Information Theory for Data Science

Assignment 1

Introduction to Information Theory and application to Classifiers

Draft version 0.1

Exercises:

1. Entropy of a random variable with 3 outcomes (pt. X)
2. Entropy of a random variable from a data series (pt. X)
3. Application of the principle of maximum entropy (pt. X)

4. …

Exercise 1 - Entropy of a random variable with 3 outcomes

1. Given a random variable with 3 outcomes, write a program to plot the entropy as a function of all possible probability vectors
2. Start with a probability vector where one of the elements is significantly higher than the others. Apply an iterative averaging procedure (for example, replace each element with the average of itself and its neighbors, followed by normalization). For each updated vector, compute the entropy and plot its value on the figure generated in step 1. Show that as the probability distribution approaches the uniform distribution, the entropy approaches its maximum value. Finally, discuss the results.

Exercise 2 - Entropy of a random variable from a data series

1. Identify a data series and estimate the probabilities of the outcomes based on their occurrences, updating the probabilities at each time step.
2. At each time step, compute the entropy, plot its behavior, and discuss the results

Note: In the presentation, include a link to the source of the data series

Exercise 3 - Application of the principle of maximum entropy

Exercise 2.a

1. Invent an exercise where you have a random variable X with an alphabet Ω_X with 2 outcomes with integer values.
2. Show some examples of the the probability distribution P(X) for different values of the mean value.
3. Discuss the results

Exercise 2.b

1. Invent an exercise where you have a random variable X with alphabet Ω_X with at least 4 outcomes, where each outcome has an inteher value (“cost”).
2. Fix the mean value bigger than the arithmetic average of the costs, and apply the principle of maximum entropy to find the probability distribution P(X)
3. Plot P(X)
4. Repeat with a mean value equal to the arithmetic average and plot the result
5. Repeat with other values of the mean value and plot the results
6. Comment the results

You must numerically solve the equation generated by the Lagrange optimization.

As an example , for Matlab you can use

```
syms x
eqn = ( . . .  )*mu == ( . . . );
V = vpasolve(eqn,x,[0 10])
```

Important

Final version assigned on XX/10/2024

Delivery by

- XX/10/2024, 11.59 PM: +2 points
- XX/XX/2024, 11.59 PM: +1 point
- XX/XX/2024, 11.59 PM: 0 points
- **Later: not accepted**