**LIMITATIONS IN AI BASED DISEASE PREDICTION SYSTEMS**

**Jainam Shah 2019A7PS0096P**

**Tanmay Parab 2019A7PS0044P**

**Aneesh Kabra 2021A7PS0442P**

* **Introduction**

The recent advancements in technology and Artificial Intelligence as well as the introduction of AI’s research in the field of medicine has shown people the excellent prospects of the usage of AI in healthcare. Deep learning has shown great potential in the field of disease prediction and drug response prediction. With development in technology, our learning models applied in the field of medicine have also undergone the same progress and this has been seen by the continuous improvement in the accuracy of such models in medical disease prediction as well as the overall performance in all aspects.

However, this progress has some impeditions that have slowly emerged with time. Concerns such as privacy of patients, limited data availability for rare diseases, stability and explainability of the model have become major obstacles that have prevented this technology from becoming widespread among various medical institutions. With the recent development in XAI(Explainable Artificial Intelligence) predictive techniques there is a solution for explainability but the other limitations are still mostly unresolved.

Below are three significant limitations and possible solutions to these problems.

* **Limitations**

1. Stability and edge use-cases:

The popular machine learning models, such as Logistic Regression, Decision Trees, Random Forests, SVMs, and more complicated models, are heavily restricted to generalized use cases and work efficiently only for specific diseases the model has been trained on. However, in practical cases, the patients might have various symptoms and multiple diseases which can only be currently diagnosed by doctors. The models and methods given in the research papers are very limited. Also, even though the models are quite helpful and have boosted their accuracy over the years, diagnosing diseases is a subject with little room for error. It may cost lives by registering incorrect treatment. For those diseases which are rare and uncharacteristic, an accuracy of even 95% in prediction is not enough. For example, for Alkaptonuria, or “black urine disease,” a one-in-a-million genetic disorder, what if the model fails to recognize correctly in a positive patient? This is where the precision and recall statistics of the model being used come into question. There also exist other cases where the machine learning approach comes up short. Doctors have a contextual understanding of the patient's life and medical history, which is crucial in providing an accurate diagnosis. They can consider the patient's medical history, family history, lifestyle habits, and environmental factors, which may not be apparent to an AI model. Also, doctors can ask patients specific questions to gather more information about their symptoms and medical history. This can help them to rule out potential causes and make a more accurate diagnosis. They can adapt their diagnosis based on new information that arises during the examination. They can also adjust their approach based on the patient's reaction and feedback, which is currently in effect for an AI model.

1. Limited Data Availability:

Medical information is very private, which imposes a huge challenge in collecting such data. However, such data is essential to train our models to perform better and improve their accuracy. This greatly restricts the quantity of data one can assemble to improve our learning models. This, in turn, affects the precision and accuracy of such models.

Healthcare data is rarely stored in one place, and for using it in such situations, accessing and gathering data from multiple sources is often very challenging. This results in making AI models much more difficult because we do not have enough data to train our model in the first place.

Limited data availability is a common challenge in developing AI models for disease prediction, particularly for rare diseases or underrepresented populations. For instance, it's possible that there aren't enough patients with uncommon disorders to produce adequate data.

1. Privacy:

Medical histories and genetic information are some things everyone wishes to keep private. Using such information obviously concerns everyone because information like this is very sensitive. This makes it the top priority of the AI to use such information to keep information like this private.

The fact that our models contain private data also makes them a huge target for theft, breaches, or exploitations. This poses a major concern with using AI in healthcare. To gain information from different sources and improve our learning models, there often are exchanges of such data between various healthcare providers. This also increases the dangers of data privacy breaches.

Our current models are effective and have shown great progress from the start. However, in cases of medical diagnoses, one cannot afford to be even 95% accurate because of the huge stakes on the line. This poses another problem highlighting the possibility of inaccurate diagnosis or treatments. If such problems are not solved immediately, it raises the possibility of a breach of the privacy or safety of a patient.

* **Solutions & Proposed Methodology**

1. Stability & Edge Use Cases:

AI-based models may be enhanced in a number of ways to support medical professionals' diagnostic procedures better. One way is to enhance their natural language processing skills. Understanding the context of the patient's symptoms and their medical history can be done better by the AI models if their natural language processing abilities are improved. Through this, the model will be able to pose more pertinent and instructive queries. This will further result in the model being able to produce much better and more precise diagnoses. Creating the AI models in such a way that they also include patient feedback can further increase the precision of diagnosis.

Using symptom information and the patient's treatment response, our AI model can provide better advice for individual therapy. However, before we can do all this, it is essential that they are properly trained first. This also includes specialized training and testing with medical specialists like physicians and nurses. This ensures that the medical practitioners' needs are looked after and also ensures that our models are a useful addition to the diagnosis process.

Multi-label classification:

Multi-label classification is a type of machine-learning task where an algorithm is trained to assign multiple labels or categories to a given input. In other words, the algorithm needs to predict more than one label for each instance of data it encounters. Multi-label classification is commonly used in applications such as text classification, where a given document may be relevant to multiple categories, or in bioinformatics, where a given gene or protein may be associated with multiple functions or diseases.

Using machine learning algorithms such as decision trees or deep learning algorithms such as neural networks, we can train our multi-label classification algorithms. Depending on the specific purpose and the data complexity, we can choose a specific algorithm.

Methodology:

Let us consider the “Application of Machine Learning in Disease Prediction” paper authored by Pahulpreet Singh Kohli. He talks about the prediction accuracy of different classification algorithms for diabetes and breast cancer. For these two diseases, he chooses the SVM classification model but we need to create a model considering the possibility of the patient having both these diseases. Multi-label classification can be applied to SVM-based models for predicting diabetes and breast cancer by modifying the models to output multiple labels or categories for each input.

For example, in the case of diabetes prediction, the SVM-based model could be modified to output multiple labels indicating the presence or absence of different risk factors associated with diabetes. These risk factors could include age, body mass index, blood pressure, family history, and glucose levels. We can have 2 kinds of outputs. One kind, let us say "1" can be used to indicate that a risk factor may be present. The other output, which can be "0" can be used to show that there is no risk factor present. Similarly, in the case of breast cancer prediction, the SVM-based model could be modified to output multiple labels indicating the presence or absence of different types of breast cancer and their stages.

We need a dataset constructed in such a manner that includes multiple categories for each type of input. This is required to train the multi-label SVM models. One way to do this is by using a clustering algorithm to group similar kinds of inputs together and assigning labels to the groups. Once the work of training the SVM models is done, we can predict the presence or absence of multiple risk-factors or different types of breast cancer for a given input. This provides us with a lot more information than traditional binary classifiers. With this kind of information available to us, we can start developing more targeted treatment plans and interventions for patients.

Neural Networks:

A form of machine learning model called a neural network is motivated by the structure and operation of the human brain. They use layers of interconnected nodes, or "neurons," that collaborate to process and analyze data in order to identify patterns and links in the data. In a neural network, each neuron takes input data and processes it using a mathematical function to generate an output. The following layer of neurons receives these outputs and uses them to carry out their own calculations and generate further outputs. This process keeps on until the last layer of neurons generates the final prediction or output of the model.

Methodology:

Neural network machine learning models can be used to incorporate patient feedback in AI-based disease prediction systems. One way to incorporate patient feedback is to use a type of neural network known as a recurrent neural network (RNN), which is designed to process sequences of inputs over time. In the context of disease prediction, patient feedback could be treated as a sequence of inputs, with each input corresponding to a different time point or stage in the disease.

In a system, where we are trying to predict the possibility of diabetes in a patient, feedback from them could include their glucose levels at varied times throughout the day. Diet and exercise habits could be some other feedback given to the AI. This can then be used in an RNN model, which would use such inputs and update its prediction accordingly.

Another possibility is to use another type of a neural network known as Deep Belief Network (DBN) for patient feedback which models complex patterns in high-dimensional data. We can take feedback as a high-dimensional input, with a different type of feedback for each kind of input. Certain examples for feedback include glucose measurements, blood pressure readings, and family history.

1. Limited Data Availability;

Data augmentation:

Data augmentation is a technique where we create data and expand our dataset using the existing data we already have. The goal is to increase the diversity of our dataset which can also improve the generalization ability of our machine learning model. By creating more data, machine learning algorithms can learn more effectively, and this can help to mitigate the effects of limited data.

Methodology:

For this, we can use GANs- Generative Adversarial Networks. Such networks have two neural networks to create synthetic but valid data. One network works on creating such data, while the other attempts to differentiate between the real data we have and the synthetic data created. If we ensure that the statistical properties of our original data are captured, we can obtain highly detailed, realistic synthetic data.

Transfer Learning:

Transfer learning entails using models that have already been trained for one task to another. Healthcare companies may use transfer learning to improve the performance of their own models even with little data by employing pre-existing models developed on comparable data.

From the pre-trained model, we take the weights and architecture of it. This in turn is used as the foundation for our new model. This new model is then retrained on a new dataset for a new task.

This greatly minimizes the quantity of data required to train the new model while also increasing the model's precision and accuracy. Since we don't have a lot of data in this field, transfer learning is really helpful. It also cuts down on the time and computing resources required to train a new model from scratch.

Active learning:

The most useful data samples are iteratively chosen for the model to learn from in active learning, a machine learning approach. A tiny labeled dataset is used to train the model in the beginning using active learning. The process then repeatedly chooses fresh samples from a pool of unlabeled data for expert labeling. The training dataset is then enlarged to include the labeled samples, and the model is retrained. Once the model's performance is sufficient or there is no more data available for labeling, the process is repeated repeatedly.

By using the most informative samples for labeling, we end up reducing the number of labeled samples actually required to obtain a high accuracy and such is the goal of active learning. This in turn also minimizes the overhead cost of labeling large amounts of data.