big data and deep learning method for failure diagnosis of system," 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 4747-4751, doi: 10.1109/BigData.2018.8622644.

[13]. I. Shamshurin and J. Saltz, "Will Deep Learning Change How Teams Execute Big Data Projects?," 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 2813-2817, doi: 10.1109/BigData.2018.8622337.

[14]. G. Sismanoglu, M. A. Onde, F. Kocer and O. K. Sahingoz, "Deep Learning Based Forecasting in Stock Market with Big Data Analytics," 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 2019, pp. 1-4, doi: 10.1109/EBBT.2019.8741818.

[15]. T. T. Chhowa, M. A. Rahman, A. K. Paul and R. Ahmmed, "A Narrative Analysis on Deep Learning in IoT based Medical Big Data Analysis with Future Perspectives," 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 2019, pp. 1-6, doi: 10.1109/ECACE.2019.8679200.

[16]. S. Hegde and M. R. Mundada, "A Hybrid Approach of Deep Learning with Cognitive Particle Swarm Optimization for the Big Data Analytics," 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2018, pp. 1-5, doi: 10.1109/ICCCNT.2018.8494200.

[17]. Y. Yang and T. Chen, "Analysis and Visualization Implementation of Medical Big Data Resource Sharing Mechanism Based on Deep Learning," in IEEE Access, vol. 7, pp. 156077-156088, 2019, doi: 10.1109/ACCESS.2019.2949879.