

## Chapter 7

### Conclusion and Further Work

#### 7.1 Introduction

Chapter 6 described the implementation of proposed software to predicts students' final result with respect to basic attributes identified. This Chapter discusses a summary of the study along with its concept and findings. Then it presents the recommendations deriving from the study.

#### 7.2 Conclusion

This thesis discusses the application of data mining technique to students' data of Faculty of Information Technology to predict the final grade for a particular subject. This experiment provides a model to find out which students gain higher grades, average grades and marginal performances in advance. This prediction will help students to provide with early warning about their final examination result/grade for a selected subject. This will help to identify the students who need special attention to maintain their grades and reduce the fail rate. A timely appropriate warning to students at a risk could help prevent failing the exam or reducing a student getting marginal performance.

The developed methodology consist of following steps:

- Use decision tree techniques to build an understandable model including the inputs and class variables.
- Identification of classification rules among several factors and class variables.
- Introducing modification to classification rules to increase the accuracy of prediction and performance
- Develop a computer software with the identified classification rules to predict students final grade for a selected subject and its confidence.
- Using this new system it is possible to predict more than 75% accurate predictions.

### 7.3 Further work

This research can be extend to predict final result of subjects using some more attributes like his parents level of education, their family income, family background, the number of attempts of his Advanced Level etc.

This research can be extend to other post graduate programs to check their success or dropper base on their basic degree, age and some other personal attributes.

Further this kind of research can be extend to evaluate the performance of online or distance mode students performance in the final examinations and their successfulness of the degree program.

It is possible to use this type of prediction system for school examinations like GCE/OL and GCE/AL to predict students final results for particular subjects.

It is possible to conduct research using data mining techniques to predict the final result of the degree program using GPA statistics.



## References

- [1] R.Kohavi and T.Provost. "Confusion Matrix". Internet :[http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion\\_matrix/confusion\\_matrix.html](http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html), 1998 [Jun. 22,2008]
- [2] P.Smith. "Principals of data mining". Internet:[http://media.wiley.com/product\\_data/excerpt/24/04712285/0471228524.pdf](http://media.wiley.com/product_data/excerpt/24/04712285/0471228524.pdf),2001 [Jan. 25, 2008].
- [3] A. Costea, T. Eklund and J. Karlsson. "A Framework for Predictive Data Mining in the Telecommunications Sector", in Intl Conf. WWW/Internet,2002,pp. 38-46.
- [4] G. Holmes, R. Kirkby, and B. Pfahringer. "Mining data streams using option trees (revised edition, 2004)." Technical Report, The University of Waikato, Department of Computer Science, Hamilton, New Zealand, 2004.
- [5] Two Crows Corporation. "Introduction to Data Mining and Knowledge Discovery," in Data Mining Concept, 3rd ed, Vol. 1, Two Crows Corporation, Ed. USA: Potomac, 2005, pp. 1-2
- [6] <http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining>
- [7] <http://onlinelibrary.wiley.com/doi/10.1002/widm.1075/abstract>
- [8] K. Sreeram and A. Murthy, "Automatic construction of decision tree from data", Siemens corporate research, United Kingdom,2004
- [9] S. R. Safavian and D. Landgrebe,"A survey of decision tree classifier methodology", M.A. thesis, Purdue University, United States, 2005
- [10] G. P. Shapiro, "Discovery, analysis, and presentation of strong rules," in G. P. Shapiro & W. J. Frawley, eds, 'Knowledge Discovery in Databases', AAAI/MIT Press, Cambridge, MA, 1991
- [11] <https://thenewstack.io/six-of-the-best-open-source-data-mining-tools/>
- [12] <https://www.r-project.org/about.html>
- [13] B. Minaei-Bidgoli, D.A. Kashy, G. Kortmeyer and W.F.Punch "Predicting student performance: an application of data mining methods with an educational web based system" in Frontiers in Education,2003 FIE 2003 33rd Annual 2003, vol. 1,p.T2A-13
- [14] M.N. Quadri and Dr. N.V. Kalyankar-drop our feature of student data for academic performance using decision tree, Global Journal of Computer Science and Technology Vol.10 Issue 2(Ver 1.0), April 2010.

- [15] Zlatko J, Kovacic, John Steven Green, Predictive working tool for early identification of 'at risk students', new Zealand.
- [16] Anupama Kumar S. and Dr. Vijayalakshmi M.N. "Efficiency of Decision Tree in Predicting Student's Academic Performance" Bangalore India 2011
- [17] G. Naga Raja Prasad and Dr. A. Vinaya Bahu "Mining Previous Marks Data to Predict Students Performance in their Final Year Examination" International Journal of Engineering Research and Technology Vol.2, February 2013.

# Appendix A

## Set of Rules for Subject -3[c]

Accuracy=62 to 78%

1. AL=ARTS^CA=A  
PR=B^C(A)=0.00^C(B)=100.00^C(F)=0.00^C(C)=0.00^C(D)=0.00
2. AL=BIOLOGY^CA=A  
PR=A^C(A)=71.43^C(B)=21.43^C(F)=0.00^C(C)=7.14^C(D)=0.00
3. AL=BIOLOGY^CA=A^DISTRICT=ANURADHAPURA  
PR=B^C(A)=71.43^C(B)=21.43^C(F)=0.00^C(C)=7.14^C(D)=0.00
4. AL=BIOLOGY^CA=A^DISTRICT=KURUNEGALA^GENDER=F  
PR=B^C(A)=71.43^C(B)=21.43^C(F)=0.00^C(C)=7.14^C(D)=0.00
5. AL=ICT/MATHS^CA=A  
PR=B^C(A)=100.00^C(B)=00.00^C(F)=0.00^C(C)=0.00^C(D)=0.00
6. AL=/MATHS^CA=A  
PR=A^C(A)=93.33.00^C(B)=6.67^C(F)=0.00^C(C)=0.00^C(D)=0.00
7. AL=MATHS^CA=A^DISTRICT=BATTICOLOA  
PR=B^C(A)=93.33.00^C(B)=6.67^C(F)=0.00^C(C)=0.00^C(D)=0.00
8. AL=ARTS^CA=B  
PR=C^C(A)=14.29^C(B)=14.29^C(F)=0.00^C(C)=57.14^C(D)=14.29
9. AL=BIOLOGY^CA=B  
PR=B^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
10. AL=BIOLOGY^CA=B^DISTRICT=KANDY  
PR=A^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
11. AL=BIOLOGY^CA=B^DISTRICT=KURUNEGALA  
PR=A^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
12. AL=BIOLOGY^CA=B^DISTRICT=GALLE  
PR=A^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
13. AL=BIOLOGY^CA=B^DISTRICT=GAMPAHA  
PR=A^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
14. AL=BIOLOGY^CA=B^DISTRICT=TRINCOMALEE  
PR=C^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
15. AL=BIOLOGY^CA=B^DISTRICT=MATARA PR=A

$$^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00$$

16. AL=BIOLOGY^CA=B^DISTRICT=MATARA^GENDER=F  
PR=A^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
17. AL=BIOLOGY^CA=B^DISTRICT=PUTTALAM  
PR=A^C(A)=44.23^C(B)=50.00^C(F)=0.00^C(C)=5.77^C(D)=0.00
18. AL=COMMERCE^CA=B  
PR=B^C(A)=0.00^C(B)=40.00^C(F)=0.00^C(C)=40.00^C(D)=20.00
19. AL=COMMERCE^CA=B^DISTRICT=MATHALE  
PR=C^C(A)=0.00^C(B)=40.00^C(F)=0.00^C(C)=40.00^C(D)=20.00
20. AL=ICT/MATHS^CA=B  
PR=A^C(A)=100.00^C(B)=00.00^C(F)=0.00^C(C)=0.00^C(D)=0.00
21. AL=MATHS^CA=B  
PR=A^C(A)=58.62^C(B)=32.76^C(F)=0.00^C(C)=8.62^C(D)=0.00
22. AL=MATHS^CA=B^DISTRICT=KEGALLE  
PR=B^C(A)=58.62^C(B)=32.76^C(F)=0.00^C(C)=8.62^C(D)=0.00
23. AL=MATHS^CA=B^DISTRICT=NUWARAELIYA  
PR=B^C(A)=58.62^C(B)=32.76^C(F)=0.00^C(C)=8.62^C(D)=0.00
24. AL=MATHS^CA=B^DISTRICT=KALUTHARA  
PR=B^C(A)=58.62^C(B)=32.76^C(F)=0.00^C(C)=8.62^C(D)=0.00
25. AL=MATHS^CA=B^DISTRICT=BATTICOLOA  
PR=B^C(A)=58.62^C(B)=32.76^C(F)=0.00^C(C)=8.62^C(D)=0.00
26. AL=MATHS^CA=B^DISTRICT=MONERAGALA  
PR=B^C(A)=58.62^C(B)=32.76^C(F)=0.00^C(C)=8.62^C(D)=0.00
27. AL=MATHS^CA=B^DISTRICT=JAFFNA^GENDER=F  
PR=B^C(A)=58.62^C(B)=32.76^C(F)=0.00^C(C)=8.62^C(D)=0.00
28. AL=ARTS^CA=C  
PR=C^C(A)=0.00^C(B)=20.00^C(F)=0.00^C(C)=50.00^C(D)=30.00
29. AL=ARTS^CA=C^DISTRICT=KURUNEGALA  
PR=B^C(A)=0.00^C(B)=20.00^C(F)=0.00^C(C)=50.00^C(D)=30.00
30. AL=ARTS^CA=C^DISTRICT=COLOMBO PR=D  
^C(A)=0.00^C(B)=20.00^C(F)=0.00^C(C)=50.00^C(D)=30.00
31. AL=ATRS^CA=C^DISTRICT=TRINCOMALEE  
PR=B^C(A)=0.00^C(B)=20.00^C(F)=0.00^C(C)=50.00^C(D)=30.00

32.  $AL=ARTS \wedge CA=C \wedge DISTRICT=HAMBANTHOTA$   
 $PR=D \wedge C(A)=0.00 \wedge C(B)=20.00 \wedge C(F)=0.00 \wedge C(C)=50.00 \wedge C(D)=30.00$
33.  $AL=BIOLOGY \wedge CA=C$   
 $PR=B \wedge C(A)=6.67 \wedge C(B)=70.00 \wedge C(F)=0.00 \wedge C(C)=20.00 \wedge C(D)=3.33$
34.  $AL=BIOLOGY \wedge CA=C \wedge DISTRICT=BADULLA$   
 $PR=C \wedge C(A)=6.67 \wedge C(B)=70.00 \wedge C(F)=0.00 \wedge C(C)=20.00 \wedge C(D)=3.33$
35.  $AL=COMMERCE \wedge CA=C$   
 $PR=C \wedge C(A)=0.00 \wedge C(B)=00.00 \wedge C(F)=0.00 \wedge C(C)=0.00 \wedge C(D)=100.00$
36.  $AL=MATHS \wedge CA=C$   
 $PR=B \wedge C(A)=16.13 \wedge C(B)=67.74 \wedge C(F)=3.23 \wedge C(C)=12.90 \wedge C(D)=0.00$
37.  $AL=MATHS \wedge CA=C \wedge DISTRICT=VAVVUNIYA$   
 $PR=C \wedge C(A)=16.13 \wedge C(B)=67.74 \wedge C(F)=3.23 \wedge C(C)=12.90 \wedge C(D)=0.00$
38.  $AL=MATHS \wedge CA=C \wedge DISTRICT=RATHNAPURA \wedge GENDER=M$   
 $PR=A \wedge C(A)=16.13 \wedge C(B)=67.74 \wedge C(F)=3.23 \wedge C(C)=12.90 \wedge C(D)=0.00$

### Set of Rules for subject 2

Accuracy=75.36 to 80.32%

39.  $CA=A \wedge AL=BiologY \wedge District=Badulla \wedge Gender=F \rightarrow Results=B$   
 $\wedge C(A)=0.72 \wedge C(B)=0.27 \wedge C(C)=0.00 \wedge C(D)=0.00 \wedge CON(F)=0.01$
40.  $CA=A \wedge AL=BiologY \wedge District=Anuradhapura \wedge Gender=F \rightarrow$   
 $Results=B \wedge C(A)=0.72 \wedge C(B)=0.27 \wedge C(C)=0.00 \wedge C(D)=0.00 \wedge CON(F)=0.01$
41.  $CA=A \wedge AL=BiologY \wedge District=Moneragala \rightarrow Results=B$   
 $\wedge C(A)=0.72 \wedge C(B)=0.27 \wedge C(C)=0.00 \wedge C(D)=0.00 \wedge CON(F)=0.01$
42.  $CA=A \wedge AL=BiologY \rightarrow Results=A$   
 $\wedge C(A)=0.72 \wedge C(B)=0.27 \wedge C(C)=0.00 \wedge C(D)=0.00 \wedge CON(F)=0.01$
43.  $CA=A \wedge AL=ICT/Maths \wedge Results=A$   
 $\wedge C(A)=1.00 \wedge C(B)=0.00 \wedge C(C)=0.00 \wedge C(D)=0.00 \wedge CON(F)=0.00$
44.  $CA=A \wedge AL=Maths \wedge District=Ampara \rightarrow Results=B$   
 $\wedge C(A)=0.83 \wedge C(B)=0.12 \wedge C(C)=0.03 \wedge C(D)=0.01 \wedge CON(F)=0.01$
45.  $CA=A \wedge AL=Maths \wedge District=Anuradhapura \rightarrow$

Results= $B^C(A)=0.83^C(B)=0.12^C(C)=0.03^C(D)=0.01^CON(F)=0.01$

46. CA=A ^ AL=Maths ^ District=Mannar →

Results= $B^C(A)=0.83^C(B)=0.12^C(C)=0.03^C(D)=0.01^CON(F)=0.01$

47. CA=A ^ AL=Maths →

Results= $A^C(A)=0.83^C(B)=0.12^C(C)=0.03^C(D)=0.01^CON(F)=0.01$

48. CA=B ^ AL=Biology ^ District=Jaffna → Results=B^

$C(A)=0.7^C(B)=0.2^C(C)=0.07^C(D)=0.00^CON(F)=0.03$

49. CA=B ^ AL=Biology → Results=

$A^C(A)=0.7^C(B)=0.2^C(C)=0.07^C(D)=0.00^CON(F)=0.03$

50. CA=B ^ AL = Maths^ District= Jaffna → Results=B^

$C(A)=0.68^C(B)=0.29^C(C)=0.03^C(D)=0.00^CON(F)=0.00$

51. CA=B ^ AL=Maths ^ District=Vavuniya → Results=B^

$C(A)=0.68^C(B)=0.29^C(C)=0.03^C(D)=0.00^CON(F)=0.00$

52. CA=B ^ AL=Maths → Results= A

$^C(A)=0.68^C(B)=0.29^C(C)=0.03^C(D)=0.00^CON(F)=0.00$

53. CA=C → Results= B ^C(A)=0.00^C(B)=1.00^C(C)=0.00^C(D)=0.00^CON(F)=0.00

### Set of Rules for subject 1[w]

Accuracy=62 to 78 %

54. CA=A ^ AL=Biology ^ District=Puttalam → Results=B^

$^C(A)=0.7^C(B)=0.3^C(C)=0.00^C(D)=0.00^CON(F)=0.00$

55. CA=A ^ AL= Biology ^ District=Rathnapura →

Results=B^C(A)=0.7^C(B)=0.3^C(C)=0.00^C(D)=0.00^CON(F)=0.00

56. CA=A ^ AL=Biology → Results=A ^

$C(A)=0.7^C(B)=0.3^C(C)=0.00^C(D)=0.00^CON(F)=0.00$

57. CA=A ^ AL=Maths → Results=A

$^C(A)=0.68^C(B)=0.28^C(C)=0.04^C(D)=0.00^CON(F)=0.00$



58. IF CA=B ^ AL=Biology ^ Disstrict=Ampara → Results=C  
 $C(A)=0.11^C(B)=0.74^C(C)=0.11^C(D)=0.05^CON(F)=0.00$
59. IF CA=B ^ AL=Biology ^ Disstrict=Polonnaruwa → Results=C  
 $C(A)=0.11^C(B)=0.74^C(C)=0.11^C(D)=0.05^CON(F)=0.00$
60. IF CA=B ^ AL=Maths ^ Disstrict=Colombo → Results=A  
 $C(A)=0.24^C(B)=0.66^C(C)=0.1^C(D)=0.00^CON(F)=0.00$
61. IF CA=B ^ AL=Maths ^ Disstrict=Nuwa Eliya → Results=C  
 $C(A)=0.24^C(B)=0.66^C(C)=0.1^C(D)=0.00^CON(F)=0.00$
62. CA=B ^ AL=Biology → Results=B ^  
 $C(A)=0.11^C(B)=0.74^C(C)=0.11^C(D)=0.05^CON(F)=0.00$
63. CA=B ^ AL=Maths → Results=B  
 $C(A)=0.24^C(B)=0.66^C(C)=0.1^C(D)=0.00^CON(F)=0.00$
64. CA=C ^ AL=Biology ^ District=Colombo → Results=B  
 $C(A)=0.00^C(B)=0.38^C(C)=0.62^C(D)=0.00^CON(F)=0.00$
65. CA=C ^ AL=Biology ^ District=Kurunegala → Results=B ^  
 $C(A)=0.00^C(B)=0.38^C(C)=0.62^C(D)=0.00^CON(F)=0.00$
66. CA=C ^ AL=Biology → Results=C ^  
 $C(A)=0.00^C(B)=0.38^C(C)=0.62^C(D)=0.00^CON(F)=0.00$
67. CA=C ^ AL=Maths ^ District=Mannar → Results=C  
 $C(A)=0.00^C(B)=0.47^C(C)=0.29^C(D)=0.24^CON(F)=0.00$
68. CA=C ^ AL=Maths ^ District=Mulathivu → Results=C ^  
 $C(A)=0.00^C(B)=0.47^C(C)=0.29^C(D)=0.24^CON(F)=0.00$
69. CA=C ^ AL=Maths ^ District=Polonnaruwa → Results=C  
 $C(A)=0.00^C(B)=0.47^C(C)=0.29^C(D)=0.24^CON(F)=0.00$
70. CA=C ^ AL=Maths Results=B  
 $C(A)=0.00^C(B)=0.47^C(C)=0.29^C(D)=0.24^CON(F)=0.00$

LIBRARY / UOM		
20	18	H
20		
20		
20		

