A REVIEW ON MACHINE LEARNING: TRENDS AND FUTURE PROSPECTS

Manish Kumar Aery[1], Chet Ram[2]
Assistant Professor, Dept. of MCA,IET Bhaddal, Ropar, Punjab, India.
Reserch Scholar, Dept. of CSE,IET Bhaddal, Ropar, Punjab, India.

Abstract

Machine learning addresses the question of how to build computers that improve automatically through experience. It is one of today's most rapidly growing technical fields, lying at the intersection of computer science and statistics, and at the core of artificial intelligence and data science. Recent progress in machine learning has been driven both by the development of new learning algorithms and theory and by the ongoing explosion in the availability of online data and low-cost computation. The adoption of data-intensive machine-learning methods can be found throughout science, technology and commerce, leading to more evidence-based decision-making across many walks of life, including health care, manufacturing, education, financial modeling, policing, and marketing.

Key Words: Machine, Computer Science, statics, technology

1. INTRODUCTION

The Machine Learning field evolved from the broad field of Artificial Intelligence, which aims to mimic intelligent abilities of humans by machines. In the field of Machine Learning one considers the important question of how to make machines able to "learn". Learning in this context is understood as inductive inference, where one observes examples that represent incomplete information about some "statistical phenomenon". In unsupervised learning one typically tries to uncover hidden regularities (e.g. clusters) or to detect anomalies in the data (for instance some unusual machine function or a network intrusion).In supervised learning, there is label associated with each example. It is supposed to be the answer to a question about the example. If the label is discrete, then the task is called classification problem– otherwise, for real-valued labels we speak of aregression problem. Based on these examples (including the labels), one is particularly interested to predict the answer for other cases before they are explicitly observed. Hence, learning is not only a question of remembering but also of generalization to unseen cases.

2. NEED MACHINE LEARNING

To better understand the uses of machine learning, consider some of the instances where machine learning is applied: the self-driving Google car, cyber fraud detection, online recommendation engines—like friend suggestions on Facebook, Netflix showcasing the movies and shows you might like, and "more items to consider" and "get yourself a little something" on Amazon—are all examples of applied machine learning.

All these examples echo the vital role machine learning has begun to take in today's data-rich world. Machines can aid in filtering useful pieces of information that help in major advancements, and we are already seeing how this technology is being implemented in a wide variety of industries.

With the constant evolution of the field, there has been a subsequent rise in the uses, demands, and importance of machine learning. Big data has become quite a buzzword in the last few years; that's in part due to increased sophistication of machine learning, which helps analyze those big chunks of big data. Machine learning has also changed the way data extraction, and interpretation is done by involving automatic sets of generic methods that have replaced traditional statistical techniques.

The process flow depicted here represents how machine learning works

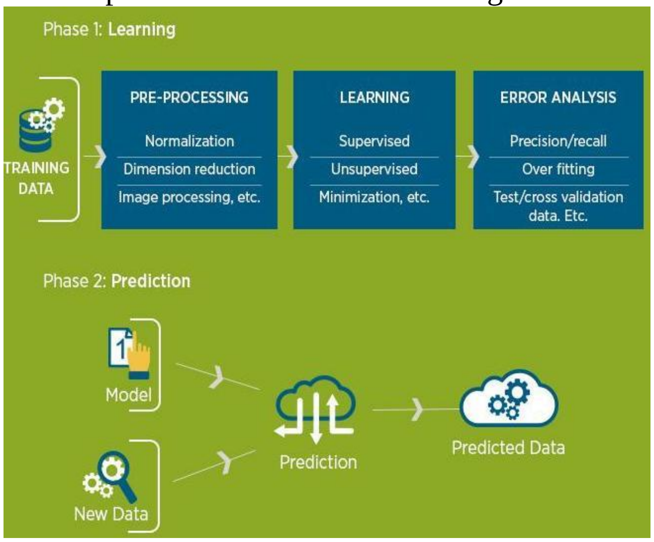

Fig. 1 Machine Learning Process

1. Why Machine Learning Matters?

With the **rise in big data, machine learning has become a key technique for solving problems in areas, such as:**

- **Computational finance, for credit scoring and algorithmic trading**
- **Image processing and computer vision, for face recognition, motion detection, and object detection**
- **Computational biology, for tumor detection, drug discovery, and DNA sequencing**
- **Energy production, for price and load forecasting**
- **Automotive, aerospace, and manufacturing, for predictive maintenance**
- **Natural language processing, for voice recognition applications**

2. Machine Learning Working.

Machine learning uses two types of techniques: supervised learning, which trains a model on known input and output data so that it can predict future outputs, and unsupervised learning, which finds hidden patterns or intrinsic structures in input data.

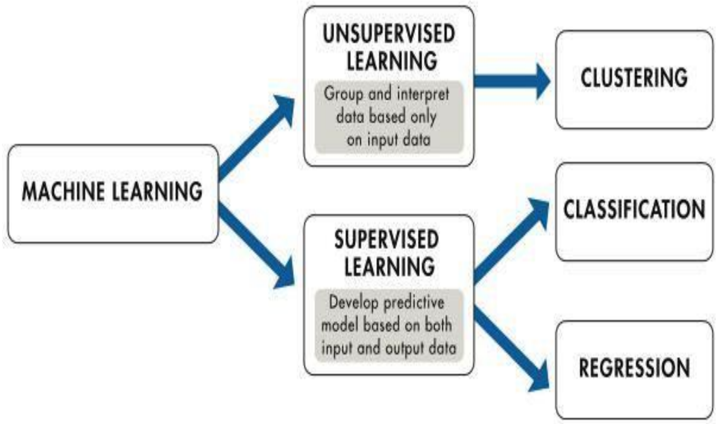

Fig. 2: Machine learning techniques include both unsupervised and supervised learning.

3 SUPERVISED LEARNING

Supervised machine learning builds a model that makes predictions based on evidence in the presence of uncertainty. A supervised learning algorithm takes a known set of input data and known responses to the data (output) and trains a model to generate reasonable predictions for the response to new data. Use supervised learning if you have known data for the output you are trying to predict. Supervised learning uses classification and regression techniques to develop predictive models.

Classification techniques

predict discrete responses for example, whether an email is genuine or spam, or whether a tumor is cancerous or benign. Classification models classify input data into categories. Typical applications include medical imaging, speech recognition, and credit scoring.

Use classification if your data can be tagged, categorized, or separated into specific groups or classes. For example, applications for hand-writing recognition use classification to recognize letters and numbers. In image processing and computer vision, unsupervised pattern recognition techniques are used for object detection and image segmentation.

Common algorithms for performing classification include support vector machine (SVM), boosted and bagged decision trees, *k*-nearest neighbor, Naïve Bayes, discriminant analysis, logistic regression, and neural networks.

Regression techniques

predict continuous responses—for example, changes in temperature or fluctuations in power demand. Typical applications include electricity load forecasting and algorithmic trading.

Use regression techniques if you are working with a data range or if the nature of your response is a real number, such as temperature or the time until failure for a piece of equipment.

Common regression algorithms include linear model, nonlinear model, regularization, stepwise regression, boosted and bagged decision trees, neural networks, and adaptive neuro-fuzzy learning.

4 UNSUPERVISED LEARNING

Unsupervised learning finds hidden patterns or intrinsic structures in data. It is used to draw inferences from datasets consisting of input data without labeled responses.

Clustering

is the most common unsupervised learning technique. It is used for exploratory data analysis to find hidden patterns or groupings in data. Applications for cluster analysis include gene sequence analysis, market research, and object recognition.

For example, if a cell phone company wants optimize the locations where they build cell phone towers, they can use machine learning to estimate the number of clusters of people relying on their towers. A phone can only talk to one tower at a time, so the team uses clustering algorithms to design the best placement of cell towers to optimize signal reception for groups, or clusters, of their customers.

Common algorithms for performing clustering include k-means and k-medoids, hierarchical clustering, Gaussian mixture models, hidden Markov models, self-organizing maps,fuzzy c-means clustering, and subtractive clustering.

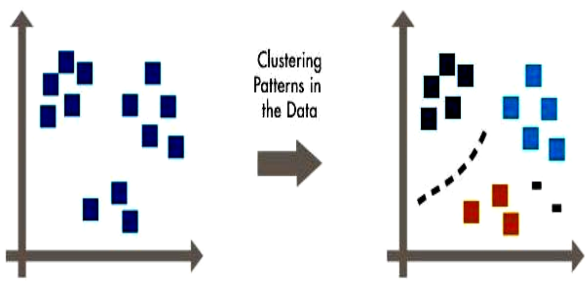

Fig. 3 Clustering finds hidden patterns in your data.

How Do You Decide Which Machine Learning Algorithm to Use?

Choosing the right algorithm can seem overwhelming—there are dozens of supervised and unsupervised machine learning algorithms, and each takes a different approach to learning. There is no best method or one size fits all. Finding the right algorithm is partly just trial and error—even highly experienced data scientists can't tell whether an algorithm will work without trying it out. But algorithm selection also depends on the size and type of data you're working with, the insights you want to get from the data, and how those insights will be used. Here are some guidelines on choosing between supervised and unsupervised machine learning:

- Choose **supervised learning** if you need to train a model to make a prediction--for example, the future value of a continuous variable, such as temperature or a stock price, or a classification—for example, identify makes of cars from webcam video footage.
- Choose **unsupervised learning** if you need to explore your data and want to train a model to find a good internal representation, such as splitting data up into clusters.

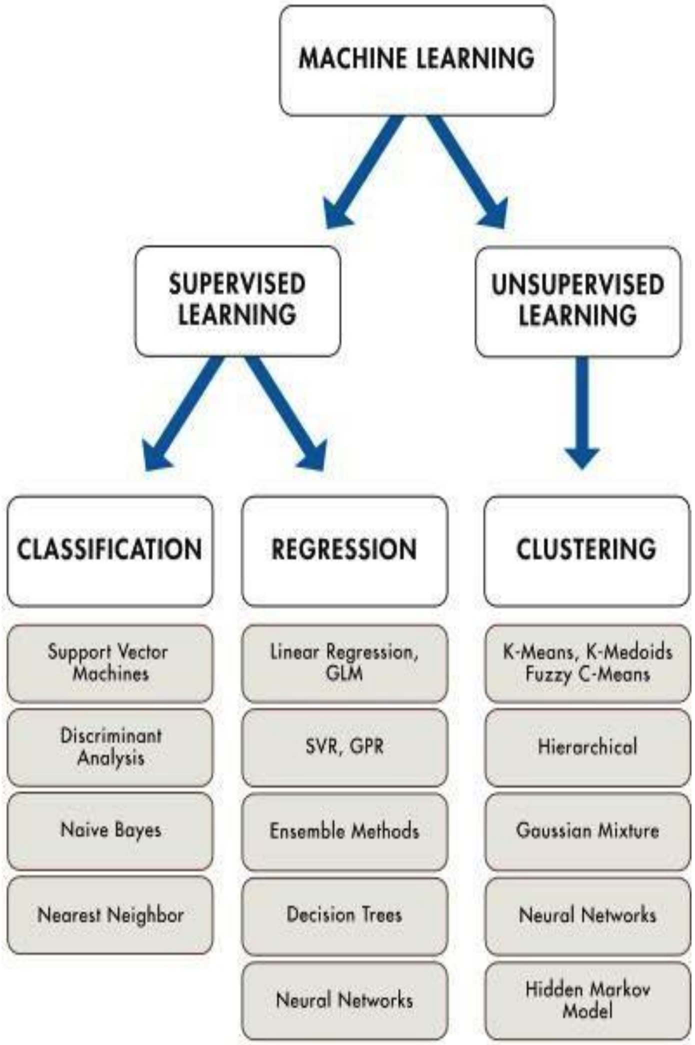

Fig. 4: Machine learning techniques.

Steps used in Machine Learning.

There are 5 basic steps used to perform a machine learning task:

1. **Collecting data**: Be it the raw data from excel, access, text files etc., this step (gathering past data) forms the foundation of the future learning. The better the variety, density and volume of relevant data, better the learning prospects for the machine becomes.
2. **Preparing the data**: Any analytical process thrives on the quality of the data used. One needs to spend time determining the quality of data and then taking steps for fixing issues such as missing data and treatment of outliers. Exploratory analysis is perhaps one method to study the nuances of the data in details thereby burgeoning the nutritional content of the data.

3. **Training a model**: This step involves choosing the appropriate algorithm and representation of data in the form of the model. The cleaned data is split into two parts – train and test (proportion depending on the prerequisites); the first part (training data) is used for developing the model. The second part (test data), is used as a reference.
4. **Evaluating the model**: To test the accuracy, the second part of the data (holdout / test data) is used. This step determines the precision in the choice of the algorithm based on the outcome. A better test to check accuracy of model is to see its performance on data which was not used at all during model build.
5. **Improving the performance**: This step might involve choosing a different model altogether or introducing more variables to augment the efficiency. That's why significant amount of time needs to be spent in data collection and preparation.

5 APPLICATIONS OF MACHINE LEARNING

1. **Financial Trading**

Many people are eager to be able to predict what the stock markets will do on any given day — for obvious reasons. But machine learning algorithms are getting closer all the time. Many prestigious trading firms use proprietary systems to predict and execute trades at high speeds and high volume. Many of these rely on probabilities, but even a trade with a relatively low probability, at a high enough volume or speed, can turn huge profits for the firms. And humans can't possibly compete with machines when it comes to consuming vast quantities of data or the speed with which they can execute a trade.

2. **Data Security**

Malware is a huge — and growing — problem. In 2014, Kaspersky Lab said it had detected 325,000 new malware files every day. But, institutional intelligence company Deep Instinct says that each piece of new malware tends to have almost the same code as previous versions — only between 2 and 10% of the files change from iteration to iteration. Their learning model has no problem with the 2–10% variations, and can predict which files are malware with great accuracy. In other situations, machine learning algorithms can look for patterns in how data in the cloud is accessed, and report anomalies that could predict security breaches.

3. **Personal Security**

If you've flown on an airplane or attended a big public event lately, you almost certainly had to wait in long security screening lines. But machine learning is proving that it can be an asset to help eliminate false alarms and spot things human screeners might miss in security screenings at airports, stadiums, concerts, and other venues. That can speed up the process significantly and ensure safer events.

4. **Healthcare**

Machine learning algorithms can process more information and spot more patterns than their human counterparts. One study used computer assisted diagnosis (CAD) when to review the early mammography scans of women who later developed breast cancer, and the computer spotted 52% of the cancers as much as a year before the women were officially diagnosed. Additionally, machine learning can be used to understand risk factors for disease in large populations. The company Medecision developed an algorithm that was able to identify eight variables to predict avoidable hospitalizations in diabetes patients

5. **Marketing Personalization**

The more you can understand about your customers, the better you can serve them, and the more you will sell. That's the foundation behind marketing personalization. Perhaps you've had the experience in which you visit an online store and look at a product but don't buy it — and then see digital ads across the web for that *exact* product for days afterward. That kind of marketing personalization is just the tip of the iceberg. Companies can personalize which emails a customer receives, which direct mailings or coupons, which offers they see, which products show up as "recommended" and so on, all designed to lead the consumer more reliably towards a sale.

6. **Fraud Detection**

Machine learning is getting better and better at spotting potential cases of fraud across many different fields. PayPal, for example, is using machine learning to fight money laundering. The company has tools that compare millions of transactions and can precisely distinguish between legitimate and fraudulent transactions between buyers and sellers.

7. **Recommendations**

You're probably familiar with this use if you use services like Amazon or Netflix. Intelligent machine learning algorithms analyze your activity and compare it to the millions of other users to determine what you might like to buy or binge watch next. These recommendations are getting smarter all the time, recognizing, for example, that you might purchase certain things as gifts (and not want the item yourself) or that there might be different family members who have different TV preferences.

8. **Online Search**

Perhaps the most famous use of machine learning, Google and its competitors are constantly improving what the search engine understands. Every time you execute a search on Google, the program watches how you respond to the results. If you click the top result and stay on that web page, we can assume you got the information you were looking for and the search was a success. If, on the other hand, you click to the second page of results, or type in a new search string without clicking any of the results, we can surmise that the search engine didn't serve up the results you wanted — and the program can learn from that mistake to deliver a better result in the future.

9. **Natural Language Processing (NLP)**

NLP is being used in all sorts of exciting applications across disciplines. Machine learning algorithms with natural language can stand in for customer service agents and more quickly route customers to the information they need. It's being used to translate obscure legalese in contracts into plain language and help attorneys sort through large volumes of information to prepare for a case.

10. **Smart Cars**

IBM recently surveyed top auto executives, and 74% expected that we would see smart cars on the road by 2025. A smart car would not only integrate into the Internet of Things, but also learn about its owner and its environment. It might adjust the internal settings — temperature, audio, seat position, etc. — automatically based on the driver, report and even fix problems itself, drive itself, and offer real time advice about traffic and road conditions.

6. CONCUSION

Machine learning is an incredible breakthrough in the field of artificial intelligence. While it does have some frightening implications when you think about it, these applications are just several of the numerous ways this technology can improve our lives.Like any discipline, machine learning has a lot of "folk wisdom" that can be hard to come by, but is crucial for success. This article summarized some of the most salient items

REFERENCES

1. Blankertz, G. Dornhege, C. Schafer, R. Krepki, J. Kohlmorgen, K.-R. Muller, V. Kunzmann, F. Losch, and G. Curio. BCI bit rates and error detection for fast-pace motor commands based on single-trial EEG analysis.IEEE Transactions on Neural Systems and Rehabilitation Engineering, 11:127–131, 2003.
2. E. Bauer and R. Kohavi. Empirical comparison ofvoting classification algorithms: Bagging, boosting and variants.MachineLearning, 36:105–142, 1999.
3. A Few Useful Things to Know about Machine Learning Pedro Domingos Department of Computer Science and Engineering University of Washington Seattle, WA .S.A.pedrod@cs.washington.edu.
4. Zien, G. Ratsch, S. Mika, B. Scholkopf, T. Lengauer, and K.-R. Muller.Engineering Support Vector Machine Kernels That Recognize Translation Initiation Sites.BioInformatics,16(9):799–807, September 2000.
5. Yom-Tov. An introduction to pattern classification. In U. von Luxburg, O. Bousquet, and G. Ratsch, editors,Advanced Lectures on Machine Learning, volume 3176 of LNAI , pages 1–23. Springer, 2004.
6. V.N. Vapnik.The nature of statistical learning theory. Springer Verlag, New York, 1995. U. von Luxburg, O. Bousquet, and G. R̈atsch, editors.Advanced Lectures on Machine Learn-ing, volume 3176 ofLNAI. Springer, 2004
7. A Brief Introduction into Machine Learning Gunnar Ratsch Friedrich Miescher Laboratory of the Max Planck Society, Spemannstraße 37, 72076 Tubingen, Germany.
8. Trends in extreme learning machines: a review, by Huang, G., Huang, G., Song, S., & You, K. (2015)
9. Machine Learning: A study of algorithms that learn from data and experience.
10. A Visual Introduction to Machine Learning (2015). Courtesy of Stephanie yee and Tony Chu