

The Success of the students: A study based on test results

Andrés Felipe Téllez Rodríguez
KM 8 Via San Antonio – La Ceja
Parcelacion La Selva – Casa 10
+57 3006234499
aftellezr@eafit.edu.co

Jonathan Andrés Bedoya
Zapata
Cra 110 #38 - 25 Interior 501
(San Javier - 20 de Julio)
+57 312 5820198
jabedoyaz@eafit.edu.co

ABSTRACT

The predictive value that data has, will allows us to predict the possibility to succeed of a student, based on their chosen career, and the success he had on the ICFES tests. For that with usage of decision trees, and different strategies to keep the information short and understandable

Keywords

Students, prediction, ICFES, trees.

1. INTRODUCTION

Trying to gauge the possibility of success of recently graduated students is a task that could bear great fruits, as it gives an edge on the knowledge of the possible careers that the student should take. Of course, measurement of such things should be supported by other studies. And thus, this group has searched for different problems that bear 4 similar processes, and their respective solutions, to give ourselves a proper baseline to start developing the program that would allow us to accomplish our objectives.

1.1 Problem 1

[1] The first problem to solve was the prediction of the behavior of the vote based on personality and sociopolitical elements. By itself is a hard problem due to the different interpretations that the data of the results could give.

1.1.1 Solution

The solution given was divide the results into three different variables. The first a sociological one, the second a psychological one, and the third a political. Read each value and see who they voted for. If anyone with the same parameters appears then the prediction is done.

1.1.2 Graph of the Solution

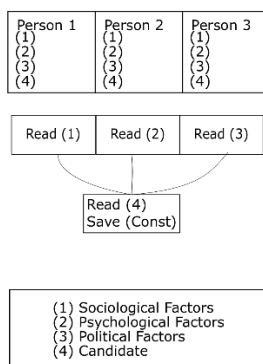


Image 1. The graph

1.2 Problem 2

[2] The second problem is making a prediction model of the satisfaction with the school and physical education. That is because of the relation that the school satisfaction is related to either things like depression, or scholar desertion. On the other hand, it is also related to the social interactions, the stress levels, and the academic performance.

1.2.1 Solution

The solution was given by the usage of several preexisting measurement units to classify, and organize the answers given, and then by using some formulas that are related to the aforementioned measurement units.

1.2.2 Graph

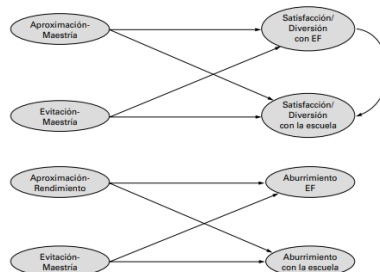


Image 2. The Measurements used, and their structure

1.3 Problem 3

[3] The third problem is establishing a deeper relation between economics, and the crime. Because of their long-time relationship, it allows us to have a deeper understanding of the crime workings. Also, it allows to see a better picture for possibilities. Like whom

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

could fall easier to crime or end up dealing with illegal procedures.

1.3.1 Solution

The solution is using several criteria to reunite what makes crime and economy so close, and then seeing what relation it has with the real life.

1.3.2 Graphic

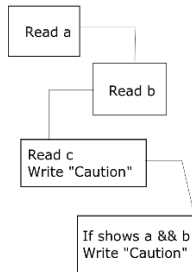


Image 3. Graphic of the solution

2. Design of the data structure

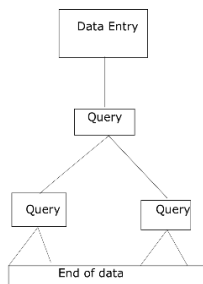


Image 4. Graphic of the solution

The design is simple, the user gives a query, and the algorithm separates the data in two branches, and keeps repeating until all the

data is consumed. The data now separated is classified and allows us to predict results with it.

3. Calculus of the time complexities

Time complexities allows us to predict how the time an operation will take in relation to the data provided. Ideally the complexity of an algorithm in the worst case is constant, meaning that no matter the amount of data you introduce the time that is used to resolve the algorithm is still the same. That case is represented as $O(1)$.

In the algorithm used to solve the problem that we considered; the Time Complexity is $O(\log(M/N))$. Where M is the rows and N is the columns.

4. Criteria for the design of the data structure

Looking for a starting point, we found an algorithm that gave us a solid base to begin coding. From there we continued using what we thought was the best method.

5. Results obtained in the final solution

Using data that contained 500 rows and 158 columns and running it 100 times and found out that the average time required to run the program was 1.6601 seconds.

6. Conclusion

To sum the project up, what we wanted to do in this project is to predict the success of a student. We have done this by sorting data and using queries to fish out the desired results. This solution could be used in the three problems that were considered. With the results of this project, the interested parties can filter the students that are candidates for a scholarship, and with that information select the students better suited for a full or partial scholarship.

7. Acknowledgement

We thank for assistance with [] to [] for comments that greatly improved the manuscript

8. References

- [1] González M. and Darías M. 1998 *PREDICCIÓN DE LA CONDUCTA DE VOTO Personalidad y factores sociopolíticos*
- [2] Baena-Extremera Antonio and Granero-Gallegos Antonio. 2015. *Modelo de predicción de la satisfacción con la educación física y la escuela*
- [3] Ramires de Garay Luis David. 2014. *Crimen y economía: una revisión crítica de las explicaciones económicas del crimen*