DATA COMPRESSION USING HUFFMANN CODING ALGORITHM

Major Project report submitted in partial fulfilment of the requirement for the award of the degree of
BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING

Submitted by:

Mridul Gupta (206301038),
Shivam Kumar (216303001),
Chandan Kumar Singh (206301082),
Siddharth Sharma(206301112)

Under the guidance of:
Dr. Suyash Bhardwaj
Assistant professor

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING FACULTY OF ENGINEERING AND TECHNOLOGY GURUKUL KANGRI UNIVERSITY
2023-2024

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who have contributed help towards the successful completion of this Major Project “Data Compression using Huffman Coding Algorithm **“**. Their support, guidance, and encouragement were invaluable throughout this journey.

We are all deeply thankful to Dr. Suyash Bhardwaj, our project supervisor, for their continuous support and guidance. Their expertise in data structures and algorithms and insightful feedback greatly enriched the quality of this research.

This project would not have been possible without info to high-quality algos. We extend our appreciation to Google for providing the researchs of different institiutes that formed the foundation of our project.

We would like to acknowledge the support of the Department of Computer Science and Engineering for providing the necessary resources and infrastructure essential for the successful execution of this project.

CERTIFICATE

This is to certify that the content of this Major Project report is compiled and written by us, and we have not copied them from anywhere. This work is done by us, and we have not presented it anywhere else.

Mridul Gupta	Shivam Kumar	Chandan Kumar Singh	Siddharth Sharma
(206301038)	(216303001)	(206301082)	(206301112)

CERTIFICATE

This is to certify that the project work embodied in this project file is done under my supervision and is not being copied from anywhere or presented anywhere for any sort of benefit.

Dr. Suyash Bhardwaj
Assistant Professor
Department of Computer Science & Engineering
Faculty of Engineering and Technology

ABSTRACT

Huffman coding is a widely used technique for lossless data compression, particularly efficient for files with non-uniform probability distributions. The implementation utilizes a binary tree-based approach to generate optimal prefix codes for each symbol in the input data. Through Huffman coding, the algorithm achieves compression by replacing frequently occurring symbols with shorter codewords and less frequent symbols with longer codewords. The process involves constructing a Huffman tree based on the frequency of symbols in the input data, followed by encoding the data using the generated Huffman codes. The encoded data is then decoded using the same Huffman tree to reconstruct the original input. The efficiency of the compression depends on the frequency distribution of symbols in the input data, with higher compression ratios achievable for datasets with more redundancy. The implementation provides a versatile and efficient solution for data compression applications, offering a balance between compression ratio and computational complexity. Performance evaluation demonstrates the effectiveness of the Huffman coding technique in reducing the size of the data while preserving its integrity. Overall, this abstract highlights the significance of Huffman coding as a fundamental method for data compression in various computational applications.