01URZSM DATA ETHICS AND DATA PROTECTION

COMPUTER SCIENCE PART (DATA ETHICS)

Exam 30 June 2023

A. CASE (7,5points). PowerSchool.

PowerSchool is a US-based company that provides software for lower secondary school (equivalent to "scuola media" in Italy) and it holds data on more than 45 million children. Some of this data is used to create risk-assessment scores aimed at predicting students' failures. It has been found that the most relevant predictors for high-risk scores were the following variables:

- Attendance in current semester: number of days
- Attendance in last year: number of days
- Behavior in current semester: number of referrals
- Behavior in last year: number of referrals
- Past graduation grade: score from 1 to 5
- Limited English Proficiency: yes/no, indicating whether English is not the primary language of the child and he/she have difficulty communicating effectively in English
- Free lunch: yes/no, whether the child has right to free lunches at school due to low income of his/her family
- Failure (target variable): yes/no

An audit for checking the tool against discrimination towards some protected attributes was performed with predictions on 2000 children. Overall, 159 failures were predicted, divided within the following subgroups:

		Total number	**Nr. Predicted at high risk**
Gender	Male	900	71
	Female	1100	88
Race	White	1200	72
	Black	240	48
	Hispanic/latino	390	28
	Asian	140	10
	American Indian	20	1
	Pacific Islander	10	0

1) **Analyze and comment the results of the audit with a fairness criterion at your choice: towards which category a relevant disparate impact of the prediction tool occurred? Why? (5 p.).**
2) **How should the table look like if results were fairer? Identify which values of the table you would change and explain the modification(s). Be sure all the numbers will be consistent with the change introduced. (2,5 p.)**

N.B. Answer concisely to the questions: give a separate answer for each point.
In doing that, clearly state your own hypotheses, and any other information that you suppose in addition to the provided data, to coherently support your line of reasoning.

***

B (7,5 points). Define the three approaches of normative ethics (3p.). Then identify which of the following three elements of the list belong to which approach, explaining why (4,5p.):
1. Fairness criteria computation; 2. Fairness qualitative assessment; 3. ACM code of ethics.

Indications for a possible solution

A. CASE (7,5points). PowerSchool.

The following comments focus on the interpretation of the case. It should be used as a guide and not as the unique solution, which might be partially dependent on the reasoning presented and the hypotheses made (if valid). The exam, given the available time, requires synthetic (but precise and logically coherent) answers, therefore the length of the analysis reported herein should not be taken as a reference.

The exam test is based on a real tool (link 1) with the problem highlighted in link 2:

1) https://www.documentcloud.org/documents/21175358-overview-presentation-school-district-u-46
2) https://www.documentcloud.org/documents/21175361-variable-weights-school-district-u-46

Answer to question 1)

The only applicable fairness criterion is Independence, because the table contains only data about high-risk predictions (R=1): to compute it, it is sufficient to compute the ratio Nr. Predicted at high risk/ Total number.

		Total number	**Nr. Predicted at high risk**	**P(R=1)**
Gender	Male	900	71	8%
	Female	1100	88	8%
Race	White	1200	72	6%
	Black	240	48	20%
	Hispanic/latino	390	28	7%
	Asian	140	10	7%
	American Indian	20	1	5%
	Pacific Islander	10	0	0%

As it can be noticed from the results, Black children are predicted as high risk with more than double percentage than children of other races. Based on the inequality patterns explained in the lectures (slideset demographic disparities), the most relevant proxies is "free lunch": it is a welfare benefit for low-income families, which are composed with higher percentage of black people. Poor school performances in US can also be easily tracked back to difficult socio-economic conditions, therefore also to black category. Results on Pacific islander children can be ignored due to very small sample size.

Answer to question 2)

A fairer prediction would result in a high-risk prediction rate for black like the other races: for example, by changing the number of predicted high risks for black to 20, the ratio will be 8%, which is very similar to the other races. To keep the total numbers consistent, male and females' high risks should be decreased accordingly: for example, respectively to 50 and 81. Table is reported below, with changed values in red.

		Total number	**Nr. Predicted at high risk**
Gender	Male	900	50
	Female	1100	81
Race	White	1200	72
	Black	240	20
	Hispanic/latino	390	28
	Asian	140	10
	American Indian	20	1
	Pacific Islander	10	0

B. Question on theory (7,5points)

Check slides and video-lectures to check definitions and understand the correct mapping, which was: **Fairness criteria computation → consequentialism, Fairness qualitative assessment → virtue ethics, ACM code of ethics → deontology**.