Indian Institute of Technology Kharagpur
Intelligent Systems (CS60036)
End-Semester Examination, Spring 2017-18

Full Marks: 100   **Time: 3 hours**

All parts of the same question must be answered together
Be precise in your answers, and state any assumptions made
If there are multiple ways to perform a computation, state which one you are using

Question 1 [2+3+6+4 = 15 marks]
   (a) Even without the training data being biased, how can Machine Learning algorithm be biased?
   (b) The following contingency table shows the predictions made by a recidivism risk (whether a defendant is going to recommit crime within two years) prediction tool and their corresponding ground truths. Is the tool biased?

Black Defendants	Low	High		**White Defendants**	Low	High
Survived	990	805		Survived	1139	349
Recidivated	532	1369		Recidivated	461	505

   (c) Define Disparate Treatment, Disparate Impact and Disparate Mistreatment. How are they different from each other?
   (d) How can we utilize constrained optimization frameworks to incorporate fairness in algorithmic decision?

Question 2 [3+4+8 = 15 marks]
   (a) Define explainable discrimination and bad discrimination.
   (b) Derive the expression of bad discrimination in the light of optimal acceptance rate.
   (c) Consider the following case and propose two mechanisms step by step to remove bad discrimination from data.
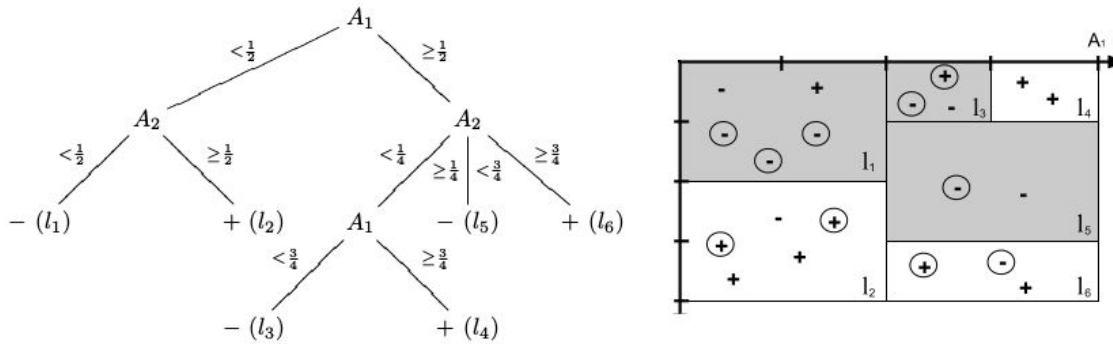
| | medicine | | computer | |
	female	male	female	male
number of applicants	800	200	200	800
acceptance rate	15%	25%	35%	45%
accepted (+)	120	50	70	360

Question 3 [3+3+4 = 10 marks]
   (a) Define a measure to quantify indirect discrimination.
   (b) Define extended lift for direct discrimination scenario.
   (c) How can you prevent direct and indirect discrimination simultaneously?

Question 4 [4+4+3+4+5 = 20 marks]

(a) What is Information Gain of a Class (IGC)? How is it different from Information Gain with respect to a sensitive attribute (IGS)?

(b) How can we map leaf relabeling to a 0/1 knapsack problem?

(c) Consider the following diagram

A_1

$< \frac{1}{2}$ A_2 $\geq \frac{1}{2}$ A_2

$< \frac{1}{2}$ $- (l_1)$ $\geq \frac{1}{2}$ $+ (l_2)$ A_1 $< \frac{1}{4}$ $\geq \frac{1}{4}$ $< \frac{3}{4}$ $- (l_5)$ $\geq \frac{3}{4}$ $+ (l_6)$

$< \frac{3}{4}$ $- (l_3)$ $\geq \frac{3}{4}$ $+ (l_4)$

The left is a decision tree where each leaf is a datapoint in the original database.
Find the discrimination introduced by the decision tree classifier.

(d) Using max(IGC-IGS) as splitting criteria reconstruct the above classifier.

(e) Considering the above example, find out which leaves need to be relabeled in order to remove discrimination using a pseudo polynomial approximation algorithm for 0/1 knapsack.

Question 5 [3+4+4+6+3 = 20 marks]

(a) What is Lipschitz condition for statistical parity?

(b) In what condition individual fairness imply group fairness?

(c) Discuss the issues with simultaneously maintaining Lipschitz condition and statistical parity.

(d) Propose a mechanism to overcome the issue by data transformation.

(e) Estimate the cost of the transformation proposed in the answer to (d).

Question 6 [4+6 = 10 marks]

(a) Given a data, define the steps to construct a Suppes-Bayes Causal Network (SBCN).

(b) Construct a SBCN based on the table below

Male		Female		
Admitted	Denied	Admitted	Denied	Department
512	313	89	19	A
313	207	17	8	B
120	205	202	391	C
138	279	131	244	D
53	138	94	299	E
22	351	24	317	F

Consider the example below where "sex" is a sensitive attribute.

Sex	Ethnicity	Highest degree	Job type	Cl.
M	Native	H. school	Board	+
M	Native	Univ.	Board	+
M	Native	H. school	Board	+
M	Non-nat.	H. school	Healthcare	+
M	Non-nat.	Univ.	Healthcare	−
F	Non-nat.	Univ.	Education	−
F	Native	H. school	Education	−
F	Native	None	Healthcare	+
F	Non-nat.	Univ.	Education	−
F	Native	H. school	Board	+

(a) Calculate the Discrimination on the overall dataset.
(b) Use Massaging technique and relabel examples and reconstruct the dataset to remove discrimination.
(c) Use Reweighting to relabel examples and reconstruct the dataset to remove discrimination.