MACHINE LEARNING

Tan Kian Hua

Adjunct Professor of Management and International Business, and Cybersecurity

Abstract: Machine learning (ML) is the study of computer algorithms that can improve automatically through experience and by the use of data artificial intelligence is associated with it. Machine learning algorithms create a model based on training data to make predictions or judgments without having to be explicitly programmed to do so. Machine learning algorithms are utilized in a wide range of applications, including medicine, email filtering, speech recognition, and computer vision, where developing traditional algorithms to do the required tasks is difficult or impossible.

Introduction

Machine Learning (ML) has revolutionized the computing world by allowing computers to learn as they go with enormous datasets, avoiding many of the pitfalls and dead ends of the previous programming. When exposed to huge volumes of data, Machine Learning algorithms can self-teach and evolve. The combination of this unique technology and AI applications has the potential to be tremendously powerful. Smart robots will soon be doing all of our jobs – much faster, much more accurately, and even self-improving. Will intelligent people soon become obsolete, or will self-aware robots take our place? What are the biggest Machine Learning trends for 2017?

Machine learning and artificial intelligence applications in the financial sector have exploded in popularity in recent years. To increase efficiency and improve customer experience, these institutions have utilized their considerable capacity to deliver business solutions in front-end and back-end operations. This paper will discuss the applications of machine learning and artificial intelligence in the banking industry, as well as their applicability in many functional areas and how these institutions use computational intelligence to improve their businesses. Traditional financial institutions are catching up to computational intelligence technology with products like Chatbot, while fintech firms, which appear to have embraced A.I., are trailing behind. Its creativity played a vital role in financial intelligence many years ago. To summarize, machine learning and artificial intelligence are fast altering the banking business, and it appears that there is little we can do about it.

Machine learning is closely related to computational statistics, which focuses on making predictions with computers; nevertheless, statistical learning is not all machine learning. The discipline of machine learning benefits from the study of mathematical optimization since it provides tools, theory, and application domains. Data mining is a similar branch of research that focuses on unsupervised learning for exploratory data analysis. Data and neural networks are used in some machine learning implementations to replicate the functioning of a biological brain. Machine learning is also known as predictive analytics when it is used to solve business challenges.

The rule-based machine learning method of association rule learning is used to uncover associations between variables in huge databases. Its purpose is to use some form of "interestingness" metric to uncover strong rules in databases. To detect correlations between variables in huge databases, the association rule learning approach uses rule-based machine learning. Its goal is to find strong rules in databases by using some measure of "interestingness."

Machine Learning is fast moving away from abstractions and toward commercial problem-solving in research disciplines, thanks to AI and Deep Learning. Forbes forecasts that theoretical research in machine learning will progressively pave the way for corporate problem solutions in What Is the Future of Machine Learning. Smart (ML) algorithms may now simply use massive amounts of both static and dynamic data to learn and improve for greater performance, thanks to Big Data's resurgence in a mainstream enterprise.

Machine learning arose from the search for artificial intelligence as a scientific pursuit. Some academics were interested in making machines learn from data in the early days of AI as an academic discipline. They tried different symbolic methods as well as what was then referred to as "neural networks," which were largely perceptrons and other models that were subsequently discovered to be reinventions of generalized linear models of statistics. Probabilistic reasoning was also used, particularly in automated medical diagnosis software.

At a high level of abstraction, the field of machine learning is introduced. There is a discussion of supervised and unsupervised learning, regression, and classification. The balance of bias, variance, and model complexity is an important guiding principle in learning.

Machine learning may create models such as neural networks (feed-forward and recurrent), support vector machines, random forests, self-organizing maps, and Bayesian networks. Following that, the fundamental concepts of partitioning a dataset into training, testing, and validation sets, as well as cross-validation, are discussed. The role of the domain expert in keeping the project on track is explored next, followed by a discussion of how to assess the model's usefulness. The chapter concludes with some pointers on how to end a machine learning project.

However, as the emphasis on logical, knowledge-based approaches has grown, a schism has emerged between AI and machine learning. Theoretical and practical data gathering and representation issues plagued probabilistic systems. Expert systems had dominated AI by 1980, and statistics had fallen out of favor. Work on symbolic/knowledge-based learning continued inside AI, leading to inductive logic programming, but the more statistical line of research in pattern recognition and information retrieval was now outside the discipline of AI proper. Around the same time, AI and computer science had abandoned neural network research. Researchers from other disciplines, such as Hopfield, Rumelhart, and Hinton, continued this path as "connectionism" outside of the AI/CS field. Their biggest breakthrough came in the mid-1980s when they reinvented backpropagation.

- **2017 Trends in Machine Learning Application Development**

The integrated AI and advanced ML approach, which began four years ago and has remained unaffected, will dominate Artificial Intelligence application development in 2017, according to Gartner's Top 10 Technology Trends for 2017. As a result of this fatal combination, more systems will "understand, learn, forecast, adapt, and potentially act autonomously." Low-cost hardware, memory, and storage technologies, as well as increasing processing power, smarter algorithms, and massive data streams, will all promote the development of ML-powered AI applications. Ml-powered AI applications will continue to emerge in fields like preventive healthcare, banking, finance, and media. For businesses, this implies fewer human checkpoints and more automated activities. 2017 Forrester's predictions Sensors and smart apps are taking over every area of life. Artificial Intelligence and Machine Learning Clouds will increasingly rely on IoT data as part of our daily lives.

Researchers from other disciplines, such as Hopfield, Rumelhart, and Hinton, continued this path as "connectionism" outside of the AI/CS field. Their biggest breakthrough came in the mid-1980s when they reinvented backpropagation.

Machine learning (ML), which was established as a separate field in the 1990s, began to develop. The goal of the field shifted from artificial intelligence to practical problems that could be solved. It turned its focus away from the symbolic approaches it had received from AI and toward statistics, fuzzy logic, and probability theory methodologies and models.

The distinction between machine learning and artificial intelligence is frequently misinterpreted. ML learns and predicts based on passive observations, whereas AI refers to an agent that interacts with the environment to learn and take actions that increase its chances of attaining its objectives.

Many sites still claim that machine learning is a subfield of AI in 2020. Others argue that not all machine learning is AI, and that only an "intelligent subset" of machine learning should be deemed AI.

- **Machine Learning Democratization in the Cloud**

Cloud technology, open standards, and the algorithm economy will continue to democratize AI and ML. The expanding trend of deploying prebuilt machine learning algorithms to provide self-service business intelligence and analytics is a step forward in the democratization of machine learning. The author of Google Says Machine Learning is the Future advocates for the democratization of machine learning through idea-sharing. Google's Tensor Flow, for example, has promoted the necessity for open standards in Machine Learning. According to this article, practically anyone with a laptop and Internet access may now claim to be a Machine Learning expert if they have the correct mindset.

- **Optimization**

Many learning issues are phrased as minimization of some loss function on a training set of instances, which relates machine learning to optimization. The difference between the model's predictions and the actual problem instances is expressed by loss functions (for example, in classification, one wants to assign a label to instances, and models are trained to correctly predict the pre-assigned labels of a set of examples).

Cloud-based IT services were already a solid step toward making complex Data Science a popular activity, and now, thanks to Cloud and packaged algorithms, mid-sized and smaller enterprises will have access to Self-Service BI and Analytics, which was before simply a pipe dream. Furthermore, in data-centric business systems, mainstream corporate users will eventually take an active role. Trends in Machine Learning According to Future AI, in 2017, more businesses will take advantage of the Machine Learning Cloud and advocate for democratized data technology.

- **Generalization**

The goal of generalization distinguishes optimization from machine learning: whereas optimization techniques can minimize a loss on a training set, machine learning is concerned with decreasing loss on unseen samples. Characterizing the generalization of various learning algorithms, particularly deep learning algorithms, is a hot area in current research.

In 2017, the Platform Wars will reach their pinnacle.

In 2017, the platform war between IBM, Microsoft, Google, and Facebook to be the leader in machine learning will reach a climax. According to Where Machine Learning Is Headed, 2017 will see a huge increase in smart apps, digital assistants, and mainstream adoption of Artificial Intelligence. Although many ML-enabled AI systems have shown to be successful, self-driving cars may be doomed to fail.

Statistics

In terms of methodologies, machine learning and statistics are similar, but their main goals are different: statistics derive population inferences from a sample, whereas machine learning looks for generalizable predictive patterns. Machine learning ideas, from methodological principles to theoretical tools, have a lengthy history in statistics, according to Michael I. Jordan. He also proposed the phrase data science as a catch-all word for the entire field.

Leo Breiman distinguished two statistical modeling paradigms: data model and algorithmic model, the latter referring to machine learning methods such as Random forest.

Some statisticians have integrated machine learning approaches, resulting in a hybrid field known as statistical learning.

Data Science and Machine Learning demand-supply gaps will widen.

The business world is quickly approaching 2018, when, according to McKinsey, the first hole in data technology knowledge will be felt in the United States, followed by the rest of the world. Until university programs and industry workshops begin to generate a skilled workforce, the demand-supply gap in Data Science and Machine Learning skills will continue to widen. As a result of the widening demand-supply gap, more businesses and academic institutions will partner to train future Data Scientists and Machine Learning professionals. This type of instruction will compete with traditional Data Science classrooms by emphasizing practical skills over theoretical knowledge. The KD Nuggets will continue to test themselves, and the inquiring mind by releasing articles like 10 Algorithms that Machine Learning Engineers Should Know. In 2017, KDNugget and Kaggle will continue to contribute to alternative training for future Data Scientists and Machine Learning specialists through actual skill development.

Depending on the type of the "signal" or "feedback" available to the learning system, machine learning systems are generally categorized into three major categories:

• Supervised learning: A "teacher" presents the computer with sample inputs and desired outputs, to learn a general rule that maps inputs to outputs.

• Unsupervised learning: The learning algorithm is given no labels and is left to find structure in its data on its own. Unsupervised learning can be a goal in and of itself (finding hidden patterns in data) or a means to an end (finding hidden patterns in data) (feature learning).

• Reinforcement learning: A computer program interacts with a dynamic environment to complete a task (such as driving a vehicle or playing a game against an opponent). The software is given input in the form of incentives as it navigates its issue space, which it strives to maximize.

Learning under supervision

Supervised learning algorithms create a mathematical model of a set of data that includes both inputs and outputs. The data is referred to as training data, and it is made up of a set of training examples. Each training example has one or more inputs and a supervisory signal as the desired output. Each training sample is represented by an array or vector, sometimes referred to as a feature vector, in the mathematical model, and the training data is represented by a matrix. Supervised learning techniques develop a function that may be used to predict the output associated with fresh inputs by iteratively optimizing an objective function. The algorithm will be able to accurately estimate the output for inputs that were not part of the training data if it uses an optimum function. An algorithm that learns to complete a task increases the accuracy of its outputs or predictions over time.

Active learning, classification, and regression are examples of supervised learning algorithms. When the outputs are limited to a small set of values, classification techniques are employed, and regression algorithms are used when the outputs can have any numerical value within a range. An incoming email, for example, would be the input to a classification algorithm that filters emails, and the output would be the name of the folder to file the email.

The purpose of similarity learning, which is closely connected to regression and classification, is to learn from examples using a similarity function that quantifies how similar or related two items are. Ranking, recommendation systems, visual identification tracking, face verification, and speaker verification are some of the uses.

A supervised learning model called a support-vector machine separates data into areas separated by a linear border. The black circles are separated from the white circles by a linear barrier.

- **Unsupervised Learning**

Unsupervised learning methods take a collection of data with only inputs and detect structure in it, such as data point grouping or clustering. As a result, the algorithms learn from unlabeled, unclassified, and uncategorized test data. Unsupervised learning algorithms discover commonalities in the data and react depending on the existence or lack of such commonalities in each new piece of data, rather than responding to feedback. The area of density estimation in statistics, such as calculating the probability density function, is a key application of unsupervised learning. Unsupervised learning, on the other hand, comprises various fields that require summarizing and explaining data aspects.

Cluster analysis is the division of a set of data into subsets (called clusters) based on one or more predetermined criteria, with observations obtained from different clusters being distinct. Different clustering approaches make different assumptions about the structure of the data, which is commonly characterized by some similarity metric and evaluated, for example, by internal compactness, the similarity between cluster members, separation, or the difference between clusters. Other methods rely on density estimation and graph connectedness.

- **Semi-supervise learning**

Unsupervised learning (without any labeled training data) and supervised learning are the two types of learning (with completely labeled training data). Although some of the training examples lack training labels, several machine-learning researchers have discovered that unlabeled data, when combined with a modest amount of labeled data, can enhance learning accuracy significantly.

The training labels in weakly supervised learning are noisy, limited, or imprecise, yet they are generally cheaper to obtain, resulting in larger effective training sets.

- **Learning via reinforcement**

Reinforcement learning is a branch of machine learning that studies how software agents should behave in a given environment to maximize some metric of cumulative reward. Because of its broad scope, game theory, control theory, operations research, information theory, simulation-based optimization, multi-agent systems, swarm intelligence, statistics, and genetic algorithms are all investigated in the topic. The environment is generally represented as a Markov decision process in machine learning (MDP). Dynamic programming approaches are used in many reinforcement learning systems. When exact mathematical models of the MDP are not achievable, reinforcement learning procedures are applied. Reinforcement learning algorithms are employed in driverless vehicles and in teaching humans to play games.

- **Humans and Machines will live in harmony.**

Since 2012, data technologies have seen parabolic growth and extensive dissemination in the global business world. Humans will finally recognize that it is time to stop fearing machines and start cooperating with them. According to the InfoWorld article Application Development, Docker, and Machine Learning Are Top Tech Trends for 2017, humans and machines will collaborate rather than compete. In this context, readers should read the DATAVERSITY® article The Future of Machine Learning: Trends, Observations, and Forecasts, which reminds readers that as businesses become more reliant on pre-built ML algorithms for Advanced Analytics, the need for Data Scientists and large IT departments may decrease.

- **Consider the following ideas:**

If the threat of intelligent machines replacing Data Scientists is as real as it appears, 2017 is likely to be the year when the global Data Science community takes a fresh look at the capabilities of so-called "smart machines." The repeated failures of self-driving cars have proven one thing: even learning robots cannot match the natural reasoning abilities granted by nature to humans. If self-driving or self-guided machines are to be useful to humans, current Artificial Intelligence and Machine Learning research should focus on recognizing the limits of machine power, assigning tasks that are appropriate for machines, and including more human intervention at necessary checkpoints to avoid failure. Machines can handle repetitive, routine activities well, but any situation that is out of the norm will require human intervention.

The Algorithm Economy will be in the spotlight. Businesses will use prepackaged algorithms for all data-centric tasks such as BI, Predictive Analytics, and CRM in the next year or two. According to Forbes, the algorithm economy will usher in a marketplace where all data businesses will fight for a piece of the pie. Self-Service BI will become more popular in 2017, as will algorithmic business solutions and machine learning on the cloud. 2017 could see two separate sorts of algorithm economies in terms of algorithm-driven commercial decision making. Average firms, on the one hand, will rely on prepackaged algorithmic models for their operational and customer-facing operations. Proprietary machine learning algorithms, on the other hand, will become a market differentiation among huge, competing businesses.

REFERENCES

[1]. *Mitchell, Tom* (1997). *Machine Learning*. McGraw Hill, New York, ISBN 0-07-042807-7, OCLC 36417892

[2]. Arthur Samuel, who invented the term "machine learning" in 1959, is typically credited with the definition "without being expressly programmed," but the phrase is not found verbatim in this publication, and could be a paraphrasing. Confer "How can computers learn to solve problems without being explicitly programmed?" asks Arthur Samuel (1959) in Koza, John R.; Bennett, Forrest H.; Andre, David; Keane, Martin A. (1996). Genetic Programming is used to automate the topology and sizing of analog electrical circuits. Design '96: Artificial Intelligence. 151–170 in Springer, Dordrecht. doi:10.1007/978-94-009-0279-4_9.

[3]. "Voronoi-Based Multi-Robot Autonomous Exploration in Unknown Environments through Deep Reinforcement Learning," IEEE Transactions on Vehicular Technology, 2020.

[4]. Go to:a, b, c, d Pattern Recognition and Machine Learning, Springer, ISBN 978-0-387-31073-2, Bishop, C. M.

[5]. Pattern recognition and machine learning "They can be seen as two sides of the same coin."

[6]. "Tutorial: Polynomial Regression in Excel," facultystaff.richmond.edu, Christopher Stevenson. The date was January 22, 2017.

[7]. Ghelani, D., & Hua, T. K. (2022). Conceptual Framework of Web 3.0 and Impact on Marketing, Artificial Intelligence, and Blockchain. *International Journal of Information and Communication Sciences*, *7*(1), 10.

[8]. *Ghelani, D., & Hua, T. K. A Perspective Review on Online Food Shop Management System and Impacts on Business.*

[9]. Ghelani, D., Hua, T. K., & Koduru, S. K. R. (2022). A Model-Driven Approach for Online Banking Application Using AngularJS Framework. *American Journal of Information Science and Technology*, *6*(3), 52-63.

[10]. Dr. John Ughulu. The role of Artificial intelligence (AI) in Starting, automating and scaling businesses for Entrepreneurs.. *ScienceOpen Preprints.* DOI: 10.14293/S2199-1006.1.SOR-.PP5ZKWJ.v1

[11]. *Ughulu, J. Entrepreneurship as a Major Driver of Wealth Creation.*

[12]. Oak, R., Du, M., Yan, D., Takawale, H., & Amit, I. (2019, November). Malware detection on highly imbalanced data through sequence modeling. In *Proceedings of the 12th ACM Workshop on artificial intelligence and security* (pp. 37-48).

[13]. Hua, T. K., & Biruk, V. (2021). *Cybersecurity as a Fishing Game: Developing Cybersecurity in the Form of Fishing Game and What Top Management Should Understand*. Partridge Publishing Singapore.

(8) Sneed, J. (2017) Predicting ESP Lifespan using Machine Learning, Unconventional Resources Technology Conference, Austin, Texas, USA. Unconventional Resources Technology Conference, SPE/AAPG/SEG, Austin, July 24-26.

https://doi.org/10.15530/URTEC-2017-2669988

(9) Mehryar Mohri; Afshin Rostamizadeh; Ameet Talwalkar (2012). Machine Learning Foundations. MIT Press, Cambridge, Massachusetts, ISBN 9780262018258.

(10) Go to: a b Alpaydin, Ethem (2010). Machine Learning: An Overview The MIT Press, London, ISBN 978-0-262-01243-0. The 4th of February, 2017.

(11) van Otterlo, M.; Wiering, M. (2012). Reinforcement learning and markov decision processes. Reinforcement Learning. Adaptation, Learning, and Optimization. Vol. 12. pp. 3–42. *doi*:*10.1007/978-3-642-27645-3_1*. *ISBN 978-3-642-27644-6.*

(12) science.sciencemag.org/content/290/5500/2323

(13)Towardsdatascience.com/all-machine-learning-models-explained-in-6-minutes-9fe30ff6776a