



PROJECT REPORT

Implementation Analysis of Synthetic Data Vault for Medical Workforce Number Prediction

BINAR EKO HARIJANTO

18.K1.0046

**Faculty of Computer Science
Soegijapranata Catholic University**

2022

HALAMAN PENGESAHAN



Judul Tugas Akhir: : Implementation Analysis of Synthetic Data Vault for Medical Workforce
Number Prediction

Diajukan oleh : Binar Eko Harijanto

NIM : 18.K1.0046

Tanggal disetujui : 21 Juni 2022

Telah setuju oleh

Pembimbing : Rosita Herawati S.T., M.I.T.

Penguji 1 : Yonathan Purbo Santosa S.Kom., M.Sc

Penguji 2 : Rosita Herawati S.T., M.I.T.

Penguji 3 : Yulianto Tejo Putranto S.T., M.T.

Penguji 4 : Hironimus Leong S.Kom., M.Kom.

Penguji 5 : Y.b. Dwi Setianto S.T., M.Cs.

Penguji 6 : R. Setiawan Aji Nugroho S.T., MCompIT., Ph.D

Ketua Program Studi : Rosita Herawati S.T., M.I.T.

Dekan : Dr. Bernardinus Harnadi S.T., M.T.

Halaman ini merupakan halaman yang sah dan dapat diverifikasi melalui alamat di bawah ini.

sintak.unika.ac.id/skripsi/verifikasi/?id=18.K1.0046

DECLARATION OF AUTHORSHIP

I, the undersigned:

Name : Binar Eko Harijanto

ID : 18.K1.0046

Declare that this work, titled "Implementation Analysis of Synthetic Data Vault for Medical Workforce Number Prediction", and the work presented in it is my own. I confirm that:

- 1 This work was done wholly or mainly while in candidature for a research degree at Soegijapranata Catholic University
- 2 Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- 3 Where I have consulted the published work of others, this is always clearly attributed.
- 4 Where I have quoted from the work of others, the source is always given.
- 5 Except for such quotations, this work is entirely my own work.
- 6 I have acknowledged all main sources of help.
- 7 Where the work is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Semarang, June, 29, 2022



Binar Eko Harijanto

18.K1.0046

DECLARATION OF RESEARCH PUBLICATION FOR ACADEMIC PURPOSE

I, the undersigned:

Name : Binar Eko Harijanto
Major : Information Engineering
Faculty : Computer Science
Type of Work : Final Project

Agreed to give Soegijapranata Catholic University Non-exclusive Free Royalty Rights over scientific work titled “Implementation Analysis of Synthetic Data Vault for Medical Workforce Number Prediction”. With Non-exclusive Free Royalty Rights, Soegijapranata Catholic University is allowed to save, move/make change, manage in the form of database, maintain, and publish this final project as long as my name is included as the writer/owner of the copyright.

Semarang, July, 16, 2022



Binar Eko Harijanto

18.K1.0046

ABSTRACT

The Covid 19 pandemic has proven us the importance of medical workforce distribution, as insufficiency of health professionals may lead into patient abandonments, eventually casualties. Moreover, healthcare would be more effective if the government have access to future medical workforce numbers. Unfortunately, the implementation of prediction algorithm within the field is not yet present, and high quality medical workforce data in Indonesia are rare.

This research approaches said problem by utilizing Support Vector Regression, and Random Forest algorithm to predict future numbers of medical workforce within Semarang city. To fight data scarcity, Synthetic Data Vault technique is implemented to substitute the real dataset. The results are in the form of time series data prediction and accuracy tests using MSE (Mean Square Error) and MAPE (Mean Absolute Percentage Error) to compare the performance of presented methods.

Keywords:

Prediction Algorithm, Synthetic Data, Medical Workforce

TABLE OF CONTENTS

COVER	i
APPROVAL AND RATIFICATION PAGE	ii
DECLARATION OF AUTHORSHIP	iii
DECLARATION OF RESEARCH PUBLICATION	iv
ABSTRACT	v
TABLE OF CONTENTS	vi
LIST OF FIGURES	viii
LIST OF TABLES	ix
1. INTRODUCTION	1
1.1. Background	1
1.2. Problem Formulation	2
1.3. Scope	2
1.4. Objectives	2
2. LITERATURE STUDY.....	3
3. RESEARCH METHODOLOGY.....	7
4. ANALYSIS AND DESIGN	9
4.1 Data Acquisition and Preprocessing	9
4.2 Workflow	11
4.3 Methods and Formula	13
5. IMPLEMENTATION AND RESULT	15
5.1 Implementation	15
5.2 Result	20
6. CONCLUSION.....	26
REFERENCE.....	277



LIST OF FIGURES

Figure 1	20
Figure 2	21
Figure 3	25



LIST OF TABLES

Table 1	9
Table 2	10
Table 3	19
Table 4	22
Table 5	22
Table 6	23
Table 7	23
Table 8	23
Table 9	24
Table 10	24

