HIGH SCHOOL EDITION
Journal of Student Research

Volume 13 Issue 1 (2024)

Analysis of Artificial Neural Networks on Predicting Dementia

Tanish Potula[1] and Ian Custar[#]

[1]Innovation Academy, USA
[#]Advisor

ABSTRACT

Dementia, a chronic neurological disorder characterized by memory loss and impaired cognitive ability, has affected tens of millions of people globally and projections estimate this number to rise in the coming decade. Despite its widespread impact, dementia remains one of many diseases with no true cure or definitive method of diagnosis. This study aims to explore the question of how artificial neural networks can be used to predict dementia. Analyzing current research on neural networks for disease diagnosis highlights a notable gap, known as the "black box conundrum," indicating the challenge of tracing steps leading to a given output. Another gap in the literature is the absence of comparative data on factors influencing successful neural network algorithms. Conducting a comprehensive cross-continental comparative analysis across North America, Asia, and Europe, and utilizing an experimental design with the UCI ML Repository's Parkinson's Database, this study specifically aims to fill the identified gaps. The findings reveal key elements for successful neural networks: a data sample exceeding 100, clinically intelligible data values, and a backpropagation model. In the experimental phase, a comparison between a backpropagation model and logistic regression, paired with Shapley value analysis, elucidated a 94.87% accuracy for the former, surpassing the latter's 92.3%. These findings extend beyond dementia, providing valuable insights for neural network applications in all healthcare settings. Future research endeavors should focus on refining and optimizing neural networks with the ultimate goal of achieving 100% accuracy and maximized efficiency.

Introduction

Dementia, a chronic neurological disorder that commonly results in memory loss and impaired cognitive ability, remains one of the most prominent and mysterious diseases. Currently, "more than 7 million people ages 65 or older had dementia in 2020" and this number is projected to grow to "more than 9 million Americans" by 2030 and 12 million by 2040 (Population Reference Bureau). Despite these large numbers, a cure for the disease remains undiscovered due to the disease's complex biological nature. In addition, these unknown characteristics of "the disease's underlying biology" not only have prevented a true cure for the disease but have also led to many difficulties in diagnosis (Gotz, 2018). The future of dementia research or healthcare research in general is influenced heavily by technology or more specifically artificial intelligence (AI). Over the past decade, the debate around technology and machine learning (ML), a form of AI characterized by its ability to process and analyze data by learning and adapting without human interference, has grown exponentially, and with this growth has come many new innovations in machine learning that have great potential to improve the healthcare field. One specific instance of ML that will be discussed in this paper is artificial neural networks (ANNs), which as defined by Hardesty in an article for the Massachusetts Institute of Technology as "a means of doing machine learning, in which a computer learns to perform some tasks by analyzing training examples" (Hardesty, 2017). In essence, their structure mimics one of the human brains by consisting of multiple complex layers of interconnected notes. In addition, similarly to the brain, everything must be aligned in a structured order in order for the network to function, as shown by the fact that Michael Meighu, an IT director in the field of AI and ML discusses how even one component being out of line could cause the neural network to have difficulty in processing the patterns (Siwicki, 2022). To answer the question of how artificial neural networks can be used to predict dementia it is necessary to understand the history of these topics.

History of Dementia

Dementia was originally discovered in the late 19th century and was followed by the discovery of Alzheimer's Disease (AD), the most common form of dementia. At this time, dementia used to be categorized into pre-senile and senile which distinguished between dementia before the age of 65 and dementia after the age of 65, but in recent decades, this ideology has changed due to the fact that post-65 dementia cases are "much more common", showing a correlation between aging and the disease (Ballenger, 2017). Unfortunately, dementia cases have continued to rise in the recent decades which makes the necessity for new innovations much more important.

History of ML/ANN

While dementia was first discovered in the late 19th century, AI was originally discovered in the mid-20th century, but it "originally found its way into healthcare" in the early 1970s with its use in biomedical applications (Quest et al., 2020). About two decades after AI's original discovery came the discovery of artificial neural networks which despite their discovery in 1980, became common and popular as recently as 2012 when the creation of AlexNet showed that neural networks were far superior to existing AI algorithms. Despite ANNs superior capabilities, hesitations still remain in regard to its applications to medicine and patient care.

Literature Review

Machine Learning in Healthcare

While machine learning has become more common in healthcare in recent years, its application in the field is lagging due to a few barriers to entry. This can be seen through the fact that only 1 in every 1280 jobs in healthcare requires AI-skills (Goldfarb & Teodoridis, 2022). Avi Goldfarb, Ph.D. and Rotman Chair in AI and Healthcare at the University of Toronto and Florenta Teodoridis, who has a Ph.D. is a Professor of Management and Organization, analyze these barriers of entry and came to the conclusion that four main factors are prohibiting ML's use in healthcare: algorithmic limitations, regulations, data access, and misaligned incentives. The most essential component of a neural network or any machine learning algorithm is the data. In order for the machine to provide accurate and efficient results, the data set must have a large sample set of quality data. However, as Goldfarb & Teodoridis (2022) mention, the issue in healthcare is that most of the data is either "difficult to access" or "is often incomplete". In addition, most current neural network models contain algorithmic limitations in regard to the black-box conundrum. In essence, this conundrum occurs when "it is difficult to understand how a specific prediction was generated" (Goldfarb & Teodoridis, 2022). This interpretation issue presents a major liability for hospitals. Therefore, these limitations have perpetuated further research on specific metrics of neural networks known as "partial dependency plots, Shapley values, and LIME Analysis" which improve the algorithm's clinical intelligibleness and understanding in order to minimize these liabilities and make the technology more accessible into the healthcare field (Quest et al., 2020). Despite the algorithmic and data limitations that come with ML, the overarching limiting factor comes from the perception and trust of AI. As studied by PixelPlex, a well-known IT consultant company that has worked with Fortune

500 companies in the field of artificial intelligence, many people "fear the idea that computer algorithms will be taking decisions about their treatment" while others even fear that these algorithms will replace their jobs (Belova, 2020). It is not a replacement for people, "like machines on an assembly line" as although AI is independent and adapts on its own, it still requires human intervention to manage errors, faults, and also to check the program's results in order to create a more accurate and precise system of patient care (Quest et al., 2020). Therefore, although AI and ML are currently limited in their use in healthcare applications, with more research backing which will lead to improved trust amongst the public, these algorithms have the potential to shape the future of medicine in a positive way.

Diagnosis of Dementia

One of the most prominent areas in which ML can shape the future of society comes in its possibility to improve the diagnosis system of dementia. The diagnosis of dementia is a multi-step process that is very complicated due to the disease's nature. Dementia is a "progressive, insidious disorder" which means that symptoms at first may seem minimal and hard to distinguish but become more and more prominent with time (Buntinx et al., 2011). This theory of progression is echoed by Jurgen Gotz, a director for aging dementia research, who claims that symptoms of dementia or Alzheimer's in specific develop over time at different rates for different people which means that diagnosis of dementia often occurs years after an individual has already experienced brain damage" (Gotz, 2018). Buntinx characterizes the process of diagnosing dementia as a four-step process with the first phase being the trigger phase or the phase where dementia is suspected, then this is followed by the disease diagnosis which requires the use of medical imaging, and then following this is the care diagnosis which in essence is the determination of what type of treatment will be needed. The last part of the process is the carer assessment which is when doctors must determine if the patient has people around him who can take care of them and if these people require any assistance. Often this last step is overlooked when analyzing dementia despite the fact that the carer of the patient may be put through tremendous amounts of stress which can lead to health issues for themselves. The analysis of dementia diagnosis can be understood from many perspectives as the analysis by Gotz focuses more prominently on the treatment aspect of dementia diagnosis while the analysis of the study by Buntinx focuses more prominently on the aspect of treatment planning. By comparing the two separate perceptions about the current state of dementia diagnosis, it can be seen that the common conclusion is that the process is time-consuming, inefficient, and often ineffective.

Medical Imaging

Despite the inefficiencies in current treatment options, some important components of the process remain necessary even with the possibility of new technology. For instance, medical imaging is a key component in ML as it can provide the necessary data for the algorithm to process and produce an output diagnosis. The common forms of medical imaging that take place in the diagnosis of dementia are X-rays, MRIs, CT scans, PET and SPECT scans. All are similar in the fact that they are able to gather images about the injured/affected area but they each have some differences as shown by the comparative analysis conducted by Kasban et al. (2015). As Kasban et al. (2015) outlines, CT scans and MRI scans are similar in regard to the fact that they are used for maximizing spatial resolution in the manner of X-rays and magnetic fields, respectively. On the other hand, PET Scans and SPECT scans examine the metabolism of the disease through the use of positron emissions and gamma rays, respectively. In essence, each of the tests serves their own purpose in diagnosing a wide variety of medical diseases, but each have limitations when it comes to diagnosing complex diseases such as dementia. While currently in the status quo, these advanced forms of medical imaging are unable to pinpoint dementia diagnosis in an effective and accurate manner, when the resulting data from these tests are processed through complex neural network algorithms the possibilities for the future of diagnosis are endless. Despite this, there still remain some issues with neural networks that have yet to be addressed fully. Through this analysis of the current situation of machine learning in healthcare and the current state of dementia diagnosis, a gap that can be identified in the literature is that current ML algorithms that pertain to dementia and other medical diseases have difficulties meeting the criteria of being efficient and accurate, while also being able to be easily interpreted and implementable into real-world situations. This is due to the fact that they lack the ability to retrace their steps in regard to how the output, or the diagnosis, was produced. The second gap in the literature that can be seen is that no real comparative data has been collected on the factors that lead to the greatest success for ML algorithms.

Methodology

The goal of this subsection of the research is to answer the question of "How can we compare various neural networks to determine the most important metrics and how can we use this comparison to create an accurate and applicable neural network for the purpose of predicting dementia status?". In order to achieve this research goal, the research was conducted through the use of two distinctive methodologies: comparative analysis and experimental design.

Comparative Analysis

A comparative analysis can be defined as a "method of determining and quantifying connections between two or more variables by monitoring distinct groups" (Mollah, 2021). For the purpose of the comparative analysis, I will be analyzing the research surrounding neural networks in three continents: North America, Asia, and Europe. The purpose of diversifying the sources across these continents was to ensure that any conclusions made from the analysis were respective of the global population rather than simply the American population. These continents, in particular, were chosen as the focus of this analysis due to their superiority in AI research. As stated by the Harvard Business Review which analyzed a study from Tsinghua University, the country with the greatest number of annual AI research publications in 2017 was China with a little under 40,000 papers followed by the U.S, India, and the UK who had approximately 24,000, 11,000 and 7500, respectively. In addition, "the number of Chinese AI firms has reached 1189" which is second to the U.S. which has 2000 AI Firms (Li et al. 2021). The reason for the three continents of choice can be further justified through the analysis of the AI Readiness Index by the Oxford Insights and International Development Research Center which ranks countries based on their "preparedness to implement AI into public service" ("Government AI", 2022). The top 10 countries are the U.S., Singapore, U.K., Finland, Netherlands, Sweden, Canada, Germany, Denmark and Korea. Of these 10, 2 are from North America, 2 are from Asia, and the remaining 6 are from Europe. Through this data, it can be concluded that North America, Asia, and Europe are by far the three continents with the greatest AI capabilities and therefore will provide the most relevant information. This method of choice aligns with the overall research question as a study of the current research will provide a greater understanding of how neural networks can be used to predict different diseases, dementia in particular. In addition, the information gathered from this analysis will provide valuable information about the accuracy of neural network models and how they differ based on different programming and data factors. The steps I will take to complete the comparative analysis is to initially gather 5 academic scholarly sources for each of the three continents culminating to a total of 15 sources. After these 15 sources are gathered and read, they will be analyzed in a qualitative and quantitative manner through the use of the machine learning checklist written by Scott et al. (2021). The Scott et al. (2021) checklist contains 10 parameters/questions that should be met/asked in order to ensure an ML algorithm is fit for healthcare applications. After this data is gathered from the existing studies through the comparative analysis method, the data will be compared across the 15 sources to determine nuanced differences in the neural networks and how they affected the overall performance of the algorithm. One possible limitation to this research study is the fact that the broad category of ML means that there are many ways to achieve the same feat (prediction of disease) and therefore comparing two studies with different programming techniques could yield results that don't provide much help to the overall research. This data analysis will lead to the experimental design methodology.

Experimental Design

The second methodology I will be using is the experimental design methodology in which I will be creating a neural network for the purpose of predicting dementia in individuals. An experimental design is defined as a controlled research experiment that ensures that "precision is maximized, and specific conclusions can be drawn regarding a hypothesis statement" (Bell, 2009). In the case of my experiment, the controlled conditions will be the parameters for the code of the neural network, and I will draw conclusions to support or deny the hypothesis that a neural network can predict dementia in individuals over the age of 65 with at least an 85% accuracy. This methodology of choice aligns directly with the overall research question as it will help determine the accuracy of an artificial neural network in predicting dementia. This research methodology can be justified through the scholarly works of Chatterjee et al. (2011) and Kim & Lim (2021). Chatterjee et al. (2011) uses a similar methodology of creating a neural network but rather does it for the purpose of diagnosing neonatal disease. In the end, the study concluded that the "overall predictive accuracy required was 75%" after the model was tested (Chatterjee et al., 2011). In regard to the Kim & Lim (2021), the authors similarly create a deep neural network for the purpose of predicting dementia. At the end of their research, they were able to conclude that the optimal predictive accuracy was 81.9%. Despite the difference in what disease was being predicted, they both were able to gather data from the neural network that furthered the conversation around machine learning in healthcare, therefore justifying the use of this methodology. Although Kim & Lim (2021) had a similar goal of predicting dementia, the algorithm in this experiment will differ in the way it predicts dementia and the way it was coded.

Data Set

The first step in the process is to locate a database of deidentified dementia data that contains patient data along with their dementia status. For this specific experiment, the data will be collected from the UCI Machine Learning Repository's Parkinson's Database. This database consists of approximately 6 audio recordings from each of the 31 participants yielding a total of 195 different samples of data. Of the 31 individuals, 23 had Parkinson's while the remaining 8 did not. The data set consisted of 23 attributes, each correlating to a different metric that helped in diagnosing if the patient had dementia or not. The complete description of these attributes and their relevance is listed in Appendix A.

Code

The program will be written in Python and will be coded using sickit-learn database and the Jupyter Notebook as the IDE. Other libraries that will be used include Pandas and NumPy. After the completion of the comparative analysis, the most effective neural network model will be concluded and will be used as the model of choice for the experimental design. Then, after the neural network is coded the data will be analyzed in conjunction with the comparative analysis data analysis to determine if the neural network was better suited for medical application than currently existing models. Conclusions will be made on if the gap in the research was filled along with if the hypothesis was accepted. One possible limitation to the findings gathered from the study could be that a small sample size of the database could result in inaccurate and irreplicable results. Despite this, the methods of analysis align with the research question as it will lead to a conclusion about what factors impact artificial neural networks and the accuracy at which neural networks can predict dementia, in turn answering the research question.

Findings/Results

Comparative Analysis

At the completion of the comparative analysis methodology, three distinct data tables were produced with each one representing one of the three continents that were used in the analysis. In each analysis, four different features in particular were examined: Data Sample, Type of Model, Performance, and Clinical Intelligibility. Data sample represents the population size of the data, which was used to train the algorithm, type of model refers to the name of the ML algorithm itself, performance refers to statistical metrics such as accuracy, AUC, Precision, Recall, etc, and lastly clinically intelligibility refers to if the program consisted of an algorithm which represented which factors were most important in the decision-making process. These four factors in particular were narrowed down as the most important for the purpose of the analysis as they directly correlate to the applicability of a ML algorithm into a real-world healthcare setting and are the only ones from the Scott et al. (2021) checklist which are objective. From the data gathered it can be seen that the majority of artificial neural networks follow a certain model known as backpropagation. Furthermore, on average it can be seen that most currently existing neural networks have an accuracy in the 70-90% range and consist of a data sample of at least 100 people. Lastly, one important trend that can be noted from the aspect of clinical incelebrity is the fact that a majority of the North American and Asian neural network designs were clinically intelligible while a majority of European ones were not. This trend aligns with the fact that North America and Asia are more developed in AI research and have more existing publications and developments, emphasizing the importance of clinically intelligible data in regard to avoiding black boxes. The information gathered from the three data tables was used to determine how each of the factors impacted the performance of the data and its applicability and was used to determine the most optimal way to code a dementia-based neural network.

Table 1. Neural Network Analysis North America

North America	Data Sample Population	Type of ML model	Performance/ Measure of Interpretability	Clinically Intelligible?
Tomov & Tomov (2018)	303	Backpropagation DNN	99% Accuracy MCC = 0.98 0.99 Precision 1.0 Recall 0.99 F1	Yes
Ercal et al. (1994)	326	Feedforward Backpropagation	Max Accuracy: 95.6	No
Qiu et al. (2022)	4550	CNN imaging model	0.6 Mean Accuracy AUC = 0.945	Yes
Arnaout et al. (2021)	107,283 Images	CNN imaging model	AUC = 0.99 95% Sensitivity 100% NPV (Negative Predictive Value)	No
Esteva et al. (2017)	129,450 Images	CNN model	72.1% Accuracy – 3 class model 55.4% Accuracy – 9 class model	Yes

			AUC = 0.91	

Table 2. Neural Network Analysis Europe

Asia	Data Sample Population	Type of ML model	Performance/ Measure of Interpretability	Clinically Intelligible?
Kim & Lim (2021)	7031	Backpropogation DNN	Accuracy:81.9% Specificity: 82.1% Recall: 68.6% AUC: 0.855	Yes
Chatterjee et al. (2011)	94	MLP backpropagation	75% Predictive Accuracy	Yes
Bhagyasree et al. (2017)	466	Jrip	70.31% Accuracy	Yes
Khan et al. (2019)	12,636	MLP Backpropogation	Accuracy > 93% NPV > 94%	No
Agarwal & Saxena (2018)	569	Logistic Regression	97.2%	No

Table 3. Neural Network Analysis Asia

Europe	Data Sample Population	Type of ML model	Performance/ Measure of Interpretability	Clinically Intelligible?
Olaniyi & Khashman (2015)	345	Radial Basis	70% Recognition Rate	No
Heden et al. (1997)	1120	MLP backpropagation	AUC = 0.86	No
Olaniyi et al. (2015)	303	MLP backpropagation	85% recognition Rate	No
Karan et al. (2012)	456	3-layer MLP backpropagation	1.045 error value	No
Orukwo & Kabari (2020)	100	3-layer MLP backpropagation	90% Accuracy	No

Experimental Design

Using the data from the comparative analysis it can be concluded that the most optimal neural network mechanism was a backpropagation neural network consisting of a data sample greater than 100 and clinically intelligible data values. In order to confirm that a backpropagation model was the most effective, the procedure of the experiment was altered slightly to incorporate a comparison between a backpropagation model and a Logistic Regression Model. Before proceeding to the results of these two models, it is important to understand the differences between them. Firstly, a back-propagation model is essentially a guess and check mechanism in which random weights are initially assigned to each node in the neural network. The output value is received and then compared to the target output value. Based on the error, the model works backwards recalibrating the weights of the nodes until the target output is very similar to the actual output. In essence, the model is checking possibilities to determine how to map proper inputs to their respective outputs. On the other hand, a logistic regression model is similar to the line of best-fit idea implemented in linear correlation except uses what is known as a sigmoid function, an S-shaped curve. Based on this curve, the algorithm finds "a relationship between [the] features and probability of a particular outcomes" or in layman's terms it takes in an input and determines which of two outcomes is more likely (Agrawal 2017). Looking at the results, Figure 1 illustrates a confusion matrix for the Logistic Regression model and Table 4 represents the performance metrics associated with the confusion matrix. From Table 4 it can be seen that four major metrics were analyzed from the neural network: accuracy, precision, recall and F1. The respective definitions for each of these terms can be seen in Appendix B. The same data was tested with a Backpropagation model and the results can be seen in Figure 2 and Table 5. The data shows that the Backpropagation model slightly surpassed the Logistic Regression model in terms of each of the four performance measures.

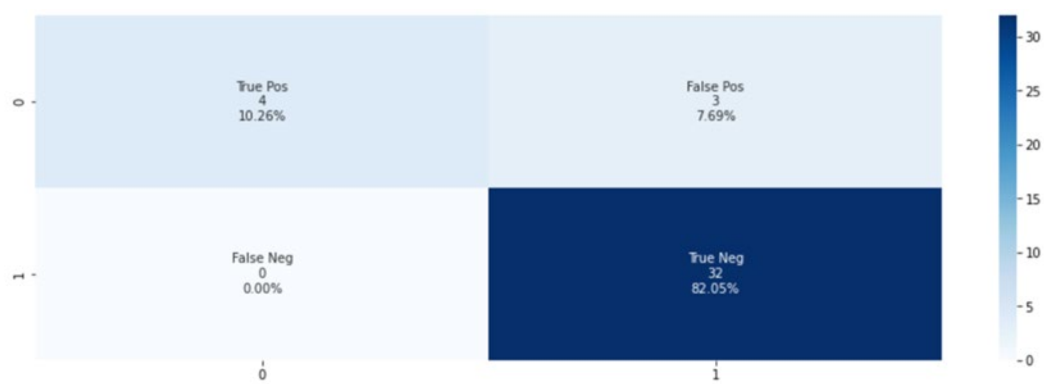

Figure 1. Confusion Matrix Logistic Regression

Table 4. Performance Metrics Logistic Regression

Accuracy	Precision	Recall	F1
0.923077	0.914286	1.0	0.955224

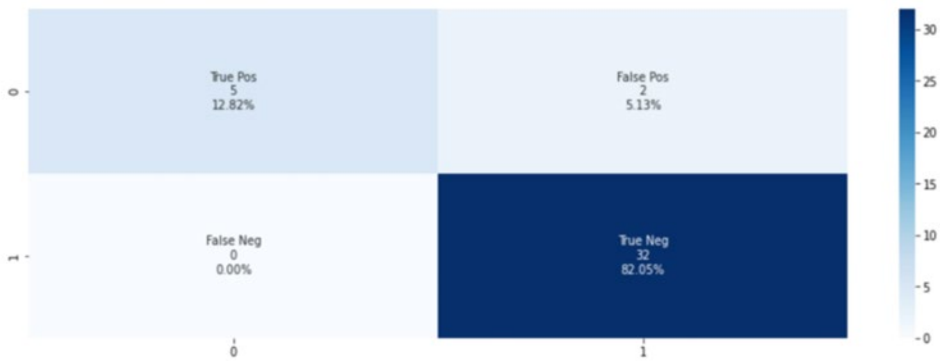

Figure 2. Confusion Matrix Backpropogation

Table 5. Performance Metrics Backpropogation

Accuracy	Precision	Recall	F1
0.948718	0.941176	1.0	0.969697

In addition to the performance of the neural network, in order to ensure the data was clinically intelligible, a SHAPLEY value algorithm was used to determine which factors most heavily impacted the decision-making. The data from the Logistic Regression Model can be seen in Figure 3, and the data from the Backpropagation model can be seen in Figure 4. Although both of the models had relatively similar performance measures, the weight of each factor in the decision-making was very different for both models as the Logistic Regression Model weighed MDVP:Jitter(%) and other measures of variation in fundamental frequency higher, while the Backpropagation model weighed spread2 along with other non-linear measures of frequency higher. From this we can conclude that both models are relatively similar in their effectiveness in diagnosing dementia, but have great differences in the process by which the diagnosis was determined.

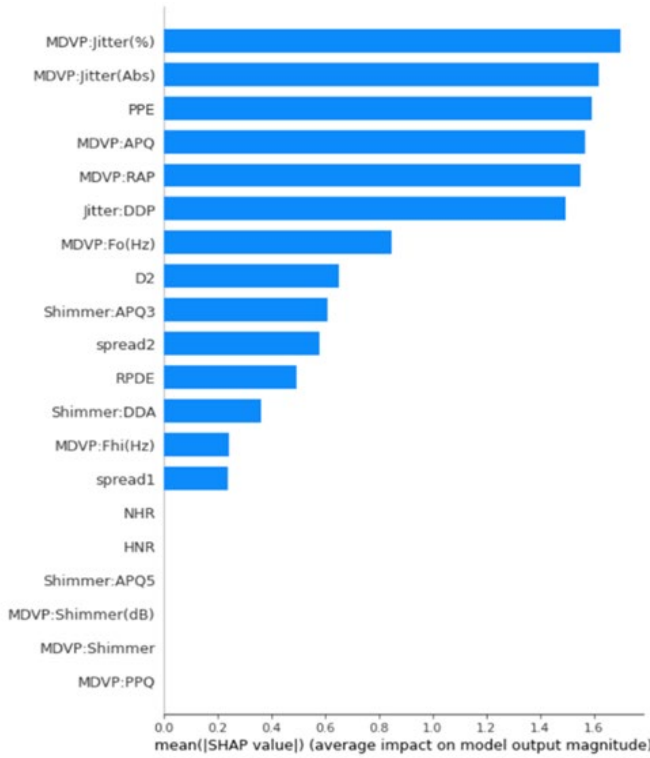

Figure 3. Shapley Value Analysis Logistic Regression

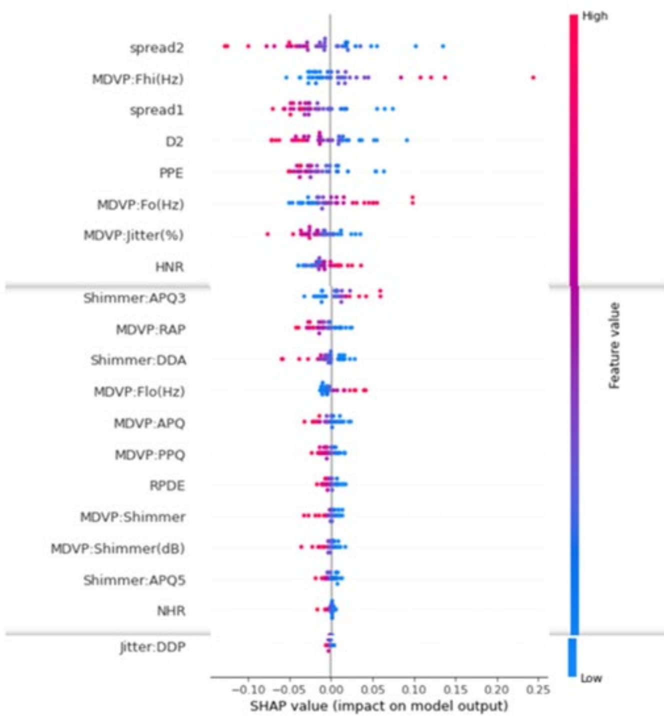

Figure 4. Shapley Value Analysis Backpropogation

Discussion

Significance of Data

From the comparative analysis methodology, a few major conclusions can be drawn that connects to the overarching research question. For one it can be concluded that data samples with a greater abundance of data were better fit to be applied for a real-world healthcare application. The likely reason for this conclusion is the fact that ANNs are a very complex algorithmic process and with limited amounts of data, accuracy is often minimized as a result of a lack of pattern being identified by the algorithm. If we connect this back to what artificial intelligence in nature relies on, data, we can understand that without the necessary amount of data, accuracy is either low or metrics of accuracy are not valid and reliable as a result of a skewed sample. In addition, after thorough analysis of the 15 academic journals, one other theoretical conclusion can be made from the data which is that a backpropagation research model is the most common as well as the most accurate and precise. The likely reason for this conclusion is that artificial neural networks that function around a backpropagation model are more accurately aligned to the values in the model in comparison to a logistic regression model that functions around a curve of best-fit principle. Therefore, there is a bit more room for error in other types of models than a backpropagation one. These conclusions can be connected to other applications of neural networks and can be used to better promote disease diagnosis and improve the accuracy of neural networks.

In terms of the experimental design methodology, the metrics from the confusion matrix illustrate an effective neural network. From Figure 1 it can be seen that 32 true negatives were identified correctly, 0 false negatives, 3 false positives, and 4 true positives. From Figure 2 it can be seen that 32 true negatives, 0 false negatives, 2 false positives, and 5 true positives were identified. The issue with both of the models was the fact that 3 false positives were identified. False positives are a very prominent issue because aside from the mental and emotional damage it can cause the patient, it also leads to a utilization of hospital resources which wastes money and time. So based on this information, we can make some important conclusions to answer the research question. For one, artificial neural networks do have the ability to predict cases of dementia in individuals and secondly, they can do this accurately in a way that can be applied to a real-world healthcare situation. This can be validated by the Scott et al. (2021) checklist for healthcare application. Examining the ten factors in the checklist, it can be seen that only one criterion was not met by the neural network, the sufficiency of the data which can easily be improved with more resources.

Fulfillment of Gap

The first gap that was identified as a part of the literature review was the fact that most created neural networks contain a gap in clinically intelligible data. This can be simply defined as data that prevents the black box conundrum or more specifically data that ensures that conclusions generated from the algorithm can be explained to both the user of the algorithm and the patient. As seen in the 15 sources from the comparative analysis methodology, over half of the academic journals created neural networks without clinically intelligible data. The neural network created in the experimental design partially fulfilled the gap in the literature and can be seen in Figures 3 and s. From the tables, it can be seen that both models have different mean Shapley values for each of the attributes. These Shapley values demonstrate which factors the algorithm used in terms of importance level to make its conclusions. This ensures that when the algorithm is used, the medical professional using the algorithm can easily understand the data and explain the condition to the patient to ensure that they are aware of how their diagnosis was made. Despite this, the gap was not completely fulfilled because some research already exists consisting of Shapley value algorithms for other neural network models, and there still remains the issue that the steps of the entire diagnosis process cannot be retraced. The second gap that was identified as part of the literature review was that no research had been conducted that compared different neural networks to analyze the most optimal neural network conditions. This gap was effectively fulfilled through the comparative analysis methodology as three major factors were identified.

Limitations

Many small limitations were present across the research process, but a few major limitations were present. The first limitation that can be identified comes from the comparative analysis methodology in terms of access to sources. Many of the research papers, that I was initially exploring were blocked by a paywall or blocked the data information because of confidentiality agreements. This limitation prevented the use of these academic journals in the comparative analysis methodology and resulted in 15 sources being the number that could be achieved. Transitioning to the experimental design methodology there were many limitations in data access. The first limitation was that the initial data collection plan for the research was to get data from dementia patients over the age of 65. As a high school student, I was unable to access many of the data sources that held this information because of their expenses. Therefore, the data I was able to find was from the UCI Machine Learning Repository only pertained to Parkinson's Disease in particular and did not have an age restriction as part of the data. Parkinson's is a specific classification of dementia but is less prominent than others such as Alzheimer's Disease. Another limitation with this data set is that the data sample is limited in terms of Parkinson's patients so more data would be needed to better validate the model and ensure its applicability. If expense capabilities were expanded and more access to data was available, the research could have better answered the research question and added to the scholarly conversation surrounding neural networks and their prediction of diseases.

Conclusions

Implications

From the analysis of background research sources along with the independent methodologies and experiments, we can conclude that artificial neural networks are an effective way to predict dementia. This leads to the question of what this conclusion means for the future of healthcare. As seen from the analysis of scholarly sources, current healthcare practices have very limited involvement of artificial intelligence due to many limitations such as interpretability, data access, and other ethical concerns. Despite these previous concerns, through this research, it can be seen that artificial neural networks can be created in a form where its results can be partially interpreted to understand the process of diagnosis. In addition, despite the fact that the experimental design methodology utilized limited data sources, the analysis of academic scholarly sources from the comparative analysis methodology confirms the idea that accessing data is possible with ample resources. Lastly in regard to ethical concerns, the findings from this research study have shown that although neural networks aren't perfect, their accuracy is extremely high and exceeds that of professional doctors, especially for the diagnosis of difficult diseases such as dementia. Therefore, from this study we can conclude that artificial neural networks are the key to revolutionizing healthcare diagnosis by improving process efficiency and limiting the extent of misdiagnoses.

Future Research and Reflection

Throughout this research process, I have gained great knowledge about the world of artificial intelligence, machine learning, and dementia, but more importantly I was able to develop essential skills such as data collection and analysis that can be applicable to any future career field I choose to pursue. Furthermore, although this research study, effectively produced a neural network to predict Parkinson's disease, a subset of dementia, and also confirmed the notion that ANNs can be used as an effective diagnosis system for dementia, there are still a great amount of research to be conducted on this topic before it can become revolutionizing technology. For instance, one key attribute this research is limited is that it only detects Parkinson's disease when the specific tests are conducted and tracked. In order to expand the scope of the ANN, it is important that it should be able to detect a diagnosis as well as the type of dementia the patient has through more general tests which are easier to conduct. In order for this to occur, future research must incorporate a much larger data set with a more complex algorithmic process. Another major area for future research on ANNS and dementia is treatment. While diagnosis is very important and a critical factor in the medical process, the diagnosis is meaningless without proper treatment. Therefore, ANNs should be programmed to not only diagnose dementia but also to create effective treatment plans coordinated to the patient's diagnosis. This will allow for dementia cases to be reduced as well as for individuals with dementia to live longer and healthier lives. Overall, the potential that artificial intelligence and neural networks have in the medical field is endless, and with the right creation and proper implementation, it has the potential to revolutionize patient care for generations upon generations.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

Agrawal, A. (2017, March 31). *Logistic regression. Simplified.* Medium. Retrieved April 26, 2023, from https://medium.com/data-science-group-iitr/logistic-regression-simplified-9b4efe801389

Agarwal, A., & Saxena, A. (2018). *Malignant tumor detection using machine learning ... - student learning*. Retrieved March 21, 2023, from http://studentlearning.in/wp-content/uploads/2020/08/Malignant-Tumor-Detection-Using-Machine-LearningArushi-Agarwal.pdf

Arnaout, R., Curran, L., Zhao, Y., Levine, J. C., Chinn, E., & Moon-Grady, A. J. (2021). An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. *Nature Medicine*, *27*(5), 882–891. https://doi.org/10.1038/s41591-021-01342-5

Ballenger, J. F. (2017, July 1). *Framing Confusion: Dementia, Society, and History*. Journal of Ethics | American Medical Association. https://journalofethics.ama-assn.org/article/framing-confusion-dementia-society-and-history/2017-07

Bell, S. (2009). *Experimental design*. Experimental Design - an overview | ScienceDirect Topics. Retrieved December 4, 2022, from https://www.sciencedirect.com/topics/earth-and-planetary-sciences/experimental-design#:~:text=Experimental%20design%20is%20the%20process,has%20on%20a%20dependent%20variable.

Belova, Kira. (2020, December 7). *How Artificial Intelligence Shapes Healthcare Industry*. PixelPlex. https://pixelplex.io/blog/how-ai-shapes-healthcare-industry/

Bhagyashree, S. I., Nagaraj, K., Prince, M., Fall, C. H., & Krishna, M. (2017). Diagnosis of dementia by machine learning methods in epidemiological studies: A pilot exploratory study from South India.

Social Psychiatry and Psychiatric Epidemiology, *53*(1), 77–86. https://doi.org/10.1007/s00127-017-1410-0
Bittrich, S., Kaden, M., Leberecht, C., Kaiser, F., Villmann, T., & Labudde, D. (2019). Application of an interpretable classification model on early folding residues during protein folding. *BioData Mining*, *12*(1). https://doi.org/10.1186/s13040-018-0188-2
Buntinx, F., de Lepeleire, J., Paquay, L., Iliffe, S., & Schoenmakers, B. (2011). Diagnosing dementia: No easy job. *BMC Family Practice*, *12*(1). https://doi.org/10.1186/1471-2296-12-60
Chatterjee, M., Chowdhury, D. R., & Samanta, R. K. (2011, January). *An Artificial Neural Network model for neonatal disease diagnosis*. Retrieved March 20, 2023, from http://www.cscjournals.org/library/manuscriptinfo.php?mc=IJAE-57
Dementia Diagnosis. Stanford Health Care (SHC) - Stanford Medical Center. (2018, August 1). https://stanfordhealthcare.org/medical-conditions/brain-and-nerves/dementia/diagnosis.html
Ercal, F., Chawla, A., Stoecker, W. V., Hsi-Chieh Lee, & Moss, R. H. (1994). Neural network diagnosis of malignant melanoma from color images. *IEEE Transactions on Biomedical Engineering*, *41*(9), 837–845. https://doi.org/10.1109/10.312091
Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with Deep Neural Networks. *Nature*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056
Goldfarb, A.,&Teodoridis.(2022,March 9). *Why is AI adoption in health care lagging?* Brookings. Retrieved September 16, 2022, from https://www.brookings.edu/research/why-is-ai-adoption-in-health-care-lagging/#:~:text=AI%20has%20received%20a%20great,and%20a%20misalignment%20of%20incentives
Götz, J. (2018, January 18). *Why it's so hard to treat dementia*. CNN. https://edition.cnn.com/2018/01/18/health/dementia-alzheimers-drug-treatment-partner
Government AI readiness index 2021. Oxford Insights. (2022, January 21). Retrieved December 5, 2022, from https://static1.squarespace.com/static/58b2e92c1e5b6c828058484e/t/61e95661c567937d21998d14/1642681965033/gov_ai_readiness_2021.pdf
Hardesty, L.(2017, April 14). *Explained: Neural networks*. MIT News | Massachusetts Institute of Technology. https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414
Hedén Bo, Öhlin Hans, Rittner, R., & Edenbrandt, L. (1997). Acute myocardial infarction detected in the 12-lead ECG by Artificial Neural Networks. *Circulation*, *96*(6), 1798–1802. https://doi.org/10.1161/01.cir.96.6.1798
Karan, O., Bayraktar, C., Gümüşkaya, H., & Karlık, B. (2012). Diagnosing diabetes using neural networks on small mobile devices. *Expert Systems with Applications*, *39*(1), 54–60. https://doi.org/10.1016/j.eswa.2011.06.046
Kasban, H., El-bendary, M., & Salama, D. (2015). A Comparative Study of Medical Imaging Techniques. *International Journal of Information Science and Intelligent System, 4*(2), 37–58.
Khan, M. T., Kaushik, A. C., Ji, L., Malik, S. I., Ali, S., & Wei, D.-Q. (2019). Artificial Neural Networks for prediction of tuberculosis disease. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.00395
Kim, J., & Lim, J. (2021). A deep neural network-based method for prediction of dementia using Big Data. *International Journal of Environmental Research and Public Health*, *18*(10), 5386. https://doi.org/10.3390/ijerph18105386
Li, D., Tong, T. W., & Xiao, Y. (2021, February 18). *Is China emerging as the global leader in AI?* Harvard Business Review. Retrieved December 4, 2022, from https://hbr.org/2021/02/is-china-emerging-as-the-global-leader-in-ai.

Mollah, M. (2021, October 8). *Comparative analysis -definition, overview, methods, benefits, templates*. Real Wealth Business. Retrieved December 4, 2022, from https://realwealthbusiness.com/comparative-analysis/.

Olaniyi, E. O., Oyedotun, O. K., Helwan, A., & Adnan, K. (2015). Neural network diagnosis of heart disease. *2015 International Conference on Advances in Biomedical Engineering (ICABME)*. https://doi.org/10.1109/icabme.2015.7323241

Olaniyi, E., & Khashman, A. (2015, November). *Liver Disease Diagnosis Based on Neural Networks* . Retrieved March 21, 2023, from https://www.researchgate.net/profile/Ebenezer-Olaniyi/publication/318468361_Liver_Disease_Diagnosis_Based_on_Neural_Networks/links/596cd4eeaca272042d37f90b/Liver-Disease-Diagnosis-Based-on-Neural-Networks.pdf.

Orukwo, J. O., & Kabari, L. G. (2020). Diagnosing diabetes using artificial neural networks. *European Journal of Engineering Research and Science*, *5*(2), 221–224. https://doi.org/10.24018/ejers.2020.5.2.1774

Population Reference Bureau. (n.d.). *Fact sheet: U.S. Dementia Trends*. PRB. Retrieved September 16, 2022, from https://www.prb.org/resources/fact-sheet-u-s-dementia-trends/

Qiu, S., Miller, M. I., Joshi, P. S., Lee, J. C., Xue, C., Ni, Y., Wang, Y., De Anda-Duran, I., Hwang, P. H., Cramer, J. A., Dwyer, B. C., Hao, H., Kaku, M. C., Kedar, S., Lee, P. H., Mian, A. Z., Murman, D. L., O'Shea, S., Paul, A. B., … Kolachalama, V. B. (2022). Multimodal Deep Learning for Alzheimer's Disease Dementia Assessment. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-31037-5

Quest, D. J., Upjohn, D., Pool, E., Menaker, R., Hernandez, J. S., & Poole, K. (2020, November 9). *Demystifying ai in healthcare: Historical perspectives and current considerations*. American Association for Physician Leadership . Retrieved from https://www.physicianleaders.org/articles/demystifying-ai-healthcare-historical-perspectives-and-current-considerations

Scott, I., Carter, S., & Coiera, E. (2021). Clinician checklist for assessing suitability of machine learning applications in Healthcare. *BMJ Health & Care Informatics*, *28*(1). https://doi.org/10.1136/bmjhci-2020-100251

Siwicki, B. (2022, February 21). *Dispelling the mysteries around neural networks in healthcare*. Healthcare IT News. https://www.healthcareitnews.com/news/dispelling-mysteries-around-neural-networks-healthcare

Tomov, S., & Tomov, S. (2018). On Deep Neural Networks for Detecting Heart Disease. *University of Tennessee*. https://doi.org/10.13140/RG.2.2.11232.12806