

**GRADGROOM: AN INTEGRATED FRAMEWORK FOR  
AN EDUCATIONAL EARLY WARNING SYSTEM  
WITH MENTOR MATCHING**

DVDS Abeysinghe

219177F

Master of Science in Computer Science, Specialization in Cloud  
Computing

Department of Computer Science Engineering  
Faculty of Engineering

University of Moratuwa  
Sri Lanka

June 2024

**GRADGROOM: AN INTEGRATED FRAMEWORK FOR  
AN EDUCATIONAL EARLY WARNING SYSTEM  
WITH MENTOR MATCHING**

DVDS Abeysinghe

219177F

Thesis submitted in partial fulfillment of the requirements for the degree  
Master of Science in Computer Science, Specialization in Cloud  
Computing

Department of Computer Science Engineering  
Faculty of Engineering

University of Moratuwa  
Sri Lanka

June 2024

## Declaration

I declare that this is my own work and this thesis/dissertation does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature:

Date: 27/06/2024

DVDS Abeysinghe

The above candidate has carried out research for the Masters thesis under my supervision. I confirm that the declaration made above by the student is true and correct.

Name of Supervisor: Prof. Shantha Fernando

Signature of the Supervisor:

Date: 2024-06-28

## **Dedication**

My academic journey has been propelled by the unfailing support, encouragement, and sacrifices of my parents, husband, and all of my teachers and lecturers, to whom I dedicate my thesis. My biggest source of inspiration has been your confidence in me, and I will always be appreciative of your love, wisdom, and sacrifices, which have enabled me to do this.

Thank you.

Dharani Abeysinghe

## **Acknowledgement**

I would like to convey my sincere appreciation to my supervisor, Professor Shantha Fernando, Senior Lecturer, University of Moratuwa, for his unwavering support, wisdom, and insightful feedback given to me during this entire process of my MSc final research. His expertise, wisdom, encouragement, and commitment played a pivotal role in paving the direction and quality of this research.

I am also immensely thankful to the University of Moratuwa for being my alma mater, who shaped me academically and professionally during my period of studying for my Masters. Further, I would like to express my appreciation and thanks to my working institution for supporting me in doing this study.

Also I extend my heartfelt appreciation to all my work colleagues for the immense support extended in many forms.

Furthermore, I am indebted to my family, my parents, husband, siblings and in-laws for their unconditional love, encouragement, and understanding during the demanding journey of pursuing my MSc degree.

Last but not least, I extend my heartfelt thankfulness to all the participants for their generous contribution with their time and insights to this research project.

This thesis would not have been conducted to this success without the support and encouragement of all those mentioned above, and for that, I am profoundly grateful.

Dharani Abeysinghe

## Abstract

“Education is the key to success”, one of the most heard motivational statements by all of us. People engage in education at different phases of our lives in different ways. Among them, university education plays a vital role in our academic and professional lives. During university education many undergraduates will be facing several challenges demanding from educational matters to socio-economic problems. In many such situations undergraduates tend to abandon the degree programs halfway leaving them incomplete. Hence creating an Educational Early Warning Systems (EEWS) to predict and identify at-risk undergraduate students in the early stages of the degree programs will improve the graduating ratio against the dropouts. Further, mentoring is another aspect in education where it can be used greatly in the undergraduate studies to speak closer to the students addressing their curiosity and guiding them in the correct direction. There exists many separate frameworks for EEWS and mentoring. Despite an integrated framework that would be capable of identifying students at-risk and then match with the most matched mentor would be impactful and effective to handle by the universities. This study has proposed an integrated framework namely as “GRADGROOM” as a solution to the identified lacuna by extending EEWS framework with mentor matching which perform at-risk undergraduate student prediction and mentor-mentee matching for the identified at-risk undergraduate students. In the study, the proposed framework has been evaluated with two case studies. The framework has compared the use of KNN, SVM and RF algorithms while merging them to a stacking ensemble model with decision tree classifier as the base classifier in at-risk undergraduate student prediction. Further, from the predictions, the framework has extended the work up to logic based mentor-mentee matching and has conducted a successful mentoring program to validate the results. In the study the algorithms KNN, SVM, RF and the stacking ensemble model have given accuracies of 96%, 100%, 92% and 96% respectively proving the success of the prediction at-risk in the EEWS framework. Overall through the case studies and feedback gatherings, study has emphasized the importance of extending the at-risk prediction through EEWS followed with a mentoring process. Moreover, the study has concluded that a proper mentoring process conducted immediately after being identified as at-risk undergraduate students will be highly beneficial to reshape their study patterns to align with the correct route of studying.

**Keywords:** At-risk undergraduate student prediction, Educational Early Warning Systems (EEWS), Mentor-mentee matching, Virtual mentoring.

# Table of Contents

Declaration .....	i
Dedication .....	ii
Acknowledgement.....	iii
Abstract .....	iv
Table of Contents .....	v
List of Figures .....	viii
List of Tables.....	ix
List of Abbreviations.....	x
CHAPTER 1 .....	1
INTRODUCTION .....	1
1.1    Background .....	1
1.1.1    Mentoring and virtual mentoring .....	1
1.1.2    Requirement of mentoring in education.....	1
1.1.3    Role of mentoring in university education.....	2
1.1.4    Early warning systems in education.....	2
1.1.5    Requirement of Early warning systems in university education.....	3
1.1.6    Existing systems for virtual mentoring and early warning .....	3
1.2    Problem statement .....	4
1.3    Research Question .....	5
1.4    Research aim and Objectives.....	5
1.5    Research contribution and scope .....	5
1.6    Thesis Organization.....	6
CHAPTER 2 .....	7
LITERATURE REVIEW.....	7
2.1    Introduction .....	7
2.2    Related works .....	7
2.2.1    Impact of EEWS in higher education.....	7
2.2.2    Impact of Mentoring in higher education.....	10
2.2.3    Machine learning and Artificial Intelligence based EEWS .....	11

2.2.4	Use of existing theories with technology for EEWS .....	15
2.2.5	Knowledge based Mentor matching systems.....	17
2.2.6	Integration between EEWS and mentoring.....	18
2.3	Summary of literature review .....	19
2.4	Conclusive remarks .....	21
CHAPTER 3	.....	23
METHODOLOGY AND CONCEPTUAL FRAMEWORK	.....	23
3.1	Planning.....	23
3.2	Requirement gathering and anlysis .....	24
3.3	Framework development.....	24
3.4	Conceptual framework development.....	25
3.5	Testing and validation of the framework: A case study .....	32
3.6	Present findings: Results analysis .....	32
3.7	Conclusive remarks .....	33
CHAPTER 4	.....	34
TESTING AND VALIDATIONOF FRAMEWORK : CASE STUDIES.....		34
4.1	Case study 1- Predicting at-risk undergraduate students at first year .....	36
4.2	Case study 2- Predicting at-risk undergraduate students at second year.....	37
4.3	Conclusive remarks .....	37
CHAPTER 5	.....	38
RESULTS AND DISCUSSION	.....	38
5.1	Results of case study 1 .....	38
5.1.1	Algorithm-wise prediction results.....	38
5.1.2	Ensemble model (Stacking) prediction results.....	41
5.2	Results of case study 2 .....	43
5.2.1	Algorithm-wise prediction results.....	43
5.2.2	Ensemble model (Stacking) prediction results.....	46
5.3	Overall prediction results analysis.....	48
5.4	Success evaluation of mentoring.....	49
5.4.1	Mentor-mentee matching .....	49
5.4.2	At-risk undergraduate student mentoring and not mentoring impact ..	52
5.5	Conclusive remarks .....	54

CHAPTER 6 .....	56
CONCLUSION AND FUTURE WORKS .....	56
6.1    Conclusive remarks .....	57
6.1.1    Research problem.....	57
6.1.2    Research objectives.....	57
6.2    Ethical considerations.....	59
6.3    Future works.....	59
References .....	60
APPENDIX- A: Data gathering Questionnaires .....	65
APPENDIX- B: ML and Python codes.....	68

## List of Figures

<b>Figure</b>	<b>Description</b>	<b>Page</b>
<b>Figure 1:</b>	<b>The methodology of the study .....</b>	<b>23</b>
<b>Figure 2:</b>	<b>Process of framework implementation .....</b>	<b>25</b>
<b>Figure 3:</b>	<b>Basic data gathered about students .....</b>	<b>34</b>
<b>Figure 4:</b>	<b>Economical factors related to the student .....</b>	<b>35</b>
<b>Figure 5:</b>	<b>Details about study and Internet usage .....</b>	<b>35</b>
<b>Figure 6:</b>	<b>At-risk prediction results obtained from KNN algorithm.....</b>	<b>39</b>
<b>Figure 7:</b>	<b>Confusion matrix for KNN algorithm .....</b>	<b>39</b>
<b>Figure 8:</b>	<b>At-risk prediction results obtained from SVM algorithm.....</b>	<b>40</b>
<b>Figure 9:</b>	<b>Confusion matrix for SVM algorithm .....</b>	<b>40</b>
<b>Figure 10:</b>	<b>At-risk prediction results obtained from RF algorithm .....</b>	<b>40</b>
<b>Figure 11:</b>	<b>Confusion matrix for RF algorithm.....</b>	<b>41</b>
<b>Figure 12:</b>	<b>Structure of the ensemble model.....</b>	<b>41</b>
<b>Figure 13:</b>	<b>At-risk prediction results obtained from ensemble model .....</b>	<b>42</b>
<b>Figure 14:</b>	<b>Confusion matrix for ensemble model.....</b>	<b>42</b>
<b>Figure 15:</b>	<b>At-risk prediction of ensemble model- case validation .....</b>	<b>43</b>
<b>Figure 16:</b>	<b>Confusion matrix of ensemble model- case validation.....</b>	<b>43</b>
<b>Figure 17:</b>	<b>At-risk prediction results obtained from KNN algorithm.....</b>	<b>44</b>
<b>Figure 18:</b>	<b>Confusion matrix for KNN algorithm .....</b>	<b>44</b>
<b>Figure 19:</b>	<b>At-risk prediction results obtained from SVM algorithm.....</b>	<b>45</b>
<b>Figure 20:</b>	<b>Confusion matrix for SVM algorithm .....</b>	<b>45</b>
<b>Figure 21:</b>	<b>At-risk prediction results obtained from RF algorithm .....</b>	<b>45</b>
<b>Figure 22:</b>	<b>Confusion matrix for RF algorithm.....</b>	<b>46</b>
<b>Figure 23:</b>	<b>At-risk prediction results obtained from ensemble model .....</b>	<b>46</b>
<b>Figure 24:</b>	<b>Confusion matrix for ensemble model.....</b>	<b>47</b>
<b>Figure 25:</b>	<b>At-risk prediction of ensemble model- case validation .....</b>	<b>47</b>
<b>Figure 26:</b>	<b>Confusion matrix of ensemble model- case validation.....</b>	<b>48</b>
<b>Figure 27:</b>	<b>Output of mentor-mentee matching algorithm .....</b>	<b>51</b>

## List of Tables

<b>Table</b>	<b>Description</b>	<b>Page</b>
<b>TABLE 1:</b>	<b>Related works- EEWS and Mentoring in Higher Education.....</b>	<b>19</b>
<b>TABLE 2:</b>	<b>Related works- Integration of EEWS and Mentoring.....</b>	<b>20</b>
<b>TABLE 3:</b>	<b>Related works- ML based EEWS .....</b>	<b>20</b>
<b>TABLE 4:</b>	<b>Academic parameters for at-risk prediction in EEWS.....</b>	<b>27</b>
<b>TABLE 5:</b>	<b>Demographic parameters for at-risk prediction in EEWS .....</b>	<b>27</b>
<b>TABLE 6:</b>	<b>LMS parameters for at-risk prediction in EEWS.....</b>	<b>28</b>
<b>TABLE 7:</b>	<b>Behavioral characteristics for at-risk prediction in EEWS .....</b>	<b>28</b>
<b>TABLE 8:</b>	<b>Surface level parameters for mentor-mentee matching .....</b>	<b>30</b>
<b>TABLE 9:</b>	<b>Interest identification parameters for mentor-mentee matching .....</b>	<b>30</b>
<b>TABLE 10:</b>	<b>Deep level parameters for mentor-mentee matching.....</b>	<b>31</b>
<b>TABLE 11:</b>	<b>Summary of obtained results from models.....</b>	<b>48</b>
<b>TABLE 12:</b>	<b>Mentee data gathered from the questionnaires after labeling .....</b>	<b>50</b>
<b>TABLE 13:</b>	<b>Mentor data gathered from the questionnaires after labeling .....</b>	<b>51</b>
<b>TABLE 14:</b>	<b>Final matched results of mentors and mentees .....</b>	<b>52</b>

## **List of Abbreviations**

<b>Abbreviation</b>	<b>Description</b>
<b>ALM</b>	Additive Logistic Model
<b>CRRNN</b>	Convolution Residual Recurrent Neural Network
<b>CV</b>	Cross-Validation
<b>DA</b>	Deferred Acceptance
<b>DT</b>	Decision Tree
<b>EEWS</b>	Educational Early Warning Systems
<b>EWS</b>	Early Warning Systems
<b>FN</b>	False Negative
<b>FP</b>	False Positive
<b>GPA</b>	Graded Point Average
<b>LM</b>	Logistic Model
<b>LMS</b>	Learning Management System
<b>ML</b>	Machine Learning
<b>NLP</b>	Natural Language Processing
<b>RF</b>	Random Forest
<b>STEM</b>	Science, Technology, Engineering, and Mathematics
<b>SVM</b>	Support Vector Machines
<b>TE</b>	Technology and Engineering
<b>TN</b>	True Negative
<b>TP</b>	True Positive
<b>UTAUT</b>	The Unified Theory of Acceptance and Use Of Technology