Machine Learning and Data Science for Process Industries

Yejur Dube[1], Vinamra Soni[2], J. Satheesh Kumar[3]

[1,2]Student, Department of Electronics and Instrumentation Engineering, Dayananda Sagar College of Engineering, Bangalore, India.
[3]Associate Professor, Department of Electronics and Instrumentation Engineering, Dayananda Sagar College of Engineering, Bangalore, India.

Email: [1]yejur59@gmail.com, [2]vinamrasoni01@gmail.com, [3]jsatheeshngl@gmail.com

Abstract

In the last few years, Machine Learning has drawn a lot of attention. One of the reasons it is getting popular today is that, it provides a united domain for dealing with problems with intelligent-decision making skills. From small to big firms, Machine Learning is now used to extract the best data models which ultimately benefits the company. This paper lays the foundation of a thorough investigation of the machine learning world and study what exactly is the reason behind the huge surge of this new phenomenon in the industry today. Moreover, it discusses the use of Data Science, improvements in several leading industries, some case studies and the impact of its growth in usage in the world today. Based on many graphs and business models, why and how some leading companies are incorporating Machine Learning and Data Science in their business plans and whether it has benefitted them in the long run, have been analysed.

1. Introduction

The re-kindled fascination in Machine Learning (ML), observed over the last few decades, has surfaced in various areas of study, including engineering. The software principles, analytical techniques and data science that creates an ML model which has, in the late, re-constructed the paradigm of the industry equation, have been utilized [1]. Data Science (DS) means studying data and finding the meaning in it. It involves applying analytical techniques, scientific principles and ML which gives the required insights into a given data. The studied data helps in increasing efficiency, finding new business opportunities, marketing and selling better [2]. The only difference between ML and DS is

that while DS extracts meaningful data, ML understands and builds methods that utilise data to improve overall performance. Below graph depicts the after results of various industries with implementation of Data Science & Machine Learning.
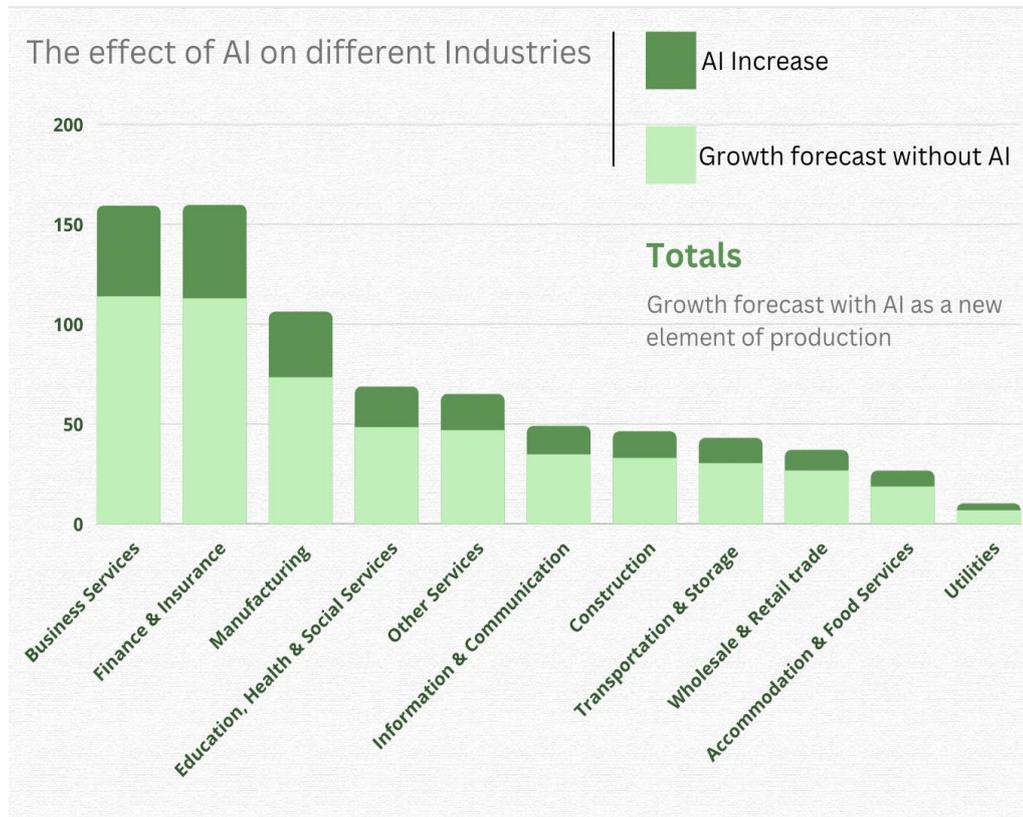

Figure 1. Growth forecast of various industries after implementation of AI-ML [22]

2. Data Science without Machine Learning

Machine Learning is the heart of Artificial Intelligence (AI). Without it, computers were used for explicit and unnecessarily vast programming. But today, these computers learn, grow and innovate by themselves. The newly found ability to quickly apply mathematical calculations to enormous data is getting popular today. It has hooked its way everywhere - from self-drive cars to everyday use search engines. The answer to why ML and DS are needed is simple: 'High-value predictions that can guide better decisions and smart actions in real-time without human intervention' [3]. Data Science is impossible to use when there are significant and unique sets of data and it is for this reason, big data gets criticised for being overhyped. It gets progressively more difficult making predictive models that work. Machine Learning, which provides smart alternatives in analysing vast data, with accurate and precise results has proved to be a big step forward in software as well as technology industries. Figure 2 shows the relationship between ML and DS.
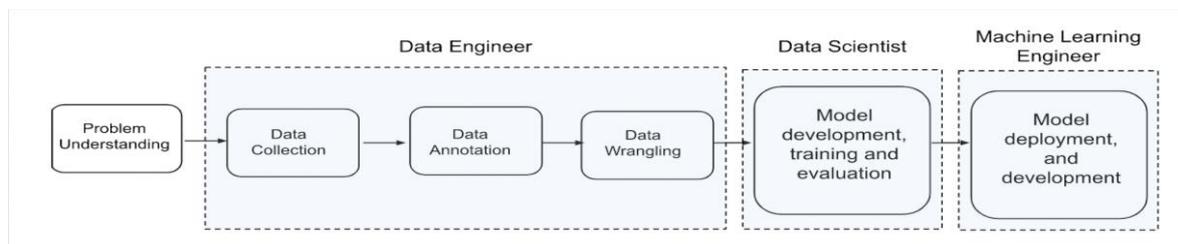
Figure 2. Relation between a Data Scientist and a Machine Learning Engineer [4]

3. Requirement of Data Science and ML in Industries

The reasons why Data Science and ML are needed are:

- Assurance of quality and performance.

- Demand and throughout forecasting.

- Enhancement of supply chain.

- Sustaining and achieving energy efficiency. [5]

4. Industrial Applications that are implemented with DS and ML

Data Science and Machine Learning have impacted many aspects of the world, and here some of the many ways the world has benefitted from them are discussed:

- Fraud Detection: In the early days, banks were under heavy stress regarding the huge pile of paperwork and ocean of data which caused the banks to suffer huge losses. So, to rescue themselves from such atrocity, data scientists were hired who learned to divide and conquer data via customer profiling, past expenditures and other variables to analyse the data in their hands in the best way possible in order to understand and overcome the daily frauds banks used to suffer.

- Drug Development: Forming a new drug has always been a tedious process for medical industries as it takes around a decade to launch a new medicine and not to mention the various kinds of disciplines which are involved. Data Science has shortened this process via data management; from initial drug testing to its future success rate probability, everything is now controlled. Using mathematical modelling and simulations, medical industries do not need "lab experiments" to predict how a drug will act inside a human body. [16]

- Quality Control: Machine learning is being used in industries for the inspection of the quality of products and their control. ML-based computers are being used to identify a

good product from a faulty one. These computers have built-in algorithms for detecting common defects, and applying that algorithm, gets rid of 99 percent of faulty products.

- Predictive Maintenance: ML enables maintenance by predicting future errors in the machine and scheduling timely check-ups which save manufacturers hours of time. Such equipment failure predictions come out 92 percent true, thus increasing efficiency in various businesses as it promises reliability and subpar quality. [17]

5. Outcomes of Data Science and ML in the Industry Sector

Modern-day tools have changed industries drastically. The following work delves into some companies' pasts and observes how these tools have benefitted them, particularly from Machine Learning and Data Science.

5.1 The Healthcare Industry

Not long ago, without ML and Data Science, the healthcare industries did not have the most optimum and sophisticated diagnostic and treatment options. Data Science and ML have revolutionised the industry by helping doctors make better diagnoses of their patients, thanks to a much better database. The FDA has approved a few new systems which indicate the emergence of new technology [6]. The AI in healthcare funding hit a historic high in the last few years - starting from $50M in 2013, reaching $400M in just 2 years and finally hitting $600M in 2018. The below figure depicts the investments done by the healthcare industries in AI-ML.
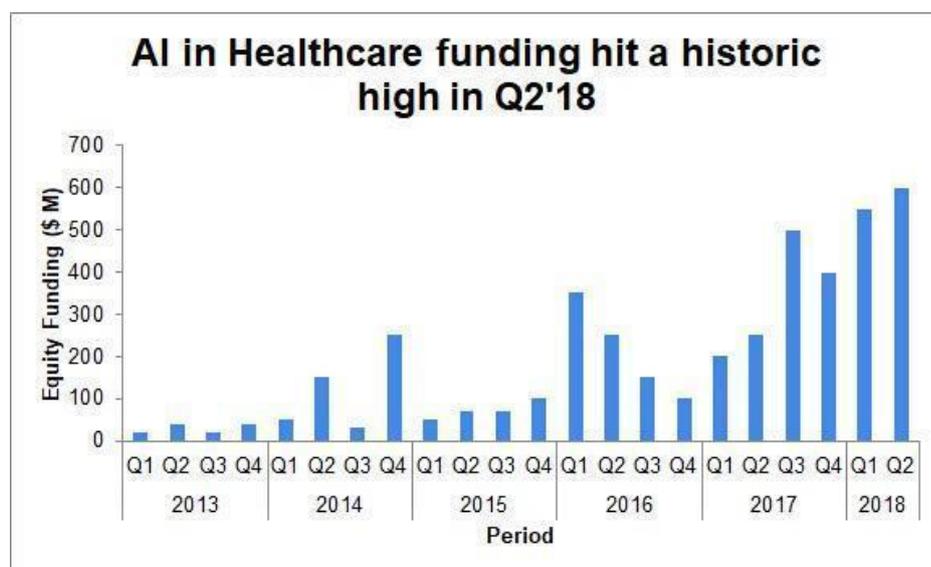

Figure 3. Funding in AI by healthcare industries [7]

5.2 The Manufacturing Industry

The manufacturing industry is one of the first industries ever to make use of automation and ML starting as early as the 1960s. Technology and the industry have grown since then. Robots are getting used to performing daily mundane tasks better than their human counterparts who require constant supervision. For every 1000 workers, the work replaced by a robot helps the company save from a quarter to a half percent of revenue. With the use of AI and ML, robots are getting smarter and more cost-effective. Over the last 24 months, hardware upgrades were increased by 8.2%, advanced process control was increased by 1.9% and finally, AI assets were increased by 6.5% [8].

5.3 The Finance Industry

It's hard to imagine the impact of AI-ML in the finance industry, especially when it comes to customer satisfaction. But, it is promised by these industries that these technologies will have a good grasp of the industry's roots. AI-based models can take on a much more vast database which helps in various ways including security and fraud detection, algorithm trading, loan and insurance underwriting and much more. It is further stated by these industries that analytics is the key factor in decision-making followed by decisions made based on data making up to 49%, key strategic initiatives at 16%, customer relations at 10%, risk-taking abilities at 9%, financial performance at another 9%, response towards trends at 5% and creation of new products comprising 1% [9].

6. Case Studies

Forefront companies are investing in AI as they have realised the importance of their business aspect and the fact that one day it will replace people due to technological surges and digital disruption. Below are some companies that are way ahead in the AI game.

6.1 WALMART

A 60s company has been continuously on its toes to increase customer satisfaction, so it is not a surprise that they are consistent with innovation and new technology. As of today, Walmart is the second-largest online retailer with sales up to 37% just last year. Walmart is already planning to install self-scanning robots in 50 additional stores, not to forget intelligent cameras, interactive displays and a massive data centre. Not only would it ease the workload of human personnel but then people can work more on intricate details [10]. As per sales

growth in the span of 16 years (2006-2022), it started with a 3% rise in 2006, suffering a minor pitfall during 2011-2014 (dropping down to -1.5%) but eventually, the sales skyrocketed to a whopping 8.7% in 2021. The below figure shows the sales growth of Walmart from 2006 to 2022.
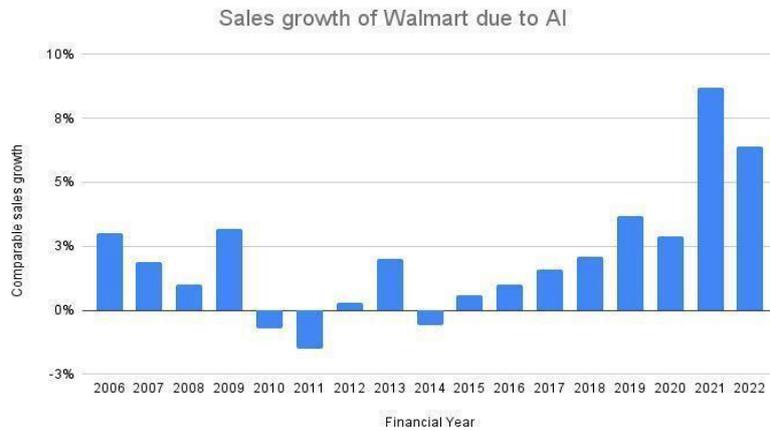

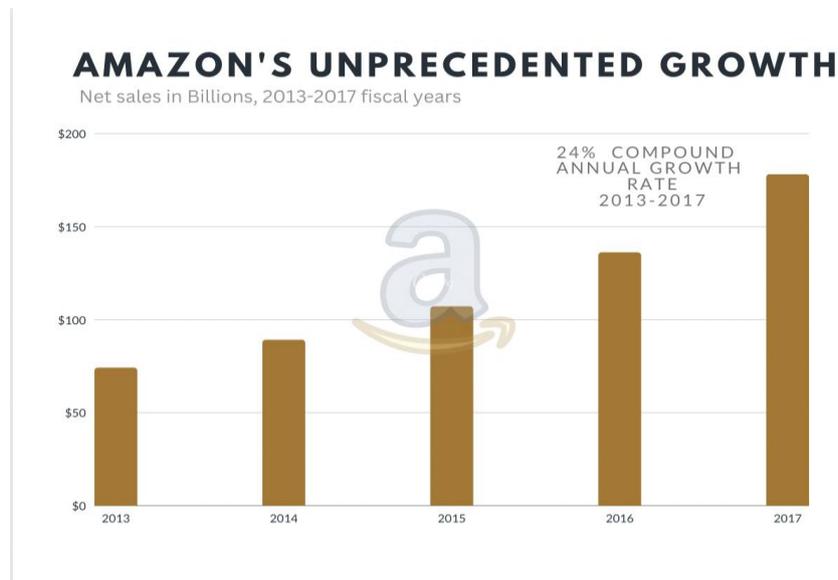**Figure 4.** The total sales growth of Walmart from 2006 to 2022 [11]

6.2 AMAZON

Figure 5. Amazon's incredible long-term growth [12]

Since 2009, Amazon has jumped to the second position in the fortune 500 list, just behind Walmart. In Amazon, AI-ML is not confined to just the business segments, it is everywhere. While there's no denying the power of AI in Amazon, the delivery component is entirely dependent on warehouse operations. Amazon's AI-based model makes recommendations to its customers and that alone makes up 35% of Amazon's total sales. This

tech giant is going to rely heavily on AI to study and understand better what its customers search for and if they are getting the best results. As per sales growth, the company's baby years started with earning its first $50b in 13 years which ensured that Amazon has a bright future which was later proved by the $469.8b it made in net sales in 2021 alone. Figure 5 below shows the growth of Amazon's long-term growth.

6.3 APPLE

Apple's drop in sales last year (total revenue: $55 billion) caused it to suffer certain setbacks but they were not more than a mere tint in the company's overall impact in the world today. When it comes to AI, Apple uses it extensively and everywhere: from the audio-to-text translator and prediction engine to its cameras. Apple's AI utilisation is so strong that it can even differentiate between an accidental draw with the Apple pencil and a deliberate input, and every Apple user cannot deny the impressive feature in their iPhone – Siri. Apple is only expected to go up and thrive in the world. Apple has one of the most impressive revenue profile in the business world - up to the time the iPad was launched in 2010. Apple had made $50b already but that's not all, what's even more impressive is that by the time Apple watch was released in 2015, the revenue of Apple was just shy of a $250b margin and the exponential rise in its growth eventually passed the $300b mark by 2021. Figure 6 shows the growth of Apple in the 21st century.
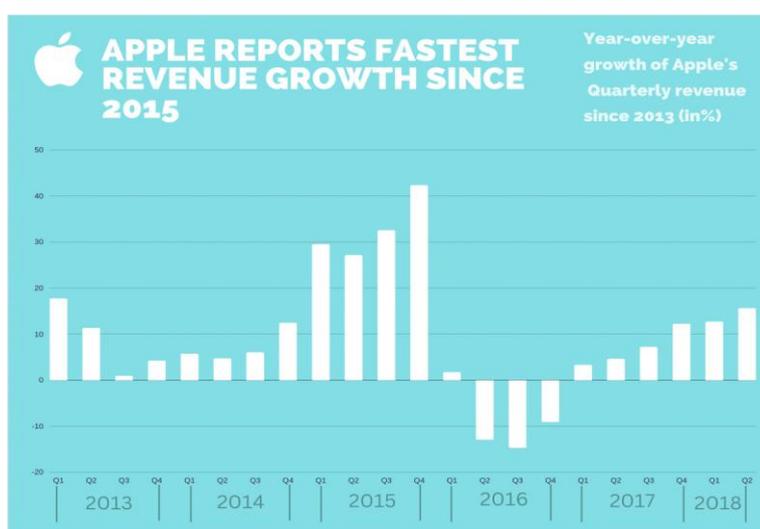

Figure 6. Apple's incredible 21st century growth [13]

7. The Impact of AI and ML on the Manufacturing Industry

The manufacturing industry has transformed in the last few years. The industry has gone through many technological revolutions in the past, and as of now, this industry is

adapting new tech models to improve traditional manufacturing processes. The Big Data Analytics services in the Industry sector are expected to grow rapidly up to USD 4.55 billion by 2025, with a 30.9% CAGR between 2020 – 2025. As per Forrester, data-driven associations have shown 30% growth annually along with better profits and improved customer satisfaction. Figure 7 depicts the usage and importance of ML and Data Science grown over the past years [14].
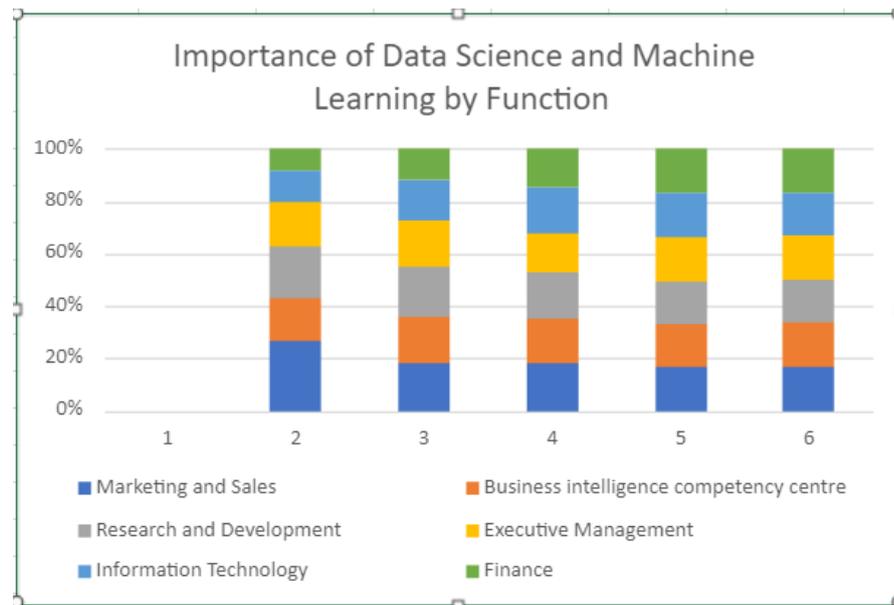

Figure 7. The impact of Data Science and Machine Learning on the Manufacturing Industry

8. Primary Requirement of Making a Neural Network-based Process Industry

AI is being pushed in every industry in the world to increase effectiveness and efficiency, and understanding every subpart of AI will become crucial in the coming future. Here, the requirements of making a Neural Network which is a Machine Learning implementation is studied. So, the first step is Data Sourcing which basically means feeding normal data to the Neural network which, in turn, involves finding the source of data and converting it to be used globally. The second step is Data Labelling, which is a time-consuming process because in order to categorize inputs, the initial data is required to be categorized beforehand. Next is Data Versioning, which is data notation for scientists to decide which data has the best results.

Now, a cycle of steps which are the most crucial part of building a neural network: first, Model Architecture involves the developer deciding the purpose of the network and the layers which will go into creating the model; secondly, Model Training which, as the name suggests, is training of the model, where various inputs are fed into the model and the output

is observed, then each of the output is compared to the labels designated during Data labelling stage, and if there's a difference between output and label, it shall be reduced. The final stage in the cycle is Model Evaluation which involves feeding new data to the model which was not used in the training stage to correctly predict if the model can truly identify patterns within the training data. The mountain peak of creating a Neural network has been covered, and what is left to do is the Model Versioning to ensure product quality assurance in the future and Model Deployment which differs based on the use case [18]. Table 1 represents the attainment of neural networks in industries.

Table 1. Attainment of the present Neural Network Models on Industries

Sl. No.	Domain Area	Summary Points
1.	Psychiatric patient length of stay	It predicts the duration of stay of a psychiatric patient based on data recorded at the time of admission. The network scored 74% in the accuracy of prediction of data under extended tolerance [19].
2.	Stock market forecasting	Network models are being used to predict the performance of a particular stock in the market. One criticism was levelled at such networks that although neural nets generate good models, it is still hard to obtain a good structure of the resulting model.
3.	Energy	New energy sources can be located. Electric companies can predict demand load.
4.	Manufacturing	Neural Nets can be used in robots that prevent everyday hazards. Neural models would be highly efficient in logistics [20].

9. Discussion Regarding Common Research Problems Observed from Previous Works

In order for Data Science and Machine Learning to grow, certain propositions should be made regarding the problems industries have encountered while embracing these technologies and these shortcomings are not limited to one or two but a horizon of fields. Some of these problems are stated below:

- Deep Learning Algorithms: The world has appreciated Learning algorithms but an understanding of why deep learning work so well haven't been grabbed still. The

sophisticated Mathematics that goes into it and the model it creates have not been understood. The fact that deep learning will create accurate models irrespective of input is unaware.

- Casual Reasoning: Although Machine Learning has changed the whole game for many in major industries, a question regarding its effect has surfaced which puts forward an argument that ML-based models are answerable to casual questioning. Such questions interfere in the face of big data and economists are trying new methods in order to make casual questioning flexible and efficient.

- Privacy: Many times, in order to create an efficient model which is big on data, multiple parties pool together their respective data but often comes the question of privacy and its preservation. One of the examples is hospitals; if multiple hospitals share the data of one patient, then the best decision can be taken for the patient, but due to legal reasons, hospitals are not allowed to share data due to the fear of privacy breaches. Industries and the Government are devising methods using cryptography and statistics to achieve unbreakable data transfer.

- Ethics: Data Science generates new ethical questions as to how data is generated and how AI interprets it. The ethical principle stated in the Belmont report and the Menlo report suggests that people should always be aware of, when they are chatting with a chatbot. These issues raise new issues for the Data Science community.

Hence, Data Science and Machine Learning should be viewed as separate fields, and time and money should be invested into their development. The future of these technologies is uncertain and that depends primarily on the younger generation; but, as far as these fields are concerned, they have strong promises and it will be beneficial for the world to embrace them. [21]

10. Conclusion and Future Scope

The principles of Machine Learning have proven to be successful and highly efficient. Its impact is global and revolutionary and it's still growing. After observing numerous graphs and data, it is understood why some Fortune 500 companies have thrived and are ahead of others in their field, which is only because such firms placed a strong emphasis on their business model for the constant improvement of their digital requirements. Machine Learning has brought a new paradigm of solving problems which is time efficient, cost-friendly and

reduces man labour by a significant amount. It is growing daily while improving its already impressive structure [15]. Although it is a skill which requires a lot of practice, particularly Mathematics, no one can deny the fact that it is a necessity if someone aims to thrive in today's digital surge. Industries are looking everywhere to find room to improve and it is proven by the reach of their existing technologies. From Google's self-driving cars to intelligent AI robots in Japan's restaurants, shortly the newly hired members in the tech industry will appreciate honing this skill.

References

[1] Michael Frank, Dimitris Dikakis, Vassilis Charissis, "Machine Learning Methods for computational Science and Engineering, MPDIi, journal, computation, 3rd March 2020.

[2] Max Mowbray, Mattia Valerio, Carlos Perez-Galvan, Dongda Zhang, Antonio Del Rio Chanona, Fransiso J. Navarro Brull, "Industrial Data Science- a review of Machine Learning applications for chemical and process industries", Royal Society of Chemistry, Journal, 2022.

[3] Raffaele Cioffi, Marta Travaglioni, Giuseppina Piscitelli, Antonella Petrillo, Fabio De Felice, "Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and directions", MPDI, Sustainability, journal, Jan. 2020.

[4] Patrick Brus, "Why you Should Learn Machine Learning and not Data Science", blog, Sept.2022.

[5] Xiuquan Li, Hongling Juan, "Artificial Intelligence Technology and Engineering Applications", ACES Journal, May 2017.

[6] Kayla Matthews, "4 Industries Being Transformed by Machine Learning and Robotics", Productivity Bytes in AI, Automation, Industry, Machine Learning, Robotic, Blog, Aug. 2017.

[7] Lucy Manole, "How can AI and Machine Learning impact Healthcare Industry", blog, Oct.2019.

[8] Louis Columbus, "10 Ways Machine Learning is Revolutionizing Manufacturing in 2019", Forbes, Blog, Aug.2019.

[9] "Data Science in Finance: 6 Use Cases That Changed The Industry", Blog, Apr. 2021.

[10] Kayla Matthews, "4 Industries Being Transformed by Machine Learning and Robotics", Productivity Bytes in AI, Automation, Industry, Machine Learning, Robotics, Blog, Aug. 2017.

[11] T. Ozbun, "United States: Sales Growth of Walmart 2006-2022", Statistica, Mar. 2022.

[12] "Amazon Strategy Teardown:Amazon's Barreling Into Physical Retail,Financial Services,Health-care and AI- Led Computing", Report, CB Insights, Mar. 2018.

[13] Felix Richter, "Apple Reports Fastest Revenue Growth Since 2015", Journal, Statistica, May 2018.

[14] Louis Columbus, "State of AI and Machine Learning", Forbes, Blog, Sept. 2019.

[15] Maryam M Najafabadi, Flavio Villanustre, Taghi M Khoshgoftaar, Naeem Seiya, Randall Wald, Edin Muharemagic, "Deep Learning Applications and Challenges in Big Data Analytics", Journal of Big Data, 2015.

[16] Upasana, "Top 10 Data Science Applications", Blog, Nov. 2022.

[17] "6 Benefits of Machine Learning in Manufacturing", Blog, IT Convergence, 2022.

[18] "Developing Neural Networks to Fulfil Business Requirements", Smart Bridge, Blog, Nov. 2019.

[19] Kevin Gurney, "An Introduction of Neural Networks", UCL Press, Taylor and Francis E-Library, 2004.

[20] Simon Jones, "What are The Industry Applications of Neural Networks", Blog, Sept. 2015.

[21] Jeannette M. Wing, "10 Research Challenges Areas in Data Science", Journal, HDSR, Sept. 2020.

[22] "The Future of Artificial Intelligence in Manufacturing Industries", Blog, Plant Automation Technology, 2022.