**Criteria for Successful Integration of Machine Learning Tools in a Medical Setting**

**A Research Paper submitted to the Department of Engineering and Society**

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Pamela Flintsch Medina
Spring, 2020

On my honor as a University Student, I have neither given nor received
unauthorized aid on this assignment as defined by the Honor Guidelines
for Thesis-Related Assignments

# Criteria for Successful Integration of Machine Learning Tools in a Medical Setting

## *Fears Surrounding Technology*

The opening of the 1984 science fiction film *The Terminator* begins, "Their [the machines] war to exterminate mankind has raged for decades, but the final battle would not be fought in the future. It would be fought here, in our present" (Cameron 1984). These remarks embody the commonly held fear of technological dependence resulting in negative consequences, particularly if the technology is seen as self-adaptive. Fears, such as these, inspire distrust in in novel technologies, such as machine learning, despite their potential benefits. In the case of applications of machine learning in medical field, this fear leads to the unrealized potential for these to improve patient care. The purpose of this research paper is to evaluate necessary criteria for successful implementation of machine learning technology in the medical field.

## *Research Question and Methods*

In order for machine learning to provide its intended utility, the technology must be properly applied and accepted by the system in which it is intended to be integrated. This paper seeks to define specific criteria that if met allow for the successful integration of machine learning within the medical field.

In order to answer the research questions, data was collected from multiple sources and is analyzed using three methodologies. Historical case studies describing the implementation of machine learning technologies in assessing disease risk, tumor detection, and drug delivery are used to emphasize the immense utility of the technology (Bazazeh and Shubair 2016; Fox and Kriegl 2006; Manolio 2010). A historical case study analyzing the switch from written to digitized medical records to exemplify a transition from manual to computational methods within the medical field (Kupwade Patil and Seshadri 2014). Additionally, Discourse analysis of the film *Gattaca* and television series *Black Mirror,* are also utilized to determine to establish how advanced medical technologies and their associated risks are currently perceived (Charlie Brooker n.d.; Niccol 1997). In addition, this same mode of analysis is used to interpret an interview with an established expert in the field of artificial intelligence, Stuart Russel (Fears of an AI pioneer | Science n.d.). This interview serves to establish guidelines on how machine learning technology should be constructed in order for it intended benefit to be realized. Policy analysis of the United States the Health Insurance Portability and Accountability Act is also conducted in order to determine what legislation currently exists and needs to be present prior to implementation of machine learning technologies to ensure patient privacy protection (Annas 2003).

***Defining Machine Learning***

Since its development by Arthur Samuel, in 1959, machine learning has become increasingly prominent in day-to-day life (Brink, Richards, and Mark Fetherol n.d.). Originally created to automate a game of checkers, machine learning is now used for many applications, from traffic prediction to computational finance and product recommendation (Boyarshinov 2005; Garcia Esparza, O'Mahony, and Smyth 2012; Rzeszótko and Nguyen 2012) Another application of machine learning, Google Lens, allows for users to search images based on the content of the images (Vanitha, Jeevaa, and Shriman 2019). Machine learning, as its number of applications continues to grow, has proved useful across several industries and fields.

Machine learning has one primary focus: deriving patterns from big data (Brink, Richards, and Mark Fetherol n.d.). Unlike traditional computational methods, machine learning does not require the researcher to know exactly what trends for which they are looking for.

Instead, machine learning uses an algorithmic approach to search for patterns within data sets and analyzes how well these patterns describe that given data set. This “learning” is particularly useful with big data sets, as patterns are often not clearly visible or obstructed by the inclusion of unnecessary data points.

### *Medical Applications of Machine Learning*

Machine learning’s utility in analyzing large data sets makes it of particular interest in the medical field. In a paper published in *World Neurology*, Cacciola et al. describes the use of big data in the medical field and the current lack of physician proficiency in interpreting it (Cacciola, Conti, and Tomasello 2019). They argue machine learning provides immense utility in the finding of non-obvious patterns in clinical data and serves to remedy the physicians’ shortcomings.

Many applications of machine learning in the medical field are being heavily explored and show promising results. Within radiology, machine learning has been used to optimize assessment of tumor treatment areas and dosage calculations to prevent unnecessary patient exposure to radiation (Garcia-Sanchez et al. 2019). Machine learning has also been implemented in the automation of differentiation between benign and malignant breast cancer tumors using MRI data (Ji et al. 2019). The model was able to differentiate with a sensitivity of 99.5% and recommended 9.6% fewer biopsies, indicating implementation would improve patient care. Additionally, machine learning has proved a useful tool in the study of proteomics, genomics, and drug delivery (Fox and Kriegl 2006; Leung et al. 2016; Sonsare and Gunavathi 2019). These applications are just a few of the many ways in which machine learning is currently being used to improve patient care.

### *STS Framework*

Actor-Network Theory (ANT), along with the concepts of technological momentum and paradigm shifts were used to guide research and to avoid bias created as a consequence of reliance on only one. ANT has been criticized for giving equal weight to both human and non-human factors but provides superiority in its ease of identifying key stakeholders (Latour 1996; Murdoch 1998). This is controversial, specifically in health care situations, where you could be equating a patient's interaction with a highly trained professional, to the patient's interaction with a piece of technology only made useful through the experts interpretation of its outputs. Use of this framework in conjunction with that of technological momentum allows for emphasis on human factors while still benefiting from the ease and clarity ANT supplies. Additionally, the switch from reliance on manual expert assessment to reliance on automated computational assessment was analyzed as a paradigm shift.

***Results***

A technology's success is directly correlated to its level of utility to the user. For any technology to provide its intended benefit users must be willing and able to use the technology over the current system. Machine learning technologies hold promise in assisting patient care, however, without acceptance of the technology within the field this promise cannot be achieved. This research paper proposes the following criteria for enabling integration of machine learning technologies within the medical field:

1. The technology must be proven consistently reliable and superior to the current metric.
2. All individuals who come in to contact with the technology must be first educated on its capabilities.
3. The technology must have a defined goal with defined capabilities.

4. A system of protection for information derived from the technology must be in place prior to the collection of data.

The achievement of these criteria would enable the acceptance of the technology and the realization of the advantages it has to offer. This paper analyzes existing machine-learning technologies and provided benefits to emphasize the motivation behind their implementation. Additionally, societal perspective of novel technologies is explored in order to define the obstacles faced during integration of machine learning into the medical field. Current policy surrounding data collected by these technologies and the weaknesses of these policies are identified to determine what form of future legislation needs to be set forth to ensure patient protection.

*Combating Technophobia*

Many fields, including medicine are undergoing a paradigm shift, using computational methods in place of manual analysis. One of the most prominent examples of this paradigm shifts, is the clinical transition from written to digitized medical records (Kupwade Patil and Seshadri 2014). The analysis of these digitized records, or "Big Data", has shown promise in creating a more proactive, instead of reactive, healthcare approach(Cacciola, Conti, and Tomasello 2019; Kupwade Patil and Seshadri 2014). Additionally, computational technologies have shown to be more success in differentiating between benign and malignant tumors in breast cancer than the current clinical alternative (Giger n.d.). By analyzing the replacement of prior clinical standards by their computational equivalent as a paradigm shift, it becomes apparent that the driving force in the shift is held within the technology's apparent superiority over current methodologies. Despite the advantageous nature of this technologies acting as a driving force, obstacles still exist to resist this paradigm shift. For successful integration of these novel technologies, society must be willing to use it.

A major obstacle to the integration of machine learning in the medical field is a generalized fear of the new and complex technology referred to as technophobia (Ha, Page, and Thorsteinsson 2011; La Torre et al. 2019; Meuter et al. 2003). A study concerning autonomous vehicles showed that social perception had a significant impact on an individual's willingness to interact with the technology (Koul and Eydgahi 2020). This fear is a consequence of not only the novelty of the technology, but also a popularly seen narrative in media that stresses the dangers of technological dependence. Award winning shows, like *Black Mirror* paint a vivid picture of dystopian societies that could result from embedding technology within our everyday life (Charlie Brooker n.d.). Images like these embed deep fear and discourage acceptance of machine learning technology. However, granting the users full of the instruction on the capabilities of the machine learning technology and its limitations has the power to curve these fears. If individuals are aware of a technology's exact capabilities an understanding of risk is achieved. By working to clearly define and vocalize risk, the notion of never-ending unintended consequence associated with machine learning technologies can be combated.

### *Defining Capabilities*

Fear of the uncertainty machine learning technology may hold, however, is not only held by the layman. Stuart Russel, professor of computer science at UC Berkley, argues that these fears are valid but can be avoided by adhering to strict criteria during the construction of said technology (Fears of an AI pioneer | Science n.d.). Russell states that these technologies should be made to fill a specific need and not in an effort to manufacture a new means of intelligence. Following these criteria ensures that the technophobic ideal of an artificial intelligence takeover cannot be realized. For the integration of a machine learning technology to be successful within the medical field, it should follow the criteria set forth by Russell, as to ensure the patient is being provided with a net benefit. Russell also makes a point to state that all fields have risks, in the case of machine learning these risks just tend to be more prevalently attached to its name. In an interview Russell stated, “No one in civil engineering talks about ‘building bridges that don't fall down.’ They just call it ‘building bridges.’” The quote provides an interesting perspective of looking at machine learning without the bias derived from societal technophobia.

*Data Security*

The overarching goal of machine learning technology in the medical field is to improve patient care. However, in doing so, it must be ensured that the patient is being provided with a net benefit. An important factor in ensuring this benefit is the protection of patient data. It has already been suggested that the collection of data during genetic screenings, which is often seen used in machine learning applications of analyzing disease risk, could be used by insurance companies as a metric for setting rates (Nill, Laczniak, and Thistle 2019). Applications of data like these present a risk of hindering patient care instead of advancing it. For this reason, strict legislation has to be set on how data can be used prior to the data’s collection.

This issue of unintended risk of development can be analyzed using a technological momentum framework. In the case of applications of machine learning that predict disease risk using genetic information, this technology was created to fill a societal need (Bazazeh and Shubair 2016; Leung et al. 2016). The intention of the technology was to benefit society by allowing for the necessary monitoring or treatment associated with mediating the disease under consideration. However, now that this technology was created, it has the opportunity to shape the way society interacts with itself. The establishment of an individual as “at risk” for certain ailments and diseases has many unintentional consequences.  For example, risk diagnoses for cancers have been found to have an intense emotional impact on the patient (C and Rt 1996). Additionally, the film *Gattaca* embodies the possible unintentional consequences of genetic screenings, depicting a society in which data derived from these screening has become a platform for discrimination (Niccol 1997). Technology must therefore, be created, adapted, and controlled in a way which mitigates these risks, while still allowing for it to provided its intended utility. In this way, technology follows the fame-work of technological dependence. The technology was shaped by society, but in turn the technology has a direct impact on the way society interacts.

Within the United States the Health Insurance Portability and Accountability Act sets standards for how patient data is protected (Annas 2003). It ensures patient data will not be shared with non-essential personnel as to protect the patient's privacy. In machine learning, this is often seen in the form of de-identification of patient data. However, with opt-in screenings that are not necessarily conducted for medical reasons this same confidentiality is not always ensured. By creating technologies that can take data from screenings like these and use it to derive useful information, developers are creating a potential risk for the patient. Through tight regulation of how and where these machine learning technologies can be applied risk is mediated. By regulating how the technology is used a net benefit for the patient can be ensured.

*Framework Utility in Analysis*

Actor-Network Theory (ANT), technological momentum, and the concept of paradigm shifts all provide specific utility while analyzing the research collected. ANT identifies key stake-holders, such as, medical professionals (physicians, radiologist, nurses, etc.), hospital administrators, and patients. This is important in identifying the limitations of our research in considering important perspective. The ANT framework also allows for the examination of interaction of machine learning systems with non-human factors, such as the computing power the hospital has available to them. This comes into play when analyzing how these systems should be constructed. Once stake-holders are identified, the theory of technological momentum serves as a tool for analyzing the complex relationship between society and machine learning within the medical field. Additionally, the transitions of the medical field from dependence on professional assessment to reliance on automated technologies is assessed as a paradigm shift. Evaluation of the relationship between society and the technology through these frameworks results in a better understanding of what steps need to be taken in order to integrate machine learning into the medical field. A task, if achieved, has the potential to save lives.

***Limitations***

There are several limitations to this analysis that should be mentioned. Firstly, the technology has only recently started being adapted into the medical field so documentation of a patient perspective is severely lacking. Patients must be comfortable with the technology in order for it to provide its intended benefit. Additionally, due to limitations presented with the ongoing COVID-19 situation, it was not possible to conduct interviews with health care professionals at this time. These experts would provide critical input into what metrics are necessary in order for the machine learning technology to be preferred over the current clinical standard.

***Next Steps***

This analysis provides interesting implications for future research. Most prominently, a patient perspective of applications of machine learning within the medical field needs to be analyzed in order to gauge comfortability with the technology. This is important because although a generalized societal technophobia is known, more specific research should be done to take into account the specific circumstances surrounding the patient experience. Additionally, more interviews with medical professionals across several specialties should be conducted to better understand an expert perspective of the technology, as they will e the individuals actually interacting with the technology.

***Conclusion***

Novel applications of machine learning in new fields have become increasingly prevalent. Within the medical field these applications have shown promise in advancing the medical field. The benefits machine learning has to offer for patient care can only be realized if the technology is successfully integrated into the medical field. In order to do so expert opinion must favor the technology over the current metric. If the technology holds obvious superiority over the current clinical standard and proves reliable, this must be clearly presented to experts in order for them to embrace the technology. Once the technology is embraced by experts, it must then be embraced by society, in order to be considered successfully integrated. Generalized technophobia must be combatted through education of the capabilities of the technology in order to achieve this goal. Additionally, to ensure that the technology is actually providing the patient with a net benefit, the information derived from the technology must be protected through strict legislation. I argue that by meeting these three criteria society can successfully take advantage of the benefits machine learning has to offer the medical field.

## References

Annas, G. J. (2003). HIPAA regulations—A new era of medical-record privacy? *New England Journal of Medicine*, *348*(15), 1486–1490. https://doi.org/10.1056/NEJMlim035027

Bazazeh, D., & Shubair, R. (2016). Comparative study of machine learning algorithms for breast cancer detection and diagnosis. *2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA)*, 1–4. https://doi.org/10.1109/ICEDSA.2016.7818560

Boyarshinov, V. (2005). *Machine learning in computational finance* [Phd]. Rensselaer Polytechnic Institute.

Brink, H., Richards, J. W., & Mark Fetherol. (n.d.). *Chapter 1. What is machine learning? - Real-world machine learning*. Retrieved January 31, 2020, from https://learning.oreilly.com/library/view/real-world-machine-learning/9781617291920/kindle_split_000.html

C, L., & Rt, C. (1996). Emotional and behavioral responses to genetic testing for susceptibility to cancer. *Oncology (Williston Park, N.Y.)*, *10*(2), 191–195, 9; discussion 200.

Cacciola, A., Conti, A., & Tomasello, F. (2019). Big Data analysis: The leap into a new science methodology. *World Neurosurgery*, S1878875019325604. https://doi.org/10.1016/j.wneu.2019.04.268

Cameron, J. (1984, October 26). *The Terminator*.

Charlie Brooker. (n.d.). *Black Mirror*. Netflix.

*Fears of an AI pioneer | Science*. (n.d.). Retrieved February 22, 2020, from https://science.sciencemag.org/content/349/6245/252?casa_token=lQBNcdrxsPYAAAAA:G_64bPmrXDEAOsknOVLirNb_zdbZVm5U6nJdrz6LwfShRIzKMAoWmCDpFNYq_Zu2s52ReZON001qTkI

Fox, T., & Kriegl, J. M. (2006). *Machine learning techniques for in silico modeling of drug metabolism* [Text]. https://www.ingentaconnect.com/content/ben/ctmc/2006/00000006/00000015/art00003

Garcia Esparza, S., O'Mahony, M. P., & Smyth, B. (2012). Mining the real-time web: A novel approach to product recommendation. *Knowledge-Based Systems*, *29*, 3–11. https://doi.org/10.1016/j.knosys.2011.07.007

Garcia-Sanchez, A.-J., Garcia Angosto, E., Llor, J. L., Serna Berna, A., & Ramos, D. (2019). Machine learning techniques applied to dose prediction in computed tomography tests. *Sensors (Basel, Switzerland)*, *19*(23). https://doi.org/10.3390/s19235116

Giger, N. A. and B. Q. H. and M. L. (n.d.). *A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets | Kopernio*. Retrieved April 9, 2020, from https://kopernio.com/viewer?doi=10.1002%2Fmp.12453&token=WzEyMDY5NjAsIjEwLjEwMDIvbXAuMTI0NTMiXQ.4JtK_oB-6QrZqwB0Scdki424KT0

Ha, J.-G., Page, T., & Thorsteinsson, G. (2011). A study on technophobia and mobile device design. *International Journal of Contents*, *7*(2), 17–25. https://doi.org/10.5392/IJoC.2011.7.2.017

Ji, Y., Li, H., Edwards, A. V., Papaioannou, J., Ma, W., Liu, P., & Giger, M. L. (2019). Independent validation of machine learning in diagnosing breast Cancer on magnetic resonance imaging within a single institution. *Cancer Imaging*, *19*. https://doi.org/10.1186/s40644-019-0252-2

Koul, S., & Eydgahi, A. (2020). The impact of social influence, technophobia, and perceived safety on autonomous vehicle technology adoption. *Periodica Polytechnica Transportation Engineering*, *48*(2), 133–142. https://doi.org/10.3311/PPtr.11332

Kupwade Patil, H., & Seshadri, R. (2014). Big Data security and privacy issues in healthcare. *2014 IEEE International Congress on Big Data*, 762–765. https://doi.org/10.1109/BigData.Congress.2014.112

La Torre, G., Esposito, A., Sciarra, I., & Chiappetta, M. (2019). Definition, symptoms and risk of techno-stress: A systematic review. *International Archives of Occupational and Environmental Health*, *92*(1), 13–35. https://doi.org/10.1007/s00420-018-1352-1

Latour, B. (1996). On actor-network theory: A few clarifications. *Soziale Welt*, *47*(4), 369–381. JSTOR.

Leung, M. K. K., Delong, A., Alipanahi, B., & Frey, B. J. (2016). Machine learning in genomic medicine: A review of computational problems and data sets. *Proceedings of the IEEE*, *104*(1), 176–197. https://doi.org/10.1109/JPROC.2015.2494198

Meuter, M. L., Ostrom, A. L., Bitner, M. J., & Roundtree, R. (2003). The influence of technology anxiety on consumer use and experiences with self-service technologies. *Journal of Business Research*, *56*(11), 899–906. https://doi.org/10.1016/S0148-2963(01)00276-4

Murdoch, J. (1998). The spaces of actor-network theory. *Geoforum*, *29*(4), 357–374. https://doi.org/10.1016/S0016-7185(98)00011-6

Niccol, A. (1997, October 24). *Gattaca*.

Nill, A., Laczniak, G., & Thistle, P. (2019). The use of genetic testing information in the insurance industry: An ethical and societal analysis of public policy options. *Journal of Business Ethics*, *156*(1), 105–121. https://doi.org/10.1007/s10551-017-3554-y

Rzeszótko, J., & Nguyen, S. H. (2012). Machine learning for traffic prediction. *Fundamenta Informaticae*, *119*(3–4), 407–420. https://doi.org/10.3233/FI-2012-745

Sonsare, P. M., & Gunavathi, C. (2019). Investigation of machine learning techniques on proteomics: A comprehensive survey. *Progress in Biophysics and Molecular Biology*, S0079610719301385. https://doi.org/10.1016/j.pbiomolbio.2019.09.004

Vanitha, C. N., Jeevaa, V. N., & Shriman, S. P. (2019). Image and face recognition using CV lens machine learning android application. *2019 5th International Conference on Advanced Computing Communication Systems (ICACCS)*, 972–975. https://doi.org/10.1109/ICACCS.2019.8728435