University of Kerbala
College of Computer Science & Information Technology
Computer Science Department

Intelligent Decision Support Model For Healthcare Data Management Based on Fog Computing

A Thesis
Submitted to the Council of the College of Computer Science and Information Technology / University of Kerbala in Partial Fulfillment of the Requirements for the Master Degree in Computer Science

Written by :
Oras Abdulkhudhur Hussein Radi

Supervised by :
Assist. Prof. Dr. Ayad Hameed Mousa

2023 A.D. 1444 A.H.

بسم الله الرحمن الرحيم

(وَيَسْأَلُونَكَ عَنِ الرُّوحِ ۖ قُلِ الرُّوحُ مِنْ أَمْرِ رَبِّي وَمَا أُوتِيتُم مِّنَ الْعِلْمِ إِلَّا قَلِيلًا)

صدق الله العلي العظيم

(سورة الاسراء آية 85)

Supervisor Certification

I certify that the thesis entitled (**Intelligent Decision Support Model for Healthcare Data Management Based on Fog Computing**) was prepared under my supervision at the department of Computer Science/College of Computer Science & Information Technology/ University of Kerbala as partial fulfillment of the requirements of the degree of Master in Computer Science.

Signature:

Supervisor Name: Assist. Prof. Dr. Ayad Hameed Mousa

Date: 5 / 3 /2023

The Head of the Department Certification

In view of the available recommendations, I forward the thesis entitled "Intelligent Decision Support Model for Healthcare Data Management Based on Fog Computing" for debate by the examination committee.

Signature:

Assist. Prof. Dr. Muhannad Kamil Abdulhameed

Head of Computer Science Department

Date: 5 /3 /2023

Certification of the Examination Committee

We hereby certify that we have studied the dissertation entitled (**Intelligent Decision Support Model for Emergency Management Based on Fog Computing**) presented by the student (**Oras Abdulkhudhur Hussein**) and examined her in its content and what is related to it, and that, in our opinion, it is adequate with (**Very Good**) standing as a thesis for the degree of Master in Computer Science.

Signature:
Name: Mehdi Ebady Manaa
Title: Assist. Prof. Dr.
Date: 27/ 2/ 2023
(Chairman)

Signature:
Name: Elham Mohammed Thabit
Title: Assist. Prof. Dr.
Date: 12 / 2 / 2023
(Member)

Signature:
Name: Akqeel Abdulkareem Farhan
Title: Assist. Prof. Dr.
Date: 20/2 / 2023
(Member)

Signature:
Name: Ayad Hameed Mousa
Title: Assist. Prof. Dr.
Date: 12 / 2 / 2023
(Member and Supervisor)

Approved by the Dean of the College of Computer Science & Information Technology, University of Kerbala.

Signature:
Title: Assist. Prof. Dr.
Name: Ahmed Abdulhadi Ahmed
Date: 12 / 3 / 2023
(Dean of Collage of Computer Science & Information Technology)

Dedication

To the memory of My Father.

Acknowledgement

I would like to praise and thank **Allah** for his many gifts to me and for giving me the patience and strength to complete my work with honesty, diligence and progress.

And I would like to thank my supervisor, Assist. Prof. Dr. (**Ayad Hameed Mousa**) for his continuous guidance to my work with accuracy, wisdom, and eloquence, he has been my teacher for a long time and he always gives me confidence, strength, and makes me Special, I can't stop thanking **My family**, **My mother**, **My brothers**, **My sisters**, and I would like to produce special thanks to my gorgeous nephews (**NoorAlhuda, and Nabaa**). And all **My loving friends** who always supported me.

Oras Abdulkhudhur Hussein

2023

Abstract

Recently, several healthcare emergencies have occurred globally, which has required and sometimes forced relevant decision-makers to pay overmuch attention to the decision-making process. In the Decision Support System (DSS) architecture, data is considered a backbone component due to DSS bringing together data and knowledge from different areas and sources, processing them, and storing them in the cloud to provide relevant users with meaningful information to support the decision-making process. Hence, DSS deals with cloud data and conventional ways of warehousing huge data in the cloud have a limitation, dealing with such massive volumes of data increases mistakes, packet drop, and the likelihood of data bottlenecks. As a result, it is suitable and essential to use Machine Learning (ML) algorithms in the design and development of an intelligent decision support system in the context of fog-computing. This research proposed an intelligent model for a decision support system in healthcare data management based on fog computing environment (IDSMFOG) using (K-means+ SVM) algorithms, and design instrument for measuring usability of the proposed model named questionnaire usability for DSS (QU-DSS). A DSS based on the proposed model was developed, and later it was also evaluated in terms of usability by QU-DSS instrument. The collected results confirmed that the accuracy of the developed system is (99.87 for USA example, and 99.97 for Iraq example). As well as the results of QU-DSS confirmed that the developed system is usable and support the decision-making process.

Declaration Associated with this Thesis

1. Congratulations - your paper #1570847145 ('IDSMFOG: An Intelligent Decision Support Model for Emergency Management in the F') for ICDSIC2022 has been **accepted** .from icdsic@uokerbala.edu.iq ,icdsic2022-chairs@edas.info

2. Paper ID# 29276, It is my great pleasure to inform you that your paper entitled **"Toward Measuring the Usability of Decision Support Applications in Fog Computing Environment"** has been initially **ACCEPTED** and will be PUBLISHED on the (https://www.scopus.com/so Indonesian Journal of Electrical Engineering and Computer Science, a Scopus urceid/21100799500) and ScimagoJR (https://www.scimagojr.com/journalsearch.php?q=21100799500&tip=sid&clean=0) indexed journal. **Congratulations**!

Table of Contents

Dedication i
Acknowledgement ii
Abstract iii
Table of Contents v
List of Tables vii
List of Figures viii
List of Abbreviations ix

CHAPTER ONE INTRODUCTION 11

1.1 Introduction 12
1.2 Problem Statement 13
1.3 The Aim of the Thesis 14
1.4 Research Contribution 14
1.5 Related Work 15
1.6 Thesis Organization 21

CHAPTER TWO THEORETICAL BACKGROUND 23

2.1 introduction 24
2.2 DSS's Brief History 25
2.3 Taxonomies of DSS 26
2.4 Architectures of DSS 27
2.5 Intelligent Decision Support System 30
2.6 Fog Computing 31
2.8 Data Management 34
2.9 Data Management in Fog Computing 39
2.9.1 Data Gathering 39
2.9.1.1 Data Converting 40
2.9.1.2 Data Annotation 40
2.9.2 Data Preprocessing 40
2.9.2.1 Data Fusion and Integration 40
2.9.2.2 Data Cleaning 41
2.9.2.3 Data Normalization 41
2.9.3 Artificial Intelligence (AI) in Fog Computing 42
2.9.3.1 Support Vector Machines (SVM) 45

2.9.3.2 K-means ... 46
2.9.3.3 Evaluation Performance Metrics ... 46
2.9.3.4 Network Performance Metrics ... 49
2.9.3.5 Data Visualization ... 49
2.10 Usability Measurement ... 49
CHAPTER THREE PROPOSED METHODOLOGY ... 51
3.1 introduction ... 52
3.2 Model Development ... 52
3.2.1 Data Gathering for Decision Support ... 54
3.2.1.1 Offline Data ... 54
3.2.2 Data Preprocessing for Decision Support ... 56
3.2.3 Intelligent Decision Support System (IDSS) ... 56
3.3 Development of the System ... 57
3.3.1 Pseudocode for the System (IDSMFOG) ... 58
3.4. Instrument Design and Development ... 60
CHAPTER FOUR RESULTS AND DISCUSSION ... 61
4.1 introduction ... 62
4.2 K-SVM Model ... 62
4.2.1 K-SVM Building Results ... 63
4.3 IDSMFOG Home Page ... 65
4.4 QU-DSS Design Findings ... 67
4.4.1 Pilot Test: Measuring Instrument Goodness ... 69
4.4.1.1 Subject Selection ... 69
4.4.1.2 QU-DSS Validation ... 69
4.4.1.3 The Factor Analysis Procedure ... 70
4.5 Use QU-DSS for Usability Measurement ... 73
CHAPTER FIVE CONCLUSION AND FUTURE WORK ... 80
5.1 Introduction ... 81
5.2 Conclusion ... 81
5.3 Future Work ... 82
REFERENCESS ... 83

List of Tables

Table 1.1: Related Works……………..…………………………………………….……19
Table 4. 1:The proposed Instrument (First Draft)..67
Table 4. 2: Finding of Measuring Reliability...70
Table 4. 3: The KMO Overall Findings..71
Table 4. 4: The Factor Loading Overall Findings...71
Table 4. 5: Usefulness Dimension Findings 74
Table 4. 6: Simplicity Dimension Findings 75
Table 4. 7:Reliability Dimension Findings 76
Table 4. 8: Decision-Support Dimension Findings..77
Table 4. 9: Flexibility Dimension Findings 78

List of Figures

Figure 2.1:DSS's Brief History Adapted from 26
Figure 2. 2 :The Five Distinct Components of a DSS 28
Figure 2 .3: Architecture of a Decision Support System 30
Figure 2. 4: Fog Computing 32
Figure 2. 5: A Three-layer Fog Diagram with A Data View 37
Figure 2. 6 : Types of ML 44
Figure 2 .7 : SVM classifier Illustration 45
Figure 2. 8: The confusion matrix...48
Figure 3 .1 The Own Proposed IDSMFOG Model for Evaluation online / Offline Data…...53
Figure 3. 2: Datasets Source Used Throughout the p-IsDSMFOG...55
Figure 3 .3 Data after Annotation. 56
Figure 3. 4: Evaluation Instrument Design (Main Phases) 60
Figure 4. 1: Accuracy of SVM with sample of Iraqi Cites Cases and Deaths. 63
Figure 4. 2: Decision Making (unlockdown) Iraqi Cities 64
Figure 4. 3: Accuracy of SVM with sample of USA Cites Cases and Deaths........................64
Figure 4. 4: Decision Making (unlock down) USA Cities..65
Figure 4. 5 IDSMFOG Home Page with weekly Growth Cases of COVID-19. 66
Figure 4. 6: IDSMFOG Home Page with Daily Growth Cases of COVID-19.......................66
Figure 4. 7: Usefulness Dimension Findings..74
Figure 4. 8: Simplicity Dimension Findings...75
Figure 4. 9: Reliability Dimension Findings...76
Figure 4. 10: Decision-Support Dimension Findings 77
Figure 4. 11: Flexibility Dimension Findings 78

List of Abbreviations

Abbreviation	Description
AI	**A**rtificial **I**ntelligence
ANN	**A**rtificial **N**eural **N**etworks
CDS	**C**ashing **D**ecision **S**ystem
DBMS	**D**ata **B**ase **M**anagement **S**ystem
DL	**D**eep **L**earning
DSS	**D**ecision **S**upport **S**ystem
FLX	**Flex**ibility
IDSMFOG	**I**ntelligent **D**ecision **S**upport **M**odel based on **Fog** Computing Environment
IDSS	**I**ntelligent **D**ecision **S**upport **S**ystem
IoT	**I**nternet **o**f **T**hing
KMO	**K**aiser-**M**eyer-**O**lkin
K-SVM	**K-**means with **SVM** algorithms
M	**M**ean

Abbreviation	Description
MBMS	**M**odel-**b**ased **M**anagement **S**ystem
ML	**M**achine **L**earning
p-IDSMFOG	**p**rototype based on **I**ntelligent **D**ecision **S**upport **M**odel based on **Fog** Computing Environment
QU-DSS	**Q**uestionaire **U**sability**-** **D**ecision **S**upport **S**ystem
REL	**Rel**iability
RI	**R**efining **I**nformation
SD	**S**tandard **D**eviation
SLR	**S**ystematic **L**itreture **R**eview
SMP	**S**i**mp**licity
SVM	**S**upport **V**ector **M**achine
UFL	**U**se**f**u**l**ness

CHAPTER ONE

INTRODUCTION

1.1 Introduction

A Decision Support System (DSS) is a computer-based system that simulates humanoid decision-making and thinking, the two are able of processing information from users, and offering options that are alike to those presented by experts. DSS can greatly assist in analyzing various maintenance options and selecting the most dependable and cost-effective solutions logically and professionally [1].

A process of decision-making is a result of human activities with potentially beneficial consequences, Scientists are seeking to enhance the decisions' goodness by using computerized systems to strengthen and increase human capacities. Artificial Intelligence has recently achieved this level of realism in a variety of applications. In several industries, including medical services, finances, advertising, online security, and e-commerce, these Intelligent DSS are increasingly being utilized to manage and assist top management in decision-making processes [2].

In the context of pandemics, intelligent decision support systems have supported top management in defining and simplifying any individual or organization's difficulties. Particularly in the decision-making process, prompting relevant fields to pay more attention to emergency management and decision-making processes [3], by applying new ways that reduce the likelihood of people being dead and infrastructure being damaged as a result of man-made and natural disasters [4].

Most people have recently utilized cloud computing technology, which is a network-based paradigm that allows cloud users to access computer resources from any location, at any time. Cloud computing is a massive data storage facility that can be accessed from anywhere on the planet. As a result, most DSS may use the cloud to gather data and make responsible decisions. End-to-end latency occurs when transferring a big amount of data from a faraway data generator to the cloud, which has a critical impact on DSS, especially in emergency cases. Fog Computing was introduced by Cisco as a technology that transports cloud computing to the network's edge, this is because emergency management systems may have latency concerns because of the centralized nature of cloud Computing [5].

Fog Computing is a distributed, decentralized platform where all devices at the network's edge connect to provide services such as communication, storage, and soft computation [6]. Fog enabled emergency management techniques utilized systems with intelligence and data analytics to notify people about the problem, hence it was deemed a good option for disaster management [7] .

A variety of Machine Learning (ML) algorithms, including Supervised Learning, and Unsupervised Learning, aid in the extraction of various parameters for the decision-making and prediction model [6].

In this thesis the main goal is to develop intelligent model concerned with healthcare data management in fog computing environment.

1.2 Problem Statement

Artificial intelligence (AI) is the backbone behind effective decision support systems. A DSS helps facilitate decision-making for a team or business based on data. The Internet of Things **(**IoT**)** devices periodically produce row data that include useless, noisy, or repetitive records, but dealing with such huge amounts of data leads to increase errors. In addition, processing and storing repetitive records or unclean data causes waste the resource without any gain [5].

Transferring huge amounts of data to the Cloud leads to increased errors, as well as long-distance data transferring between the source of data and data centers causes end-to-end delays, especially in near real-time data processing. Consequently, this may give a negative impact on the processes of decision-making.

Data source constraints and lack of aggregated data in each IoT device make the difficulty of processing, managing, and storing generated data and finally delivering these meaningful data to DSS [8].In line with the above situations, the model, approach, and method to process data are necessary and needed to utilize in designing intelligent DSS in Fog Computing environment.

1.3 The Aim of the Thesis

The key objective of this research is to propose an intelligent decision support model enabling healthcare data management in Fog Computing environment (IDSMFOG). The following particular objectives are defined to accomplish the main objective:

- To develop an intelligent decision support system model for healthcare data management based on fog computing (IDSMFOG).
- To test the developed system in terms of usability and decision-support.

1.4 Research Contribution

This study contributes globally to the body of knowledge with the definite part of decision support systems in the Fog Computing environment. The detailed list of contributions is as follows:

- An intelligent decision support model in the Fog Computing Environment (IDSMFOG) evaluated by online data.
- An Instrument to measure the usability, and decision support for the developed system based on the IDSMFOG model.
- In addition, this research will add to body of knowledge by using the proposed model as a guideline for DSS developers to follow in

constructing a workable DSS in the Fog Computing environment, as a result of a thorough literature analysis and empirical findings.

1.5 Related Work

The introduction of the information era witnessed the rise of PCs and IoT innovation, which altered people's life. As of late, several emergency issues have arisen throughout the world that must be controlled to preserve money, people's lives, reputations, and natural surroundings, this obliged the associated department to give further thought to decision-making and emergency management. Many decision-making processes may address the problem by selecting the best option from a variety of options. As a result, an environment-based IDSMFOG has a promising future.

Various researchers have been encouraged in recent years to write about the decision-making processes for healthcare data management in the Fog computing environment, while also utilizing various intelligent algorithms and data management processes to assist top management in making urgent decisions.

1) According to, A. Kumari, *et al* [9] fog, cloud and IoT techniques are in introducing continuous context-aware services to the patients whenever they are needed. They employed real- time data aggregation and filtering, as well as the ZIP compression tool to compress data before transferring it to the Cloud.

2) S. Shukla, *et al*. [10], the goal is to minimize access time among healthcare IoT, end employers, and Cloud. In an Iot-FC context, for data packet apportionment and choice, the suggested bsrainy Foggy analyzing

model and methods utilize a Fuzzy Inference System joint with Reinforcement Learning, and NN evolution approaches.

3) G. Yoon, *et al*. [11] fog node was used as monitoring node. The monitoring component is used in the Iot environment to quickly regulate IoT devices, reduce network traffic, and lighten server stress. In the context of disaster management, Operation and Management server's associated data pattern is used to monitor the normalcy of data to each sensor apparatus. Instead of transmitting all data to the Only data external of the regular data range is sent to the server by the monitoring node with patterns in data.

4) In the context of digital services, to decrease data transmission overhead and processing time, data analytics can be conducted close to where the data is created. This will lead to a new sort of hierarchical data analytics, F. Mehdipour, *et al*. [12]. Fog layer serving as the first layer and the cloud serving as the last layer. It is feasible to allow IoT applications with the capacity of on-premise processing utilize thier suggested solution, Fog-engine, which results in many benefits such as decreased data size, reduced data transfer, and cheaper cloud costs.

5) In Fog-based clinical decision support, X. Liu, *et al*.[13] used a lightweight data mining technique to assess patients' medical services status in real-time using cloud computing. Their design included numerous layers of neural networks to let the cloud server safely perform any activation operations to improve the precision of the fog server's categorization of freshly discovered abnormal symptoms.
6) In a Fog environment, A. M. Ghosh and K. Grolinger [14] introduced a utility model for data that studies the effect of data source goodness on

expected results and permits for process improvement. They used activities for enhancing data goodness, and activities for converting data to support or accelerate data analysis by reduction, cleaning, and encryption in DSS data.

7) M. Ahmed, *et al*.[15]suggested a novel analysis methodology for massive amounts of healthcare data obtained from IoT wearable devices or archival patient medical records by using both fog and Cloud computing.

8) To address the issues raised by online and offline data processing, data storage, and data categorization, H. Alshammari, *et al*.[16] suggested architecture makes use of both Fog Computing and cloud platforms. Furthermore, it ensured solid and secure knowledge of patient medical data.

9) In the context of healthcare the study, A. Sutagundar and P. Sangulagi,[17] suggested an approach that made use of sensor cloud and fog Computing functionality to better categorize and preserve information while minimizing latency. The Random Forest classifier, in conjunction with the Genetic Algorithm, is used to categorize information and save just the necessary quantity of information onto the cloud server with priority.

10) In the industrial context, N. Bv and R. M. R [18] deployed defect detection machines built on their operation noise employing Fog Computing construction. In the industrial context, a micro data center is employed as a fog server to analyze and categorize machine sounds as usual and unusual. They use the Fog server to place supervised machine

learning models to filter and detect problematic equipment based on the operational sound used SVM classifier.

11) D. Aishwarya and R. I. Minu, [19] used Fog Computing data centers in storing management areas to receive, analyze, and research data for demand forecast and decision-making. It combined PCA, K-means, and Reinforcement Learning algorithm to create a DSS model for saving supermarkets.

12) N. Al Mudawi [20] offered a smart IoT-based Intensive Care Unit, a health monitoring system to help medical professionals and facilities continuously watch patients and take appropriate action. The suggested technique could take the patient's body variables in real-time, including fever, heartbeat, and others. If the patient's variables are abnormal, the doctor gets an alert about who has been assigned to care for them as well as the hospital's Emergency Unit. Through suggested technology, clinicians may similarly keep an eye on the patient from a distance.

13) The goal of, V. Colombo, *et al*. [21] work is to solve the difficulty of correctly and effectively closely watching the Fog that used a self-adaptive peer-to-peer surveillance solution that uses a small rule-based expert system to quickly change its behavior in response to the data it has collected.

This research analyzes existing and relevant studies on data management based on the Fog Computing environment for the past six years. The main purpose of this analysis is to find out the glitches and limitations of these studies, the comparisons and limitations of each of the studies have been tabled and displayed in Table 1.1.

Table1. 1: Related works

No.	Author/Year	Main Description	Limitation	Results
1	A. Kumari, *et al*.2018	Used Healthcare 4.0. in fog envirnment for clinical DSS	More data analytics done in the Cloud layer	a 3-layer patient-driven Healthcare It gives insights to the end users for the applicability of fog devices and gateways in Healthcare 4.0 .
2	S. Shukla, et al.2019	Model for analyzing data for DSS in fog environment to reduce latency.	The approach is tested on simulators iFogSim (Net-Beans)	The obtained results indicated the better performance of the proposed approach compared with existing methods
3	G.Yoon, *et al*.2019	Data management using a fog environment	Didn't use AI techniques	The monitoring node can control the operations of sensor devices remotely according to the requests of the OM server.
4	F.Mehdipour, *et al*.2019	Data analytics in fog environment	This study used a fog-engine as an analytical tool. Didn't use AI techniques	fog-engine (FE) can be deployed in the traditional centralized data analytics platform and how it enhances existing system capabilities.
5	X. Liu,*et al*. 2019	Used NN in fog environment for clinical DSS	using simulations.	It can be proven that the method achieves the goal of patient health status monitoring without privacy leakage to unauthorized parties by balancing real-time and high-accurate prediction .
6	A. M. Ghosh and K. Grolinger.2019	Used model for data analytics in fog by utilizing Deep Learning	Used edge,fog, Cloud	Results show that 50% data reduction did not have a significant impact on the classification accuracy and 77% reduction only caused 1% change.
7	M. Ahmed, *et al*. 2020	Data management in fog environment using ML(SVM)	The method simulated in the iFogSim toolkit	It affords advantages, such as reduced load on the central server by locally processing the data Additionally, it would offer the scalability of the system in the IoT network.

No.	Author/Year	Main Description	Limitation	Results
8	H. Alshammari, *et al*.2021	An analytical model for data preprocessing in fog-cloud computing	Cloud is used for more data analysis.	It enables the use of both FC and cloud platforms to handle the problems faced through online and offline data processing, data storage, and data classification.
9	A. Sutagundar and P. Sangulagi,2021	An approach to reduce the time of processing data using AI in fog environment	Used WSN to save traveling	The proposed work is working far better than conventional methods in terms of classification accuracy, latency, packet delivery ratio, energy consumption and network lifetime.
10	N. Bv and R. M. R.2021	Industry machine monitoring in fog environment using AdaB classifier model	Didn't sprite noise, not work with all audio features	The experimental results show the performance of machine learning models for the machines sound recorded with different Signal to Noise Ratio levels for normal and abnormal operations.
11	D. Aishwarya and R. I. Minu.2021	Activity recognition in a fog environment used CNN	It worked at the edge layer and used lightweight object detection	The proposed architecture gives an alert whenever a suspicious activity is detected.
12	N. Al Mudawi 2022	Used DL for detecting diseases in IoT-Fog environment.	Used FogBus	The proposed method can measure patient's body parameters in real-time and in case of anomalous values of the patient's body parameters, the device will send a notification to the assigned doctor and the Emergency Care Unit (ECU) of the hospital
13	V. Colombo, *et al*. 2022	This study provided decision about the system behavior using a lightweight rule-based expert system in fog	Used at the edge of the network	Empirical results show that adaptation can improve monitoring accuracy, while reducing network and power consumption at the cost of higher memory consumption.

In line with the above situations, the IoT devices periodically generate row data, these data must be transferred to the cloud or fog for processing or storage, due to IoT devices constraints that are not skilled in handing out, managing, and storing produced data and finally delivering meaningful data to DSS, as well as, these data contain useless, noisy, or repetitive records, hence, dealing with such huge amounts of data leads to increased mistakes, and end-to-end delay specially in a real-time data processing. Consequently, this will give a negative impact on the decision-making process especially in emergencies situations. In addition, Processing and storing noisy or repetitive data wastes resources without providing any benefit.

Consequently, it can be indicated from Table 1.1 that the step-by-step IDSS model for healthcare data management in a Fog Computing environment is missing, as well as the need for a model to manage and prepare data near to the end devices for DSS applications is critical. To bridge this gap, the IDSMFOG model was proposed.

1.6 Thesis Organization

There are five chapters in this thesis. The following is a list of the following each chapter's whole contents:

Chapter 2: This chapter shows the theoretical background for DSS, as well as data management in a fog environment.

Chapter 3: Before identifying the suitable proposed model components, a systematic and in-depth analysis of concepts and theories is required. It is critical to ensure that the proposed healthcare data management model in a fog environment corresponds to the components required and meets the study objectives. Thus, this chapter includes extensive discussions of the principles and theories underpinning this work.Model Development- In Chapter 3, the work required in accomplishing objectives one, two, and three are covered in depth. It describes the steps involved in determining the proper phases and components of the proposed model. It also goes through the stages required in developing the suggested model. Then the system based on the proposed model was developed, as well as, tested in the term of usability.This chapter also discusses real-world DSS solutions and how they apply to this study.

Chapter 4: This chapter discusses the results of the developed system and test usability findings.

Chapter 5: This chapter shows the thesis conclusions, findings, and suggestion for further work.

CHAPTER TWO

THEORETICAL BACKGROUND

2.1 Introduction

In truth, today's decision-making is more difficult than it was in the past due to two major factors. To begin with, advances in technology and communication networks have resulted in a higher number of viable solution options from which a decision-maker might choose. Second, the increasing structural complexity of today's challenges might result in a cost-magnification chain reaction if a mistake occurs [22].

The term Decision Support System (DSS) has a variety of definitions that vary depending on the perspective of the author. It is available in a variety of forms and may be used in a variety of ways. A DSS is described as "a computer-based system that supports the decision-making process" on the one hand. A DSS is more clearly characterized as: "an interactive, flexible, and adaptive computer-based information system designed to aid in the solution of a non-structured management problem so that better decisions may be made. It is data-driven, has a user-friendly interface, and allows for the decision maker's insights". Other interpretations lay halfway in between these two extremes. Keen and Scott Morton founded DSS in 1978 [23].

Human decision, which involves deductive reasoning supported by experience, information, and knowledge, is used to make administrative choices in general. The decision-making process can be somewhat supported by computer-aided automation to compensate for the influence of human mistake. Unless correctly processed data and an optimal model are given, the final system cannot be fully automated. DSS is used to simulate human reasoning and decision-making; both are capable of taking data from users, digesting them, and providing solutions that are similar to those offered by human experts. DSS can greatly assist in analyzing various maintenance options and selecting the most reliable and cost-effective solutions in a methodically and transparently manner [22].

DSS blends people's intellect with computer applications to improve the quality of outcomes. "For managers working with semi-structured challenges, DSS is a computer-assisted decision-making tool." DSS is still a relevant and broad term for a variety of systems of information that aid in decision-making. Interactive computer-based solutions that help decision-makers address unstructured problems with data and models, the name "DSS" is still applicable and inclusive of a wide range of information systems that help with decision-making [24].

2.2 DSS's Brief History

In the late 1950s and early 1960s, research papers on logistic decision creation were conducted at Carnegie-Institute-of-Technology. In the 1960s, technological research on interactive computer systems was predominantly conducted at the Massachusetts Institute of Technology. The concept of DSS is thought to have originated as a distinct field of study in the mid-1970s, before growing faster in the 1980s. Between mid- and late-1980s, single-user and model-based DSS gave way to Organizational Decision Support Systems (ODSS), Group Decision Support Systems (GDSS), and Executive Information Systems (EIS), see Figure 2.1.

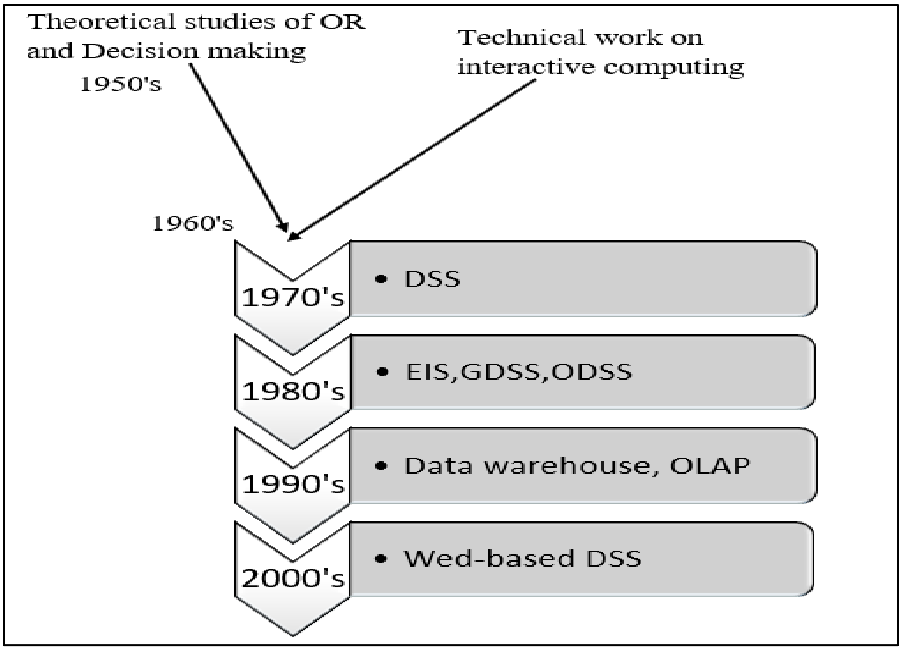

Figure 2.1:DSS's Brief History Adapted from [23]

DSS exists within a multidisciplinary ecosystem that includes (but is not limited to) data science, artificial intelligence, human-computer interaction, simulation methods, software engineering, and telecommunication [23].

2.3 Taxonomies of DSS

There is no all-encompassing DSS classification, just as there is no all-encompassing definition. Various studies propose various classifications. At the user level, researchers distinguish between inactive, interactive, and supportive DSS. An inactive DSS aids in decision-making but does not offer explicit choice proposals or results. Such decision conceptions or approaches can be generated via interactive DSS. An interactive DSS enables the decisions - makers (or its advisor) to comprehend, amend, or boost the system's suggested decisions before returning them to it for confirmation. Before sending the decision maker's thoughts for approval, the system completes, completes, and enhances them [25].

After then, the technique is repeated until a consolidated solution is reached. On a conceptual level, DSS separates as: 1) Communication-Driven DSS, 2) Data-Driven DSS, 3) Document-Driven DSS, 4) Knowledge-Driven DSS, and 5) Typical-Driven DSS, access to and modification of a numerical, financial, optimization, or simulation model are prioritized by a Typical-driven DSS, while not necessarily data-intensive, depend on Decision-makers are assisted by the data and criteria given by DSS members in comprehending state. Amazon is an example of a Data-Driven DSS that enables top management to make decisions [26].

Entree to and alteration of a time sequence of inner corporate data as well as, in some circumstances, external data are the goals of data-driven DSS or data-oriented DSS. There are technological distinctions between PC DSS and enterprise-wide DSS. Several managers at a corporation are served by enterprise-wide DSS that link to massive data warehouses. Simple systems, one-user desktop DSS run on a single manager's PC [26].

2.4 Architectures of DSS

Once again, different studies identify different components in a DSS. DSS is made up of three primary components:

- The database management system (DBMS).
- The model-based management system (MBMS).
- The dialog generation and management system (DGMS).

Researcher and practitioners have debated the four important DSS components:

- the user interface.

- The database.
- The model and analytical tools.
- The DSS networking and structure.

DSS has 5 elements:

- Users of the decision-making process with a variety of positions or tasks (decision-maker, advisor, domain-expert, system-expert, data collector).
- The environment of clear and precise decision-making.
- A goal system that adequately captures most desires.
- Outward source of data, information database, working database, data-warehouses, and meta-database; accurate model and approaches; actions; implication and search engines; managerial applications; and reporting system.
- A location where decision-making options can be created, examined, and recorded.

D. Coleman proposed a generic architecture comprised of five distinct components as shown in Figure 3.2 [27].

- The data management system.
- The model management system.
- The knowledge engine.
- The user interface.
- The user(s).

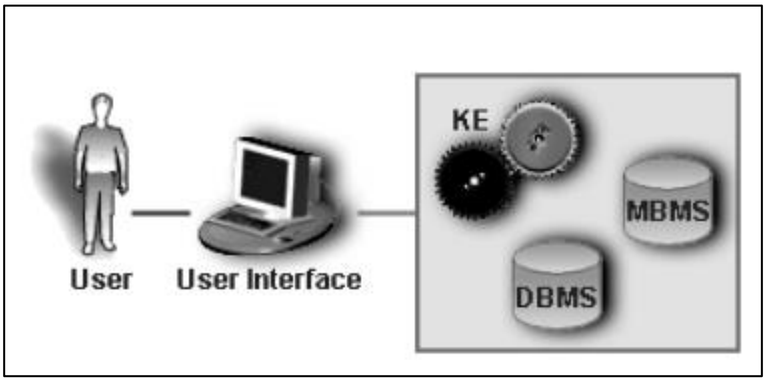

Figure 2. 2 :The Five Distinct Components of a DSS [27]

As indicated in Figure 2.2, a schematic of the fundamental architecture of a DSS incorporates the system's decision-maker as a component. a knowledge basis, a database, and a model base are examples of input.

Processing entails utilizing decision models to simulate or investigate several states to determine the optimum answer given the restrictions. Processing feedback can give extra ideas that can be modernized in real-time to assist in issue resolution. Forecasts and explanations may be generated as output to substantiate recommendations and give guidance. After receiving the findings, the decision-maker can use the system to ask questions or provide further details.

The definition of decision support has recently been expanded to include a variety of decision-making tools, including analytics, business intelligence, and knowledge management systems, which may or may not include direct interaction with the decision-maker. The abilities of these decisions are typically increased by utilizing AI features, such as aggregating widely scattered data and drawing insights from vast distributed datasets referred to as big data. Personalization for decision maker preferences, can be included in such systems, which can even mimic human decision making. They provide strong new tools for solving very complicated issues and are developing future trends [24]. In the context of this research, the proposed model (IDSMFOG) architecture is based on the following architecture, see Figure 2.3.

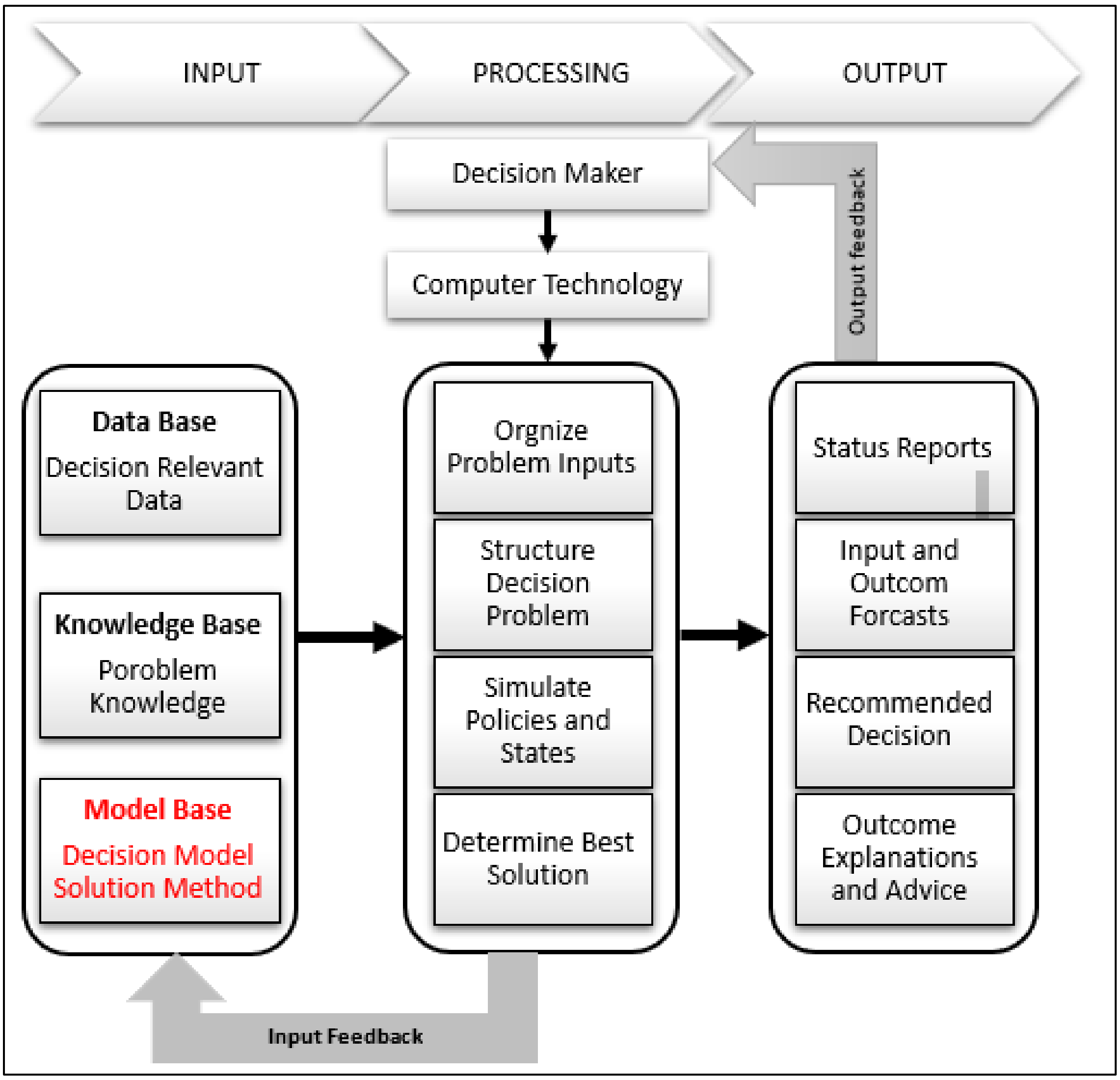

Figure 2 .3: Architecture of a Decision Support System [24]

2.5 Intelligent Decision Support System

Intelligent Decision Support Systems (IDSS) use Artificial Intelligent (AI) approaches to enhance and improve decision-making process. IDSS, as one of the most essential components of artificial intelligence and big data, maybe a valuable tool for risk analysis, management, and control. Its capabilities include not only vast information retrieval and processing, mining, and integration, but also intelligent data processing for risk detection, assessment, and reaction [6].

Fuzzy set theory, case-based reasoning, evolutionary algorithms, artificial neural networks (ANN), and machine learning techniques are some examples of AI tools that, when combined with DSS, can be extremely helpful in solving challenging applied problems that are frequently real-time, involve massive amounts of distributed data, and demand complex reasoning. One theory of intelligence holds that it is mainly concerned with logical behavior, leading an intelligent system to choose the optimum course of action in a given situation. IDSS would be DSS that exhibit certain traits indicative of intelligent behavior, including:

- Studying experience.
- Make meaning of contradictions or ambiguities.
- A new scenario in a timely and suitable manner.
- Using logic to address issues and making logical inferences.
- Navigating complex situations.
- The use of information to comprehend or alter the environment.
- Recognizing the proportional weight that different considerations have while making a choice.

IDSS, which applies a range of AI approaches, is developing as beneficial for practical and essential applications such as (Neural networks, Expert systems, Genetic algorithms, Case-based algorithms, and Fuzzy logic). Data is viewed as the IDSS's backbone since it can be used to improve human decision making in a variety of contexts, including corporate choices and healthcare support Apps [4].

2.6 Fog Computing

Cisco coined the term Fog Computing [8] to describe the process of extending Cloud computing capabilities to the network's edge. Fog

Computing enhances Cloud settings by processing data closer to the data source at the network's edge. As seen in Figure 2.4.

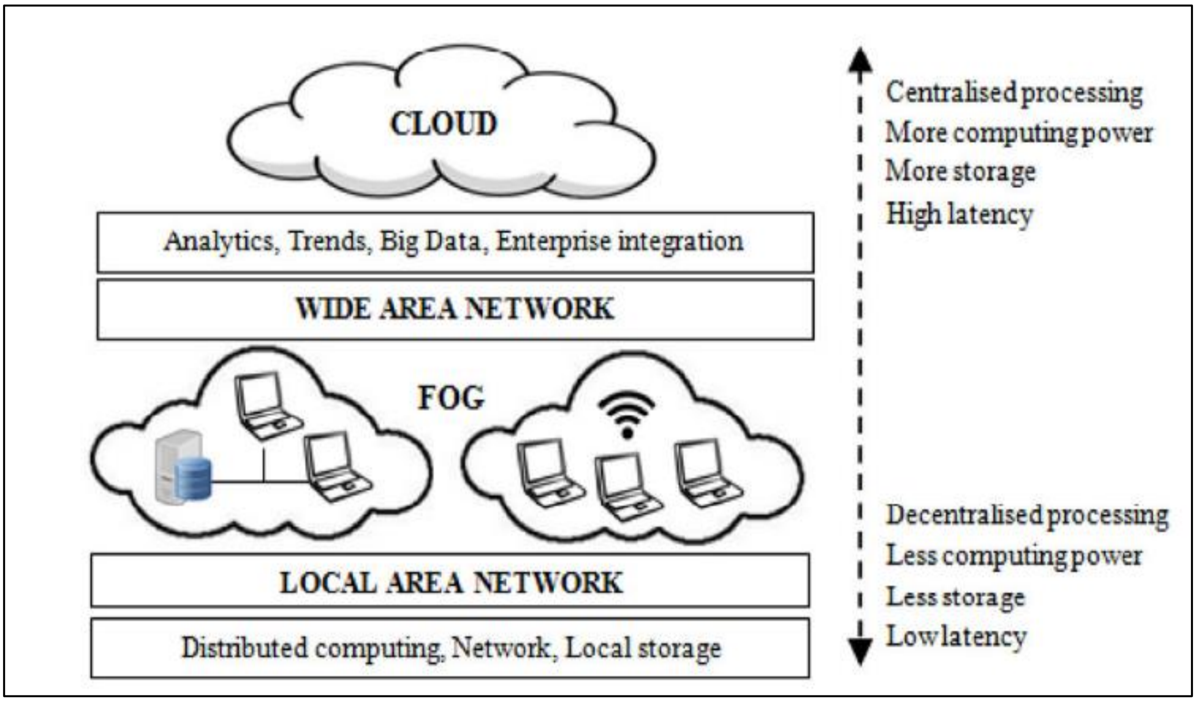

Figure 2. 4: Fog Computing [8]

Fog Computing does not displace Cloud computing; instead, the two approaches work best together. By processing data near the client, latency, bandwidth, and even data transmission costs can be reduced. This also moves the necessity for processing to the Cloud. The traditional Cloud computing architecture has been expanded upon by the Fog Computing idea.

The fog layer aims to provide minimum computing, storage, and networking capabilities near the network's edge for quick data processing, but the Cloud will still be used for large datasets and historical data analysis. By sending the data closer to the client rather than having it go via a central server in the cloud, fog computing solves the issue of continuous data transfer. These fog devices could communicate internally to optimize the amount of processing power on the local network by doing some tasks on the fog layer rather than utilizing bandwidth for every device [8].

The development of IoT has enabled any item with an internet connection to connect and interact in real-time. It has enabled the smart operation of a far broader range of physical goods, including automobiles, electric meters, household equipment, medical appliances, traffic signals, and street lighting. This leads to the continual creation of information at an enormous rate, necessitating the use of analytics to extract relevant information.

The cloud paradigm applies ideal analytical methods to extract valuables from collected data to better decision-making in future activities [28].

Cloud computing has significant disadvantages, such as high latency, to address this constraint, CISCO offered a vision of Fog Computing that may run applications directly on billions of linked devices at the network edge in a decentralized, distributed framework [8].

To offer useful data to DSS, Fog is used as an appropriate alternative to the Cloud. Users may use the Fog framework to design, control, and operate their applications to take advantage of the edge of services. They created an open-source platform to attract additional developers to build apps on the network's edge.

As a result, several applications for fog environments have already been presented to improve the future of the sector that it belongs [28]. High-resource PCs may be used as a Fog Computing environment since they deal with managing and processing data [29].

2.8 Data Management

As the world's digitization gains traction, the development, gathering, examination, and storing of data that drives advancement every day exists. The development of decisions-making methods built on pertinent facts, as well as both, have grown in significance.

IoT technology, which may detect sensor networks, systems, or people capable of watching the real-world, gathering and processing data, and making decisions, has occurred in parallel with the proceedings. This technology includes public Cloud and database systems. The device that makes up this skill may interconnect with each other and then exchange material online [11].

IoT technology is effectively being used in health, home automation, intelligent buildings, energy management systems, smart farming, and surrounding industries systems as a result of these capabilities. IoT systems cannot filter data as it is being generated, despite the importance of this data in the healthcare sector the right healthcare professional must get the data produced by IoT devices instantly and precisely.

The initial thing where data may drive be handled beforehand being delivered to the Cloud is at the IoT edges. To design and amend better data and integrate criteria, as well as to effectively address pertinent challenges in the data life cycle, it is necessary to review data characteristics [30].

People are overwhelmed with data as a consequence of the internet and social media of thing. The popularity of social media and the IoT has directed to rapid growth of big data, which is a collection of enormous volumes of data. Terabytes to petabytes of large data have been generated. Volume, velocity, variety, and veracity are the four main features that set big data apart. The velocity characteristics deal with information in motion, i.e., streaming the data to reply in milliseconds rather than seconds, the varied attributes represent information at rest in the terabytes to Exabytes range [13],[14], covers data in a variety of formats, including text, multimedia, structured and unstructured data, and the truth qualities deal with data that is unclear owing to data inconsistency.

Large data sets have characteristics that make it challenging for businesses to control and make use of them. In today's environment, information overload is an issue, yet comprehension is lacking. Ninety percent of the data on the planet today was created over for the last two years on various social networking sites.

Big data challenges traditional warehouses that gather and store enormous amounts of internal and external data as its relevance rises. The use of data from these archives enhances organizational performance, organizational effectiveness, and decision-making [31].

Before data is delivered to the cloud, it must be filtered since failing to do so limits the speed and accuracy of Cloud services, and a lack of synchronization or precision can result in patients getting medical care and losing their lives. Speed and precision are therefore two essential factors to take into account [32].

Analysis of big data tries to swiftly extract crucial data that can be used to generate predictions, identify trends, track down personal information, and eventually reach conclusions. The handling of big data benefits greatly from diverse and enormous datasets. This isn't always the case, though, since additional information might create confusion and problems [33].

IoT data must be handled quickly and in large volumes, since it is frequently unreliable and generated at high rates. The data produced by IoT cannot be effectively interpreted using conventional data mining techniques. Data must be properly linked and pre-processed to yield useful findings.

Massive amounts of raw data are continuously being collected via IoT, therefore, it is crucial to develop creative methods for turning raw data into information that can be used. For instance, sensors can give pertinent raw data streams in the healthcare industry by detecting crucial elements including human breathing, eating, drinking, heart rate, heart rate, and sugar levels [34].

Millions of sensors are expected to generate a large amount of data. For a variety of purposes, these digital data are employed in diverse ways. To ensure privacy and security, it is necessary to reveal the sources of the data and how it is used. Even after processing and evaluation, the acquired vast data may still be useless if it isn't understood. To extract information from the moment raw data is created, data mining tools are the most frequently suggested methods [35].

How to gather important data from many complex observable environments at different times and provide relevant results is a crucial problem. Data must be reviewed using suitable data-mining procedures to find pertinent trends in IoT data. To eliminate the need to change data mining rules each time a sensor is included or removed, IoT data-mining algorithms must also be capable of adapting to dynamic settings or shifting data streams. A branch of artificial intelligence called machine learning seeks to simulate human learning on machines without the need for explicit programming.

For IoT data mining, ML methods are useful. This is because algorithms for machine learning have many of characteristics that make them appropriate for IoT data analysis. If a new smart device is introduced to the network, for example, the machine learning algorithms can continue to learn new rules. Data mining is among the most advantageous analytical techniques even if many others have been designed to make the IoT smarter [32].

The interim data storage, initial processing, and analytics are handled by fog, which acts like a conduit for connecting devices to the cloud. Data from IoT devices are produced, processed, and maybe briefly stored by fog before being swallowed by cloud apps, receiving the appropriate response from fog or the cloud, and being delivered to a thing. A fog diagram with three layers and a data view as shown in Figure 2.5.

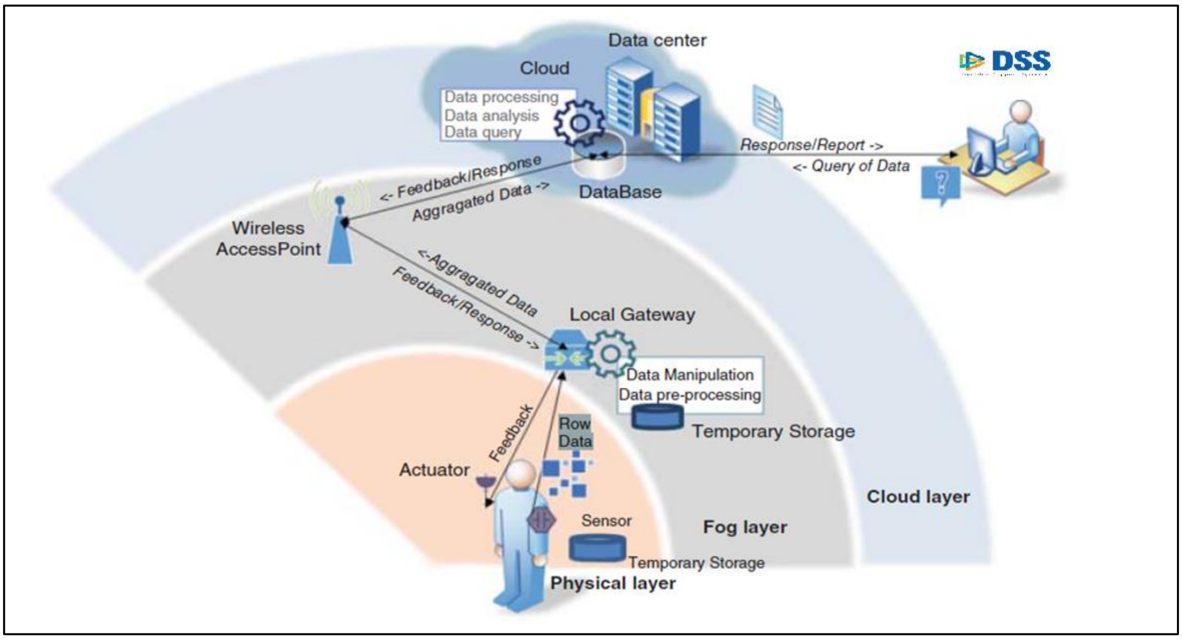

Figure 2. 5: A Three-layer Fog Diagram with A Data View [30]

The three main stages of a common Fog Computing architecture are the device layer (also known as the physical layer), fog level (also known as the Edge-network layer), and cloud-layer. A reference model concept for Fog-computing is described in the paper. The purpose of a three-layer design was to gather and handle data in fog computing models. It included gateways, IoT middleware, and sensor nodes.

The fog layer handles data management, processing, virtualization, and service delivery.

A Fog-based schematic for data analytics was suggested. A Fog-based construction with three horizontal and three vertical layers was created for crowdsensing applications.

A programming framework was provided to develop the processing paradigm for data streams. Another multitier Fog computing system for data processing, which is also used to monitor, administer, and maintain systems, was introduced in edge nodes. The handling of data and related ideas, such as data aggregation techniques, data filtering tactics, data placement, data privacy, and so forth, are the focus of fog data management [16].

According to the three-layer fog architecture shown in Figure 2.5, sensor and acquired data are transferred to the higher layer and should be controlled accordingly. End-to-end latency and network traffic are two of the main justifications for employing fog computing, as was previously stated. Local data management has benefits including greater productivity and privacy.

The main advantages of data management in fog computing are as follows:

• Increasing efficiency: Local data processing and the deletion of erroneous, tedious, or unnecessary data in the fog layer minimize the network load and boost the efficiency of the network. Reducing the volume of data will also lower Cloud processing and storage needs because data transported to the Cloud must be processed, stored, and analyzed there.

• Increasing the level of privacy: A challenge with IoT and Cloud computing is protecting data privacy. IoT sensors may generate and transmit private and personal data, but doing so unaltered or unencrypted raises the risk of exposure. Furthermore, difficult mathematical operations cannot be completed by computers with limited resources. Techniques for protecting end-device privacy, including encrypting algorithms, might not be practical. As a result, a fog layer may be used for encryption, privacy, and data modification. Nevertheless, another topic that will be explored further is fog device protection.

• Increasing data quality: Data quality might be improved by eliminating duplicate, corrupted, or noisy data, as well as by integrating incoming data into a Fog layer.

• Reducing the end-to-end latencies: Network latency and processing time are issues that must be taken into consideration when collecting data from the Cloud in IoT scenarios because of the inherent nature of networks. The end-to-end delay may be decreased by placing the data pre-processing close to the devices in the fog layer.

• Decreasing cost: Internet use, Cloud processing, and storage expenses are reduced via local data processing and compression in a fog layer. But a trade-off must be made and the expenses of fogging devices must be considered [30].

2.9 Data Management in Fog Computing

The operations that make up the fog data include data gaining at the physical layer, where data is created, handed out and storing in higher levels, and send to the DSS. There are several methods for managing and processing fog data, and these methods may be divided into 3 categories: data collection, data preprocessing, and data distribution to DSS. The following elements can be included in the three phases:

2.9.1 Data Gathering

This means that data collect from different sensors and send to a fog node or gateway for processing to generate meaningful data for DSS [32]**.**

2.9.1.1 Data Converting

The process of changing digital data from one specific format to another is called data conversion. The procedures involved in converting data from one application to another are detailed in this sample process flow. For this publication, converted data refers to any data that has been migrated from a legacy system to a new system. The data conversion phases are as follows: [36].

2.9.1.2 Data Annotation

Data annotation is a process in which a human data annotator adds groupings, labels, and other contextual components to a raw data collection so that computers can read and act on the information [16].

2.9.2 Data Preprocessing

To extract meaningful information from raw data, data processing must be achieved, this process contains sub-process as in the following sections.

2.9.2.1 Data Fusion and Integration

Data fusion is the integration of readings generated by sensors and basic information that happens as near the data source as possible. The main aims of data fusion are as follows:

- to minimalize the uncertainty by fusing various data sources together and performing condition calculations;
- to support real-time decision-making; and
- to decrease the security risks related to sending data over the network [37].

2.9.2.2 Data Cleaning

Sensors generate row data that generally contain unclean and useless data that could be errors, missing, and repetitive values, Further, the large quantity of unwanted and useless data can lead to high computation costs and the overutilization of resources, so there is a necessity to fill in the missing value with (min. max, mean, median), removing repetitive, and inconsistent data [38],[39].

2.9.2.3 Data Normalization

In general, data normalization is the process of transforming the mapping of data variables from one uniform range to another. It also addresses a variety of issues, such as variable distribution, while other solutions additionally fix and optimize for range and distribution [33].

There are several normalizing methods, including Min-Max, Z-score, and Decimal scaling normalization.

Normalization is scaling the data to be analyzed to a specific range such as (0.0, 1.0) to provide better results.

- Minimax-Maximum Normalization

The original data is transformed linearly in this method of data normalization. The minimum and maximum value from the data is retrieved, and each value is changed using the formula 2.1, [33]:

$$\boldsymbol{V'} = (\mathbf{V} - \mathbf{min}\ (\mathbf{A}))/(\mathbf{max}\ (\boldsymbol{A}) - \mathbf{min}\ )\boldsymbol{A})) \ (\boldsymbol{new}\ \_\mathbf{max}\ (\boldsymbol{A}) - \boldsymbol{new}\ \_\mathbf{min}\ (\boldsymbol{A})) + \boldsymbol{new}\_\mathbf{min}\ (\boldsymbol{A}) \qquad (2.1)$$

Such that A is the feature data,

The absolute values of A's lowest and maximum are given by Min(A) and Max(A), respectively.

The updated value for each data item is denoted by "V'."

V represents the previous value of each data item.

The new _max and _min variables represent the maximum and minimum values of the range, respectively (i.e., the needed range boundary values).

2.9.3 Artificial Intelligence (AI) in Fog Computing

AI is a rapidly evolving subject with several practical applications in everyday life and active research areas. Before the first computer with programming was made, many people pondered if such machines would ever become intelligent. People now use intelligent software to automate daily jobs, identify audio or visual data, make medical diagnosis judgments, and perform many types of scientific studies [15],[40].

By connecting the internet to the real world, the burgeoning IoT age enables a smart society. At the same time is spreading rapidly over a wide range of economic industries and enterprises. Big data, improvements in ML algorithms, and high-performance computing and storage capabilities in the Cloud are primarily to blame for the transformation of human daily life brought on by many of new applications in computer vision, gaming, speech recognition, medical diagnostics, and other fields [41].

The foundation for accomplishing the objective of ubiquitous intelligence is the combination of IoT with AI, specifically AI-enabled IoT. Many IoT devices provide enormous amounts of data for ML, which results in intelligent choices, analytics, and other data-added services. On the other side, extremely large amounts of data have the potential to overwhelm storage systems. In many IoT applications, processing near or at the source of the data is highly sought. In time-sensitive applications, such as emergencies situations, these IoT apps can leverage local processing to decrease latency. To that end, a system-level framework for allocating resources and activities along the edge-Cloud border might be provided by the new Fog Computing paradigm [12],[6].

ML algorithms are a branch of AI that can identify patterns in data and learn from them to build predictive models and make data-driven choices. Due to enhanced computer processing power and the availability of vast amounts of data, mostly generated by the IoT, their usage in additional applications is expanding [42].

Big data technologies and high-performance computing have developed along with ML, opening up new options in a range of operational contexts to justify, evaluate, and obtain data-intensive processes.

A scientific field known as ML gives computers the ability to grasp among other things without being fully programmed. Every year, more academic fields are being included in machine learning, such as bioinformatics, biochemistry, medical meteorology, economics, aquaculture, chemo-ecology, robotics, and climatology. ML algorithms are in the tens of thousands, and hundreds are developed annually.

The three components of each machine learning technique are representation, evaluation, and optimization.

There are three forms of ML as follow:

- Supervised learning is training data that includes wished-for outputs.
- Unsupervised learning is training data that has no real outputs.
- Reinforcement learning [6]. See Figure 2.6.

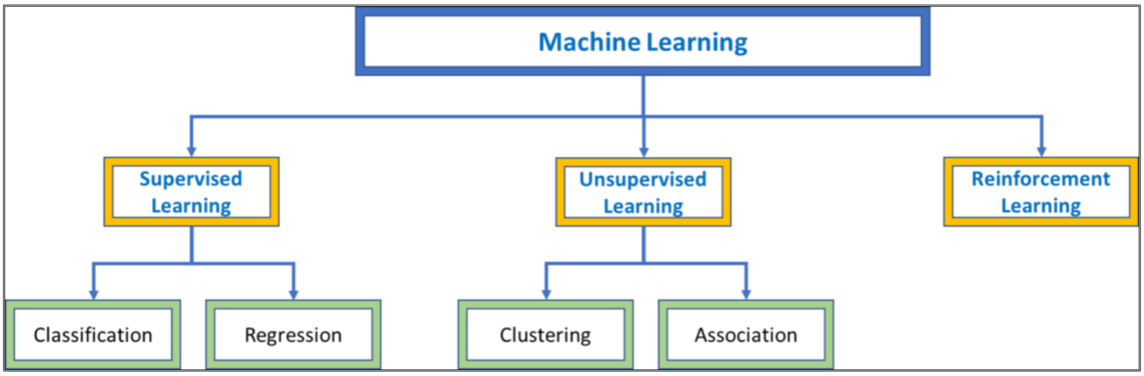

Figure 2. 6 : Types of ML [6]

The ML models and their associated algorithms for IoT big data analysis, such as clustering, feature extraction, and classification, incorporate support vector machines (SVM), k-means, and k-nearest neighbors (K-NN). The resource-constrained environment introduces new challenges to the design of ML algorithms.

There are certain generic algorithms that respond to distribution and resource limitations, such as k-NN and some particular neural network approaches.

However, the latter typically needs tedious adjustment of the granularity of the environment model and environmental sensor data [41]. The Fog-level IoT system allows for quicker and more effective transferring of data to the Cloud, as well as the utilization of edge intelligence to decrease data transmission and storage requirements on the cloud [43].

2.9.3.1 Support Vector Machines (SVM)

SVM is a well-known supervised ML model based on statistical learning theory that can be used for classification and regression issues [18]. SVM's purpose in linear separable and binary classification is to identify an optimal hyperplane that separates the two classes with a maximum separating margin. The geometrical distance of blank space between the two species is denoted as the margin [44]. Figure 2.7 illustrates a SVM classifier.

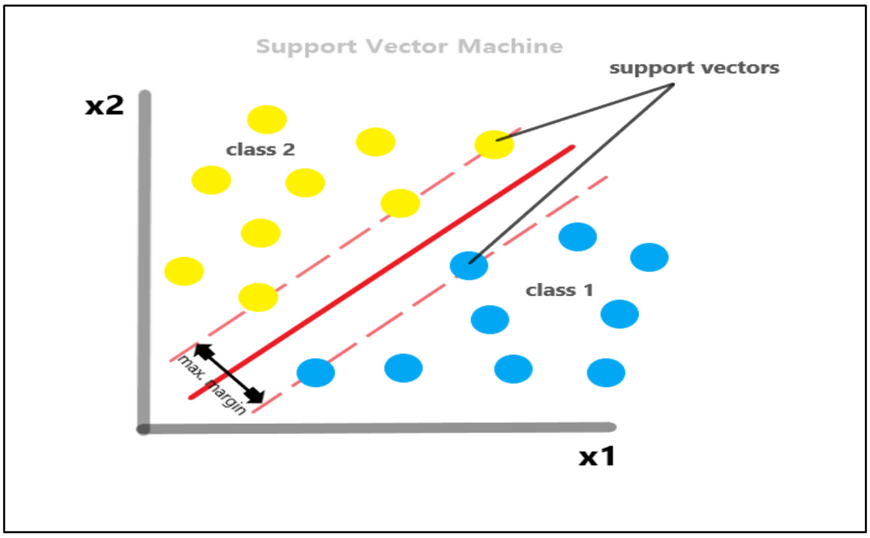

Figure 2 .7 : SVM classifier Illustration

The bigger the margin, the better the SVM classifier's generalization ability. This margin's bounds are parallel hyperplanes to the separating hyperplane. In classification assume you have a training set as in formula 2.2 [44]:

$$(x1, y1), \ldots, (x_i, y_i) \in R^{n} \times \{+1, -1\}, \qquad (2.2)$$

These are n-dimensional designs. the labels for the (vectors) xi are yi. A label with a value of +1 indicates that the vector belongs to class +1, whereas a label with a value of -1 indicates that the vector belongs to class -1. So, we look for a function $f(x) = y: R\, n \rightarrow \{+1, -1\}$.It not only accurately identifies the forms in the training data (a process that is reasonably easy), but also correctly classifies previously unnoticed patterns. We refer to this as a generalization.

The VC theory (Vapnik-Chervonenkis), also known as statistical learning theory, shows that it is essential to restrict the class of functions that our computer is allowed to learn; otherwise, learning the underlying function is not possible. SVMs are therefore built using the class of hyperplanes, as in formula 2.3 [44].

$$(\mathbf{w} \cdot \mathbf{x}) + b = 0;\ w \in R^{n},\ b \in R, \qquad (2.3)$$

2.9.3.2 K-means

K-means is a traditional clustering technique and one of the unsupervised ML algorithms that classify data based on their distance or dissimilarity, categorizing similar samples into one group and dissimilar samples into different clusters.

2.9.3.3 Evaluation Performance Metrics

This section contains metrics for determining how successful a classifier is in predicting the class label of records. Some well-known classifier assessment measures include accuracy (also known as recognition rate), sensitivity (or recall), specificity, precision, and F1. Although accuracy is a specific statistic, the term "accuracy" is sometimes used to refer to a classifier's prediction ability in general. Using training data to generate a classifier and then estimating the accuracy of the resulting learned model might lead to false overoptimistic predictions due to the learning algorithm's overspecialization of the data. Instead, assess the classifier's accuracy on a test set of class-labeled tuples that were not used to train the model [45].

The data should be divided in such a way that neither of them is too high, which is more dependent on the amount of data you have. If your data is too small then no split will give you satisfactory variance so you will have to do cross-validation but if your data is huge then it doesn’t really matter whether you choose an 75:25spli, 80:20 split or a 90:10 split (indeed you may choose to use less training data as otherwise, it might be more computationally intensive) [46].

There is a need to understand four terms that serve as "building blocks" in the computation of numerous evaluation measures:

- **True positives (TP):** These are the positive records that the classifier properly labeled. The number of true positives is denoted by TP.
- **True negatives (TN):** These are the negative records that the classifier successfully identified. The number of true negatives is denoted by TN.
- **False positives (FP):** These are the negative records that were classified wrongly as positive. The number of false positives is denoted by FP.
- **False negatives (FN):** These are the positive records that were classified incorrectly as negative. The number of false negatives is denoted by FN.

A confusion matrix in Figure 2.13 summarizes these phrases. The confusion matrix is a great tool for determining how effectively your classifier recognizes records from various classes. TP and TN indicate when the classifier is correct, but FP and FN indicate when the classifier is incorrect (i.e., mislabeling). A confusion matrix is a table of at least m by m dimensions given m classifications (where m = 2). A $CM_{i,j}$ item in the first m rows and m columns indicates the number of tuples of class I categorized as class j by the classifier. For a classifier to be accurate, most of the records should be represented along the confusion matrix's diagonal, from entry

$CM_{1,1}$ to entry $CM\ m, m$, with the rest of the entries being 0 or close to zero. That is, FP and FN should ideally be close to zero [19].

		Predicted class		
		yes	*no*	Total
Actual class	*yes*	TP	FN	P
	no	FP	TN	N
	Total	P'	N'	$P + N$

Figure 2. 8: The confusion matrix.

Additionally, P' is the total number of records with a positive classification (TP + FP) and N' is the total number of records with a negative classification (TN + FN). The total records are TP+TN+FP+TN, commonly abbreviated as P+N or P' + N'. Even though the confusion matrix shown is for a binary classification assignment, confusion matrices for many classes may easily be constructed in a similar way.

Recall that the percentage of true positive cases to the sum of true positives and false negatives is represented. Recall was defined as in formula 2.4[15]

$$\text{Recall} = \frac{T_P}{TP+FN} \qquad (2.4)$$

Precision was defined as the proportion of true positive cases to all positive instances. Precision was defined as in formula 2.5, [15]:

$$\text{Precision} = \frac{T_P}{TP+FP} \qquad (2.5)$$

The F-score was a mix of precision and recall. The F-score was defined as in formula2.6, [15]:

$$\text{F-score} = \frac{2*Recall*Precision}{Recall+Precision} \qquad (2.6)$$

The performance analysis of all experiments in this study was based on the most frequent assessment metric used for statistical testing, known as accuracy. The accuracy is the percentage of correctly anticipated labels. The confusion matrix is used to calculate accuracy. Accuracy is measured by multiplying the genuinely positive and negative values by the total number of samples. Accuracy was defined as in formula 2.7,[15]:

$$\text{Accuracy} = \frac{T_P + TN}{TP + TN + FP + FN} \qquad (2.7)$$

2.9.3.4 Network Performance Metrics

When you're evaluate your performance metrics there are several different metrics that you can analyze. Network performance can be affected by a number of different factors. Using a network performance monitoring solution, your enterprise can search for these factors and understand how they're hurting your network's performance. There are some of essential network performance metrics as bandwidth usage, throughput, latency, packet loss, and connectivity [47].

2.9.3.5 Data Visualization

The term "data visualization" refers to the visual display of information and data. Data visualization tools make it easy to observe and comprehend patterns, anomalies, and shapes in data by employing visual components like diagrams, graphs, and maps. We now have a lot of data in our hands, so data visualization methods and approaches are critical to studying large amounts of data and making data-driven decisions in the era of Big Data [48].

2.10 Usability Measurement

According to [49],[50], identifying the usability metrics is easier than collecting them. Often, usability is measured based on users' satisfaction with a given application. Usability measures are based on the extent to which the functional and non-functional requirements are met for the intended application. Accordingly, proposing and developing a usability test instrument will also depend on the extent to which the services for this application are met for user needs, and of course, according to the nature of the intended application [51]. As it is evident that the applications have become customized and each application is designed and developed to solve a specific problem or for pre-defined purposes, so the usability test for this application also depends on the main purpose of this application. Accordingly, and in the context of this study, the proposed instrument for measuring the usability of DSS applications will focus on five important dimensions [52]. Besides, Usability testing focuses on the actual's users' experience with the intended application, as well as their comments and feedback during the use of such applications in their environments [49].

CHAPTER THREE

PROPOSED METHODOLOGY

3.Introduction

The building procedure for IDSMFOG suggested design is discussed in the third chapter. The IDSMFOG phases and components are designed after relative research and then evaluation of the linked studies. All of these stages and elements are combined to form the suggested design.

3.2 Model Development

To achieve objective 1, the IDSMFOG model for healthcare data management in a Fog Computing environment is developed. Through the offered functionality, this model serves as the overall framework. This design model includes elements acquired from current Fog models, techniques, and methods. Many previous Fog models, techniques, and methodologies, as well as other real-time, DSS solutions, have been investigated to discover the phases, components, and model functions required for this model.

The goal of proposing the IDSMFOG model is to provide healthcare data management for support decision-making in fog comput ing environments depending on the organization's needs, as well as to assist and guide DSS developers in the design of DSS. Similarly, the major goal of the IDSMFOG model is to assist decision makers, particularly in emergencies situations, by generating meaningful data from raw data provided by IoT devices and delivering it to DSS.

As a consequence, raw data is unclean data, To decrease processing and data retrieval time, a DSS should contain just the necessary data. The most challenging component of this study is the construction phase, which leads to the key contribution, which is the IDSMFOG design model technique for emergency management based on a Fog computing environment and its built system.

The design process yields the outcomes of the proposed IDSMFOG Design Model. Hence, the suggested IDSMFOG is made up of three major phases that are combined to form the proposed model.

As a result, the proposed IDSMFOG model comprises three major phases: data collection for decision support, data preprocessing for decision support, and an intelligent decision support system. Each phase of the IDSMFOG model, as described in the following sections, contains a collection of processes and sub-processes or components. Figure 3.1 shows own proposed model.

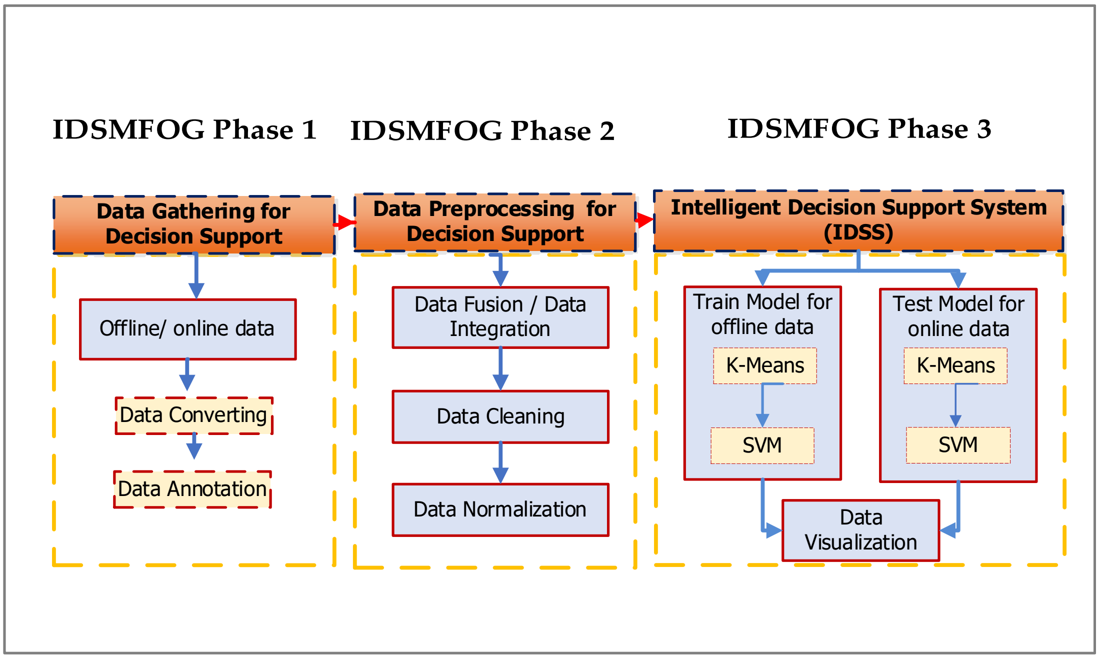

Figure 3 .1 The Own Proposed IDSMFOG Model for Evaluation online / Offline Data

After model develpment it will be evaluated on online/offline data in own website application.

3.2.1 Data Gathering for Decision Support

This phase is the first phase of the IDSMFOG model, data collection is an essential component of every project or project management attempt. All components and processes of this phase are as following:

3.2.1.1 Offline Data

In this component data are previously collected, and datasets for IDSS of COVID-19 can be gathered from more than one source (Johns Hopkins, Our World In Data , New York Times, and The Humanitarian Data Exchange).

In this research, the datasets (COVID-19.CSV files) were collected from github.com. This dataset has about 19,800 records and their attributes are (Cases, Deaths, Recovered, Vaccinated people, Population, Lon, Lat, Fips, and Date).

After removing the missing rows from the dataset, the dataset's size became about 19,761 records only three attributes (Cases, Deaths, and Population) are used to detect Which city in (Iraq, USA) is need to be on lockdown or afflicted city, the others are used to visualize the number of Cases on the world map for every country and the report of (Confirmed, Recovered, Deaths, Vaccinated, Death Rate) all over the world (since the beginning of COVID-19), for last three months, last month, last week and time series, Figure 3.2 shows data sources.

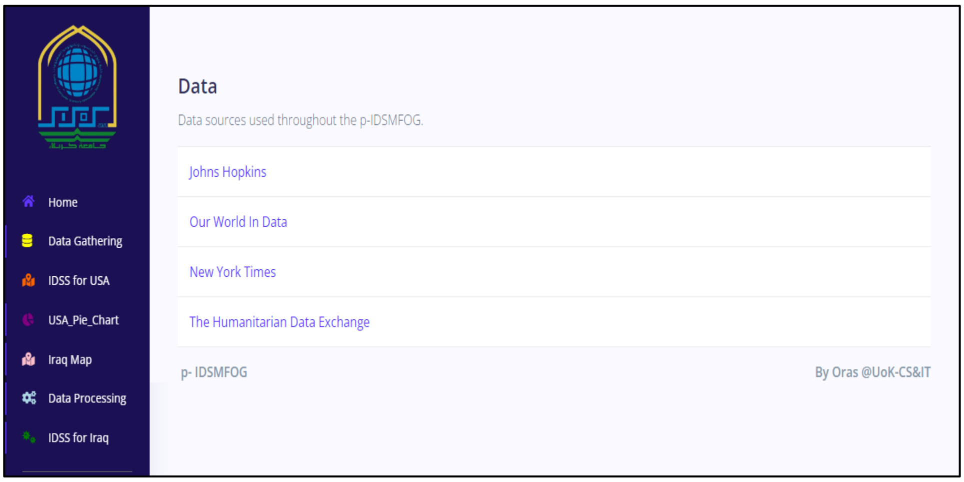

Figure 3. 2: Datasets Source Used Throughout the p-IDSMFOG

The process of changing digital data from a specific format to another is called data converting. The procedures involved in converting data from one application to another are detailed in this sample process flow.

After gathering raw data, you should now feed that data into AI systems, so they can carry out human-like actions. The issue: These machines can only act according to the conditions you established for the dataset. The main method for bridging the gap between sample data and AI/ML is data annotation.

Data annotation is a process in which a human data annotator adds groupings, labels, and other contextual components to raw data collection so that computers can read and act on the information, as shown in Figure 3.3.

```
iso_code,continent,location,date,total_cases,new_cases,new_cases_smoothed,total_deaths,new_deaths,new_deaths_smoothed,total_cases_per_million,new_cases_per_million,new_cases_smoothed_per
_million,total_deaths_per_million,new_deaths_per_million,new_deaths_smoothed_per_million,reproduction_rate,icu_patients,icu_patients_per_million,hosp_patients,hosp_patients_per_million,w
eekly_icu_admissions,weekly_icu_admissions_per_million,weekly_hosp_admissions,weekly_hosp_admissions_per_million,total_tests,new_tests,total_tests_per_thousand,new_tests_per_thousand,new
_tests_smoothed,new_tests_smoothed_per_thousand,positive_rate,tests_per_case,tests_units,total_vaccinations,people_vaccinated,people_fully_vaccinated,total_boosters,new_vaccinations,new_
vaccinations_smoothed,total_vaccinations_per_hundred,people_vaccinated_per_hundred,people_fully_vaccinated_per_hundred,total_boosters_per_hundred,new_vaccinations_smoothed_per_million,ne
w_people_vaccinated_smoothed,new_people_vaccinated_smoothed_per_hundred,stringency_index,population,population_density,median_age,aged_65_older,aged_70_older,gdp_per_capita,extreme_pover
ty,cardiovasc_death_rate,diabetes_prevalence,female_smokers,male_smokers,handwashing_facilities,hospital_beds_per_thousand,life_expectancy,human_development_index,excess_mortality_cumula
tive_absolute,excess_mortality_cumulative,excess_mortality,excess_mortality_cumulative_per_million
AFG,Asia,Afghanistan,2020-02-24,5.0,5.0,,,,,0.126,0.126,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,8.33,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-02-25,5.0,0.0,,,,,0.126,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,8.33,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-02-26,5.0,0.0,,,,,0.126,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,8.33,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-02-27,5.0,0.0,,,,,0.126,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,8.33,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-02-28,5.0,0.0,,,,,0.126,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,8.33,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-02-
29,5.0,0.0,0.714,,,,0.126,0.0,0.018,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,8.33,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
01,5.0,0.0,0.714,,,,0.126,0.0,0.018,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,27.78,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
02,5.0,0.0,0.0,,,,0.126,0.0,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,27.78,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
03,5.0,0.0,0.0,,,,0.126,0.0,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,27.78,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
04,5.0,0.0,0.0,,,,0.126,0.0,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,27.78,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
05,5.0,0.0,0.0,,,,0.126,0.0,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,27.78,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
06,5.0,0.0,0.0,,,,0.126,0.0,0.0,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,27.78,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
07,8.0,3.0,0.429,,,,0.201,0.075,0.011,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,27.78,39835428.0,54.422,18.6,2.581,1.337,1803.987,,597.029,9.59,,,37.746,0.5,64.83,0.511,,,,
AFG,Asia,Afghanistan,2020-03-
```


Figure 3 .3 Data after Annotation.

3.2.2 Data Preprocessing for Decision Support

In this phase data fusion / integration, merging data from different data sources leads generally unclean and useless data that could be errors, missing, and repetitive values, Further, the large quantity of unwanted and useless data can lead to high computation costs and the overutilization of resources, so there is a necessity to removing repetitive, and null values. Some libraries in python are used to clean and scale the data in a specific range.

3.2.3 Intelligent Decision Support System (IDSS)

After data is gathered and processed, data must be visualized to make decisions that help decision makers and stockholders.

Integration modern tools such as AI andML in the data analysis process in fog Computing environment automate most of the regular work to minimize the interaction between people and new application, that serves the emergencies situations like the COVID-19 pandemic by reducing the time need to make a decision. IDSMFOG acts as an intelligent model that utilized ML techniques to support decision-making process in the fog computing environment.

In this research, K-Means cluster algorithm is used as an unsupervised ML technique with two clusters (0,1) to predict a class to the dataset of two examples, Iraqi (7690 records) and USA (2311 record) cites depending on two features (Cases, Deaths).

SVM classifier is used in this research to classify the Iraqi and USA cites that need to be lockdown (0) or unlockdown (1), the training data is equal to 75% of all dataset, test data is equal to 25% and the class of the datasets is from the K-Means cluster, the SVM model fit the train data and testing the test data, and then calculate the accuracy of the SVM model. After training and testing the SVM model, the model saved and loaded in the developed system to predict new data on USA example.

Deploying Ml model after loaded it to IDSMFOG system to predict new data, the result of that visual display of information and data Choroplethmapbox and world map functions in python are used to visualize the final decisions.

The primary result of these phases is an IDSMFOG conceptual model for near real-time decision-making in a fog computing environment.

3.3 Development of the System

The system was constructed based on the suggested IDSMFOG model to validate the proposed model. The developed system is known as p-IDSMFOG, its concerned with COVID-19 Dashboard, and decides which (Iraq, USA) cities need to be lockdown. It was created in three stages based on the IDSMFOG model: Data Gathering for decision support, Data preprocessing for decision support, Intelligent decision support system.

Data sources are the fundamental building blocks in system development in the context of this study. As a result, the specifics of p-IDSMFOG development in practice will be explored in the next section, a high-level Python web framework (Django) is adapted with CSS, HTML, and JS codes to develop p-IDSMFOG.

In line with the development of decision support systems that use massive amounts of data and deal with extremely complex processes, this study prepared a set of functions that were used as guidance for the developer during the IDSMFOG development phase during the first phase of system development. As a result, some components of the system will be highlighted in the next sections.

3.3.1 Pseudocode for the System (IDSMFOG)

Before starting the design and development of p-IDSMFOG system, it is necessary to know about the dataset used, and data collected perversely, in p-IDSMFOG data of Confirmed, Recovered, Deaths, vaccinated cases of COVID-19 are from the repositories. In line with model phases and component, p-IDSMFOG development process used all component of the model that is mentioned previously with the related function of the python Django framework and pseudocode to help the clinician with the decision-making process.

3.1. Pseudocode illustrate the Preprocessing, and train model phases for IDSMFOG

3.1. Pseudocode of IDSMFOG

--------------------*Preprocessing Phase*----------------------

```
INPUT
        D = {x1,x2,…..,xn}

OUTPUT
        SVM = {xn+1}    // SVM classifier prediction on new data

BEGIN
        F= number of features
        For each i in F
           If X[0] == ' '  Then
              X[0] = 'annotation'

        For i = 1 to n
           For j = 1 to n
              If X[i][j] == NAN Then
                 Dropna(X[i][j])
```

-------------------*Train Model Phase*----------------------

K-means Algorithm

```
        Normalize(F)
        C=2
        kmeans = KMeans(C)
        kmeans.fit(F)
        clusters = kmeans.fit_predict(F)
        D1= D add clusters   //Merage D with clusters
```

SVM Algorithm

```
        X_train, X_test, y_train, y_test = train_test_split(F, Cluster,
        test_size=0.25, random=true)

        classifier = SVC(kernel='linear', random_state=0)
        classifier.fit(X_train, y_train)
        SVM = classifier.predict(F)
        pickle.dump(save classifier)
        loaded_classifier=pickle.load(open classifier)
        F1=import online data with F featuers
        SVM = classifier.predict(F1)
         For i = 1 to n
            If SVM[i] = 0
               SVM = 'lockdown'
            Else
               SVM = 'unlockdown'
           plot on map SVM, F1
END
```

3.4. Instrument Design and Development

To measure users' experience towards measuring usability for DSS applications. In this study, the questionnaire as an evaluation instrument for DSS applications has been used for its reliability and wide adoption. To meet the study objectives, the questionnaire based-instrument is designed to elicit comments, suggestions, and feedback regarding DSS usability named (QU-DSS). Since the proposed instrument (QU-DSS) has concerns to usability of DSS, the QU-DSS usability dimensions and their items have identified through a systematic literature review (SLR).

Figure 3.4 presents the main phases of the evaluation instruments design.

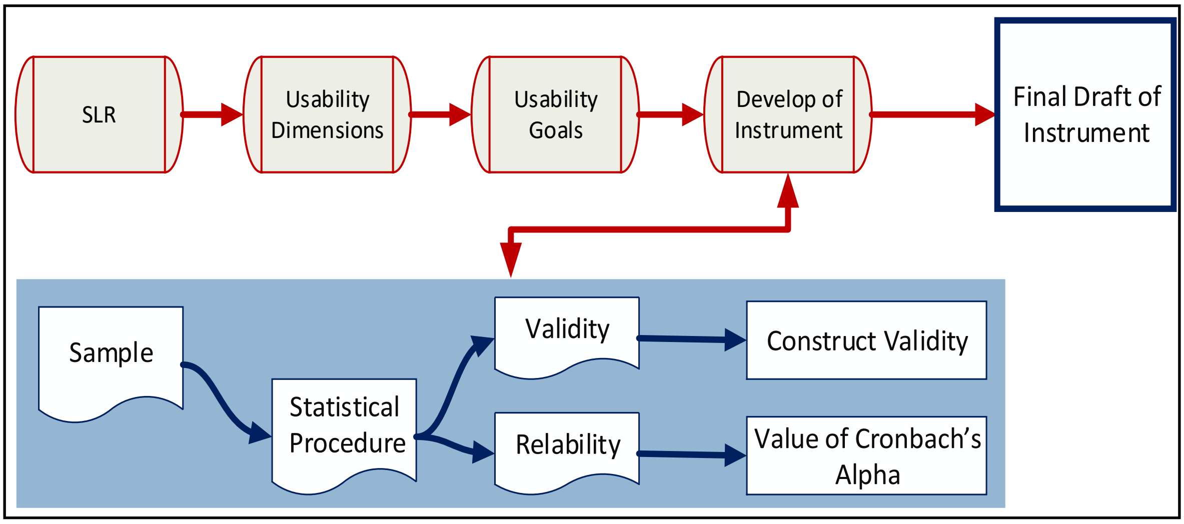

Figure 3. 4: Evaluation Instrument Design (Main Phases)

As indicated in Figure 3.4, the designing procedure of the QU-DSS start with SLR to identify the instrument dimensions and their items. Consequently, the findings of SLR depict that the QU-DSS evaluation instrument has five dimensions named: Usefulness, Simplicity, Reliability, Flexibility, and Decision-Support.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

This chapter is demonstrated the experimental results of the developed system named p-IDSMFOG, with its SVM model and evaluation, and then the p-IDSMFOG usability dimensions findings will be discussed.

The p-IDSMFOG is a website application concerned with COVID-19 Dashboard, and decides which (Iraq, USA) cities are need to be lockdown.It was created in three stages based on the IDSMFOG model: Data Gathering for decision support, Data preprocessing for decision support, Intelligent decision support system. The laptop is used as a fog environment with intel processor CORE i5, RAM 8GB, HD 500GB SSD, and windows 10 (64-bit) OS.

Django web framework python 3 is used with Html, CSS, and JS codes in PyCharm Community Edition 2021.3.2 to develop p-IDSMFOG, after training the K-SVM (k-means + SVM) model in Jupyter anaconda3, after model building and saving, its loaded and used in the website application to predict new data in USA example.

4.2 K-SVM Model

In K-SVM model building process, K-Means with k equal two is used to get the class to the dataset, such that the cites who have a high number of Cases and Deaths puts under cluster 1 and other cites under cluster 0, after get a class of dataset, the SVM classifier has been built.

4.2.1 K-SVM Building Results

After applying K-Means and getting a class of dataset, the dataset is split into 75% train and 25% test data the model train and test data of Iraqi cites about (7690 records) and saved by using the pickle library in python and then load it in the IDSMFOG system to predict new data of Iraqi cites about COVID-19 to decide if any city needs to be on lockdown or not, the model achieved 99.94 % accuracy, as shown in Figure 4.1

```
predection class= [0 0 0 ... 0 1 0]
Accuracy of SVM= 99.94797086368366
F-1 Score: [99.97132205 99.71988796]
                    city   Cases  Deaths   active recover        date  Clusters
2                  ANBAR   23830      97   23,729       4  2022-09-16         0
3                BABYLON   70152    1072   69,024      56  2022-09-16         0
4          BAGHDAD-KARKH  383360    1900  381,228     232  2022-09-16         1
5         BAGHDAD-RESAFA  320577    3470  317,074      33  2022-09-16         1
6                 Basrah  241362    1565  239,754      43  2022-09-16         1
...                  ...     ...     ...      ...     ...         ...       ...
7686              NINEWA       6       0        6       0  2020-05-11         0
7687         SALAH AL-DIN       9       0        4       5  2020-05-11         0
7688         SULAYMANIYAH     160       4      156       0  2020-05-11         0
7689             THI-QAR      72       3       58      11  2020-05-11         0
7690              WASSIT      37       2       33       2  2020-05-11         0

[7688 rows x 7 columns]
```

Figure 4. 1: Accuracy of SVM with sample of Iraqi Cites Cases and Deaths.

The k-svm result of Iraq example of IDSMFOG system, the results plot on Iraqi map and circulate the cities that need to be lockdown with red circles, and others with green circles, as shown in Figure 4.2.

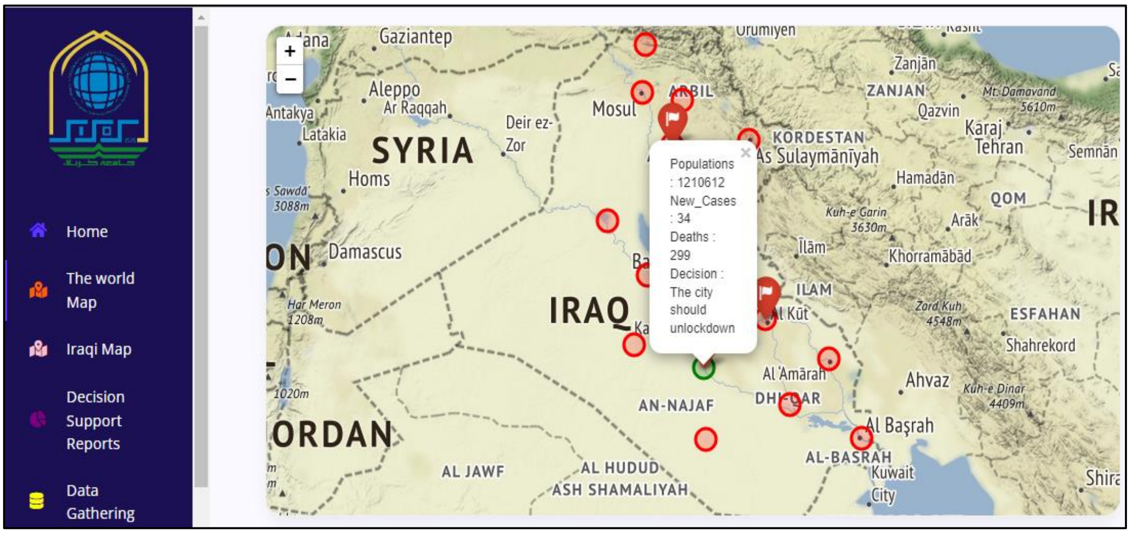

Figure 4. 2: Decision Making (unlockdown) Iraqi Cities

Another K-SVM example on the USA dataset (3082 records) build to predict which city needs to be lockdown, this model achieved 99.87 % accuracy, as shown in Figure 4.3

```
 0 0 1 1 1 1 1 1 0 1 0 0 0 1 1 0 1 1 0 1 0 0 1 1 0 1 0 1 0 0 0 1 1 0 1 1 1
 1 1 1 1 1 1 0 0 1 1 1 1 0 1 0 1 0 0 1 1 0 1 1 0 0 1 1 1 1 1 0 0 0 0 0 0 1
 0 1 0 0 1 0 0 1 1 1 0 1 1 0 1 1 1 0 0 0 1 0 0 0 0 1 1 0 1 0 1 1 0 0 0 1 1
 0 1 1 1 0 1 0 0 1 0 1 1 0 1 0 0 0 1 1 1 0 1 0 1 1 0 0 1 1 1 0 1 1 0 0 1 1
 1 0 1 1 1 0 0 1 1 0 0 0 1 1 0 1 0 0 1 1 1 1 1 1 0 0 0 1 0 1 1 1 1 1 1 0 1
 0 0 1 1 0 1 0 0 1 0 1 1 1 1 1 0 1 0 1 1 0 0 0 1 1 0 1 1 1 1 1 0 1 0 1 0 1
 1 1 1 0 1 0 1 1 1 1 0 1 1 1 1 1 1 0 0 0 1 1 1 1 1 0 0 1 1 1 0 0 1 1 1 1 0
 1 0 1 1 1 1 1 1 0 0 1 0 0 1 1 1 1 0 1 0 0 1 0 0 0 1 0 1 1 1 1 1 0 1 1 1 0
 0 1 0 1 1 1 1 0 1 1 1 0 0 1 1 1 1 1 1 0 1 1 1 1 0 0 1 1 1 1 1 0 0 1 0 1 0
 1 1 0 0 0 1 0 1 0 1 0 0 1 0 1 0 0 1 1 0 1 0 1 1 1 0 0 0 1 0 0 1 0 0 1 0 0
 1 1 1 0 0 1 1 1 0 0 0 1 1 1 0 1 1 0 0 0 0 0 0 1 1 1 1 1 1 1 0 1 1 1 0 0 1
 1 1 1 0 0 1 1 1 0 0 1 0 1 1 1 1 0 1 0 1 1 1 0 0 1 1 0 1 0 1 0 1 0 1 1 0 0
 0 1 1 1 0 0 1 1 1 1 1 1 0 0 1 1 0 0 0 0 0 0 1 0 0 1 1 1 1 0 0 1 1 0 0 1 0
 0 1 1 1 0 1 1 1 1 1 0 0 0 0 0 0 1 0 1 0 1 0 0 1 0 0 1 1 1 1 1 0 1 0 1 0 1
 0 1 0 1 1 1 0 1 0 0 1 1 0 1 1 0 1 0 0 0 0 1 1 1 1 1 0 1 0 1 1 0]
Accuracy of SVM= 99.87046632124353
F-1 Score: [99.84962406 99.88623436]
            date     county     state   fips   cases  deaths  Population  \
0     2022-09-18       Ohio   Indiana  18115    1638    19.0        5978
1     2022-09-18       Lake   Indiana  18089  123921  1894.0      498558
2     2022-09-18    Barbour   Alabama  01005    6873   101.0       24964
3     2022-09-18      Bacon   Georgia  13005    3996    80.0       11079
4     2022-09-18      Boone  Illinois  17007   16181   132.0       53159
...          ...        ...       ...    ...     ...     ...         ...
3080  2022-09-18       Hill     Texas  48217    9715   198.0       36471
3081  2022-09-18    Hidalgo     Texas  48215  248226  3991.0      880356
3082  2022-09-18  Henderson     Texas  48213   17465   432.0       83667
3083  2022-09-18   Hemphill     Texas  48211    1327     7.0        3271
3084  2022-09-18       Hays     Texas  48209   73204   532.0      255397

      Clusters
```

Figure 4. 3: Accuracy of SVM with sample of USA Cites Cases and Deaths.

The k-svm result of USA example of IDSMFOG system, the results plot on USA map and shows message by mouse move to the city that needs to be lockdown or not with the number of Cases and Populations of it, as shown in Figure 4.4.

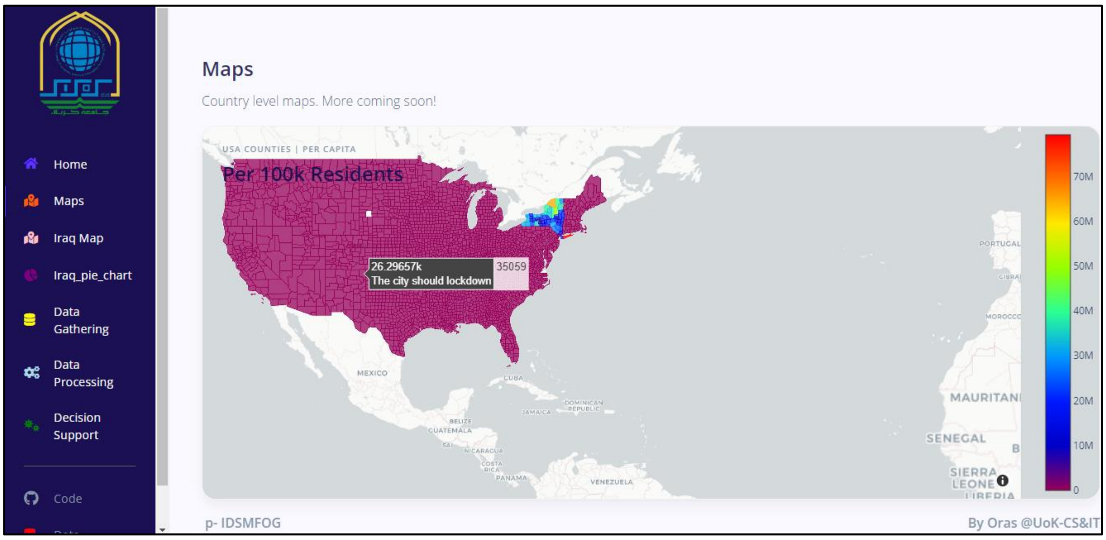

Figure 4. 4: Decision Making (unlock down) USA Cities

4.3 IDSMFOG Home Page

To get data from more than one data source, and put it in one place or file, data integration is utilized in the data processing phase to get data from multiple data sources about people around the world who were infected with COVID-19 (Confirmed cases), Recovered, Deaths, Vaccinated, and death rate, COVID-19 Dashboard homePage is shown in Figure 4.5.

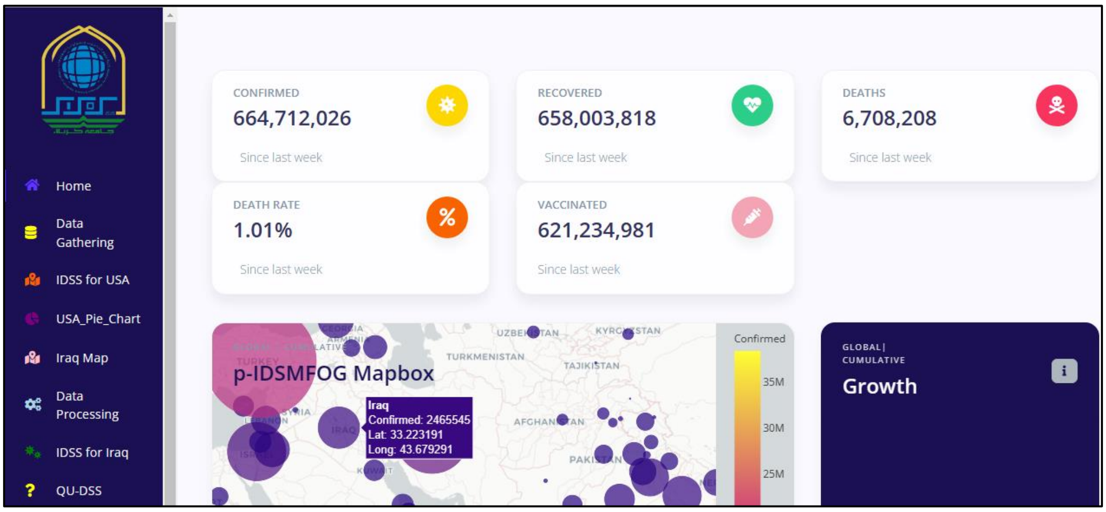

Figure 4. 5 IDSMFOG Home Page with weekly Growth Cases of COVID-19.

Figure 4.6, shows number of (Confirmed Cases and Deaths) for (week, month, three months, and time series), and sorted list of world countries with number of Cases daily updated

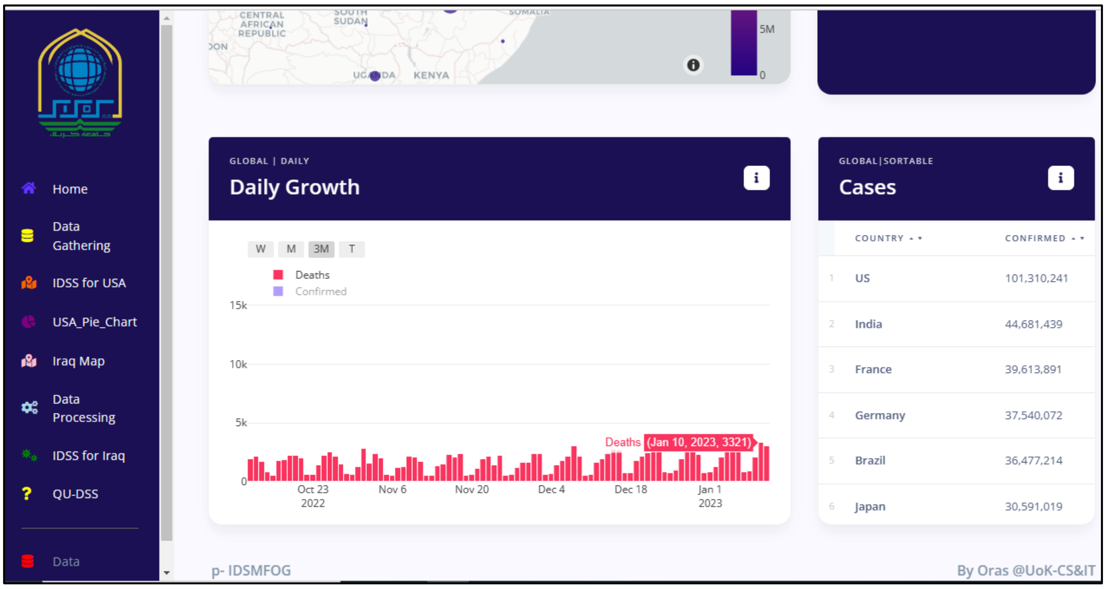

Figure 4. 6: IDSMFOG Home Page with Daily Growth Cases of COVID-19.

4.4 QU-DSS Design Findings

After the design of QU-DSS in chapter three, and as indicated in Table 4.1, The proposed instrument formed a sequence of items (questions) with a suitable answer for each question. Accordingly, the QU-DSS instrument was utilized to measure the usefulness, simplicity , flexibility, reliability, as well as decision support under usability for DSS applications based on a fog computing environment.

Table 4. 1:The proposed Instrument (First Draft)

No	QU-DSS Dimensions	QU-DSS Items	Author and Year
		QU-DSS-UFL-1	
		QU-DSS-UFL-2	
		QU-DSS-UFL-3	
1	Usefulness (UFL) 7 Items	QU-DSS-UFL-4	
		QU-DSS-UFL-5	
		QU-DSS-UFL-6	
		QU-DSS-UFL-7	
		QU-DSS-SMP-1	
		QU-DSS-SMP-2	
		QU-DSS-SMP-3	
2	Simplicity (SMP) 6 Items	QU-DSS-SMP-4	
		QU-DSS-SMP-5	
		QU-DSS-SMP-6	
		QU-DSS-REL-1	
		QU-DSS-REL-2	
3	Reliability (REL) 7 Items	QU-DSS-REL-3	

		QU-DSS-REL-4
		QU-DSS-REL-5
		QU-DSS-REL-6
		QU-DSS-REL-7
4	Decision Support (DS) 7 Items	QU-DSS-DS-1
		QU-DSS-DS-2
		QU-DSS-DS-3
		QU-DSS-DS-4
		QU-DSS-DS-5
		QU-DSS-DS-6
		QU-DSS-DS-7
5	Flexibility (FLX) 6 Items	QU-DSS-FLX-1
		QU-DSS-FLX-2
		QU-DSS-FLX-3
		QU-DSS-FLX-4
		QU-DSS-FLX-5
		QU-DSS-FLX-6

The QU-DSS was piloted to measure its reliability and validity before delivering it to in real environment in measuring DSS Applications based on a fog computing environment. In this study, researchers used “somewhat agree” instead of the “neutral” option to force respondents to choose a side. Thus, a 5-point Likert-type scale has been utilized in the research, and the interpretations of the scales: Strongly Agree, Agree, Somewhat Agree Disagree, and Strongly Disagree.

4.4.1 Pilot Test: Measuring Instrument Goodness

To evaluate the whole QU-DSS instrument under survey conditions. A pilot study was conducted.

By pilot testing, the instrument's problems are determined before utilizing the full examination. Pilot testing focuses on each question in the proposed instrument and exam the validity. It checks whether the item is interpreting and representing the information it’s intended to measure.

4.4.1.1 Subject Selection

In the context of this study, the pilot test conducted with 84 respondents was obtained among the college of computer science postgraduate students as well as CS lecturers. The respondent number is sufficient to generate credible results in the statistical test.

The total number of sample participants was 84. The pilot test data showed 50 females (59%) and 34 males (41%). Most of the participants revealed their age to be above 30.

4.4.1.2 QU-DSS Validation

In the context of this study, two well-known methods were adopted to validate the QU-DSS instrument, the first one is Content Validity and the second one is Interitem Consistency Analysis. For content validity, face validity was utilized . The justification for conducting face validity is to confirm that the QU-DSS proposed instrument includes a sufficient collection of measurement elements of the desired dimensions. Accordingly, this study employed seven experts face-to-face and through e-mail to assess the validity of the QU-DSS items.

The review finding of the experts indicates that some QU-DSS components weren't appropriate for usage and did not satisfy the required specifications of QU-DSS dimensions. Consequently, the QU-DSS's first draft was modified in terms of repositioning, rewording, and sometimes discarding some irrelevant items. To confirm consistency, the reliability test was conducted as a second method for QU-DSS instrument.

Thus, consistency is measured by the dependability test The Cronbach's alpha coefficient (α) was determine. This research tested and found ($\alpha>0.6$) to be significant. Table 4.2 depict the findings of the reliability test of the QU-DSS measurement items, the findings indicate that the QU-DSS instrument is consistent and significant, and can be adapted to utilize for measuring the usability and decision support for DSS application in for computing environment.

Table 4. 2: Finding of Measuring Reliability

No	QU-DSS Dimensions	No. of Items	Cronbach's alpha
1	QU-DSS-UFL	7	0.783
2	QU-DSS-SMP	6	0.725
3	QU-DSS-REL	7	0.714
4	QU-DSS-DS	7	0.791
5	QU-DSS-FLX	6	0.708

4.4.1.3 The Factor Analysis Procedure

Factor analysis was used to confirm the level of importance of QU-DSS components. The acceptable items are chosen using the Kaiser-Meyer-Olkin (KMO) and Bartlett's tests, as well as the Factor loading value. Consequently, the KMO test was used to arrange the data for factor loading. As indicated in Table 4.3, the obtained findings of the KMO test have been accepted based on the condition of the KMO test ≥ 0.50.

Table 4. 3: The KMO Overall Findings

No	QU-DSS Dimensions	No. of Items	KOM	The Value of Bartlett's test
1	QU-DSS-UFL	7	0.783	0.000
2	QU-DSS-SMP	6	0.725	0.000
3	QU-DSS-REL	7	0.714	0.000
4	QU-DSS-DS	7	0.791	0.000
5	QU-DSS-FLX	6	0.708	0.000

As indicated in Table 4.3, the outcome of the Bartlett's test of sphericity is (0.000) for all dimensions, indicating that the second criterion is satisfied (p 0.05). As a result, this stimulates the data to be willing and able to undergo factor loading analysis testing. Table 3.4 details the factor loading test.

Table 4. 4: The Factor Loading Overall Findings

No	QU-DSS Dimensions	QU-DSS Items	Factor Loading Value
		QU-DSS-UFL-1	0.587
		QU-DSS-UFL-2	0.661
		QU-DSS-UFL-3	0.625
1	Usefulness (UFL) 7 Items	QU-DSS-UFL-4	0.554
		QU-DSS-UFL-5	**0.362***
		QU-DSS-UFL-6	0.511
		QU-DSS-UFL-7	0.531
		QU-DSS-SMP-1	0.686
2	Simplicity (SMP) 6 Items	QU-DSS-SMP-2	0.534

No	QU-DSS Dimensions	QU-DSS Items	Factor Loading Value
		QU-DSS-SMP-3	0.541
		QU-DSS-SMP-4	0.661
		QU-DSS-SMP-5	0.672
		QU-DSS-SMP-6	0.572
3	Reliability (REL) 7 Items	QU-DSS-REL-1	0.645
		QU-DSS-REL-2	**0.387***
		QU-DSS-REL-3	0.615
		QU-DSS-REL-4	0.609
		QU-DSS-REL-5	0.694
		QU-DSS-REL-6	0.621
		QU-DSS-REL-7	0.566
4	Decision Support (DS) 7 Items	QU-DSS-DS-1	0.551
		QU-DSS-DS-2	0.686
		QU-DSS-DS-3	0.538
		QU-DSS-DS-4	0.546
		QU-DSS-DS-5	0.665
		QU-DSS-DS-6	0.673
		QU-DSS-DS-7	0.578
5	Flexibility (FLX) 6 Items	QU-DSS-FLX-1	0.689
		QU-DSS-FLX-2	0.534
		QU-DSS-FLX-3	0.545
		QU-DSS-FLX-4	0.662
		QU-DSS-FLX-5	**0.373***
		QU-DSS-FLX-6	0.571

As can be seen in Table 4.4, all items in QU-DSS are usable and represent the respective dimensions except the three items marked with an asterisk (*) are excluded from the test and whose factor loading values are less than 0.50.

This chapter begins with a complete explanation of IDSMFOG, which is split into three phases for decision support (data gathering for decision support, data preprocessing for decision support, and an intelligent decision support system). Furthermore, with IDSMFOG, each of these stages is divided into building component and sub-component elements and guiding component elements. These elements and features were developed by a series of comparative studies on current DSS models, approaches, methodologies, and guidelines, and real-world DSS operations. In addition, the proposed steps for creating IDSMFOG were defined.

4.5 Use QU-DSS for Usability Measurement

After confirming the validity of the developed tool for use, it was used to measure the usability of a clinical decision support system developed previously by the researcher.

The actual users of the system in three government hospitals (27 clinicians) as well as the sixth-stage students belonging to the Faculty of Medicine (73) total (100), were asked to evaluate the developed system in terms of usability. Moreover, to sum up or to describe the characteristics of the selected dataset or samples, the mean M and standard deviation STD were calculated and carried out as descriptive statistics. The following tables and figures show the results obtained for usefulness,simplicity, reliability, decision support, and flexibility through the evaluation process.

Table 4. 5: Usefulness Dimension Findings

No	Usefulness Items	STD	M	1	2	3	4	5
				Frequency & Percentage				
1	QU-DSS-UFL1	1.023	4.541	2.0 %	1.0 %	3 %	77 %	17 %
2	QU-DSS-UFL2	1.045	4.672	0.0 %	1.0 %	3 %	78 %	18 %
3	QU-DSS-UFL3	1.101	4.826	2.0 %	1.0 %	3 %	77 %	17 %
4	QU-DSS-UFL4	1.087	4.911	2.0 %	1.0 %	3 %	77 %	17 %
5	QU-DSS-UFL5	1.102	4.541	2.0 %	0.0 %	2 %	78 %	18 %
6	QU-DSS-UFL6	1.031	4.441	3.0 %	0.0 %	3 %	77 %	17 %
			Average	**2**	**1**	**3**	**77**	**17**

Note: 1,2,3,4,and 5 means : Strongly Disagree, Disagree, Somewhat Agree, Agree, Strongly Agree

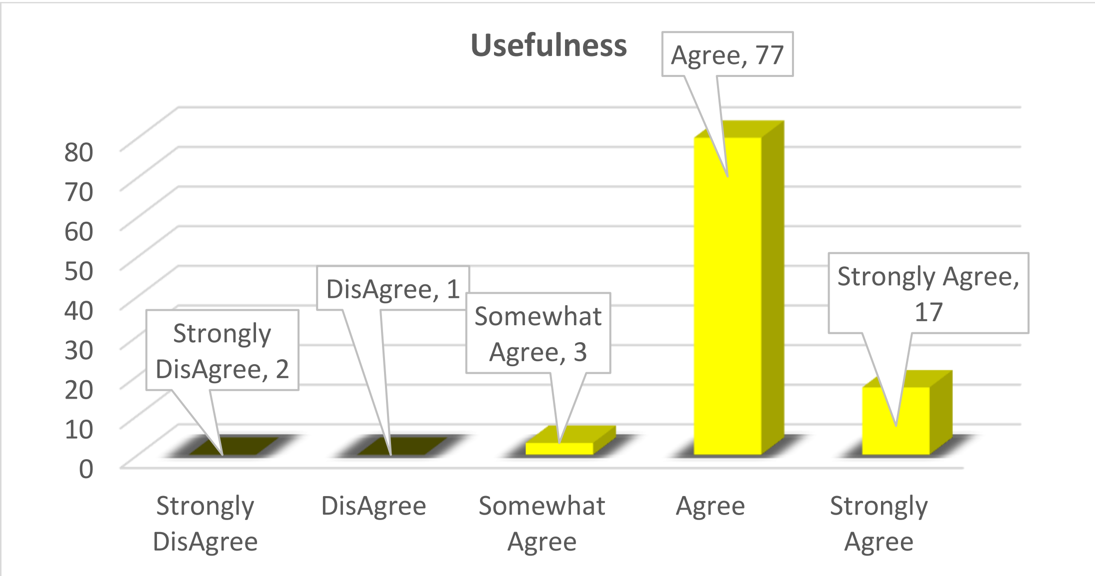

Figure 4. 7: Usefulness Dimension Findings

In terms of usefulness, the details of the findings are illustrated in Table 4.5 and Figure 4.7 above. The majority of participants either strongly agree or agree and few of them somewhat agree. Thus, the DSS application is measured by QU-DSS is useful.

Table 4. 6: Simplicity Dimension Findings

No	Simplicity Items	STD	M	1	2	3	4	5
				Frequency & Percentage				
1	QU-DSS-SMP1	1.103	4.740	1.0 %	0.0 %	6 %	77 %	16 %
2	QU-DSS-SMP2	1.145	4.274	0.0 %	2.0 %	7 %	74 %	17 %
3	QU-DSS-SMP3	1.055	4.427	1.0 %	2.0 %	7 %	74 %	17 %
4	QU-DSS-SMP4	1.387	4.713	0.0 %	2.0 %	7 %	75 %	16 %
5	QU-DSS-SMP5	1.215	4.642	1.0 %	2.0 %	5 %	74 %	18 %
6	QU-DSS-SMP6	1.065	4.148	1.0 %	0.0 %	7 %	74 %	16 %
			Average	**1**	**2**	**7**	**74**	**16**

Note: 1,2,3,4,and 5 means : Strongly Disagree, Disagree, Somewhat Agree, Agree, Strongly Agree

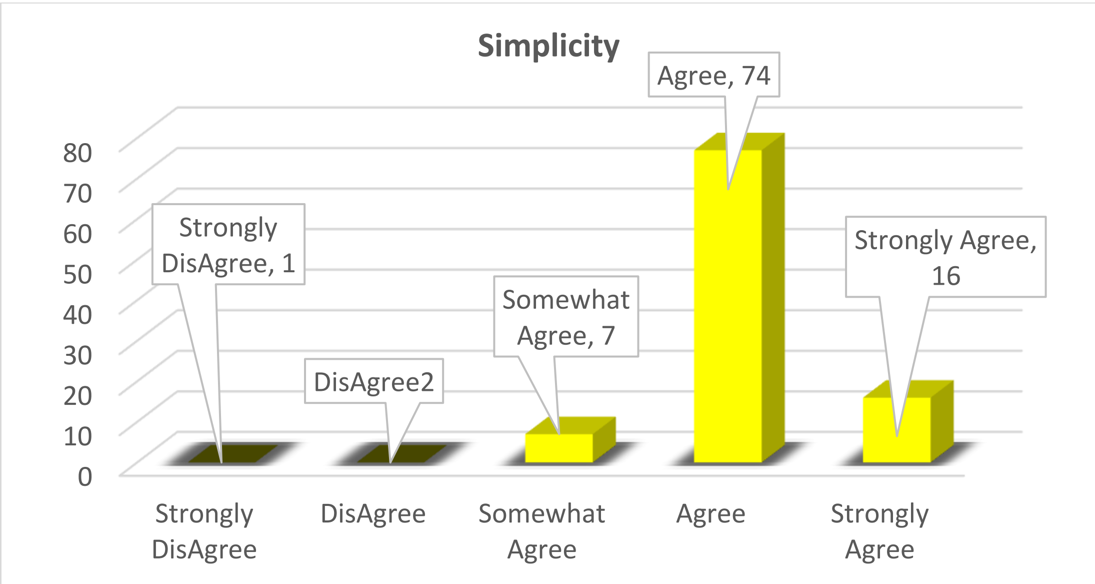

Figure 4. 8: Simplicity Dimension Findings

In terms of simplicity, the findings details are illustrated in Table 4.6 and Figure 4.8 above. The majority of participants either strongly agree or agree and few of them somewhat agree. Thus, the DSS application is measured by QU-DSS has the characteristic of simplicity.

Table 4. 7:Reliability Dimension Findings

No	Reliability Items	STD	M	1	2	3	4	5
				Frequency & Percentage				
1	QU-DSS-REL1	1.103	4.740	0.0 %	3.0 %	15 %	47 %	35 %
2	QU-DSS-REL2	1.105	4.274	0.0 %	0.0 %	15 %	50 %	35 %
3	QU-DSS-REL3	1.005	4.427	0.0 %	3.0 %	15 %	47 %	35 %
4	QU-DSS-REL4	1.307	4.713	0.0 %	0.0 %	15 %	50 %	35 %
5	QU-DSS-REL5	1.205	4.642	0.0 %	3.0 %	15 %	47 %	35 %
6	QU-DSS-REL6	1.005	4.148	0.0 %	0.0 %	15 %	50 %	32 %
			Average	**0**	**3**	**15**	**47**	**35**

Note: 1,2,3,4,and 5 means : Strongly Disagree, Disagree, Somewhat Agree, Agree, Strongly Agree

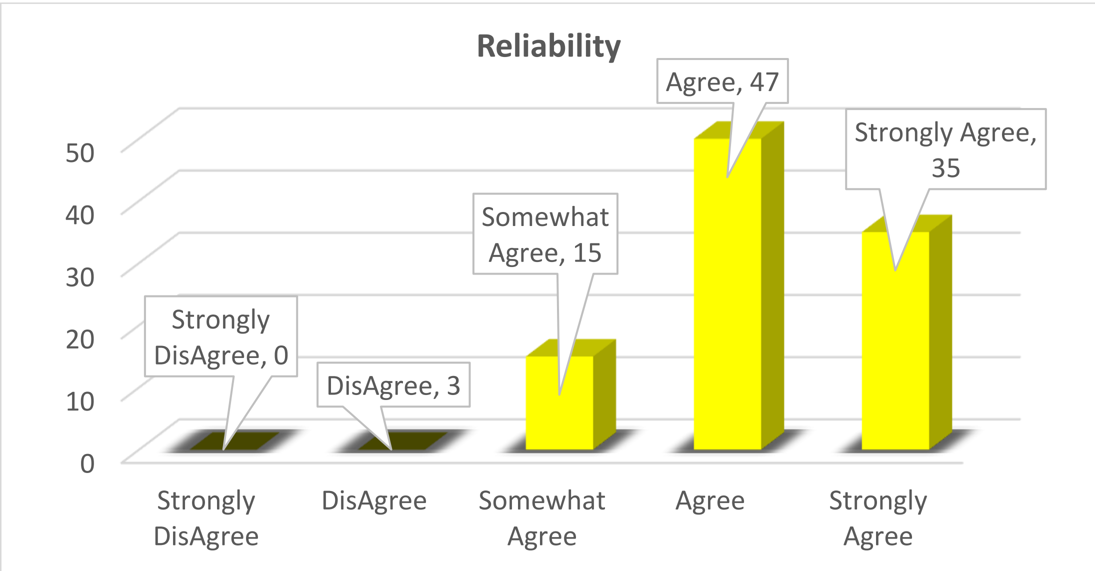

Figure 4. 9: Reliability Dimension Findings

In terms of reliability, the findings details are illustrated in Table 4.7 and Figure 4.9 The majority of participants either strongly agree or agree and few of them somewhat agree. Thus, the DSS application is measured by QU-DSS confirmed is reliable.

Table 4. 8: Decision-Support Dimension Findings

No	Decision-Support Items	STD	M	1	2	3	4	5
				Frequency & Percentage				
1	QU-DSS-DS1	1.001	4.942	2.0 %	1.0 %	10 %	62 %	25 %
2	QU-DSS-DS2	1.504	4.775	2.0 %	1.0 %	10 %	62 %	25 %
3	QU-DSS-DS3	1.115	4.827	2.0 %	1.0 %	11 %	62 %	24 %
4	QU-DSS-DS4	1.117	4.419	2.0 %	1.0 %	10 %	62 %	25 %
5	QU-DSS-DS5	1.212	4.547	2.0 %	1.0 %	10 %	62 %	25 %
6	QU-DSS-DS6	1.106	4.348	0.0 %	0.0 %	10 %	67 %	23 %
			Average	**2**	**1**	**10**	**62**	**25**

Note: 1,2,3,4,and 5 means : Strongly Disagree, Disagree, Somewhat Agree, Agree, Strongly Agree

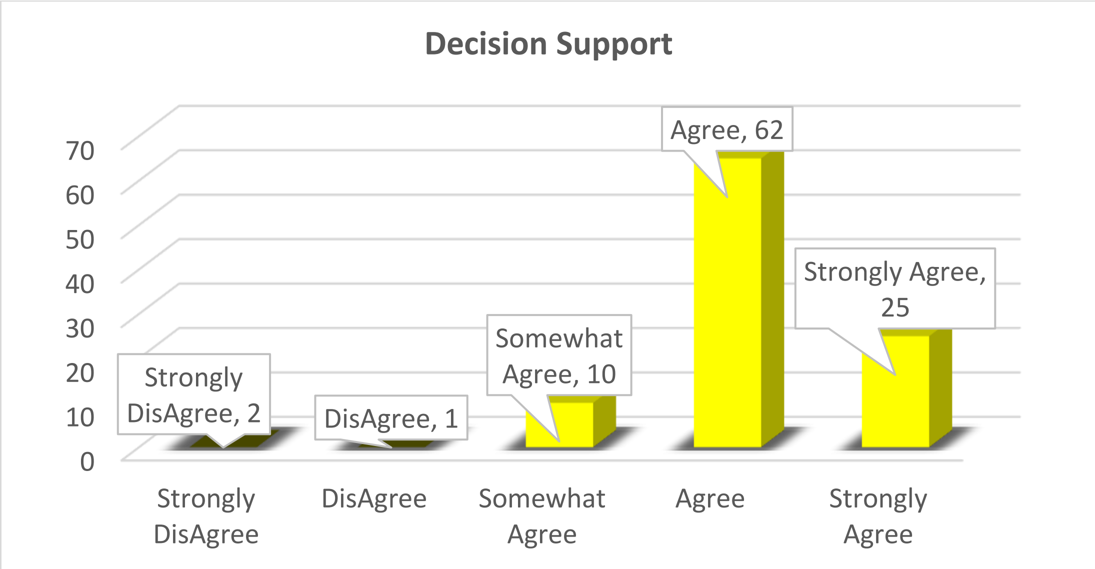

Figure 4. 10: Decision-Support Dimension Findings

In terms of decision-support, the findings details are illustrated in Table 4.8 and Figure 4.10. The majority of participants either strongly agree or agree and few of them somewhat agree. Thus, the DSS application is measured by QU-DSS can support decision-making process.

Table 4. 9: Flexibility Dimension Findings

No	Flexibility Items	STD	M	1	2	3	4	5
				Frequency & Percentage				
1	QU-DSS-FLX1	1.452	4.142	2.0 %	2.0 %	11 %	66 %	19 %
2	QU-DSS-FLX2	1.511	4.275	2.0 %	2.0 %	12 %	66 %	18 %
3	QU-DSS-FLX3	1.712	4.327	2.0 %	2.0 %	11 %	66 %	19 %
4	QU-DSS-FLX4	1.602	4.479	2.0 %	2.0 %	11 %	66 %	19 %
5	QU-DSS-FLX5	1.801	4.547	0.0 %	2.0 %	11 %	68 %	19 %
6	QU-DSS-FLX6	1.105	4.168	0.0 %	2.0 %	11 %	70 %	17 %
			Average	**2**	**2**	**11**	**66**	**19**

Note: 1,2,3,4,and 5 means : Strongly Disagree, Disagree, Somewhat Agree, Agree, Strongly Agree

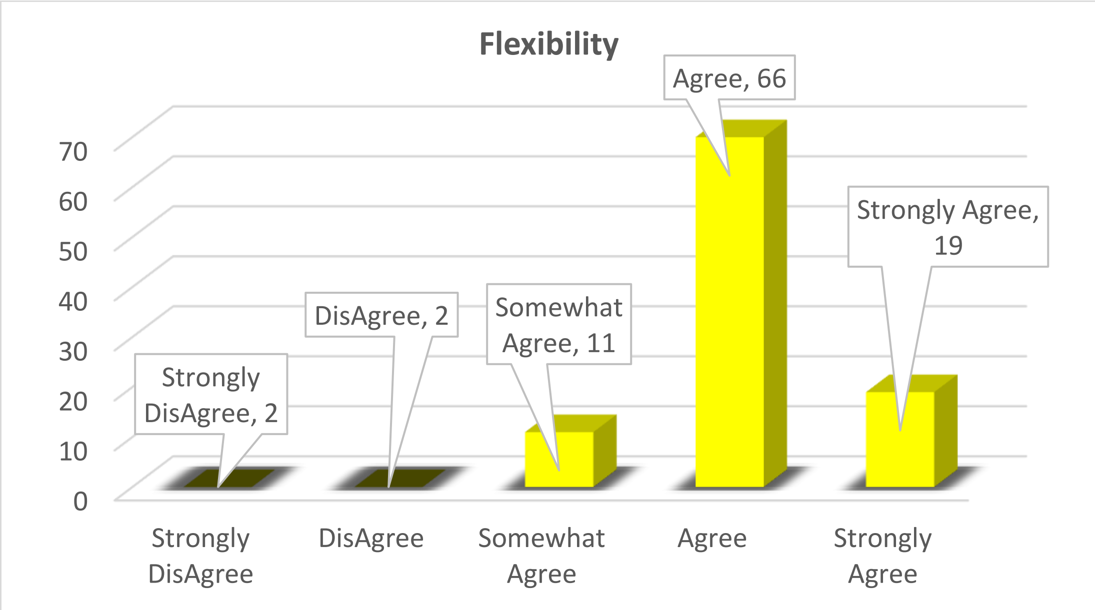

Figure 4. 11: Flexibility Dimension Findings

In terms of flexibility, the findings details are illustrated in Table 4.9 and Figure 4.11. The majority of participants either strongly agree or agree, and few of them somewhat agree. Thus, the DSS application is measured by QU-DSS are flexible.

In line with the above situation, the findings confirm that QU-DSS are usable in terms of flexibility, reliability, simplicity, usefulness, and decision-support.

The model validation and usability measurement outcomes demonstrated that all of the objectives of the study was met through the conceptual design model and the tests that were reviewed by users' experience and experts.

CHAPTER FIVE

CONCLUSION AND FUTURE WORK

5.1 Introduction

This chapter discusses the specifics of the research aims. The study focuses on creating an intelligent decision support model for healthcare data management using fog computing environment (IDSMFOG).

5.2 Conclusion

The primary determination of this research was to create data management model used for healthcare data management in a Fog computing environment that can ensure the value of organizational decision-making. This goal was set to meet the study's expectations.

The fog computing environment may be viewed as an acceptable and suitable option to manage and handle the enormous volume of data gathered from the IoT environment and prepare them for real-time DSS in applications.

To effectively handle the huge amount of data and take advantage of the capabilities of artificial intelligence algorithms, it is also vital and required to implement some artificial intelligence algorithms in fog computing, in this thesis k-means with two cluster and SVM algorithms are used and sore accuracy (99.87,99.97) for USA and Iraq example respectivelly.

As can be inferred from the research findings, the proposed model can be used as both a theoretical and practical guideline for developers in healthcare field to improve the performance of their decision support systems.

According to the findings, the suggested IDSMFOG model might be used to build and develop DSS systems that assist decision-making in a fog environment by providing selection with near real-time data.

5.3 Future Work

As a future work, It could be to improve the functioning of the existing healthcare system securely by applying encryption, as well as developing and assessing an algorithm of an enhanced scheme in a real-world healthcare environment. Also there is a plan to study the possibilities of homomorphism security mechanism to build secure healthcare application for the upcoming future generation. By managing, processing, and delivering healthcare data to be used by DSS applications in the fog computing environment, this research offers direction to data management developers for facing design, delivery, operational, and integration challenges for DSS data. They can also use other cluster algorithms, such as DBSCAN, compute network evaluation metrics and, deal with scalability issues.

REFERENCES

[1] S. Roy, R. Bose, and D. Sarddar, "A Fog-Based DSS Model for Driving Rule Violation Monitoring Framework on the Internet of Things," *Int. J. Adv. Sci. Technol.*, vol. 82, pp. 23–32, 2015, doi: 10.14257/ijast.2015.82.03.

[2] G. Phillips-Wren, "AI tools in decision making support systems: A review," *Int. J. Artif. Intell. Tools*, vol. 21, no. 2, 2012, doi: 10.1142/S0218213012400052.

[3] L. Wang, "Intelligent decision support system of emergency language based on fog computing," *Math. Probl. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/6611501.

[4] A. Id Hadiana, "Fog Computing Architecture for Indoor Disaster Management," *Int. J. Informatics, Inf. Syst. Comput. Eng.*, vol. 1, no. 1, 2020, doi: 10.34010/injiiscom.v1i1.4022.

[5] A. R. Dar, D. Ravindran, and S. Islam, "Fog-based spider web algorithm to overcome latency in cloud computing," *Iraqi J. Sci.*, vol. 61, no. 7, pp. 1781–1790, 2020, doi: 10.24996/ijs.2020.61.7.27.

[6] G. Neelakantam, D. D. Onthoni, and P. K. Sahoo, "Fog computing enabled locality based product demand prediction and decision making using reinforcement learning," *Electron.*, vol. 10, no. 3, pp. 1–16, 2021, doi: 10.3390/electronics10030227.

[7] A. Yousefpour *et al.*, "All one needs to know about fog computing and related edge computing paradigms: A complete survey," *J. Syst. Archit.*, vol. 98, no. February, pp. 289–330, 2019, doi: 10.1016/j.sysarc.2019.02.009.

[8] V. Hurbungs, V. Bassoo, and T. P. Fowdur, "Fog and edge computing: concepts, tools and focus areas," *Int. J. Inf. Technol.*, vol. 13, no. 2, pp. 511–522, 2021, doi: 10.1007/s41870-020-00588-5.

[9] A. Kumari, S. Tanwar, S. Tyagi, and N. Kumar, "Fog computing for Healthcare 4.0 environment: Opportunities and challenges," *Comput. Electr. Eng.*, vol. 72, pp. 1–13, 2018, doi: 10.1016/j.compeleceng.2018.08.015.

[10] S. Shukla, M. F. Hassan, M. K. Khan, L. T. Jung, and A. Awang, *An analytical model to minimize the latency in healthcare internet-of-things in fog computing environment*, vol. 14, no. 11. 2019.

[11] G. Yoon, D. Choi, J. Lee, and H. Choi, "Management of IoT Sensor Data Using a Fog Computing Node," *J. Sensors*, vol. 2019, 2019, doi: 10.1155/2019/5107457.

[12] F. Mehdipour, B. Javadi, A. Mahanti, and G. Ramirez-Prado, "Fog computing realization for big data analytics," *Fog Edge Comput. Princ. Paradig.*, pp. 261–290, 2019, doi: 10.1002/9781119525080.ch11.

[13] X. Liu, R. H. Deng, Y. Yang, H. N. Tran, and S. Zhong, "Hybrid privacy-preserving clinical decision support system in fog–cloud computing," *Futur. Gener. Comput. Syst.*, vol. 78, pp. 825–837, 2018, doi: 10.1016/j.future.2017.03.018.

[14] A. M. Ghosh and K. Grolinger, "Deep Learning: Edge-Cloud Data Analytics for IoT," *2019 IEEE Can. Conf. Electr. Comput. Eng. CCECE 2019*, 2019, doi: 10.1109/CCECE.2019.8861806.

[15] M. Ahmed, R. Mumtaz, S. M. H. Zaidi, M. Hafeez, S. A. R. Zaidi, and M. Ahmad, "Distributed fog computing for internet of things (Iot) based ambient data processing and analysis," *Electron.*, vol. 9, no. 11, pp. 1–20, 2020, doi: 10.3390/electronics9111756.

[16] H. Alshammari, S. A. El–Ghany, and A. Shehab, "Big IoT Healthcare Data Analytics Framework Based on Fog and Cloud Computing," *J. Inf. Process.*

Syst., vol. 16, no. 6, pp. 1238–1249, 2020, doi: 10.3745/JIPS.04.0193.
[17] A. Sutagundar and P. Sangulagi, “Fog computing based information classification in sensor cloud-agent approach,” *Expert Syst. Appl.*, vol. 182, no. May, p. 115232, 2021, doi: 10.1016/j.eswa.2021.115232.
[18] N. Bv and R. M. R. Guddeti, “Fog-based Intelligent Machine Malfunction Monitoring System for Industry 4.0,” *IEEE Trans. Ind. Informatics*, vol. 3203, no. c, pp. 1–10, 2021, doi: 10.1109/TII.2021.3056076.
[19] D. Aishwarya and R. I. Minu, “Edge computing based surveillance framework for real time activity recognition,” *ICT Express*, vol. 7, no. 2, pp. 182–186, 2021, doi: 10.1016/j.icte.2021.04.010.
[20] N. Al Mudawi, “Integration of IoT and Fog Computing in Healthcare Based the Smart Intensive Units,” *IEEE Access*, vol. 10, pp. 1–1, 2022, doi: 10.1109/access.2022.3179704.
[21] V. Colombo, A. Tundo, M. Ciavotta, and L. Mariani, *Towards Self-Adaptive Peer-to-Peer Monitoring for Fog Environments*, vol. 1, no. 1. Association for Computing Machinery, 2022.
[22] M. J. Druzdzel and R. R. Flynn, “Decision support systems,” *Underst. Inf. Retr. Syst. Manag. Types, Stand.*, pp. 461–472, 2011, doi: 10.1177/0193841x8500900105.
[23] Maria Rashidi, Maryam Ghodrat, Bijan Samali and Masoud Mohammadi “ Decision support system factors,” pp.19-38, 2018 , doi:10.5772/intechopen.79390.
[24] A. Kaklauskas, “Intelligent decision support systems,” *Intell. Syst. Ref. Libr.*, vol. 81, pp. 31–85, 2015, doi: 10.1007/978-3-319-13659-2_2.
[25] D. J. Power, “Specifying An Expanded Framework for Classifying and DescribingDecision Support Systems,” *Commun. Assoc. Inf. Syst.*, vol. 13, no. February, 2004, doi: 10.17705/1cais.01313.
[26] S. Gupta, S. Modgil, S. Bhattacharyya, and I. Bose, “Artificial intelligence for decision support systems in the field of operations research: review and future scope of research,” *Ann. Oper. Res.*, vol. 308, no. 1–2, pp. 215–274, 2022, doi: 10.1007/s10479-020-03856-6.
[27] D. Coleman, “Decision support systems,” *Data Process.*, vol. 26, no. 8, pp. 35–36, 1984, doi: 10.1016/0011-684X(84)90199-0.
[28] D. R. Bestak and D. S. Smys, “Big Data Analytics for Smart Cloud-Fog Based Applications,” *J. Trends Comput. Sci. Smart Technol.*, vol. 2019, no. 02, pp. 74–83, 2019, doi: 10.36548/jtcsst.2019.2.001.
[29] S. Tuli *et al.*, “HealthFog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments,” *Futur. Gener. Comput. Syst.*, vol. 104, pp. 187–200, 2020, doi: 10.1016/j.future.2019.10.043.
[30] T. S. Nikoui, A. M. Rahmani, and H. Tabarsaied, “Data management in fog computing,” *Fog Edge Comput. Princ. Paradig.*, pp. 171–190, 2019, doi: 10.1002/9781119525080.ch8.
[31] K. P. Yadav, S. Sharma, A. S. KIET, and Ghaziabad, “2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT),” pp. 0–3, 2019.
[32] Ş. M. Kaya, A. Erdem, and A. Güneş, “A Smart Data Pre-Processing Approach to Effective Management of Big Health Data in IoT Edge,” *Smart Homecare Technol. TeleHealth*, vol. Volume 8, pp. 9–21, 2021, doi: 10.2147/shtt.s313666.
[33] A. Ramesh, S. Rajkumar, and J. L. M. Livingston, “Disaster management in smart cities using IoT and big data,” *J. Phys. Conf. Ser.*, vol. 1716, no. 1, 2021,

doi: 10.1088/1742-6596/1716/1/012060.
[34] B. Bin Ma, S. Fong, and R. Millham, "Data stream mining in fog computing environment with feature selection using ensemble of swarm search algorithms," *2018 Conf. Inf. Commun. Technol. Soc. ICTAS 2018 - Proc.*, pp. 1–6, 2018, doi: 10.1109/ICTAS.2018.8368770.
[35] R. Shehab, M. Taher, and H. K. Mohamed, "L IVE B IG D ATA A NALYTICS R ESOURCE M ANAGEMENT T ECHNIQUES IN F OG C OMPUTING FOR T ELE - HEALTH A PPLICATIONS," vol. 07, no. 01, pp. 89–103, 2021.
[36] K. K. Al-jabery, T. Obafemi-Ajayi, G. R. Olbricht, and D. C. Wunsch II, "Data preprocessing," *Comput. Learn. Approaches to Data Anal. Biomed. Appl.*, pp. 7–27, 2020, doi: 10.1016/b978-0-12-814482-4.00002-4.
[37] R. Dautov, S. Distefano, and R. Buyya, "Hierarchical data fusion for Smart Healthcare," *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0183-6.
[38] A. I. Montoya-Munoz and O. M. C. Rendon, "An approach based on fog computing for providing reliability in iot data collection: A case study in a colombian coffee smart farm," *Appl. Sci.*, vol. 10, no. 24, pp. 1–16, 2020, doi: 10.3390/app10248904.
[39] I. Ali, I. Ahmedy, A. Gani, M. Talha, M. A. Raza, and M. H. Anisi, "Data collection in sensor-cloud: A systematic literature review," *IEEE Access*, vol. 8, pp. 184664–184687, 2020, doi: 10.1109/ACCESS.2020.3029597.
[40] A. Mijuskovic, R. Bemthuis, A. Aldea, and P. Havinga, "An Enterprise Architecture based on Cloud, Fog and Edge Computing for an Airfield Lighting Management System," *Proc. - IEEE Int. Enterp. Distrib. Object Comput. Work. EDOCW*, vol. 2020-Octob, pp. 63–73, 2020, doi: 10.1109/EDOCW49879.2020.00021.
[41] Z. Zou, Y. Jin, P. Nevalainen, Y. Huan, J. Heikkonen, and T. Westerlund, "Edge and Fog Computing Enabled AI for IoT-An Overview," *Proc. 2019 IEEE Int. Conf. Artif. Intell. Circuits Syst. AICAS 2019*, pp. 51–56, 2019, doi: 10.1109/AICAS.2019.8771621.
[42] J. Queiroz, P. Leitão, J. Barbosa, and E. Oliveira, "Distributing intelligence among cloud, fog and edge in industrial cyber-physical systems," *ICINCO 2019 - Proc. 16th Int. Conf. Informatics Control. Autom. Robot.*, vol. 1, no. Icinco, pp. 447–454, 2019, doi: 10.5220/0007979404470454.
[43] P. K. S. P, M. S. Niharika, M. P. Niharika, B. Harshitha, and N. R. T. G, "Survey on Fog Computing in Healthcare Monitoring," vol. 10, no. 05, pp. 647–652, 2021.
[44] Anupriya and A. Singhrova, "Enhanced whale optimization based traffic forecasting for SDMN based traffic," *ICT Express*, vol. 7, no. 2, pp. 143–151, 2021, doi: 10.1016/j.icte.2021.05.005.
[45] M. Abdel-Basset, G. Manogaran, A. Gamal, and V. Chang, "A Novel Intelligent Medical Decision Support Model Based on Soft Computing and IoT," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4160–4170, 2020, doi: 10.1109/JIOT.2019.2931647.
[46] A. Hassouneh, A. M. Mutawa, and M. Murugappan, "Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG based on machine learning and deep neural network methods," *Informatics Med. Unlocked*, vol. 20, p. 100372, 2020, doi: 10.1016/j.imu.2020.100372.
[47] C. F. Zucker, C. Ghidini, A. Idrissou, F. Van Harmelen, and P. Van Den Besselaar, "Network metrics for assessing the quality of entity resolution between multiple datasets," *Semant. Web*, vol. 12, no. 1, pp. 21–40, 2020, doi: 10.3233/SW-200410.

[48] C. I. Nwakanma, J. W. Kim, J. M. Lee, and D. S. Kim, “Edge AI prospect using the NeuroEdge computing system: Introducing a novel neuromorphic technology,” *ICT Express*, vol. 7, no. 2, pp. 152–157, 2021, doi: 10.1016/j.icte.2021.05.003.

[49] N. S. Aziz, N. S. Sulaiman, W. N. I. T. M. Hassan, N. L. Zakaria, and A. Yaacob, “A Review of Website Measurement for Website Usability Evaluation,” *J. Phys. Conf. Ser.*, vol. 1874, no. 1, 2021, doi: 10.1088/1742-6596/1874/1/012045.

[50] Z. T. Al-Ars and A. Al-Bakry, “A web/mobile decision support system to improve medical diagnosis using a combination of K-mean and fuzzy logic,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 6, pp. 3145–3154, 2019, doi: 10.12928/TELKOMNIKA.v17i6.12715.

[51] D. L. Berry *et al.*, “Development and evaluation of the personal Patient Profile-Prostate (P3P), a Web-based decision support system for men newly diagnosed with localized prostate cancer,” *J. Med. Internet Res.*, vol. 12, no. 4, pp. 1–13, 2010, doi: 10.2196/jmir.1576.

[52] Q. Zhou, Z. Yin, Q. Ying, and S. Wang, “Intelligent Data Mining and Decision System for Commercial Decision Making,” *TELKOMNIKA Indones. J. Electr. Eng.*, vol. 12, no. 1, 2014, doi: 10.11591/telkomnika.v12i1.3977.

الخلاصة

في الآونة الأخيرة ، حدثت العديد من حالات الطوارئ المتعلقة بالرعاية الصحية على مستوى العالم ، والتي تطلبت وأجبرت في بعض الأحيان صانعي القرار المعنيين على إيلاء اهتمام كبير لعملية صنع القرار. في بنية نظام دعم القرار ، تعتبر البيانات مكونًا أساسيًا بسبب نظام دعم القرار الذي يجمع البيانات والمعرفة من مختلف المجالات والمصادر ومعالجتها وتخزينها في السحابة لتزويد المستخدمين المعنيين بمعلومات مفيدة لدعم القرار, ومن ثم ، يتعامل نظام دعم القرار مع البيانات السحابية والطرق التقليدية لتخزين البيانات الضخمة في السحابة لها قيود ، حيث يؤدي التعامل مع مثل هذه الأحجام الهائلة من البيانات إلى زيادة الأخطاء وإسقاط الحزم واحتمال حدوث اختناقات في البيانات. نتيجة لذلك ، من المناسب والضروري استخدام خوارزميات التعلم الآلي في تصميم وتطوير نظام دعم اتخاذ القرار الذكي في سياق الحوسبة الضبابية. اقترح هذا البحث نموذجًا ذكيًا لنظام دعم القرار في إدارة بيانات الرعاية الصحية استنادًا إلى بيئة الحوسبة الضبابية (IDSMFOG) باستخدام خوارزميات (K-mean + SVM) ، وأداة تصميم لقياس قابلية استخدام النموذج المقترح المسمى استبيان قابلية استخدام النظام (QU- DSS)تم تطوير نظام دعم القرار على أساس النموذج المقترح ، وبعد ذلك تم تقييمه أيضًا من حيث قابليته للاستخدام بواسطة أداة (QU-DSS). وأكدت النتائج التي تم جمعها أن دقة النظام المطور في المثالين الولايات المتحدة والعراق هي (99.87 و 99.97) على التوالي ، وكذلك أكدت نتائج (QU-DSS) أن النظام المطور قابل للاستخدام ويدعم عملية اتخاذ القرار.

جامعة كربلاء

كلية علوم الحاسوب وتكنولوجيا المعلومات

قسم علوم الحاسوب

نموذج دعم القرار الذكي لإدارة البيانات في القطاع الصحي بالإعتماد على الحوسبة الضبابية

رسالة ماجستير

مقدمة الى مجلس كلية علوم الحاسوب وتكنولوجيا المعلومات في جامعة كربلاء وهي جزء من متطلبات نيل درجة الماجستير في علوم الحاسوب

كتبت بواسطة

أوراس عبد الخضر حسين راضي

بإشـراف

أ.م.د. أياد حميد موسى

1444 هـ **2023 م**