ISSN: 2321-7782 (Online)
e-ISJN: A4372-3114
Impact Factor: 7.327

Volume 6, Issue 9, September 2018

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

Machine Learning and its applications

Dr. Pooja Nagpal
Assistant Professor,
Hindu Institute of Management & Technology,
Rohtak, India.

Abstract: ***Machine learning (ML) equips computers to learn and interpret without being explicitly programmed to do so. Here, as the "computers", also referred as the "models", are exposed to sets of new data, they adapt independently and learn from earlier computations to interpret available data and identify hidden patterns. This involves data analysis and automation of analytical model-building using numerous ML algorithms. ML enables computers and computing machines to search for and identify hidden insights, without being programmed for where to look for, when exposed to new data sets.***

Keywords: Machine Learning, Deep Learning, Big Data, Artificial Intelligence.

I. INTRODUCTION

Machine learning is different from what it used to be in the past, due to the emergence of advanced computing technologies. Initially, it had gained momentum due to pattern recognition and the fact that computers did not have to be programed to execute certain tasks to learn. Many researchers who were interested in Artificial Intelligence (AI) investigated this area further to find out whether computers could really learn from data or not.The focus here is on iterative learning. Machines begin to adapt to new data that they are exposed to, over a period. Based on the patterns and computations that are previously created, machines learn to repeat decisions made in the past, in similar situations. This aspect of machines' ability to learn from the existing patterns, is now gaining huge momentum.

Today, people are sitting up and taking notice of the fact that machines are now able to apply complicated mathematical calculations to areas, such as big data, at a much faster rate. Consider Google Car for instance, which is primarily built on the crux of machine learning. Another important use of machine learning can be found in regular recommendations that are rolled out by companies like Netflix and Amazon - an example of machine learning in everyday life. Next, ML can also be combined with linguistic rules creation. This application is implemented by Twitter, where you will know what customers say about you. And not to forget, machine learning is significantly being used to detect fraud in various industry sectors.

To get the maximum value from big data, businesses must know exactly how to pair the right algorithm with a particular tool or process and build machine learning models based on iterative learning processes. Some of the key machine learning algorithms is -

- Random forests
- Neural networks
- Discovery of sequence and associations
- Decision trees
- Mapping of nearest neighbor

- Supporting vector machines
- Boosting and bagging gradient
- Self organizing maps
- Multivariate adaptive regression
- SEO
- Analysis of principal components

Machine Learning Methods

Machine Learning is an application of Artificial Intelligence that provides computers and systems the capability to learn automatically from the previous data without being specifically programmed. It is currently the hot field both for research and thesis. There are algorithms designed for machine learning process such that the systems can identify patterns from data-sets and make decisions accordingly. Machine Learning algorithms are categorized into following types:

Supervised Learning

These algorithms are trained using labeled examples, in different scenarios, as an input where the desired outcome is already known. An equipment, for instance, could have data points such as "F" and "R" where "F" represents "failed" and "R" represents "runs".

A learning algorithm will receive a set of input instructions along with the corresponding accurate outcomes. The learning algorithm will then compare the actual outcome with the accurate outcome and flag an error, if there is any discrepancy. Using different methods, such as regression, classification, gradient boosting, and prediction, supervised learning uses different patterns to proactively predict the values of a label on extra unlabeled data. This method is commonly used in areas where historical data is used to predict events that are likely to occur in the future. For instance, anticipate when a credit card transaction is likely to be fraudulent or predict which insurance customers are likely to file their claims.

- Supervised learning can be further classified into:
    - **Classification:** A problem in which an output variable is a category.
    - **Regression:** Output variable is a real value.

Unsupervised Learning

This method of ML finds its application in areas were data has no historical labels. Here, the system will not be provided with the "right answer" and the algorithm should identify what is being shown. The main aim here is to analyze the data and identify a pattern and structure within the available data set. Transactional data serves as a good source of data set for unsupervised learning.

For instance, this type of learning identifies customer segments with similar attributes and then lets the business to treat them similarly in marketing campaigns. Similarly, it can also identify attributes that differentiate customer segments from one another. Either ways, it is about identifying a similar structure in the available data set. Besides, these algorithms can also identify outliers in the available data sets.

- It can be further divided into:
    - **Clustering:** In this type of problem, data with similar patterns are grouped together.

Association: In this problem, rules are discovered that describes the data

Some of the widely used techniques of unsupervised learning are -

- k-means clustering
- self-organizing maps
- value decomposition
- mapping of nearest neighbor

Semi-supervised Learning

This kind of learning is used and applied to the same kind of scenarios where supervised learning is applicable. However, one must note that this technique uses both unlabeled and labeled data for training. Ideally, a small set of labeled data, along with a large volume of unlabeled data is used, as it takes less time, money and efforts to acquire unlabeled data. This type of machine learning is often used with methods, such as regression, classification and prediction. Companies that usually find it challenging to meet the high costs associated with labeled training process opt for semi-supervised learning.

Reinforcement Learning

This is mainly used in navigation, robotics and gaming. Actions that yield the best rewards are identified by algorithms that use trial and error methods. There are three major components in reinforcement learning, namely, the agent, the actions and the environment. The agent in this case is the decision maker, the actions are what an agent does, and the environment is anything that an agent interacts with. The main aim in this kind of learning is to select the actions that maximize the reward, within a specified time. By following a good policy, the agent can achieve the goal faster.

Hence, the primary idea of reinforcement learning is to identify the best policy or the method that helps businesses in achieving the goals faster. While humans can create a few good models in a week, machine learning is capable of developing thousands of such models in a week.

II. APPLICATIONS MACHINE LEARNING

The value of machine learning technology has been recognized by companies across several industries that deal with huge volumes of data. By leveraging insights obtained from this data, companies are able work in an efficient manner to control costs as well as get an edge over their competitors. This is how some sectors / domains are implementing machine learning -

Financial Services

Companies in the financial sector are able to identify key insights in financial data as well as prevent any occurrences of financial fraud, with the help of machine learning technology. The technology is also used to identify opportunities for investments and trade. Usage of cyber surveillance helps in identifying those individuals or institutions which are prone to financial risk, and take necessary actions in time to prevent fraud.

Marketing and Sales

Companies are using machine learning technology to analyze the purchase history of their customers and make personalized product recommendations for their next purchase. This ability to capture, analyze, and use customer data to provide a personalized shopping experience is the future of sales and marketing.

Government

Government agencies like utilities and public safety have a specific need FOR Ml, as they have multiple data sources, which can be mined for identifying useful patterns and insights. For example sensor data can be analyzed to identify ways to minimize costs and increase efficiency. Furthermore, ML can also be used to minimize identity thefts and detect fraud.

Healthcare

With the advent of wearable sensors and devices that use data to access health of a patient in real time, ML is becoming a fast-growing trend in healthcare. Sensors in wearable provide real-time patient information, such as overall health condition, heartbeat, blood pressure and other vital parameters. Doctors and medical experts can use this information to analyze the health condition of an individual, draw a pattern from the patient history, and predict the occurrence of any ailments in the future. The technology also empowers medical experts to analyze data to identify trends that facilitate better diagnoses and treatment.

Transportation

Based on the travel history and pattern of traveling across various routes, machine learning can help transportation companies predict potential problems that could arise on certain routes, and accordingly advise their customers to opt for a different route. Transportation firms and delivery organizations are increasingly using machine learning technology to carry out data analysis and data modeling to make informed decisions and help their customers make smart decisions when they travel.

Oil and Gas

This is perhaps the industry that needs the application of machine learning the most. Right from analyzing underground minerals and finding new energy sources to streaming oil distribution, ML applications for this industry are vast and are still expanding.

Some Machine Learning Algorithms and Processes

If you're studying machine learning, you should familiarize yourself with these common machine learning algorithms and processes: neural networks, decision trees, random forests, associations and sequence discovery, gradient boosting and bagging, support vector machines, self-organizing maps, k-means clustering, Bayesian networks, Gaussian mixture models, and more.

Other tools and processes that pair up with the best algorithms to aid in deriving the most value from big data include:

- Comprehensive data quality and management
- GUIs for building models and process flows
- Interactive data exploration and visualization of model results
- Comparisons of different machine learning models to quickly identify the best one
- Automated ensemble model evaluation to identify the best performers
- Easy model deployment so you can get repeatable, reliable results quickly
- Integrated end-to-end platform for the automation of the data-to-decision process

III. Conclusion

In this paper we looked at what machine learning is, how it was first introduced into the world of remote sensing, what a typical workflow is like, and what kind of problems are being solved using machine learning. Machine learning has a bright future because more and more people are learning the basics of machine learning and applying it in their regular jobs and researches. New algorithms are cropping up every other day, and the accuracy rate of classifications is improving along with it. Those problems in remote sensing (mapping land cover) which seemed difficult and sometimes impossible are being solved by new algorithms every single day. It is not far-fetched to say that most analysis work done in the world today will be done by machine learning algorithms in near future.

References

1. Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31, doi:10.1016/j.isprsjprs.2016.01.011
2. Rodriguez-Galiano, V. F., & Chica-Rivas, M. (2014). Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. International Journal of Digital Earth, 7(6), 492-509, doi:10.1080/17538947.2012.748848
3. Shang, X., & Chisholm, L. A. (2014). Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(6), 2481-2489, doi:10.1109/JSTARS.2013.2282166
4. Nath, R. K., & Deb, S. K. (2010). Water-body area extraction from high resolution satellite images-an introduction, review, and comparison. International Journal of Image Processing (IJIP), 3(6), 353-372,
5. Li, Y., Tao, C., Tan, Y., Shang, K., & Tian, J. (2016). Unsupervised multilayer feature learning for satellite image scene classification. IEEE Geoscience and Remote Sensing Letters, 13(2), 157-161, doi:10.1109/LGRS.2015.2503142
6. Huang, X., & Jensen, J. R. (1997). A machine-learning approach to automated knowledge-base building for remote sensing image analysis with GIS data. Photogrammetric engineering and remote sensing, 63(10), 1185-1193.
7. Mitra, P., Shankar, B. U., & Pal, S. K. (2004). Segmentation of multispectral remote sensing images using active support vector machines. Pattern recognition letters, 25(9), 1067-1074, doi:10.1016/j.patrec.2004.03.004
8. Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on geoscience and remote sensing, 42(8), 1778-1790, doi:10.1109/TGRS.2004.831865
9. Kubat, M., Holte, R. C., & Matwin, S. (1998). Machine learning for the detection of oil spills in satellite radar images. Machine learning, 30(2-3), 195-215, doi:10.1023/A:1007452223027
10. Maggiori, E., Tarabalka, Y., Charpiat, G., & Alliez, P. (2017). Convolutional neural networks for large-scale remote-sensing image classification. IEEE Transactions on Geoscience and Remote Sensing, 55(2), 645-657, doi:10.1109/TGRS.2016.2612821
11. Pedregosa, F., Varoquaux, G., Gramfort, A. & Michel, V. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
12. Jensen, R. R., Hardin, P. J. and Yu, G. (2009), Artificial Neural Networks and Remote Sensing. Geography Compass, 3: 630–646. doi:10.1111/j.1749-8198.2008.00215.x.

