

**A MULTI-CRITERIA RANKING SYSTEM FOR EVALUATING GENE
REGULATORY NETWORK INFERENCE METHODS**

**KANONTE, MARIAM
(21PCG02288)**

**B.Sc Software Engineering, Institut de Formation et de Recherche en
Informatique, Université d'Abomey-Calavi, Abomey-Calavi – Benin Republic**

APRIL, 2024

**A MULTI-CRITERIA RANKING SYSTEM FOR EVALUATING GENE
REGULATORY NETWORK INFERENCE METHODS**

BY

**KANONTE, MARIAM
(21PCG02288)**

**B.Sc Software Engineering, Institut de Formation et de Recherche en
Informatique, Université d'Abomey-Calavi, Abomey-Calavi – Benin
Republic**

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

APRIL, 2024

ACCEPTANCE

This is to attest that this dissertation has been accepted in partial fulfilment of the requirements for the award of the degree of Master of Science in Computer Science in the Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria.

Miss Adefunke F. Oyinloye
(Secretary, School of Postgraduate Studies)

Signature and Date

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

Signature and Date

DECLARATION

I, **KANONTE, MARIAM (21PCG02288)** declare that this dissertation is a representation of my work and is written and implemented by me under the supervision of Dr. Itunuoluwa M. Isewon of the Department of Computer and Information Sciences, Covenant University, Ota, Nigeria. I attest that this dissertation has in no way been submitted either wholly or partially to any other university or institution of higher learning for the award of a master's degree. All information cited from published and unpublished literature has been duly referenced.

KANONTE, MARIAM

Signature and Date

CERTIFICATION

We certify that the thesis titled “**A MULTI-CRITERIA RANKING SYSTEM FOR EVALUATING GENE REGULATORY NETWORK INFERENCE METHODS**” is the original work carried out by **KANONTE, MARIAM, (21PCG02288)** in the Department of Computer and Information Sciences, Covenant University, Ota, Ogun State, Nigeria under the supervision of Dr. Itunuoluwa M. Isewon. We have examined and found the work acceptable for the award of the degree of Master of Science (M.Sc) in Computer Science.

Dr. Itunuoluwa M. Isewon
(Supervisor)

Signature and Date

Prof. Olufunke O. Oladipupo
(Head of Department)

Signature and Date

Prof. Adebukola S. Onashoga
(External Examiner)

Signature and Date

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

Signature and Date

DEDICATION

This research work is dedicated to God Almighty, the custodian of all wisdom, knowledge, and understanding, for His grace and favour throughout the execution of this research.

ACKNOWLEDGEMENTS

I am deeply grateful to the Almighty for the grace and privilege of being where I stand today.

I extend my sincere appreciation to the Chancellor of this esteemed university, Dr David O. Oyedepo. Through his teachings, he imparts invaluable values essential for a well-rounded and successful life. I am indebted to him for instilling in me the unwavering belief that excuses have no place in the pursuit of success. I also wish to convey my gratitude to the Vice-Chancellor Prof. Abiodun H. Adebayo for wholeheartedly embracing the Chancellor's vision and for his inspiring leadership. My heartfelt thanks also extend to the Dean of the School of Postgraduate Studies, Prof. Akan B. Williams, whose profound commitment, and leadership foster an environment conducive to student success. I am equally grateful to all members of the administration for their dedication to the University's mission.

I express my gratitude to the Covenant Applied Informatics and Communication Africa Centre of Excellence (CApIC-ACE) for generously sponsoring my education. I also wish to acknowledge the head of the Department of Computer and Information Sciences, Prof. Olufunke O. Oladipupo, for her steadfast dedication and commitment to the success of his students. My profound appreciation goes to my supervisor, Dr. Isewon Itunuoluwa, a paragon of excellence and inspiration, for her invaluable guidance, patience, and for instilling in me the values of diligence and meticulousness.

I extend my thanks to Dr. Ratheil Houndji, my mentor, whose passion for programming, data, and meticulous work ethic has inspired not only me but also countless other students. His unflinching support and encouragement have been a constant throughout this journey. I am equally grateful to Vincent-Beni Zossou, my steadfast companion and lifelong friend, whose unwavering belief in me and unflagging support have been a beacon of strength through every challenge.

To my parents, Kanoute Irake and Adjanohoun Evelyne, I owe a debt of gratitude for their boundless love, presence, steadfast encouragement, and unshakeable faith in every decision I make. Their prayers, sacrifices, and spirited determination to see me succeed in all my endeavors are deeply appreciated. I also extend my heartfelt thanks to my brother, Kanoute Moussa, for his support, belief in me, and constant encouragement.

I am immensely grateful for the opportunity to have crossed paths with remarkable individuals whom I now proudly call my brothers and sisters from different corners of the globe: Patrick, Joaquim, Myriam, Daniella, Sonia, Triumphant, Stephen, Daniel. My heartfelt gratitude goes to these wonderful individuals who have been steadfast companions throughout this journey, as well as to Emmanuel, Faith, Olajumoke, Chioma, with whom I share the workspace. I am grateful to Samuel for his friendship and formidable support throughout this adventure. To my dear friends Vershima and Danielle, your love and support have been a constant source of strength. You are simply the best. I love you all.

TABLE OF CONTENTS

CONTENTS	PAGES
COVER PAGE	i
TITLE PAGE	ii
ACCEPTANCE	iii
DECLARATION	iv
CERTIFICATION	v
DEDICATION	vi
ACKNOWLEDGEMENTS	vii
TABLE OF CONTENTS	ix
LIST OF FIGURES	xii
LIST OF TABLES	xiii
LIST OF ABBREVIATION	xiv
ABSTRACT	xvi
CHAPTER ONE : INTRODUCTION	1
1.1 Background of the study	1
1.2 Statement of the research problem	3
1.3 Aim and objectives of the study	5
1.4 Significance of the study	5
1.5 Scope of the study	5
CHAPTER TWO : LITERATURE REVIEW	6
2.1 Preamble	6
2.2 Multi-Criteria Decision Model	6
2.2.1 Analytic Hierarchy Process	6
2.2.2 VIKOR	7
2.2.3 PROMETHEE	8
2.3 Gene regulatory network inference methods	9
2.3.1 Correlation-based methods	11
2.3.2 Bayesian-based methods	12
2.3.3 Regression-based methods	12
2.3.4 Machine learning based methods	13
2.4 Types of datasets	14
2.4.1 Ribonucleic sequencing data	14
2.4.2 Microarray gene expression data	15
2.4.3 ChIP-seq data	16
2.4.4 Spatial transcriptomics data	17

2.5	Existing gene regulatory inference methods	18
2.5.1	Regression based methods	18
2.5.2	Bayesian-based methods	22
2.5.3	Correlation based methods	24
2.5.4	Machine learning based methods	26
2.6	Evaluation of gene regulatory network inference methods	28
2.7	Existing web-based gene regulatory networks inference methods	29
2.7.1	CORNET (CORrelation NETworks)	29
2.7.2	AraNet	30
2.7.3	SoyNet	31
2.8	Existing benchmarking framework for the inference gene regulatory network	31
2.8.1	NetBenchmark	32
2.8.2	BEELINE	33
2.9	Existing web-based benchmarking frameworks for gene regulatory inference methods	34
2.9.1	GeNeCK	35
2.9.2	GRNbenchmark	35
2.9.3	Summary of findings	37
CHAPTER THREE : METHODOLOGY		38
3.1	Preamble	38
3.2	Develop a multi-criteria ranking algorithm for evaluating gene regulatory network inference methods	38
3.3	Design of the framework for benchmarking and ranking GRN inference methods	39
3.3.1	Benchmarking framework	39
3.3.2	Architecture of the framework	44
3.4	Implementation of the framework	44
3.4.1	Machine learning methods	44
3.4.2	New GRN inference methods identified	46
3.5	Containerization of the framework	53
3.5.1	Adding a new GRN inference method	53
3.5.2	Adding a new evaluation technique	54
3.5.3	Integration of visualisation tools	55
3.6	Development of the web application	55
3.6.1.	Front-end	55

3.6.2	Back-end	56
3.6.3	Integrated Development Environment	57
3.6.4	Hosting the web application	57
3.7	Conclusion	59
CHAPTER FOUR : RESULTS AND DISCUSSION		61
4.1	Preamble	60
4.2	Develop a multi-criteria ranking algorithm for evaluating GRNi methods	60
4.3	Design a framework for benchmarking and ranking GRNi methods	63
4.4	Implement and containerize the framework	64
4.4.1	Containerization of the GRN Inference Methods	64
4.4.2	Data visualisation	65
4.5	Develop a web application using the framework	65
4.5.1	Home page	66
4.5.2	Login and registration pages	67
4.5.3	Inference page	68
4.5.4	Evaluation page	69
4.5.5	Visualisation page	70
4.5.6	Dashboard page of the user	71
CHAPTER FIVE : CONCLUSION AND RECOMMENDATIONS		74
5.1	Summary	73
5.2	Conclusion	74
5.3	Contribution to knowledge	74
5.4	Recommendation	74
REFERENCES		75

LIST OF FIGURES

FIGURES	TITLE OF Figure S	PAGES
1.1	Types of MCDM	2
2.1	Overview of the Net Benchmark Framework	33
2.2	Overview of the GRNbenchmark Framework	37
3.1	Overview of the Beeline Framework	42
3.2	GReNaDIne workflow	43
3.3	PRISMA Flowchart for the selection of the machine learning methods	46
3.5	PFBNET workflow	49
3.6	PFBNET workflow	50
3.7	Venn diagram of the methods identified	53
3.8	The workflow of GRN Evaluator	58
4.2	Repository of the images on docker hub	65
4.3	Home page of GRN Evaluator	66
4.4	Login page of GRN Evaluator	67
4.5	Registration page of GRN Evaluator	68
4.6	Inference page of GRN Evaluator	69
4.7	Evaluation page of GRN Evaluator	70
4.8	Visualisation page of GRN Evaluator	71
4.9	Dashboard page of GRN Evaluator. a) Truncated upper view b) Extended view	72

LIST OF TABLES

TABLES	TITLE OF Table	PAGES
2.1	MCDM approaches along with their domains of application	10
2.2	Regression-based GRN inference methods	22
2.3	Bayesian-based GRN inference methods	24
2.4	Correlation-based GRN inference methods	26
2.5	Machine learning based GRN inference methods	28
3.1	Inclusion and exclusion criteria	45
3.2	Concise Description of the methods identified	51

LIST OF ABBREVIATION

AHP	Analytic Hierarchy Process
ARIMA	Autoregressive integrated moving average
AUPRC	The Area Under the Precision-Recall Curve
AUROC	Area Under the Receiver Operating Characteristics Curve
BEELINE	Benchmarking gEnE reguLatory network Inference from siNgle-cEll transcriptomic data
BN	Bayesian network
CGH	Comparative genomic hybridization
CPU	Central processing unit
CSS	Cascading style sheets
DBN	Dynamic Bayesian Network
DDos	Distributed denial of service
DNA	Deoxyribonucleic acid
EP	Early precision
EPR	Early precision ratio
GReNaDine	Gene Regulatory Network Data-driven Inference
GRN	Gene Regulatory Network
GRNi	Gene Regulatory Network inference
GUIs	Graphical user interfaces
HTTP	Hypertext transport protocol
IDE	Integrated Development Environment
KS	Kolmogorov Smirnov
LR	Logistic Regression

MADM	Multi-attribute decision making
MCDM	Multi-Criteria Decision Model
MI	Mutual information
ML	Machine Learning
MLR	Multiple Linear Regression
MODM	Multi-objective decision making
PLS	Partial least squares
PRISMA	Preferred reporting items for systematic reviews and meta-analysis
PROMETHEE	Preference Ranking Organization METHOD for Enrichment Evaluations
RDBMS	Relational database management system
RNA-seq	Ribonucleic acid sequencing
scRNA-seq	Single cell ribonucleic acid sequencing
SNP	Single nucleotide polymorphism
SQL	Structured Query Language
TF	Transcription factors
VIKOR	Vise Kriterijumska Optimizacija I Kompromisno Resenje

ABSTRACT

In this study, a multi-criteria ranking algorithm was introduced for evaluating gene regulatory network inference (GRNi) methods called RankGRN. This algorithm integrates a technique for determining the weights of evaluation criteria based on priorities defined by the user for ranking GRNi methods. Built upon the Weighted Sum Model, RankGRN, first generates the weight of each evaluation criteria considering the user-specified priority, after which it calculates the score of each GRNi method from all the selected weighted evaluation criteria, then it ranks the GRNi methods and identifies the best method according to the priority of each evaluation criteria and the overall best method. Subsequently, a web-based framework was developed for benchmarking and ranking GRNi methods called GRN Evaluator. The newly developed multi-criteria ranking algorithm was integrated in GRN Evaluator for ranking the GRNi methods after benchmarking. GRN Evaluator is a novel framework designed to address limitations in existing GRN benchmarking frameworks such as BEELINE, GReNaDIne, NetBenchmark, etc. Key advancements in GRN Evaluator include the incorporation of machine learning methods for GRN inference, addition of more datasets (both single-cell and bulk RNA-Seq data), inclusion of additional evaluation metrics, and integrating visualization tools for better interpretation of gene networks. Additionally, GRN Evaluator offers a user-friendly web interface, enhancing accessibility and usability. The systematic approach to evaluating GRNi methods across multiple datasets used in this study demonstrates their performance in various contexts. The framework effectively ranks these methods, providing valuable insights for researchers. The findings from this study serve as a guideline for selecting appropriate GRNi methods based on users' specific needs.

Keywords: gene regulatory network inference methods, machine learning algorithms, ranking algorithm, multi-criteria decision making, single cell RNA-Seq, bulk RNA-Seq, benchmarking frameworks.