

Pembangunan Aplikasi Chatbot Informasi Akademik berbasis *Cosine Similarity* dan *Library Sastrawi Stemmer* (Studi Kasus: Teknik Informatika IT PLN)

Rosida Nur Aziza¹; Tiara Sukma Ardanti²; Efy Yosrita³, Rahma Farah Ningrum⁴

^{1,2,3,4} Teknik Informatika, IT PLN
rosida@itpln.ac.id

ABSTRACT

The fast delivery of accurate academic information for students is one of the problems faced by the Program Studi Teknik Informatika IT PLN. Information received by students from social media belongs to the department, as well as from fellow students is often inaccurate and unclear, or requires further explanation, while questions asked to the department office via chat are not always answered on time. To overcome this problem, a chatbot application was built to produce automatic answers to incoming questions. On the other hand, chat applications, such as WhatsApp, LINE, and Telegram, have provided Application Programming Interface (API) facilities for sending and receiving messages that allow the development of applications to be in the form of chats. Therefore, the Chatbot proposed in this study is implemented on Telegram app and was developed using the Stemmer Literature as a library for the message preprocessing, the TF-IDF method for word weighting, and the Cosine Similarity method for examining similarities between objects. Based on the results of the tests carried out...

Keywords: chatbot, academic information, cosine similarity

ABSTRAK

Penyediaan informasi akademik yang tepat dan cepat bagi mahasiswa yang membutuhkan menjadi salah satu masalah yang dihadapi oleh Prodi Teknik Informatika IT PLN. Informasi yang diterima oleh mahasiswa dari media sosial milik prodi, maupun dari kawan sesama mahasiswa seringkali kurang akurat dan tidak jelas, atau memerlukan penjelasan yang lebih lanjut, sedangkan pertanyaan yang diajukan ke prodi melalui chat tidak selalu dijawab tepat waktu. Untuk mengatasi permasalahan tersebut, maka dibangunlah suatu aplikasi chatbot yang bertugas memberikan jawaban secara otomatis terhadap pertanyaan-pertanyaan yang masuk. Di sisi lain, aplikasi chat, seperti WhatsApp, LINE, dan Telegram, telah menyediakan fasilitas Application Programming Interface (API) untuk pengiriman dan penerimaan pesan yang memungkinkan pengembangan aplikasi untuk dalam bentuk chat. Oleh karena itu, Chatbot yang dibuat pada penelitian ini beroperasi pada aplikasi chat Telegram dan dikembangkan menggunakan library Sastrawi Stemmer sebagai pustaka untuk prapemrosesan pesan, metode TF-IDF untuk pembobotan kata, dan metode Cosine Similarity untuk menghitung kemiripan antar objek. Berdasarkan hasil pengujian yang dilakukan...

Kata kunci: chatbot, informasi akademik, cosine similarity

1. PENDAHULUAN

Cara penyebaran informasi akademik di Program Studi (Prodi) S1 Teknik Informatika IT PLN adalah melalui akun media sosial resmi dari prodi dan forum komunikasi yang beranggotakan perwakilan beberapa mahasiswa dari berbagai angkatan. Apabila mahasiswa perlu penjelasan lebih lanjut dari informasi atau pengumuman yang diberikan, mereka akan mengirimkan pesan ke media

sosial tersebut atau mengajukan pertanyaan langsung via aplikasi chat ke Prodi. Mahasiswa acapkali harus menunggu dalam hitungan jam untuk mendapatkan jawaban dari Prodi. Di sisi lain, Prodi juga merasa kesulitan menjawab semua pertanyaan terkait akademik yang diajukan oleh mahasiswa secara cepat. Berdasarkan wawancara yang dilakukan dengan pihak Prodi, dapat disimpulkan bahwa Prodi memerlukan suatu sistem yang dapat memberikan informasi akademik, sekaligus menjawab pertanyaan-pertanyaan dari mahasiswa dengan otomatis menggunakan *Chatbot*. Hasil kuesioner yang disebarkan ke mahasiswa juga menunjukkan bahwa 66% mahasiswa, dari 138 responden, menyatakan sangat setuju dan setuju sekali atas rencana pembuatan aplikasi *chatbot* informasi akademik ini.

Chatbot adalah aplikasi yang memberikan layanan percakapan ke pengguna dengan cara efisien dan menerapkan metode kecerdasan buatan (*Artificial Intelligence*, AI) [1]. Aplikasi ini dibangun untuk membantu pengguna dalam memberikan jawaban atas pertanyaan-pertanyaan yang masuk secara otomatis dan *real time*, serta dapat digunakan di berbagai bidang, seperti bidang pemasaran [2], pendidikan [3], hiburan [4], layanan kesehatan [5], dan sektor layanan publik [6]. Di bidang Pendidikan, *chatbot* dapat digunakan untuk membantu peserta didik dalam proses belajar mengajar, seperti pada penelitian [7] yang mengembangkan *chatbot* untuk membantu mahasiswa dalam pelaksanaan kegiatan di laboratorium. Selain itu, teknologi ini juga dapat digunakan untuk mendapatkan informasi akademik di universitas [8].

Salah satu tahapan dalam pengembangan *chatbot* adalah penggunaan metode pencocokan. Metode *cosine similarity* adalah salah satu metode yang dapat dipakai untuk mencocokkan pertanyaan-pertanyaan yang diajukan dengan basis pengetahuan yang dimiliki *chatbot* dengan akurasi yang tinggi dan waktu eksekusi yang cepat [9], [10]. Penggunaan dari metode ini, umumnya, Bersama dengan metode TF-IDF dan memiliki prinsip mencari kesamaan dari dua vektor berdasarkan nilai perkalian sudut *cosine* keduanya [11]. Pada penelitian ini, *Library Sastrawi Stemmer* digunakan pada proses *stemming*, yaitu pencarian kata dasar dalam Bahasa Indonesia. *Library NLP (Natural Language Processing)* ini dipilih karena menawarkan kemudahan dan memiliki API yang sederhana. Penelitian oleh Rosid dkk., membandingkan penggunaan dua *library* untuk *stemmer*, yaitu Porter *Stemmer* dan Sastrawi terhadap 50 sampel teks dalam Bahasa Indonesia. Dari penelitian tersebut, diperoleh kesimpulan bahwa *Library Sastrawi* menghasilkan akurasi lebih baik, yaitu 92%, dan waktu pemrosesan lebih cepat dibandingkan Porter *stemmer* [12].

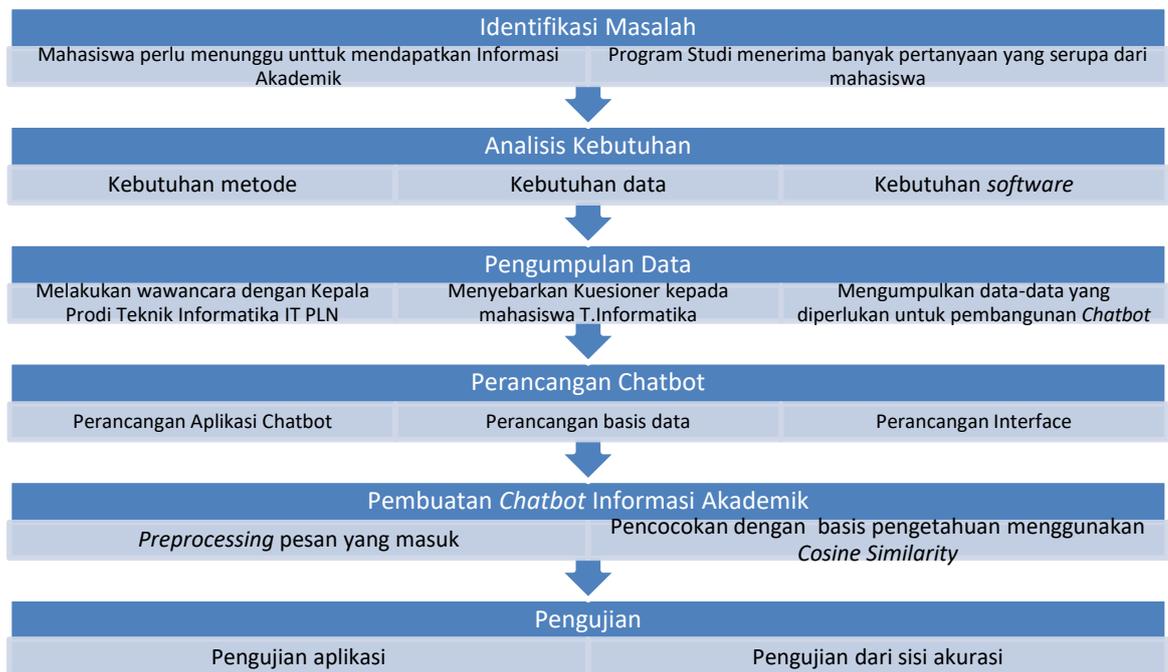
2. METODE/PERANCANGAN PENELITIAN

2.1 Tahapan Penelitian

Gambar 1 menunjukkan tahapan penelitian yang dilakukan untuk pembangunan aplikasi *chatbot* Informasi Akademik yang diimplementasikan menggunakan aplikasi chat Telegram. Aplikasi ini dipilih karena termasuk aplikasi *chat* dengan jumlah pengguna yang terbanyak dan telah dilengkapi dengan fasilitas untuk pembangunan *chatbot* [13].

Pada tahapan Identifikasi Masalah, penulis menemukan masalah yaitu adanya waktu tunggu dari mahasiswa yang mengajukan pertanyaan-pertanyaan terkait informasi akademik kepada pihak Prodi. Dari sisi prodi, banyak mahasiswa yang mengajukan pertanyaan yang serupa dalam kurun waktu tertentu. Pada tahapan Analisis Kebutuhan, penulis mencoba merumuskan penyelesaian dari masalah yang dihadapi dan mencoba menganalisis kebutuhan, dari sisi metode yang digunakan, data yang dipakai, dan perangkat lunak yang akan dipakai, Analisis kebutuhan ini melibatkan pihak program studi S1 Teknik Informatika Institut Teknologi PLN sebagai penyedia layanan. Dibutuhkan suatu sistem yang membantu pihak prodi untuk memberi informasi akademik secara efisien dalam

menjawab pertanyaan seputar akademik secara otomatis dengan *Chatbot* dan dinamis dalam pembaruan basis pengetahuan dari *Chatbot*.



Gambar 1. Tahapan Penelitian Pembangunan *Chatbot* Informasi Akademik

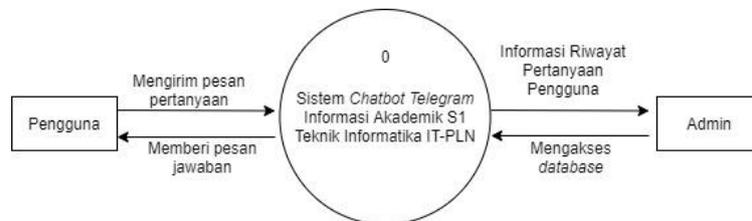
Pada awal penelitian, pengumpulan data dilakukan dengan mengadakan wawancara dengan pihak Prodi dan menyebarkan kuesioner kepada mahasiswa sebagai instrumen kebutuhan awal penelitian. Dari hasil wawancara diperoleh 163 data sebagai basis pengetahuan dan daftar pertanyaan yang sering diajukan oleh mahasiswa ke Prodi. Sedangkan pada penyebaran kuesioner, terdapat 10(sepuluh) pernyataan yang diberikan dan responden memberikan tanggapan dengan jawaban: Sangat Tidak Setuju, Tidak Setuju, Kurang Setuju, Setuju, dan Sangat Setuju. Adapun kesepuluh pernyataan yang diajukan adalah sebagai berikut :

- 1) Semua Informasi seputar akademik Prodi S1 Teknik Informatika IT-PLN disalurkan dengan tepat kepada mahasiswa S1 Teknik Informatika IT-PLN.
- 2) Mahasiswa S1 Teknik Informatika IT-PLN mudah dalam mendapatkan informasi seputar akademik.
- 3) Informasi akademik yang diminta oleh mahasiswa S1 Teknik Informatika IT-PLN diberikan dengan sesuai dan tepat.
- 4) Informasi akademik yang ditanyakan oleh mahasiswa S1 Teknik Informatika IT-PLN didapatkan secara langsung saat itu juga atau *realtime* kapanpun termasuk di luar jam kerja.
- 5) Informasi akademik yang didapatkan dari teman teman terkadang kurang tepat dan masih membingungkan.
- 6) Ketika mahasiswa S1 Teknik Informatika IT-PLN menanyakan informasi akademik, waktu yang diperlukan untuk mendapatkan jawaban selalu kurang dari 24 jam.
- 7) Pertanyaan yang ditanyakan oleh mahasiswa S1 Teknik Informatika mengenai informasi akademik sebagian besar adalah pertanyaan yang sama dan ditanyakan secara berulang kepada pihak program studi S1 Teknik Informatika IT-PLN.

- 8) Informasi akademik disajikan melalui media alternatif seperti *instagram* dan *whatsapp* yang digunakan oleh pihak Program Studi S1 Teknik Informatika IT-PLN sudah mampu membantu mahasiswa S1 Teknik Informatika IT-PLN memperoleh informasi yang dibutuhkan.
- 9) Prodi S1 Teknik Informatika IT-PLN menyediakan wadah atau *platform* khusus bagi mahasiswa S1 Teknik Informatika IT-PLN untuk mendapatkan informasi akademik.
- 10) Apakah aplikasi *Chatbot* di *Telegram* dapat membantu mahasiswa S1 Teknik Informatika mendapatkan informasi akademik secara *realtime* dan tepat.

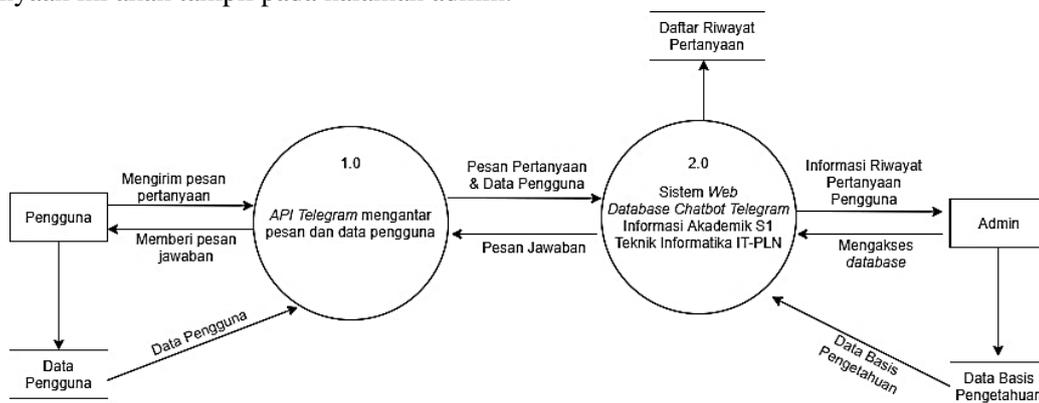
Data yang digunakan dalam pembangunan *chatbot* merupakan data yang didapatkan dari catatan internal Prodi S1 Teknik Informatika IT PLN. Data ini meliputi data dosen, termasuk nama dosen, NIDN (Nomor Induk Dosen Nasional), NIP (Nomor Induk Pegawai), kalender akademik, *link* surat menyurat, pelaksanaan remedial, jadwal kegiatan, dan lain sebagainya.

2.2 Data Flow Diagram (DFD) dari Aplikasi Chatbot Informasi Akademik



Gambar 2. Diagram Konteks sistem Chatbot Informasi Akdemik

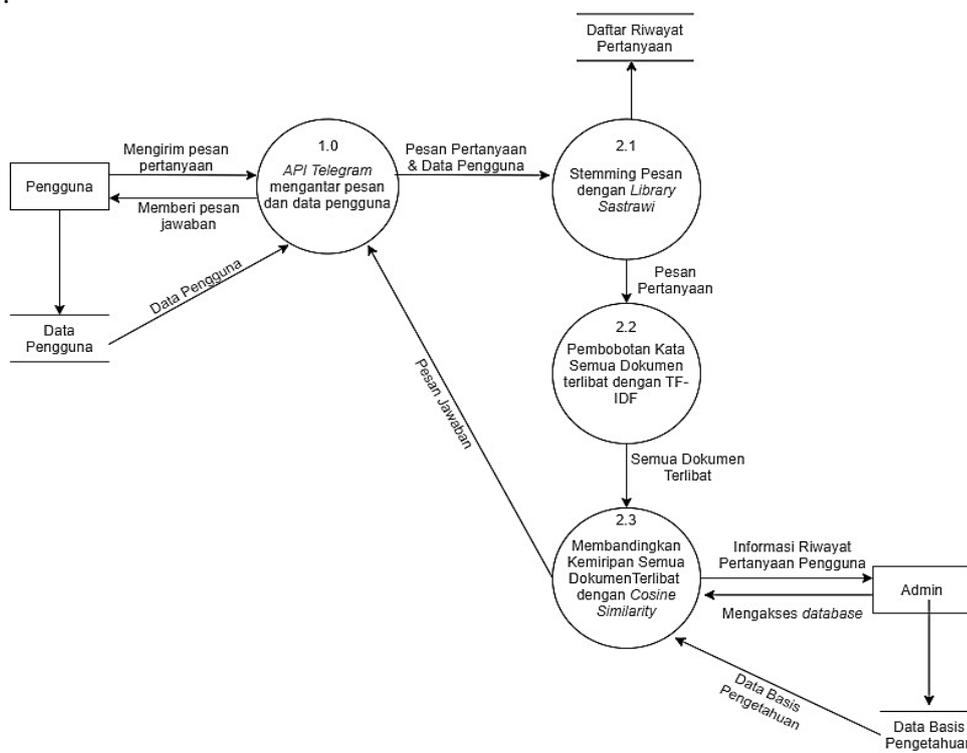
Dalam membuat *Data Flow Diagram (DFD)* langkah pertama yaitu pembuatan diagram konteks dari sistem. Diagram konteks berada pada level 0 dan hanya menggambarkan sistem secara umum dengan menampilkan aliran data pada sistem. Gambar 2 menunjukkan proses yang dilakukan oleh sistem *Chatbot Telegram* Informasi Akademik Program Studi S1 Teknik Informatika IT-PLN, dimulai dari pengguna yang mengirim pesan pertanyaan ke sistem, lalu pertanyaan tersebut tersimpan sebagai data *store* berisikan informasi riwayat pertanyaan pengguna . Informasi Riwayat pertanyaan ini akan tampil pada halaman admin.



Gambar 3. Diagram Nol level 1

Gambar 4 merupakan diagram rinci 2.0 yang termasuk ke dalam DFD level 2. Pada diagram memecah proses yang ada pada proses 2.0 yaitu *Sistem Web Database Chatbot Telegram* Informasi

Akademik S1 Teknik Informatika IT-PLN. Pemecahan proses dimulai dari pesan pertanyaan pengguna melalui *API Telegram* masuk ke sistem, kemudian terjadi proses *stemming* pesan dengan *Library Sastrawi Stemmer*, lalu pesan diproses dengan semua data yang ada pada sistem basis pengetahuan untuk selanjutnya diberi pembobotan yang dilakukan menggunakan *TF-IDF*. Kemudian setelah pembobotan semua dokumen atau data yang terlibat akan dibandingkan menggunakan *Cosine Similarity*.



Gambar 4. Diagram Rinci 2.0 level 2

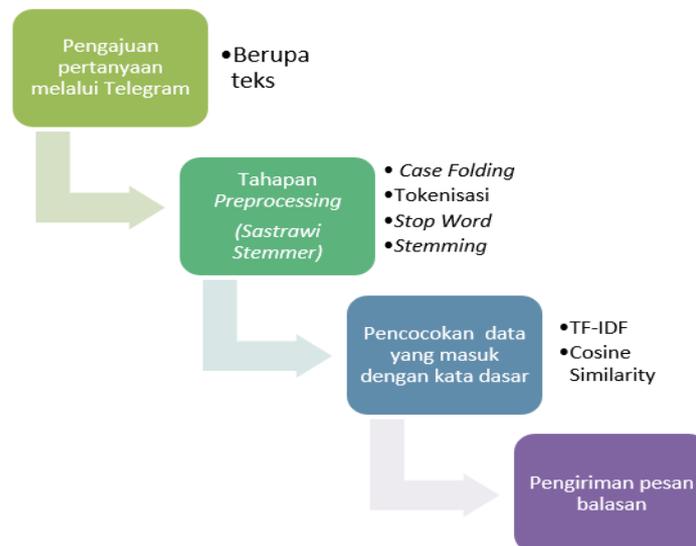
2.3 Tahapan Pemrosesan Teks pada Chatbot Telegram Informasi Akademik

Deskripsi alur sistem *chatbot* dapat dilihat pada gb.5. Pada proses penerimaan pesan, sistem menerima pesan masuk yang dikirimkan melalui *instant messaging* (Telegram). Teks yang masuk tersebut kemudian diteruskan ke tahapan *preprocessing* untuk mengubah pesan menjadi data yang sesuai dengan format pemrosesan yang diperlukan. Tahapan selanjutnya adalah proses pencocokan pola dengan kata dasar yang ada di dalam basis pengetahuan (*knowledge base*) menggunakan metode *Cosine Similarity*. Untuk menghitung kemiripan antar objek yaitu pesan yang telah di-*stemming* dengan basis pengetahuan dan diberi bobot dengan *TF-IDF*. Lalu setelah semua proses berjalan maka sistem akan memberikan pesan balasan sesuai dengan pertanyaan yang diajukan oleh pengguna.

Pada tahapan *Preprocessing*, pesan yang masuk akan diubah menjadi kata dasar melalui proses *case folding*, *tokenisasi*, *stop word*, dan *stemming* dengan bantuan *Library Sastrawi*. *Sastrawi* merupakan pustaka PHP yang digunakan untuk mereduksi kata-kata infleksi Bahasa Indonesia ke bentuk kata dasar [14]. Adapun penjelasan dari tiap proses tersebut adalah sebagai berikut:

- 1) *Case Folding*, untuk mengubah semua huruf dalam teks menjadi huruf kecil dan menghapus karakter yang tidak diperlukan.

- 2) *Tokenisasi*, yaitu proses memecah aliran teks menjadi kata, frasa, simbol, atau elemen bermakna lainnya yang disebut token atau *term*.
- 3) *Stop Word*, yaitu tahapan menghapus kata berulang pada dokumen yang tidak memiliki makna karena berfungsi sebagai penghubung kata-kata pada kalimat.
- 4) *Stemming*, yaitu proses menggabungkan beberapa *term* menjadi suku kata.



Gambar 5. Bagan Alur Proses *Chatbot* Informasi Akademik

Algoritma pada pustaka sastrawi adalah algoritma *stemming* Nazief dan Adriani. Hasil *Stemming* didapat berdasarkan beberapa proses seperti berikut [15] :

- 1) Mencocokkan kata yang dicocokkan dengan basis data, dimana *Library Sastrawi Stemmer* adalah pustaka *PHP* untuk mengubah kata-kata bahasa Indonesia menjadi bentuk dasarnya. Disaat terdapat kata yang cocok maka proses berhenti.
- 2) Menghapus sufiks atau imbuhan yang ditempatkan di akhir kata (“-lah”,”-kah”,”-tah”, atau “-pun”). Proses ini diulangi lagi untuk menghilangkan sufiks *inflectional* (“-ku”,”-mu”, atau “-nya”).
- 3) Menghapus sufiks *derivational* (“-i”,”-an”, atau “-kan”), bila imbuhan akhir adalah “-an” dan huruf terakhirnya adalah “-k” maka “-k” dihapus. Apabila tidak ditemukan pada pustaka proses akan berhenti, namun apabila tidak ditemukan maka proses menghapus (“-i”,”-an” atau”-kan”) tidak dilakukan
- 4) Menghapus prefiks *derivational* (“be-“,”di-“,”ke-“,”me-“,”pe-“,”se-“, dan “te-“) proses ini akan berhenti apabila terdapat kondisi seperti adanya kombinasi imbuhan awal dan imbuhan akhir yang tidak diperbolehkan. Kemudian terdapat imbuhan awal yang sama dengan awalan yang telah terhapus. Dan menghilangkan 3 awalan.
- 5) Apabila dari proses yang dilakukan tidak memberikan kata dasar untuk kata tersebut maka kata asli sebelum *stemming* akan dikembalikan.

Proses *Stop Word* pada tabel 1 menghapus kata berulang dan kata yang tidak memiliki makna serta angka untuk kemudian dideklarasikan pada program. Kata yang akan dihilangkan saat *Stop Word*, antara lain: ‘yang’, ‘untuk’, ‘pada’, ‘ke’, ‘para’, ‘dsb’, ‘dst’, ‘dll’, ‘dahulu’, ‘dulunya’, ‘anu’, ‘demikian’, ‘tapi’, ‘ingin’, ‘juga’, ‘nggak’, ‘mari’, ‘nanti’, ‘ok’, ‘ok’, ‘seharusnya’, ‘sebetulnya’,

'setiap', 'setidaknya', 'sesuatu', 'pasti', 'saja', 'ya', 'walau', 'tolong', 'tentu', 'amat', 'apalagi', 'bagaimanapun', '0', '1', '2', '3', '4', '5', '6', '7', '8', '9', '!', '@', '#', '\$', '%.

Tabel 1. Contoh Hasil *Pre-Processing* Pesan

Proses	Teks
Tanpa <i>Pre-Processing</i> pesan	Kapan Pendaftaran dan Pengumpulan berkas sidang skripsi gelombang 2 ?
<i>Case Folding</i>	kapan pendaftaran dan pengumpulan berkas sidang skripsi gelombang 2
<i>Tokenisasi</i>	kapan;pendaftaran;dan;pengumpulan; berkas; sidang; skripsi; gelombang; 2
<i>Stop Word</i>	kapan;pendaftaran;pengumpulan;berkas;sidang; skripsi;gelombang
<i>Stemming</i>	kapan daftar kumpul berkas sidang skripsi gelombang

3. HASIL DAN PEMBAHASAN

3.1 Penerapan *Library Sastrawi* dan Metode *Cosine Similarity*

Library Sastrawi, metode *TF-IDF*, dan metode *Cosine Similarity* diterapkan dalam menjawab pertanyaan seputar informasi akademik yang ditanyakan oleh pengguna *Chatbot Telegram* Informasi Akademik Program Studi S1 Teknik Informatika Institut Teknologi – PLN dimana peneliti memilih pertanyaan “Saya Mau Membuat Surat Riset?” Dengan jumlah data pada basis pengetahuan sebanyak 163 data atau dokumen. Untuk menampilkan penerapan metode peneliti menggunakan sampel perhitungan dengan 11 dokumen terlibat, yaitu 1 pertanyaan dan 10 sampel data atau dokumen dari perhitungan sebenarnya. Penerapan metode direalisasikan melalui beberapa tahap sebagai berikut:

1. *Processing* Pesan

Pertama menentukan nilai N yaitu jumlah semua dokumen yang terlibat, dengan pertanyaan (Q) “Saya Mau Membuat Surat Riset?” dan 163 dokumen lainnya yang terlibat maka nilai N=164. Kemudian seperti yang dapat dilihat pada gambar, beberapa tahapan *preprocessing* pesan, yaitu: *Case Folding*, *Tokenisasi*, *Stop Word*, dan *Stemming*. Sebagai contoh diambil beberapa pernyataan atau pertanyaan dari 163 data yang ada di basis pengetahuan, yaitu D2, D25, D53, D54, D55, D56, D101, D111, D149, D153).

- a. D2 : Berapa NIDN dosen Abdurrasyid., S.Kom., MMSI ?
- b. D25 : Berapa NIDN dosen Muhammad Fadli Prathama, S.Si., MMSI ?
- c. D53 : Membuat surat aktif kuliah
- d. D54 : Membuat surat magang
- e. D55 : Membuat surat riset
- f. D56 : Berapa lama proses pengerjaan surat
- g. D101 : Kapan pelaksanaan sidang magang gelombang 2 semester genap 2020/2021?
- h. D111 : Kapan pelaksanaan sidang skripsi gelombang 2?
- i. D149 : Berapa NIP dosen Eka Putra, S.Kom., M.Kom. ?
- j. D153 : Berapa NIP dosen Yudhy Setyo Purwanto, SS. M. Hum., MM, MOS?

Hasil dari *Case Folding*, *Tokenisasi*, *Stop Word*, dan *Stemming* dari data di atas adalah sebagai berikut:

Tabel 2. Contoh Hasil *Preprocessing* dar

Data	Term yang mewakili data
Q	saya mau buat surat riset
D2	berapa nidn dosen abdurrazyid kom mmsi
D25	berapa nidn dosen muhammad fadli prathama mmsi
D53	buat surat aktif kuliah
D54	buat surat magang
D55	buat surat riset
D56	berapa lama proses kerja surat
D101	kapan laksana sidang magang gelombang semester genap
D111	kapan laksana sidang skripsi gelombang
D149	berapa nip dosen eka putra kom
D153	berapa nip dosen yudhy setyo purwanto hum mos

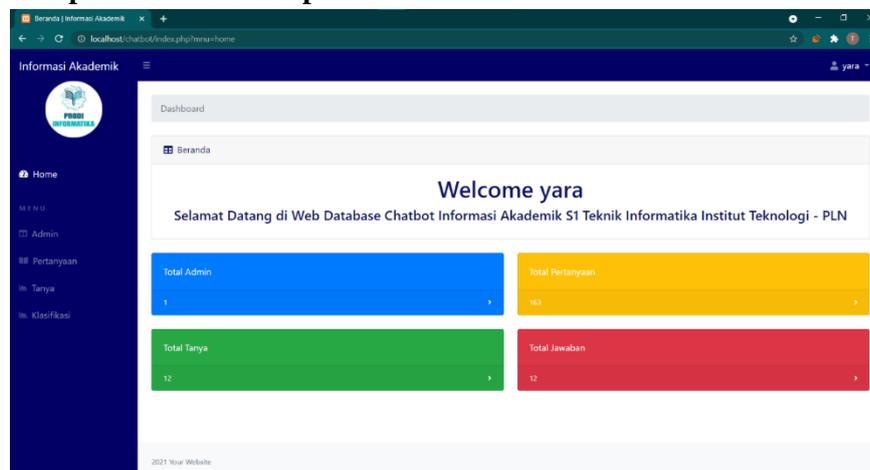
2. Pembobotan setiap *term* dari 164 dokumen, dengan menerapkan metode *Term Frequency – Inverse Document* atau TF-IDF.
3. Menghitung kemiripan dokumen antara pertanyaan yang diajukan (Q) dengan data pada basis pengetahuan (D) menggunakan rumus metode *cosine Similarity*. Berdasarkan perhitungan diperoleh hasil bahwa Q paling mirip dengan dokumen 55(D55), seperti ditunjukkan pada tabel 3.

Tabel 3. Urutan dokumen paling mirip

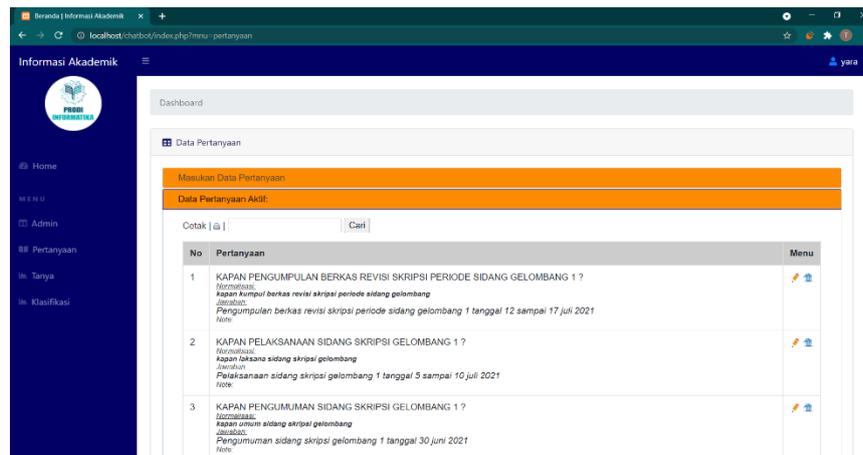
1	2	3	4	5	6	7	8	9	10
D55	D54	D53	D56	D2	D25	D101	D111	D149	D153

Dengan pertanyaan yang masuk (Q) adalah “Saya Mau Membuat Surat Riset?”, kemudian dinormalisasikan menjadi “saya mau buat surat riset “. Pertanyaan yang mirip pada dokumen basis pengetahuan adalah “Membuat surat riset “. Mekan jawaban yang diberikan (oleh *chatbot*) adalah “Isi tautan berikut untuk membuat surat aktif kuliah <https://bit.ly/SuratAktifKuliahInformatika> pembuatan surat diproses dalam 3 hari setelah pengisian data, kemudian surat dapat diunduh pada tautan berikut <http://bit.ly/SuratMahasiswaInfor> “.

3.2 Tampilan Aplikasi Web dan Aplikasi Chatbot

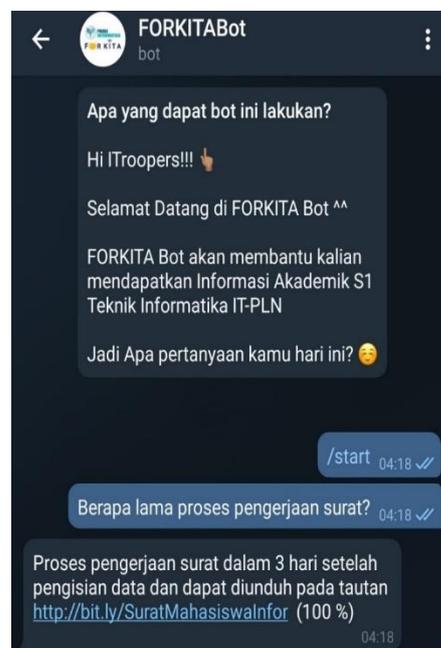


Gambar 6. Tampilan Halaman Beranda (*Home*)



Gambar 7. Tampilan Halaman Data Pertanyaan

Aplikasi web yang dibangun ini ditujukan untuk admin untuk pengelolaan *chatbot* Telegram Informasi Akademik di Prodi S1 Teknik Informatika IT PLN. Gambar 6 merupakan tampilan halaman beranda untuk *Web Database Chatbot* yang menampilkan total admin yang terdaftar, total pertanyaan dalam basis pengetahuan, total tanya dari pengguna, dan total jawaban yang dijawab oleh sistem. Ada 163 pertanyaan yang tersimpan sebagai basis pengetahuan, namun penambahan pertanyaan baru oleh admin dimungkinkan dengan cara memasukkan pertanyaan atau pernyataan baru melalui web atau memasukkan dokumen excel. Gambar 7 merupakan halaman daftar data pertanyaan aktif yang memiliki dua menu yaitu menu hapus dan ubah untuk mengubah data pertanyaan. Gambar 8 berikut adalah tampilan saat *bot FORKITA* telah ditambahkan lalu memulai percakapan.



Gambar 8. Tampilan FORKITABot pada Saat Menjawab Pertanyaan

4. KESIMPULAN DAN SARAN

Dari serangkaian proses perencanaan pembuatan, penerapan metode *Cosine Similarity* dengan *Library Sastrawi Stemmer* hingga pengujian akurasi pada penelitian ini dapat diambil beberapa kesimpulan yaitu:

1. Dalam pembuatan *Chatbot Telegram* Informasi Akademik program studi S1 Teknik Informatika Institut Teknologi PLN penulis melalui beberapa proses mulai dari pembuatan *Data Flow Diagram* (DFD), perancangan serta pembuatan basis data, perancangan tampilan tatap muka, pelatihan datalatih pada database, pembuatan *bot*, dan penghubungan *web database* dengan *bot*.
2. Analisis kebutuhan untuk *Chatbot Telegram* Informasi Akademik program studi S1 Teknik Informatika Institut Teknologi PLN adalah sistem *Chatbot Telegram* menggunakan bahasa pemrograman *PHP* dengan *MySQL* sebagai manajemen basis datanya, *chatbot* dapat diakses oleh pengguna melalui aplikasi pihak ketiga *Telgram*, sistem hanya bisa memproses pesan berupa teks dengan pemahaman bahasa Indonesia untuk percakapannya, dan pembaruan basis pengetahuan dilakukan oleh admin dengan mudah pada *Web Database Chatbot*.
3. Penerapan dari metode *Cosine Similarity* dengan *Library Sastrawi Stemmer* pada *Chatbot Telegram* memiliki hasil akurasi ketepatan yang cukup tinggi yaitu sebesar 90% dengan rata-rata akurasi kemiripan dokumen pada setiap data uji sebesar 82,424%. Hal ini dipengaruhi oleh penulisan pertanyaan oleh pengguna dan isi dari basis pengetahuan.

UCAPAN TERIMAKASIH

Penulis mengucapkan terima kasih kepada LPPM Institut Teknologi PLN yang telah memberi dukungan dana hibah untuk membantu pelaksanaan penelitian dan atau penulisan artikel.

DAFTAR PUSTAKA

- [1] A. Heryandi, "Developing Chatbot for Academic Record Monitoring in Higher Education Institution," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 879, no. 1, 2020, doi: 10.1088/1757-899X/879/1/012049.
- [2] E. Adamopoulou and L. Moussiades, *An Overview of Chatbot Technology*, vol. 584 IFIP, no. June. Springer International Publishing, 2020.
- [3] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Comput. Educ. Artif. Intell.*, vol. 2, p. 100033, 2021, doi: 10.1016/j.caeai.2021.100033.
- [4] S. Garcia-Mendez, F. De Arriba-Perez, F. J. Gonzalez-Castano, J. A. Regueiro-Janeiro, and F. Gil-Castineira, "Entertainment Chatbot for the Digital Inclusion of Elderly People without Abstraction Capabilities," *IEEE Access*, vol. 9, pp. 75878–75891, 2021, doi: 10.1109/ACCESS.2021.3080837.
- [5] N. Bhirud, S. Tataale, S. Randive, and S. Nahar, "A Literature Review On Chatbots In Healthcare Domain," *Int. J. Sci. Technol. Res.*, vol. 8, no. 7, pp. 225–231, 2019.
- [6] T. Makasi, A. Nili, K. C. Desouza, and M. Tate, "A Typology of Chatbots in Public Service Delivery," *IEEE Softw.*, vol. 39, no. 3, pp. 58–66, 2022, doi: 10.1109/MS.2021.3073674.
- [7] T. Nadu, "Education," no. Icicv, pp. 1317–1322, 2021.
- [8] J. R. Carrizales, Y. J. Ramirez, J. A. Armas, and E. E. Grandón, "Cognitive services to improve user experience in searching for academic information based on chatbot," pp. 1–4, 2019.

- [9] M. Kowsher, M. A. Alam, M. J. Uddin, M. R. Islam, N. Pias, and A. R. M. Saifullah, "Bengali Informative Chatbot," *5th Int. Conf. Comput. Commun. Chem. Mater. Electron. Eng. IC4ME2 2019*, pp. 1–7, 2019, doi: 10.1109/IC4ME247184.2019.9036585.
- [10] R. Rismanto, Y. Yunhasnawa, and R. A. Bhakti, "Penerapan Metode Cosine Similarity Dalam Aplikasi Chatbot Layanan Wisata Di Wilayah Malang," *Semin. Inform. Apl. Polinema*, pp. 1–8, 2019.
- [11] A. R. Lahitani, A. E. Permanasari, and N. A. Setiawan, "Cosine similarity to determine similarity measure: Study case in online essay assessment," *Proc. 2016 4th Int. Conf. Cyber IT Serv. Manag. CITSM 2016*, 2016, doi: 10.1109/CITSM.2016.7577578.
- [12] M. A. Rosid, A. S. Fitriani, I. R. I. Astutik, N. I. Mulloh, and H. A. Gozali, "Improving Text Preprocessing for Student Complaint Document Classification Using Sastrawi," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 874, no. 1, 2020, doi: 10.1088/1757-899X/874/1/012017.
- [13] T. S. Gunawan, A. B. Falelmula Babiker, N. Ismail, and M. R. Effendi, "Development of Intelligent Telegram Chatbot Using Natural Language Processing," *Proceeding 2021 7th Int. Conf. Wirel. Telemat. ICWT 2021*, no. April 2022, 2021, doi: 10.1109/ICWT52862.2021.9678471.
- [14] I. K. P. Pinajeng, I. M. Sukarsa, and I. M. S. Putra, "Perbaikan Kata pada Sistem Chatbot dengan Metode Jaro Winkler," *JITTER J. Ilm. Teknol. dan Komput.*, vol. 1, no. 2, pp. 86–95, 2020.
- [15] A. C. Herlingga, I. P. E. Prisma, D. R. Prehanto, and D. A. Dermawan, "Algoritma Stemming Nazief & Adriani dengan Metode Cosine Similarity untuk Chatbot Telegram Terintegrasi dengan E-layanan," *J. Informatics Comput. Sci.*, vol. 2, no. 01, pp. 19–26, 2020, doi: 10.26740/jinacs.v2n01.p19-26.