The Integration of Machine Learning and Artificial Intelligence In Healthcare—Opportunities and Challenges

Ashley Zapata Minero

DSCI 529: Security and Privacy

Abstract

I wrote this research paper to study the medical field's current infrastructure, understand the terminology, the influence of big data on the medical field, and current types of machine learning implementations. I evaluated the biases of these models and what could be done to prevent them from happening. I introduced essential concepts in the machine learning field, such as algorithms being used today in the medical. I also explored the integration of AI and ML systems into the medical field and its implications. Such implications were things like bias and how bias could harm the current policies, such as HIPAA. Also, on the ethical side, it can cause harm to patients' lives, and even a human constructing the system introduces their own bias. I also explored how the medical field has evolved into Medical Big Data. I propose a more holistic approach to better fairness and reduce bias within already existing frameworks, including techniques and human integration behind these applications.

The Integration of Machine Learning and Artificial Intelligence In Healthcare—Opportunities and Challenges

Years ago, if someone told you that you would see self-driving cars on the streets, would you have believed it? Even further, post-humanism is not that far-fetched and may become tangible if we compare the past to today. With the arrival of big data, many opportunities have emerged, but with those opportunities also come the challenges of dealing with all this identifiable information in the era of Big Data. Countless industries are joining in this Artificial Intelligence(AI) and Machine Learning(ML) surge as companies fight to be the next ones to advance in both ML and AI applications. For instance, one such case is NVIDIA beating Microsoft with their NVIDIA DGX-1 built on Tesla's P100 GPUs in 2024. Tesla's P100 GPUs exemplify how competitive and fast-paced the AI race has become.

While discussing the technological side of things, such as transportation, maybe one spectrum that we deem positive and exciting, using Machine Learning(ML) and Artificial Intelligence(AI) in the medical field raises red warning signs in every person's head, despite the potential benefits of AI and ML applications in the medical field, such as improved diagnostics and personalized treatments. These technologies can introduce significant privacy, data security, and ethical use risks. This paper argues that while ML and AI applications can significantly enhance healthcare for patients, their safe and fair integration depends on addressing the key issues of bias, privacy, and outdated regulatory frameworks. To better understand how these issues can occur, it is essential first to break down what we mean when discussing AI and ML in the healthcare context.

First, defining key terms in this intersection of medicine and statistical computation is essential, which encompasses machine learning. As Aldenderfer (2025) states, "Artificial

intelligence (AI) is the ability of a computer to perform tasks typically associated with living creatures." In particular, many subfields exist within AI, one being Machine Learning(ML). Another term Aldenderfer (2025) mentions is ML, "the study of algorithms that improve themselves by experience." Within ML, there are different types of processes in handling data and instructing the machine on what to do. These processes include supervised, unsupervised, reinforcement, semi-supervised, and deep-learning ML algorithms. Understanding these algorithm's framework is essential, especially when talking about the medical field.

By breaking down the fundamental ML algorithm types and how they actually function, we can understand how the medical field incorporates ML into its field. Supervised learning consists of x features(little x is an attribute) in an X feature vector (holding attributes such as age, height, and weight), and we have a y-target (the variable we are predicting). In contrast, in unsupervised learning, we only have an X feature vector and no y-target since we only look for unseen patterns in data and the relationships between inputs and outputs. According to Aldenderfer (2025), "Reinforcement Learning (used in ChatGPT and self-driving cars) involves itself in how software agents should take action in their environments, tying the output with a reward function, and the system tries to maximize this function." A third algorithm, as Chen(2025) explains, semi-supervised learning uses both supervised and unsupervised machine learning. Initially, it uses two datasets to train: a labeled dataset then an unlabeled dataset. It leverages unsupervised learning to its advantage for unlabeled data because there is usually a vast amount of it to classify. A fifth method, as Holdsworth & Scapicchio (2025) state, Deep Learning, is used in supervised learning but more commonly in unsupervised learning and relies on multilayered artificial neural networks, commonly referred to as deep neural networks, that try to emulate human neurons. These networks contain three or more layers, but in real-world

applications, they often consist of hundreds or thousands of layers to train highly complex models. These types of processes are directly applicable to the medical field.

These machine learning methods are not just theoretical; they are already being integrated into the daily operations of hospitals and clinics, helping to shape a new era of what many call 'Medical Big Data.' As we know, millions of people go to these facilities multiple times. As a result, millions of data points would be stored in these healthcare facilities' databases, awakening the term Medical Big Data. As a result, the medical field has adopted the Data Science definition of Big Data. In practice, Big Data is defined by 'three V's': volume (large amounts of data), velocity (high speed of access and analysis), and variety (variance across individuals and data types)(Price II & Cohen, p.37). In the medical field today, with the incorporation of these ML and AI applications, Mooney and Pejaver(2018) mention that the kinds of Big Data being collected on patients are "genomic/biological data sets (e.g., whole exome profiling), geospatial (e.g., neighborhood characteristics), electronic health records (e.g., Records of all patients with skin conditions), personal monitoring (e.g., Daily GPS records, Fitbit readings), and effluent data (e.g. Google search results, Reddit)" (p.96). This collection of patient data includes genealogical and non-genealogical data, such as things patients do outside of the scope of the medical facility, which raises concerns of perhaps infringements of the patient's privacy by aggregating this data using ML and AI frameworks. This kind of data aggregation hints at data management policies not being in place, allowing this to occur not just by accident but because the current policy enacted allows it.

When comparing the US privacy laws to the EU, it is generally recognized that the EU protects its citizens' data more than the US, which usually only cares about governmental abuse of that data and the entity in charge of the data(e.g., big business, data brokers). The EU

broadens its scope to those outside the EU who can access the data collection of EU residents/citizens. In terms of the EU in the handling of medical big data of its EU residents/citizens, as Price II and Cohen(2019) state, the EU puts no bounds on what classifies as health data, not caring about the format of the data is in, how it is collected, or who is in ownership of that data. The EU cares more about the data as long as it is related to a person's physical or mental health and health care services, as that data shall not reveal information about the EU person's health(p.38). In stark contrast to the EU, the US laws/policies enforced for data protection, as mentioned by Price II and Cohen(2019), tend to focus on physicians, health systems, and their business associates(p.38). In the United States, the law established for data privacy and protection is called the Health Insurance Portability and Accountability Act (HIPAA). The HIPAA is an outdated framework still being used in the medical field. While ML and AI algorithms are collecting this agglomerative collection of data, data leakage may occur within the frameworks of HIPAA. Perhaps even the ML and AI models themselves are taught to ignore these frameworks since they are not that strong in the first place to prevent all this agglomerative action from happening. HIPAA's policy itself may be an indicator of what data infringements can be done with Medical Big data.

Within the HIPAA law's outdated nature, many loopholes exist while claiming to protect data privacy. HIPAA's broad definition of identifiable information is one of its downfalls. As Price II and Cohen(2019) state, HIPAA's protected health information(PHI) is broad, including individually identifiable health information, and 'covered entities' (such as "health care providers, health insurance companies, and 'health clearing houses'") are prohibited from using or disclosing this data except in specific circumstances. HIPAA is also said to have an overprotective and underprotective spectrum. It is overprotective because it allows for the use of

PHI for healthcare treatment, which includes "'quality improvement' operations, payment, public health, and law enforcement." Still, it does not allow the use of PHI without an Institutional Review Board(IRB) waiver or patient authorization for research. It is under protective because HIPAA protects information security through a separate security rule(p.38). HIPAA's outdated nature and broad criteria for personal health information, along with its separate reign of security in a different place, make it a not-so-strong policy. We can see that because of HIPAA, all this health information and non-health information is being collected by ML and AI algorithms because many medical institutions can use these loopholes, such as stating 'quality of improvement' services while giving your consent, which perhaps you did not read, and they can steal all your non-health data using that loophole. A more plausible chain of actions to prevent this, since today we are not only concerned about Medical Big Data but also Big Data and AI and ML applications of this data, would be to establish laws governing AI and ML applications, such as a standard for acceptance, or specific number successful trials before establishing it into actual implementation. Perhaps never leaving just the algorithm, but having a person behind the scenes to make the final decision, especially when discussing ethics, which a machine lacks. Adding privacy laws that apply to AI and ML and successive creations that stem from both these disciplines can prevent medical big data and its aggregation of non-medical data from happening. Even people in this disciplinary field looking into sci-fi can prepare for this new era and learn how to deal with its implications. That said, HIPAA uses some methods to prevent privacy violations.

Even though HIPAA tries to prevent privacy violations through data de-identification, HIPAA's efforts fall short in today's medical big data-filled interconnected datasets. HIPAA uses de-identifying strategies to keep patients safe. According to Price II and Cohen(2019), one of

HIPAA's strategies for protecting patients' data from privacy violations while still allowing them to share their data is de-identifying it by removing 18 specified identifiers. These identifiers include names and email addresses. However, a privacy risk is de-identified data may become reidentifiable through data triangulation from other datasets. This is further emphasized because HIPAA regulates actors, not the data itself(p.39). This being said data triangulation can occur when training a model, multiple data sources are used, and although the dataset's respective column names may be removed, individuals can be identified through similar qualities. This poses a risk to the safety of patients because ML and AI algorithms use extra information that could allow them to correlate patients' features by aggregating these different datasets, which, although unnamed, could result in high correlations. Although unintentional, if this type of data triangulation were used for bad reasons, it could jeopardize someone's life, such as removing them from an opportunity to get hired if it was discovered they had a costly disease that companies did not want to be obligated to pay for. In these situations, adding noise to a dataset may help. Another possible solution could be reducing the dimensions of the dataset or even considering adding encryption to the data that only the person who owns the data and or the subject of the data has access to. On the technical side of things, processes such as a correlation matrix may help, where values close to 1 are highly correlated features, while those closer to 0 mean no correlation. Having someone oversee the correlation matrix results by checking them manually would greatly help. While the issue of the de-identification approach of HIPAA is a pressing issue, there is a more stagnant problem at hand with HIPAA.

The most stagnant problem with HIPAA is that the majority of health data today is not covered by HIPAA. HIPAA was made long ago before the era of big data came to fruition. HIPAA was enacted by Congress in 1966, but at this time, no one knew a digital revolution

would eventually arrive in the medical field. Back then, it was when all health data was recorded on health records, no one thought that data could be intangible like it is today. Since they associated medical data with physically documented data, Congress focused on healthcare providers and other covered entities. Hence, today, HIPAA does not cover large volumes of data unrelated to health. This is why there has been a surge of big companies, such as Google, Apple, and IBM, which can operate outside HIPAA's mandate and access all health data since HIPAA's reign ended after the introduction of big data. This insinuates that HIPAA's scope is no longer what it used to be, even more so when it comes to AI and ML, which have no clauses prohibiting them from doing so, which is why I mentioned my earlier solutions to this. I also see a necessary call to action of having a renovated HIPAA, as seen with California's move to update CPPA, with CPRA to be instated in this day and age. As an earlier example, HIPAA's implications for hiring procedures affected by health data do not stop at companies' non-prohibitions.

HIPAA's implications on patients' daily lives could be irreversible. The United States' established work policies are determinant of the impact that patients' data could have on their lives. According to Price II and Cohen(2019), in the United States, if sensitive data about patient information were to be released to employers or insurers, "such as a debilitating or expensive disease, they may wish not to employ or insure that person" for that reason. This is important since health insurance is typically tied to employment in the United States(p.39). Similar to an earlier example, this also shows that the implications of HIPAA could be irreversible because of its lack of adequate privacy implementations, which is due to its outdated nature. The company gaining access to that patient's medical data is not only an infringement on the privacy of that individual but also could forever restrict that individual from getting access to a job if that data were to be leaked, which is not a far-off prediction in this digital age. Some solutions to consider

may include creating policies specifically for AI and ML that use a data cycle, where, just like GDPR, only retains data for a limited amount of time; this policy, too, should do the same and supply an officer to make sure the policy is being followed. While these policy suggestions are good to keep in mind, some policies try to prevent this type of discrimination from happening, although minimally.

While non-discrimination policies in hiring and services aim to address specific issues, they also present notable drawbacks. According to Price II and Cohen(2019), in general, applying to all Americans, the Genetic Information Non-discrimination Act(GINA) prohibits some genetic discrimination. In the work field, there is also "the Americans with Disabilities Act(ADA), which prohibits employers from discriminating against work candidates" based solely on medical history or conditions that are disabilities. In the medical field, the Patient Protection and Affordable Care Act(PPACA) ensures no discrimination in healthcare services(p.39). It is also worth noting that ADA, GINA, and PPACA "limit consequences to accessing of data rather than protecting data themselves" (Price II and Cohen, p.41). These policies attempt to protect patients' rights to their ability to be hired or service rights, which is something everyone should have and is essential, but at the same time, they also situate the same problem HIPAA has, which is limiting the consequences of accessing data. This makes it redundant to set these policies in the first place if the consequences are lenient, and essentially, they are letting the data flow since little protection is established. A solution to this would be updating and enforcing business policies regarding these issues. A good role model for this type of enactment would be APRA, whose policies are strictly enforced. It is vital to have a situation we can simulate to see the gravity of the situation and the scope of integration of AI and ML in

the medical data at hand. A ‘distributional shift’ example shows how AI and ML are currently working in the medical field.

A real-world example of a ‘distributional shift’ occurred in an AI and ML application designed to monitor eye diseases. In the medical field, the influence of big data has created a medical terminology dictionary for the problems of interacting with these statistically optimized machines. As Challen et al. state, ‘distributional shift’ is an erroneous ‘out-of-sample’ prediction due to a mismatch between the data and the environment it was trained on. This erroneous ‘out-of-sample’ prediction could be due to bias in the training set, change over time, or a system used in a different population. This can also be influenced further by a mismatch in training and operational data or when a trained ML system is inappropriately used in an unanticipated patient context. This issue was particularly seen when De Fauw et al. discovered that their AI and ML incorporated system worked well on scans from an Optimal Coherence Tomography(OCT) machine(“non-invasive imaging tests that use light waves to take cross-section pictures of your retina” (Turbert, 2024)). However, this was not the case on another machine, which necessitated a process to normalize the data coming from each machine before a diagnostic prediction could be made(p.233, p.234). Having medical concepts to define these errors and measures of bias already seen in the medical field tells us that precautions are necessary with AI and ML integrations. If human eyes are not monitoring the accuracy of the model’s predictions, it could bring false positives, such as saying that someone has an eye disease, perhaps costing them thousands or millions of dollars. Even further, if this disease were costly and the data were to be leaked by the previous non-protective clauses, we have seen that it would be detrimental to this individual, like removing their chance of getting a job, because the US health coverage is related to hiring. Hence, data governance and ML and AI governance are needed to prevent this. To

make it even stronger, setting up an official to regulate the data's relevance to its application could prevent this bias. It was also good that De Fauw et al. were able to spot this before it could cause any further harm. Like 'distributional shift,' another similar terminology exists in the medical field's AI and ML dictionary of bias occurrences.

Bias is one of the most ethically concerning problems that raise warning signs in most people's heads when they hear about ML and AI implementations in the medical field because biases, like in statistics, skew the data, hence why it is essential to know these medical terminologies when it concerns life or death. As Challen et al.(2024) reported, 'insensitivity to impact' is where a system makes predictions that fail to consider the impact of false positives(incorrectly predicted as true) or false negatives(incorrectly predicted as false) predictions within the clinical context of use. Take, for example, although both humans and machines find it difficult to discriminate between benign tumors (which are noncancerous tumors, but as the cells multiply, they form tumors) and malignant melanocytic lesions(which are moles, but can also be cancerous ones), humans tend to go to the side of caution. However, ML and AI machines are only trained on the Y target result(benign or malignant) and cannot replicate a human's side of caution— that is based entirely on intuition. Furthermore, one solution that was offered by Challen et al. was to measure the cost of both potential missed diagnoses(false negatives) and overdiagnosis(false positives) that would, in effect, reduce the ML systems being optimized for the wrong tasks and apply it to more real-life scenarios (p.233). In hindsight, Challen et al. solution for 'insensitivity to impact' would be deemed efficient; however, it is necessary to understand that all this would be doing is making more computational statistic optimization and perhaps would not take into account all cases because, again in statistics, not everything is 100% accurate as seen by the errors we have viewed from AI and ML

incorporations and their inaccuracy presented. When putting lives at risk, it is better to have a statistician who monitors all of this statistical incorporation rather than just giving the machine some mathematical optimization and not supervising it, which may jeopardize someone's chance at life or death. In other words, bias will not be gone just with the integration of statistical optimization. Let's talk about another medical AI and ML bias terminology that is equally important to know.

Complex algorithms can be challenging for non-statisticians in the medical field to interpret. As Challen et al.(2024) indicated, 'black box decision making' is when "a system's predictions are not open to inspection or interpretation and can only be judged as correct based on the final outcome." An example of 'black box decision making' is an AI and ML system based on artificial neural networks(ANN), which would make it impossible to understand predictions for these algorithms, making it even harder to detect errors or biases found in the system. If we apply this to a real-life situation, take, for instance, a US Army-based image analysis system that was supposed to detect armored vehicles. In this AI and ML system, the machine had learned "to discriminate between images of sunny and cloudy days rather than to find the armored vehicles" that were situated, hidden in trees, versus the empty forests they were initially trying to test. This showed that the machine had introduced bias in the training set, which is another example of a 'distributional shift' mentioned earlier, because it was trying to be trained on a different situation, which did not result in the expected outcome, perhaps due to bias. Furthermore, this makes it 'black box decision making' since, in the end, it was only able to tell apart images from sunny vs cloudy days. In addition, a solution offered by Challen et al. for fixing 'black box decision making' is for AI and ML systems to produce 'saliency maps,' which show the most influential parts of a machine learning model. This system can, for example,

identify the areas of skin lesions or chest X-rays that most contributed to their prediction. An alternative they offered was a statistical analysis of the system's behavior by changing the inputs(p.233). The impact of 'black box decision making' is crucial, especially in this era of big data, and the respective ANN that these AI and ML systems apply, there are Feed Forward Neural Networks(FFNN), Convolutional Neural Networks(CNN), among others, it is imperative to choose the adequate one. Not only that, but the aspects to keep in mind are the initialized hyperparameters (the ones the programmer decides). These include the activation function(which statistical application you want to incorporate into the AI and ML system, such as ReLU, Sigmoid, Tanh, and LeakyReLU), the number of layers, and the number of neurons per layer. All these factors are essential when working with ANN. This gets even more complicated when there are thousands of layers, especially in medical data. I believe Challen et al.'s solution is only surface level. For this undetectable bias, what needs to be done is to study not just through a 'saliency map' but a thorough overview of the programming steps the machine took to reach its path. In that inference, Challen et al.'s statistical analysis could be one of the steps, but another suggestion would be a tool to visualize this in real-time. Such as the node layers and the data, and having programmers collaborate with statisticians could prove fruitful. Let us look at another AI and ML bias terminology.

AI and ML systems often encounter more failures than successes in various medical applications. These models hallucinate when not provided with adequate information. Per Challen et al.(2024), 'unsafe failure mode' is when a system produces a prediction without confidence in the prediction accuracy or insufficient information to make the prediction(p.234). An example would be if an individual of color were to be analyzed based only on data from a non-person of color; the model would have no confidence in making a prediction. In such a

scenario, Challen et al.(2024) suggested that a possible solution would be to incorporate a system that, if it sees that its confidence level is low, would use its fail-safe backup system/mechanism(that ensures no harm is caused to people) and refuse to make a prediction. This way, clinicians would know whether a system's prediction is trustworthy. They also suggested that the fail-safe mechanism should be applied if the system has insufficient input information or detects an 'out-of-sample' situation, like De Fauw et al. (p.234). In this instance, I think incorporating a fail-safe mechanism is very important, just as Challen et al. stated. However, I think the reliance on the system could be dangerous because what if the fail-safe system, which is also an AI and ML system, contains bias, then bad results would happen. Regardless, I feel that having a human technician at the backend, monitoring what is happening, would be more fruitful. Perhaps a fail-safe system that gives the technician a warning, and they could check out what it is and authorize the fail-safe mechanism, or not authorize it based on their analysis of what the machine presents and their knowledge of what is acceptable. Wholly relying on machines to automate these tasks at the point we are in ML and AI systems is not ethical in the medical field, so as of right now, human presence is better. Let us look at what exactly can make this 'safe-failure-mode' happen.

Data varies widely in form and source, often selected by individuals, which can introduce bias. Keeping this in mind when dealing with medical big data is crucial to address bias in the ML and AI systems. As Challen et al.(2024) state, the performance of ML algorithms depends on the model's training data composition and other parameters selected during training(hyperparameters). Other factors are the limited availability of high-quality data for training, correctly labeled with the outcome of interest, or when the data was collected as 'interesting cases' and is not representative of the normal(p.232, p.233). These factors mentioned

above are also very dependent when considering the approach of 'safe-failure mode,' hence, not only would applying a human for the oversight of these processes prove to be crucial, but also someone who particularly monitors to ensure the accurate implementation of these parameters is necessary. This would impede these biases from occurring at the human level. Currently, everything seems to be handled at the machine level. We should not only be looking at the machine level but also the human level. Hence, all factors should be considered when working with medical big data in ML and AI systems. Now, let us look at a more human, relatable topic in AI and ML systems.

'Fairness' is a societal ideal, and it has become a relatively new topic in the evolving discourse on ML and AI systems. 'Fairness' is especially important for medical data and the subgroups of populations in medical big data. As Price II and Cohen(2019) state, marginalized populations are missing tons of health and non-health data. With the incorporation of AI and ML systems having free rein of data, it is not unusual to see medical data having access to non-medical data. Hence, when it is missing things such as credit card use or internet history, it may lead to biases; when it comes to medical data, things such as genomic databases or EHRs are affected due to a lack of health insurance, the ability to access health care, and other reasons. The 'distributional shift' talked about earlier is a consequence of this lack of inclusion of big data, which is complex. Sometimes, it may favor those, but in other instances, it may disfavor those whose data is missing. For example, if a collection of data were occurring and showed a minority group responding less well to the study than other groups, and if the information for the minority population was not included, then it might lead to give the minority group more priority than if the data had been included, and vise versa. Price II and Cohen proposed a solution to provide better healthcare services for the underserved population. Another solution was applying

statistical adjustment for significant data gaps, which may help alleviate the problem somewhat(p.39). I believe that Price II and Cohen's proposed solutions are outstanding. Still, I would like to add another solution, which is incorporating AI and ML system libraries that try to promote 'fairness' by creating these libraries. Such libraries are like the ones seen in the scikit-learn fairlearn package, which aim to address fairness issues when training a model.

Scikit-learn's Fairlearn offers tools like fairness assessment and equalized odds to address ongoing fairness challenges in machine learning, which aim to identify and reduce bias across demographic groups in the training data. The fairness assessment is a process that evaluates the model's performance across different groups to see if there are any sensitive features(ethnicity, age, orientation) that could cause the data to skew. Holistically, it identifies any disparities and looks for biases found in the dataset, where output 1 indicates a potential disparity, and 0 means none. In equalized odds, which allows for identifying uses a model's true positive and false negative rates to ensure all are equal across all demographic groups, particularly the largest disparity between two groups, with an output of 0 indicating no disparity and 1 that there is disparity. However, just like Price II and Cohen's solutions, they will not entirely mitigate the issue of fairness that the authors of fairlearn define as a "socio-technical" issue. If more of these packages are made and implemented within the medical field, and we have a person supervising this process, perhaps this bias issue in ML and AI systems will be somewhat improved. Let us look at a more technical bias in the medical field.

Beyond healthcare policy and data management, programmers also have technical challenges in designing ML and AI systems. These technical challenges also raise critical issues, particularly as these systems are increasingly applied in medical contexts. Since we have seen the adoption of Data Science terms in the medical field, it is not strange that they have

recognized the ‘curse of dimensionality’ in implementing ML and AI systems. However, data scientists and programmers who make these systems are primarily concerned with this issue. According to Lomsadze (2023), the ‘curse of dimensionality’ states that “as the number of features(dimensions) grows, the amount of data needed to generalize accurately grows exponentially.” Per Aldenderfer (2025), what this means in practice is that when working with datasets that have large amounts of features, it is in the best interest of programmers to reduce the number of dimensions, which can be done by removing features or by transforming(normalizing with unsupervised ML, and standardizing with supervised ML). I agree with this method of dealing with the curse of dimensionality. Still, I also recognize that in the medical field, with the use of medical big data, perhaps this is a challenge many programmers must deal with, especially with the ML and AI systems aggregating even non-medical data, especially when dealing with wide data sets that must be converted to tall datasets. According to Mooney and Pejaver(2018), “medical data requires tall datasets and small and/or biased training sets can lead to overfitting, which limits the problem current machine-learning methods can address” (p.105). Perhaps this reference to the medical big data needing to be converted and how even working with small or biased datasets can lead to overfitting(gives the model too much data to train on, becoming strict with criteria it makes decisions on) – raises an issue with applying generalized machine learning models such as decision trees, K Nearest Neighbors, Linear Regression, and other models to medical big data applications. Instead of using these non-medical-specific models, one solution would be to create more medical field-specific models, which can also help find better results. Another solution to this would be assembling statisticians and medical specialists to collaboratively find a middle ground to better enhance the ML and AI field with medical-specific types of models. Let us

further explore what medics currently wish to see with this new hybrid of the medical field and statistical computation, which is AI and ML in the medical field.

As ML and AI technologies evolve daily, the medical field continues to explore future collaborations of health care with AI and ML systems, with their aspirations rooted in both promise and practicality for future use. As Mooney and Pejaver(2017) state in their article, in the future, many more applications involving reinforcement learning and deep neural networks are the long-term goal of the medical field, such as those also being used in natural language processing or image classification, this is in the foreseeable future where computations cost get lower than they currently are(p.104). Based on what is known about the current results and the rise of terminology, it is good that there is a future vision of what this new age of big medical data with AI and ML systems will bring. Furthermore, we have seen some "successful" implementations incorporating reinforcement learning types. Based on Challen et al.(2024) findings, "a sepsis system has successfully been tested in different contexts of the community hospital despite being trained on intensive care, a potential distributional shift, and thus shows some capability of adaptation through 'transfer learning'" (p.233). This shows potential growth in the medical field and how, in the future, it may be possible to fully implement these systems while still having full awareness of their present dangers. It may not be bad to have a system where people can opt in or out of being a system subject and have a human counterpart option since it is vital to cater to every individual. This brings us to the ending remark.

In conclusion, ML and AI implications are not the only ones to be concerned but how, in the future, ML and AI implications will be impacted by the United States' outdated policies in health care, causing privacy implications associated with the key issues found in models themselves which are bias, and also themselves can cause privacy implications. I offered many

solutions and built upon existing ones for a more rounded approach. I talked about key terms of Artificial Intelligence, Machine learning, and the subsets of tools in Machine Learning, such as supervised learning, unsupervised learning, and deep learning. I talked about collecting medical and non-medical data, medical big data, and its implications. I compared the EU's medical big data handling to the United States' HIPAA policies. I expanded on HIPAA, talking about its outdated nature and what constitutes Personal Health Information, and I offered solutions to HIPAA. I spoke about one of HIPAA's de-identification strategies and its implications. I spoke of HIPAA's implications to not only health data but also other data that big companies can use, due to HIPAA being outdated. I discussed the implications of leaked patient health data concerning the United States' established work policies. I spoke about ADA, GINA, and PPACA, which help Americans in different sectors maintain their rights to non-discrimination through health data. However, it does not necessarily function as a good policy, as it limits consequences. I spoke about different definitions incorporated in the medical field, such as 'distributional shift,' 'insensitivity to impact,' 'black box decision making,' and 'unsafe failure mode,' giving examples and solutions to the bias introduced to them. I also spoke about other reasons bias is introduced into ML and AI systems, such as the limited ability of high-quality information or 'interesting cases' as training data for ML and AI models. I also spoke about the concept of 'fairness' in the ML and AI field, which should be essential when lots of health data are not fully representative of all populations, often missing underprivileged populations, which introduces bias. I talked about bias on a human margin of error, which should also be considered when working with ML and AI models. Finally, I discussed professionals' aspirations for ML and AI to be further integrated into the medical field. In my response to all this going on in the medical field, I think many efforts should be made, and it should be collaborative between everyone,

because the medical field in AI and ML will keep increasing and incorporate into many people's visits to the hospital. Even when working with an AI and ML machine, it is imperative to be aware of human inconsistencies. May AI and ML endeavors be fruitful while always prioritizing human lives.

References

Aldenderfer, K. (2025, March 12). *Intro to Machine Learning | ML: Linear Regression* [Lecture notes]. University of Southern California.

Aldenderfer, K. (2025, February 19). *ML: Methods of Analysis 1* [Lecture notes]. University of Southern California.

Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019, March 1). Artificial Intelligence, Bias and clinical safety. BMJ Quality & Safety. https://qualitysafety.bmj.com/content/28/3/231.full

Chen, M. (2024, October 24). Lots of unlabeled data? consider semi-supervised learning for your AI project. Semi-Supervised Learning Explained | Oracle Suomi. https://www.oracle.com/fi/artificial-intelligence/machine-learning/semi-supervised-learning/#:~:text=Semi%2Dsupervised%20learning%20uses%20a,with%20an%20unlabeled%20data%20set.

Grzybowski, A., Brona, P., Lim, G., Ruamviboonsuk, P., Tan, G. S. W., Abramoff, M., & Ting, D. S. W. (2019, September 5). Artificial Intelligence for Diabetic retinopathy screening: A Review. Eye (London, England). https://pmc.ncbi.nlm.nih.gov/articles/PMC7055592/

Holdsworth, J., & Scapicchio, M. (2024, June 17). *What is deep learning?*. IBM. https://www.ibm.com/think/topics/deep-learning

Lomsadze, S. S. (2023, March 12). *Common data science interview Q&A*. Medium. https://salomelomsadze.medium.com/common-data-science-interview-q-a-806df6717630

Mooney, S. J., & Pejaver, V. (2018, April 1). Big Data in public health: Terminology, Machine Learning, and privacy. Annual Review of Public Health.

https://www.annualreviews.org/content/journals/10.1146/annurev-publhealth-040617-014208

Price, W. N., & Cohen, I. G. (2019, January 7). Privacy in the age of Medical Big Data. Nature News. https://www.nature.com/articles/s41591-018-0272-7

Turbert, D. (2024, September 26). *What is optical coherence tomography?*. American Academy of Ophthalmology. https://www.aao.org/eye-health/treatments/what-is-optical-coherence-tomography