

**SIMULATION OF A FEDERATED DEEP LEARNING APPROACH FOR
PREDICTING COVID-19 SEVERITY OF A PATIENT**

JUWE, JESSE IKECHUKWU
(19PCG02029)

B.Sc., Computer Science, Caleb University, Imota, Lagos State

n

SEPTEMBER, 2021

**SIMULATION OF A FEDERATED DEEP LEARNING APPROACH FOR
PREDICTING COVID-19 SEVERITY OF A PATIENT**

BY

**JUWE, JESSE IKECHUKWU
(19PCG02029)**

B.Sc., Computer Science, Caleb University, Imota, Lagos State

**A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE
STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
AWARD OF MASTER OF SCIENCE (M.Sc.) DEGREE IN COMPUTER SCIENCE IN
THE DEPARTMENT OF COMPUTER AND INFORMATION SCIENCE, COLLEGE
OF SCIENCE AND TECHNOLOGY, COVENANT UNIVERSITY.**

SEPTEMBER, 2021

ACCEPTANCE

This is to attest that this dissertation was accepted in partial fulfillment of the requirements for the award of Master of Science (M.Sc) degree in Computer Science in the Department of Computer and Information Science, College of Science and Technology, Covenant University, Ota, Ogun State, Nigeria.

Mr. John A. Philips
(Secretary, School of Postgraduate Studies)

Signature and Date

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies)

Signature and Date

DECLARATION

I hereby declare that this dissertation titled **SIMULATION OF A FEDERATED DEEP LEARNING APPROACH FOR PREDICTING COVID-19 SEVERITY IN PATIENTS** was carried out by **JUWE, JESSE IKECHUKWU** with matriculation number **19PCG02029**. This work is centered on an original study in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, under the supervision of Prof. Olufunke O. Oladipupo. The concepts for this research project are the outcome of my own research. Other scholars' ideas have also been completely acknowledged.

JUWE, JESSE IKECHUKWU

Signature and Date

CERTIFICATION

This is to certify that this dissertation titled **SIMULATION OF A FEDERATED DEEP LEARNING APPROACH FOR PREDICTING COVID-19 SEVERITY OF A PATIENT** was carried out by **JUWE, JESSE IKECHUKWU** and was supervised by and submitted to the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Ogun State.

Prof. Olufunke O. Oladipupo
(Supervisor)

Signature and Date

Prof. Olufunke O. Oladipupo
(Head of Department)

Signature and Date

Dr. Victor T. Odumuyiwa

Signature and Date

Prof. Akan B. Williams
(Dean, School of Postgraduate Studies)

Signature and Date

DEDICATION

I dedicate this project to God Almighty, who has provided me with abundant grace, wisdom, and knowledge throughout the completion of my Master's Degree program.

ACKNOWLEDGEMENTS

First and foremost, I want to express my gratitude to God Almighty for providing me with the opportunity to complete this program successfully and on schedule. Second, I want to express my gratitude to my great parents, Mr. and Mrs. Juwe, for their unwavering physical, financial, spiritual, and moral support from the beginning to the end of the program.

My gratitude goes to Covenant University's Chancellor, Dr. David Oyedepo, for accepting Covenant University's vision and mission, which has now become a reality, and to Covenant University's entire Management for ensuring that the vision has been realized by achieving great excellence through the master's academic program.

I particularly would like to appreciate Prof. Olufunke O. Oladipupo, the Head of the Department of Computer and Information Sciences (CIS), who also happens to be my amazing supervisor. I am grateful to her for her patience, encouragement, guidance, and supervision throughout my master's program. Special recognition for Dr. Aderonke Oni, the PG Coordinator, and the entire faculty for their insightful observations during the several presentations we had during the course of this project.

Finally, I want to express my gratitude to my friends; Oladapo Alagbe, Ayobami Shofadekan, and my great colleagues for their tremendous contributions to my project. I appreciate everything you've done for me, and may God continue to bless you.

TABLE OF CONTENTS

ACCEPTANCE	iii
DECLARATION	iv
CERTIFICATION	v
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
LIST OF FIGURES	x
LIST OF TABLES	xii
ABSTRACT	xiii
CHAPTER ONE: INTRODUCTION	1
1.1. Background Information	1
1.2. Statement of the Problem	3
1.3. Aim and Objectives of the Study	3
1.4. Research Methodology	4
1.5. Significance of the Study	5
1.6. Motivation	5
1.7. Scope of the Research	6
CHAPTER TWO: LITERATURE REVIEW	7
2.1. Introduction	7
2.2. Covid-19	7
2.3. Covid-19 Prediction with Artificial Intelligence Techniques	8
2.4. Traditional Machine Learning in Medical Prediction	8
2.5. Electronic Health Records and Artificial Intelligence in Medical Care	9
2.6. Federated Learning	10
2.6.1. Benefits of Using Federated Learning	13
2.6.2. Types of Federated Learning	14
2.6.3. Categorization of Federated Learning Techniques	17
2.6.4. Challenges in Federated Learning	27
2.6.5. Applications of Federated Learning	28
2.6.6. Federated Learning Process	30
2.6.7. Federated Averaging Algorithm (Fedavg)	32
2.6.8. Open-Source Frameworks in Federated Learning	33
2.7. Federated Learning in Medical Prediction	34
2.8. Deep Learning	35

2.9. Related Works	36
CHAPTER THREE: RESEARCH METHODOLOGY	38
3.1. Introduction	38
3.2. Adopted Tensorflow Federated Model	39
3.3. Algorithm of Adopted Tensorflow Federated Model	40
3.4. Development of the Models	42
3.4.1. Deep Learning Model	42
3.4.2. Federated Learning Model	43
3.5. Methodology Workflow	45
3.5.1. Data Collection and Preprocessing	45
3.5.2. Model Training and Distribution	47
3.5.3. Model Aggregation	48
3.5.4. Model Evaluation	48
3.6. Visualization Report Findings	49
CHAPTER FOUR: RESULT AND DISCUSSION	50
4.1. Introduction	50
4.2. Screenshot of Code Snippets	51
4.2.1. Data Collection and Preprocessing	51
4.2.2. Model Training and Distribution	53
4.2.3. Model Aggregation	54
4.2.4. Model Evaluation	54
4.3. Model Evaluation Result	63
4.4. Visualization	64
4.5. Result Interpretations	66
CHAPTER FIVE: CONCLUSION AND RECOMMENDATION	69
5.1. Summtry	69
5.2. Conclusion	70
5.3. Contribution to Knowledge	70
5.4. Recommendations	71
REFERENCES	72
APPENDIX	78

LIST OF FIGURES

Figures	Title of Figures	Pages
2.1:	Diagrammatic Representation of Federated Learning	11
2.2:	Diagrammatic representation of Traditional Machine Learning Architecture	12
2.3:	Diagrammatic representation of Federated Learning Architecture	12
2.4:	Cross-device Federated Learning	15
2.5:	Cross-silo Federated Learning	16
2.6:	Fully-decentralized in P2P Topology	19
2.7:	Vertical Federated Learning	21
2.8:	Horizontal Federated Learning	22
2.9:	Federated Transfer Learning	23
2.10:	Federated Learning Process Flow	32
2.11:	Architecture of TensorFlow Federated (TFF)	33
2.12:	Architecture of FATE	34
2.13:	Architecture of PySyft	34
3.1:	Workflow of Methodology	38
3.2:	Federated Learning Framework	39
3.3:	Activity Diagram of FL Approach	41
4.1:	Jupyter Notebook (run from Anaconda)	50
4.2:	Google Colab	51
4.3:	Screenshot of dataset sample	51
4.4:	Setting up the TFF environment	52
4.5:	Importing necessary libraries	52
4.6:	Checking for missing values	52
4.7:	Summary of dataset	52
4.8:	Correlation map of dataset	53
4.9:	Federating the dataset	53
4.10:	Distributing model to clients for training	53
4.11:	Federated Averaging Algorithm	54
4.12:	Accuracy of TFF model in TensorBoard	54

4.13:	AUC of TFF model in TensorBoard	55
4.14:	Precision of TFF model in TensorBoard	55
4.15:	Recall of TFF model in TensorBoard	56
4.16:	Loss of TFF model in TensorBoard	56
4.17:	Screenshot of the homepage (Jupyter Notebook)	57
4.18:	Screenshot of the homepage (Google Colab)	57
4.19:	Snippet of code for loading TensorBoard and importing libraries	58
4.20:	Snippet of code for loading dataset	58
4.21:	Snippet of code for creating federated dataset from loaded dataset	59
4.22:	Snippet of code for creating federated dataset from loaded dataset	59
4.23:	Snippet of code for creating federated dataset from loaded dataset	60
4.24:	Snippet of code for creating federated dataset from loaded dataset	60
4.25:	Snippet of code for creating a model with TensorFlow Keras	61
4.26:	Snippet of code for training the model on federated data	61
4.27:	Snippet of code for displaying model metrics in TensorBoard	62
4.28:	Count distribution of the labels	64
4.29:	Count distribution of countries in dataset	64
4.30:	Value count group of class “Severity_Severe”	65
4.31:	Histogram of dataset	65

LIST OF TABLES

Table	Title of Tables	Pages
2.1:	Categorization of Federated Learning	17
3.1:	Hardware Requirements of the Deep Learning Model Development Tool	43
3.2:	Software Requirements of the Deep Learning Model Development Tool	43
3.3:	Hardware Requirements of the Federated Learning Model Development Tool	44
3.4:	Software Requirements of the Federated Learning Model Development Tool	45
3.5:	Summary of Dataset Features	46
4.1:	Evaluation Metrics of TFF Model After N Training Rounds	63
4.2:	Evaluation Metrics of TFF Model Against Deep Learning Model	63

ABSTRACT

The new SARS-COV-2, also known as the Coronavirus or COVID-19, is a contagious virus that causes respiratory issues and symptoms such as cough, shortness of breath, fever, pain, weariness, and parosmia (loss or altering of sense of smell and taste). This virus was responsible for the COVID-19 pandemic outbreak in 2019 and since then several methods have been suggested to prevent the re-occurrence of the outbreak. Predicting COVID-19 severity of patient is an important aspect of the solutions suggested for providing better health care and diagnosis. One limitation of the predictive solutions is the data availability because, for security and personal reason not every data can be collected. However, most predictive algorithms do not perform optimally with small training dataset which in turn affect the accuracy of the deployed predictive model. In bridging this gap, a number of federated deep learning approaches for battling COVID-19 in various aspects do exist in literature with medical images like chest CT scans and blood works but yet not on predicting COVID-19 patient severity using dataset. Hence, this study aims at investigating the performance of federated deep learning approach for building machine learning predictive model for determining COVID-19 severity in patients with non image datasets spread across several endpoints or institutions so as to handle scarcity of data, without compromising the security of private data. For this study, data repositories like Kaggle, was beneficial in getting public dataset for building the initialized model. Computations using the TensorFlow Federated (TFF) framework was used to implement the dispersion of the initialized model to the participants for training the model, and client computations was done using Stochastic Gradient Descent (SGD). A Federated Averaging Algorithm (FedAVG) was used on the server to aggregate and average the individual updates received from clients. For evaluation, the global model was benchmarked against a deep learning model (ANN) and a traditional machine learning model (RandomForest) using evaluation metrics such as: Accuracy, AUC, Precision, Recall, F-Measure. The Deep Learning Model (ANN) had a good accuracy of 90% which performed better than the RandomForest algorithm, having an accuracy of 87%. This just goes to show how well Deep Learning models handle complex datasets better than regular machine learning models. Also, the Deep Learning model spent less time during computation than the traditional machine learning algorithm, RandomForest. The TFF model had an accuracy of 91%, which was slightly higher than the Deep Learning model, according to the results of the training done on all three models.

Keywords: *COVID-19, Federated Learning, TensorFlow Federated, Deep Learning*