**Penerapan *Natural Language Processing* Pada Aplikasi *Chatbot* Info Layanan Kantor Menggunakan *Naïve Baiyes Algorithm***

**Rakhmat Kurniawan1, Armansyah2, Kesuma Dwi Ningtyas3**

1,2,3Ilmu Komputer, Universitas Islam Negeri Sumatera Utara Medan

Email: 1rakhmat.kr@uinsu.ac.id, 2armansyah@uinsu.ac.id, 3dkesuma5@gmail.com

Email Penulis Korespondensi: dkesuma5@gmail.com

| **Article History:**Received Des 19th, 2022Revised Des xth, 202xAccepted Des xth, 202x |  | **Abstrak** |
| --- | --- | --- |
|  | Informasi layanan kantor lura adalah bagian penting dalam kehidupan sehari-hari, dimana informasi layanan kantor lurah tersebut diperoleh salah satunya dengan konsultasi langsung dengan *customer service.* Berdasarkan hasil wawancara yang dilakukan terhadap beberapa masyarakat, masyarakat memperoleh informasi layanan kantor lurah dengan cara berkunjung ke kantor lurah dan bertanya langsung kepada *customer* service. Peyampaian informasi layanan kantor lurah tersebut dirasa kurang karena keterbatasan oleh waktu kerja pegawai. Pada penelitian ini solusi yang diusulkan untuk masalah tersebut salah satunya dengan cara membangun sebuah aplikasi *chatbot* info layanan kantor dengan penerapan  *Natural Language Processing* dengan menggunakan *Naïve Bayes Algorithm.* **Kata Kunci :** *Chatbot, Natural Language Processing, Naïve Bayes Algorithm, Artificial Intelligence,* Informasi Kantor |
| ***Abstract*** |
| *Information on village head office services is an important part of everyday life, where information on village head office services is obtained by consulting directly with customer service. Based on interviews conducted with several communities, the community obtained information on the services of the lurah office by visiting the lurah office and asking directly to customer service. Submission of information on the services of the lurah's office is felt to be lacking due to limitations by the employee's working time. In this study, one of the proposed solutions to this problem is by building an office service info chatbot application by implementing Natural Language Processing using the Naïve Bayes Algorithm.****Keyword :*** *Chatbot, Natural Language Processing, Naïve Bayes Algorithm, Artificial Intelligence, Office Information* |

**1. PENDAHULUAN**

Penyampaian informasi melalui layanan *customer service*  saat ini mulai berkembang pesat. Seperti halnya pada perkembangan layanan *customer service* di media social dunia. Layanan *customer service* dapat dilakukan dengan menjawab langsung pertanyaan oleh *user*  atau dapat menggunakan mesin untuk menjawab pertanyaan yang diajukan secara otomatis. Layanan public merupakan suatu media yang disediakan oleh pemerintah untuk dapat memberikan informasi terkini kepada massyarakat.

*Chatbot*  merupakan suatu program *computer* yang melakukan percakapan dengan menggunakan metode auditoria atau tekstual. [1] *Chatbot*  adalah teknologi yang tujuan utamanya adalah untuk berinteraksi dengan pengguna manusia dengan memproses masukan bahasa alamai dan menghasilkan keluaran *relative* melalui mesin yang digerakkan oleh aturan atau mesin kecerdasan buatan. [2] Proses *chatbot* dimulai dengan masukan dari pengguna menggunakan bahasa alami dan system akan menjawab dengan respon yang masuk akal atau bisa dikatakan cerdas untuk bahasa yang sebenarnya.[3] *Chatbot*  berpotensi disebut sebagai bentuk interaksi manusia-mesin yang paling menjanjikan dan canggih.[4]

Kecerdasan buatan atau disebut *Artificial Intelligence,* disingkat AI adalah suatu system atau computer yang memiliki kecerdasan buatan didalamnya, dimana terdapat algoritma dan keputusan-keputusan untuk mengambil sebuah keputusan, kecerdasan buatan tidak lepas dari campur tangan manusia dan manusia lah yang menanamkan sebuah kecerdasan tersebut ke dalam suatu *system* teknologi dan diatur serta dikembangkan dalam konteks ilmiah dan matematis. [5]

Teorema ini digabungkan dengan “*naïve”* dimana didefenisikaan kondisi antar atribut saling bebas. [6] *Naïve Bayes Classifier* adalah metode yang popular dan cocok untuk masalah yang diselesaikan dengan algoritma *naïve bayes.* [7]

Bahasa natural adalah bahasa yang secara umum digunakan oleh manusia dalm berkomunikasi satu sama lain. [8] Bahasa yang diterima oleh computer butuh untuk diproses dan dipahami terlebih dahulu supaya maksud dari *user* bisa dipahami dengan baik oleh *computer .* [9]Tujuan dari NLP adalah memberikan jawaban atau respon yang sesuai berdasarkan pemahaman makna bahasa manusia yang dilakukan oleh mesin. [10] *Natural Language Processing* menggunakan tahapan *tokenizing, filtering* dan analisis serta menerapkan algoritma *nuth morris prrat.* [11]Sehingga dengan NLP pembangunan aplikasi *chatbot* dapar melakukan konversi kata dan memberikan respon sesuai dengan inputan *user* secara alami saat berkomunikasi dengan *chatbot.* [12]Dari *chatbot* ini nanti diharapkan dapat membantu kantor lurah kota Medan untuk berinteraksi dengan masyarakat dalam memberikan informasi kepengurusan surat.

**2. METODOLOGI PENELITIAN**

**2.1 Tahapan Penelitian**



Gambar 1. Tahapan Penelitian

**2.2.1 Teknik Pengumpulan Data**

### Observasi

Di Kantor Lurah Medan Tenggara di Jl. Menteng VII, Medan Tenggara, Kec. Medan Denai, Kota Medan, Sumatera Utara 20228

* + - 1. Studi Pustaka (Informasi dari beberapa Jurnal).

1. **HASIL DAN PEMBAHASAN**

Di tahap simulasi ini dijelaskan bagaimana cara mendapatkan data *input* berdasarkan pengumpulan data yang sudah dilakukan dan bagaimana cara mengolah data tersebut untuk mendapatkan kategori pertanyaan, data pertanyaan dan kata kunci. Ada beberapa langkah untuk mendapatkan kategori pertanyaan dalam penelitian ini adalah:

* 1. Mengumpulkan pertanyaan-pertanyaan yang sering ditanyakan masyarakat kepada *customer service* di Kantor Lurah

Tabel 1. Kategori Pertanyaan dan Pertanyaan

|  |  |
| --- | --- |
| Kaltegori Pertalnyalaln | Pertalnyalaln (P) |
| Pengurusaln KTP | * + - 1. Buk alpal saljal syalralt pengurusaln KTP?
			2. Alpal saljal yalng halrus dibalwal untuk pembualtaln KTP?
 |
| Pengurusaln Alktal Kelalhiraln | 1. Alpal saljal syalralt untuk bualt alktal kelalhiraln?
2. Bu, alnalk salyal lalhir dilualr kotal, talpi salyal alsli di keluralhaln ini, alpal alnalk salyal bisal mengurus alktal kelalhiralnnyal di keluralhaln ini?
 |

* 1. Melakukan analisis terhadap informasi apa yang ingin diambil dari pertanyaan masyarakat untuk setiap kategori

Tabel 2. Kategori Pertanyaan Berdasarkan Kata Kunci

|  |  |
| --- | --- |
| Kaltegori Pertalnyalalaln | Kaltal Kunci |
| Pengurusaln KTP | pengurusaln KTP |
| Pengurusaln Alktal Kelalhiraln | alktal kelalhiraln |

* 1. Membuat *rule* jawaban yang digunakan *chatbot* untuk menjawab pertanyaan

Tabel 3. Pengenalan Etitas

|  |  |  |
| --- | --- | --- |
| Kaltegori pertalnyalaln | Kaltal Kunci | *Rule* |
| Pegurusaln KTP | Pengurusaln KTP | K1R1 |
| Pegurusaln KTP | Pengurusaln KTP | K1R2 |
| Pengurusaln Alktal Kelalhiraln | Alktal Kelalhiraln | K2R1 |

* 1. Pengenalan pertanyaan
		+ 1. *Text Preprocessing* yaitu langkah penting sebelum melakukan proses klasifikasi kategori pertanyaan
			2. Pembobotan kata pada penelitian ini menggunakan pembobotan TF-IDF

Perhitungan manual TF

$TF\left(d,t\right)=f(d,t)$ (1)

Ket:

F (d,t) : frekuensi kemunculan *term* t pada data d

Perhitungan manual IDF

$IDF\left(t\right)=log\left(\frac{Nd}{df(t)}\right)$ (2)

Ket:

Nd : Jumlah seluruh data

Df (t) : Jumlah data yang mengandung *term* t

Perhitungan manual bobot (w)

$$W\left(t,d\right)=TF\left(d,t\right)×IDF$$

$ =TF\left(d,t\right)×\left(log\left(\frac{Nd}{df(t)}\right)\right)$ (3)

Ket:

W (d,t) : Bobot *term* t pada data d

TF (d,t) : Total kemunculan *term* data d

Nd : Jumlah seluruh data

Df(t) : jumlah data yang memiliki *term* t

* + - 1. Tahapan *naïve bayes* dilakukan setelah *prerocessing* untuk mentukan kategori pertanyaan. Tahapan ini memiliki dua bagian, yaitu *training* dan *testing.*
* *Training*

Menghitung nilai *prior probabilities* P© dengan menggunakan *maximum likehood* dan Langkah selanjutnya menghitung nilai *conditional propabilities* dengan menggunakan *laplace smoothing*.

$P\left(C\right)=\frac{Nc}{N}$ (4)

Ket:

P© : Probabilitas kemunculan kata dapa data

Nc : jumlah data dalam kelas c

N : jumlah data yang dipilih

Tahapan *laplace smoothing*  bertujuan untuk menghilangkan nilai 0 (nil) pada akhir proses pelatihan

$P\left(c\right)=\frac{Wct+1}{\left(ΣW^{'}ϵV Wct^{'}\right)+B'}$ (5)

Ket:

P(tk|c) : *conditional probabilities term* yang memiliki kategori

Wct : nilai pembobotan TF-IDF atau bobot dari kata t pada kelas C

ΣW^' ϵV Wct^ : junlah bobot dari seluruh kata yang ada di kelass C

B` : jumlah bobot kata unik di semua kelas

* *Testing*

Setelah hasil data *text preprocessing* didapat, selanjutnya melakukan Langkah *matvhing* dan menghitung probabilitas.

Tabel 4. Data *Testing*

|  |  |
| --- | --- |
| Jenis Pertalnyalaln | Isi Pertalnyalaln |
| ? | Jikal KTP hilalng, alpal yalng dibutuhkaln untuk mengurusnyal? |

Hasil *text preprocessing* untuk data *testing* dapat dilihat pada table 5.

Tabel 5. Hasil *Text Preprocessing* Data *Testing*

|  |
| --- |
| Testing |
| ktp |
| hilalng |
| butuh |
| urus |

Setelah hasil data *text preprocessing* didapat, selanjutnya melakukan Langkah *matching* dan menghitung probabilitas.

Tabel 6. Hasil *Matching Data Training* Dan *Data Testing*

|  |
| --- |
| Testing |
| ktp |
| butuh |
| urus |

Tabel 7. Data perhitungan manual probabilitas *testing*

|  |  |  |
| --- | --- | --- |
| Term | TFTesting | LS |
| Pengurusaln KTP | Pengurusaln Alktal Kelalhiraln | Pengurusaln Suralt Keteralngaln Tidalk Malmpu | Pengurusaln Suralt Keteralngaln Kemaltialn | Pengurusaln Suralt Keteralngaln Usalhal | Pengurusaln Suralt Pengalntalr Kalrtu Kelualrgal |
| ktp | 1 | 0,134417483 | 0,071394289 | 0,069892179 | 0,071394289 | 0,071394289 | 0,069892179 |
| butuh | 1 | 0,075593954 | 0,071394289 | 0,069892179 | 0,126949845 | 0,071394289 | 0,069892179 |
| urus | 1 | 0,098350001 | 0,071394289 | 0,090931821 | 0,071394289 | 0,092886112 | 0,069892179 |

Tabel 8. Data Perhitungan Manual Pemangkatan LS dan TF

|  |  |  |
| --- | --- | --- |
| TF | TFTesting | LS |
| Pengurusaln KTP | Pengurusaln Alktal Kelalhiraln | Pengurusaln Suralt Keteralngaln Tidalk Malmpu | Pengurusaln Suralt Keteralngaln Kemaltialn | Pengurusaln Suralt Keteralngaln Usalhal | Pengurusaln Suralt Pengalntalr Kalrtu Kelualrgal |
| ktp | 1 | 0,134417483 | 0,071394289 | 0,069892179 | 0,071394289 | 0,071394289 | 0,069892179 |
| butuh | 1 | 0,075593954 | 0,071394289 | 0,069892179 | 0,126949845 | 0,071394289 | 0,069892179 |
| urus | 1 | 0,098350001 | 0,071394289 | 0,090931821 | 0,071394289 | 0,092886112 | 0,069892179 |
| Halsil perkallialn | 0,000999349 | 0,000363907 | 0,000444194 | 0,000647082 | 0,000473454 | 0,000341417 |
| Perkallialn *prior probalbility* | 0,000166891 | 6,07725E-05 | 7,41805E-05 | 0,000108063 | 7,90668E-05 | 5,70167E-05 |

## Pengujian

Halaman utama *website*



Gambar 2. Halaman utama aplikasi

1. Halaman otentikasi *admin*



Gambar 3. Halaman otentikasi *admin*

1. Halaman beranda *admin*



Gambar 4. Halaman *dashboard admin*

1. Tampilan pengujian data *training/* pelatihan *bot*



Gambar 5. Halaman *training chatbot*

1. Tampilan pengujian system notifikasi



Gambar 6. Notifikasi aplikasi

1. Tampilan pengujian system obrolan



Gambar 7. Pengujian system *chatbot*

1. **KESIMPULAN**

Berdasarkan pembahasan dan hasil yang dijelaskan maka didapat kesimpulan *chatbot* hanya dapat menjawab pertanyaan berdasarkan kategori dan kata kunci yang telah disimpan didalam *knowledge base chatbot* . *Natural Language Processing*  dan  *Naïve Bayes algorithm* dapat membantu *chatbot* untuk memahami maksud pertanyaan berdasarkan bahasa alami dengan nilai akurasi klasifikasi kategori pertanyaan sebesar 0,000166891.

**UCAPAN TERIMA KASIH**

Syukur Alhamdulillah kepada Allah atas selesainya skripsi dan jurnal ini. Dan tak lupa pula terimakasih khususnya kepada orangtua saya.

**DAFTAR PUSTAKA**

[1] E. W. and P. Pebriantara, ““Rancangan Bangun Aplikasi Pembelajaran dengan Memanfaatkan Chatbot API Dialogflow dan Moodle Berbasis Android Pada SMA IT ALIA Tangerang,” vol. 3, hal. 328–335, 2018.

[2] F. Z. M. Indrayani, L. M., Amalia, R. M., & Hakim, “Emotive Expressions on Social Chatbot,” *J. Sosioteknologi*, vol. 18, hal. 509, 2020.

[3] S. J. S. Tyas, M. Febianah, F. Solikhah, A. L. Kamil, dan W. A. Arifin, “Analisis Perbandingan Algoritma Naive Bayes Dan C.45 Dalam Klasifikasi Data Mining Untuk Memprediksi Kelulusan,” *J. Teknol. Inf. Dan Komun.*, vol. 8, no. 1, hal. 86–99, 2021.

[4] F. Battineni, G., Chintalapudi, N., & Amenta, “AI Chatbot Design during an Epidemic like the Novel Coronavirus,” *Healthcare*, vol. 8, hal. 154, 2020.

[5] N. Ratama, *Sistem Penunjang Keputusan Dan Sistem Pakar Dengan Pemahaman Studi Kasus*. Tangerang Selatan: Penerbit Uwais Inspirasi Indonesia, CV, 2019.

[6] D. Ginantra, Ni luh Wiwik Sri Rahayu., Arifah, Fatimah Nur., Wijaya, Anggi Hadi., Septarini, Ri Sabti., *Data Mining dan Penerapan Algoritma*. Medan: Yayasan Kita Menulis, 2021.

[7] A. Khan, R., & Das, “Build Better Chatbots,” 2018.

[8] A. Y. Chandra, D. Kurniawan, dan R. Musa, “Perancangan Chatbot Menggunakan Dialogflow Natural Language Processing (Studi Kasus: Sistem Pemesanan pada Coffee Shop),” *J. Media Inform. Budidarma*, vol. 4, no. 1, hal. 208, 2020, doi: 10.30865/mib.v4i1.1505.

[9] B. et al Sugara, “Perbandingan Akurasi Algoritma C4.5 dan Naïve Bayes untuk Deteksi Dini Gangguan Autisme pada Anak,” *J. IKRA-ITH Inform.*, vol. 3, no. 1, hal. 119–128, 2019.

[10] Y. Alamanda, R., Suhery, C., & Brianorman, “Aplikasi Pendeteksi Plagiat Terhadap Karya Tulis Berbasis Web Menggunakan Natural Language Processing Dan Algoritma Knuth-Morris-Pratt,” *J. Coding, Sist. Komput. Untan*, vol. 4, hal. 33–34, 2016.

[11] A. Amrizal, V., Munandar, A., & Arini, “Identifikasi Matan Hadits Menggunakan Natural Language Processing Dan Algoritma Knuth Morris Pratt Berbasis Web,” *J. CoreIT J. Has. Penelit. Ilmu Komput. Dan Teknol. Inf.*, vol. 5, hal. 56, 2019.

[12] Y. Yunefri dan Y. Ersan Fadrial, “Chatbot Pada Smart Cooperative Oriented Problem Menggunakan Natural Language Processing Dan Naive Bayes Classifier Chatbot on Smart Cooperative Oriented Problems Using Natural Language Processing and Naive Bayes Classifier,” *J. Inf. Technol. Comput. Sci.*, vol. 4, no. 2, hal. 2021, 2021.