Nature of the Assignment: This simple take home focuses on basic network concepts and foundational tools. 1) question deals with data compression, which is critical for communication systems. 2) pertains fault tolerance in distributed systems. 3) focuses on basic mathematical concepts, which are critical for error detection and correction (e.g., CRC) in computer networks. 4) is about well-known distance metrics to measure similarity for vectors. 5) is on the Markov chains that is one of the most basic tools to analyze the behavior of probabilistic network systems. This assignment aims at extending your knowledge on the complementary aspects of the topics covered during the class (so it goes beyond the material simply put on the slides, as described and emphasized before).

Requirements: Assignment will be completed by each student individually (no collaboration). Directly borrowing (e.g., copy-paste) from any material and putting in solutions (e.g., from online solutions, Wikipedia, or research papers) is plagiarism (see Syllabus for its corresponding actions). Please cite very carefully each resource you use, but citing a solution does not give a license to directly put it as an answer. All of your answers must be in your in own words and interpretations.

Format: Assignment should be prepared by a text editor (e.g., Microsoft Word or Latex). Handwritten submissions are not accepted.

Grading: Note that this is an all or nothing assignment, meaning that your submission will not be accepted unless you solve all the questions.

Extra Credit: This is an extra-credit assignment, and it will be added on top of your overall score at the end of the term. This means that not doing it will not impact your overall score, but doing so will support your grade.

Further Extra Credit: Latex is the de-facto tool for scientific writing. To encourage the use of this critical tool, we will provide further extra credit if you type your assignment with Latex. You can use any package of your choice, but may also consider IEEE conference format.

1) a) Explain the basic idea behind Huffman Coding and where it can be used.
   b) Explain how encoding and decoding works for Huffman Coding
   c) For characters and their occurrence probabilities given below, encode the characters with Huffman Coding and find compression ratio to fixed-length encoding.

<table>
<thead>
<tr>
<th>Characters</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability</td>
<td>0.3</td>
<td>0.3</td>
<td>0.13</td>
<td>0.12</td>
<td>0.1</td>
<td>0.05</td>
</tr>
</tbody>
</table>
2) Explain what is Byzantine Fault Tolerance?

3) Explain what is a Galois Field? Why Galois Field arithmetic offers fast operations?
Please provide your answer mathematical in nature (no hand waiving or unclear verbal descriptions), and be brief but precise.

4) a) Give a brief description of each of the following distances:
   - Euclidean distance
   - Edit distance
   - Hamming distance

   b) Find the Euclidean distance between these two vectors:
   \[ x = [4 \ 1 \ 9 \ 5 \ 3 \ 6], \quad y = [3 \ 9 \ 6 \ 1 \ 0 \ 7] \]

   c) Find the Hamming distance between these bit strings
   \[ a = [01101010], \quad b = [11011011] \]

   d) Find the Edit distance between these two strings: "electrical engineering", "computer science"

5) a) Formally define what is a Markov chain.

   b) Consider the Markov chain with three states, \( S = \{1,2,3\} \), that has the following transition matrix
   \[
   \begin{bmatrix}
   \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\
   \frac{1}{3} & 0 & \frac{2}{3} \\
   \frac{1}{2} & \frac{1}{2} & 0
   \end{bmatrix}
   \]

   c) Given \( P(X1 = 1) = P(X1 = 2) = \frac{1}{4} \), find \( P(X1 = 3, X2 = 2, X3 = 1) \)