MODIFIED GATED RECURRENT UNIT BASED ON UPDATE-PRIME GATE FOR HEART DISEASE PREDICTION

IRFAN JAVID

A thesis submitted in
fulfilment of the requirement for the award of the
Doctor of Philosophy in Information Technology

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia

JULY 2023

DEDICATION

This Doctoral thesis ought to be dedicated to my parents, who were always there for me and their love, encouragement, and support throughout my journey. I would also like to dedicate my research work to my supervisor Prof. Dr. Rozaida Binti Ghazali, who had been incredibly supportive throughout and made it possible for me to achieve my goal successfully.

ACKNOWLEDGEMENT

In the name of ALLAH, the most gracious and the most merciful. I want to pay my profound gratitude to my Almighty Allah for bestowing all the amenities and blessings that enabled me to accomplish my research work.

First of all, I would like to thank my respected supervisor Prof. Dr. Rozaida Binti Ghazali, for being incredibly supportive, generous, and encouraging. I want to be thankful to her for her guidance, inspiration, and motivation and be available in every challenging situation throughout my research. Her supervision depicts an amalgam of professionalism and compassion that always inspired me and motivated me toward hard work and sincerity. I consider myself fortunate to have her as my supervisor, and I will always appreciate her support and guidance. Furthermore, I have learned a lot from her expertise. Notably, anytime I was overcome with sadness, hopelessness, or worry during the entire endeavor, she acted like a true mentor and avoided obstacles.

I want to thank University Tun Hussein Onn Malaysia for giving me a great working atmosphere and many resources. The dedicated and deliberate efforts of the University Tun Hussein Onn Malaysia administration and staff to provide cutting-edge facilities for delivering high-quality education in all subject areas are impressive. Because of their dedicated efforts and commitment, we have been able to do research work under the supervision of capable Information technology researchers who are well-known in Malaysia and beyond. I also like to thank the University of Poonch Rawalakot for its support and for allowing me to get my doctorate.

Sincere gratitude to all my friends, especially Mr. Zulqarnain, for their generosity and moral support throughout my research journey. Finally, I want to express my heartfelt appreciation to my beloved parents, my wife, son, brothers, and sisters for their unending love, prayers, and encouragement.

ABSTRACT

There is a growing interest in accurate methods to detect heart disease since this is one of the leading causes of mortality worldwide. This interest has spurred advancements in various fields, including the application of deep learning algorithms for heart disease detection. One such algorithm is the Gated Recurrent Unit (GRU), which falls under the category of Recurrent Neural Network (RNN) that comprises an update gate and reset gate. GRU is considered one of the most efficient prediction approaches, particularly on time series datasets. However, when GRU was used to address heart disease prediction, it encountered three significant problems such as failure of data dimensionality reduction, slow convergence rate and high computational cost. Therefore, this research proposed a model named Chi-square Gated Recurrent Unit (χ^2-GRU) to solve the dimensionality reduction issues. However, still, the χ^2-GRU model has a slow convergence rate. This problem is tackled by enhancing the χ^2-GRU model to Chi-square U-prime Gated Recurrent Unit (χ^2-U$^{/}$GRU) by dividing the update gate functionality into two parts. Finally, the $tanh$ activation function in the candidate state is replaced by a Tunable Swish function in order to minimize the computational complexity of the proposed χ^2-U$^{/}$GRU. The proposed χ^2-GRU and χ^2-U$^{/}$GRU models have been evaluated on four benchmark heart disease datasets and compared with five prediction approaches, including GRU, LSTM, CNN, SVM, and Random Forest. Based on the performance evaluation factors, including accuracy, precision/recall, F-Measure, specificity, and Friedman statistical test, it has been found that the proposed models perform significantly better than other prediction approaches and demonstrate the improved effectiveness and efficiency of the developed models. The proposed model exhibited a remarkable accuracy rate of 93.1%, 88.9%, 89.3% and 86.0%, showcasing its ability to make accurate predictions.

ABSTRAK

Terdapat minat yang semakin meningkat dalam kaedah mengesan penyakit jantung kerana penyakit ini adalah salah satu punca utama kematian pada seluruh dunia. Isu ini telah menarik minat dalam pelbagai bidang, termasuk pengunaan aplikasi algoritma pembelajaran mendalam untuk mengesan penyakit jantung. Salah satu algoritma yang digunakan dalam kajian ini ialah Unit Berulang Berpagar (GRU), yang termasuk dalam kategori Rangkaian Neural Berulang (RNN) yang terdiri daripada mengemas kini dan menetapkan semula pintu. GRU dianggap sebagai salah satu pendekatan ramalan yang paling effektif, terutamanya pada set data. Walau bagaimanapun, apabila GRU dicadangkan untuk menangani penyakit jantung ini terdapat tiga masalah penting dalam GRU, seperti kegagalan pengurangan dimensi data, kadar penumpuan yang perlahan dan kos pengiraan yang tinggi. Oleh itu, penyelidikan ini mencadangkan satu model yang dinamakan Unit Berulang Berpagar Chi-square (χ2-GRU) untuk menyelesaikan isu pengurangan dimensi. Walau bagaimanapun, model χ2-GRU mempunyai kadar penumpuan yang perlahan. Masalah ini diatasi dengan meningkatkan model χ2-GRU kepada Unit Berulang Berpagar U-kuadrat U (χ2-U/GRU) dengan membahagikan kefungsian mengemas kini (*update gate*) kepada dua bahagian. Akhir sekali, fungsi pengaktifan $tanh$ dalam keadaan calon digantikan dengan fungsi Tunable Swish baharu untuk meminimumkan pengiraan χ2-U/GRU yang dicadangkan. Model χ2-GRU dan χ2-U/GRU yang dicadangkan dalam penyelidikan ini akan dinilai pada empat set data penyakit jantung dan dibandingkan dengan lima pendekatan ramalan, termasuk GRU, LSTM, CNN, SVM dan Random Forest. Berdasarkan faktor penilaian prestasi, ketepatan, ketepatan/ingat, F-Measure, kekhususan, dan ujian statistik Friedman, didapati bahawa model yang dicadang akan memberikan prestasi yang lebih baik daripada pendekatan ramalan yang lain, memberi peningkatan keberkesanan dan kecekapan yang dibangunkan. Model yang dicadangkan memberi kadar ketepatan yang luar biasa iaitu 93.1%, 88.9%, 89.3% dan 86.0%, memberi keupayaan untuk membuat ramalan yang tepat.

CONTENTS

	TITLE	**i**
	DECLARATION	**ii**
	DEDICATION	**iii**
	ACKNOWLEDGEMENT	**iv**
	ABSTRACT	**v**
	ABSTRAK	**vi**
	CONTENTS	**vii**
	LIST OF TABLES	**xii**
	LIST OF FIGURES	**xiv**
	LIST OF SYMBOLS AND ABBREVIATION	**xvi**
	LIST OF APPENDICES	**xviii**
	LIST OF PUBLICATIONS	**xix**
CHAPTER 1	**INTRODUCTION**	**1**
	1.1 Background Study	1
	1.2 Problem Statement	5
	1.3 Research Objectives	8
	1.4 Research Scope	8
	1.5 Research Significance	9
	1.6 Organization of the Thesis	11
CHAPTER 2	**LITERATURE REVIEW**	**13**
	2.1 Introduction	13
	2.2 Heart Disease	14

	2.2.1 Types of Heart Disease	15
	2.2.2 Risk Factors of Heart Disease	16
2.3	Application of Machine Learning in Healthcare	17
2.4	Related works on Heart Disease Prediction	22
2.5	Dimension Reduction	26
2.6	Dimension Reduction Techniques	27
	2.6.1 Mutual Information Gain	27
	2.6.2 Relief Algorithm	28
	2.6.3 Analysis of Variance Method	29
	2.6.4 Chi-square Test	30
2.7	Neural Networks	32
2.8	Deep Neural Networks	33
	2.8.1 Convolutional Neural Networks (CNN)	34
	2.8.2 Recurrent Neural Network	36
	2.8.3 Long Short-Term Memory (LSTM)	38
2.9	Gated Recurrent Unit (GRU)	41
	2.9.1 Architecture and Properties of GRU	42
	2.9.2 Application of GRU in Heart disease prediction	46
2.10	Transfer Function in Neural networks	48
	2.10.1 Sigmoid Activation Function	49
	2.10.2 Tanh Activation Function	50
	2.10.3 Rectified Linear Unit (ReLU)	50
	2.10.4 Swish Activation Function	51
2.11	Back Propagation Through Time	51
2.12	Discussion Research gaps	53
2.13	Chapter Summary	55

CHAPTER 3 RESEARCH METHODOLOGY **56**
3.1 Introduction 56
3.2 Research Framework 56
3.3 Phase 1: Data Preparation 59
3.3.1 Data Acquisition 59
3.3.1.1 Cleveland Heart Disease 59
3.3.1.2 SPECT 60
3.3.1.3 Statlog 60
3.3.1.1 Framingham Heart Dataset 60
3.3.2 Data pre-processing 61
3.3.2.1 Removing Inconsistent Data 61
3.3.2.2 Selecting the features 62
3.3.3 Data Transformation 63
3.3.4 Data Partitioning 64
3.4 Phase 2: Proposed Research Model 64
3.4.1 Model Development 65
3.4.2 The Proposed χ^2-U/GRU 65
3.5 Phase 3: Network Models Training Procedure 67
3.5.1 Experimental Setup 67
3.5.2 Parameters Setting 67
3.5.3 Training of the networks 69
3.6 Phase 4: Results and Analysis 70
3.6.1 Performance Evaluation 71
3.6.1.1 Prediction Accuracy 71
3.6.1.2 Error Rate 72
3.6.1.3 Precision 72
3.6.1.4 Sensitivity/ Recall 72
3.6.1.5 Specificity 73
3.6.1.6 F-measure 73
3.7 Comparison Analysis 74

3.8 Chapter Summary 74

CHAPTER 4 THE PROPOSED MODEL: CHI-SQUARE GATED RECURRENT UNIT (χ^2-GRU) 75

4.1 Introduction 75

4.2 The Chi-square Method for Irrelevant Features Elimination 76

4.3 The χ^2-GRU Structure 80

4.4 Chapter Summary 85

CHAPTER 5 THE PROPOSED MODEL: CHI-SQUARE U-PRIME GATED RECURRENT UNIT (χ^2-U$^/$GRU) 86

5.1 Introduction 86

5.2 The Proposed χ^2-U$^/$GRU Structure 87

5.2.1 Modified update gate 88

5.2.2 Significance of T-Swish Function 90

5.3 χ^2-U$^/$GRU Mathematical Model 91

5.4 Chapter Summary 97

CHAPTER 6 RESULTS AND DISCUSSION 99

6.1 Introduction 99

6.2 Experimental Design 100

6.3 Results on Accuracy for GRU based on Feature Selection Techniques 100

6.4 χ^2-U$^/$GRU Based Features Reduction 103

6.5 Results on Accuracy Based on various Learning Rate 104

6.6 Comparative Analysis 107

6.6.1 Convergence Rate 107

6.6.2 Results on the Accuracy for All Network Models 110

6.6.3 Models Performance in terms of Precision

		Recall, F-measure, and Specificity	112
	6.7	Statistical Performance Evaluation	114
	6.8	Time Complexity	117
	6.9	Discussion	119
	6.10	Chapter Summary	121
CHAPTER 7	**CONCLUSION AND FUTURE WORK**		**122**
	7.1	Introduction	122
	7.2	Objectives Accomplished	123
		7.2.1 First Objective	123
		7.2.2 Second Objective	124
		7.2.3 Third Objective	124
		7.2.4 Fourth Objective	125
	7.3	Research Contribution	125
	7.4	Future work	127
	7.5	Concluding Remarks	128
	REFERENCES		**129**
	APPENDICES		**157**
	VITA		**161**

LIST OF TABLES

2.1	Summary of Machine Learning Research Work on Different Data Types	21
2.2	Summary of Related Work on Heart Disease Prediction	25
2.3	Advantages and Disadvantages of Feature Selection Techniques	31
2.4	Performance of CNN in heart disease prediction	36
2.5	Performance of RNN in heart disease prediction	38
3.1	Dataset Properties	61
3.2	Parameters setting of the proposed and comparative models	69
4.1	Example of Chi-square score calculation	78
6.1	Performance of ANOVA-GRU model on benchmark datasets	101
6.2	Performance of MI-GRU model on benchmark Datasets	101
6.3	Performance of Relief-GRU model on benchmark Datasets	101
6.4	Performance of χ^2-GRU model on benchmark Datasets	102
6.5	Percentage of features reduction for each technique on benchmark datasets	102
6.6	χ^2-U$^/$GRU-based feature reduction results in all datasets	104
6.7	Heart Disease Prediction Accuracy (%) on all datasets Using different learning rate	106
6.8	Heart Disease Prediction performance (%) of the proposed and comparative models on the Cleveland heart disease dataset	112
6.9	Heart Disease Prediction performance (%) of the proposed and comparative models on the SPECTF	

	heart disease dataset	113
6.10	Heart Disease Prediction performance (%) of the proposed and comparative models on the Statlog heart disease dataset	113
6.11	Heart Disease Prediction performance (%) of the proposed and comparative models on the FHS heart disease dataset	113
6.12	Ranking Based on Accuracy for Each Model on Different Datasets	114
6.13	Post-Hoc Procedure Based on Pairwise Comparisons Method	115
6.14	Average Improvement Between χ^2-U$^{\prime}$GRU and χ^2-GRU	116
6.15	Average Improvement Between X^2-U$^{\prime}$GRU and LSTM	117
6.16	Running time comparison of different models on heart disease prediction datasets	119

LIST OF FIGURES

1.1	Summary of research objectives, studies, and contribution of this research.	10
2.1	Sketch (A) depicts the place of the heart in the human body Sketch (B) depicts the blood vessels with normal flow of blood. Sketch (C) depicts a heart blood vessel contracted by blockage.	15
2.2	The standard CNN architecture	35
2.3	RNN Architecture	37
2.4	LSTM Architecture	39
2.5	The architecture of standard GRU	45
2.6	Demonstration of Activation Functions	49
2.7	Training parameters through BPTT	52
3.1	Proposed Research Framework	58
3.2	Data pre-processing steps	62
3.3	The proposed χ^2-U$^{/}$GRU architecture	66
4.1	Hyperparameter Tuning using Grid Search strategy	80
4.2	Features Selection using Chi-square Method	81
4.3	The χ^2-GRU Structure	82
4.4	Chi-square GRU Working Flow	85
5.1	The proposed χ^2-U$^{/}$GRU Architecture	88
5.2	The internal layered architecture of the proposed χ^2-U$^{/}$GRU	96
5.3	χ^2-U$^{/}$GRU step-by-step Training Procedure	97
6.1	Performance comparison of Feature Selection Techniques with GRU in terms of Accuracy	103
6.2	Convergence rate of all models on Cleveland Heart Disease dataset	108

6.3	Convergence rate of all models on SPECTF Heart Disease dataset	108
6.4	Convergence rate of all models on Statlog Heart Disease dataset	109
6.5	Convergence rate of all models on FHS Heart Disease dataset	109
6.6	Prediction accuracy of all approaches on Cleveland heart disease dataset	110
6.7	Prediction accuracy of all approaches on SPECTF heart disease dataset	110
6.8	Prediction accuracy of all approaches on Statlog heart disease dataset	111
6.9	Prediction accuracy of all approaches on FHS heart disease dataset	111

LIST OF SYMBOLS AND ABBREVIATION

CVD	-	Cardiovascular Disease
WHO	-	World Health Organization
GBD	-	Global Burden of Disease
AI	-	Artificial Intelligence
DNN	-	Deep Neural Network
ECG	-	Electrocardiogram
CDSS	-	Clinical decision support systems
SVM	-	Support Vector Machine
CART	-	Classification and Regression
NLP	-	Natural Language Processing
NYHA	-	New York Heart Association
ANN	-	Artificial Neural Network
RNN	-	Recurrent Neural Network
EHR	-	Electronic Health Record
LSTM	-	Long Short-Term Memory
GRU	-	Gated Recurrent Unit
U'GRU	-	U-prime Gated Recurrent Unit
χ^2-GRU	-	Chi-square Gated Recurrent Unit
CNN	-	Convolutional Neural Network
HD	-	Heart Disease
CAD	-	Coronary Artery Disease
BMI	-	Body Mass Index
ROC	-	Receiver Operating Characteristic
AUC	-	Area under the Curve
PACS	-	Picture Archiving and Communication Systems
WSI	-	Whole slide imaging
EEG	-	Electroencephalograms

EOG	-	Electrooculography
PAD	-	Peripheral arterial disease
RF	-	Random Forest
DBN	-	Deep Belief Network
MRI	-	Magnetic resonance imaging
HDPS	-	Heart disorders prediction system
WEKA	-	Waikato Environment for Knowledge Analysis
MLP	-	Multilayer perceptron
SAS	-	Statistical Analysis System
NN	-	Neural Network
MI	-	Mutual Information gain
ANOVA	-	Analysis of Variance
GMDH	-	Group Method of Data Handling
BP	-	Backpropagation
DL	-	Deep Learning
ML	-	Machine Learning
BPTT	-	Backpropagation Through Time
SPECTF	-	Single Proton Emission Computed Tomography Images
Statlog	-	Statistical and Logical Learning Algorithm
FHS	-	Framingham Heart Study
UCI	-	University of California Irvine
CHD	-	Congenital Heart Defects
TP	-	True Positive
TN	-	True Negative
FP	-	False Positive
FP	-	False Negative

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	Accuracy of Proposed and comparative models based on feature reduction techniques using benchmark heart disease datasets	157

LIST OF PUBLICATIONS

Journals:

(i) **Javid, I.,** Zager Alsaedi, A. K., Ghazali, R., Mohmad Hassim, Y. M., & Zulqarnain, M. (2022). Optimally organized GRU-deep learning model with Chi2feature selection for heart disease prediction. Journal of Intelligent and Fuzzy Systems, 42(4), 4083–4094. https://doi.org/10.3233/JIFS-212438. **(Impact Factor = 1.73)**

(ii) **Javid, I.,** Ghazali, R., Zulqarnain, M., & Hassan, N. (2023). Data pre-processing for cardiovascular disease classification: A systematic literature review. Journal of Intelligent and Fuzzy Systems, 44(1), 1525–1545. https://doi.org/10.3233/JIFS-220061. (**Impact Factor = 1.73)**

(iii) **Javid, I.,** Alsaedi, A. K. Z., & Ghazali, R. (2020). Enhanced accuracy of heart disease prediction using machine learning and recurrent neural networks ensemble majority voting method. International Journal of Advanced Computer Science and Applications, 11(3), 540–551. https://doi.org/10.14569/ijacsa.2020.0110369. **(Scopus Indexed)**

Conferences:

(i) **Javid, I**., Ghazali, R., Zulqarnain, M., Husaini, N.A. (2022). "Deep Learning GRU Model and Random Forest for Screening Out Key Attributes of Cardiovascular Disease. *Recent Advances in Soft Computing and Data Mining. SCDM 2022*. vol 457, https://doi.org/10.1007/978-3-031-00828-316. **(Scopus Indexed)**

(ii) **I. Javid,** R. Ghazali, I. Syed, N. A. Husaini and M. Zulqarnain, "Developing Novel T-Swish Activation Function in Deep Learning," *2022 International Conference on IT and Industrial Technologies (ICIT)*, 2022, pp. 1-7, doi: 10.1109/ICIT56493.2022.9989151. (**IEEE Proceeding)**

CHAPTER 1

INTRODUCTION

1.1 Background Study

Currently, studies show that in various countries, the percentage of heart disease is increasing daily and has been considered one of the prevalent causes of death in both genders (men and women) above age sixty. According to the studies (Amini *et al.*, 2021; Townsend *et al*., 2022), heart disease in most countries is regarded as a "second epidemic," a prominent cause of death after any infectious disease. Cardiovascular disease (CVD) is also known as ischemic heart disease, stroke, cardiac insufficiency, peripheral artery disease, and several other vascular disorders. It has a significant contribution to world mortality and has become a crucial cause of reducing the age of life. Heart disease caused an estimated 17.8 million deaths around the globe in 2017. Furthermore, around 330 million years of life were lost, and almost 35.6 million years were lived with infirmity (James *et al*., 2018; Kyu *et al*., 2018). Approximately 80% of worldwide heart failure deaths occur in lower economies and developing countries which ramp up heart disease and risk factors due to continuing epidemiological transition (Anand *et al*., 2020). Middle-class economies commonly confront heart disease compared to high or low-economic countries. According to the report launched by the world health organization (WHO) and World Bank, the Global Burden of Disease (GBD) research (World Bank, 1993) to deal with this challenge in 6 cycles of GBD appraisal published over the years 1999-2004, 2010, 2013, 2015, 2016 and 2017 (Murray & Lopez, 2013; Murray & Lopez, 2017). Furthermore, with advancements in clinical practice, training, and research in worldwide health, the active involvement of the cardiology community is crucial. The availability of up-to-date statistics on heart disease is beneficial in these efforts. The numerous reports by the Institute of Medicine promote CVD health and the preventive methods for the epidemic of heart disease in developing countries (Teo & Rafiq, 2021). Prompt analysis and curing actions against disease may reduce the risk of the condition becoming more severe. Gaining a thorough grasp of risk and preventive variables and improving diagnostic accuracy is essential.

Two methods are used for heart disease diagnosis in hospitals or cardiac centers: (i) Invasive Methods and (ii) Non-Invasive Methods. Electrocardiogram, magnetic resonance imaging, echocardiogram, single-photon emission computer tomography, and exercise stress testing are all non-invasive diagnostic methods. However, their results are uncertain and unreliable as angiography (Kumar & Kumar, 2021). Angiography is an invasive diagnostic method, but it is a highly expensive, intrusive, and technically demanding operation. Therefore, it cannot be used for large-scale population screening or treatment monitoring (Verma *et al*., 2018). Furthermore, these procedures consume significant resource expenditures, such as time, an exclusive laboratory setup, and specific instruments and techniques. Due to these limitations of diagnostic methods, hospitals or cardiac research centers are looking for other less costly and non-invasive approaches to diagnose heart diseases, such as artificial intelligence techniques, which can lead to a straightforward identification of heart disease without going through angiography.

Furthermore, in the field of medicine, Artificial Intelligence (AI) is expected to be utilized for health analysis, curing actions, forecasting upcoming risks, and recognizing appropriate medication (Gulshan *et al*., 2016). Additionally, AI implementation is ramping up at a remarkable rate in medicine. AI can be categorized depending on the machine's capacity for predicting future decisions based on past experiences. There are four leading AI technologies: machine learning, cognitive computing, automation and robotics, and machine vision. The increasing collection of medical data has increased the potential for clinicians to enhance patient analysis. In the current era, physicians have enhanced their use of computer technology to progress in result-creating support. Statistical and machine-learning approaches have been more popular in medical diagnostics over the last few decades (Verma *et al*., 2018; Richens *et al*., 2020).

Moreover, machine learning has become a top-rated tool in the medical sector for the initial analysis of diseases. Machine learning is an analytical method utilized when a task is outsized and complicated to program, for example, transforming medical records into knowledge, predicting pandemics, and analyzing genetic data (Shalev-Shwartz & Ben-David, 2014; Hastie *et al*., 2017). In recent times, different machine-learning approaches have been utilized in research studies to recognize and predict various heart problems. An automated classifier with congestive heart failure has been used for classifying the patients into two categories, i.e., patients with low risks and patients with high risks (Melillo *et al*., 2013). To come up with the optimal features and outclass-performing results, a deep neural network (DNN) has been used for the categorization of electrocardiogram (ECG) data (Rahhal *et al*., 2016). A clinical decision support system (CDSS) is designed to assess the failure of the heart. The authors also inspected the effectiveness of numerous machine learning approaches, including a support vector machine (SVM), a CART-based system with fuzzy rules, neural networks, and Random forests (Guidi *et al*., 2014). Natural language processing (NLP) and the rule-based method have been utilized to find New York Heart Association classes to identify heart failure from amorphous medical records (Zhang *et al*., 2017). To analyse heart disease in diabetic patients, an SVM approach accurately predicts features like blood sugar, blood pressure, and age (Parthiban & Srivatsa, 2012). Contemporary research studies have revealed that a substantial role has been played by different neural network models in forecasting and dealing with several classification problems. In the healthcare domain, deep learning methods have played a vital role in retrieving expedient facts and predicting diseases such as diabetes, heart disease, and brain diseases by surveying collected biomedical data (Shickel *et al*., 2018).

Additionally, deep learning models are the enhanced version of conventional artificial neural networks (ANN) that process difficult tasks. When more layers and units are added to existing neural networks, the network's expressional power increases, resulting in increased cost function complexity. Developing deep learning models has overcome the limitations and drawbacks of standard neural networks. Deep learning models have been utilized in different fields, including medical text classification (Hughes *et al*., 2017), natural language processing (George & Joseph, 2014), transfer learning (Long *et al*., 2017), and computer vision (Voulodimos *et al*., 2018). The petite rate of computer equipment, vigorous processing proficiencies, and remarkable revolution in machine learning methods have driven the motivation to develop and implement autoencoders, denoising autoencoders and stacked denoising autoencoders deep learning techniques. Therefore, Recurrent Neural Network (RNN) has also played an important role in predicting future diseases by employing fast patient representation of electronic health records (EHR) (Miotto *et al*., 2016).

Furthermore, Recurrent Neural Network has been widely employed in various data mining applications in the last few years to demonstrate improved classification performance. RNN can demonstrate robust semantic composition techniques for sentiment categorization as well as grasp temporal relationships in sequence information (Liu, 2016). The primary advantage of RNN is that it can extract sequential temporal data with varied duration, resulting in flexibility in analyzing reviews of varying lengths. There are several forms of RNN, including Matrix-Vector RNNs (Baly *et al*., 2017), Recursive Neural Networks (Irsoy & Cardie, 2014), Recursive Neural Tensor Networks (Socher *et al*., 2013), Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber, 1997), and Gated Recurrent Unit (GRU) (Cho *et al*., 2014). Above all, LSTM and GRU are well-known recurrent architectures. GRU is a variation of the RNN series, the most recent form of sophisticated LSTM cell architecture. GRU is frequently employed the same way as other RNN nodes, especially when the input data contains noise and other methods fail to identify the data points.

Although LSTM and GRU have the same goal of classifying input based on prior time steps, their functioning mechanisms are fundamentally different. LSTM architecture comprises three gates and is more complicated than GRU, whereas the GRU is a state-of-the-art and simple model that comprises two gates an update gate and a reset gate. The responsibility of the update gate is to determine the amount of the preliminary data from the preceding time step *t* should be updated and sent for future usage. While the reset gate functions the opposite way as the update gate, it is used in the standard GRU model to determine the amount of prior data from the preceding hidden state that may be ignored. GRU has three layers and fewer parameters, all described by a simple set of equations, using far fewer processing resources. The input, hidden, and output layers are the three layers utilized to learn statistical characteristics more quickly (Xing & Xiao, 2019). Medical field data is highly dimensional and has many attributes from the raw data. As a result, in current deep learning research, multiple layers have been utilized to extract the most important attributes from raw data. Even though GRU is one of the supreme operative approaches for addressing a variety of disease forecasting issues. Numerous researchers have improved the performance and accuracy of GRU for disease prediction in the past, but there are still substantial shortcomings and limits in the GRU as it currently necessitates further research.

The research's primary goal is to augment the basic GRU architecture by reducing complexity and increasing performance. The major participants of this research are divided into three stages. The first stage is to improve the performance of the GRU model with the Chi-square (χ^2) method for dimensionality reduction to solve heart disease prediction tasks. In the second stage, this research work introduces the update-prime gate ($U^{/}$-gate), which divides the existing update gate into two parts. Each part of the update gate will be responsible for implementing one task at a time, so there will be a significant performance improvement. Finally, in the third stage, the research substitutes the hyperbolic tangent activation function (tanh) with the novel Tunable Swish (T-Swish) activation function in the candidate equation. Hence, this modification also improves disease prediction accuracy and reduces the overall computational time.

1.2 Problem Statement

Genetic variants and climatic influences work together to cause common diseases. Because of this complexity, predicting whether or not a person will inherit diseases is extremely difficult. In developing countries, it is more difficult to diagnose heart disease as qualified medical professionals, equipment used for diagnosis, and other resources are scarce for identifying and treating individuals with heart problems (Verma *et al.*, 2018). Algorithmic models can help diagnose diseases when they are trained with the appropriate data (Javed Mehedi Shamrat *et al.*, 2020). Various machine learning models have been developed to predict risk levels in heart disease. Most of these approaches use publically available datasets for model training and evaluation. These datasets have increased the efficacy of machine learning-based predicting approaches and provided new research paradigms for academics to construct cutting-edge algorithms for forecasting heart disease risk. Such datasets incorporate various statistical facts based on risk factors and the level of the patient's disease. Pre-processing is essential when constructing a predicting approach for heart disease because the available clinical datasets are redundant and inconsistent (Amin *et* *al*., 2019). Furthermore, the dimensionality reduction and disease prediction stage is focused on increasing the capacity of heart disease prediction algorithms. In a detailed manner, the characteristics of forecasting algorithms that directly influence the quality of the solution technique while solving these challenges are the motivational factors for the conducted research.

Also, this research has focused on various challenges to improve the effectiveness of a disease prediction approach. Even though the Gated Recurrent Unit is considered one of the most prevailing techniques in resolving different categories of disease prediction issues, however, there are some deficiencies in the standard GRU model that need to be addressed by constructing an approach that comes up with problem-independence and high resolution for these difficult problems. Like other disease prediction models, GRU has also executed the heart disease prediction assignment, which involves three main phases: extraction of the feature, reduction of the feature, and finally, prediction (Lu *et al*., 2019). For feature reduction, the Chi-square feature selection method is commonly used in certain scenarios due to its specific advantages like independence assumption, feature relevance to target variable, univariate analysis, computational efficiency and suitability for categorical data or discrete variables.

The first deficiency in the standard GRU model is that dimensionality reduction in data according to the input is beyond its capacity, regardless of the application (Hao *et al*., 2019). Because of this deficiency, the approach has become unable to collect the intricate points necessary for obtaining relevant information. Thus, the disease prediction problem is solved in a low-quality manner.

Second, the disadvantage was discovered based on experimentation to overcome the slow convergence rate issue in traditional GRU architecture in resolving heart disease prediction problems. The core cause of this inadequacy is the execution of two different functionalities of the update gate simultaneously during the training phase. Breaking down the functionality of the update gate into two parts to overcome the cohesion problem in the standard GRU model will enhance the model's learning time. Cohesion problem refers to the measure of how closely the responsibilities of a module or component are related to each other. It indicated the degree to which the functions or operations within a module or class are interrelated. A cohesion problem arises when a module or component is responsible for multiple unrelated tasks or has low cohesion, meaning that its responsibilities are not closely related.

REFERENCES

Acharya, U. R., Fujita, H., Oh, S. L., Raghavendra, U., Tan, J. H., Adam, M., Gertych, A., & Hagiwara, Y. (2018). Automated identification of shockable and non-shockable life-threatening ventricular arrhythmias using convolutional neural network. *Future Generation Computer Systems*, *79*, 952–959. https://doi.org/10.1016/J.FUTURE.2017.08.039

Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adeli, H. (2018). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in Biology and Medicine*, *100*, 270–278. https://doi.org/10.1016/J.COMPBIOMED.2017.09.017

Afzal, N., Sohn, S., Abram, S., Scott, C. G., Chaudhry, R., Liu, H., Kullo, I. J., & Arruda-Olson, A. M. (2017). Mining peripheral arterial disease cases from narrative clinical notes using natural language processing. *Journal of Vascular Surgery*, *65*(6), 1753–1761. https://doi.org/10.1016/J.JVS.2016.11.031

Agarap, A. F. M. (2018). A neural network architecture combining gated recurrent unit (GRU) and support vector machine (SVM) for intrusion detection in network traffic data. *ACM International Conference Proceeding Series*, *1*, 26–30. https://doi.org/10.1145/3195106.3195117

Akhoondi, R., Hosseini, R., & Branch, S. (2016). A Fuzzy Expert System For Prognosis Of The Risk Of Development Of Heart Disease. *Journal of Advances in Computer Research*, *7*(2), 101–114.

Alapat, D. J., Menon, M. V., & Ashok, S. (2022). A Review on Detection of Pneumonia in Chest X-ray Images Using Neural Networks. Journal of Biomedical Physics and Engineering, 12(6), 551–558. https://doi.org/10.31661/jbpe.v0i0.2202-1461

Alassaf, M., & Qamar, A. M. (2022). Improving Sentiment Analysis of Arabic Tweets by One-way ANOVA. Journal of King Saud University - Computer and Information Sciences, 34(6), 2849–2859. https://doi.org/10.1016/j.jksuci.2020.10.023

Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., & Khan, J. A. (2019). An Automated Diagnostic System for Heart Disease Prediction Based on χ2 Statistical Model and Optimally Configured Deep Neural Network. *IEEE Access*, *7*, 34938–34945. https://doi.org/10.1109/ACCESS.2019.2904800

Alkhafaji, M. J. A., Aljuboori, A. F., & Ibrahim, A. A. (2020, June 1). Clean medical data and predict heart disease. *HORA 2020 - 2nd International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*. https://doi.org/10.1109/HORA49412.2020.9152870

Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., Alrashed, S., & Olatunji, S. O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in Biology and Medicine*, *109*, 101–111. https://doi.org/10.1016/J.COMPBIOMED.2019.04.017

Alsalamah, M. (2017). *Heart Diseases Diagnosis Using Artificial Neural Networks*. PhD. Thesis. Coventry University.

Al-wajih, E., & Ghazali, R. (2022). Threshold center-symmetric local binary convolutional neural networks for bilingual handwritten digit recognition. Knowledge-Based Systems, 259, 110079. https://doi.org/10.1016/J.KNOSYS.2022.110079

Amin, M. S., Chiam, Y. K., & Varathan, K. D. (2019). Identification of significant features and data mining techniques in predicting heart disease. *Telematics and Informatics*, *36*, 82–93. https://doi.org/10.1016/J.TELE.2018.11.007

Amini, M., Zayeri, F. & Salehi, M (2021). Trend analysis of cardiovascular disease mortality, incidence, and mortality-to-incidence ratio: results from global burden of disease study 2017. BMC Public Health 21, 401.

https://doi.org/10.1186/s12889-021-10429-0

Amma, N. G. B. (2012). Cardiovascular disease prediction system using genetic algorithm and neural network. *2012 International Conference on Computing, Communication and Applications, ICCCA 2012*. https://doi.org/10.1109/ICCCA.2012.6179185

Anand, S., Bradshaw, C. & Prabhakaran, D (2020). Prevention and management of CVD in LMICs: why do ethnicity, culture, and context matter?. BMC Med 18, 7. https://doi.org/10.1186/s12916-019-1480-9

Aydin, I., Karakose, M., & Akin, E. (2011). A multi-objective artificial immune algorithm for parameter optimization in support vector machine. *Applied Soft Computing Journal*, *11*(1), 120–129. https://doi.org/10.1016/j.asoc.2009.11.003

Baccouche, A., Garcia-Zapirain, B., Olea, C. C., & Elmaghraby, A. (2020). Ensemble deep learning models for heart disease classification: A case study from Mexico. *Information (Switzerland)*, *11*(4), 1–28. https://doi.org/10.3390/INFO11040207

Ballester, P., & Araujo, R. M. (2016). On the Performance of GoogLeNet and AlexNet Applied to Sketches. *Proceedings of the AAAI Conference on Artificial Intelligence*, *30*(1), 1124–1128. https://doi.org/10.1609/AAAI.V30I1.10171

Baly, R., Hajj, H., Habash, N., Shaban, K. B., & El-Hajj, W. (2017). A sentiment treebank and morphologically enriched recursive deep models for effective sentiment analysis in Arabic. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *16*(4). https://doi.org/10.1145/3086576

Bandyopadhyay, S. K. (2020). Stacked Bi-directional LSTM Layer Based Model for Prediction of Possible Heart Disease during Lockdown Period of COVID-19. *Journal of Advanced Research in Medical Science & Technology*, *07*(02), 10–14. https://doi.org/10.24321/2394.6539.202006

Barushka, A., & Hajek, P. (2018). Spam filtering using integrated distribution-based balancing approach and regularized deep neural networks. *Applied Intelligence 2018 48:10*, *48*(10), 3538–3556. https://doi.org/10.1007/S10489-018-1161-Y

Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J. A. W. M., Hermsen, M., Manson, Q. F., Balkenhol, M., Geessink, O., Stathonikos, N., Van Dijk, M. C. R. F., Bult, P., Beca, F., Beck, A. H., Wang, D., Khosla, A., Gargeya, R., … Venâncio, R. (2017). Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA*, *318*(22), 2199–2210. https://doi.org/10.1001/JAMA.2017.14585

Benhar, H., Idri, A., & L Fernández-Alemán, J. (2020). Data preprocessing for heart disease classification: A systematic literature review. *Computer Methods and Programs in Biomedicine*, *195*, 105635. https://doi.org/10.1016/J.CMPB.2020.105635

Betancur, J., Commandeur, F., Motlagh, M., Sharir, T., Einstein, A. J., Bokhari, S., Fish, M. B., Ruddy, T. D., Kaufmann, P., Sinusas, A. J., Miller, E. J., Bateman, T. M., Dorbala, S., Di Carli, M., Germano, G., Otaki, Y., Tamarappoo, B. K., Dey, D., Berman, D. S., & Slomka, P. J. (2018). Deep Learning for Prediction of Obstructive Disease From Fast Myocardial Perfusion SPECT: A Multicenter Study. *JACC. Cardiovascular Imaging*, *11*(11), 1654–1663. https://doi.org/10.1016/J.JCMG.2018.01.020

Bote-Curiel, L., Muñoz-Romero, S., Gerrero-Curieses, A., & Rojo-álvarez, J. L. (2019). Deep Learning and Big Data in Healthcare: A Double Review for Critical Beginners. *Applied Sciences 2019, Vol. 9, Page 2331*, *9*(11), 2331. https://doi.org/10.3390/APP9112331

Boukhatem, C., Youssef, H. Y., & Nassif, A. B. (2022). Heart Disease Prediction Using Machine Learning. *2022 Advances in Science and Engineering Technology International Conferences, ASET 2022*. https://doi.org/10.1109/ASET53988.2022.9734880

Chaabene, S., Bouaziz, B., Boudaya, A., Hökelmann, A., Ammar, A., & Chaari, L. (2021). Convolutional neural network for drowsiness detection using eeg signals. Sensors, 21(5), 1–19. https://doi.org/10.3390/s21051734

Chadha, R., & Mayank, S. (2016). Prediction of heart disease using data mining techniques. *CSI Transactions on ICT 2016 4:2*, *4*(2), 193–198. https://doi.org/10.1007/S40012-016-0121-0

Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, *2*, 100016. https://doi.org/10.1016/J.HEALTH.2022.100016

Chen, C., Seff, A., Kornhauser, A., & Xiao, J. (2015). *DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving*. 2722–2730. http://deepdriving.cs.princeton.edu

Chen, R., Sun, N., Chen, X., Yang, M., & Wu, Q. (2018). Supervised Feature Selection with a Stratified Feature Weighting Method. *IEEE Access*, *6*, 15087–15098. https://doi.org/10.1109/ACCESS.2018.2815606

Chen, Z. Y., Fan, Z. P., & Sun, M. (2012). A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data. *European Journal of Operational Research*, *223*(2), 461–472. https://doi.org/10.1016/J.EJOR.2012.06.040

Cheng, L., Zang, H., Ding, T., Sun, R., Wang, M., Wei, Z., & Sun, G. (2018). Ensemble recurrent neural network based probabilistic wind speed forecasting approach. *Energies*, *11*(8). https://doi.org/10.3390/en11081958

Cheng, Y., Wang, F., Zhang, P., & Hu, J. (2016). Risk prediction with electronic health records: A deep learning approach. *16th SIAM International Conference on Data Mining 2016, SDM 2016*, 432–440. https://doi.org/10.1137/1.9781611974348.49

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). *On the Properties of Neural Machine Translation: Encoder-Decoder Approaches*. http://arxiv.org/abs/1409.1259

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*. http://arxiv.org/abs/1412.3555

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2015). Gated Feedback Recurrent

Neural Networks. *32nd International Conference on Machine Learning, ICML 2015*, *3*, 2067–2075. https://doi.org/10.48550/arxiv.1502.02367

Das, R., Turkoglu, I., & Sengur, A. (2009). Effective diagnosis of heart disease through neural networks ensembles. *Expert Systems with Applications*, *36*(4), 7675–7680. https://doi.org/10.1016/J.ESWA.2008.09.013

De Vine, L., Zuccon, G., Koopman, B., Sitbon, L., & Bruza, P. (2014). Medical semantic similarity with a neural language model. *CIKM 2014 - Proceedings of the 2014 ACM International Conference on Information and Knowledge Management*, 1819–1822. https://doi.org/10.1145/2661829.2661974

Deng, L., Li, J., Huang, J. T., Yao, K., Yu, D., Seide, F., Seltzer, M., Zweig, G., He, X., Williams, J., Gong, Y., & Acero, A. (2013). Recent advances in deep learning for speech research at Microsoft. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 8604–8608. https://doi.org/10.1109/ICASSP.2013.6639345

Dietterich, T., Bishop, C., Heckerman, D., Jordan, M., & Kearns, M. (2010). *Introduction to Machine Learning*. 2nd Ed. The MIT Press Cambridge, Massachusetts London, England.

Dong, B., & Wang, X. (2016). Comparison deep learning method to traditional methods using for network intrusion detection. *Proceedings of 2016 8th IEEE International Conference on Communication Software and Networks, ICCSN 2016*, 581–585. https://doi.org/10.1109/ICCSN.2016.7586590

Du, L. H., Liu, W., Zheng, W. L., & Lu, B. L. (2017). Detecting driving fatigue with multimodal deep learning. *International IEEE/EMBS Conference on Neural Engineering, NER*, 74–77. https://doi.org/10.1109/NER.2017.8008295

Dulce G. Pereira, Anabela Afonso & Fátima Melo Medeiros (2015) Overview of Friedman’s Test and Post-hoc Analysis, Communications in Statistics - Simulation and Computation, 44:10, 2636-2653, doi: 10.1080/03610918.2014.931971

Dutta, A., Batabyal, T., Basu, M., & Acton, S. T. (2020). An efficient convolutional neural network for coronary heart disease prediction. *Expert Systems with Applications*, *159*, 113408. https://doi.org/10.1016/J.ESWA.2020.113408

Ebrahim, Qasem, Saleh. (2022). *Threshold Center-Symmetric local binary convolutional neural networks for bilingual handwritten digit recognition*. PhD. Thesis. Universiti Tun Hussein Onn Malaysia.

El-Bialy, R., Salamay, M. A., Karam, O. H., & Khalifa, M. E. (2015). Feature Analysis of Coronary Artery Heart Disease Data Sets. *Procedia Computer Science*, *65*(Iccmit), 459–468. https://doi.org/10.1016/j.procs.2015.09.132

Eric, A. (2018). E-swish: Adjusting activations to different network depths. arXiv preprint arXiv:1801.07145.

Ershkovich, A., & Israelevich, P. (2013). Heaviside predicted Cherenkov radiation. *Physics Today*, *66*(8), 8. https://doi.org/10.1063/PT.3.2058

Escamilla, A. K. G., El Hassani, A. H., & Andres, E. (2019). Dimensionality reduction in supervised models-based for heart failure prediction. *ICPRAM 2019 - Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods*, *Icpram*, 388–395. https://doi.org/10.5220/0007313703880395

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017a). Dermatologist-level classification of skin cancer with deep neural networks. *Nature 2017 542:7639*, *542*(7639), 115–118. https://doi.org/10.1038/nature21056

Fraiwan, L., & Lweesy, K. (2017). Neonatal sleep state identification using deep learning autoencoders. *Proceedings - 2017 IEEE 13th International Colloquium on Signal Processing and Its Applications, CSPA 2017*, 228–231. https://doi.org/10.1109/CSPA.2017.8064956.

Gárate-Escamila, A. K., Hajjam El Hassani, A., & Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked*, *19*, 100330. https://doi.org/10.1016/J.IMU.2020.100330

Genders, T. S. S., Steyerberg, E. W., Hunink, M. G. M., Nieman, K., Galema, T. W., Mollet, N. R., De Feyter, P. J., Krestin, G. P., Alkadhi, H., Leschka, S., Desbiolles, L., Meijs, M. F. L., Cramer, M. J., Knuuti, J., Kajander, S., Bogaert, J., Goetschalckx, K., Cademartiri, F., Maffei, E., … Laule, M. (2012). Prediction model to estimate presence of coronary artery disease: Retrospective pooled analysis of existing cohorts. *BMJ (Online)*, *344*(7862). https://doi.org/10.1136/BMJ.E3485

Geng, L., Wang, H., Xiao, Z., Zhang, F., Wu, J., & Liu, Y. (2019). Fully convolutional network with gated recurrent unit for hatching egg activity classification. *IEEE Access*, *7*, 92378–92387. https://doi.org/10.1109/ACCESS.2019.2925508

George K, S., & Joseph, S. (2014). Text Classification by Augmenting Bag of Words (BOW) Representation with Co-occurrence Feature. *IOSR Journal of Computer Engineering*, *16*(1), 34–38. https://doi.org/10.9790/0661-16153438

Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep Sparse Rectifier Neural Networks. *Artificial Intelligence and Statistics*. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, PMLR* 15:315-323, 2011.

Gonzalez, G., Ash, S. Y., Vegas-Sánchez-Ferrero, G., Onieva, J. O., Rahaghi, F. N., Ross, J. C., Dáz, A., Estépar, R. S. J., & Washko, G. R. (2018). Disease Staging and Prognosis in Smokers Using Deep Learning in Chest Computed Tomography. *American Journal of Respiratory and Critical Care Medicine*, *197*(2), 193–203. https://doi.org/10.1164/RCCM.201705-0860OC

Graves, A., Mohamed, A. R., & Hinton, G. (2013). Speech Recognition with Deep Recurrent Neural Networks. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 6645–6649. https://doi.org/10.48550/arxiv.1303.5778.

Gu, Y., Wang, C., Di, Y., Zhang, Y., Wang, S., & Zhang, Y. (2012). Representative Multiple Kernel Learning for Classification in Hyperspectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(7). https://doi.org/10.1109/TGRS.2011.2176341

Gudadhe, M., Wankhade, K., & Dongre, S. (2010). Decision support system for heart disease based on support vector machine and artificial neural network. *2010 International Conference on Computer and Communication Technology, ICCCT-2010*, 741–745. https://doi.org/10.1109/ICCCT.2010.5640377

Guidi, G., Pettenati, M. C., Melillo, P., & Iadanza, E. (2014a). A machine learning system to improve heart failure patient assistance. *IEEE Journal of Biomedical and Health Informatics*, *18*(6), 1750–1756. https://doi.org/10.1109/JBHI.2014.2337752

Guidi, G., Pettenati, M. C., Melillo, P., & Iadanza, E. (2014b). A machine learning system to improve heart failure patient assistance. *IEEE Journal of Biomedical and Health Informatics*, *18*(6), 1750–1756. https://doi.org/10.1109/JBHI.2014.2337752

Guidi, G., Pettenati, M. C., Melillo, P., & Iadanza, E. (2014c). A machine learning system to improve heart failure patient assistance. *IEEE Journal of Biomedical and Health Informatics*, *18*(6), 1750–1756. https://doi.org/10.1109/JBHI.2014.2337752

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P. C., Mega, J. L., & Webster, D. R. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, *316*(22), 2402–2410. https://doi.org/10.1001/JAMA.2016.17216

Hajar, R. (2016). Framingham Contribution to Cardiovascular Disease. *Heart Views : The Official Journal of the Gulf Heart Association*, *17*(2), 78. https://doi.org/10.4103/1995-705X.185130

Han, J., & Moraga, C. (1995). The influence of the sigmoid function parameters on the speed of backpropagation learning. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *930*, 195–201. https://doi.org/10.1007/3-540-59497-3_175/COVER

Hancer, E., Xue, B., & Zhang, M. (2018). Differential evolution for filter feature selection based on information theory and feature ranking. *Knowledge-Based Systems*, *140*, 103–119. https://doi.org/10.1016/J.KNOSYS.2017.10.028

Hao, Y., Sheng, Y., & Wang, J. (2019). Variant Gated Recurrent Units With Encoders to Preprocess Packets for Payload-Aware Intrusion Detection. *IEEE Access*, *7*, 49985–49998. https://doi.org/10.1109/ACCESS.2019.2910860

Hassim, Y. M. M., & Ghazali, R. (2014). Optimizing functional link neural network learning using modified bee colony on multi-class classifications. *Lecture Notes in Electrical Engineering*, *279 LNEE*, 153–159. https://doi.org/10.1007/978-3-642-41674-3_23/COVER

Hastie, T., Tibshirani, R., & Friedman, J. (2017). *Springer Series in Statistics The Elements of Statistical Learning Data Mining, Inference, and Prediction*.

He, B., Guan, Y., & Dai, R. (2018). Convolutional Gated Recurrent Units for Medical Relation Classification. *Proceedings - 2018 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2018*, 646–650. https://doi.org/10.48550/arxiv.1807.11082

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(9), 1904–1916. https://doi.org/10.1109/TPAMI.2015.2389824

Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, *18*(7), 1527–1554. https://doi.org/10.1162/NECO.2006.18.7.1527

Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (2012). *Improving neural networks by preventing co-adaptation of feature detectors*. https://doi.org/10.48550/arxiv.1207.0580

Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2018). NSE Stock Market Prediction Using Deep-Learning Models. *Procedia Computer Science*, *132*, 1351–1362. https://doi.org/10.1016/J.PROCS.2018.05.050

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(8), 2554–2558. https://doi.org/10.1073/PNAS.79.8.2554

Hoppensteadt, F. C., & Izhikevich, E. M. (1999). Oscillatory Neurocomputers with Dynamic Connectivity. *Physical Review Letters*, *82*(14), 2983. https://doi.org/10.1103/PhysRevLett.82.2983

Hughes, M., Li, I., Kotoulas, S., & Suzumura, T. (2017). Medical Text Classification Using Convolutional Neural Networks. *Studies in Health Technology and Informatics*, *235*, 246–250. https://doi.org/10.3233/978-1-61499-753-5-246

Imani, M., & Ghassemian, H. (2015). Feature Extraction Using Weighted Training Samples. *IEEE Geoscience and Remote Sensing Letters*, *12*(7), 1387–1391. https://doi.org/10.1109/LGRS.2015.2402167

Irsoy, O. ˙, & Cardie, C. (2014). Deep Recursive Neural Networks for Compositionality in Language. *Advances in Neural Information Processing Systems*, 2096–2104.

Ivakhnenko, A. G. (1968). The group method of data handling, A rival of the method of stochastic approximation. *Soviet Automatic Control*, *13*(3), 43–55. https://cir.nii.ac.jp/crid/export-ddb6bf81-69b3-4a11-9679-4f243a1d05c7.bib

Ivakhnenko, A. G. (1971). Polynomial Theory of Complex Systems. *IEEE Transactions on Systems, Man and Cybernetics*, *1*(4), 364–378. https://doi.org/10.1109/TSMC.1971.4308320

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, Z., Abera, S. F., Abil, O. Z., Abraha, H. N., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., Accrombessi, M. M. K., … Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Javed Mehedi Shamrat, F. M., Abu Raihan, M., Sazzadur Rahman, A., Mahmud, I., & Akter, R. (2020). An Analysis On Breast Disease Prediction Using Machine Learning Approaches. *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, *9*, 2. www.ijstr.org

Javid, I., Zager Alsaedi, A. K., Ghazali, R., Mohmad Hassim, Y. M., & Zulqarnain, M. (2022). Optimally organized GRU-deep learning model with Chi 2 feature selection for heart disease prediction. *Journal of Intelligent & Fuzzy Systems*, *42*(4), 4083–4094. https://doi.org/10.3233/JIFS-212438

Jovic, A., Brkic, K., Bogunovic, N., A review of feature selection methods with applications, in: 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), May 2015, pp. 1200–1205, doi: 10.1109/MIPRO.2015.7160458 .

Justus, D., Brennan, J., Bonner, S., & McGough, A. S. (2018). Predicting the Computational Cost of Deep Learning Models. *Proceedings - 2018 IEEE International Conference on Big Data, Big Data 2018*, 3873–3882. https://doi.org/10.48550/arxiv.1811.11880

Kahramanli, H., & Allahverdi, N. (2008). Design of a hybrid system for the diabetes and heart diseases. *Expert Systems with Applications*, *35*(1–2), 82–89. https://doi.org/10.1016/J.ESWA.2007.06.004

Kaur, S., Aggarwal, H., & Rani, R. (2020). Hyper-parameter optimization of deep learning model for prediction of Parkinson’s disease. *Machine Vision and Applications 2020 31:5*, *31*(5), 1–15. https://doi.org/10.1007/S00138-020-01078-1

Kaushik, S., Choudhury, A., Jatav, A., Dasgupta, N., Natarajan, S., Pickett, L., & Dutt, V. (2019). Comparative Analysis of Features Selection Techniques for

Classification in Healthcare. *Lecture Notes in Computer Science*, *In press*(May).

Keogh, E., & Mueen, A. (2017). Curse of Dimensionality. *Encyclopedia of Machine Learning and Data Mining*, 314–315. https://doi.org/10.1007/978-1-4899-7687-1_192

Kim, H. C. (2021). Epidemiology of cardiovascular disease and its risk factors in Korea. *Global Health & Medicine*, *3*(3), 134. https://doi.org/10.35772/GHM.2021.01008

Koppel, A., Warnell, G., Stump, E., & Ribeiro, A. (2019). Parsimonious online learning with kernels via sparse projections in function space. *Journal of Machine Learning Research*, *20*, 1–44.

Krishnan, S., Magalingam, P., & Ibrahim, R. B. (2020). Advanced recurrent neural network with tensorflow for heart disease prediction. *International Journal of Advanced Science and Technology*, *29*(5 Special Issue), 966–977.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. 1097–1105. http://code.google.com/p/cuda-convnet/

Kumar, N., & Kumar, D. (2021). Machine Learning based Heart Disease Diagnosis using Non-Invasive Methods: A Review. Journal of Physics: Conference Series, 1950(1). https://doi.org/10.1088/1742-6596/1950/1/012081

Kuulasmaa, K., Tunstall-Pedoe, H., Dobson, A., Fortmann, S., Sans, S., Tolonen, H., Evans, A., Ferrario, M., & Tuomilehto, J. (2000). Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations. *The Lancet*, *355*(9205), 675–687. https://doi.org/10.1016/S0140-6736(99)11180-2

Kyu, H. H., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, M., Abebe, Z., Abil, O. Z., Aboyans, V., Abrham, A. R., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., … Murray, C. J. L. (2018). Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1859–1922. https://doi.org/10.1016/S0140-6736(18)32335-3

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature 2015 521:7553*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Li, Z., Wang, C., Han, M., Xue, Y., Wei, W., Li, L. J., & Fei-Fei, L. (2018). Thoracic Disease Identification and Localization with Limited Supervision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 8290–8299. https://doi.org/10.1109/CVPR.2018.00865

Linnainmaa, S. (1976). Taylor expansion of the accumulated rounding error. *BIT Numerical Mathematics 1976 16:2*, *16*(2), 146–160. https://doi.org/10.1007/BF01931367

Lipton, Z. C., Kale, D. C., Wetzel, R., & Whittier, L. K. (2016). *Modeling Missing Data in Clinical Time Series with RNNs*. https://doi.org/10.48550/arxiv.1606.04130

Liu, H., & Cocea, M. (2017). Semi-random partitioning of data into training and test sets in granular computing context. *Granular Computing*, *2*(4), 357–386. https://doi.org/10.1007/S41066-017-0049-2/FIGURES/11

Liu, H., & Motoda, H. (1998). Feature Extraction, Construction and Selection. *Feature Extraction, Construction and Selection*. https://doi.org/10.1007/978-1-4615-5725-8

Liu, L. (2016). Hierarchical learning for large multi-class network classification. *Proceedings - International Conference on Pattern Recognition*, *0*, 2307–2312. https://doi.org/10.1109/ICPR.2016.7899980

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11–26. https://doi.org/10.1016/J.NEUCOM.2016.12.038

Loganathan, S., Sujatha, C. M., Guru Nivash, R., Krish Srinivas, R., Niveddita, J., & Nivedha, V. (2022). Implementation of Convolutional Neural Network for Epileptic Seizure Detection. Proceedings of 2022 International Conference on Intelligent Innovations in Engineering and Technology, ICIIET 2022, 106–110. https://doi.org/10.1109/ICIIET55458.2022.9967535

Long, M., Zhu, H., Wang, J., & Jordan, M. I. (2017). Deep transfer learning with joint adaptation networks. *34th International Conference on Machine Learning, ICML 2017*, *5*, 3470–3479.

Louridi, N., Douzi, S., & El Ouahidi, B. (2021). Machine learning-based identification of patients with a cardiovascular defect. *Journal of Big Data*, *8*(1), 1–15. https://doi.org/10.1186/S40537-021-00524-9/TABLES/7

Lu, B., Fu, L., Nie, B., Peng, Z., & Liu, H. (2019). A Novel Framework with High Diagnostic Sensitivity for Lung Cancer Detection by Electronic Nose. *Sensors 2019, Vol. 19, Page 5333*, *19*(23), 5333. https://doi.org/10.3390/S19235333

Luo, L. xia. (2018). Network text sentiment analysis method combining LDA text representation and GRU-CNN. *Personal and Ubiquitous Computing 2018 23:3*, *23*(3), 405–412. https://doi.org/10.1007/S00779-018-1183-9

Madani, A., Arnaout, R., Mofrad, M., & Arnaout, R. (2018). Fast and accurate view classification of echocardiograms using deep learning. *Npj Digital Medicine 2018 1:1*, *1*(1), 1–8. https://doi.org/10.1038/s41746-017-0013-1

Majumdar, A., & Ward, R. (2017). Robust greedy deep dictionary learning for ECG arrhythmia classification. *Proceedings of the International Joint Conference on Neural Networks*, *2017-May*, 4400–4407. https://doi.org/10.1109/IJCNN.2017.7966413

Malhotra, P., Vishnu, T. V., Vig, L., Agarwal, P., & Shroff, G. (2017). TimeNet: Pre-trained deep recurrent neural network for time series classification. *ESANN 2017 - Proceedings, 25th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 607–612. https://doi.org/10.48550/arxiv.1706.08838

Mateus, B. C., Mendes, M., Farinha, J. T., Assis, R., & Cardoso, A. M. (2021). Comparing LSTM and GRU models to predict the condition of a pulp paper press. *Energies*, *14*(21), 1–21. https://doi.org/10.3390/en14216958

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics 1943 5:4*, *5*(4), 115–133. https://doi.org/10.1007/BF02478259

Mchugh, M. L. (2013). The Chi-square test of independence Lessons in biostatistics. *Biochemia Medica*, *23*(2), 143–152. https://doi.org/10.11613/BM.2013.018

Medhekar, D. S., Bote, M. P., & Deshmukh, S. D. (2013). Heart Disease Prediction System using Naive Bayes. *INTERNATIONAL JOURNAL OF ENHANCED RESEARCH IN SCIENCE TECHNOLOGY & ENGINEERING*, *2*(3). www.erpublications.com

Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T., & Masood, M. (2021). Prediction of Heart Disease Using Deep Convolutional Neural Networks. *Arabian Journal for Science and Engineering 2021 46:4*, *46*(4), 3409–3422. https://doi.org/10.1007/S13369-020-05105-1

Melillo, P., De Luca, N., Bracale, M., & Pecchia, L. (2013a). Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability. *IEEE Journal of Biomedical and Health Informatics*, *17*(3), 727–733. https://doi.org/10.1109/JBHI.2013.2244902

Melillo, P., De Luca, N., Bracale, M., & Pecchia, L. (2013b). Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability. *IEEE Journal of Biomedical and Health Informatics*, *17*(3), 727–733. https://doi.org/10.1109/JBHI.2013.2244902

Metwally, A. A., Yu, P. S., Reiman, D., Dai, Y., Finn, P. W., & Perkins, D. L. (2019). Utilizing longitudinal microbiome taxonomic profiles to predict food allergy via Long Short-Term Memory networks. *PLOS Computational Biology*, *15*(2), e1006693. https://doi.org/10.1371/JOURNAL.PCBI.1006693

Miotto, R., Li, L., & Dudley, J. T. (2016). Deep Learning to Predict Patient Future

Diseases from the Electronic Health Records. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9626*, 768–774. https://doi.org/10.1007/978-3-319-30671-1_66

Moorthy, U., & Gandhi, U. D. (2020). A novel optimal feature selection technique for medical data classification using ANOVA based whale optimization. *Journal of Ambient Intelligence and Humanized Computing 2020 12:3*, *12*(3), 3527–3538. https://doi.org/10.1007/S12652-020-02592-W

Murray, C. J. L., & Lopez, A. D. (2013). Measuring the Global Burden of Disease. *New England Journal of Medicine*, *369*(5), 448–457. https://doi.org/10.1056/NEJMRA1201534/SUPPL_FILE/NEJMRA1201534_DISCLOSURES.PDF

Murray, C. J. L., & Lopez, A. D. (2017). Measuring global health: motivation and evolution of the Global Burden of Disease Study. In *The Lancet* (Vol. 390, Issue 10100, pp. 1460–1464). Lancet Publishing Group. https://doi.org/10.1016/S0140-6736(17)32367-X

Nagappan, M., Sam, S., Sangeetha, S., Nithya Priya, S., Suguna, N., & Scholar, U. G. (2019). Heart Disease Prediction Using Data Mining Technique. *Shodhshauryam, International Scientific Refereed Research Journal © 2019 SHISRRJ*, *2*(10), 2581–6306. www.shisrrj.com

Nair, V., & Hinton, G. E. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *International Conference on Machine Learning*.

Narayan Panda, R., & Zaheera, F. (2019). Prediction of Heart Disease using Deep Convolutional Neural Networks. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *10*(3), 1141–1148. https://turcomat.org/index.php/turkbilmat/article/view/12724.

Nilashi, M., Ibrahim, O., & Bagherifard, K. (2018). A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques. *Expert Systems with Applications*, *92*, 507–520. https://doi.org/10.1016/j.eswa.2017.09.058.

Pal, A., & Singh, K. P. (2022). R-GRU: Regularized gated recurrent unit for handwritten mathematical expression recognition. *Multimedia Tools and Applications, 81*(22), 31405–31419. https://doi.org/10.1007/s11042-022-12889-x

Parthiban, G., & K. Srivatsa, S. (2012). Applying Machine Learning Methods in Diagnosing Heart Disease for Diabetic Patients. *International Journal of Applied Information Systems*, *3*(7), 25–30. https://doi.org/10.5120/ijais12-450593

Patel, S. B. (2016). Heart Disease Prediction using Machine Learning and Data Mining Techniques. *October*. https://doi.org/10.090592/IJCSC.2016.018

Pathak, & Kumar, A. (2020). *Feature selection using heuristic approaches for heart disease classification*. PhD. Thesis. Anna University; 2020.

Paul, A. K., Shill, P. C., Rabin, M. R. I., & Murase, K. (2018). Adaptive weighted fuzzy rule-based system for the risk level assessment of heart disease. *Applied Intelligence*, *48*(7), 1739–1756. https://doi.org/10.1007/s10489-017-1037-6

Pham, T., Tran, T., Phung, D., & Venkatesh, S. (2016). DeepCare: A deep dynamic memory model for predictive medicine. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9652 LNAI*, 30–41. https://doi.org/10.1007/978-3-319-31750-2_3/COVER/

Ponikowski, P., Voors, A. A., Anker, S. D., Bueno, H., Cleland, J. G. F., Coats, A. J. S., Falk, V., González-Juanatey, J. R., Harjola, V. P., Jankowska, E. A., Jessup, M., Linde, C., Nihoyannopoulos, P., Parissis, J. T., Pieske, B., Riley, J. P., Rosano, G. M. C., Ruilope, L. M., Ruschitzka, F., … Davies, C. (2016). 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of. *European Heart Journal*, *37*(27), 2129-2200m. https://doi.org/10.1093/EURHEARTJ/EHW128

RBC Capital Markets | The healthcare data explosion. (2021.). Retrieved December 23, 2022, from *https://www.rbccm.com/en/gib/healthcare/episode/the_* *healthcare_data_explosion*

Rahhal, M. M. A., Bazi, Y., Alhichri, H., Alajlan, N., Melgani, F., & Yager, R. R. (2016a). Deep learning approach for active classification of electrocardiogram signals. *Information Sciences*, *345*, 340–354. https://doi.org/10.1016/j.ins.2016.01.082

Rahhal, M. M. A., Bazi, Y., Alhichri, H., Alajlan, N., Melgani, F., & Yager, R. R. (2016b). Deep learning approach for active classification of electrocardiogram signals. *Information Sciences*, *345*, 340–354. https://doi.org/10.1016/j.ins.2016.01.082

Rajendran, G., Chitturi, B., & Poornachandran, P. (2018). Stance-In-Depth Deep Neural Approach to Stance Classification. *Procedia Computer Science*, *132*, 1646–1653. https://doi.org/10.1016/J.PROCS.2018.05.132

Rajeswari, D., & Thangavel, K. (2020). The Performance of Data Normalization Techniques on Heart Disease Datasets. *International Journal of Advanced Research in Engineering and Technology*, *11*(12), 2350–2357. https://doi.org/10.34218/IJARET.11.12.2020.222

Ramachandran, P., Zoph, B., & Le Google Brain, Q. V. (2017, October 16). Searching for Activation Functions. *6th International Conference on Learning Representations, ICLR 2018 - Workshop Track Proceedings*. https://doi.org/10.48550/arxiv.1710.05941

Richens, J. G., Lee, C. M., & Johri, S. (2020). Improving the accuracy of medical diagnosis with causal machine learning. *Nature Communications 2020 11:1*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-17419-7

Rodríguez, J., Prieto, S., & Ramírez López, L. J. (2019). A novel heart rate attractor for the prediction of cardiovascular disease. *Informatics in Medicine Unlocked*, *15*, 100174. https://doi.org/10.1016/J.IMU.2019.100174

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature 1986 323:6088*, *323*(6088), 533–536. https://doi.org/10.1038/323533a0

Samarawickrama, A. J. P., & Fernando, T. G. I. (2018). A recurrent neural network approach in predicting daily stock prices an application to the Sri Lankan stock market. *2017 IEEE International Conference on Industrial and Information Systems, ICIIS 2017 - Proceedings, 2018-Janua*, 1–6. https://doi.org/10.1109/ICIINFS.2017.8300345

Sankara Babu, B., Suneetha, A., Charles Babu, G., Jeevan Nagendra Kumar, Y., & Karuna, G. (2018). Medical disease prediction using Grey Wolf optimization and auto encoder based recurrent neural network. *Periodicals of Engineering and Natural Sciences (PEN)*, *6*(1), 229–240. https://doi.org/10.21533/PEN.V6I1.286

Sarra, R. R., Dinar, A. M., Mohammed, M. A., & Abdulkareem, K. H. (2022). Enhanced Heart Disease Prediction Based on Machine Learning and χ2 Statistical Optimal Feature Selection Model. *Designs*, *6*(5), 87. https://doi.org/10.3390/designs6050087

Seah, J. C. Y., Tang, J. S. N., Kitchen, A., Gaillard, F., & Dixon, A. F. (2019). Chest radiographs in congestive heart failure: Visualizing neural network learning. *Radiology*, *290*(3), 514–522. https://doi.org/10.1148/RADIOL.2018180887/RY180887SUPPM2.MP4

Septiawan, W. M., & Endah, S. N. (2018). Suitable Recurrent Neural Network for Air Quality Prediction with Backpropagation Through Time. *2018 2nd International Conference on Informatics and Computational Sciences, ICICoS 2018*, 196–201. https://doi.org/10.1109/ICICOS.2018.8621720

Shah, D., Patel, S., & Bharti, S. K. (2020). Heart Disease Prediction using Machine Learning Techniques. *SN Computer Science 2020 1:6*, *1*(6), 1–6. https://doi.org/10.1007/S42979-020-00365-Y

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. http://www.cs.huji.ac.il/~shais/Understanding MachineLearning

Sharif, W., Samsudin, A., Deris, M. M., & Aamir, M. (2017). Discriminative Criterion Feature Ranking Technique for Text Classification. *International Journal of*

Artificial Intelligence, *15*(2), 61–78.

Sharma, S., Sharma, S., & Athaiya, A. (2020a). Activation Functions in Neural Networks. *International Journal of Engineering Applied Sciences and Technology*, *04*(12), 310–316. https://doi.org/10.33564/ijeast.2020.v04i12.054

Sharma, S., Sharma, S., & Athaiya, A. (2020b). □ACTIVATION FUNCTIONS IN NEURAL NETWORKS. *International Journal of Engineering Applied Sciences and Technology*, *4*, 310–316. http://www.ijeast.com

Shatravin, V., Shashev, D., & Shidlovskiy, S. (2022). Sigmoid Activation Implementation for Neural Networks Hardware Accelerators Based on Reconfigurable Computing Environments for Low-Power Intelligent Systems. *Applied Sciences (Switzerland)*, *12*(10). https://doi.org/10.3390/app12105216

Shawe-Taylor, J., & Sun, S. (2011). A review of optimization methodologies in support vector machines. *Neurocomputing*, *74*(17), 3609–3618. https://doi.org/10.1016/J.NEUCOM.2011.06.026

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis. *IEEE Journal of Biomedical and Health Informatics*, *22*(5), 1589–1604. https://doi.org/10.1109/JBHI.2017.2767063

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature 2016 529:7587*, *529*(7587), 484–489. https://doi.org/10.1038/nature16961

Singh, R., & Rajesh, E. (2019). Prediction of Heart Disease by Clustering and Classification Techniques. *International Journal of Computer Sciences and Engineering*, *7*(5), 861–866. https://doi.org/10.26438/ijcse/v7i5.861866

Singhal, S., Kumar, H., & Passricha, V. (2018). Prediction of Heart disease using DNN. *American International Journal of Research in Science, Technology,* *Engineering & Mathematics*, *1*(23), 257–261.

Sivaranjani, R., & Yuvaraj, N. (2019). Artificial intelligence model for earlier prediction of cardiac functionalities using multilayer perceptron. *Journal of Physics: Conference Series*, *1362*(1). https://doi.org/10.1088/1742-6596/1362/1/012062

Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. *Empirical Methods in Natural Language Processing*, 1631–1642. http://nlp.stanford.edu/

Solé, X., Ramisa, A., & Torras, C. (2014). Evaluation of Random forests on large-scale classification problems using a bag-of-visual-words representation. *Frontiers in Artificial Intelligence and Applications*, *269*, 273–276. https://doi.org/10.3233/978-1-61499-452-7-273

Soman, K., Muralidharan, V., & Chakravarthy, V. S. (2018). An oscillatory neural autoencoder based on frequency modulation and multiplexing. *Frontiers in Computational Neuroscience*, *12*, 52. https://doi.org/10.3389/FNCOM.2018.00052/BIBTEX

Sowri, N., Pillai, R., Kamurunnissa Bee, K., & Kiruthika, J. (2019). PREDICTION OF HEART DISEASE USING RNN ALGORITHM. *International Research Journal of Engineering and Technology*. www.irjet.net

Spencer, R., Thabtah, F., Abdelhamid, N., & Thompson, M. (2020). Exploring feature selection and classification methods for predicting heart disease. *Digital Health*, *6*, 2055207620914777. https://doi.org/10.1177/2055207620914777

Srinivas, K., Raghavendra Rao, G., & Govardhan, A. (2010). Analysis of coronary heart disease and prediction of heart attack in coal mining regions using data mining techniques. *ICCSE 2010 - 5th International Conference on Computer Science and Education, Final Program and Book of Abstracts*, 1344–1349. https://doi.org/10.1109/ICCSE.2010.5593711

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014).

Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.

Sudhakar, K., & Manimekalai, D. (2014). *Study of Heart Disease Prediction using Data Mining | Semantic Scholar*. International Journal of Advanced Research in Computer Science and Software Engineering. https://www.semanticscholar.org/paper/Study-of-Heart-Disease-Prediction-using-Data-Mining-Sudhakar/1ef3c246e87ceb8a24e76ed5855ba3b4bae17ce9

Takci, H. (2018). Improvement of heart attack prediction by the feature selection methods. *Turkish Journal of Electrical Engineering and Computer Sciences*, *26*(1), 1–10. https://doi.org/10.3906/elk-1611-235

Tang, Y. X., Tang, Y. B., Peng, Y., Yan, K., Bagheri, M., Redd, B. A., Brandon, C. J., Lu, Z., Han, M., Xiao, J., & Summers, R. M. (2020). Automated abnormality classification of chest radiographs using deep convolutional neural networks. Npj Digital Medicine, 3(1). https://doi.org/10.1038/s41746-020-0273-z

Tariq, Z., Shah, S. K., & Lee, Y. (2022). Feature-Based Fusion Using CNN for Lung and Heart Sound Classification. Sensors, 22(4). https://doi.org/10.3390/s22041521

Teo, K. K., & Rafiq, T. (2021). Cardiovascular Risk Factors and Prevention: A Perspective From Developing Countries. Canadian Journal of Cardiology, 37(5), 733–743. https://doi.org/10.1016/j.cjca.2021.02.009

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*. https://doi.org/10.1038/s41591-018-0300-7

Townsend, N., Kazakiewicz, D., Lucy Wright, F. (2022) Epidemiology of cardiovascular disease in Europe. Nat Rev Cardiol 19, 133–143. https://doi.org/10.1038/s41569-021-00607-3

Trigka, M., & Dritsas, E. (2023). Long-Term Coronary Artery Disease Risk Prediction with Machine Learning Models. Sensors, 23(3). https://doi.org/10.3390/s23031193

Turian, J., Bergstra, J., & Bengio, Y. (2009). Quadratic Features and Deep Architectures for Chunking. *NAACL HLT*, 245–248.

UCI. 2020. UCI Machine Learning Repository: Heart Disease Data Set. [online] Available at: http://archive.ics.uci.edu/ml/datasets/heart+disease [Accessed 20 Jun. 2020].

Unas, A. , Deepmind, G. G., Munos, R., Deepmind, G., Danihelka, I., Lanctot, M., & Graves, A. (2016). Memory-Efficient Backpropagation Through Time. *Advances in Neural Information Processing Systems*, *29*.

Urbanowicz, R. J., Meeker, M., La Cava, W., Olson, R. S., & Moore, J. H. (2018). Relief-based feature selection: Introduction and review. Journal of Biomedical Informatics, 85(July), 189–203. https://doi.org/10.1016/j.jbi.2018.07.014

V.Sabarinathan, & V.Sugumaran. (2016). Prediction of Heart Disease Using Decision Tree. *International Journal of Advanced Research in Computer Science and Software Engineering*, *6*(3), 530–532.

Vembandasamyp, K., Sasipriyap, R. R., & Deepap, E. (2015). Heart Diseases Detection Using Naive Bayes Algorithm. *IJISET-International Journal of Innovative Science, Engineering & Technology*, *2*(9). www.ijiset.com

Verma, L., Srivastava, S., & Negi, P. C. (2018). An intelligent noninvasive model for coronary artery disease detection. *Complex & Intelligent Systems*, *4*(1), 11–18. https://doi.org/10.1007/s40747-017-0048-6

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, *2018*. https://doi.org/10.1155/2018/7068349

Wang, D. L. L., & Brown, G. J. (1999). Separation of speech from interfering sounds based on oscillatory correlation. *IEEE Transactions on Neural Networks*, *10*(3), 684–697. https://doi.org/10.1109/72.761727

Wang, D. L., & Terman, D. (1995). Locally Excitatory Globally Inhibitory Oscillator Networks. *IEEE Transactions on Neural Networks*, *6*(1), 283–286.

https://doi.org/10.1109/72.363423

Wang, D., & Terman, D. (1997). Image segmentation based on oscillatory correlation. *Neural Computation*, *9*(4), 805–836. https://doi.org/10.1162/NECO.1997.9.4.805

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*, 3462–3471. https://doi.org/10.1109/CVPR.2017.369

Wang, X., Xu, J., Shi, W., & Liu, J. (2019). OGRU: An Optimized Gated Recurrent Unit Neural Network. *Journal of Physics: Conference Series*, *1325*(1). https://doi.org/10.1088/1742-6596/1325/1/012089

Wang, Z., & Qu, Z. (2018). Research on Web text classification algorithm based on improved CNN and SVM. *International Conference on Communication Technology Proceedings, ICCT*, *2017-Octob*, 1958–1961. https://doi.org/10.1109/ICCT.2017.8359971

Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. *Doctoral Dissertation, Harvard University*. https://cir.nii.ac.jp/crid/export-db42a68b-3acd-4df4-9951-4b0db420c7f9.bib?lang=en

Werbos, P. J. (1981). Applications of advances in nonlinear sensitivity analysis. *System Modeling and Optimization*, 762–770. https://doi.org/10.1007/BFB0006203

Wettschereck, D., & Dietterich, T. G. (1995). An experimental comparison of the nearest-neighbor and nearest-hyperrectangle algorithms. *Machine Learning*, *19*(1), 5–27. https://doi.org/10.1007/bf00994658

World Bank. (1993). *World development report 1993 : investing in health.* Published for the World Bank, Oxford University Press.

Wu, Q., Teney, D., Wang, P., Shen, C., Dick, A., & van den Hengel, A. (2017). Visual question answering: A survey of methods and datasets. *Computer Vision and Image Understanding*, *163*, 21–40. https://doi.org/10.1016/J.CVIU.2017.05.001

Xia, B., Li, Q., Jia, J., Wang, J., Chaudhary, U., Ramos-Murguialday, A., & Birbaumer, N. (2015). Electrooculogram based sleep stage classification using deep belief network. *Proceedings of the International Joint Conference on Neural Networks*, *2015-Septe*. https://doi.org/10.1109/IJCNN.2015.7280775

Xie, R., Hao, K., Huang, B., Chen, L., & Cai, X. (2020). Data-Driven Modeling Based on Two-Stream Gated Recurrent Unit Network With Soft Sensor Application; Data-Driven Modeling Based on Two-Stream <inline-formula><tex-math. *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, *67*(8). https://doi.org/10.1109/TIE.2019.2927197

Xing, Y., & Xiao, C. (2019). A GRU Model for Aspect Level Sentiment Analysis. *Journal of Physics: Conference Series*, *1302*(3). https://doi.org/10.1088/1742-6596/1302/3/032042

Yan, Z., Yang, K., Wang, Z., Yang, B., Kaizuka, T., & Nakano, K. (2019). Time to lane change and completion prediction based on Gated Recurrent Unit Network. *IEEE Intelligent Vehicles Symposium, Proceedings*, *2019-June*, 102–107. https://doi.org/10.1109/IVS.2019.8813838

Yang, M. S., & Nataliani, Y. (2018). A Feature-Reduction Fuzzy Clustering Algorithm Based on Feature-Weighted Entropy. *IEEE Transactions on Fuzzy Systems*, *26*(2), 817–835. https://doi.org/10.1109/TFUZZ.2017.2692203

Yang, X., Chen, A., PourNejatian, N., Shin, H. C., Smith, K. E., Parisien, C., Compas, C., Martin, C., Costa, A. B., Flores, M. G., Zhang, Y., Magoc, T., Harle, C. A., Lipori, G., Mitchell, D. A., Hogan, W. R., Shenkman, E. A., Bian, J., & Wu, Y. (2022). A large language model for electronic health records. Npj Digital Medicine, 5(1), 1–9. https://doi.org/10.1038/s41746-022-00742-2

Zafar, M., Sharif, M. I., Sharif, M. I., Kadry, S., Bukhari, S. A. C., & Rauf, H. T. (2023). Skin Lesion Analysis and Cancer Detection Based on Machine/Deep

Learning Techniques: A Comprehensive Survey. Life, 13(1), 1–18. https://doi.org/10.3390/life13010146

Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks arXiv:1311.2901v3 [cs.CV] 28 Nov 2013. *Computer Vision–ECCV 2014*, *8689*(PART 1), 818–833. https://doi.org/10.1007/978-3-319-10590-1_53

Zhang, J., Liu, F., Xu, W., & Yu, H. (2019). Feature Fusion Text Classification Model Combining CNN and BiGRU with Multi-Attention Mechanism. *Future Internet 2019, Vol. 11, Page 237*, *11*(11), 237. https://doi.org/10.3390/FI11110237

Zhang, N., Yang, G., Gao, Z., Xu, C., Zhang, Y., Shi, R., Keegan, J., Xu, L., Zhang, H., Fan, Z., & Firmin, D. (2019). Deep Learning for Diagnosis of Chronic Myocardial Infarction on Nonenhanced Cardiac Cine MRI. *Radiology*, *291*(3), 606–607. https://doi.org/10.1148/RADIOL.2019182304

Zhang, Q., Yang, L. T., Chen, Z., & Li, P. (2018). A survey on deep learning for big data. *Information Fusion*, *42*, 146–157. https://doi.org/10.1016/J.INFFUS.2017.10.006

Zhang, R., Ma, S., Shanahan, L., Munroe, J., Horn, S., & Speedie, S. (2017). Automatic methods to extract New York heart association classification from clinical notes. *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017*, *2017-Janua*, 1296–1299. https://doi.org/10.1109/BIBM.2017.8217848

Zheng, W. L., Zhu, J. Y., Peng, Y., & Lu, B. L. (2014). EEG-based emotion classification using deep belief networks. *Proceedings - IEEE International Conference on Multimedia and Expo*, *2014-Septe*(Septmber). https://doi.org/10.1109/ICME.2014.6890166

Zheng, Y., Liu, Q., Chen, E., Ge, Y., & Zhao, J. L. (2014). Time series classification using multi-channels deep convolutional neural networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8485 LNCS*, 298–310. https://doi.org/10.1007/978-3-319-08010-9_33/COVER/

Zhu, X., Zheng, W. L., Lu, B. L., Chen, X., Chen, S., & Wang, C. (2014). EOG-based drowsiness detection using convolutional neural networks. *Proceedings of the International Joint Conference on Neural Networks*, 128–134. https://doi.org/10.1109/IJCNN.2014.6889642.

Zulqarnain, M. (2020). *An Enhanced Gated Recurrent Unit with Auto-Encoder for solving Text Classification Problems*. PhD. Thesis. Universiti Tun Hussein Onn Malaysia.

Zulqarnain, M., Ghazali, R., Hassim, Y. M. M., & Rehan, M. (2020). Text classification based on gated recurrent unit combines with support vector machine. *International Journal of Electrical and Computer Engineering*, *10*(4), 3734–3742. https://doi.org/10.11591/ijece.v10i4.pp3734-3742.

VITA

The author of this thesis is Irfan Javid. He was born in October 17, 1987, in Azad Kashmir, Pakistan. He went to Umar Boys High School, Rawalakot, AJK, Pakistan for his secondary school. He pursued his degree at the University of Azad Jammu & Kashmir Muzaffarabad, Pakistan, and graduated with the BS (Hons) in Computer Science in 2009. Upon graduation, he worked as Web Designer in Pakistan Engineering Council, Islamabad, Pakistan. He then enrolled at the Bahria University Islamabad, Pakistan, in 2010, where he was awarded the MS in Software Engineering in 2012. Thereafter, he taught Computer Science as well as Software Engineering at the Department of Computer Science & IT, University of Poonch Rawalakot, Pakistan. He did his PhD study at the faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia (UTHM), Johor, Malaysia, in order to discover his interest for research and advance expertise in machine learning and deep learning. He has also published a number of research articles and conference proceedings. His area of interest in study include artificial intelligence, medicine, machine learning, deep learning, and disease prediction.