Machine Learning Approaches to Standard Prediction Rules

Dr. RAMANA NAGAVELLI
Assistant Professor, Department of Computer Science & Engineering, University College of Engineering,
Kakatiya University, Warangal, Telangana, India.

Abstract:

Machine Learning (ML) is a sub-field of Artificial Intelligence (AI) which concerns with developing computational theories of learning and building learning machines. The goal of machine learning, closely coupled with the goal of AI, is to achieve a thorough understanding about the nature of learning process (both human learning and other forms of learning), about the computational aspects of learning behaviors, and to implant the learning capability in computer systems. Machine learning has been recognized as central to the success of Artificial Intelligence, and it has applications in various areas of science, engineering and society. 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.

Key points: Machine Learning, Artificial Intelligence, Learning Tools,

I. Introduction

Machine Learning (ML) is an automated learning with little or no human intervention. It involves programming computers so that they learn from the available inputs. The main purpose of machine learning is to explore and construct algorithms that can learn from the previous data and make predictions on new input data. The **input** to a learning algorithm is training data, representing experience, and the **output** is any expertise, which usually takes the form of another algorithm that can perform a task. The input data to a machine learning system can be numerical, textual, audio, visual, or multimedia. The corresponding output data of the system can be a floating-point number, for instance, the velocity of a rocket, an integer representing a category or a class, for example, a pigeon or a sunflower from image recognition. it would be useful to discuss the issue "what is learning". Learning is a phenomenon and process which has manifestations of various aspects. Roughly speaking, learning

Process includes (one or more of) the following:

(1) Acquisition of new (symbolic) knowledge. For example, learning mathematics is this kind of learning. When we say someone has learned math, we mean that the learner obtained descriptions of the mathematical concepts, understood their meaning and their relationship with each other. The effect of learning is that the learner has acquired knowledge of mathematical systems and their properties, and that the learner can use this knowledge to solve math problems. Thus this kind of learning is characterized as obtaining new symbolic information plus the ability to apply that information effectively.

(2) Development of motor or cognitive skills through instruction and practice. Examples of this kind of learning are learning to ride a bicycle, to swim, to play piano, etc. This kind of learning is also called skill refinement. In this case, just acquiring a symbolic description of the rules to perform the task is not sufficient, repeated practice is needed for the learner to obtain the skill. Skill refinement takes place at the subconscious level.

(3) Refinement and organization of knowledge into more effective representations or more useful form. One example of this kind of learning can be reorganization of the rules in a knowledge base such that more important rules are given higher priorities so that they can be used more easily and conveniently.

(4) Discovery of new facts and theories through observation and experiment. For example, the discovery of physics and chemistry laws. The general effect of learning in a system is the improvement of the system's capability to solve problems. It is hard to imagine a system capable of learning cannot improve its problem-solving performance. A system with learning capability should be able to do self-changing in order to perform better in its future problem-solving.

Concepts of Learning

Learning is the process of converting experience into expertise or knowledge. Learning can be broadly classified into three categories, as mentioned below, based on the nature of the learning data and interaction between the learner and the environment.

- Supervised Learning
- Unsupervised Learning
- Semi-supervised Learning
- Reinforcement machine learning

Supervised Learning

This algorithm consists of a target or outcome or dependent variable which is predicted from a given set of predictor or independent variables. Using these set of variables, we generate a function that maps input variables to desired output variables. The training process continues until the model achieves a desired level of accuracy on the training data. Examples of Supervised Learning - Regression, Decision Tree, Random Forest, KNN, Logistic Regression etc.

Unsupervised Learning

In this algorithm, there is no target or outcome or dependent variable to predict or estimate. It is used for clustering a given data set into different groups, which is widely used for segmenting customers into different groups for specific intervention. Apriority algorithm and K-means are some of the examples of Unsupervised Learning.

Semi-supervised machine learning

This Algorithm falls 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

It 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.

Reinforcement Learning

Using this algorithm, the machine is trained to make specific decisions. Here, the algorithm trains itself continually by using trial and error methods and feedback methods. This machine learns from past experiences and tries to capture the best possible knowledge to make accurate business decisions.

II. Challenges and Limitations of Machine learning

The primary challenge of machine learning is the lack of data or the diversity in the dataset. A machine cannot learn if there is no data available. Besides, a dataset with a lack of diversity gives the machine a hard time. A machine needs to have heterogeneity to learn meaningful insight. It is rare that an algorithm can extract information when there are no or few variations. It is recommended to have at least 20 observations per group to help the machine learn. This constraint leads to poor evaluation and prediction.

Applications

Artificial Intelligence (AI) and Machine Learning are everywhere. Chances are that you are using them and not even aware about that. In Machine Learning (ML), computers, software, and devices perform via cognition similar to human brain.

Typical successful applications of machine learning include programs that decode handwritten text, face recognition, voice recognition, speech recognition, pattern recognition, spam detection programs, weather forecasting, stock market analysis and predictions, and so on.

Virtual Personal Assistants

Siri, Google Now, Alexa are some of the common examples of virtual personal assistants. These applications assist in finding information, when asked over voice. All that is needed is activating them and asking questions like for example "What are my appointments for today?", "What are the flights from Delhi to New York". For answering such queries, the application looks out for the information, recalls your previous queries, and accesses other resources to collect relevant information. You can even tell these assistants to do certain tasks like "Set an alarm for 5.30 AM next morning", "Remind me to visit Passport office tomorrow at 10.30 am".

Traffic Congestion Analysis and Predictions

GPS navigation services monitor the user's location and velocities and use them to build a map of current traffic. This helps in preventing the traffic congestions. Machine learning in such scenarios helps to estimate the regions where congestion can be found based on previous records.

Automated Video Surveillance

Video surveillance systems nowadays are powered by AI and machine learning is the technology behind this that makes it possible to detect and prevent crimes before they occur. They track odd and suspicious behavior of people and sends alerts to human attendants, who can ultimately help accidents and crimes.

Social Media

Face book continuously monitors the friends that you connect with, your interests, workplace, or a group that you share with someone etc. Based on continuous learning, a list of Face book users is given as friend suggestions.

Face Recognition

You upload a picture of you with a friend and Face book instantly recognizes that friend. Machine learning works at the core of Computer Vision, which is a technique to extract useful information from images and videos. Interest uses computer vision to identify objects or pins in the images and recommend similar pins to its users.

Email Spam and Malware Filtering

Machine learning is being extensively used in spam detection and malware filtering and the databases of such spam and malwares keep on getting updated so these are handled efficiently.

Online Customer Support

In several websites nowadays, there is an option to chat with customer support representative while users are navigating the site. In most of the cases, instead of a real executive, you talk to a chat boat. These bots extract information from the website and provide it to the customers to assist them. Over a period of time, the chat bots learn to understand the user queries better and serve them with better answers, and this is made possible by machine learning algorithms.

Refinement of Search Engine Results

Google and similar search engines are using machine learning to improve the search results for their users. Every time a search is executed, the algorithms at the backend keep a watch at how the users respond to the results. Depending on the user responses, the algorithms working at the backend improve the search results.

Product Recommendations

If a user purchases or searches for a product online, he/she keeps on receiving emails for shopping suggestions and ads about that product. Based on previous user behavior, on a website/app, past purchases, items liked or added to cart, brand preferences etc., the product recommendations are sent to the user.

III. Benefits and Assessments of Machine Learning

- Google and facebook are using machine learning to push relevant advertisements. That advertisement is based on users past search behavior.
- Machine learning is used to handle multi-dimensional and multi-variety data in dynamic environments.
- Machine learning allows time cycle reduction and efficient utilization of resources.
- If one wants to provide continuous quality, large and complex process environments. There are some tools present because of machine learning.
- As there are too many things that come under the practical benefit of machine learning. Also, they involve the development of autonomous computers, software programs. Hence, it includes processes that can lead to the automation of tasks.

Analysis of Machine Learning:

The process used by teachers to select appropriate reading material for their students is complicated and subjective, taking into account subject matter as well as characteristics of the text itself. For example, Fountas and Pinnell's well-known system of matching books to readers takes into account more than a dozen high level characteristics, including vocabulary, grammatical structure of phrases and sentences, use of literary devices, illustrations, and layout on the page (Fountas and Pinnell, 1999). Automatic tools cannot capture this entire range of characteristics, but a variety of methods and formulae have been developed to calculate approximations of reading level based on characteristics which are easily measured.

Many traditional formulae for reading level assessment focus on simple approximations of syntactic complexity such as sentence length. The widely-used Flesch-Kincaid Grade Level index is based on the average number of syllables per word and the average number of words per sentence in a passage of text (Kincaid et al, 1975) (as cited by Collins-Thompson and Callan (2005)). Similarly, the Gunning Fog index is based on the average number of words per sentence and the percentage of words with three or more syllables (Gunning, 1952).

These methods are quick and easy to calculate but have drawbacks: sentence length is not always an accurate measure of syntactic complexity, and syllable count does not necessarily indicate the difficulty of a word. Also, a student may be familiar with a few complex words (e.g., dinosaur names) but unable to understand complex syntactic constructions. Other measures of readability focus on semantics, which is usually approximated by word frequency with respect to a reference list or corpus. In addition to the traditional reading level metrics, researchers at Carnegie Mellon University have applied probabilistic language modeling techniques to this task. Si and Callan (2001) conducted preliminary work to classify science web pages using unigram models. More recently, Collins-Thompson and Callan manually collected a corpus of web pages ranked by grade level and observed that vocabulary words are not distributed evenly across grade levels. They developed a ‘‘smoothed unigram”. . Therefore, we can say that Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks. Suppose we want to make a system that can recognize faces of different people in an image. If we solve this as a typical machine learning problem.

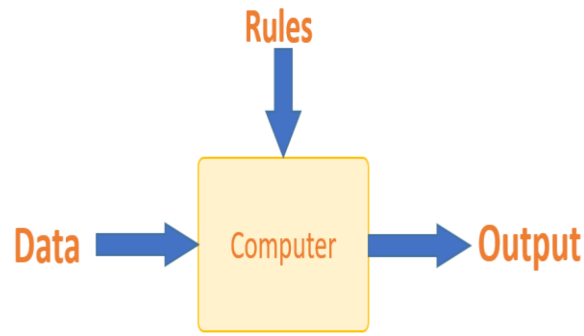

Fig 1: Traditional Learning

Classifier to better capture the variance in word usage across grade levels (Collins-Thompson and Callan, 2005). On web text, their classifier outperformed several other measures of semantic difficulty: the fraction of unknown words in the text, the mean log frequency of the text relative to a large corpus, and the Flesch-Kincaid measure. The traditional measures performed better on some commercial corpora, but these corpora were calibrated using similar measures, so it is arguably not a fair comparison. More importantly, the smoothed unigram measure worked better on the web corpus, especially on short passages.

IV. Machine Learning vs. Traditional Programming

Traditional programming differs significantly from machine learning. In traditional programming, a programmer codes all the rules in consultation with an expert in the industry for which software is being developed. Each rule is based on a logical foundation; the machine will execute an output following the logical statement. When the system grows complex, more rules need to be written. It can quickly become unsustainable to maintain.

Machine learning is supposed to overcome this issue. The machine learns how the input and output data are correlated and it writes a rule. The programmers do not need to write new rules each time there is new data. The algorithms adapt in response to new data and experiences to improve efficacy over time.

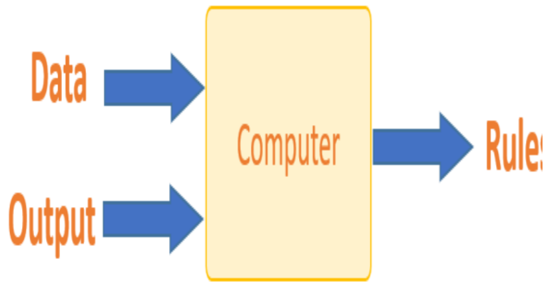

Fig 2: Machine Learning

Working Process of Machine Learning:

Machine learning is the brain where all the learning takes place. The way the machine learns is similar to the human being. Humans learn from experience. The more we know, the more easily we can predict. By analogy, when we face an unknown situation, the likelihood of success is lower than the known situation. Machines are trained the same. To make an accurate prediction, the machine sees an example. When we give the machine a similar example, it can figure out the outcome. However, like a human, if its feed a previously unseen example, the machine has difficulties to predict.

The core objective of machine learning is the **learning** and **inference**. First of all, the machine learns through the discovery of patterns. This discovery is made thanks to the **data**. One crucial part of the data scientist is to choose carefully which data to provide to the machine. The list of attributes used to solve a problem is called a **feature vector.** You can think of a feature vector as a subset of data that is used to tackle a problem on Learning Mode.

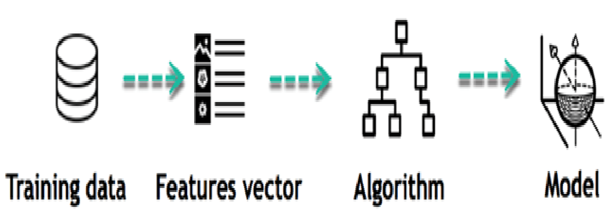

Fig 3: Working Procedure of Machine Learning Algorithm

The machine uses some fancy algorithms to simplify the reality and transform this discovery into a **model**. Therefore, the learning stage is used to describe the data and summarize it into a model. For instance, the machine is trying to understand the relationship between the wage of an individual and the likelihood to go to a fancy restaurant. It turns out the machine finds a positive relationship between wage and going to a high-end restaurant: This is the model

Inferring Mode

When the model is built, it is possible to test how powerful it is on never-seen-before data. The new data are transformed into a features vector, go through the model and give a prediction. This is all the beautiful part of machine learning. There is no need to update the rules or train again the model. You can use the model previously trained to make inference on new data.

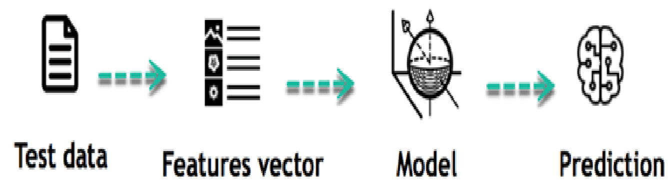

Fig 4: Working Procedure of Traditional Learning Algorithm

Traditional programming relies on hard-coded and Machine Learning relies on learning patterns based on sample data. As we go from rule-based systems to the deep learning ones, more complex features and input-output relationships become learnable.

V. Conclusion

Machine learning is quickly growing field in computer science. It has applications in nearly every other field of study and is already being implemented commercially because machine learning can solve problems too difficult or time consuming for humans to solve. To describe machine learning in general terms, a variety models are used to learn patterns in data and make accurate predictions based on the patterns it observes. First, I introduced generalization and over fitting. Both of these topics are tied to supervise learning, which uses training data to train the model. Generalization is when a machine learning model can accurately predict results from data it hasn't seen before. Over fitting happens when a model learns the training data too well and cannot generalize. Under fitting, the opposite of over fitting, can also happen with supervised learning. With under fitting, the model is unable to make accurate predictions with both training data and new data.

VI. References:

[1] E. Bauer and R. Kohavi. An empirical comparison of voting classification algorithms: Bagging, boosting and variants. Machine Learning, 36:105–142, 1999.

[2] Y. Bengio. Learning deep architectures for AI. Foundations and Trends in Machine Learning, 2:1–127, 2009.

[3] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B, 57:289–300, 1995.

[4] J. M. Bernardo and A. F. M. Smith. Bayesian Theory. Wiley, New York, NY, 1994.

[5] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. K. Warmuth. Occam's razor. Information Processing Letters, 24:377–380, 1987.

[6] W. W. Cohen. Grammatically biased learning: Learning logic programs using an explicit antecedent description language. Artificial Intelligence, 68:303–366, 1994.

[7] P. Domingos. The role of Occam's razor in knowledge discovery. Data Mining and Knowledge Discovery,3:409–425, 1999.

[8] P. Domingos. Bayesian averaging of classifiers and the overfitting problem. In Proceedings of the Seventeenth International Conference on Machine Learning, pages 223–230, Stanford, CA, 2000. Morgan Kaufmann.

[9] P. Domingos. A unified bias-variance decomposition and its applications. In Proceedings of the Seventeenth International Conference on Machine Learning, pages 31–238, Stanford, CA, 2000. Morgan Kaufmann.

[10] P. Domingos. The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World. Basic Books, New York, NY, 2015.

[11] P. Domingos and M. Pazzani. On the optimality of the simple Bayesian classifier under zero-one loss. Machine Learning, 29:103–130, 1997.

[12] G. Hulten and P. Domingos. Mining complex models from arbitrarily large databases in constant time. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 525–531, Edmonton, Canada, 2002. ACM Press.

[13] D. Kibler and P. Langley. Machine learning as an experimental science. In Proceedings of the Third European Working Session on Learning, London, UK, 1988. Pitman.

[14] A. J. Klockars and G. Sax. Multiple Comparisons. Sage, Beverly Hills, CA, 1986.

[15] R. Kohavi, R. Longbotham, D. Sommerfield, and R.Henne. Controlled experiments on the Web: Survey and practical guide. Data Mining and Knowledge Discovery, 18:140–181, 2009.

[16] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. Byers. Big data: The next frontier for innovation, competition, and productivity. Technical report, McKinsey Global Institute, 2011.

[17] T. M. Mitchell. Machine Learning. McGraw-Hill, New York, NY, 1997.

[18] A. Y. Ng. Preventing "overfitting" of cross-validation data. In Proceedings of the Fourteenth International Conference on Machine Learning, pages 245–253, Nashville, TN, 1997. Morgan Kaufmann.

[19] J. Pearl. On the connection between the complexity and credibility of inferred models. International Journal of General Systems, 4:255–264, 1978.

[20] J. Pearl. Causality: Models, Reasoning, and Inference. Cambridge University Press, Cambridge, UK, 2000.