Harnessing the Power of Machine Learning and AI to Decrease Health Expenditures Worldwide – A Review

Michele Malandruccolo

Alten, Bern, Switzerland

Correspondence to:
Michele Malandruccolo,
malandruccolom@gmail.com

Additional material is published online only. To view please visit the journal online.

Cite this as: Malandruccolo M. Harnessing the Power of Machine Learning and AI to Decrease Health Expenditures Worldwide – A Review. Premier Journal of Science 2024;1:100004

DOI: https://doi.org/10.70389/PJS.100004

Received: 24 July 2024

Revised: 9 August 2024

Accepted: 11 August 2024

Published: 10 October 2024

Ethical approval: N/a

Consent: N/a

Funding: No industry funding

Conflicts of interest: N/a

Author contribution: Michele Malandruccolo – Conceptualization, Writing – original draft, review and editing

Guarantor: Michele Malandruccolo

Provenance and peer-review: Commissioned and externally peer-reviewed

Data availability statement: N/a

ABSTRACT

Artificial Intelligence (AI) has emerged as a transformative technology in 2023, significantly influenced by the launch of ChatGPT by OpenAI on November 30, 2022. However, despite the overinflated expectations over these technologies there are still many challenges and risks associated with such technology.

In fact, despite the longstanding presence of AI technologies, numerous companies are still struggling to leverage them effectively. On the other hand, a significant shift has occurred in data analytics, fostering a more integrated data culture within organizations, driven by LLM models such as ChatGPT.

The current surge in investments from governments, educational institutions, and industries underscores the perceived high return on investment (ROI) and the focus on innovation over risk. AI approaches continue to evolve, and the potential benefits of such technology are vast. Indeed, AI algorithms have already demonstrated they can enhance operational efficiency, customer experience, and data-driven decision-making. These advancements have spurred innovative applications across finance, education, and healthcare, with notable momentum in diagnostics, potentially reducing healthcare costs globally, especially in developing countries.

The paper will initially delve into AI and machine learning algorithms, what these are and how to categorize them, to then analyse their impact, particularly on value creation and healthcare expenditures. It will then progress focusing on the challenges and bottlenecks in using these technologies, such as ethical concerns, including job displacement, misuse, environmental impact, privacy, copyright issues, and adverse patient outcomes in healthcare. Among many, navigating these challenges also require adaptation to an evolving regulatory landscape. This review article will conclude with an analysis of the current status and future expectations for AI technologies.

Keywords: Machine learning algorithms, Healthcare expenditures, Large language models, Ethical concerns in AI, Disease diagnostics

Introduction

Generative Artificial Intelligence was the thing in 2023, following ChatGPT launch by OpenAI on November 30, 2022. ChatGPT, an advanced large language model or LLM, was indeed credited for the triggering of the AI boom,[1] becoming the fastest growing application in history. But what is AI and what are the implications of using such technology? Has there been already a measurable impact on organizations?

Despite the high expectations around this "not new" technology,[2] many companies are still trying to understand how it can create value for them. It is no doubt that after November 2022 there has been a shift in the perception of data analytics with a more integrated data culture in our current society, whereby for many years organizations had failed to do so because the focus was mainly on the technology.[3]

Now, thanks primarily to ChatGPT (or Generative AI architecture) we have seen a surge in investments, whether from governments, educational institutions, or industries with many leaders feeling that they are getting a higher return on investments (ROI), focusing more on innovation rather than on the risks. Indeed, LLM models such as ChatGPT have a wide range of applications, from operations efficiency through customer experience improvements to better data-driven decision making. As these models continue to advance, also because of an increasing competitive landscape, we have seen innovative applications in finance, education, and healthcare. Regarding this latter, in diagnostics we have seen a big momentum with huge potentials to decrease health expenditure worldwide especially in the developing countries.

Nonetheless, the development and deployment of LLMs has raised several ethical concerns such as replacement of jobs, misuse of applications, environmental impact, privacy, copyright issues and undesirable patient outcomes in healthcare.[4] Therefore, such technologies will need to traverse the ever-evolving regulatory landscape before being widely deployed.

In this review article, we will initially deepen our understanding of AI and Machine Learning (ML) algorithms, describe the different categories, analyse their impact on healthcare settings and expenditures, what are the pros and cons, especially in healthcare whereby the implications on the patients may push back the progress of such technologies, to conclude with expectations and what's next for Artificial Intelligence.

Artificial Intelligence, Machine and Deep Learning

The beginning of the machine learning era dates back to 1959 when Arthur Samuel, a pioneer of the modern artificial intelligence coined this term.[2] Samuel designed a program that could self-play over time paving the way for modern models, where computers could learn and make predictions without the need of robust programming instructions. From the 'fifties to today the advancements in this field have increased substantially culminating with the rise of advanced large language models.

But what are they? Let us first get the definitions right. In the Venn diagram (Figure 1), we can see the different relationships between AI, Machine and Deep

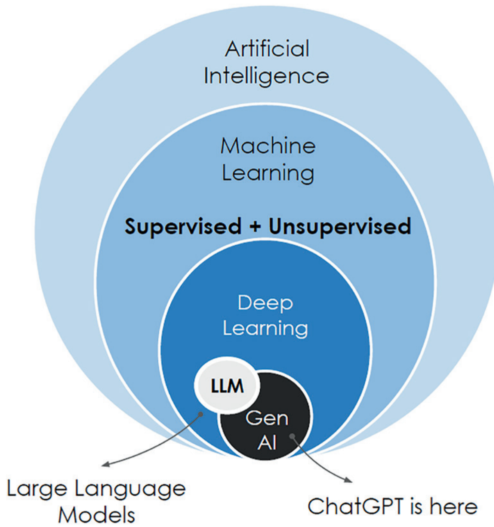

Fig 1 | Venn diagram showing the different relationships between AI, ML, DL, Gen AI and LLM

Learning (DL), Generative AI (where ChatGPT belongs) and LLMs. AI can be considered as a computer that mimics human intelligence and behaviour which through Machine Learning can predict outcomes based on what has learnt with the help of a subset algorithm within Machine Learning, Deep Learning whereby the machine learns from complex data to find patterns and simulate neural networks like those in the human brain. Ultimately, we come to ChatGPT, a cutting-edge Deep Learning Chatbot based on LLM (language models consisting of a neural network with billions of parameters) that uses Generative AI architecture to create something new through a learning process.

In ML we provide data to train the model so to recognize patterns in the data, the more data we give the more patterns will the model recognise. Database such as Kaggle have been created to store and collect such data for ML projects. Machine Learning can be divided in Supervised and Unsupervised Learning Algorithms with main objectives to group data into categories based on pre-existing patterns and to predict outcomes scanning through such patterns.

In supervised Learning the model is trained with labelled data (Figure 2) whereby the actual value of each label is known. The data is then divided in test and train where the output of the train data is our prediction whilst the patterns identified from the training data are applied on the test data for classification or prediction purposes.[5]

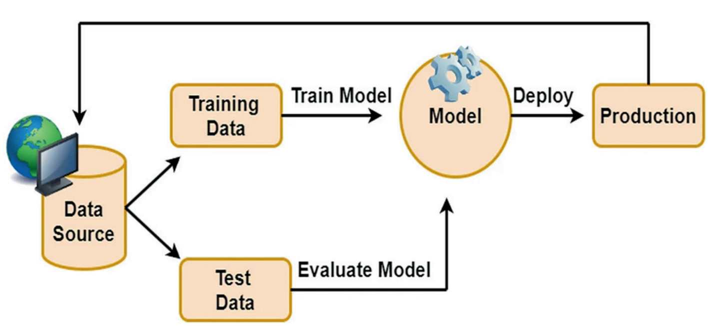

Fig 2 | Supervised learning workflow[5]

A typical example of supervised learning is a decision tree, the most popular and effective supervised learning techniques for classification problems, which works well with both categorical and continuous variables. A decision tree is characterised by nodes or decisions to take and branches which represent the choices for each node. It is a binary classification model with only two possible classes, yes or no which helps us moving to the next layer in the tree (credit cards, age, etc.) so to ultimately decide whether to give a loan or not to a customer who applied for it as represented in the example of Figure 3.

On the other hand, in unsupervised learning, we provide the model with complex data which has not been labelled. Therefore, the model has no teacher and will need to cluster the data based on similarities and learn something new every time new data is added to the model. Therefore, in the loan example, the model may generate different categories based on the similarities which can then help us making decision on which customers to target to give a loan (Figure 4).

What about DL and Generative AI? As we might have seen in the news, these AI approaches continue to evolve and the potential benefits are huge, spanning from retail through technology to healthcare. Organizations can leverage them to enhance operations, improve decision making, automate tasks and personalise experiences. In healthcare alone researchers are exploring new methods for better disease predictions. Indeed, based on data availability and data quality more insights can be gained to make promptly informed and sensible decisions.

Although, one of the challenges of such models can be the introduction of bias and prejudice potentially exacerbating social inequalities, hence the need of more regulated environments, the use of Deep Learning Algorithms can also lead to more effective and earlier disease diagnosis leading to a substantial decrease in health expenditures,[6] better resources allocation and improved patient health outcomes. This is especially true in poor regions such as Africa where resources are limited and access to healthcare not available to everyone. In fact, financial barriers are a main driver of inequality above all when we compare urban vs rural areas (Figure 5). Therefore, decreasing healthcare costs could shrink the gap between these areas, reducing inequalities.

Let us now analyse in more details how AI approaches through ML and DL can be used in healthcare settings to improve health outcomes and reduced healthcare expenditure for patients worldwide.

The Potentials of AI in Healthcare

As previously anticipated AI and ML approaches can be used in many areas such as technology, retail, or healthcare to foster innovation, operational efficiency or provide a better customer solution experience. In healthcare, ML algorithms can help with prediction

 DOI: https://doi.org/10.70389/PJS.100004 | Premier Journal of Science 2024;1:100004

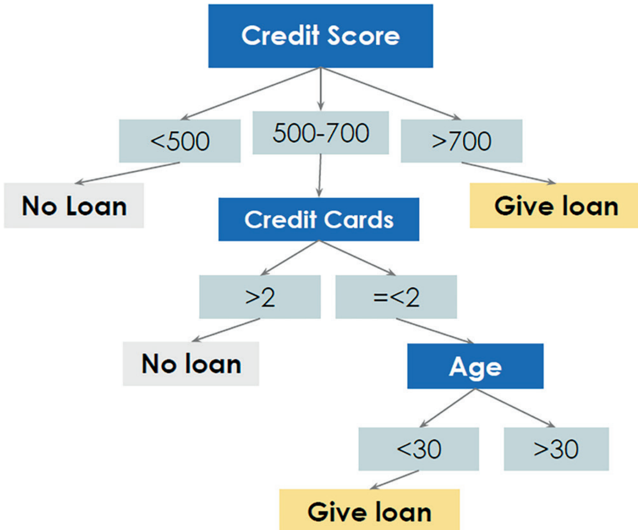

Fig 3 | Decision tree example for giving a loan to a customer

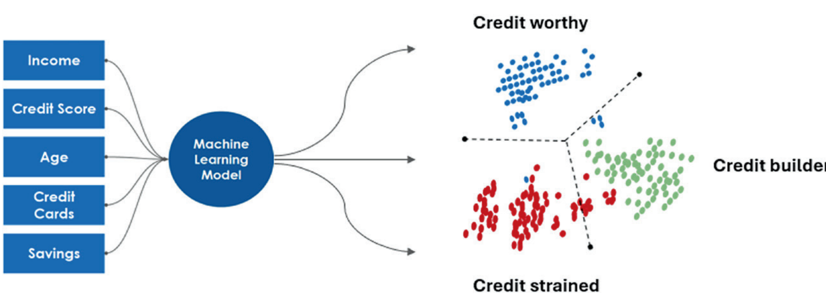

Fig 4 | Unsupervised learning loan example

of disease outbreaks, or provide a better patient risk stratification, DL can be used in drug discovery, histopathology or in image segmentation whilst Generative AI is able to create new data and lead to more personalised treatment plans (Figure 6).

Generally speaking, when we are trying to solve complex problems, various forms of data can be available, and the type of data used depends primarily on the objectives and available resources. In recent years Deep Learning techniques such as convolutional neural networks (CNNs) have become an area of active research and application in cancer prediction and malaria detection.

CNN models are well-suited for image classification (images are the type of data to run such models). In malaria detection only[8] they have shown high accuracy rates (>94%) and a low chance of missing infected patients (low number of false negatives).

Automated solutions such as CNN models have consistently shown higher accuracy than manual classification.[9] In fact, when we consider malaria infection, the two main types of diagnosis are either antigen-based rapid diagnostics tests or microscopic detection from blood cell samples,[10] with this latter the most common methodology especially in countries where resources are limited.

Then the question is, how can DL really help in improving early and accurate detection of malaria cases? Being this a classification algorithm, the CNN-base models are trained on labelled datasets containing images of parasitized and unparasitized red blood cells. The CNNs learn to extract relevant features from the cell images and classify (predict) them as infected or uninfected. To support us evaluating the best CNN

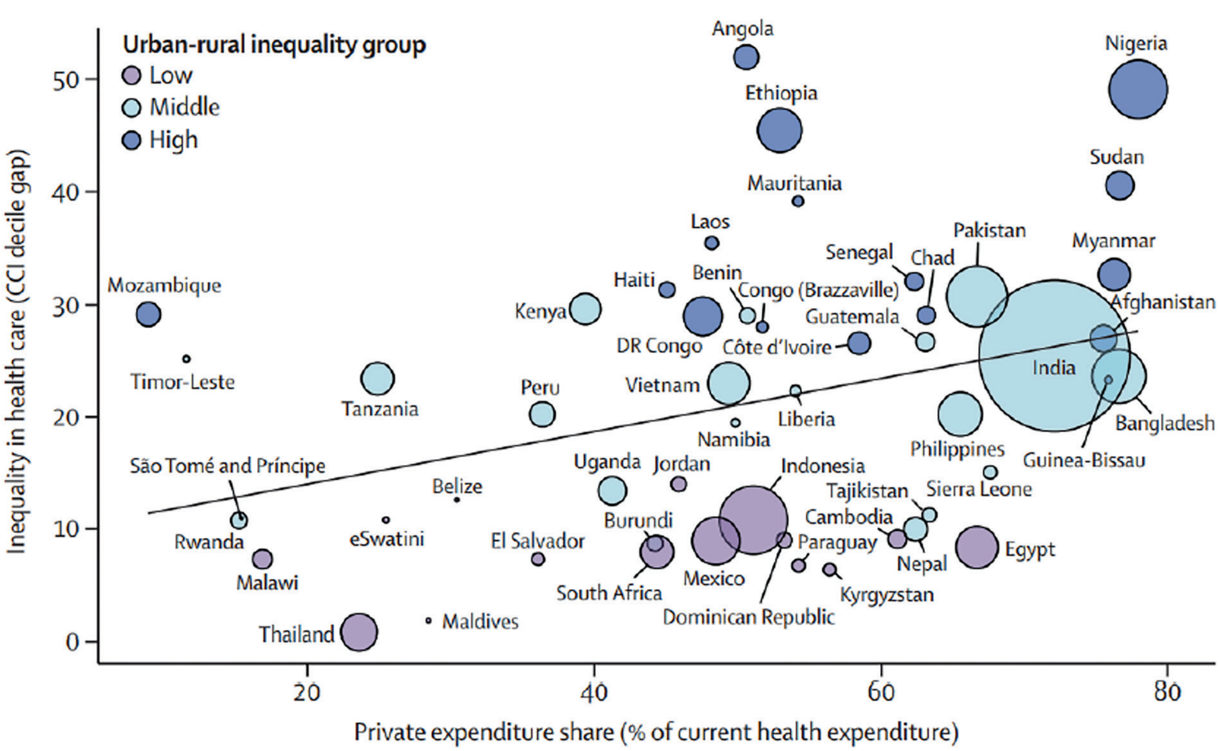

Fig 5 | Relationship between social inequality in healthcare delivery and share of private expenditures[7]

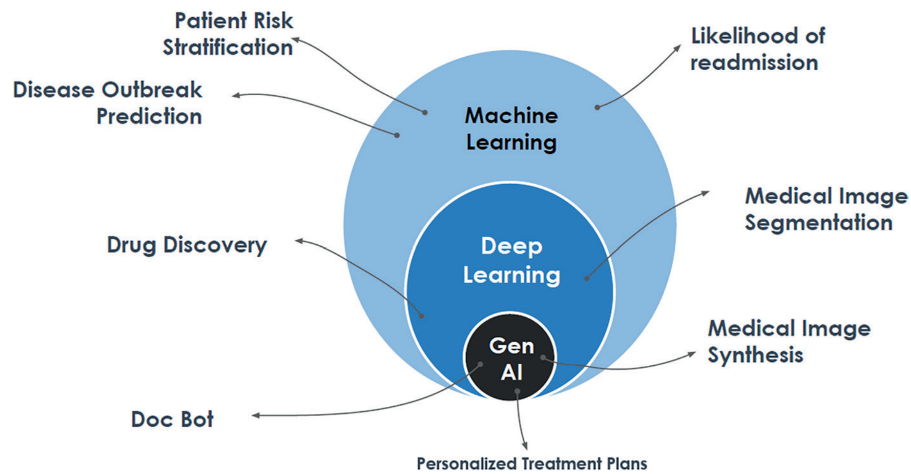

Fig 6 | Healthcare business problems that could be solved by AI

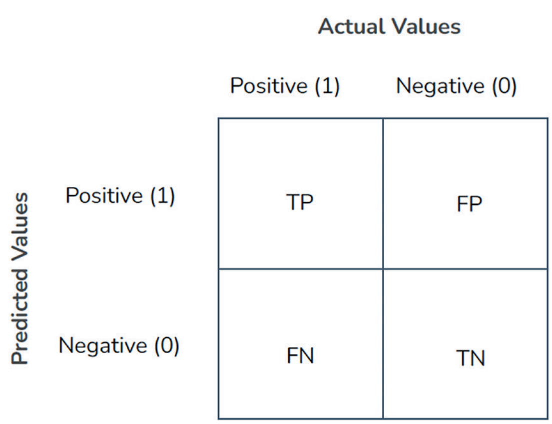

Fig 7 | Confusion matrix—True Positive (TP) and True negatives (TN) occur when the predicted and actual values match to either detect the presence or absence of infection respectively

models, confusion matrixes are built to compare actual and predicted data classes and measure the performance of such classification algorithm (Figure 7).

When building confusion matrixes there are several parameters to consider (recall, average precision, F1 score etc.) or model improvements (e.g., transfer learning, data augmentation, fine tune hyperparameters for example) that can be tried to increase model performance. Scientifically speaking, without entering the technicalities of such performance improvements, failing to treat actual malaria cases can be life threatening whilst treating patients who actually do not have malaria is less harmful, although it leads to drug resistance that can be a concern in areas where malaria is prevalent. Therefore, having a higher number of false positives (FP) (treating non-malaria cases) is less harmful than having a higher rate of false negatives (FN) (missing true malaria cases).

In 2022 there were 249 million malaria cases globally of which 94% in Africa only with approximately 400,000 deaths annually worldwide. This can pose a burden for healthcare settings especially in regions like Africa where the malaria parasites are prevalent and where there is high inequality between urban and rural areas.[7] When we consider the malaria management costs, in Rwanda only these are up to USD 300 million per year, with a median financial cost per case for diagnosis equal to USD 4.32 and for treating an episode of severe malaria equal to USD 30.26. Although serious, malaria is preventable and curable if detected early. It is clear that prompt and accurate diagnosis are crucial to reduce the progression of simple malaria cases to more severe cases and therefore substantially decrease health expenditures and reduce mortality.[6]

Other than malaria, Deep Learning solutions for better disease diagnosis have been applied successfully in radiology. This is the case of Zebra Medical Vision (Zebra) with 7 approved algorithms in Europe and 1 in the US which have been supporting doctors in making diagnosis for emphysema, breast cancer, compressed fractures, brain bleeds and other diseases. The company has been able to analyse 1mln scans (CT and X-rays) and charged hospital $1 per scan (the average cost of a CT scan in US was $896). This means substantial savings for patients and healthcare worldwide, but also a better quality of life (many patients are not well diagnosed, diagnosed too late or there is disagreement between interpretations from doctors) and a quicker, more efficient, and reliable process for doctors. Zebra has been the leader in the use of Deep Learning in radiology and with currently a more competitive landscape, they are looking to expand their portfolio to diseases such as tuberculosis or rare diseases and at the same time develop packages that could diagnose more than one condition at a time whilst their approved packages are continuously improved to become more sensitive and specific.[11]

There are many other areas whereby AI applications are currently assessed to bring down the financial burden for governments, patients or organizations. LLM models are being developed to be more aligned with the medical domains and safely answer medical questions. An example is Med-Palm 2 developed by Google or companies like Regard that are trying to reduce the level of burnout in doctors by improving the quality of clinical documentation through data. It seems that

DOI: https://doi.org/10.70389/PJS.100004 | Premier Journal of Science 2024;1:100004

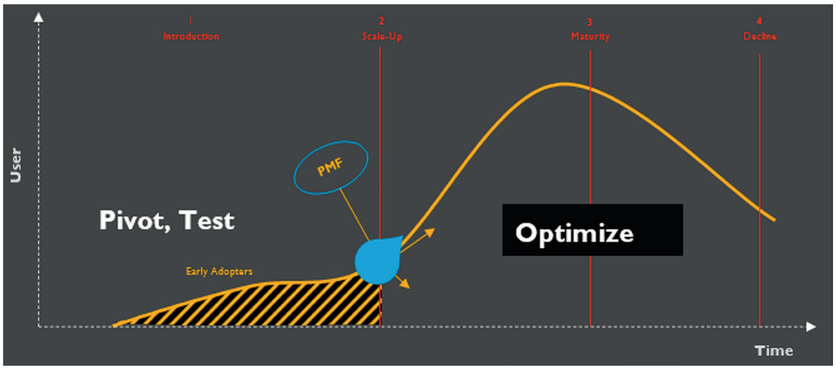

Fig 8 | Test-Measure-Learn cycle for the Product Market Fit (PMF)

their work is paying off with 30% reduction in burnout among clinicians.[12]

ML algorithms have shown enormous potential even in cancer research and other healthcare settings whereby more personalised treatments are fostered to decrease the medical costs associated with them. In cancer research, two important aspects are to gather genomic and clinical data. The first involves studying the genes, their expression patterns, and mutations whilst the second the study of medical history, treatment records, demographic information etc. Such data can be integrated in Machine Learning algorithms to provide essential insights into patients' background, identification of biomarkers that could be used for cancer diagnosis and a better understanding of disease progression to ultimately select the right treatments for the right patients (personalised approach).[13] This could be further expanded at the population level where such data could be integrated in electronic health records (HER) to identify trends and patterns related to cancer incidence.[14] Complexity could be reduced that along with better resources allocation and more efficient and effective processes can lead to additional savings. Big companies such as Novartis, Roche, and Bristol Myers Squibb (BMS) are already using AI to identify better drug candidates for a range of disorders. Bayer has partnered with Recursion Pharmaceutical in 2020 with a deal worth $1 billion to treat fibrotic diseases whilst Novartis has partnered with Microsoft to apply the Al-gorithms to large datasets. There is still a long way to go in drug discovery but the competitive advantage this technology can bring in optimization and efficiency of processes highlights that AI can soon become a critical aspect in the design of every new drug.[15]

Another instance where the potentials of ML models could lead to decrease healthcare costs is in insurance, like for patients after coronary artery bypass grafting (CABG) surgeries. In fact, such patients have two possibilities after surgeries, gradual recovery or rehospitalization due to several further complications. Since readmission has high incidence, both patients and hospital must pay the costs, with substantial and increased healthcare expenditures. Huang and colleagues using National Health Insurance Research Database (NHIRD) have built machine learning predictive models based on SVR and XGBoost algorithms to effectively reduce medical management costs and improve the quality of treatments for patients who underwent CABG in Taiwan.[16]

Discussion and Concluding Reflections

The paper has highlighted the promise of AI algorithms to bring technological advancements that could improve health patient outcomes, reduce health expenditures, and make treatments more accessible to everyone. Nonetheless there are several big challenges to be considered. First the costs in introducing and implementing such technologies in organizations. As we have seen there has been surely a shift towards a more data-driven culture accelerating the digital transformation many companies and organizations needed. However, this will require substantial investments in upskilling the workforce (e.g., trainings or further education programs) or hiring the right talents to fulfil the jobs of the future. The major world economies stand to lose $11.5 trillion in potential growth by 2028 if they cannot bridge the skills gap.[17] On the other hand, processes automation can lead to higher savings for companies.

Despite that, companies are still struggling to grasp the full potentials of AI especially because knowledge (or data) within organizations is generated and captured through several sources hence it is difficult to organize such knowledge in a way that AI solutions can be deployed systematically where needed in an effective or efficient way. This leads us to the second limitation related to the technology itself (e.g., hallucinations, factual inaccuracies, quality of the data etc.). In malaria detection for example, technology availability can really hinder the potentials of AI. Good sensors availability can introduce noise to our CNN image classification and to overcome such problems good machines are required which is likely to be not affordable for rural areas or poor countries. At the same time, sound understanding of programming is needed to select the best AI models that can help in decision-making. This must be coupled with trained clinicians who have a deep knowledge of the disease life cycle and how to interpret the data at hand. Having a higher rate of false negatives (missing patient with the disease) can increase mortality (program sensitivity) whilst on the other hand giving the drugs to patients who do not actually have the disease (false positives) can lead to drug resistance (program specificity). These technological limitations can be problematic in businesses but could be deadly in healthcare applications.

Therefore, some key questions are: will AI keep up the momentum and really solve many of the complex problems our society is facing nowadays? Will organizations like OpenAI be able to reach a stable maturity in the product market fit (PMF) curve or will they quickly decline either because the technology will not live up to overinflated expectations (Figure 8) or due to setbacks especially in the medical domain?

Lastly but not least the ethical and regulatory challenges. The impact on jobs due to the skills gap is huge and although automation could increase productivity and efficiency the effect on the labour market may be even bigger, with increase in income inequalities. These inequalities could be exacerbated by increase of carbon emissions as more computational power will be needed worldwide leading to extreme consequence, especially in the developing world. Therefore, the real need to minimize the ecological footprint of AI technologies so that the Paris Agreement targets by limiting global warming to 1.5˚C by 2050 are achieved. Then the misuse of the data and the privacy concerns whereby public information could be exposed and reveal the identity of individuals, with major concerns especially in healthcare settings. OpenAI wanted to stay true to their mission of making AI accessible to everyone, and bringing innovation in operational efficiency, customer experience and data driven decision making. However, they were confronted with many questions related to bias, fairness, transparency, and accountability.[4]

To conclude, what is next for AI and what are the expectations?

In terms of regulation expectations, we have seen already significant initiatives in the European Union with the AI Act whilst the US is taking a more AI risk management approach spread across various federal agencies. In addition, the Trade and Technology Council (TTC), a joint US-EU organization published in December 2022 a roadmap for the evaluation and Measurement of AI and associated risks for better policy development and alignment. These initiatives should lead to the early identification of risk factors associated with the use of AI. One significant risk is associated with the leakage of data that can lead to several negative impacts on enterprises. In 2023, there have been 110.8 million accounts leaked globally only in Q2. In these instances, the attacked companies keep bleeding money even afterwards (because workers still get paid when for at least 3 weeks there will not be much to work), especially when there are not any contingency plans or a robust data governance in place. These outcomes can be disastrous for small businesses and for customers whose data has been leaked. Strategies for AI governance frameworks including cybersecurity are needed so that enterprises can safely transition towards the age of AI.

Then there are other type of expectations. For example, board and CEO expectations whereby the push to implement Generative AI is happening mainly top down, with board members challenging executive leaders to build plans that employees can actually use. On the other hand, employees are first worried on how AI will impact their jobs and second on how and when these technologies can be applied in their workplaces. Many reports that they could not distinguish between Generative AI and AI, hence the increasing need of upskilling or hiring more talents. Then, the investors' expectations which reward those companies able to achieve an efficient growth. Goldman Sachs has forecasted an increase in productivity due to AI by 30% or more in the next ten years. Lastly, the costumer expectations who are looking at faster, more personalised content and a better customer experience across an ever-increasing ecosystem of touchpoints.[18]

Despite the hype around AI and the various potential applications of such technology there are still risks and challenges to overcome which will tell us to a certain extent the progress to come in the next years. Education and training are an important aspect to change cultures either within organizations or at individual level. Nonetheless, many still do not understand how AI works, the various purposes to which it can be applied and how it should be governed. Therefore, in a fast-moving digital world, the long-term vision of executive leaders should be to create new strategies, adapt or change their business models and transform the relationship between humans and machines. It is then everyone responsibility to understand the technology and make these innovations safe, sustainable, and widely accessible "for the betterment of humanity" to phrase the words of Sam Altman at OpenAI.[4]

References

1 Meredith S. A 'thirsty' generative AI boom poses a growing problem for Big Tech. CNBC.com, 2023.
2 A.L. Samuel. Some studies in machine learning using the game of checkers. IBM Journal.; 1959;3:210-223.
3 Davenport, T.H., Bean, R. Big companies are embracing analytics, but most still don't have a data-driven culture. Harvard business review. 2018.
4 Samila S., Berrone P. Open AI and the large language model market. IESE Business School. 2023.
5 Mahesh B. Machine learning algorithms – A Review. International Journal of Science and Research (IJSR); 2018;9:381-386.
6 Masimbi, O., Schurer, J.M., Rafferty E., et al. A cost analysis of the diagnosis and treatment of malaria at public health facilities and communities in three districts Rwanda. Malaria Journal, 2022;21:150.
7 Cookson, R., Griffin, S., Norheim, O.F., Culyer, A.J., Chalkidou, K., Distributional cost-effectiveness analysis comes of age, Science Direct, Value Health, 2021;24(1):118-120.
8 Solat, S.M., Bhor, A., Yelwande, R., Dalavi, A., Chikane, P., Malaria cell identification using deep learning and CNN, International journal of novel research and development, 2023;8(1):730-733.
9 Davidson, M.S., Andradi-Brown, C., Yahiya, S., Chmielewski, J., et al., Automated detection and staging of malaria parasites from cytological smears using convolutional neural networks, Biological Imaging; 2021;1:e2-1-e2-13.
10 Fikadu, M., Ashenafi, E. Malaria: An overview, infection and drug resistance, Dove Press Review, 2023;16:3339-3347.
11 Greenstein, S., Gulick, S. Zebra Medical Vision. Harvard Business School; December 5, 2019.
12 Aquino, S. How healthcare firm regard uses artificial intelligence to usher medicine into the modern age. Forbes; 2023.
13 Vadas, A., Holder, J., Siebert, A. A new generation of drug therapies requires new business strategies, Harvard Business Review; February 27, 2024.
14 Yaqoob, A., Aziz, R.M., Verma, N.K. Applications and techniques of machine learning in cancer classification: A systematic review. Springer Link, 2023;3:588-615.
15 Savage, N. Tapping into the drug discovery potential of AI. News Feature; 2021.
16 Huang, YC., L i, SJ., Chen, M., Lee., TS. The prediction model of medical expenditure applying machine learning algorithm in CABG Patients. Healthcare 2021;9(6):1-13.
17 Irving, D. The digital skills gap: what workers need for the jobs of the future, Rand Essay; 2022.
18 Bant, A., Poitevin, H., Greene, N., Brethenoux, E. 5 Forces that will drive the adoption of GenAI. Harvard Business Review; December 14, 2023.