

BAB I PENDAHULUAN

A. Latar Belakang Masalah

Kehadiran teknologi di era globalisasi saat ini tentunya sangat mempengaruhi setiap kegiatan. Hal ini dikarenakan peran teknologi tidak bisa dipisahkan dalam kehidupan sehari-hari, baik itu teknologi informasi, komunikasi sampai teknologi yang berhubungan dengan penyimpanan data. Teknologi terus berkembang seiring majunya sebuah peradaban. Dari jaman dahulu hingga kini, teknologi dibuat untuk memudahkan manusia dalam melakukan berbagai hal.

Di jaman sekarang tidak sedikit pekerjaan yang berhubungan dengan komputer dan menghasilkan data besar setiap harinya, baik itu pekerjaan dengan menggunakan software maupun aplikasi. Selain kapasitas data yang berhubungan dengan software maupun aplikasi sebagai media bekerja, kapasitas game sebagai media hiburan pun semakin besar kapasitasnya hingga mencapai puluhan GigaByte. Tentunya hal ini mempengaruhi kinerja computer, laptop maupun handphone sebagai media penggunaan yang membuat media-media tersebut menjadi lamban, hang bahkan error.

Semakin berkembangnya teknologi kebutuhan akan informasi sangatlah diperlukan oleh masyarakat umum. Dengan semakin banyaknya informasi yang perlu disimpan secara digital, secara otomatis akan meningkatkan keperluan untuk menyediakan media penyimpanan data yang lebih besar lagi, tentunya tuntutan pengiriman data baik berupa gambar, file, audio, video maupun data lainnya juga dituntut untuk serba cepat dan mudah. Namun didalam pengiriman data itu sendiri terkadang terkendala dengan kapasitas atau ukuran data yang cukup besar bahkan sangat besar, sehingga membuat data harus di perkecil terlebih dahulu kapasitas atau ukurannya tetapi hal ini menyebabkan data terutama data berupa video dan gambar menjadi blur. Oleh karena itu diperlukan suatu alternatif mekanisme penyimpanan data sehingga dengan media penyimpanan yang ada, kita dapat menyimpan data sebanyak-banyaknya dengan kapasitas yang paling kecil tanpa merusak kualitas data.

Di zaman sekarang teknologi media penyimpanan memang terus berkembang hingga saat ini, bisa dirasakan saat dahulu dan sekarang perkembangan media penyimpanan terus mempunyai inovasi dan teknologi baru, dahulu media penyimpanan pertama kali di temukan berukuran sangat besar, bisa mencapai satu ruangan ukuran sedang, dan hanya bisa menampung beberapa megabyte data saja selain ukurannya yang besar media penyimpanan zaman dahulu sangat lambat dalam hal membaca data yang sudah tersimpan pada media penyimpanan, beda hal nya dengan sekarang dimana media penyimpanan sudah canggih dimana ukurannya

sudah semakin kecil namun kapasitas yang tersedia saat ini sudah ber puluh kali lipat besarnya jika dibandingkan dengan media penyimpanan zaman dahulu, selain kapasitasnya yang beragam, ada beberapa jenis juga dan semuanya memiliki kelebihan dan kekurangannya, dan semua kelebihan dan kekurangannya itu berbanding dengan harga media penyimpanannya. Namun harga yang masih tinggi untuk media penyimpan menjadi keraguan tersendiri untuk menambahkan atau mengganti media penyimpanan dengan yang baru.

Dunia Pendidikan juga tidak terlepas dari pengaruh perkembangan teknologi maupun informasi terutama dalam hal penyimpanan data. Pasalnya, di setiap sekolah tentunya memiliki arsip data digital baik berupa dokumen, foto, audio video maupun aplikasi yang setiap tahun atau sedikitnya setiap bulan akan terus bertambah kapasitas dan jumlah datanya. Hal ini tentunya membuat media penyimpanan data yang tersedia juga akan cepat habis. Jika hal ini sudah terjadi, maka hal umum yang akan dilakukan adalah membeli media penyimpanan yang baru, tentunya hal ini sangat tidak efektif mengingat media penyimpanan data saat ini seperti Hardisk contohnya, cukup mahal dan sangat tidak mungkin jika setiap bulan harus membeli hardisk karena akan memakan biaya yang mahal sebab habisnya kapasitas penyimpanan tidak hanya terjadi di satu komputer saja.

Dalam hal ini, penelitian menekankan pada upaya memampatkan seoptimal mungkin tanpa merubah kualitas data dengan menggunakan Algoritma Huffman yang dapat digunakan dengan media mobile maupun desktop/komputer. Dimana algoritma huffman dapat digunakan untuk melakukan pemampatan atau mengecilkan data sehingga ukuran data yang di hasilkan menjadi lebih kecil dari ukuran aslinya. Algoritma ini digunakan untuk mencari kode biner yang paling efisien menggunakan pohon biner yang diurutkan berdasarkan frekuensi.

Penggunaan komputer di hampir setiap harinya selalu memakan kapasitas penyimpanan data yang cukup besar dan jika terus menerus media penyimpanan akan semakin sedikit kapasitasnya yang tersedia. Beberapa file penting seperti arsip ijazah maupun kebutuhan ujian sekolah tahunnya yang terus bertambah juga dapat mempercepat habisnya kapasitas penyimpanan yang ada. Untuk merapihkan dan menghemat kapasitas penyimpanan data tersebut bisa dengan menggunakan teknik pemampatan data agar data tidak perlu dihapus agar ruang penyimpanan data tetap dapat memenuhi kebutuhan yang ada. Pemampatan data dapat di gunakan kepada berbagai file yang sudah selesai digunan dan dijadikan arsip untuk kepentingan kearsipan sekolah nantinya. Data-data yang dapat dimampatkan atau dikompresi yaitu dokumen baik berupa text, gambar, audio, video maupun software dan aplikasi.

Pemampatan atau kompresi data merupakan salah satu metode untuk memperkecil ruang penyimpanan data pada suatu media penyimpanan. Selain berguna dalam penyimpanan data, kompresi data dapat membantu memperkecil ukuran data yang ditransmisikan di dalam suatu media jaringan, seperti internet sehingga memperkecil waktu transfer data. Metode pengecilan ukuran data memiliki dua jenis, yaitu *lossy compression* dan *lossless compression*. Kedua metode ini memiliki kelebihan dan kekurangannya masing masing. Dengan mengecilkan ukuran data bisa memotong kapasitas media penyimpanan yang dimiliki, dan sisa media penyimpanan yang ada bisa di manfaatkan untuk kebutuhan lainnya juga, seperti menginstall program lainnya dan game yang membutuhkan kapasitas penyimpanan yang besar. Aplikasi pengecilan ukuran data adalah aplikasi yang berfungsi untuk pemampatan data atau mengecilkan data dari ukuran besar menjadi lebih kecil tanpa memangkas kelengkapan isi datanya. Dengan aplikasi pemampatan data, data yang dimiliki bisa di kurangi kapasitasnya tanpa mengurangi kelengkapan isi data tersebut.

Data adalah file dan berbagai file yang terkumpul menjadi satu. Data juga termasuk bagian penting di dalam kehidupan, di dalam pekerjaan yang menggunakan komputer pasti ada saja data yang dikumpulkan, baik data yang ukuran kecil sampai ukuran besar, tergantung pekerjaan yang di jalankannya. Data juga bervariasi ukurannya mulai dari byte, kilobyte, megabyte, gigabyte, dan masih lebih besar lagi. File juga beragam macam jenis nya, mulai dari file yang biasa di gunakan di kantor seperti file microsoft office, data aplikasi komputer yang ber ekstensi exe, file audio, file video, file gambar, biasanya file besar terdapat pada aplikasi game, foto dan video editing. dan masih banyak file file lainnya juga.

Dengan salah satu algoritma yang bisa mengecilkan ukuran data, yaitu algoritma Huffman. Algoritma Huffman merupakan algoritma *lossless compression*, yaitu teknik kompresi yang tidak mengubah data aslinya. Hal tersebut yang menyebabkan algoritma ini banyak dipakai dalam proses kompresi. Algoritma Huffman bekerja dengan cara melakukan pengkodean dalam bentuk bit untuk mewakili data karakter. Algoritma Huffman di perkenalkan oleh David Albert Huffman pada tahun 1951

Karakter ↕	Frekuensi ↕	Kode ↕
(spasi)	7	111
a	4	010
e	4	000
f	3	1101
h	2	1010
i	2	1000
m	2	0111
n	2	0010
s	2	1011
t	2	0110
l	1	11001
o	1	00110
p	1	10011
r	1	11000
u	1	00111
x	1	10010

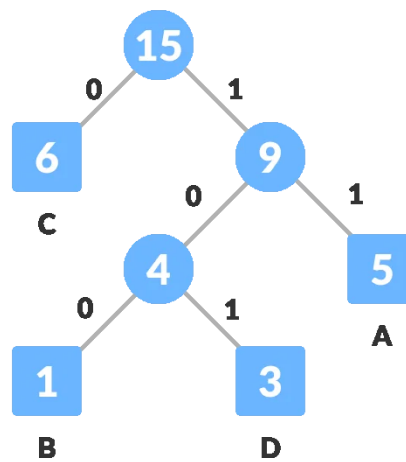
Gambar 1.1 Frekuensi Algoritma Huffman

Lossless Compression adalah teknik pengecilan data yang mampu merekonstruksi bentuk asli data. Lossless Compression ini memungkinkan untuk mengembalikan file ke bentuk aslinya. *Lossless Compression* pada algoritma Huffman dalam penelitian ini dapat di terapkan ke format file apapun yang dapat meningkatkan kinerja rasio Pengecilan data. Algoritma huffman menggunakan teknik mengecilkan data karakter American Standard Code for Information Interchange (ASCII). Chart penggunaan American Standard Code for Information Interchange (ASCII) dapat dilihat pada gambar 1.1. berikut:

Dec	Hex	Name	Char	Ctrl-char	Dec	Hex	Char	Dec	Hex	Char	Dec	Hex	Char
0	0	Null	NUL	CTRL-@	32	20	Space	64	40	@	96	60	`
1	1	Start of heading	SOH	CTRL-A	33	21	!	65	41	A	97	61	a
2	2	Start of text	STX	CTRL-B	34	22	"	66	42	B	98	62	b
3	3	End of text	ETX	CTRL-C	35	23	#	67	43	C	99	63	c
4	4	End of xmit	EOT	CTRL-D	36	24	\$	68	44	D	100	64	d
5	5	Enquiry	ENQ	CTRL-E	37	25	%	69	45	E	101	65	e
6	6	Acknowledge	ACK	CTRL-F	38	26	&	70	46	F	102	66	f
7	7	Bell	BEL	CTRL-G	39	27	'	71	47	G	103	67	g
8	8	Backspace	BS	CTRL-H	40	28	(72	48	H	104	68	h
9	9	Horizontal tab	HT	CTRL-I	41	29)	73	49	I	105	69	i
10	0A	Line feed	LF	CTRL-J	42	2A	*	74	4A	J	106	6A	j
11	0B	Vertical tab	VT	CTRL-K	43	2B	+	75	4B	K	107	6B	k
12	0C	Form feed	FF	CTRL-L	44	2C	,	76	4C	L	108	6C	l
13	0D	Carriage feed	CR	CTRL-M	45	2D	-	77	4D	M	109	6D	m
14	0E	Shift out	SO	CTRL-N	46	2E	.	78	4E	N	110	6E	n
15	0F	Shift in	SI	CTRL-O	47	2F	/	79	4F	O	111	6F	o
16	10	Data line escape	DLE	CTRL-P	48	30	0	80	50	P	112	70	p
17	11	Device control 1	DC1	CTRL-Q	49	31	1	81	51	Q	113	71	q
18	12	Device control 2	DC2	CTRL-R	50	32	2	82	52	R	114	72	r
19	13	Device control 3	DC3	CTRL-S	51	33	3	83	53	S	115	73	s
20	14	Device control 4	DC4	CTRL-T	52	34	4	84	54	T	116	74	t
21	15	Neg acknowledge	NAK	CTRL-U	53	35	5	85	55	U	117	75	u
22	16	Synchronous idle	SYN	CTRL-V	54	36	6	86	56	V	118	76	v
23	17	End of xmit block	ETB	CTRL-W	55	37	7	87	57	W	119	77	w
24	18	Cancel	CAN	CTRL-X	56	38	8	88	58	X	120	78	x
25	19	End of medium	EM	CTRL-Y	57	39	9	89	59	Y	121	79	y
26	1A	Substitute	SUB	CTRL-Z	58	3A	:	90	5A	Z	122	7A	z
27	1B	Escape	ESC	CTRL-[59	3B	;	91	5B	[123	7B	{
28	1C	File separator	FS	CTRL-\	60	3C	<	92	5C	\	124	7C	
29	1D	Group separator	GS	CTRL-]	61	3D	=	93	5D]	125	7D	}
30	1E	Record separator	RS	CTRL-^	62	3E	>	94	5E	^	126	7E	~
31	1F	Unit separator	US	CTRL-?	63	3F	?	95	5F	?	127	7F	DEL

Gambar 1.2 American Standard Code for Information Interchange (ASCII) Chart

Algoritma Huffman akan membangun pohon biner penuh untuk berbagai simbol setelah menghitung probabilitas masing-masing simbol dan menempatkannya dalam urutan menurun. Seperti pohon biner pada gambar 2.1. berikut.



Gambar 1.3 Pohon Algoritma Huffman

Dengan menggunakan aplikasi pemampatan data diharapkan bisa mengurangi ukuran kapasitas data dan bisa digunakan untuk merapihkan data arsip yang biasa di simpan setelah selesai digunakan. Selain itu pemampatan data diharapkan bisa mempercepat pengiriman data menggunakan email yang kapasitasnya sudah lebih kecil dari ukuran aslinya.

Berdasarkan latar belakang yang telah dikemukakan di atas maka penelitian ini mengambil judul ***“Penerapan Algoritma Huffman Untuk Pemampatan File”***

B. Permasalahan

Media penyimpanan merupakan perangkat yang digunakan untuk menyimpan berbagai macam data digital baik dalam bentuk dokumen, suara, gambar, video, dan lain sebagainya. Seiring berkembangnya jaman, teknologi media penyimpanan data pun kini telah mengalami perubahan. Tentunya, sebuah media penyimpanan harus memiliki kapasitas yang cukup untuk menyimpan semua data-data yang dibutuhkan.

Penggunaan media peyimpanan kian hari kian bertambah, baik itu didalam dunia industri, sampai kedalam dunia pendidikan. Sebagai contoh didalam dunia pendidikan, dimana dalam jangka Panjang setiap tahunnya, data mengenai peserta didik akan terus bertambah, baik itu data diri peserta didik, data ujian , ijazah maupun administrasi lainnya yang akan disimpan dalam bentuk arsip digital dan tentunya akan terus memenuhi kapasitas media penyimpanan yang ada.

Selama ini, untuk mengatasi permasalahan habisnya kapasitas penyimpanan yaitu dengan menghapus beberapa data yang dianggap sudah tidak diperlukan lagi pada saat ini, namun hal ini akan Kembali menjadi masalah jika data-data yang sudah dihapus akan dipakai atau diperlukan dikemudian hari. Selain menghapus data-data yang dianggap sudah tidak digunakan, Adapun cara lain yaitu dengan membeli dan menambah media penyimpanan baru, tetapi hal ini pun dirasa kurang efektif karena membutuhkan biaya yang cukup mahal.

Selain kedua cara tersebut, terdapat cara yang lebih sering digunakan yaitu dengan memampatkan data, dimana data yang dirasa memiliki ukuran kapasitas besar, akan diperkecil ukuran kapasitasnya dengan merubahnya menjadi .zip atau .Rar. Namun, untuk pemampatan data tersebut memiliki batas maksimal data yang dapat dimampatkan, yaitu 4 GigaByte.

Algoritma Huffman dapat digunakan untuk merubah ukuran file besar menjadi lebih kecil. Cara kerjanya adalah dengan mengubah isi setiap data menjadi sebuah kode ASCII dan diubah menjadi kode biner, setelah menjadi kode biner di pilih kembali kode biner yang bisa di hapus dan dilakukan penghapusan, setelah proses penghapusan dan data dimampatkan data akan menjadi lebih kecil ukurannya, selesai dimampatkan data tidak bisa dibuka dan untuk membuknya kembali harus melalui proses dekompresi, untuk mengembalikan keseluruhan data yang sudah terhapus kode binernya oleh proses pemampatan.

Berdasarkan hal yang sudah dikemukakan sebelumnya, permasalahan yang terjadi mengenai media penyimpanan yaitu besarnya ukuran data digital yang membuat kapasitas media penyimpanan yang tersedia cepat habis atau tidak dapat digunakan secara optimal. Dimana untuk mengatasi permasalahan tersebut harus menghapus beberapa data untuk mengurangi isi media penyimpanan maupun membeli dan menambah media penyimpanan baru yang cukup mahal. Berikut adalah beberapa media penyimpanan yang di ambil dari beberapa komputer staff sekolah.

Tabel 1.1 Kapasitas Penyimpanan

Nama PC	Total Kapasitas	Kapasitas Terpakai	Sisa Kapasitas
PC Staff 1	500 GB	221 GB	279 GB
PC Staff 2	250 GB	180,5 GB	52 GB
PC Staff 3	250 GB	207 GB	43 GB
PC Staff 4	250 GB	195 GB	55 GB
PC Staff 5	250 GB	210,6 GB	39,4 GB

Dari tabel di atas bisa di lihat berbagai macam komputer dengan kapasitas yang berbeda beda, dan kapasitas yang tersedia juga berbeda beda. Kapasitas itu akan terus berkurang karena pemakaian berkala.

Data yang biasanya terdapat di dunia pendidikan, sekolah khususnya biasanya menghabiskan kapasitas media penyimpanan, berikut adalah data penyimpanan pihak sekolah dapat dilihat pada tabel 1.2 berikut:

Tabel 1.2 Penggunaan Data

No	Jenis Data	Keseluruhan Data Yang digunakan
1	Data Peserta Didik	Dalam Gibabyte
	Data Ujian Sekolah	Sekitar 3-10 GB Per Tahun
	Data Ijazah	
	Data Ujian Praktik	
	Data Uji Kompetensi	
	Data Keuangan	
2	Administrasi Guru/Pegawai	Dalam Gibabyte
	Data Diri	Sekitar 2 -5 GB Per tahun
	Data Absensi	
	Data Honor/Finger	
	Data Administrasi Guru	

Data diatas yaitu data yang biasa di hasilkan setiap tahunnya dengan rutinitas yang sama, tetapi selain data di atas, ada juga data lain yang turut di simpan dalam media penyimpanan komputer.

Dengan banyaknya data yang tidak tersusun dengan baik dan besarnya kapasitas data-data tersebut tentunya menyebabkan media penyimpanan menjadi lebih cepat habis atau berkurang, selain itu penyebab lainnya yaitu penggunaan komputer yang hampir setiap harinya menyimpan data. Hal ini menjadi permasalahan yang harus segera diatasi agar kapasitas media penyimpanan digital tersebut tidak cepat habis dan dapat digunakan secara optimal. Maka, demikian dapat diidentifikasi masalah sebagai berikut:

1. Identifikasi masalah

Dari permasalahan di atas bisa di simpulkan beberapa masalah sebagai berikut :

- a. Belum optimal dalam penggunaan media penyimpanan untuk menyimpan data

b. Belum efektif di dalam proses penanganan penyimpanan data

2. Pokok Masalah

Adapun pokok permasalahan dari penelitian ini, yaitu belum optimal dan efektif di dalam pemampatan file

3. Research question

Pertanyaan penelitian yang dapat diajukan dari penelitian ini adalah :

1. Bagaimana Penerapan Algoritma Huffman Untuk Pemampatan File dapat mengoptimalkan media penyimpanan?
2. Berapa tingkat optimasi dan efektifitas Penerapan Algoritma Huffman Untuk Pemampatan File?

C. Maksud dan Tujuan Penelitian

Maksud dari penelitian ini adalah menerapkan algoritma huffman untuk pemampatan file.

Tujuan penelitian yang ingin di capai adalah :

- a) mendapatkan ukuran file yang lebih optimal
- b) mendapatkan proses yang lebih efektif di dalam memampatkan data
- c) mengembangkan prototype aplikasi pemampatan data
- d) mengukur tingkat optimalisasi dan efektifitas penerapan algoritma huffman

D. Signifikansi Penelitian

Penelitian ini dilakukan dalam rangka mengembangkan teknik komputasi pemodelan algoritma Huffman untuk memampatkan data maupun file sehingga dapat mengoptimalkan penggunaan media penyimpanan digital. Dimana akan dibuat atau dikembangkan sebuah aplikasi menggunakan algoritma Huffman yang akan membuat suatu file atau data menjadi lebih kecil ukurannya tanpa merubah kualitas data atau file itu sendiri. Dengan adanya aplikasi ini diharapkan dapat membuat file atau data menjadi lebih kecil ukurannya sehingga media penyimpanan digital yang digunakan dapat dioptimalkan penggunaannya. Manfaat dari penelitian ini yaitu:

- (a) Manfaat Teoritis, memberikan sumbangan ilmu pengetahuan, dalam Penerapan algoritma Huffman untuk memperkecil kapasitas suatu file atau data
- (b) Manfaat secara praktis, yaitu membuat suatu file atau data menjadi lebih kecil ukurannya tanpa merubah kualitas data atau file tersebut sehingga penggunaan media penyimpanan digital dapat dioptimalkan dan file maupun data tersebut dapat termanajemen dengan baik

(c) Manfaat kebijakan, dapat menjadi acuan dalam manajemen data maupun file dan dalam menggunakan media penyimpanan secara optimal.

E. Spesifikasi Yang Diharapkan

Spesifikasi hasil yang diharapkan dari penelitian ini adalah terciptanya sebuah aplikasi yang mampu memampatkan suatu data atau file untuk mengoptimalkan penggunaan media penyimpanan. Dengan adanya aplikasi ini, kapasitas media penyimpanan yang ada tidak cepat berkurang atau habis sehingga penggunaan media penyimpanan dapat dioptimalkan.

F. Asumsi Dan Keterbatasan Pengembangan

1. Asumsi

Penanganan yang sering dilakukan yaitu langsung menaruh data dengan ukuran aslinya, cara yang biasa dilakukan setiap harinya. Cara itu dapat membuat media penyimpanan cepat penuh. Dengan adanya penelitian ini ada beberapa asumsi yang ingin dicapai :

- a. Penerapan ALGORITMA HUFFMAN lebih menekankan kepada file yang berjenis docx, pptx, xlsx, pdf, png, jpg
- b. Penerapan ALGORITMA HUFFMAN tidak disarankan untuk kompresi file yang bersifat rahasia

2. Keterbatasan penelitian

Penelitian ini memiliki keterbatasan pengembangan, yaitu:

Algoritma huffman tidak bisa memastikan seberapa besar rasio pemampatan data yang dapat dimampatkan.

Algoritma Huffman tidak bisa menghasilkan estimasi waktu pada saat proses kompresi dan dekompresi file

Keterbatasan aplikasi ini adalah :

Aplikasi ini tidak bisa memilih tempat simpan untuk file yang akan di kompresi dan di dekompresi.

G. Definisi istilah

Dalam penelitian pasti selalu ada definisi istilah yang digunakan, berikut adalah beberapa definisi istilah yang di gunakan dalam penelitian ini :

1. Compression

Kompresion adalah teknik memampatkan atau mengecilkan ukuran file tanpa mengurangi isi atau memangkas isi file itu.

2. Aplikasi

Aplikasi adalah perangkat lunak yang berfungsi untuk membantu mempermudah pekerjaan manusia, aplikasi biasa disebut software yang hanya bisa di fungsikan dan tidak dapat di pegang langsung

3. Algoritma

Algoritma adalah prosedur atau langkah langkah dalam perhitungan, untuk menyelesaikan suatu masalah

4. Lossy Compression

Lossy compression adalah teknik pengecilan data yang menghilangkan beberapa data yang tidak terlihat.

5. Lossless Compression

Lossless Compression adalah teknik pengecilan databesar yang mampu merekonstruksi bentuk asli data. Lossless Compression ini memungkinkan untuk mengembalikan file ke bentuk aslinya.

6. Media penyimpanan

Media penyimpanan adalah sebuah alat yang dibuat untuk menyimpan berbagai data dengan berbagai kapasitas, dan media penyimpanan memiliki kapasitas yang bervariasi

7. dekompressi

Dekompressi adalah teknik mengembalikan data ke ukuran aslinya yang sudah melalui proses kompresi sebelumnya.