

QAS Untuk Peningkatan Kualitas Pelayanan Bidang Administrasi Pada Kantor Desa Dengan Algoritma *Cosine Similarity*

Jumantri Tua Beruntung Silaen*, Guidio Leonarde Ginting, Ryan Syahputra

Fakultas Ilmu Komputer dan Teknologi Informasi, Prodi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia
Email : *jumantrisilaen1212@gmail.com

Abstrak-Perkembangan dan penggunaan teknologi komputer sekarang sudah semakin canggih. Setiap tahun teknologi komputer semakin berkembang seiring dengan berjalannya waktu. Penggunaan komputer sangat penting untuk memperoleh informasi dalam proses pendaftaran maupun pengolahan data, seperti halnya pada Kantor Desa Silaen. *Question Answering System* dan metode *Cosine Similarity* adalah sistem berbasis komputer untuk meningkatkan kualitas pelayanan dengan sistem tanya jawab yang bertujuan memberikan alternatif-alternatif yang lebih baik. Disini penulis berusaha untuk membuat alternatif pelayanan yang memanfaatkan teknologi komputer dengan tujuan memberikan kemudahan bagi masyarakat maupun peserta didik untuk mempelajari materi pelayanan serta mempermudah untuk mengatur waktu dalam hal pengurusan administrasi sehingga mudah dipahami.

Kata Kunci: Administrasi¹, Cosine Similarity², Natural Language Processing³, Pelayanan⁴, Question Answering System (QAS)⁵

Abstract-The development and use of computer technology is now increasingly sophisticated. Every year computer technology is growing with the passage of time. The use of computers is very important to obtain information in the registration process and data processing, as is the case at the Silaen Village Office. The *Question Answering System* and the *Cosine Similarity* method are computer-based systems to improve service quality with a question and answer system that aims to provide better alternatives. learn service material and make it easier to manage time in terms of administrative management so that it is easy to understand.

Keywords: Administration¹, Cosine Similarity², Natural Language Processing³, Service⁴, Question Answering System (QAS)⁵

1. PENDAHULUAN

Question Answering System (QAS) merupakan teknologi kecerdasan buatan yang dapat diolah dengan berbagai macam bentuk, seperti ChatBot dan berbagai macam metode, salah satunya dengan menggunakan *Natural Language Processing* (NLP) atau bisa juga di kenal dengan *Artificial Matching Markup Language* (AIML) dimana keduanya menggunakan metode yang membandingkan pola-pola tertentu pada *database*[1]. Namun pada saat sekarang ini kenyataannya penyelenggara pelayanan administrasi kependudukan yang dilakukan oleh pemerintah daerah masih belum efisien serta kualitas pelayanan yang belum baik, sehingga masyarakat mengeluh atas pelayanan yang diberikan oleh penyelenggara pelayanan karena tidak sesuai dengan harapan mereka. Terutama pelayanan yang dirasakan oleh masyarakat Desa Silaen, terkhusus dalam hal pengurusan surat-surat atau pelayanan di bidang administrasi di Desa Silaen belum prima, sebagai contoh ketika akan melakukan pengurusan KTP ataupun surat-surat lainnya, masyarakat kerap bingung dengan berkas-berkas yang akan dilengkapi serta antrian yang cukup lama juga terkadang pelayan bidang administrasi di Desa Silaen tersebut cenderung tidak hadir, sehingga membuat masyarakat desa harus kembali lagi di lain hari untuk pengurusan administrasi yang sama. meskipun tuntutan tersebut sering tidak sesuai dengan harapan karena pelayanan yang terjadi selama ini masih berbelit-belit dan melelahkan. Kualitas pelayanan prima yang dimaksud yaitu pelayanan yang mendekati pemerintah kepada masyarakat. Hal ini diwujudkan dengan cara mengetahui dan menganalisis berbagai persoalan yang dihadapi oleh masyarakat, untuk kemudian menciptakan strategi pelayanan yang efisien[2].

Berdasarkan penelitian terdahulu yang dilakukan oleh Erdipa Panjaitan dalam jurnal nya yang berjudul "Peranan Pemerintah Desa dalam Meningkatkan Kualitas Pelayanan Administrasi kepada Masyarakat". Hasil penelitian menunjukkan bahwa Kualitas pelayanan di kantor Kepala Desa Aek Korsik Kecamatan Aek Kuo Kabupaten Labuhanbatu Utara dapat dinilai dari lima dimensi yaitu *Tangible*, *Reliability*, *Responsiviness*, *Assurance*, dan *Emphaty*[3].

K Purba dalam jurnal nya yang berjudul "Kualitas Pelayanan Administrasi Kependudukan Pada UPT DISDUKCAPIL Kecamatan Tampar Kota Pekanbaru". Kualitas pelayanan administrasi kependudukan pada UPT Disdukcapil kecamatan Sail masih belum optimal. Berdasarkan hasil penelitian masih ditemukan dimana unsur-unsur kualitas pelayanan publik belum berjalan sebagaimana mestinya. Unsur *reliability & responsibility* dan *tangible aspect* merupakan unsur yang harus diperhatikan dalam meningkatkan kualitas pelayanan administrasi kependudukan[4].

Rama Nur Jazuli Setiawan dalam jurnal nya yang berjudul "sistem penjawab pertanyaan publik di kecamatan kaliwates jember menggunakan metode *cosine similartiy*". Kesimpulan dari Sistem Penjawab Pertanyaan Publik dengan menggunakan metode Metode *Cosine Similarity* dalam Sistem Penjawab Pertanyaan Publik ini memiliki tingkat akurasi sebesar 85%. Metode *Cosine Similarity* sesuai diterapkan dalam Sistem Penjawab Pertanyaan Publik karena tingkat kesalahannya relatif rendah[4], [5]

Untuk menyelesaikan masalah tersebut dalam penelitian ini, maka penulis melakukan perancangan suatu aplikasi berbasis komputer yaitu *Question Answering System* Peningkatan Pelayanan Administrasi dengan menggunakan bahasa pemrograman *Pemograman Web* dengan menerapkan metode *Cosine Similarity*. *Question Answering System* Peningkatan Pelayanan Administrasi yang dirancang penulis pada penelitian ini dapat memberikan hasil yang lebih baik terhadap hasil dari proses yang dilakukan dalam peningkatan pelayanan bidang administrasi di Kantor desa dan berdasarkan perangkingan nilai dari tiap-tiap keadaan yang ditemui oleh masyarakat Desa Silaen.

2. METODOLOGI PENELITIAN

2.1 Question Answering System

Sistem yang dikenal sebagai *Question Answering System* ini mengijinkan user untuk menginputkan pertanyaan dalam bahasa natural, yaitu bahasa yang digunakan dalam percakapan sehari-hari, dan memperoleh jawaban dengan cepat serta ringkas, atau bahkan disertai dengan kalimat yang cukup untuk mendukung kebenaran dari jawaban tersebut. *Search engine* yang ada saat ini dapat mengembalikan daftar dokumen yang telah diurutkan berdasarkan tingkat relevansi dari dokumen tersebut terhadap *query user*, tetapi tidak memberikan jawaban kepada *user*[6].

2.2 Natural Language Processing

Natural language processing (NLP) adalah cabang dari kecerdasan buatan yang berhubungan dengan interaksi antara komputer dan manusia menggunakan bahasa alami. Menurut *Textmetrics*, NLP digunakan untuk mengukur sentimen dan menentukan bagian mana dari bahasa manusia yang penting[7].

2.3 Text Mining

Text mining digunakan untuk mengubah kumpulan teks menjadi numerik sehingga dapat dikomputasikan. *Text mining* adalah salah satu bidang khusus. Sesuai dengan buku *The Next Mining Handbook*. *Text Mining* dapat didefinisikan sebagai suatu proses menggali informasi dimana seorang *user* berinteraksi dengan sekumpulan dokumen dengan menggunakan *tools* analisis yang merupakan komponen-komponen, yang salah satunya adalah peningkatan dokumen[8]–[10].

Tabel 1. Data pertanyaan dan jawaban

Kode	Pertanyaan	Jawaban
D1	1. Min dikantor desa sudah bisa mengurus KTP? 2. Min sudah bisa buat KTP belum? 3. Hai min, sudah bisa apa belum ngurus KTP?	Ya, sudah
D2	1. Apa aja sih berkas untuk ngurus KTP min? 2. Min untuk ngurus KTP, perlu gak sih surat pengantar? 3. Emang surat apa aja sih buat ngurus KTP min?	Photocopy KK, surat pengantar dari Kepala Dusun
D3	1. Untuk pengurusan KTP untuk anak Sekolah sudah boleh dilakukan dan waktunya kapan ya? 2. Anak SMA Udah bisa gak sih buat KTP? 3. Min, kelas satu SMA udah bisa ya buat KTP?	Sudah, Anak berusia 17 tahun ke atas dan Waktu pengurusan KTP dimulai setiap senin-jumat sesuai jam kerja kantor
D4	1. Untuk lokasi rekaman pembuatan KTP dimana ya? 2. Lokasi ngurus KTP dimana ya min? 3. Tempat ngurus KTP emang bisa ya di Kantor camat?	Untuk rekaman KTP bisa dilakukan di kantor camat kecamatan atau langsung ke Disdukcapil

Berikut ini merupakan tabel data pertanyaan, dapat dilihat pada tabel 2 dibawah ini:

Tabel 2. Data Pertanyaan

Kode	Pertanyaan
D1	1. min dikantor desa sudah bisa mengurus ktp ? 2. min sudah bisa buat ktp belum ? 3. hai min, sudah bisa apa belum ngurus ktp ?
D2	1. apa aja sih berkas untuk ngurus ktp min ? 2. min untuk ngurus ktp, perlu gak sih surat pengantar ? 3. emang surat apa aja sih buat ngurus ktp min ?
D3	1. untuk pengurusan ktp untuk anak sekolah sudah boleh dilakukan dan waktunya kapan ya ? 2. anak sma udah bisa gak sih buat ktp ? 3. min, kelas satu sma udah bisa ya buat ktp ?
D4	1. untuk lokasi rekaman pembuatan ktp dimana ya ? 2. lokasi ngurus ktp dimana ya min ? 3. tempat ngurus ktp emang bisa ya di kantor camat ?

Berikut ini merupakan tabel hasil tokenzing, dapat dilihat pada tabel 2 dibawah ini:

Tabel 3. Hasil *tokenzing*

Teks	Hasil Tokenizing
1. min dikantor desa sudah bisa mengurus ktp	[min] [di] [kantor] [desa] [sudah] [bisa] [mengurus] [ktp]
2. min sudah bisa buat ktp belum	
3. hai min, sudah bisa apa belum ngurus ktp	[min] [sudah] [bisa] [buat] [ktp] [belum]
	[hai] [min] [sudah] [bisa] [apa] [belum] [ngurus] [ktp]
1. apa aja sih berkas untuk ngurus ktp min	[apa] [aja] [sih] [berkas] [untuk] [ngurus] [ktp] [min]
2. min untuk ngurus ktp, perlu gak sih surat pengantar	[min] [untuk] [ngurus] [ktp] [perlu] [gak] [sih] [surat] [pengantar]
3. emang surat apa aja sih buat ngurus ktp min	[emang] [surat] [apa] [aja] [sih] [buat] [ngurus] [ktp] [min]
1. untuk pengurusan ktp untuk anak sekolah sudah boleh dilakukan dan waktunya kapan ya	[untuk] [pengurusan] [ktp] [untuk] [anak] [sekolah] [sudah] [boleh] [dilakukan] [dan] [waktunya] [kapan] [ya]
2. anak sma udah bisa gak sih buat ktp	
3. min, kelas satu sma udah bisa ya buat ktp	[anak] [sma] [udah] [bisa] [gak] [sih] [buat] [ktp]
	[min] [kelas] [satu] [sma] [udah] [bisa] [ya] [buat] [ktp]
1. untuk lokasi rekaman pembuatan ktp dimana ya	[untuk] [lokasi] [rekaman] [pembuatan] [ktp] [dimana] [ya]
2. lokasi ngurus ktp dimana ya min	
3. tempat ngurus ktp emang bisa ya di kantor camat	[lokasi] [ngurus] [ktp] [dimana] [ya] [min]
	[tempat] [ngurus] [ktp] [emang] [bisa] [ya] [di] [kantor] [camat]

Berikut ini merupakan tabel hasil filtering, dapat dilihat pada tabel 2 dibawah ini:

Tabel 3. Hasil *filtering*

No	Hasil Tokenizing	Hasil Filtering
1	[min] [di] [kantor] [desa] [sudah] [bisa] [mengurus] [ktp] [min] [sudah] [bisa] [buat] [ktp] [belum] [hai] [min] [sudah] [bisa] [apa] [belum] [ngurus] [ktp]	[min] [kantor] [desa] [mengurus] [ktp] [min] [buat] [ktp] [belum] [hai] [min] [apa] [belum] [ngurus] [ktp]
2	[apa] [aja] [sih] [berkas] [untuk] [ngurus] [ktp] [min] [min] [untuk] [ngurus] [ktp] [perlu] [gak] [sih] [surat] [pengantar] [emang] [surat] [apa] [aja] [sih] [buat] [ngurus] [ktp] [min]	[berkas] [ngurus] [ktp] [min] [min] [ngurus] [ktp] [perlu] [gak] [pengantar] [emang] [buat] [ngurus] [ktp] [min]
3	[untuk] [pengurusan] [ktp] [untuk] [anak] [sekolah] [sudah] [boleh] [dilakukan] [dan] [waktunya] [kapan] [ya] [anak] [sma] [udah] [bisa] [gak] [sih] [buat] [ktp] [min] [kelas] [satu] [sma] [udah] [bisa] [ya] [buat] [ktp]	[pengurusan] [ktp] [[anak] [sekolah] [sudah] [boleh] [dilakukan] [waktunya] [kapan] [sma] [udah] [buat] [ktp] [min] [kelas] [satu] [sma] [udah] [buat] [ktp]
4	[untuk] [lokasi] [rekaman] [pembuatan] [ktp] [dimana] [ya] [lokasi] [ngurus] [ktp] [dimana] [ya] [min] [tempat] [ngurus] [ktp] [emang] [bisa] [ya] [di] [kantor] [camat]	[lokasi] [rekaman] [pembuatan] [ktp] [dimana] [lokasi] [ngurus] [ktp] [dimana] [min] [tempat] [ngurus] [ktp] [emang] [kantor] [camat]

Berikut ini merupakan tabel hasil case streaming, dapat dilihat pada tabel 2 dibawah ini:

Tabel 4. Hasil *Streaming*

No	Hasil Filtering	Hasil Steaming
1	[min] [kantor] [desa] [mengurus] [ktp] [min] [buat] [ktp] [belum] [hai] [min] [apa] [belum] [ngurus] [ktp]	[min] [kantor] [desa] [urus] [ktp] [min] [buat] [ktp] [belum] [hai] [min] [apa] [belum] [urus] [ktp]
2	[berkas] [ngurus] [ktp] [min] [min] [ngurus] [ktp] [perlu] [gak] [pengantar] [emang] [buat] [ngurus] [ktp] [min]	[berkas] [urus] [ktp] [min] [min] [urus] [ktp] [perlu] [gak] [antar] [emang] [buat] [urus] [ktp] [min]
3	[pengurusan] [ktp] [[anak] [sekolah] [sudah] [boleh] [dilakukan] [waktunya] [kapan]	[urus] [ktp] [[anak] [sekolah] [udah] [boleh] [lakukan] [waktu] [kapan]

No	Hasil <i>Filtering</i>	Hasil <i>Steaming</i>
	[sma] [udah] [buat] [ktp] [min] [kelas] [satu] [sma] [udah] [buat] [ktp]	[sma] [udah] [buat] [ktp] [min] [kelas] [satu] [sma] [udah] [buat] [ktp]
4	[lokasi] [rekaman] [pembuatan] [ktp] [dimana] [lokasi] [ngurus] [ktp] [dimana] [min] [tempat] [ngurus] [ktp] [emang] [kantor] [camat]	[lokasi] [rekam] [buat] [ktp] [mana] [lokasi] [urus] [ktp] [mana] [min] [tempat] [urus] [ktp] [memang] [kantor] [camat]

2.4 TF-IDF

TF-IDF (*Term Frequency Invers Document Frequency*) merupakan metode yang digunakan untuk menentukan seberapa jauh keterhubungan kata (*term*) terhadap dokumen dengan memberikan bobot setiap kata. Metode TF-IDF ini menggabungkan dua konsep yaitu frekuensi kemunculan sebuah kata di dalam sebuah dokumen dan inverse frekuensi dokumen yang menggunakan kata tersebut. Dalam perhitungan bobot menggunakan TF-IDF, dihitung terlebih dahulu nilai TF perkata dengan bobot masing-masing kata adalah 1 [11], [12]

Tabel 6. Penerapan TF-IDF

Pertanyaan sebelum di <i>text mining</i>	Pertanyaan setelah di <i>text mining</i>
Dimanakah Lokasi Untuk Pembuatan KTP dilaksanakan?	[mana] [lokasi] [buat] [ktp] [laksana]

Berikut ini merupakan gambar dari hasil TF-IDF yang dapat dilihat pada gambar 1 dibawah ini:

Term Q	TF	IDF	TF*DF
mana	1	1	1
lokasi	1	1	1
buat	1	1	1
ktp	1	1	1
laksana	1	1	1

Gambar 1. Hasil TF-IDF

3. HASIL DAN PEMBAHASAN

3.1 Analisa Masalah

Dalam penelitian ini, akan dilakukan analisa dan perancangan perangkat lunak sistem tanya jawab dengan menggunakan algoritma *cosine similarity*. *Question Answering System* Peningkatan Pelayanan Administrasi yang dirancang penulis pada penelitian ini dapat memberikan hasil yang lebih baik terhadap hasil dari proses yang dilakukan dalam peningkatan pelayanan bidang administrasi di Kantor desa dan berdasarkan perangkingan nilai dari tiap-tiap keadaan yang ditemui oleh masyarakat Desa Silaen.

3.2 Penerapan TF-IDF

Kesesuaian jawaban yang akan disampaikan oleh sistem sangat bergantung kepada data yang digunakan, pada penelitian tersebut sebagai contoh penulis menggunakan beberapa pertanyaan yang akan dijadikan sebagai acuan dalam menjawab pertanyaan yang akan disampaikan oleh masyarakat.

Term P	D1			D2			D3			D4			TF	IDF
	1	2	3	1	2	3	1	2	3	1	2	3		
min	1	1	1	1	1	1	0	0	1	0	1	0	8	1.176091
kantor	1	0	0	0	0	0	0	0	0	0	0	1	2	1.778151
desa	1	0	0	0	0	0	0	0	0	0	0	0	1	2.079181
urus	1	0	1	1	1	1	1	0	1	0	1	1	8	1.176091
ktp	1	1	1	1	1	1	1	1	1	1	1	1	12	1
buat	0	1	0	0	0	1	0	1	1	1	0	0	5	1.380211
beham	0	1	1	0	0	0	0	0	0	0	0	0	2	1.778151
hai	0	0	1	0	0	0	0	0	0	0	0	0	1	2.079181
apa	0	0	1	0	0	0	0	0	0	0	0	0	1	2.079181
berkas	0	0	0	1	0	0	0	0	0	0	0	0	1	2.079181
jaik	0	0	0	0	1	0	0	0	0	0	0	0	1	2.079181
antar	0	0	0	0	1	0	0	0	0	0	0	0	1	2.079181
evang	0	0	0	0	0	1	0	0	0	0	0	0	1	2.079181
anak	0	0	0	0	0	0	1	0	0	0	0	0	1	2.079181
sekolah	0	0	0	0	0	0	1	0	0	0	0	0	1	2.079181
bekel	0	0	0	0	0	0	1	0	0	0	0	0	1	2.079181
lakukan	0	0	0	0	0	0	1	0	0	0	0	0	1	2.079181
waktu	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Term P	D1			D2			D3			D4			TF	IDF
	1	2	3	1	2	3	1	2	3	1	2	3		
kaplan	0	0	0	0	0	0	1	0	0	0	0	0	1	2.079181
sma	0	0	0	0	0	0	0	1	1	1	0	0	2	1.778151
udah	0	0	0	0	0	0	1	1	1	0	0	0	3	1.60206
kelas	0	0	0	0	0	0	0	0	1	0	0	0	1	2.079181
satu	0	0	0	0	0	0	0	0	1	0	0	0	1	2.079181
lokasi	0	0	0	0	0	0	0	0	0	1	1	0	2	1.778151
rekam	0	0	0	0	0	0	0	0	0	1	0	0	1	2.079181
mana	0	0	0	0	0	0	0	0	0	1	1	0	2	1.778151
tempat	0	0	0	0	0	0	0	0	0	0	1	1	1	2.079181
memang	0	0	0	0	0	0	0	0	0	0	0	1	1	2.079181
camat	0	0	0	0	0	0	0	0	0	0	0	1	1	2.079181
perlu	0	0	0	0	1	0	0	0	0	0	0	0	1	2.079181

Gambar 2. Perhitungan TF-IDF

Berikut ini merupakan gambar dari lanjutan perhitungan TF-IDF, dapat dilihat pada gambar 3 dibawah ini:

$$\begin{aligned} \text{CosSim}(Q, D1_1) &= \frac{\sum_{i=1}^n (td_{ij} \times tq_{ik})}{\sqrt{\sum_{i=1}^n td_{ij}^2 \times \sum_{i=1}^n tq_{ik}^2}} \\ &= \frac{1}{223607*1} = 0,44721 \end{aligned}$$

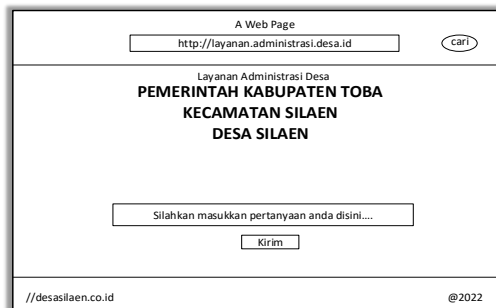
Perhitungan similaritis antara Q (pertanyaan yang di analisa) dengan D1₁ didapatkan hasil 0,44721. Begitu selanjutnya sampai didapatkan hasil pada perhitungan ke 12.

Tabel 7. Hasil dari penerapan *Cosine Similarity*

Q	1	2	3
D1	0,44721	0,62454	0,44721
D2	0,44721	0,44721	0,62454
D3	0,44721	0,62454	0,62454
D4	0,87393	0,75295	0,44721

3.4 Implementasi

Tahap implementasi merupakan tahap penciptaan perangkat lunak, tahap kelanjutan dari kelanjutan dari kegiatan perancangan sistem. Tahap ini merupakan tahap dimana sistem siap untuk dioperasikan, yang terdiri dari penjelasan mengenai kebutuhan sistem, dan implementasi antar muka.



Gambar 6. Tampilan Interface QAS

Dibawah ini dapat dilihat tampilan QAS pada gambar 7 dibawah ini:



Gambar 7. Tampilan QAS

4. KESIMPULAN

Berdasarkan dari penelitian yang telah dilakukan, maka hasil akhir dari penelitian tersebut dapat diambil kesimpulan dari pembahasan sebelumnya. Algoritma TF-IDF mampu mencocokkan pertanyaan dengan data yang sudah tersimpan pada database yang mana dapat menghasilkan jawaban yang sesuai. Dengan menrapkan pendekatan *Natural Language Processing* pada aplikasi *Question Answering System* percakapan yang terjadi layaknya antara manusia dengan manusia. Dengan membuat aplikasi *Question Answering System* ini dapat menggantikan manusia sehingga dapat meningkatkan kualitas Pelayanan di Kantor Desa Silaen terutama di bidang administrasi.

REFERENCES

- [1] R. R. R. Ferbiansyah, "Question Answering System Berbahasa Indonesia Menggunakan Metode Pattern Based Approach (Studi Kasus Bahan Ajar Mata Kuliah Sistem Operasi Prodi S1 Sistem Informasi Universitas Airlangga)," pp. 1–4, 2015.
- [2] K. Purba, "Kualitas Pelayanan Administrasi Kependudukan Pada Upt Disdukcapil Kecamatan Tampan Kota Pekanbaru," J.

- Adm. Publik*, vol. 11, no. 1, pp. 31–55, 2020, doi: 10.31506/jap.v11i1.7127.
- [3] E. Panjaitan, R. Dewi, and N. Angelia, “Peranan Pemerintah Desa dalam Meningkatkan Kualitas Pelayanan Administrasi kepada Masyarakat,” *Perspektif*, vol. 8, no. 1, p. 32, 2019, doi: 10.31289/perspektif.v8i1.2543.
- [4] M. Metode *et al.*, “untuk memenuhi kebutuhannya dalam berbagai bidang.”
- [5] D. Liu, X. Chen, and D. Peng, “Some cosine similarity measures and distance measures between q-rung orthopair fuzzy sets,” *Int. J. Intell. Syst.*, vol. 34, no. 7, pp. 1572–1587, 2019.
- [6] F. Azwary, F. Indriani, and D. T. Nugrahadi, “Question Answering System Berbasis Artificial Intelligence Markup Language,” *Kumpul. J. Ilmu Komput.*, vol. 04, no. 01, pp. 48–60, 2016.
- [7] A. N. Rohman, E. Utami, and S. Raharjo, “Deteksi Kondisi Emosi pada Media Sosial Menggunakan Pendekatan Leksikon dan Natural Language Processing,” *Eksplora Inform.*, vol. 9, no. 1, pp. 70–76, 2019, doi: 10.30864/eksplora.v9i1.277.
- [8] H. Sari, G. L. Ginting, and T. Zebua, “Penerapan Algoritma Text Mining dan TF-IDF Untuk Pengelompokan Topik Skripsi Pada Aplikasi Repository STMIK Budi Darma,” vol. 2, no. 7, pp. 414–432, 2021.
- [9] D. Antons, E. Grünwald, P. Cichy, and T. O. Salge, “The application of text mining methods in innovation research: current state, evolution patterns, and development priorities,” *R&D Manag.*, vol. 50, no. 3, pp. 329–351, 2020.
- [10] T. Jo, “Text mining,” *Stud. Big Data*, 2019.
- [11] B. Herwijayanti, D. E. Ratnawati, and L. Muflikhah, “Klasifikasi Berita Online dengan menggunakan Pembobotan TF-IDF dan Cosine Similarity,” *Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 1, pp. 306–312, 2018.
- [12] S.-W. Kim and J.-M. Gil, “Research paper classification systems based on TF-IDF and LDA schemes,” *Human-centric Comput. Inf. Sci.*, vol. 9, pp. 1–21, 2019.