

Job Fair 2023

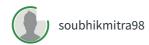
**Engineering Mathematics** 

Discrete Mathematics

Digital Logic and Design

Computer Organiz

# **Introduction to Data Compression**



Read

**Discuss** 

In this article, we will discuss the overview of Data Compression and will discuss its method illustration, and also will cover the overview part entropy. Let's discuss it one by one.

### Overview:

One important area of research is <u>data compression</u>. It deals with the art and science of storing information in a compact form. One would have noticed that many compression packages are used to compress files. Compression reduces the cost of storage, increases the speed of algorithms, and reduces the transmission cost. Compression is achieved by removing redundancy, that is repetition of unnecessary data. Coding redundancy refers to the redundant data caused due to suboptimal coding techniques.

#### Method illustration:

- To illustrate this method let's assume that there are six symbols, and binary code is used to assign a unique address to each of these symbols, as shown in the following table
- Binary code requires at least three bits to encode six symbols. It can also be observed
  that binary codes 110 and 111 are not used at all. This clearly shows that binary code
  is not efficient, and hence an efficient code is required to assign a unique address.

Symbols	W1	W2	W3	W4	W5	W6
Probability	0.3	0.3	0.1	0.1	0.08	0.02
Binary code	000	001	010	011	100	101

• An efficient code is one that uses a minimum number of bits for representing any information. The disadvantage of binary code is that it is fixed code; a Huffman code is

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our <u>Cookie Policy</u> & <u>Privacy Policy</u>

Got It!

• Coding techniques are related to the concepts of entropy and information content, which are studied as a subject called information theory. Information theory also deals with uncertainty present in a message is called the information content. The information content is given as

$$log_{2 (1/pi)}$$
 or  $-log_{2 pi}$ .

## **Entropy:**

• Entropy is defined as a measure of orderliness that is present in the information. It is given as follows:

• Entropy is a positive quantity and specifies the minimum number of bits necessary to encode information. Thus, coding redundancy is given as the difference between the average number of bits used for coding and entropy.

coding redundancy = Average number of bits - Entropy

• By removing redundancy, any information can be stored in a compact manner. This is the basis of data compression.

Last Updated: 27 Jul, 2021

#### 4

# Similar Reads

1. Difference between Lossy Compression and Lossless Compression

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our <u>Cookie Policy</u> & <u>Privacy Policy</u>

- 3. Data Compression With Arithmetic Coding
- 4. Shannon-Fano Algorithm for Data Compression
- 5. PGP Compression
- 6. Compression of IPv6 address
- 7. Difference between Inter and Intra Frame Compression
- 8. DjVu Compression in Computer Network
- 9. LZW (Lempel-Ziv-Welch) Compression technique
- 10. Difference between Data Privacy and Data Security

Previous

# **Article Contributed By:**



soubhikmitra98 soubhikmitra98

# Vote for difficulty

Current difficulty: Easy

Easy

Normal

Medium

Hard

Expert

Article Tags: Computer Networks

Practice Tags: Computer Networks

Improve Article

Report Issue



We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our <u>Cookie Policy</u> & <u>Privacy Policy</u>

### feedback@geeksforgeeks.org

Company Explore

About Us Job Fair For Students

Careers POTD: Revamped

In Media Python Backend LIVE

Contact Us Android App Development

Terms and Conditions DevOps LIVE

Privacy Policy DSA in JavaScript

Copyright Policy

Third-Party Copyright Notices

Advertise with us

Languages Data Structures

Python Array

Java String

C++ Linked List

GoLang Stack

SQL Queue

R Language Tree

Android Tutorial Graph

Algorithms Web Development

Sorting HTML

Searching CSS

Greedy JavaScript

Dynamic Programming Bootstrap

Pattern Searching ReactJS

Recursion AngularJS

Backtracking NodeJS

Data Science & ML Interview Corner

Data Science With Python Company Preparation

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our <u>Cookie Policy</u> & <u>Privacy Policy</u>

**GfG School** 

Write & Earn

Maths For Machine Learning Experienced Interview
Pandas Tutorial Internship Interview
NumPy Tutorial Competitive Programming
NLP Tutorial Aptitude

**Python** 

Python Tutorial CBSE Notes for Class 8

Python Programming Examples CBSE Notes for Class 9

Django Tutorial CBSE Notes for Class 10

Python Projects CBSE Notes for Class 11

Python Tkinter CBSE Notes for Class 12

OpenCV Python Tutorial English Grammar

UPSC/SSC/BANKING

SSC CGL Syllabus Write an Article

SBI PO Syllabus Improve an Article

IBPS PO Syllabus Pick Topics to Write

UPSC Ethics Notes Write Interview Experience

UPSC Economics Notes Internships

UPSC History Notes Video Internship

@geeksforgeeks, Some rights reserved