

MACHINE LEARNING (ML)

The trending topic in information technology field

^[1]Afroz Labbai, ^[2]Arbaaz Iqbal Nakhwa, ^[3]Khan Ibrahim Firoz, ^[4]Uzair Khalil Faquih

^[1]Student at ai-arkp, ^[2]Student at ai-arkp, ^[3]Student at ai-arkp, ^[4]Student at ai-arkp.

^[1]Computer Department of Ai-arkp.

^[1]Anjuman-I-Islam's Abdul Razzaq Kalsekar Polytechnic ,Navi-mumbai,India.

I. Abstract: *This paper is an information article on Machine Learning. Machine learning is an trending topic of IT industry. It is developed in mostly python, java, R, c++, javascript, scala language. ML is used in a wide range in our day to day life. And ML is going to change many things in the future. Mostly ML is used for BigData problem solution, Pattern learning, etc. ML is purely algorithm based working.*

II. Keywords: *AI, IoT, VR, AR, Big Data, Chatbot, Analytics, Algorithm, ML.*

III. Introduction to ML(Machine Learning):

Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" (e.g., progressively improve performance on a specific task) with data, without being explicitly programmed.

The name *machine learning* was coined in 1959 by Arthur Samuel.^[5] Evolved from the study of pattern recognition and computational learning theory in artificial intelligence^[6], machine learning explores the study and construction of algorithms that can learn from and make predictions on data – such algorithms overcome following strictly static program instructions by making data-driven predictions or decisions, through building a model from sample inputs. Machine learning is employed in a range of computing tasks where designing and programming explicit algorithms with good performance is difficult or infeasible; example applications include email filtering, detection of network intruders, and computer vision.

Machine learning is an application of artificial intelligence^[6] (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly

I. Some machine learning methods:

- **Supervised machine learning algorithms** can apply what has been learned in the past to new data using labelled examples to predict future events
- **Semi-supervised machine learning algorithms** fall somewhere in between supervised and unsupervised learning, since they use both labelled and unlabelled data for training – typically a small amount of labelled data and a large amount of unlabelled data.

- **Reinforcement machine learning algorithms** is a learning method that interacts with its environment by producing actions and discovers errors or rewards

II. How MI and AI changed The industry

Artificial Intelligence and Machine Learning have been forecasted to be the game-changers of the coming decade. Most of the tech trends discussed really fall under the umbrella field of Data Science, as they directly or indirectly affect Data Management. However, some specific innovations like Augmented Reality^[9] (AR), Virtual Reality^[10] (VR), or Digital Twins also indicate hardware enhancements and overlapping of the physical and digital worlds.

In a blog post titled *Forrester's Top Emerging Technologies To Watch: 2017-2021*, Forrester made strong assertions about an impending disruptive technology world, where suddenly an avalanche of ground-breaking digital technologies such as mobile, social, Big Data, Cloud, IoT^[8], Ai, and ML will all emerge and converge to turn the global businesses upside down.

In another Forrester's post titled *The Top Technology Trends To Watch: 2016 To 2018*, the author claims that Big Data and IoT^[8] have suddenly catapulted the need for Intelligent Analytics systems powered by robotic assistants and self-teaching algorithms. What this trend indicates is that the future business owners, leaders, and managers will have more time to invest on critical business issues as day-to-day Data Management tasks will be taken over by intelligent machines. The additional layer of ML-driven AI capabilities has taken the afore-mentioned technologies to the ultimate era of technological freedom, where man and machine co-habitate happily. The transformative outcomes of such a "connected business world" are participative customer relationships, data-powered decision making, implementation of innovative risk and security strategies, and superfast technology maturity driven by global customer obsession with everything *tech*. The businesses, big or small should take advantage of this era of *digital disruptions* and empower their workforces and customers alike.

In the next five years, say by 2021, the fully automated, intelligent Data Management systems may actually end up chopping about 6 percent of US-based Data Science jobs. In 2021, we may even see a virtually connected, business ecosystem once again transforming the customer engagement platforms and methods.

III. What Major Industry Watchers Think About Artificial Intelligence and Machine Learning

According to the column *Gartner, IDC, and Forrester on the Future of Digital Transformation*:

"By 2018, 35% of IT resources will be spent to support the creation of new digital revenue streams and by 2020 almost 50% of IT budgets will be tied to digital transformation initiatives. Enterprises pursuing digital transformation initiatives will more than double the

size of their software development teams by 2018, focusing new hires almost entirely on digital initiatives.”

The CXO Today article *7 AI Technologies To Power The Enterprise* indicates that Ericsson has candidly claimed in its 6th Annual Trend Report that about 35 percent of advanced internet users prefer an AI advisor and about 25 percent hope to see an AI Manager next year.

In *Top Technology and IT Predictions for 2017 and Beyond*, the IT industry leader HCL proclaims that industry watchdogs like IDC, Gartner, or Forrester have all claimed that the 2017 will usher in an era of “accelerated digital disruption” making 2018 more technologically significant than ever. The top technologies that will make this era happen are AI^[6], ML, and IoT^[8]. In the newly connected business world, intelligent systems may increasingly collaborate, engage, and solve business problems over disconnected geographies. This is where the virtually connected world will triumph over physical boundaries.

Some of the common forecasts made by the above industry watchers for 2018 and the coming years are:

1. The IT trend of newer digital services with additional revenue streams has started and will gain momentum in 2018 before reaching 40 percent in 2019.
2. The tech trend of “connected” and intelligent systems has started and will gain momentum in 2018 before we see 40 percent of digital transformation and upwards of 70-80 percent of AI-powered, IoT^[8] initiatives in 2019.
3. By 2022, IoT^[8] will save businesses and customers an astronomical sum of \$1 trillion a year in maintenance and services. Some of that will start happening in 2018.
4. In 2018, more analytics vendors will offer IoT^[8]-driven capabilities which may lead global businesses to invest more in AI^[6]-powered, connected technologies. According to Forrester, automated IoT^[8] has already been adopted in transportation industry fleet management, government security and surveillance systems, retail inventory management, and manufacturing assets management.

IV. Application of Machine Learning:

1. Virtual Personal Assistants:

Siri, Alexa, Google Now are some of the popular examples of virtual personal assistants. As the name suggests, they assist in finding information, when asked over voice.

Virtual Assistants are integrated to a variety of platforms. For example:

- Smart Speakers: Amazon Echo and Google Home
- Smartphones: Samsung Bixby on Samsung S8
- Mobile Apps: Google Allo

2. Predictions while Commuting:

Traffic Predictions: We all have been using GPS navigation services. While we do that, our current locations and velocities are being saved at a central server for managing traffic. This data is then used to build a map of current traffic. While this helps in preventing the traffic and does congestion analysis, the underlying problem is that there are less number of cars that are equipped with GPS. Machine learning in such scenarios helps to estimate the regions where congestion can be found on the basis of daily experiences. **Online Transportation Networks:** When booking a cab, the app estimates the price of the ride. When sharing these services, how do they minimize the detours? The answer is machine learning.

3. Videos Surveillance:

Imagine a single person monitoring multiple video cameras! Certainly, a difficult job to do and boring as well. This is why the idea of training computers to do this job makes sense. The video surveillance systems nowadays are powered by AI that makes it possible to detect crime before they happen. They track unusual behaviour of people like standing motionless for a long time, stumbling, or napping on benches etc. The system can thus give an alert to human attendants, which can ultimately help to avoid mishaps. And when such activities are reported and counted to be true, they help to improve the surveillance services. This happens with machine learning doing its job at the backend.

4. Social Media Services:

People You May Know: Machine learning works on a simple concept understanding with experiences. Facebook continuously notices the friends that you connect with, the profiles that you visit very often, your interests, workplace, or a group that you share with someone etc.

Face Recognition: You upload a picture of you with a friend and Facebook instantly recognizes that friend. Facebook checks the poses and projections in the picture, notice the unique features, and then match them with the people in your friend list.

5. Online Customer Support:

A number of websites nowadays offer the option to chat with customer support representative while they are navigating within the site. However, not every website has a live executive to answer your queries. In most of the cases, you talk to a chatbot^[7]. These bots tend to extract information from the website and present it to the customers. Meanwhile, the chatbots^[7] advances with time. They tend to understand the user queries better and serve them with better answers, which is possible due to its machine learning algorithms. Machine learning is an application of artificial intelligence (AI^[6]) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

V. Some machine learning methods:

Machine learning algorithms are often categorized as supervised or unsupervised. Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly.

In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data.

Semi-supervised machine learning algorithms fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training – typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semi-supervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring unlabeled data generally doesn't require additional resources. Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal. Machine learning enables analysis of massive quantities of data. While it generally delivers faster, more accurate results in order to identify profitable opportunities or dangerous risks, it may also require additional time and resources to train it properly. Combining machine learning with AI and cognitive technologies can make it even more effective in processing large volumes of information.

VI. Statistics and Machine Learning at Scale

Those driverless cars we keep hearing about from Google? They're just one example of machine learning – and we're about to hear of many more. Even though the concept of machine learning has been around for decades, it's now becoming more pertinent to our daily lives because of increased amounts of data and cheaper storage. This paper, which is based on a presentation given at the Analytics Conference, introduces key machine learning concepts and introduces SAS solutions – SAS In-Memory Statistics for Hadoop and SAS Visual Statistics – that enable machine learning at scale.

VII. Some popular Machine Learning methods

Two of the most widely adopted machine learning methods are supervised learning and unsupervised learning – but there are also other methods of machine learning. Here's an overview of the most popular types.

Supervised learning algorithms are trained using labeled examples, such as an input where the desired output is known. For example, a piece of equipment could have data points labeled either "F" (failed) or "R" (runs). The learning algorithm receives a set of inputs along with the corresponding correct outputs, and the algorithm learns by comparing its actual output with correct outputs to find errors. It then modifies the model accordingly. Through methods like classification, regression, prediction and gradient boosting, supervised learning uses patterns to predict the values of the label on additional unlabeled data. Supervised learning is commonly used in applications where historical data predicts likely future events. For example, it can anticipate when credit card transactions are likely to be fraudulent or which insurance customer is likely to file a claim.

Unsupervised learning is used against data that has no historical labels. The system is not told the "right answer." The algorithm must figure out what is being shown. The goal is to explore the data and find some structure within. Unsupervised learning works well on transactional data. For example, it can identify segments of customers with similar attributes who can then be treated similarly in marketing campaigns. Or it can find the main attributes that separate customer segments from each other. Popular techniques include self-organizing maps, nearest-neighbor mapping, k-means clustering and singular value decomposition. These algorithms are also used to

segment text topics, recommend items and identify data outliers. Semi supervised learning is used for the same applications as supervised learning. But it uses both labeled and unlabeled data for training – typically a small amount of labeled data with a large amount of unlabeled data (because unlabeled data is less expensive and takes less effort to acquire). This type of learning can be used with methods such as classification, regression and prediction. Semi supervised learning is useful when the cost associated with labeling is too high to allow for a fully labeled training process. Early examples of this include identifying a person's face on a web cam. Reinforcement learning is often used for robotics, gaming and navigation. With reinforcement learning, the algorithm discovers through trial and error which actions yield the greatest rewards. This type of learning has three primary components: the agent (the learner or decision maker), the environment (everything the agent interacts with) and actions (what the agent can do). The objective is for the agent to choose actions that maximize the expected reward over a given amount of time. The agent will reach the goal much faster by following a good policy. So the goal in reinforcement learning is to learn the best policy.

Humans can typically create one or two good models a week; machine learning can create thousands of models a week.

Reference:

- [1]: Name designation, etc of the 1st author.
- [2]: Name designation, etc of the 2nd author.
- [3]: Name designation, etc of the 3rd author.
- [4]: Name designation, etc of the 4th author.
- [5]: Samuel, Arthur (1959) "Some studies in machine learning using the game of checkers". IBM journal of research and development.
- [6]: Artificial intelligence, sometimes called machine intelligence, is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals.
- [7]: A chatbot (also known as a talkbot, chatterbot, Bot, IM bot, interactive agent, or Artificial Conversational Entity) is a computer program or an artificial intelligence which conducts a conversation via auditory or textual methods.
- [8]: IoT is short for Internet of Things. The Internet of Things refers to the ever-growing network of physical objects that feature an IP address for internet connectivity, and the communication that occurs between these objects and other Internet-enabled devices and systems.
- [9]: Augmented Reality (AR) is an interactive experience of a real-world environment whereby the objects that reside in the real-world are "augmented" by computer-generated perceptual information, sometimes across multiple sensory modalities, including visual, auditory, haptic, somatosensory, and olfactory.
- [10]: Virtual reality (VR) is an interactive computer-generated experience taking place within a simulated environment, that incorporates mainly auditory and visual, but also other types of sensory feedback like haptic.