Information Theory for Data Science

Academic year 2023/2024

List of questions – Sections 1 and 4 – version 1.1

Written exam A: 4 questions taken from a list of 20 questions (time available = 1 hour, closed book exam, maximum grade = 30). List = 20 red questions.

Written exam B: 4 questions taken from a list of 50 questions (time available = 1 hours, closed book exam, maximum grade = 25). List = all 50 questions.

Section 1 and 4

1. Definition, meaning and properties of the entropy of a random variable. Definition and properties of the entropy of a binary random variable. (Proofs not required.)
2. Prove that $H(X) \leq \log_2 M$ (M the cardinality of the alphabet of X).
3. Use Lagrange optimization to find the maximum of H(X).
4. Present and discuss the principle of maximum entropy. (Lagrange formulation not required.) What distributions can you expect for different values of the mean?
5. Present an example of the application of the principle of maximum entropy to a random variable with 2 outcomes. Lagrange formulation is required. Solve it for a mean value equal to the arithmetic average of the two costs.
6. Present the definition of Renyi entropy. Prove that the limit for $\alpha \rightarrow 1$ is the Shannon entropy.
7. Present and discuss the properties of the entropy of a function of a random variable. (Proof not required.)
8. Prove the properties of the entropy of a function of a random variable.
9. Present the definition of permutation entropy.
10. Define the joint entropy H(X,Y). Discuss its meaning. Present and discuss the properties with respect to the entropies H(X) and H(Y). (Proofs not required.)
11. Proof the property of the joint entropy H(X,Y) with respect to the entropies H(X) and H(Y) when H and Y are statistically independent.
12. Proof the property of the joint entropy H(X,Y) with respect to the entropies H(X) and H(Y) when H and Y are not statistically independent.
13. Define the conditional entropy H(X|Y). Discuss its meaning. Present the link with H(X,Y). Present and discuss the properties with respect to H(X). (Proofs not required.) Present the chain rule of entropy.
14. Proof the properties of H(X|Y) with respect to H(X) when X and Y are statistically independent.
15. Proof the properties of H(X|Y) with respect to H(X) when X and Y are not statistically independent.
16. Define the information gain I(X;Y). Discuss its meaning. Present and discuss its properties with respect to H(X). Present the link with the joint and the conditional entropies. (Proofs not required.)

17. Define the Kullback-Leibler divergence. Proof that it is always positive.
18. Define the Kullback-Leibler divergence. Discuss its meaning when we compare an observed and a model distribution.
19. Discuss how and why information gain is used for tree classifiers.
20. Present and discuss the fundamental entropy properties of a secure scheme. Present and discuss the link between the key and message entropies and its consequences. (Proof not required.) Present the one-time pad scheme.
21. Proof the link between the key and message entropies for a secure scheme.
22. Present the difference between symmetric and asymmetric encryption schemes. List (no description) some examples of the two schemes.
23. Present the properties of m-sequences. Compare them against ideal random sequences.
24. Present McEliece encryption and decryption.
25. Discuss different Block Modes (EBC, CBC, CTR).

Note: the questions for Sections 2 and 3 are available on Prof. Taricco's directory.