

TTIC 31230 Fundamentals of Deep Learning

Problems for Source Coding.

Problem 1 In the early days of computing attempts were made to build hardware with three level logic — voltage levels representing 0, 1 and 2. Suppose we use codes written in strings over a three letter alphabet to select code words for a given population distribution Pop over a discrete set. We want to find a prefix-free code $c(y)$ over three letter code words for each item y .

(a) Give a function $g(\text{Pop})$ of the population distribution Pop such that we are guaranteed that there exists a code satisfying

$$E_{y \sim \text{Pop}} |c(y)| \leq g(\text{Pop})$$

where $|c(y)|$ is the number of characters in the three letter code for $c(y)$ for y . Give your answer in nats. The convention in this class is that $H(\text{Pop})$ is in units of nats.

(b) Repeat part (a) for an alphabet of size k .