FUTURE TRENDS OF DATA SCIENCE WITH MACHINE LEARNING AND DEEP LEARNING TECHNOLOGIES

K Ramesh
Department of Electrical& Electronics Engineering, Bapatla Engineering College Bapatla, India
rameshbec.nandur@gmail.com

Satya Dinesh Madasu
Department of Electrical& Electronics Engineering, Bapatla Engineering College Bapatla, India
msdinesh.nitjsr@gmail.com

Abstract
In this paper the author discusses the future prospects of data science with emerging technologies. The applications of data science is thoroughly discussed. Machine learning and the related subjects required to go through the applications of machine learning are also discussed. Classification of machine learning algorithms with relevant examples are given. Machine learning algorithms are also introduced. Electro encephalogram is described in detail with examples. Brain computer interface is also clearly explained with examples.
Keywords: *Machine learning, Electro encephalogram, Brain computer interface, supervised learning, reinforcement learning.*

1. Introduction

When somebody is mining the data, they require the knowledge of statistics and machine learning. To know what is the exact behaviour of the data to extract the knowledge on the data it is required to focus on statistics and machine learning. So where it involves in order to extract the information from the data, these two major things are initially needed. There are different types of data. We have 1-dimensional data for example any text that we are seeing. In the newspaper we will get several information. They are all in the form of only data. Suppose we are looking for a hotel, we will see the feedback about the hotel. Example of 2-dimentional data is image that we are seeing. Nowadays everybody is using Facebook or WhatsApp. We will have the two-dimensional image data. Suppose some images are given or we are looking for some person, we are identifying the matching. So, the image data is considered as the two-dimensional data. similarly, there is a multi-dimensional data. Any video data is a multi-dimensional data. Nowadays CCTV is very popular. The analytics will be done on that for example in a particular time there may be some incident or in ATM people may do the theft which is recorded in CCTV. To analyse this nowadays, CCTV is very popular. For example, if it is assumed that we have four days data with us. For complete days data with us and we are looking for one or two persons so they may appear in one or two frames. But four days video we must analyse, and we have to spend four days manually. If we have need to analyse the video, we must spend four complete days. Then in this scenario can we use machine learning?

If multi-dimensional data is with us say a video data, is it possible to mine the data? Can we find the information from these data? similarly in the case of satellite images, satellite image data is a multi-dimensional data. They have these are the form XY image and then third dimension they call a band. The band size may be 128 or 256. It is in the third dimension. It involves only the numbers or the signal processing. Any speech processing is a one-dimensional data. Apart from the text data, any speech for example the signal. When it is stored in one dimensional form, these are all the different data. Whatever problem is going to be solved in machine learning or artificial intelligence, it will be classified into one of these data. For example, it is assumed that we have the speech processing data[1]. It will be exactly like this. And assume that if some Fourier transform or the wavelet transformation is applied to convert this from time domain to frequency domain. Then there will be a frequency domain data. If it is classified as say beta, alpha, gamma and delta something like that which one form of the data. So any problem that is being solved with the help of machine learning will fall under these data. it will fall under one of these data.

If we're dealing with a signal processing, it is a one-dimensional data. If we are dealing with the image processing it is a two-dimensional data. If we are working for a video data or satellite image analysis, it's a multi-dimensional data. And all data mining researchers require data in any form. And nowadays Natural language process (NLP) is very famous which involves the text. There are all the different kinds of data that are being used in day-to-day life. Once the data is given, some tool is required to analyse it. The tools are nothing but mathematics or

statistics or machine learning. So where are all the applications?

2. Applications of Data science

Data science, where it is being applicable nowadays is Siri or Google or Alexa. The morning is starting with these devices. This is a virtual personal assistant. These are working based on natural language processing. If a model is being trained say for example you are asking where exactly you are or what is meant by IIT. We are asking the question to Alexa. Somebody might have asked previously somebody might have trained. Suppose we are asking the question what is newton's third law? somebody might have initially trained. Initially the machine does not know what exactly it is what exactly newton's third law. We are keep on training the device. We are feeding the information either in the form of speech or in the form of text. Then it is mining the data from the internet. Sometimes it will answer from the internet it finds. For example, if we want to hear some songs, it will mine from some website as well. That means it is getting the data from several sources. So, it would be trained by somebody else. Somebody might have told what Newton's third law, or it is will mine from the internet sources. Then it is replying that. Means whatever the human what we have done. How did we learn newton's third law in the school? Somebody might have taught that means we got the training and then after sometimes, when the test comes, and the question appears we are writing. The training and testing which is playing a vital role in a machine learning[2].

The virtual personal assistants is one of the applications. And predictions while traveling. Sometimes most of us are using google maps. When we're traveling, sometimes we can find if there is a traffic. Let us say from one place to another place in a particular if there is a red in the road map that means the traffic is very high. How do we get this information? Many people are using mobile nowadays. On that road many are using the internet with the data nowadays. The mobile data or the internal data occurring in that place Is very huge. Based on that it's pulling the data. In fact, it is pulling the data from the mobile and then informing that find this road. We have a huge data in our hand based on that the google is analysing. So here the traffic is very high. That is one of the applications. It is pulling the data and is calculating whether it is a dense or empty road. This is another application.

Third one is about the video surveillance. Nowadays CCTV camera is very popular. Suppose we want to analyse for example a cricket match. Video cricket match video is given to us, so we are going to mine. What are all the frames the batsman is beating sixes? Assume that if it is a one-day match or even a T20 it takes almost four hours. Total four hours if we instead of mining can you automate this? The four hours video is input data and then what are all the frames the batsman is hitting sixes[3]. Similarly in the video surveillance we want to locate some persons or there may be unwanted activities like a fighting, or some people may walk in the parking area. To analyse this the machine learning is playing a vital role. The data here is a video data. The video data is a three-dimensional data. It's like a image and every second we will have some 80 to 100 frames. Imagine 30 minutes video is given and it's a huge data. This is one of the applications in video surveillance.

The social media services, for example everybody is using Facebook. Suppose we want to search some person in the Facebook. We first observe that we can see that search some name. It will first give as a friend. Then suppose he is not in the friends list; it will go to a friend of friend and friend of friend of friend and so on. And it will try to search how far they are related to. Similarly, in LinkedIn, first connection second connection and third connection and so on. How far they are related. It's like a radius. They use the concept called graph search. In the social media they use the concept called graph search. There are several graph-based algorithms in machine learning. Tree based algorithms are also there in machine learning. They use this, similarly the search engine results refining all. Suppose if we type Bapatla Engineering College, we will get in a page, will get all the information maybe all relevant to Bapatla Engineering College. If you try some XYZ link initially we may not find. Initially first time we may not find assume that this XYZ is not very popular. And we are trying to find out the location and assume that at one point of time this XYZ is becoming popular that will be based on the number of searches made by the user. The data is being collected from by the users. So, we are going to type something this XYZ we're going to type something. Assume that many users hundreds of thousands of or hundred thousand of people are trying to type the same XYZ at one point of time. The search engine will understand that, and the users want to know about this XYZ. Then it will give start giving the optimal result. These are all the applications based on the data science and in fact the machine learning. It is a tool to solve all these problems.

4. Requirements to learn Machine learning

Before machine learning era assume that in fact it started long back in 80s itself. The people have walked on neural networks. From a given set of

data points, they try to fit a straight line which is called as a linear regression. The regression is playing a vital role. The data is being analysed and then the result is obtained. They used to give discrete data and convert it into continuous one. For this the linear regression was introduced and it was useful. The non-linear regression can be chosen for a problem based on prediction. If we must be familiar in machine learning, we must learn all these topics initially. We need basic statistics and about linear regression. Then there is a concept called logistic regression. And there is requirement of some mathematics like a linear algebra, calculus, differential equations that is involved or the application for machine learning to study. And then clustering, the difference between clustering and classification. Then it is needed to study clustering, k-means clustering, and centroid means clustering. Apart from these feature extraction algorithms called k-nearest neighbour, principal component analysis (PCA) and then several classification algorithms need to be studied. Name base is based on the probabilistic model that requires probability[4]. The mathematics one requires probability also and then there are several classification algorithms need to be studied. If one is familiar with these topics, then it is easy to know the basics of machine learning.

5. Classification of Machine learning algorithms

The machine learning algorithms are classified as supervised, unsupervised and reinforcement learning. Suppose there is a robot. And then we have trucks of fruits. The fruits are we have grapes, apples, bananas, papayas everything. The task is the robot has to classify these folks so that means let us say 70 percentage of the fruits in the track we are giving a training to the robot that means we are annotating it. We are saying that this is orange, this is apple, this is banana, and this is a mango etc. We are annotating. We are giving the class. We are labelling the class. And then how do we annotate or how do we give the training? For example, the banana is completely different from apple in terms of colour in terms of shape in terms of texture. So, this colour, shape, texture they are all nothing but the features. The features are playing the important role in machine learning. Features extract the features that's called feature extraction. Once it is extracting, the features have been 70 percent training[5]. It is extracting the features. Now we are going for the testing 30 percent training of remaining fruits. we are going for the testing to clarify whether it classify properly or not. And then we are getting the accuracy. If it is classifying exactly such algorithm is called supervised learning.

The second one is unsupervised learning. In this we don't have any class label. We are not training the robot, but can it classify? Based on all the features it must extract the futures by itself and we must do the classification. It's unsupervised. The human being, initially when we are one-year-old baby, mother is teaching the face recognition. Initially she's telling he's your father he's your brother he's your sister. After five years or even 10 years the human is becoming unsupervised. Once the child sees one time, he can be able to distinguish. That is a major difference between the supervised learning and unsupervised learning. And what is meant by reinforcement learning? Suppose a monkey is there. If it is doing a right job, it will be rewarded. Let us say a banana is a reward or a gift. If it does a job wrongly, we are going to assign a job to the monkey. If it does the job correctly, we are rewarding. If it is not doing correctly, we are giving a punishment. We are giving a beat. In the process of reward and the punishment the monkey starts working properly. Monkey starts working correctly the assigned job so we call this as reinforcement learning [6].

6. Brain computer interface

Any data is given and if we go for the machine learning algorithm, it falls one of the classification algorithms. It may be supervised; it may be unsupervised, or it may be the reinforcement learning. In another few years, we are going to enter the brain computer interface. For example, right now we are switching on the TV or switching on the AC. We use remote. Let us say after 10 to 15 years when we enter the room, and we are thinking that the AC has set to 25 degrees Celsius but if we want 19 degrees Celsius. We are just thinking. The brain is giving some signal. This signal will be converted into digital and what exactly we are thinking and accordingly it is changing to 19 degrees Celsius. Automatically will it change? and if we want to change the TV from news to sports, can it do automatically? that is called the brain computer interface. For this to collect the data from the brain.

Everybody knows about ECG. In the case of a brain, we collect the data putting this apparent as we. It's called the EEG means electro encephalogram. The brain is divided into four parts, and we can see the electrodes.

Fig1. Device used for encephalogram.

The electrodes are round in shape. These are all responsible for getting the data from the brain. The brain is divided into four parts exactly like the heart. ECG is for heart and here we have EEG for brain. The electrodes are seen here in fig1. This will fetch the data from the brain. For example, if we want to analyse some problem, in fact, these are the example of the EEG device. The electrodes will get the data. Suppose we are eating something and then it is giving some taste for example we are eating Mango, it's a sweetened taste. Sometimes it may soar. We may think if it is sweet, it will be good. What exactly the taste is? the brain is giving some information if it is soaring test. So, the mood will change and then the brain signal that will generate something. When we are touching the hot water, immediately we are taking the fingers out. How is this happening? The brain is sending a signal to the fingertip and automatically we are taking out. Based on the EEG data it can be predictable. So, the EEG device we are putting like this. A person let us say watching a video we are putting the EEG device like this and then it will generate a data. Whatever the data that EEG is going to give it will be like exact or perfect. The problems that we can solve, for example prediction of advertisement. For example, 20 YouTube songs are given and then we have to do the ranking. Usually what we do with the data? We will do it based on sentiment analysis. We can easily do it using any statistical data we can come out. But now the EEG apparatus is put it in the head and then it will give a signal. Based on this signal with sentiment analysis that means the same person is giving a data as well. We are taking the brain signal also. We are taking feedback also. It is like using EVM and Ballot both in the election. EEG will give a data and the sentiment analysis, and the feedback also will give a data. So based on these two data can be performed. Assume that 20 or 200 songs are given, and we have to analyse based on these two data. For this, machine learning is going to play a vital role. We have two different data one is sentiment analysis data another is the EEG data [7]. The architecture of the framework is shown in Fig.2.

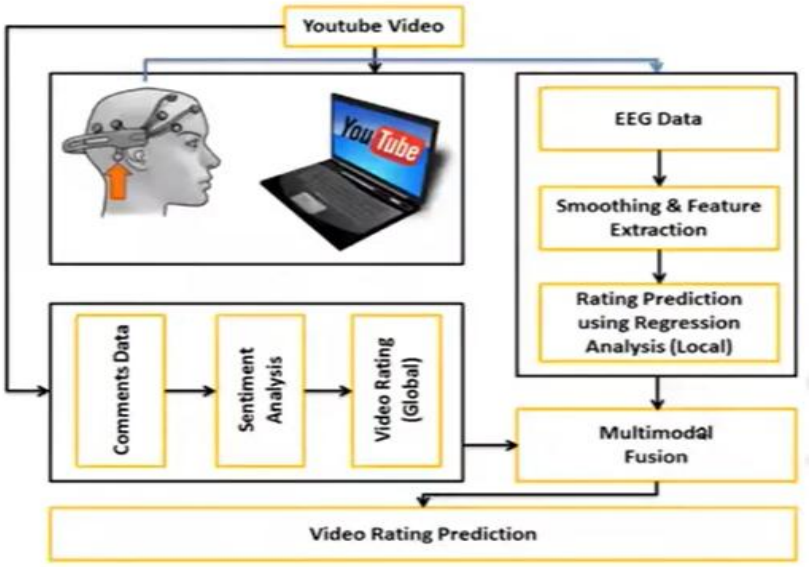

Fig.2 Architecture of the framework

In both the data we can find the feature extraction. Using any statistical model or regression model. Then the classification can be performed. Similarly, sentiment analysis also does the same. It will extract the features and do the classification. So now it is multimodal. We have an easy data with the weight plus we have a sentiment analysis data with the weight so whichever is giving higher

score, we consider that is the top. And then we can analyse based on this course. The EEG data we are applying to a machine learning algorithm and sentiment analysis we are applying a machine learningalgorithm then we call this as a multi-modal fusion or bimodal fusion. That means we are using both EEG data as well as sentiment analysis data. Combining both we get the result that is what the idea. Assume that we are using the 160 electrodes, the signal will be better than using 14 electrodes. In another 15 to 20 years, we may even can put it in the mobile. It will be like a chip in a mobile. Suppose you are getting an email from your boss saying that you have to attend the meeting at five o'clock. And you know that you are attending this talk you may not attend so you are just thinking. This can go as a reply without even typing, so this is what the brain computer interface. This technology may be possible in another few years. This is one of the applications. Usually, people use with sentiment analysis. For example, the data mining researchers suppose want to classify positive sentiment or negative sentiment they have to extract the features. By preprocessing they will extract the features, train the data set and for testing they will go for the classifier. This is a standard way. And what are all the future extraction algorithms? Suppose the signal is given the one dimensional signal, how to extract the futures for example we use this for different purposes. For example, in the banking sector now in order to do the transaction, we use only the signature. The signature data is the image data. Then we are doing the classification to know the signature whether it is a matching. If it is getting matched, they will process. And then the transaction is being happened. Suppose several users are giving the EEG signal data, how one can extract the features? for example one signal will be different from other signals, whatever the brain is generating [8]. How do one can extract the features from the data. For example, we may use the frequency domain like a Fourier transform like a wavelet transform.

All feature extraction algorithms such as independent component analysis (ICA), kernel principal component analysis, latent semantic analysis, NPI, we use this natural language processing, principle component analysis are to be studied by a data science engineer. For example, we are using the video data. Video data is very huge. So, we need to extract the information, we need to have only some vectors. Some features out of the data because we do not want to use the huge data because we may not have the computation power. In that case we extract the features from the data for further analysis. We use only the feature extracted data. These are all some of the important feature extraction algorithms.

7. Other applications

Other applications like human activity analysis or human activity recognition are also needed. The infrared images in the night what are all the actions are going on particularly in the border security force. Border intrusion people may come from border during night. At that time, we put the camera and it has to give the signal. These are all some of the applications where the machine learning algorithms or even deep learning algorithms are useful. To analyse from the video data, we require the machine learning algorithm. Similarly, many people have worked on Covid X-ray images.

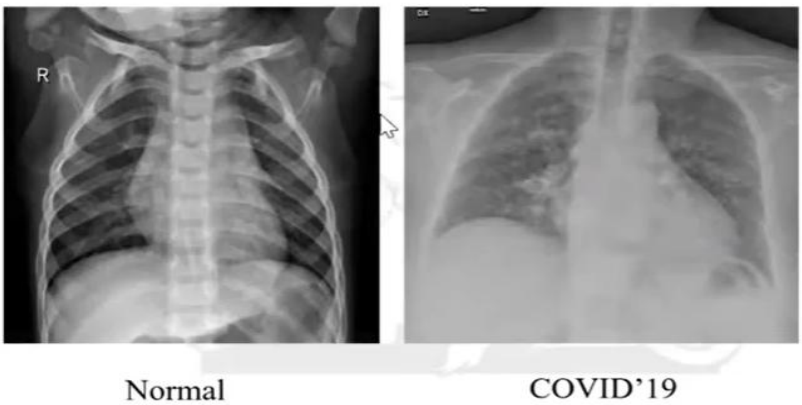

Fig.3 Binary Classification.

This is a binary class problem. The left-hand side it's a usual or in a normal person. Right hand side the person is suffering from covid. so, x-ray image data we are given on right hand side lungs and the bones are not visible whereas in the left-hand side normal person. Everything is perfect. The bones are visible, and the lungs are visible. So, we know that as a human we can be able to categorize but how the machine categorizes this. We are passing the information, so it has to extract the features and when it is extracting all the features, we are giving it training. Like this we have thousands of images

or hundred thousand of images. We are giving a training and then it will classify who are normal person and who are suffering from covid19. This class of problem is called binary class problem. Like this we have so many applications.

Sign language recognition is another application. it's American sign language where we have 26 characters. We can classify these by starting with alphabet and then go by words. then we may get the results. Machine learning algorithms are like a trial and error. Once we're familiar with the letters, we go for the words then we go for the speech. When we're putting a gadget or mobile in front of them (disabled), whatever actions they are doing the mobile has to give a reply. Similarly, we understand what exactly they want to say and we are trying to give a reply and then the mobile will give the action for them. That is called the sign language recognition [9].

Another application is in terms of computer vision problem. We can apply both machine learning and deep learning algorithms. Input is like a video and from the video we are fetching certain data. We call it as a motion history image, dynamic image or RGB motion image and we apply deep learning algorithm to extract the features[10]. The deep learning algorithms are used to extract the features and machine learning algorithm for classification. So, mixture of deep learning and machine learning can yield better results. Deep learning algorithm will give features and based on the features we apply machine learning algorithm for classification [11]. These are all some of the applications.

8. Conclusion

The deep learning plus machine learning algorithm accuracy is more. We have to measure what exactly the output that we are getting is also very important. Several applications are there in this field. We can imagine whatever data that we are finding in the world, can think of applying statistics, machine learning, deep learning and then find out what exactly we are inferring from the data that is what the main thing. If we have a data, we can use machine learning or statistics as a tool in further information that is called data science. These are all some of the important classification algorithms. Initially when we learn machine learning it will be good to learn logistic regression, decisiontree-based classification, k nearest neighbors, support vector machines. Feature extraction algorithms are all nothing but the classification algorithms.

References

1. Davenport, Thomas, and Jeanne Harris. Competing on analytics: Updated, with a new introduction: The new science of winning. Harvard Business Press, 2017.
2. Big Data Nowby O'Reilly Media, [online] Available: http://indexof.co.uk/Big-Data-Technologies/Big%20Data%20Now.pdf.
3. Amandeep Khurana, "Bringing Big Data Systems to the Cloud", IEEE Cloud Computing Published by the IEEE Computer Society, pp. 72-75, 2014.
4. Hellerstein, M. Joseph et al., "The MADlib analytics library: or MAD skills the SQL", Proceedings of the VLDB Endowment, vol. 5.12, pp. 1700-1711, 2012.
5. Robin Anil, Ted Dunning and Ellen Friedman, "Mahout in action", Manning, 2011.
6. Seung-Hwan Lim et al., "Graph Processing Platforms at Scale: Practices and Experiences", Proceedings of the 2015 IEEE International Symposium on Performance Analysis of Systems and Software, 2015.
7. Big data: The next frontier for innovation competition and productivity, [online] Available: http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation.
8. D. Coetzee, "Illiteracy in South Africa: Some preventive policies and strategies from a developmental perspective", Development Southern Africa, vol. 8, no. 2, pp. 215-231, 1991.
9. Tyagi, Amit Kumar and G Rekha, "Machine Learning with Big Data (March 20 2019)", Proceedings of International Conference on Sustainable Computing in Science Technology and Management (SUSCOM), February 26-28, 2019, [online] Available: https://ssrn.com/abstract=3356269.
10. G. Rekha, A.K. Tyagi and V. Krishna Reddy, "Solving class imbalance problem using bagging boosting techniques with and without using noise filtering method", International Journal of Hybrid Intelligent Systems, pp. 1-10, 2019.
11. Tyagi, Amit Kumar and M Shamila, "Spy in the Crowd: How User's Privacy Is Getting Affected with the Integration of Internet of Thing's Devices (March 20 2019)", Proceedings of International Conference on Sustainable Computing in Science Technology and Management (SUSCOM), February 26-28, 2019, [online] Available: https://ssrn.com/abstract=3356268.