# Youth Unemployment Bayesian Network Report

Matteo Antonio Inajetovic 0000949648

## A. Introduction

I 've decided to choose as topic for my project one of the biggest problems of our country in this period, the youth unemployment. The small Bayesian Network I 've realized (Fig.1) is a strongly simplified model for the topic, since the latter has a lot of causes and problems belonging to many complex human knowledge areas (sociology, economics, politics etc…).

Thus, I decided to pick up some relevant features, namely the gender (Female or Male), the economic condition (wealthy or poor), the provenience (South or North Italy) and the level of instruction (Graduated or not) to predict whether a young person can be employed or not. Since I could not find any specific dataset for my topic, I built up my variables' CPDs looking at the most reliable Italian sources of data in this environment (Istat, Censis, IlSole24Ore, IlMessaggero, Alamalaurea etc..).
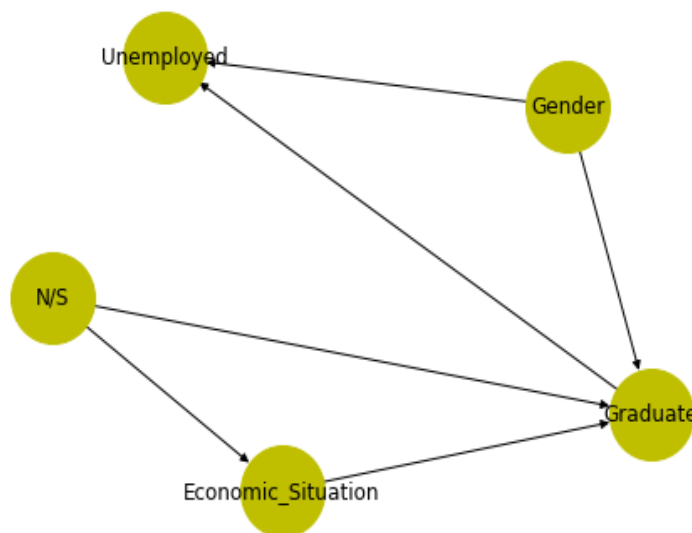
## B. Network representation



Fig.1:

Links' description:

N/S (North/South) → Economic_Situation: The poverty level is higher in the South Italy, so given the provenience of a person as evidence, is more probable to observe a better economic condition if N/S is False, namely that the person comes from North Italy.

Economic_Situation → Graduate: Nowadays universities are helping poor people to avoid huge taxes, but still, the economic situation influences the probability for a young person to graduate.

N/S → Graduate: People from the North usually graduate more than ones from the South.

Gender → Graduate: From recent studies emerged that girls graduate more than boys.

Gender → Unemployed: The female unemployment is unfortunately higher than the male one.

Graduate → Unemployed: A lot of graduates have a lot of problems in finding a job and consequentially to be employed.

## C. Experiments

Local independences, v-structures, active trails and Markov Blankets have been studied, the results are reported in the jupyter notebook linked to this report. Then, exploiting the Variable Elimination exact inference method, I made a lot of queries and observed many of the most relevant youth unemployment problems: the fact that even if women usually graduate more than men, are more likely to be unemployed, the irrelevance of graduation in finding a job and also other socio-economic differences affecting the Italian employment world.

Finally, given a query and the related reference value computed through the Variable Elimination algorithm, two approximate inference methods have been compared, the Likelihood Weighting and Rejection Sampling (Fig. 2, Fig. 3).
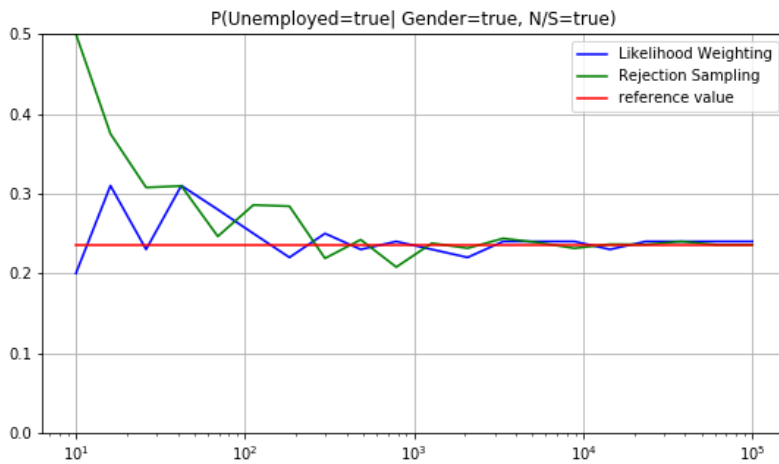


Fig.2. Query: P(Unemployed=True |Gender=True, N/S=True).



Fig.3. Absolute error on the query P(Unemployed=True |Gender=True, N/S=True).