

Analisis Algoritma Shannon-Fano Dalam Kompresi Data Pengajuan Proposal Skripsi Mahasiswa STMIK CIC Cirebon

Chairun Nas¹, Wanda Ilham², Ilwan Syafrinal³,

^{1,2,3}Manajemen Informatika, Universitas Catur Insan Cendekia
Jl. Kesambi No.202, Cirebon, Jawa Barat

Email: chairun.nas@cic.ac.id¹, wandailham@cic.ac.id², ilwan.syafrinal@cic.ac.id³

Abstrak. Skripsi merupakan tantangan terakhir bagi seorang mahasiswa dalam mencapai gelar sarjana. Dalam proses nya, mahasiswa perlu mengajukan proposal tentang skripsi yang akan dikerjakan. Sebelumnya pengajuan proposal skripsi masih dilakukan secara manual, sehingga perlu dikembangkan sebuah sistem agar pengajuan proposal skripsi dapat dilakukan secara online. Namun permasalahan yang terjadi, proposal yang diajukan diberi batasan panjang karakter karena keterbatasan ruang penyimpanan. Hal ini mengakibatkan proposal yang di ajukan mengalami pemotongan karakter sehingga isi dari proposal yang diajukan tidak jelas. Untuk mengatasi masalah tersebut, maka perlu dilakukan analisis agar dapat melakukan kompresi terhadap data proposal. Analisis yang dilakukan dalam kompresi data proposal tersebut adalah dengan menggunakan algoritma Shannon-Fano. Algoritma Shannon-fano melakukan pemampatan data terhadap karakter sehingga dapat menekan kapasitas data. Pada pengujian, data dengan ukuran 69 byte, dilakukan kompresi dengan algoritma Shannon-Fano sehingga ukuran data menjadi 35 byte dengan rasio kompresi sebesar 48,98%. Maka dengan algoritma Shannon-Fano tersebut, telah dapat melakukan kompresi terhadap data proposal pengajuan skripsi.

Kata kunci: *Kompresi, Algoritma, Shannon-Fano, Proposal.*

1 Pendahuluan

Capaian akhir seorang mahasiswa dalam mengenyam ilmu di bangku kuliah adalah untuk memperoleh gelar sarjana. Dengan memperoleh gelar sarjana, maka akan membuka peluang yang besar bagi seseorang untuk dapat bersaing di dunia kerja. Untuk mencapai gelar sarjana tersebut, maka seorang mahasiswa perlu melalui berbagai rintangan, salah satunya adalah skripsi atau tugas akhir. Skripsi atau tugas akhir merupakan karya ilmiah yang ditulis melalui kegiatan perencanaan, pelaksanaan dan hasil penelitian ilmiah oleh mahasiswa jenjang program sarjana muda atau sarjana [1]. Dalam prosesnya, sebelum melakukan kegiatan skripsi, perlu dilakukan pengajuan proposal skripsi agar pihak kampus dapat mengetahui pembimbing yang cocok untuk skripsi tersebut.

Proposal skripsi adalah bentuk gambaran, desain penelitian atau usulan penelitian dalam rancangan pembuatan skripsi yang akan diajukan. Dalam pengajuan proposal skripsi, perlu memenuhi isi kandungan seperti latar belakang masalah, pemecahan masalah, metodologi penelitian dan lain-lain. Saat ini, pada STMIK CIC Cirebon, pengajuan proposal skripsi dilakukan secara online dengan menggunakan Sistem Informasi Manajemen. Sehingga mahasiswa dapat melakukan pengajuan proposal skripsi dimanapun dalam waktu yang telah ditentukan. Namun permasalahan yang terjadi adalah, terbatasnya ruang penyimpanan, sehingga banyak karakter pada pengajuan proposal skripsi dibatasi. Hal ini berakibat pada tidak jelasnya isi yang disampaikan pada proposal skripsi yang diajukan. Selain dari itu, dengan diberinya

batasan panjang karakter, mengakibatkan terpotongnya bagian isi dari proposal. Maka untuk mengatasi permasalahan ini diperlukan sebuah analisis yang dapat melakukan kompresi data, sehingga kapasitas data yang diinputkan dapat berukuran kecil.

Analisis yang dilakukan untuk kompresi data dapat dilakukan dengan menggunakan algoritma *Shannon-Fano*. Algoritma *Shannon-Fano* merupakan algoritma kompresi yang sangat baik dalam pengkompresian data teks, dimana didasarkan pada *variable-length code* yang berarti beberapa karakter pada data yang akan dikodekan, direpresentasikan dengan kode yang lebih pendek dari karakter yang ada pada data [2]. Algoritma *Shannon-Fano* akan mencari probabilitas dari setiap simbol atau karakter yang muncul pada data [3]. Maka dapat diartikan semakin banyak karakter yang memiliki persamaan akan semakin pendek dalam pengkodean.

Pada penelitian sebelumnya, algoritma *Shannon-Fano* telah dilakukan pengujian untuk kompresi file teks. Pengujian dilakukan dengan 2 kategori, yaitu file uji dengan ukuran yang sama dan memiliki karakter yang bervariasi, dan file uji dengan ukuran yang sama dan karakter yang tidak bervariasi. Hasil pengujian diperoleh bahwa karakter yang bervariasi akan menghasilkan kompresi yang rendah, sedangkan karakter yang tidak bervariasi akan menghasilkan kompresi yang tinggi. Berdasarkan data uji sebanyak 15 file uji yang dilakukan dengan algoritma *Shannon-Fano*, diperoleh rata-rata kompresi sebesar 11,96% dengan rasio kompresi sebesar 88,04% [4].

Selain dari itu, algoritma *Shannon-Fano* juga dapat melakukan kompresi data citra pada penelitian yang telah dilakukan sebelumnya. Citra yang digunakan dibagi menjadi 2 yaitu citra *homogen* dan citra *heterogen*. Selanjutnya citra tersebut dilakukan pengujian dan dihasilkan rasio kompresi. Hasil rasio kompresi pada citra *homogen* diperoleh rata-rata sebesar yaitu 52%, sedangkan untuk citra *heterogen* memperoleh rasio kompresi rata-rata sebesar 8,6%. Pada pengujian kedua citra, algoritma *Shannon-Fano* telah dapat melakukan kompresi terhadap data citra dengan baik [5].

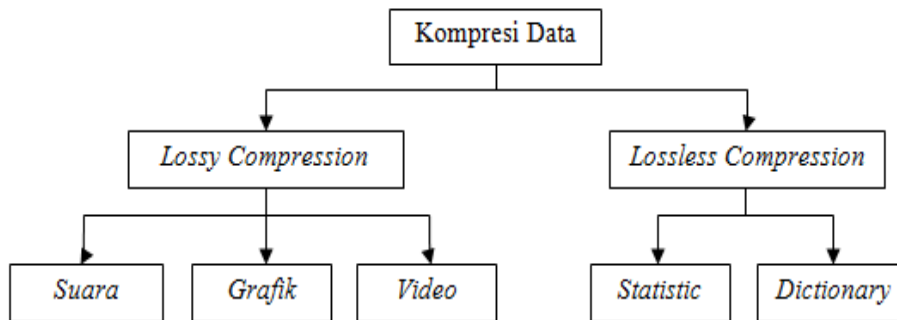
Maka dari penelitian yang telah dibahas sebelumnya, maka algoritma *Shannon-Fano* cocok digunakan dalam kompresi data pengajuan proposal skripsi pada sistem informasi manajemen STMIK CIC.

2 Kajian Pustaka

2.1 Kompresi Data

Kompresi data adalah proses mengkonversikan sebuah *input data stream* (*stream* sumber atau data mentah asli) menjadi data *stream* lainnya (*bit stream* hasil atau *stream* yang telah terkompresi) yang berukuran lebih kecil [6]. Kompresi data dilakukan terhadap data digital dengan tujuan untuk mengurangi redundansi yang terdapat dalam data sehingga dapat disimpan atau ditransmisikan secara efisien [4]. Maka dapat diartikan kompresi data merupakan sebuah proses untuk melakukan pemampatan atau memperkecil jumlah ukuran data dari ukuran asli yang sebenarnya sehingga tidak membutuhkan banyak ruang penyimpanan data.

Tujuan kompresi data adalah untuk merepresentasikan suatu data digital berupa teks, gambar, suara dan lain-lainnya dengan sesedikit mungkin *bit*, namun tetapkan mempertahankan kebutuhan minimum untuk membentuk kembali data aslinya [6]. Teknik kompresi data dibagi ke dalam dua bagian, seperti pada Gambar 1 berikut:



Gambar 1 Teknik Kompresi data

Pada gambar diatas, maka dapat dijelaskan teknik kompresi data sebagai berikut [2]:

a. Lossy Compression

Teknik kompresi data dimana rasio kompresi data antara data asli dengan data setelah di kompresi cukup tinggi. Teknik ini memiliki kelemahan dimana data yang telah di kompresi akan hilang saat akan dilakukan proses dekompresi. Teknik ini biasanya digunakan untuk data berupa audio, video, grafik dan lain-lain.

b. Lossless Compression

Teknik kompresi data dimana hampir sama dengan *Lossy Compression*, hanya saja kelebihan nya adalah data yang telah dikompresi tidak akan hilang ketika akan dilakukan proses dekompresi. Teknik ini biasanya digunakan untuk data berupa teks, zip, rar dan lain-lain.

2.2 Algoritma Shannon Fano

Algoritma *Shannon-Fano* pada dasarnya adalah metode yang mengganti setiap simbol yang ada pada data menjadi kode *biner* yang panjangnya ditentukan berdasarkan pada probabilitas simbol [7]. Algoritma ini disebut sebagai teknik kompresi data paling awal yang ditemukan oleh Claude Shannon dan Robert Fano pada tahun 1949 di MIT [8]. Algoritma *Shannon-Fano* menggunakan pendekatan *top down* dalam penyusunan *binary tree*, sehingga sangat efisien untuk mengkompresi data teks yang berukuran besar [2].

Secara garis besar, algoritma *Shannon-Fano* memiliki ketentuan sebagai berikut [3]:

- a. Setiap simbol yang berbeda memiliki kode yang berbeda.
- b. Simbol dengan probabilitas yang kecil memiliki kode panjang *bit* yang lebih panjang dan simbol dengan probabilitas yang lebih besar memiliki panjang *bit* yang lebih pendek.
- c. Meskipun kode yang dihasilkan memiliki panjang *bit* yang berbeda dengan kode pada karakter asli, tetapi dapat dikodekan secara unik.

Maka adapun langkah-langkah dalam kompresi data pada algoritma *Shannon-Fano* dapat dilihat sebagai berikut [8]:

- a. Urutkan simbol sesuai dengan karakter tabel ASCII serta bangun tabel daftar kemunculan simbol dari data (Probabilitas simbol)
- b. Urutkan tabel kemunculan simbol sesuai dengan frekuensi kemunculan dari frekuensi kemunculan paling tinggi ke paling terendah.
- c. Bagilah tabel kemunculan simbol menjadi dua bagian dengan pembagian berdasarkan pada total frekuensi bagian (bagian atas) sedekat mungkin dengan jumlah frekuensi dengan bagian yang lain (bagian bawah). Hal ini dapat dilakukan dengan membentuk pohon *biner*.
- d. Tetapkan bagian atas dengan digit *biner* 0 dan bagian bawah dengan digit *biner* 1
- e. Terapkan langkah 3 dan 4 secara berulang-ulang untuk masing-masing dua bagian. Bagilah kelompok dan tambahkan *bit* ke kode sehingga setiap simbol memperoleh daun yang sesuai dengan pohon *biner*.

2.3 Proposal

Proposal menurut KBBI adalah rencana yang dituangkan dalam bentuk rancangan kerja, perencanaan secara sistematis, matang dan teliti yang dibuat oleh peneliti sebelum melaksanakan penelitian baik penelitian berupa penelitian lapangan maupun penelitian perpustakaan [9]. Proposal skripsi dapat diartikan suatu bentuk rancangan penelitian atau usulan penelitian yang disusun oleh seorang mahasiswa tentang suatu topik masalah demi mencapai penyelesaian dari masalah tersebut. Proposal penelitian dilakukan guna untuk sebuah pengajuan sebelum memulai langkah dalam membuat tugas akhir atau skripsi. Secara umum, proposal dibedakan menjadi 4 jenis yaitu [9]:

- a. **Proposal Bisnis**, merupakan proposal yang berkaitan dengan dunia usaha baik itu perseorangan maupun kelompok. Contohnya proposal pendirian usaha, proposal kerjasama usaha.
- b. **Proposal Proyek**, merupakan proposal yang berkaitan pada serangkaian rencana bisnis atau komersial seperti proposal proyek pembangunan.
- c. **Proposal Penelitian**, merupakan proposal yang berkaitan dengan dunia akademis seperti proposal skripsi, tesis dan disertasi.
- d. **Proposal Kegiatan**, merupakan proposal yang berkaitan dengan pengajuan rencana sebuah kegiatan baik itu bersifat individu maupun kelompok seperti proposal kegiatan pementasan seni.

Proposal dibuat dengan tujuan untuk memperoleh dana ataupun untuk mencapai tujuan akhir seperti gelar kesarjanaaan.

3 Analisa Dan Pembahasan

Sebelum dilakukan implementasi kedalam Sistem, algoritma *Shannon-Fano* perlu dilakukan analisa untuk mengetahui bagaimana proses kompresi data yang dilakukan oleh algoritma. Pada pembahasan berikut data yang akan dilakukan analisa adalah berupa teks. Adapun contoh data kasus yang akan dilakukan analisa sebagai berikut:

“ PENELITIAN INI DILAKUKAN UNTUK MEMPEROLEH GELAR SARJANA ILMU KOMPUTER ”

Dari data teks berikut, selanjutnya dilakukan proses kompresi algoritma *Shannon-Fano*. Berdasarkan langkah-langkah pada algoritma *Shannon-Fano*, maka dapat dijabarkan proses analisis sebagai berikut:

- a. Mengurutkan karakter sesuai dengan tabel ASCII dan menghitung frekuensi kemunculan. Adapun pengurutan berdasarkan karakter sesuai dengan tabel ASCII dan frekuensi kemunculan karakter dapat dilihat pada Tabel 1 berikut:

Tabel 1 Kode Karakter Dalam ASCII Dan Frekuensi Kemunculan Karakter.

Karakter	Binary	Ukuran	Hexadecimal	Frekuensi Kemunculan
P	01010000	8 Bit	50	3
E	01000101	8 Bit	45	7
N	01001110	8 Bit	4E	6
L	01001100	8 Bit	4C	5
I	01001001	8 Bit	49	6
T	01010100	8 Bit	54	3
A	01000001	8 Bit	41	7
SPASI	00100000	8 Bit	20	8
D	01000100	8 Bit	44	1
K	01001011	8 Bit	4B	4
U	01010101	8 Bit	55	5
M	01001101	8 Bit	4D	4
R	01010010	8 Bit	52	4
O	01001111	8 Bit	4F	2
H	01001000	8 Bit	48	1
G	01000111	8 Bit	47	1
S	01010011	8 Bit	53	1
J	01001010	8 Bit	4A	1
Total Karakter				69

Dari pengkodean ke dalam ASCII, maka dapat diketahui besar ukuran memori yang dibutuhkan untuk menyimpan data kasus tersebut. Untuk mengetahui ukuran memori yang dibutuhkan dari data kasus tersebut adalah sebagai berikut:

$$\text{Ukuran Memori Karakter} = \text{Banyak Karakter} \times \text{Ukuran Bit (1)}$$

Maka dari rumus diatas peroleh ukuran memori dari data kasus seperti pada Tabel 2 Berikut:

Tabel 2 Ukuran Memori Setiap Karakter.

Karakter	Ukuran	Frekuensi Kemunculan	Ukuran Memori Karakter
P	8 Bit	3	24
E	8 Bit	7	56
N	8 Bit	6	48
L	8 Bit	5	40
I	8 Bit	6	48
T	8 Bit	3	24
A	8 Bit	7	56
SPASI	8 Bit	8	64
D	8 Bit	1	8
K	8 Bit	4	32
U	8 Bit	5	40
M	8 Bit	4	32
R	8 Bit	4	32
O	8 Bit	2	16
H	8 Bit	1	8
G	8 Bit	1	8
S	8 Bit	1	8
J	8 Bit	1	8
Total Ukuran Memori			552 Bit (69 Byte)

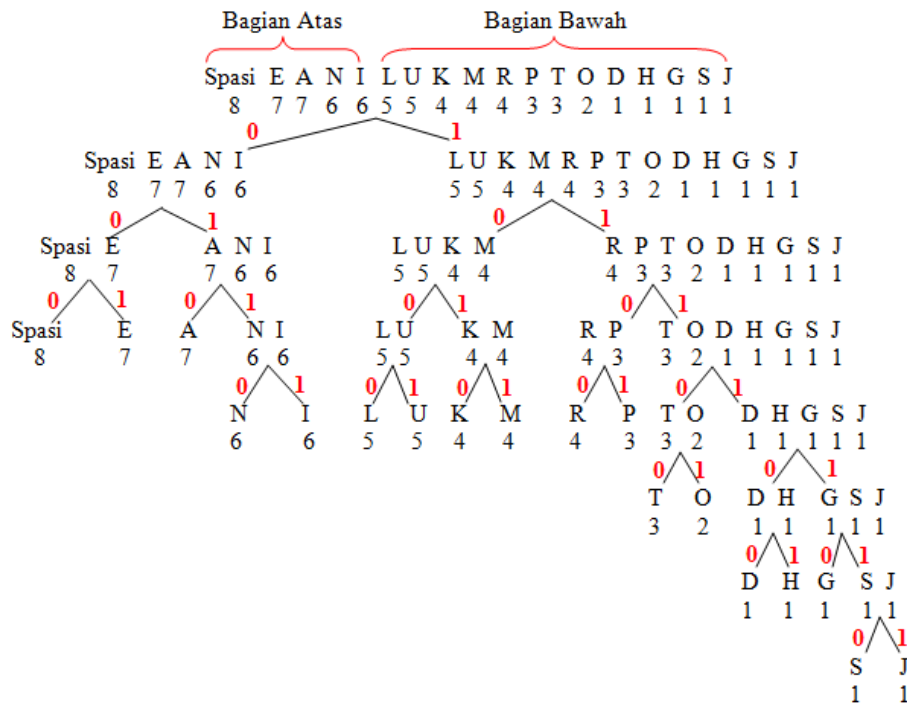
- b. Selanjutnya urutkan karakter berdasarkan frekuensi kemunculan tertinggi sampai yang terendah (*descending*) seperti pada Tabel 2 berikut:

Tabel 3 Pengurutan Frekuensi Kemunculan Karakter.

Karakter	Frekuensi
Spasi	8
E	7
A	7
N	6
I	6
L	5
U	5
K	4
M	4
R	4
P	3
T	3
O	2
D	1
H	1
G	1
S	1
J	1

- c. Proses selanjutnya adalah membagi kemunculan karakter menjadi dua bagian, dimana jumlah frekuensi bagian atas (bagian pertama) sedekat mungkin dengan jumlah frekuensi dengan bagian bawah (bagian kedua). Untuk membagi bagian tersebut, dapat digunakan dengan cara membentuk pohon *biner*. Pada pohon *biner*, bagian atas dimulai dengan

digit *biner* 0 dan bagian bawah dimulai dengan digit 1. Lakukan hal ini secara berulang sampai semua karakter memiliki digit *biner*. Adapun pohon *biner* yang terbentuk adalah seperti Gambar 2 berikut:



Gambar 2 Pohon Biner Proses Algoritma Shannon-Fano

Setelah diperoleh pohon *biner*, maka dapat ditentukan kode *Shannon-Fano* untuk setiap karakter. Kode *Shannon-Fano* diperoleh dari setiap digit *biner* karakter, seperti pada karakter **E**, untuk melihat kode *Shannon-Fano* dapat dilihat digit *biner* dari atas ke bawah yaitu **001**. Adapun kode *Shannon-Fano* untuk setiap karakter dapat dilihat pada Tabel 3 berikut:

Tabel 4 Kode Shannon-Fano Setiap Karakter.

Karakter	Kode Shannon-Fano	Ukuran
Spasi	000	3 bit
E	001	3 bit
A	010	3 bit
N	0110	4 bit
I	0111	4 bit
L	1000	4 bit
U	1001	4 bit
K	1010	4 bit
M	1011	4 bit
R	1100	4 bit
P	1101	4 bit
T	11100	5 bit
O	11101	5 bit
D	111100	6 bit
H	111101	6 bit
G	111110	6 bit

S	1111110	7 bit
J	1111111	7 bit

Untuk menghitung ukuran memori karakter yang telah dikodekan dengan algoritma *Shannon-Fano* dapat mengikuti rumus menghitung ukuran memori karakter di atas. Adapun ukuran memori dari karakter yang telah dikodekan dapat dilihat pada Tabel 5 berikut:

Tabel 5 Ukuran Memori Karakter Dari Kode *Shannon-Fano*.

Karakter	Ukuran	Frekuensi Kemunculan	Ukuran Memori Karakter
Spasi	3 bit	3	9
E	3 bit	7	21
A	3 bit	6	18
N	4 bit	5	20
I	4 bit	6	24
L	4 bit	3	12
U	4 bit	7	28
K	4 bit	8	32
M	4 bit	1	4
R	4 bit	4	16
P	4 bit	5	20
T	5 bit	4	20
O	5 bit	4	20
D	6 bit	2	12
H	6 bit	1	6
G	6 bit	1	6
S	7 bit	1	7
J	7 bit	1	7
Total Ukuran Memori			282 Bit (35,25 Byte)

Setelah dilakukan kompresi dengan algoritma *Shannon-Fano*, maka ukuran memori data kasus dari 69 Byte menjadi 35,25 byte. Selanjutnya dapat dihitung rasio pemampatan dengan menggunakan rumus:

$$Rasio = 100\% - \frac{Ukuran Memori Setelah Kompresi}{Ukuran Memori Sebelum Kompresi} \times 100\% \quad (2)$$

Maka dengan menggunakan rumus diatas, diperoleh rasio kompresi sebagai berikut:

$$100\% - \frac{282 \text{ Bit}}{552 \text{ Bit}} \times 100\% = 48,92 \%$$

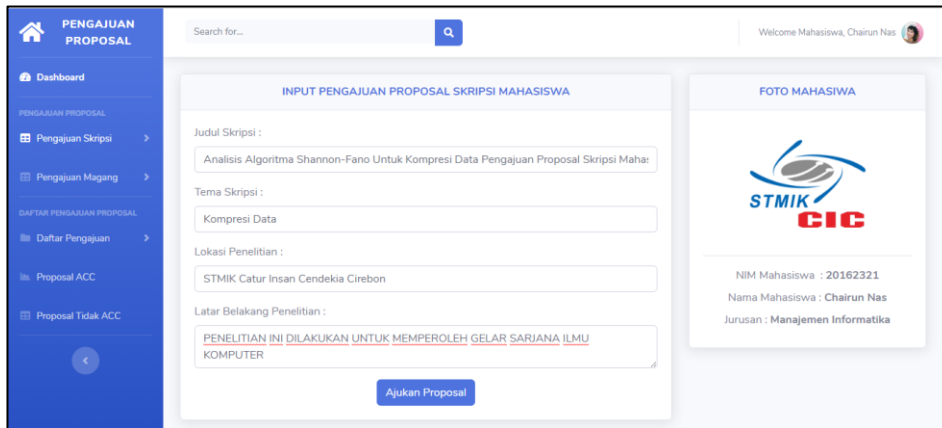
Dari proses kompresi, maka diperoleh bahwa 48,92 % ukuran memori dari ukuran awal berhasil di kompresi.

4 Hasil

Dari proses analisis diatas, selanjutnya algoritma *Shannon-Fano* diimplementasikan kedalam sistem pengajuan proposal skripsi mahasiswa. Adapun sistem pengajuan proposal skripsi mahasiswa dapat dilihat sebagai berikut:

4.1 Halaman Pengajuan Proposal Skripsi

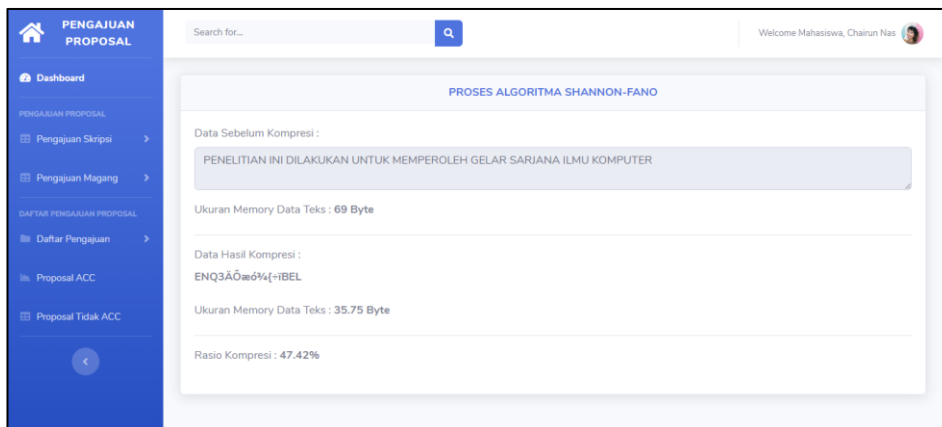
Pada halaman ini, mahasiswa menginputkan data pengajuan proposal skripsi seperti menginputkan judul skripsi, lokasi penelitian dan latar belakang penelitian. Adapun bentuk halaman pengajuan proposal tersebut dapat dilihat pada Gambar 3 berikut:



Gambar 3 Halaman Pengajuan Proposal Skripsi Mahasiswa

4.2 Halaman Proses Kompresi Data Pengajuan Proposal

Setelah data pengajuan proposal diinputkan, maka selanjutnya akan diproses dengan menggunakan algoritma *Shannon-Fano*. Data yang dikompresi dalam sistem ini adalah data input untuk field Latar Belakang Penelitian. Selanjutnya setelah data disimpan maka akan diperoleh kode hasil algoritma *Shannon-Fano* dan ukuran data yang telah dikompresi beserta dengan rasio kompresi. Adapun bentuk halaman proses kompresi data dapat dilihat pada Gambar 4 berikut:



Gambar 3 Halaman Pengajuan Proposal Skripsi Mahasiswa

Maka sistem pengajuan proposal skripsi mahasiswa ini telah dapat digunakan serta dikembangkan untuk mengkompresi data.

5 Kesimpulan Dan Saran

Berdasarkan pembahasan yang telah dijelaskan, maka dapat ditarik kesimpulan dan saran dari penelitian ini sebagai berikut:

5.1 Simpulan

Algoritma *Shannon-Fano* telah dapat melakukan proses kompresi data untuk data pengajuan proposal skripsi mahasiswa. Dari hasil proses algoritma *Shannon-Fano* didapatkan rasio kompresi sebesar 48,92%, dimana sebesar 48,92% data pengajuan berhasil dikompresikan. Maka algoritma *Shannon-Fano* dapat dipilih sebagai algoritma yang cocok untuk melakukan proses kompresi data.

5.2 Saran

Untuk penelitian selanjutnya, perlu adanya pengembangan terhadap proses kompresi data dengan menggabungkan algoritma-algoritma yang ada, serta untuk kedepannya, data yang akan diuji dapat lebih besar lagi.

6 Referensi

- [1] Soemanto, W., Pedoman Teknik Penulisan Skripsi (Karya Ilmiah), Jakarta : Bumi Aksara, 2009.
- [2] Setiawan, S., Teknik Kompresi Data menggunakan Algoritma Shannon Fano, Universitas BSI Bandung, Seminar Nasional Inovasi Dan Teknologi, pp. 66-70, 2012.
- [3] Pane, M.R., Perancangan Aplikasi Kompresi Menggunakan Metode Shannon-Fano Dan Unary Coding Pada File Teks, Majalah Ilmiah INTI, 12(3), pp. 306-311, 2017, ISSN 2339-210X.
- [4] Sutardi, Implementasi Dan Analisis Kinerja Algoritma Shannon-Fano Untuk Kompresi File Text, DINAMIKA Jurnal Ilmiah Teknik Mesin, 6(1), pp. 53-60, 2014, ISSN 2085-8817.
- [5] Risma & Zain, G.Z., Implementasi Kompresi Citra Menggunakan Algoritma Shannon-Fano, Universitas negeri Makasar, Prosiding Seminar Nasional, pp. 985-991, ISBN 978-602-9075-25-7.
- [6] Aji, Y.Y., Darwiyanto, E. & Septiana, G., Analisis Perbandingan Kompresi Dan Dekompresi Menggunakan Algoritma Shannon-Fano 2 Gram Dan Lempel Ziv Welch pada Terjemahan Hadits Shahih Muslim, e-Proceeding of Engineering, pp. 5197-5204, 2016, ISSN 2355-9365.
- [7] Rachesti, D.A., Purboyo, T.W. & Prasasti, A.L, Comparison Of Text Data Compression Using Huffman, Shannon-Fano, Run Length Encoding And Tunstall Methods, International Journal Of Applied Engineering Research, 12(23), pp. 13618-13622, 2017, ISSN 0973-4562.
- [8] Haque, M.J. & Huda, M.N, Study On Data Compression Technique, International Journal Of Computer Applications, 159(5), pp. 6-13, 2017, ISSN 0975-8887.
- [9] Hasyim, M. & Nugrahanto, G., Pelatihan Pembuatan Proposal Kegiatan Pada Remaja Dusun Ngepuh Lor, Desa Banyusidi, Pakis, Magelang, Jawa Tengah, Jurnal Motivasi Dan Kewirausahaan, 3(3), pp. 206-210, 2014, ISSN 2089-3086.