Second National Conference on Internet of Things : Solution for Societal Needs
In association with International Journal of Scientific Research in Computer Science, Engineering and Information Technology | ISSN : 2456-3307 (www.ijsrcseit.com)

# A Big Data and the Machine Learning Algorithms

**Prof. Pradip Sitaram Ingle[1], Mr. Vaibhav Subhash Borade[2], Mr. Akshay Shivaji Sonune[2]**

[1]Assistant Professor, Department of Information Technology, Anuradha Engineering College Chikhli, Maharashtra, India

[2]Student, Information Technology Department, Anuradha Engineering College, Chikhli, Maharashtra, India

## ABSTRACT

The Mayor of central banks discusses the dental institute of big data processing on big topics. Microdots are used with automated redesign applications in a variety of cases, including investigation, monetary policy and the established financier. Central banks are informed that they use big data for supervision and regulation (suptech and regtech applications). The quality, sample, and presentation of the data are important data for the central banks, even though the legal security on the lathe is due to the privacy and confidentiality of the data. Information institutes provide a variety of information that can be assessed on the basis of the infrastructure to be assessed, the infrastructure, to the humanitarian need. Cooperation between public authorities allows the capacity of central banks to be recovered, sorted, and analyzed macrodactyly.

## I. INTRODUCTION

The world is changing and so is the way it's measured. For decades, policy maker assess the state of the economy. Collecting this data requires substantial effort and publication often occurs with a delay of several months or even years. However the The last few years have seen an explosive growth in the amount of readily available data. NewData collection and dissemination models allow analysis of large amounts of data.in real time. An important factor in this development is the advent of the information age, and especially the smartphone and cloud computing: individuals and companies produce unprecedented amounts of data being stored for future use on servers of technology companies. For example, billions of Google searches every day reveal what people want to buy or where want to go for dinner. Social media post and the private sector have relied on data published by official statistical institutions to allow market participants to track the dissemination of information on social media. Companies record every step of their production or sale process, and payment transactions and e-commerce create a digital footprint. The additional catalyst in the form of creation of big data, especially financial data, has it has been the Great Financial Crisis (GFC) of 2007-09. The GFC laid bare the need for more disaggregated data: a relatively small but interconnected bank like Lehman The brothers could bring down the financial system because it was so interconnected. Regulatory and reporting requirements established after the GFC have increased the data reported to central banks and supervisory authorities, and additional work to improvement of central bank statistics is ongoing (Buch (2019)).

The advent of big data coincides with a quantum leap in the technology and software are used to analyze it:

artificial intelligence (AI) is the topic of the day and allows researchers to find significant patterns in large amounts of data. For example, natural language Processing techniques (NLP) convert unstructured text into structured data that Machine learning tools can analyze to discover hidden connections. Network analysis it can help you visualize relationships in this high-dimensional data. For the first time in history, it is possible to produce areal-time image of economic indicators such asconsumer spending, business sentiment or people's movements. These developments have spurred central banks' interest in big data. Growing The interest is reflected in the number of speeches from central banks that mention Big Data. and to do so in an increasingly positive light (Graph 1). And yet big data and machines learning pose challenges, some of them more general, some of them specific banks and supervisory authorities. This article reviews the use of big data and machine learning in the central bank community, taking advantage of a survey conducted in 2020 among members of the Irving Fischer Committee (IFC). The survey contains responses from 52 respondents across all regions and examines how central banks define and use big data, as well as what opportunities and challenges they see.

## II. WHAT IS BIG DATA

Wondered how much data is generated in the format text, email providers videos searches and music approximately generated every month by a single smartphone user in this number 5 billion smartphone users that even process is in fact this is quite a lot of computing systems and this massive amount of data is termed as big data generated from minute 2.19 snap on Snapchat 3.8 million searched on Google 1 million people Login to Facebook 4.5 million videos watched and YouTube 188 million email so how are you classify any data as big data this is possible please provide us today and value for the Healthcare industry hospitals and clinics across the world for your 2314 how can I am the danger 2314 data collected in the form of creation wickets in tests of all the status generated at a very high speed it reduces the velocity in The Dictator the various data types of structured semi structured and unstructured data examples include exam bonfires and x-ray images accuracy and trustworthiness of The generated data is termed as previously stated that it is better treatment and reduced cost this is known as the plan you think they used to process this big data store and process the spectator disturb lyrics and Spark example and see how to store hours and processes big data Hadoop uses a distributed file system known as Hadoop distributed file system the storing data if you wish you try you finally broken down into smaller chunks and store it in there is machines that when you read the file you also make copies of it would do it this way you store your pic there in a distributed then make sure that even the one machine fail you did it safe on another not produce technique is used to process big data link is broken into smart b&b Now instead of one machine Re machines Dab and thanks to the processing machine you did it safe on another not produce technique is used to process in data link is broken into smart b&b now it's a one machine Re machine take up its task and completed in a fashion and the sum of reserves and thanks to this the process becomes easy and fast this is known as parallel processing not be used toward and process of big data and analysis data from numerous applications in games like he was free and interior designers news today designer understand in which state produces fast we start work with my this insight into the river comes to your mind of the game and improve the User experience which in turn reduces the customer and made Stewart and process static data we can analyze data from numerous applications in games like industry and Call of Duty designer News Today understand it which state produces all we start work with my this is I can help them overcome story Mein

became and improved User experience which in turn reduces the customer churn rate of disaster management in 2012 he was IL picture of disaster management in 2012 he was used to gain a better understanding of the Storms effect on the east coast of the US and necessary measures were taken it could protect the Hurricane from five days late when it was impossible to hear these indications of help you can be accurately process and analyzed using video actually processed and analyzed as a question which of the following statements is not correct about attitude distributed file system in HTML is the storage layer that you did I get stored in a distributed manner in HTML c first performs processing of data stored on Fast and Furious which of the following statements is not correct about Hadoop distributed file system it at a first is the storage layer and you did I get stored in a distributed manner in htr380 performs processing of data are stored on top of that and it's in a team of the given and video actors in the comment section below 3 lakh winner will receive Amazon gift voucher plan twerk big data that is what you think will be the most significant impact of big data in the future is now in the comments if you

## III. HOW BIG DATA WORK? WHAT IS BIG DATA USED FOR?

Anything anybody does is going to have some predictive power for other things to do Stephen a form a data scientist in Google in his book everybody lies beta new Tata and what the internet can tell us about who we really are very fascinating study researchers use Microsoft print to look for people who search for diagnosed with pancreatic cancer and then people who never made that search Angel look at all the health symptoms suggest that they meet in the lead-up to the diagnosis on your diagnosis they found that they were very me a patterns of symptoms that were far more likely to suggest the future diagnosis of pancreatic cancer for example searching for in digestion in abdominal pain but evidence for pancreatic cancer was searching for just in digestion without abdominal pain without abdominal pain was much more and likely to have pancreatic cancer what we to understand people to know what they are interested in or out thinking it to do a survey to actually asked them different questions but we all know that people like people can't answer certain questions given to their loved ones this is where big data becomes very interesting things that people like not ask their friends family or doctors they ask Google data scientist at Google search data and the result is something that seems very different from what we generally tend to believe of humans they have used this data to understand suicide it's been found that one can predict how many suicide there will be an area based on Google searches so when more people are searching how to commit suicide or how to give yourself it turns out the day will be more to side in that area to use it to understand why people think of suicide on what do people search before they look more information on how to commit suicide gets in inside on why people are societal it is found that 40% of people before the search for suicide search for some health condition before so health seem to be a big reason more than financial problems or relationship problems in the number one health problem that causes suicidal thoughts is depression and human behavior was extremely Complex it is no surprise that Mini existing findings in the fields of psychology and medicine are conflicting the more data we collect the better understanding of human behavior will be instead of relying on theory driven methods as often the case in psychological studies that the approaches in crime country and let me know theories arise directly from at the sheer volume of data available available on the internet is meant in finding a way to utilize the available data to understand market trends in customer behavior it in efficient application of the collected data that is used to predict tiffin market in customer Trends Every

time We type into a search box will you been something about ourselves as millions of us are searching for questions on things to buy places to meet friends produce a map of our collective home centralized Data Analytics and strategies are made accordingly to provide a personalized experience for the customer but there are potential downsize to the collection and use of big data technological expertise in civil libertarians have consistently want about security issues and big data privacy risk if you use a variable like a smart watch that your heart but it to use Netflix to watch a show of these activities we with digital trail of data which is collected for analysis of premature Corporation Sachem Facebook apple and Amazon use Big Data Analytics to collect and analyze a daughter are Dada

## IV. DIFFERENT TYPES OF BIG DATA

That we are going to handle use amount of data on the this big data and that's why does become a subject that we are having military category of data will be coming here what will be our structured data and unstructured data at the last one is with the semi structured. Data structure determine that will you determine which can be represented in the form of rows and columns and that is known as the structured data as an example we can say that the respective web blocks that differed in data which should be generated by the sensor and say if we do in survey on overpopulation population will be getting our structured data. Structure data is example of database in unstructured data that data will be in the form of SIM, data file, PDF ,images, videos on the images sent by the satellites and different machines and data on tractor videos cheated on the category of unstructured data that is also used type structure theatre now it is where having 20% but , unstructured data is now a days is having almost 50% in case of semi structured data to some extent it is structured to some extent it is unstructured as an example we can be discussing the XML file, JSON files and soon they will be falling under the category of semi structured data.

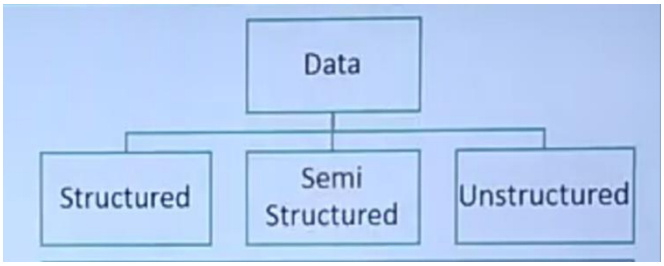


Figure 1. Types of Big Data

1. structured data
2. unstructured data and
3. semi structured data

### 1. Structured data:

Going for this database you can find that this is a good example of a structured data. When having certain columns so that is the employee ID , number, name, age, department when and salary and hate when having that list into rows containing the respective information so here when having the records and hearing the column structure data type of data which are stored already in an order and there are nearly 20% of the total existing data. Now a days and their data structure and all the data generated from sensors with long and deep Lock aunties all machine generated structured data is a weblog Elizabeth insert data and so that human generated structured data Atos Wiki are taken as information from in human like their names, addresses, Gender and date of birth and so on to the example of structured data.

### 2. Unstructured Data:

Is database concentrate is the unstructured data so, the unstructured data have no clear format in storage and we can store structured data in rows and column database path unstructured data cannot be stored like that so unstructured data cannot be stored in the form of rows and columns and when having at least 80% of the data send of the data. Now it is existing

which is feature on structure images and big data, Audio images are categorized as mission generated on structured data show the images which will be sent by the respective satellites can be treated warning in the category of unstructured data. Various types of human generated on structured data is the images ,videos ,social media, data, except for example of structured data at the text documents, PDF, images and videos and Central to this is the during on this is the concept of this on structured data.

### 3. Semi structured Data:

It is very difficult to categories this type of data from to structure your order from them they will be looking as on structure so that's why these data are known as semi structured data. We cannot stored this type of data using traditional database format and what it contain some organizational properties and examples of semi structured data are spread sheet files which you have in our excel in the work we are having the file specified is a good example of semi structured data. In the XML and JSON documents that are the Java Object Notation for extended markup language that is a full form of XML and NoSQL databases are the data items which are falling under this answer.

### Advantages And Limitations of big data

Awareness of big data first of all we are going to discuss about that when is of big data first advantage is

- **Quick decision making**

You know big data help the business executors and managers to make better and informed business decisions. Do big data that helps in quick decision making.

- **Fraud detection**

Banks and other Financial Institutions to detect frauds like Foodland purchases with credit cards even before the cardholder gets to know about it fraud detection.

- **Improved customer service**

Big data provides massive information's regarding the customers and their tastes and preferences an analysis of such data helps the company to provide better Customer services next and when it is.

- **Competitive advantage**

Big Data Analytics helps the business to gain a competitive edge over its rivals over its rivals next one is.

- **Better sales inside**

Real time analytics tells us how this is a going so big that it is very helpful to have an idea about the sales agents.

### Limitations of big data

- big data requires massive storage
- Big data may include unwanted or unstructured data
- Big data can be used for the manipulation of customer records
- Big Data Analytics may not be suitable in the short run
- Big Data Analytics may subject to technical glitches
- Big Data Analytics may subject to security threats

### Problems on big data:

### Problem 1:

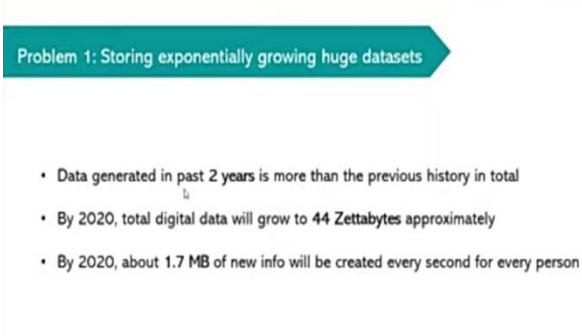


**Figure 2: Problem 1 (Big Data)**

The first problem that we have is storing exponentially growing to status with this already told you ok have a lot status it and for example if your

storage is limited to some 10 PB ok you cannot assume that the data will always be falling under 10 PB water one day the data is suddenly more falling under 10 PB water one day the data is certainly more on example for this the festivals feeling that we have the online sales so the website traffic is more than that more than you predict all right and sometimes you are not able to place your order as a user ,as a customer and sometimes the company will not be handle the lock lot of traffic the decade and a lot of people who are visiting the website then who are trying to find the products and in so many previous other great festivals and all they had to create online festivals there were so many issues that happened which actually spoilt the entire business of few famous Shoppers ok so we have to store the growing you status its, so it becomes a great challenge you have to make sure that the data is growing you know properly like a it certainly doesn't grow more and it is consistent so we cannot predict that but we should make sure that our system is capable enough to handle the traffic traffic to the same example that we saw in the initial volume data generated in past two years is more than the previous history in total so imagine if we are not ready to save or store all the data properly we would be the losers we would be not able to properly analyze it alright so my 2020 1.7 MB of new info will be created every second for every Passband I am sure it's more than that alright so this data story is practically difficult alright you have to make sure that you have a system I would say all rights you have to make sure that you have a system using which you can store your data which is like kind of dumping the data would pick as and when the data comes in you have to first dump it and then make sure which data you want in which state are you don't want rather than missing to store the data when you missed to store all the growing information you when you miss to store all the growing information you will be losing some really valuable data so this is the first problem storing exponentially growing huge data sets so as the picture depicts

**Problem 2**

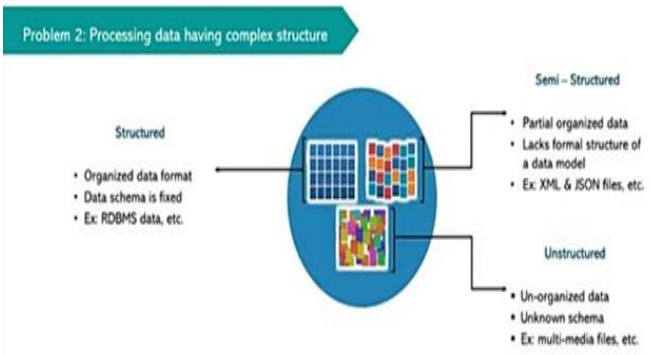


**Figure 3: Problem 2 ( Big Data)**

Second problem is processing data having Complex structure so this was the one that we were talking about a variety when the data is of different variety we saw that it was unstructured structured and semi structured ok so when the data is having a complex structure processing the data is very difficult alright so what is it so when you take structured data and the data is in an organized format so data schema fixed ok so data cleomes fixed structure to it example is RDBMS data so few see you can identify you can identify the data based on how it is stored and based on the particular schema. Example all the HR related data would be stored under the head HR schema and FEMA and accounting details could be stored under the account scheme up so you have a prop organized data format in the structure 1 so, this is a simple thing when you when the data is in a proper structure it is, it becomes quite easy to process or simpler to process because you have it all their coming to the next to which are which are probably A B complex structure as I told you it is partially organized lacks the formal structure example XML JSON except so what we do here in the semi structured data you have to convert it to structured information and that you have to process using the structured data processing engines or your RDBMS and good example is even you know

in Oracle database you can store the data in XML format and you have the option of taking the XML and processing the using some XML queries XML query is there but again that that is possible only for unlimited data when it comes to data which grows and it is very expensive to put it in your Oracle database or any other Rachis BMS and storage and processing it. Where you can process the semi structured data, so here comes the no sequel data bases are the not only sequel data bases so no sequel are Himalayas when it is schema less what happens is you can store the data which comes in a semi structured format like cheese XML JSON so you can store the data in your NOC in your no sequel and it is it comes of different data models in structured usually it is RDBMS model or relational model or you can also call it as the table and model but here you do waste on your type of data you can store it as a document you can store it and graphs and so many other options are also available depending on the nature of your data. You can choose to use the no sequel database that you want this no sequel database is mostly open source and based on the category of data you want you can use and it tries to hide the complexity from you and helps you to process the semi structured data in a better way.

**Solution on big data:**

Solve the problems now the solution for all these problems, so how can be used Hadoop as a solution for all the problems. We were discussing Hadoop it so firstly how to process Framework that allows us to store and process large data sets in parallel and distributed fashion. so to explain it further it is an open source Framework which is used for parallel and distributed computing where you can store and large data sets at the same time that is and Francis Hadoop status at the same time that is the charge that is the beauty of hello so it is not one particular software it is a framework. It is safe framework and it allows us to store and process large data sets ok and it helps you to do it in parallel as well which was the third one the third challenge that we saw that the data cannot be process faster because we keep you have a one master and mini slave architecture then it would be serving one by one only but here if you see here it is doing everything doing everything in parallel and in a distributed manner so that is distributed computing you know that the data is distributed across the different servers and then the processing happens so something like that so how does Hadoop solve a problem the main two components is HDFS map reduce which are part of which helps is too kind of small and big data from HDFS in Hadoop distributed file system which is Hadoop storage system the name implies it is a file system and it is not a database it is not a database on right so it is the file system like your Windows file system or your Linux file system it is also a file system but it is Hadoop distributed file system it the data is distributed data is not store in one Li one place it is distributed alright so because cause of distributed nature it allows to down any kind of data across the cluster the group of service bright so you can download any kind of data across the cluster using the HDFS connected servers can actually share the story on rent and this is to store the data and how do I process the data in a distributed manner is done using processing is done using map reduce so produce is used for parallel processing of data that is stored in HDFS so I have my data in HTML if I have to do some processing on HDFS I have to use map reduce to do it now let me explain you further on how this helps you to solve the problem write so two things to keep in mind now Hadoop is a frame which which allows to store and process La status in parallel and distributed and restoring is using HDFC and processing is using map .

## V. WHAT IS MACHINE LEARNING?

As you know, we are living in a world of humans and machines, humans have been evolving and learning from past experiences for millions of years, on the other hand, the age of machines and robots has just started in today's vacuum. nowadays, the theses of the machine or robots are like them. it needs to be programmed before actually following its instructions, but what is the machine that started learning itself and this is where machine learning comes into the picture? Machine learning is at the core of many futuristic technological advancements in our world today, you can see various examples or implementations of Learning Machines all around us such as Tesla's Apple Siri robot Sophia Al and many more, so what is it exactly machine learning? Machine learning is a subfield of artificial intelligence that focuses on the design of systems from which you can learn and make decisions and predictions based on experience, which is the data in the case of machines, learning allows the computers act and generate data. Decisions driven instead of being explicitly programmed to accomplish a certain task, these programs are designed to learn and improve over time when exposed to new data, let's go ahead and discuss one of the biggest confusions of people in the world. of them, machine learning and deep learning are the same, you know what is wrong, let me clarify things for you, artificial intelligence is a broader concept that machines can perform tasks in a more intelligent way, it covers everything that it allows computer to behave like humans think of a famous Turing test to determine if a computer is capable of thinking like a human or not if you are talking to Siri on your phone and you get an answer, you are already very close to that, so This was about artificial intelligence now coming to the machine learning part, so as I already said that machine learning is a subset or current application of AI, it is based on the idea that we must be able to give the machine accessing the data and letting it learn from itself is a subset of artificial intelligence that deals with the extraction of patterns from the data set, this means that the most who can not only find the rules for optimal behavior

But they can also adapt to changes in the world, many of the algorithms involved have been known for decades, even thanks to advanced computing and parallel computing, now they can scale to massive volumes of data, so this was the part of machine learning. Now deep learning is a subset of machine learning in which similar machine learning algorithms are used to train deep neural networks to achieve better precision in cases where the former was not performing correctly in the process. Map. I hope now you have understood. that Al machine learning and deep learning are all different.

### How do you solve a problem using machine learning?

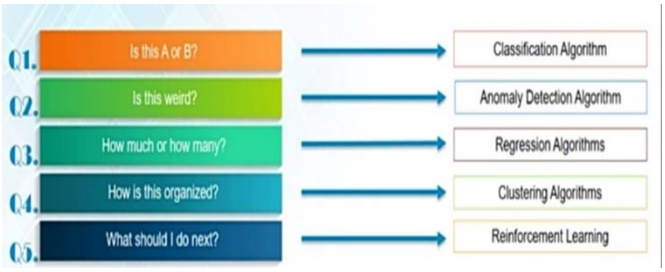


**Figure 4: How solve a pro using machine learning**

So whenever there is a problem, that boundary problem can be classified in five ways, so these five ways are like this: a or b is this, see how much or how many, how is it organized, and what should I do next to make its problem you may want an answer, but what could be is this a or b, since in your opinion you have a problem that is to ask, let's say you are differentiating between fruits in this an apple or a banana, so when you have this type of problem use classification algorithms the next category is this view analyze patterns, so when you have problems where you have to analyze the frame and where and you have to find a LOC anomaly or a strange one,

you are actually looking for an anomaly detection algorithm. The next category is how much or how many so you have to deal with the numbers. So when you want some numerical values, you want to get a certain value or for example what should be the minimum number of hours you have to spend to get a promotion. So when you have these problems and apply your expiration algorithm on that, then we have this organization. So when you have these types of questions, you use clustering algorithms because you are basically trying to figure out what the structure behind a certain data set is. So when you are trying to figure out this structure behind certain problems, you use clustering algorithms and then you have a category that says what should I do next, so well a decision has to be made and then algorithms are used for reinforcement learning.

**Machine learning types**

3 kinds of way so the 1st way Supervised Learning the 2nd way is Reinforcement Learning and then you have the Unsupervised Learning.

**1. Supervised Learning:**

The 1st kind of learning is called supervised learning. What is supervised learning so supervised if you concentrate on the word supervised, supervised means when monitoring someone or when you're constantly monitoring someone or making them understand some thing so you compare the scene with a classroom scene. So you sit in a class and the teacher explains you different concepts. The concept be learnt on your own that's why teachers is there so the way we teach machines in supervised learning is like this so we provide them with a particular set of inputs and we give the corresponding answer as for example if I am if I say that what are the parameters for deciding whether it will rain today or not so the humidity should be above some certain level temperatures should be above some certain level the brain should be in a certain direction and then if these scenarios are there it will rain right so we give a lot of inputs to the computer with this data and with each data we assign one it rains and zero it does not rain so if the temperature is high and the humidity is high and the wind is in a particular direction it will rain so we say 1 and if the humidity is low and the temperature is low as well so we say it will not rain so I will give it a certain input that today it's certified degrees Celsius the huge 97% the wind is in this direction will it rain so the machine will actually see ok so much up from my past experiences I saw that ok this was the temperature and the humidity was this it rained so it will compare that according to it and it will come up with the probability and hence come up with the solution or an answer whether it will rain or not so this is what supervised learning is so basically we are providing it with the Answers and the inputs as well all.

**2. Unsupervised Learning:**

Next topic is now unsupervised learning so what is unsupervised learning so you can understand on supervised learning by comparing it straightaway with supervised learning so in supervised learning like I said you were giving them ass well to input but an unsupervised learning you don't do that you just give them inputs now you are not giving telling your computer what will be the answer so what computer does is or what is it think logically the only thing computer can do with the inputs can do with the inputs is find a pattern behind it or find a structural in it right so this is what the computer actually does so in unsupervised learning what it does you provide inputs so for example I want my computer to I give my computer some inputs on fruit so I don't tell the computer what the flu pills actually but I give other parameters such as how big it is or what color it has say what is the taste of that fruit so when I give all these conditions or all these parameters to my computer so it groups the fruits basis on that so basically it will group on the basis of size it's a group

it be on the basis of taste it will group it be on the basis of color and then it shows us that data and then we can actually label ok so the size is big and color is this you this this fruit will be known as an apple now what kind of problems are actually there in unsupervised learning could be when we don't know whether there is correlation in the letter or whether there is a structure in the data for example so big data is nothing but a huge chunk of data so we don't know that ring that is Adam it's not structure so whenever we want to find a structure in data we use Unsupervised learning now it's the job of the algorithms to figure out what is if there's a pattern in the data and if at all there is a pattern it gives us that fact and hence decide how we can move ahead alright so if we have structure or if we have if you know how that data can be differentiated how it can be structured then supervised learning can be applied on it but If we don't know what that data structure is we use Unsupervised learning.

### 3. Reinforcement Learning

Reinforcement learning basically is when your computer is trying to take decisions. So what kind of problems can be included in reinforcement learning is say when you you're when you wanted to teach you computer how to play chess now you cannot tell your computer what to do because there are a lot of things that are like a zillion possibilities or zillion moves that can be done in chess. So you cannot tell each and every move to your computer but what you can tell is whether he did or wrong and that is what reinforcement learning. Is for another example that you can take is when you are training your dog all right so you cannot tell your dog what to do because, even not understand but you can actually reward your dog if it does and you can punish them who does wrong. So that same thing is actually applied in reinforcement learning as well so it basically the computers aim is to maximize rewards ,when it does the actions so it will come up with the solution which has the maximum rewards in place so we define if it does out sudden action you get a reward and then from its past experiences it understand. When I did this I got a reward so let me do something similar I will get mire rewards and that is what reinforcement learning.

## VI. WHAT IS ALGORITHMS?

Very generic approach so if you want to interact with the computer and tell it how it should execute a particular task the way you do it is using a program right so what is the program a program is basically logic which is wrapped around a syntax a particular syntax which is specific programming language now this programming language could be anything it could be is to have a script it could be java it could be python it could be c it could be c++ whatever right but basic things doesn't seem that is the logic algorithm the logic remains the same in every language having said that what is this logic this logic is what an algorithm is alright so in simple words and algorithm is a step by step procedure towards solving a problem in the computer world all right so let's take an example to understand this thing which you exhaust discussed all right so let's take an example so this is a problem or this is an algorithm to print numbers from 1 to 20 so let's go step by step and understand what this algorithm is doing so this is the start passion we start over here and then we see that our algorithm is initializing a variable X to 0 So initialize a variable X to 0 and then we implemented it by 1 after that we are printing that variable and we are choking it whether it's less than 20 so if it's less than 20if it's true or if it's yes it goes back and increments the value gain by 1 otherwise if it's unknown it goes on and into program Since it's a yes right now it goes back and increments the value by 1 again so we are printed 1 and now we have increment the value of X by 1 again so now the value is 2 and then we print that value so we have now 1

and 2 on the board all then again it checks whether it's less than 20 if it's true it goes back again incremented by 1 now it's 3 prints see so this process goes on until the value of X reaches 20 so when the value of x is 20 it prints the value and it checks whether it's less than 20 which unknown and then it ends the program hence you have value which are printed from 1 to 20 so this is a step by step procedure for printing values between 1 to 20 and this is in similar way you would create other algorithms as well so as complex an algorithm can be it can always be represented using a flow chart.

**Machine learning Algorithm:**

**1. Classification Algorithm**

When you have a set number of outputs so basically question like this so is it cold outside today so it's answer will be either yes or no so you have only two outputs the outputs could be either yes but the output could be known all right or the next kind of question to be will you go to work today so its either a yes, no or maybe right so you will either go for work or you'll not go for work or you maybe you say maybe I'll go to work right but there is no other on service can come up right. So when you have these kind of conditions you come up with you solve it using classification algorithm .

Now, when you have two outputs like for example yes or no it is called two class classification but when you have more than two choices as in. In a second question we have yes, no or maybe so this is called multi class classification. So, whenever there is an output which is set as in it , is either true or false or 0, 1 or yes / no whatever. So if it's fixed you use classification algorithm this is basically the gist that you should get out of this any doubts and any doubt.

**2. Anomaly Detection Algorithms**

In these algorithm you analyze a certain pattern and you get alerted whenever there's an anomaly or something which is not usual which happens all. For example if as you can see in the fig that you have apart when you have some blue men and then 7 a red kind of person comes up so this algorithm will actually flag that person because, he breaks the pattern he is something which is not expected and he becomes an anomaly and this is what anomaly detection algorithms are about .Now what is the use case for anomaly detection algorithms it could be for example in credit card companies so in credit card companies each transaction of yours is monitored and whenever there is a transaction which is not usual . which doesn't match your daily transaction pattern you get alerted for it so they might confirm with you whether you only made this transaction all. So when you have these kind of problems you use anomaly detection algorithms to solve them.

**3. regression algorithms**

So like I said, whenever you have to get the value, use regression algorithms. So for example what the correct temperature will be for tomorrow, so whatever value comes out of this will be a number. Let's say I got to 23 degrees Celsius, I got that temperature using some formulas and I got that numerical value, so every time I come up with a numeric value of every time my problem demands that I have to get a mathematical value I go with algorithms of regression. So the second example of this could be whenever I want to give my customer a discount now, how much discount should I give that customer to get more customers and at the same time I don't go and lose? I'm still making a profit, so whenever this kind of problem comes up, I go ahead and use regression algorithms.

**4. grouping algorithms**

So clustering algorithms are basically used, which is why we discussed unsupervised learning, remember, so in unsupervised learning we have clustering algorithms. In which we try to establish a structure

so that we have some unstructured data that you want to make sense of, so what we do is use a clustering algorithm and if there is a pattern that we are calculating, we see that it comes up with that Pattern. and it shows us like this, for example, I send data to my computer and my data then apply a blocking algorithm on that, so this is the type of output that I will get, so it will categorize it and group A, group B and group C and then I can make a decision about what, what can I, what I want to do with this data that I have right, if this computer does not understand anything about this data, it does not understand, maybe it is of course. maybe it's food maybe it's money. So the computer doesn't understand, but what the computer understands are numbers and that's how it relates them to each other and creates groups. Now when you come up with groups, maybe you Wang will use someone else. I'll call it quantity and decide or come up with the solution, but this is what a clustering algorithm will give you.

### 5. reinforcement algorithms

Whenever you have a decision to make, and therefore whenever you have a decision to make and your decision is based on your machine's past experiences or whatever input you have given to your machine, you use reinforcement learning. Now, for example, whenever you are going to want to train your computer to play chess, it is usually application learning and when you have learned or when you have created a model for that and your game is actually being played by the computer. every decision the computer makes. it is also based on or also taken from reinforcement learning. The other example, like A to Z, was of a temperature control system where your system had to decide that it should increase the temperature or it should decrease the temperature. Therefore, whenever such problems exist, use reinforcement or reinforcement learning algorithms.

## VII. CONCLUSIONS

Big data and machine learning algorithms are now being used in almost every industry of the economy. Central banks are also increasingly using big data for research purposes and to inform political decisions. In 2020, over 80% of central banks say they do used big data, up to from just 30% five years before. Among the institutions that currently uses big data, over 70% use them for economic research, while 40% say they use them inform political decisions. These numbers suggest that big data and machine learning offer many benefits questions and can help central banks fulfill their mandate. Current GDP and inflation or by examining spending patterns across regions and populations real-time subgroups provide just two of many examples. However, central banks also face challenges in unlocking the full potential of big data and automatic learning. A key topic of discussion is the availability of big data and tools for process it, store it and analyze it. The design of legal frame work or aspects of cyber security are also at the heart of central bankers' concerns. More practical problems they are budget constraints and difficulty in training existing staff or hiring new staff to work on big data issues. In fact, half of that there ported being interested in collaborating in many or more specific project, with there are three types of cooperation envisaged. First, by sharing knowledge among those institutions that have developed specific skills that they can be reused in other jurisdictions. These skills include general big data techniques (e.g. data visualization, network analysis, machine learning tools), more general information management problems (e.g. development of open-source code, data sharing protocols, encryption and anonymization techniques for the use of confidential information data) as well as specific applications more dedicated to the central bank community (e.g. suptech and regtech areas). Second, using big data to work on

global issues such as international spillovers, global and cross-border value chains

## VIII. REFERENCES


[1]. Altig, D, S Baker, J Barrero, N Bloom, P Bunn, S Chen, S Davis, J Leather, B Meyer, E Mihaylov and P Mizen (2020): "Economic uncertainty before and during the Covid-19 pandemic", Journal of Public Economics, vol 191, September, pp 104–274.

[2]. Angwin, J, J Larson, S Mattu and L Kirchner (2016): "Machine bias", ProPublica, 23 May.

[3]. Amstad, M, L Gambacorta, C He and D Xia (2021): "Trade sentiment and the stock market: new evidence based on big data textual analysis of Chinese media", BIS Working Papers, no 917.

[4]. Andersen, A, E Hansen, N Johannesen and A Sheridan (2020): "Consumer responses to the Covid-19 crisis: evidence from bank account transaction data", CEPR Discussion Papers, no 14809.

[5]. Athey, S and G Imbens (2019): "Machine learning methods that economists should know about", Annual Review of Economics, vol 11, pp 685–725.

[6]. Bailey, M, R Cao, T Kuchler and J Stroebel (2018): "The economic effects of social networks: evidence from the housing market", Journal of Political Economy, vol 126, no 6, pp 2224–76.

[7]. Baker, S, N Bloom and S Davis (2016): "Measuring economic policy uncertainty", Quarterly Journal of Economics, vol 131, no 4, pp 1593–636.

[8]. Baker, S, R Farrokhnia, S Meyer, M Pagel and C Yannelis (2020): "How does household spending respond to an epidemic? Consumption during the 2020 Covid-19 pandemic", NBER Working Papers, no 26949.

[9]. Battaglini, M and E Patacchini (2019): "Social networks in policy making", Annual Review of Economics, vol 11, pp 473–94.

[10]. Boissay, F, T Ehlers, L Gambacorta and H S Shin (2020): "Big techs in finance: on the new nexus between data privacy and competition", in R Rau, R Wardrop and L Zingales (eds), The Handbook of Technological Finance, Palgrave Macmillan.

[11]. Bolón-Canedo, V, N Sánchez-Maroño and A Alonso-Betanzos (2015): "Recent advances and emerging challenges of feature selection in the context of big data", Knowledge-Based Systems, vol 86, September, pp 33–45.

[12]. Broeders, D and J Prenio (2018): "Innovative technology in financial supervision (suptech) – the experience of early users", FSI Insights, no 9.

[13]. Buch, C (2019): "Welcoming remarks International Seminar on Big Data", Building Pathways for Policy-Making with Big Data, Bali, 26 July 2018, also in IFC Bulletin (2015).

[14]. Cantú, C, S Doerr, G Cheng, J Frost and L Gambacorta (2020): "On health and privacy: technology to combat the pandemic", BIS Bulletin, no 17.