1 2 9 0 Departamento de Engenharia Informática FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE D COIMBRA

Security and Privacy – MECD
Exam Normal - January 9, 2023
90 min (*sem consulta*)

Nome: **Nº Aluno:**

<u>VERY IMPORTANT</u>: the space for your responses is limited by the space available below. Any text written out of the boxes will not be considered.

1. (2.0 v) One can identify 4 basic anonymization operations: generalization, suppression, anatomization, and perturbation. Choose two of these operations and explain how they work. Use the space for diagrams or examples if it helps you.

Operation: _____

Operation: _____

2. Considering the presented dataset, answer to the following:

a) (2.0 v) Study the dataset in terms of distinction and separation

	age	sex	district	disease
1	30	Female	Lisboa	rhinitis
2	40	Female	Porto	Gastritis
3	30	Female	Lisboa	cancer
4	45	Female	Braga	Covid-19
5	32	Female	Lisboa	cancer
6	40	Female	Porto	Gastric Ulcer
7	45	Female	Braga	pneumonia
8	32	Female	Lisboa	rhinitis
9	20	Male	Coimbra	pneumonia
10	19	Male	Coimbra	Covid-19
11	20	Male	Coimbra	pneumonia
12	40	Female	Porto	Gastritis
13	19	Male	Coimbra	Covid-19

b) (2.0 v) Propose concrete operations to achieve **3-diversity (l=3)** regarding **disease**, indicating the resulting dataset. When applicable, indicate the coding models and hierarchies used and the reasoning behind them.

1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				

Nome: **Nº Aluno:**

c) (1.5 v) Indicate specific privacy issues that remain present in the resulting dataset.

3. (2.0 v) Succinctly, explain how bootstrapping allows going from somewhat homomorphic encryption to fully homomorphic encryption.

4. (2.0 v) Indicate one challenge that the data economy creates to the organizations and explain how an organization must look at that challenge considering the principles of privacy protection/fair information practices.

5. (2.0 v) Considering the methods discussed in class, indicate and describe one method to craft an adversarial attack.

6. (3.0 v) Describe 2 techniques to protect the privacy of subjects when applying deep learning for synthetic data generation. Explain how the techniques help with privacy and when they should be applied (e.g., before, during, or after training the model/synthesizing the data from the trained model).

7. (3.5 v) Classify the following sentences as true or false, correcting the ones classified as false.

a. Information security issues prevail because attacks are usually cheap.

b. Asymmetric encryption must be used to encrypt large volumes of data.

c. Elliptic curves provide the same level of security with smaller keys.

d. K-anonymity cannot deal with sparse data, but l-diversity is effective.

e. Consent to use personal data should have a limited scope but can be hard to withdraw.

f. When using differential privacy, it is very important to know the queries that are going to be performed.

g. Oblivious transfer requires that the receiver is semi-honest.