DECISION TREES IN ACADEMIC SUCCES IN SABER PRO

UNIVERSIDAD EAFIT

Team Presentation

Santiago Ochoa

Miguel Angel Zapata

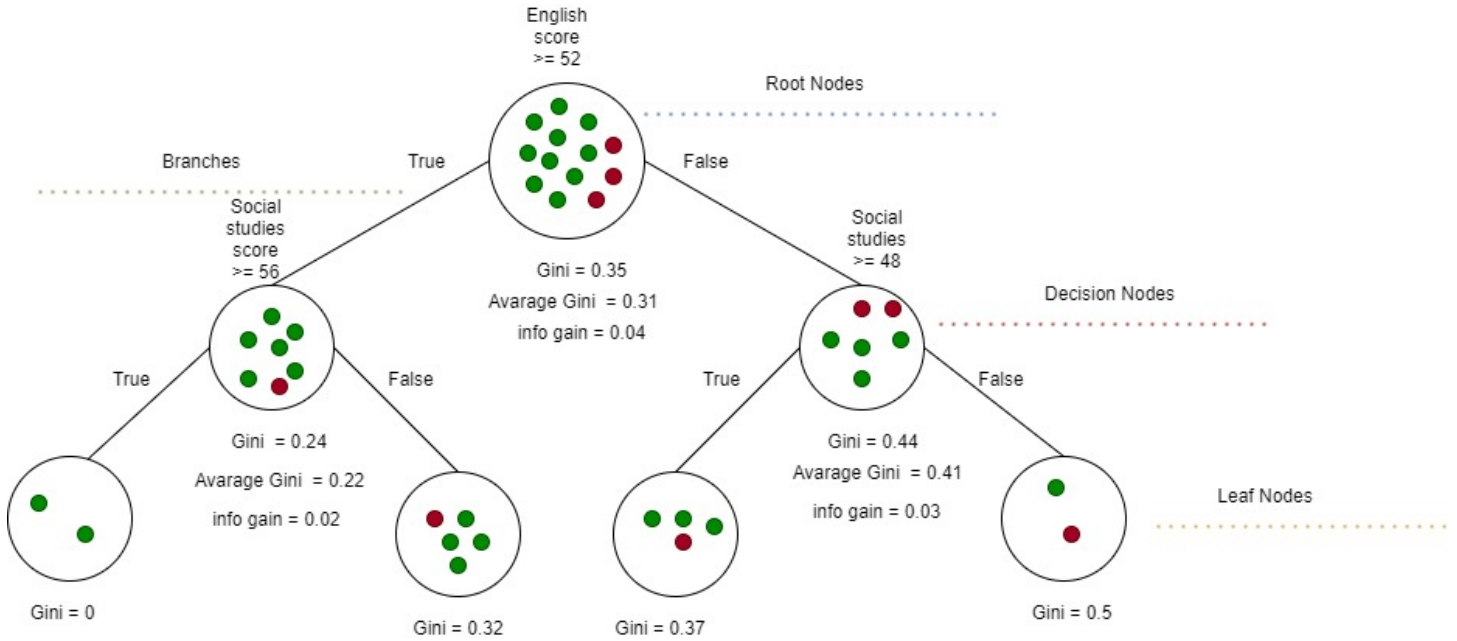
Miguel Correa

Mauricio Toro

https://github.com/sochoac1/ST0245002/tree/master/proyecto/proyecto/

UNIVERSIDAD EAFIT

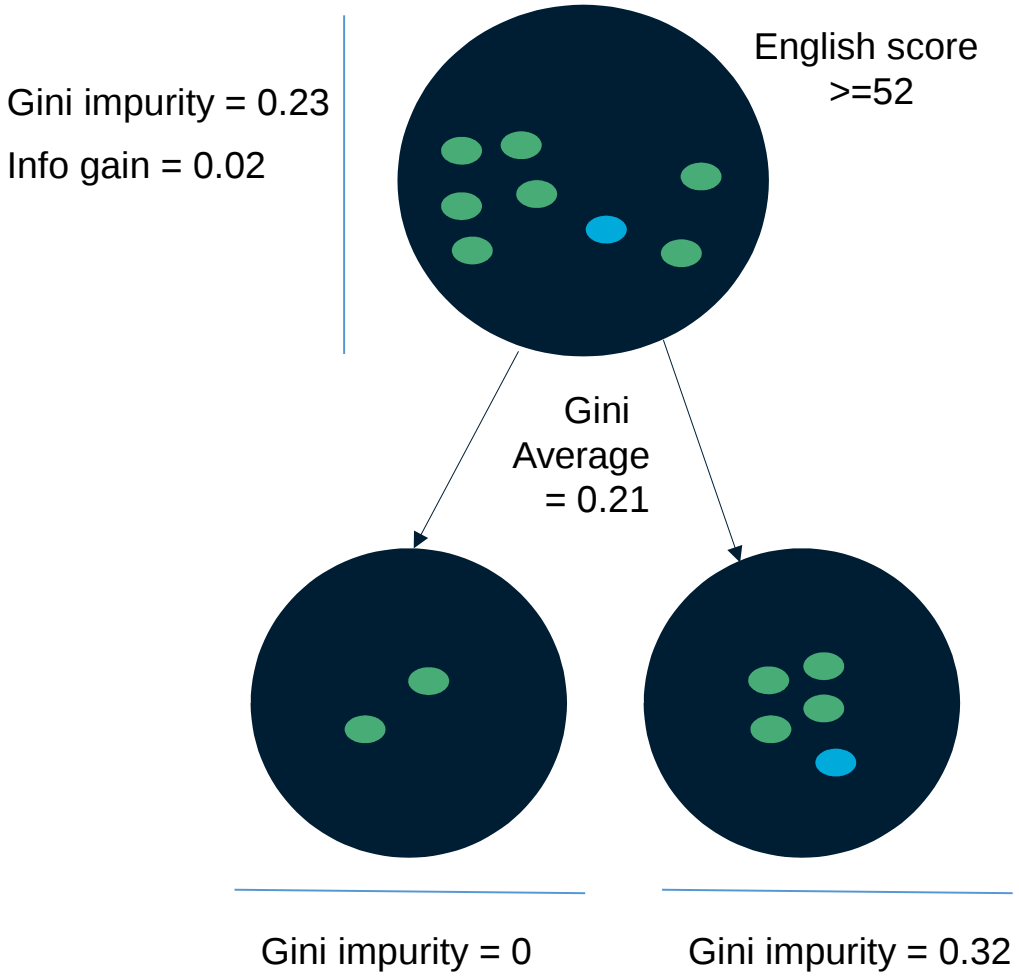
Algorithm Design

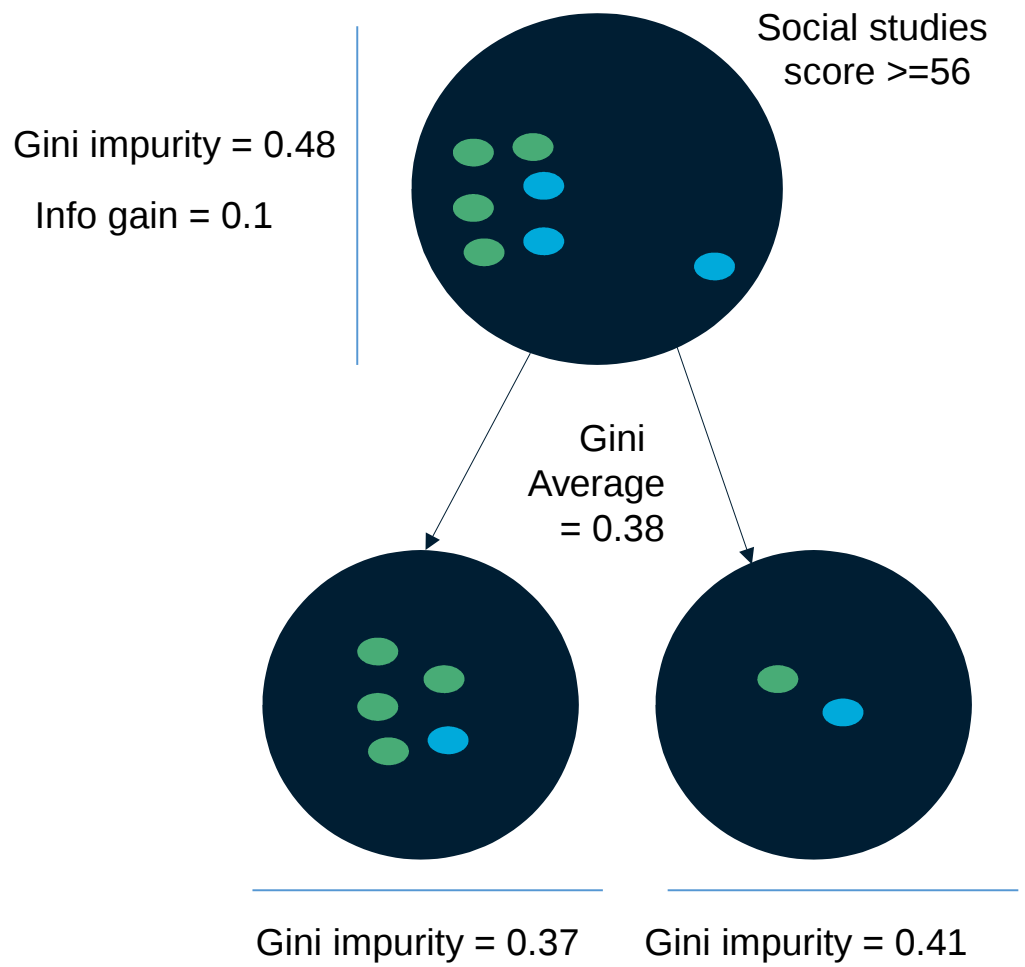


To build a binary tree the algorithm implemented was CART. In this example, we show a model to predict the academic success as the probability that a student must get a score above the average in the Saber pro.

UNIVERSIDAD EAFIT®

Node Splitting



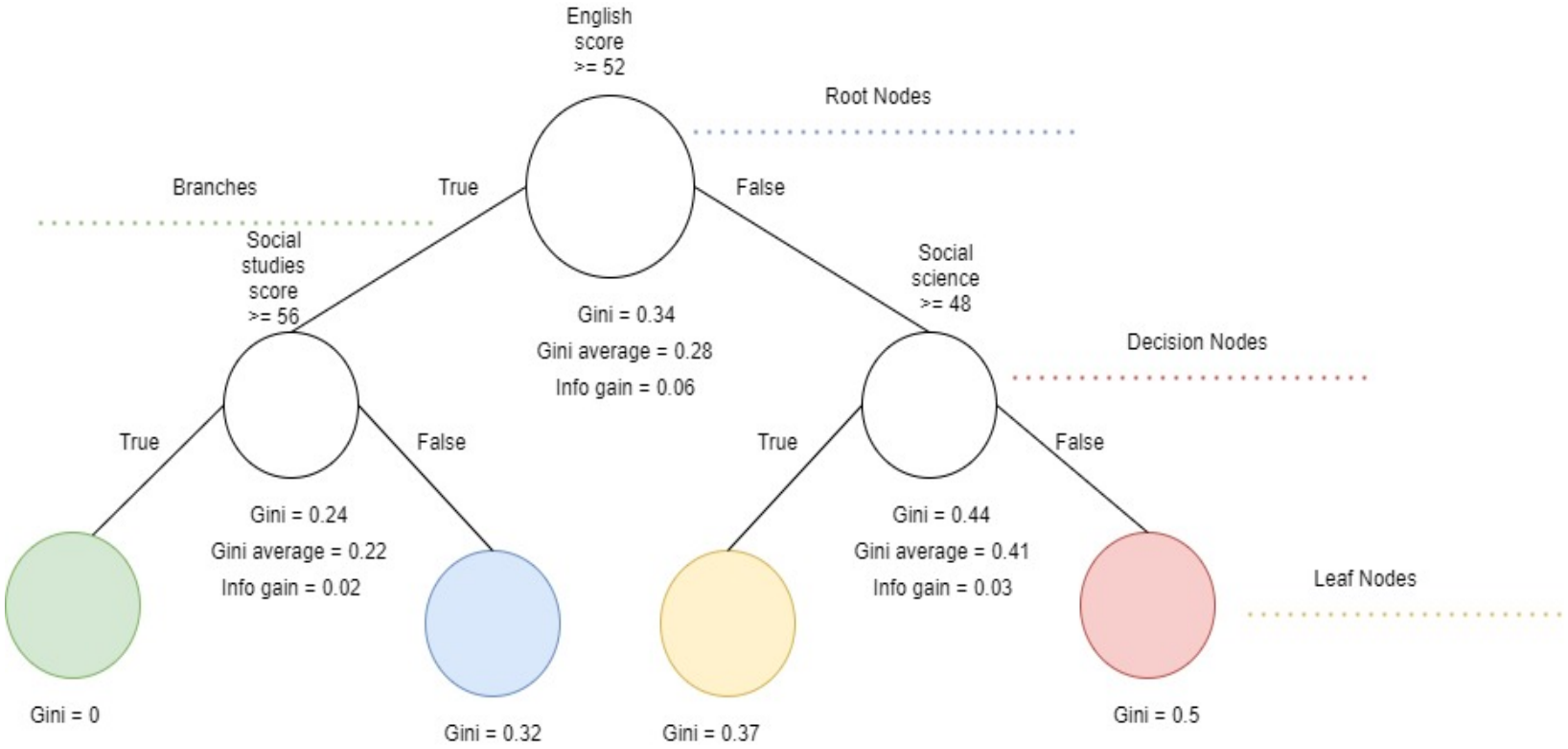


Algorithm Complexity

	Time Complexity	Memory Complexity
Training the model	$O(N^2*M*2^M)$	$O(N*M*2^M)$
Testing the Model	O(N*M)	O(1)

The variable N represents the number of rows and M represents the number of columns of a matrix which contains the training dataset to build the tree.

Decision-Tree Model

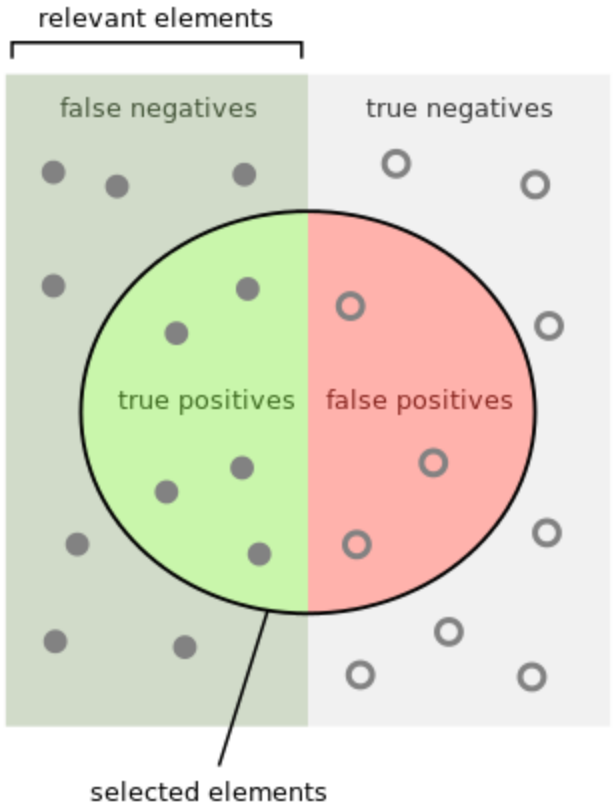


A binary decision tree to predict Saber Pro scores based on the results of Saber 11. Green nodes represent those with a high probability of success, blue ones a medium probability, violet a probability between medium and low and the red ones a low probability of success.

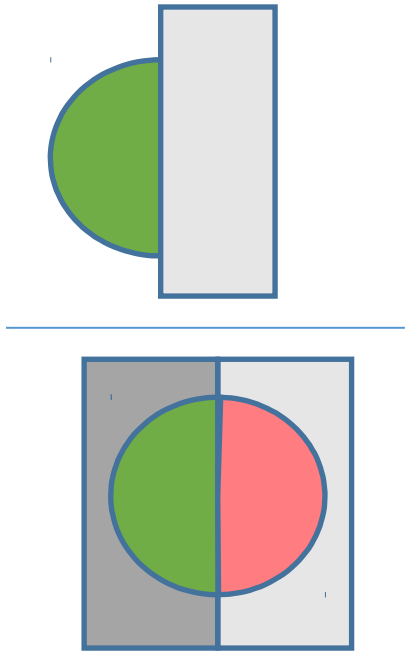
Most Relevant Features

- Social Studies
- English
- Science

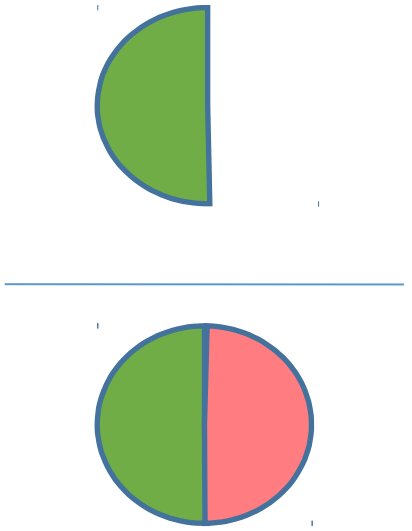
Evaluation Metrics



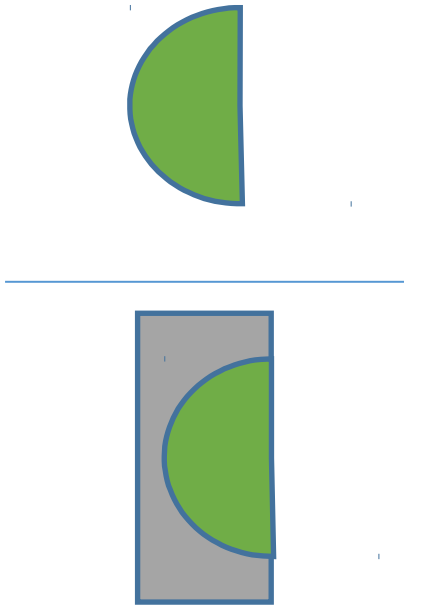
Exactitud



Precisión



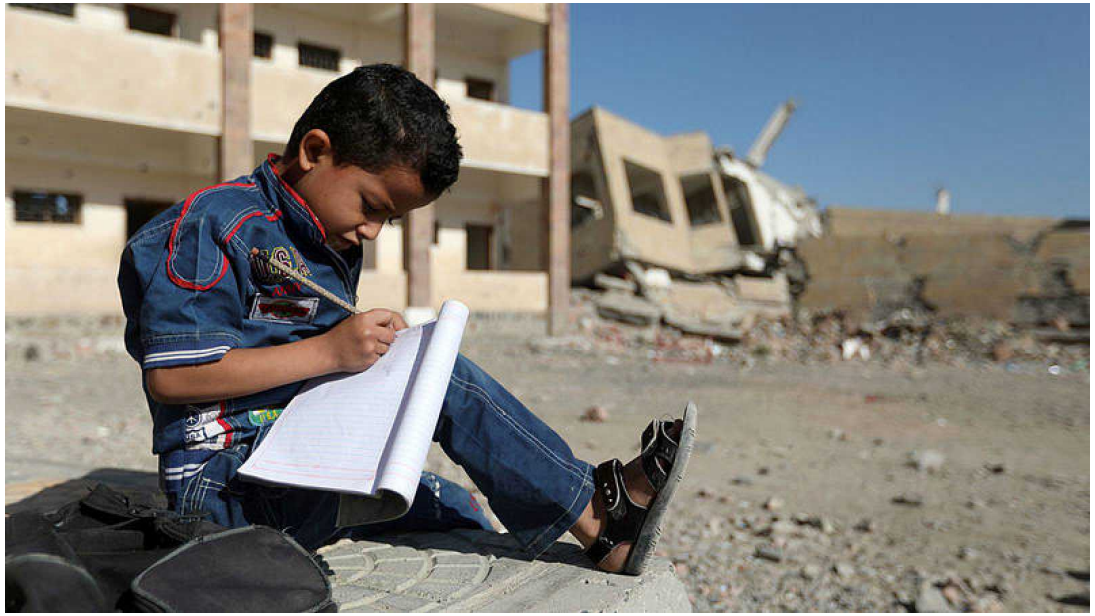
Sensibilidad



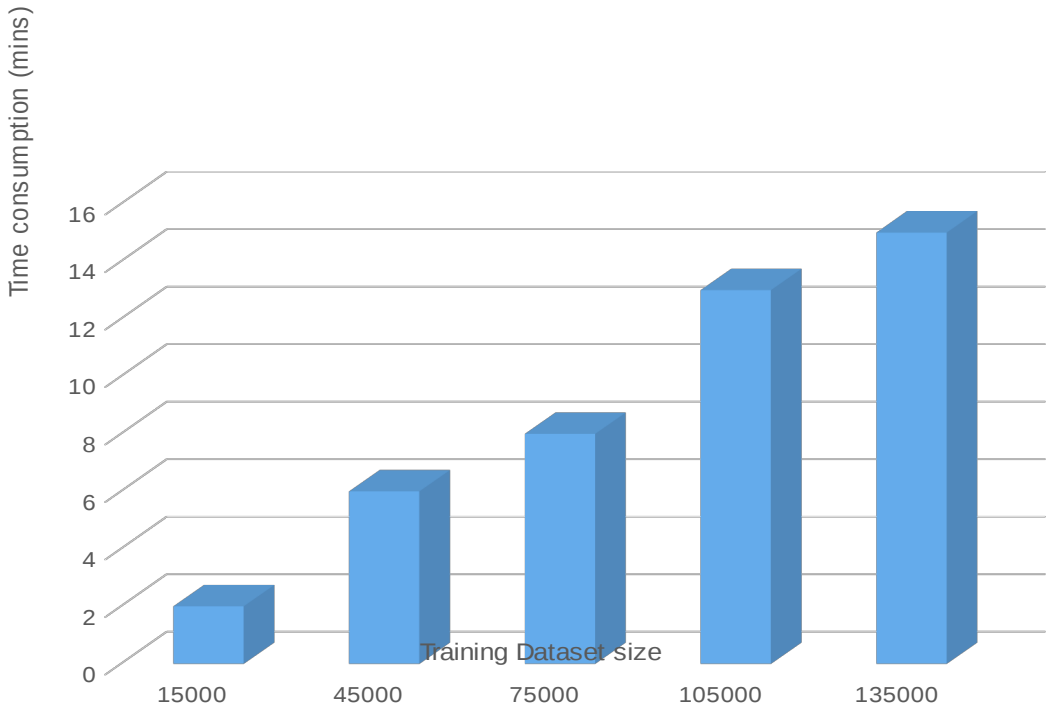
Evaluation Metrics

	Training data set	Testing data set
Accuracy	0.78	0.77
Precision	0.75	0.76
Recall	0.8	0.79

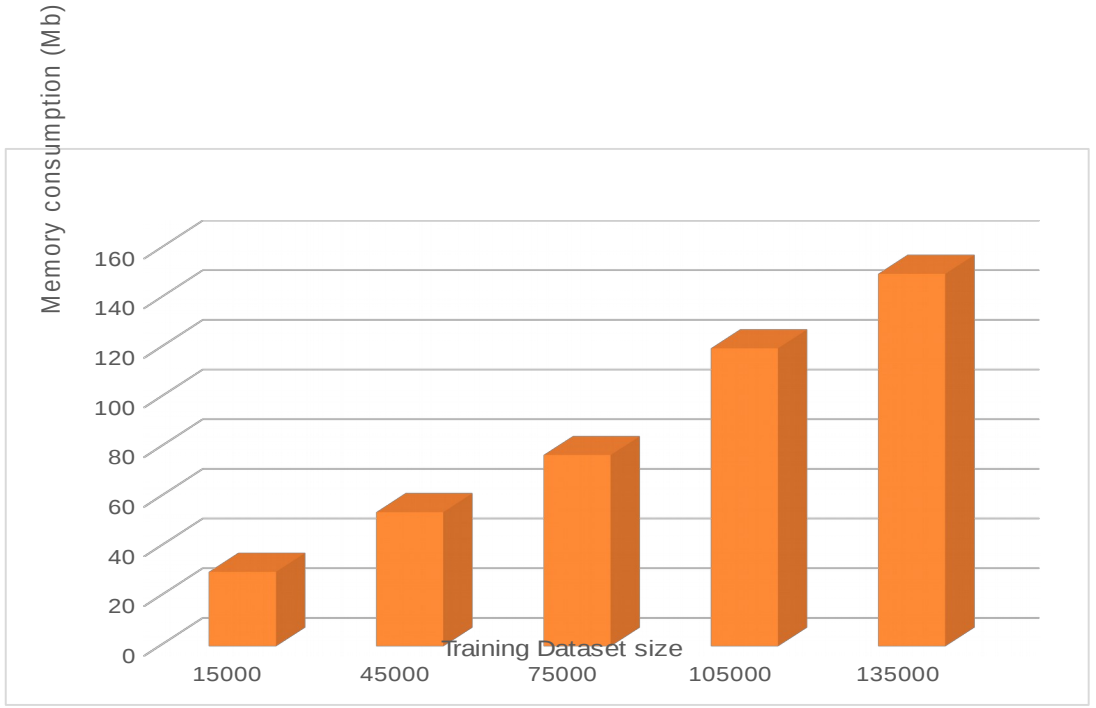
Evaluation metrics using a training dataset of 135,000 students and test dataset of 45,000 students.



Time and Memory Consumption



Time Consumption



Memory Consumption


THANK YOU!