

Agen Virtual Berbasis Text menggunakan algoritma Jaringan Syaraf Tiruan

¹Yasmine Novtiristya Hanuun, ²Dwi Cahyono
ynh025@gmail.com, ²dwik@unitomo.ac.id

ABSTRAK

Proses pencetakan citra digital atau gambar yang telah diolah oleh komputer ke atas permukaan bahan atau media fisik lainnya dengan menggunakan mesin cetak (printer) disebut dengan digital printing, yaitu teknik percetakan yang sedang berkembang saat ini. Mayoritas *customer* pada bidang digital printing ini adalah orang umum yang masih belum mengetahui produk- produk digital printing. Maka dari itu dibutuhkan seorang customer service untuk melayani pelanggan yang ingin bertanya. Namun yang menjadi ironi adalah pelanggan ingin mendapatkan respon yang cepat untuk setiap pertanyaan, sedangkan seorang customer service hanya dapat menjawab pertanyaan pelanggan satu per satu. Sehingga hal ini dinilai kurang memberikan kepuasan pelanggan. Atas dasar tersebut penulis melakukan penelitian tentang pembuatan sebuah agen virtual yang dapat membantu customer service memberi informasi kepada pelanggan dengan cepat juga menangani beberapa pelanggan secara bersamaan. Tahapan awal merupakan pengolahan text pre-processing pada dataset menggunakan pendekatan *Natural Language Processing* yang membuat agen virtual dapat memahami bahasa yang digunakan oleh manusia. Kemudian tahap selanjutnya adalah prediksi respon menggunakan algoritma Jaringan Syaraf Tiruan. Kemudian agen ini diuji dengan menggunakan 25 percakapan yang berbeda dan berhasil menjawab 24 percakapan dan mendapatkan tingkat akurasi tinggi sebesar 96%. Agen virtual ini kemudian akan diintegrasikan ke aplikasi Telegram menggunakan API Telegram.

Kata kunci : *Natural Language Processing*, Agen Virtual, Jaringan Syaraf Tiruan, *Digital Printing*

This is an open-access article under the [CC-BY-SA](#) license.



Histori Naskah

Naskah di-Kirim : 16 Feb 2023

Naskah di-Terima: 14 Apr 2023

Naskah di-Publikasin : 30 April 2023

I. PENDAHULUAN

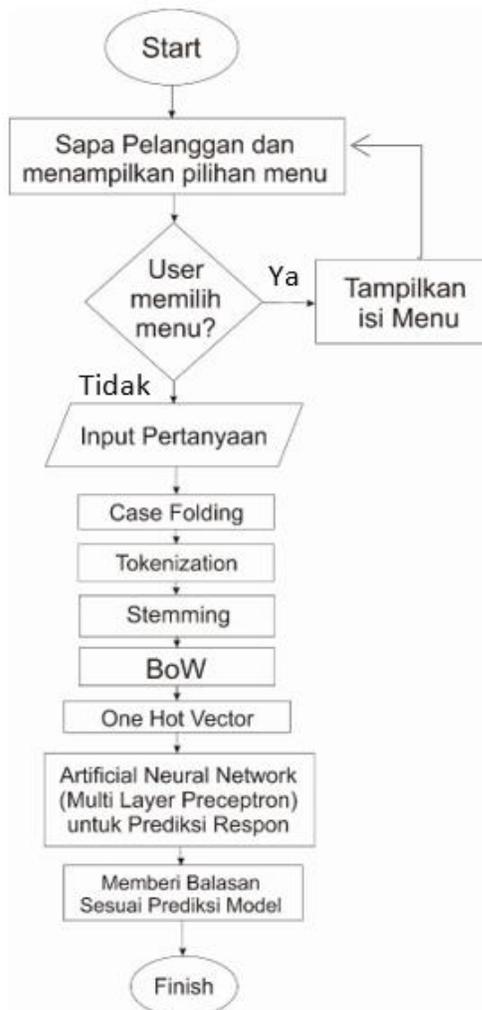
Proses pencetakan citra digital atau gambar yang telah diolah oleh komputer ke atas permukaan bahan atau media fisik lainnya dengan menggunakan mesin cetak (printer) disebut dengan digital printing, yaitu teknik percetakan yang sedang berkembang saat ini[1]. Industri percetakan pada saat ini masih sangat dibutuhkan bagi masyarakat khususnya pedagang. Banyak masyarakat yang mengurungkan niatnya untuk memilih digital printing sebagai prioritas kebutuhan dagangnya karena mengira bahwa produk- produk digital printing lumayan mahal. Padahal yang sering terjadi adalah pelanggan tidak mengetahui kebutuhan produk yang cocok untuk kebutuhan mereka. Untuk itulah, setiap perusahaan digital printing memperkerjakan seorang *customer service*. Tetapi, seorang customer service juga memiliki batasan kapasitas kerja. Ketika banyak dari pelanggan bertanya pada waktu yang bersamaan, tentunya customer service juga akan memerlukan waktu lebih lama menanggapi pertanyaan pelanggan. Namun yang menjadi ironi adalah pelanggan ingin mendapatkan respon yang cepat untuk setiap pertanyaan, sedangkan seorang customer service hanya dapat menjawab pertanyaan pelanggan satu per satu. Sehingga hal ini dinilai kurang memberikan kepuasan pelanggan.

Pemrosesan bahasa alami, juga dikenal sebagai pemrosesan bahasa alami teoretis, mengacu pada proses pengembangan berbagai teknik komputasi untuk menganalisis dan menampilkan teks dalam bahasa alami pada satu atau lebih tingkat analisis linguistik untuk mencapai tujuan manusia dalam hal bahasa, khususnya bahasa. menyelesaikan berbagai tugas dan aplikasi. [2]. Dengan adanya pendekatan Bahasa alami yang kemudian diolah menggunakan algoritma jaringan syaraf tiruan diharapkan dapat terwujud sebuah agen virtual yang memberikan layanan kepada pelanggan dengan baik.

Maka dari itu, sebuah Agen virtual dirasa merupakan jawaban dari permasalahan ini. Keuntungannya bagi perusahaan digital printing tidak perlu membayar lebih untuk menambah *staff customer service* mereka. Di pihak lain, seorang customer service juga akan merasa akan sangat diringankan karena dapat lebih menghemat waktu mereka.

II. METODE

Metodologi yang diterapkan pada penelitian ini mengacu pada Gambar 1, yang mengilustrasikan urutan operasi serta hubungan antara proses-proses tersebut dan arahnya. Untuk memberikan penjelasan yang lebih mendalam tentang bagaimana setiap tahapan program berfungsi, diagram alir dapat dibangun dengan menggunakan temuan analisis [3][4]. Pada penelitian kali ini *flowchart* akan menggambarkan alur bekerjanya sistem dalam mengolah pertanyaan (menggunakan NLP) yang kemudian akan masuk ke proses prediksi respon menggunakan JST.



Gambar 1. Metode Agen Virtual Berbasis Text

A. *Natural Language Processing (NLP)*

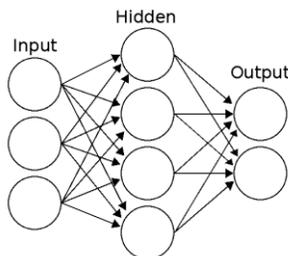
Pemrosesan bahasa alami adalah fokus utama Natural Language Processing (NLP), subbidang Kecerdasan Buatan (AI). Bahasa yang digunakan manusia untuk berkomunikasi satu sama lain disebut bahasa alami, dan itu adalah bahasa yang mereka gunakan secara alami. Bahasa yang diterima oleh komputer terlebih dahulu harus diproses dan diinterpretasikan agar komputer dapat memahami dengan benar maksud pengguna [5]. Studi dan penerapan pemrosesan bahasa alami, atau NLP, adalah cabang penelitian dan praktik yang menyelidiki bagaimana komputer dapat diprogram untuk memahami dan memanipulasi bahasa alami dalam bentuk teks atau ucapan untuk mencapai tujuan tertentu. Manipulasi teks dianggap sebagai topik penting penelitian dalam NLP. Analisis morfologi adalah langkah pertama dalam setiap sistem NLP yang bekerja dengan teks. Dalam query maupun dokumen, teks diubah untuk mendapatkan berbagai bentuk morfologis dari term yang terlibat. Pemrosesan informasi leksikal dan sintaksis membutuhkan penggunaan kamus untuk beberapa tujuan, termasuk menentukan ciri-ciri kata, mengenali bagian-bagian ucapan, menentukan kata dan frasa, dan menguraikan kalimat [6].

B. *Jaringan Syaraf Tiruan*

Jaringan syaraf tiruan adalah model yang terdiri dari neuron buatan yang saling berhubungan dan memiliki fitur yang mirip dengan jaringan saraf biologis yang ditemukan di otak manusia. Jaringan saraf biologis otak manusia berfungsi dengan menerima stimulus atau input, kemudian melakukan proses dan akhirnya menciptakan output. Agar JST dapat belajar, jaringan

harus memiliki akses ke sebagian besar informasi yang disajikan selama fase pelatihan. Mengikuti pelatihannya dengan jumlah data yang cukup besar, JST akan berusaha untuk mengklasifikasikan data berikutnya berdasarkan apa yang telah dipelajarinya dari berbagai unit yang berbeda. Selama periode pelatihan, keluaran mesin dibandingkan dengan gambaran yang diberikan oleh manusia tentang apa yang harus diamati.

Jika kedua kumpulan temuan dari perbandingan itu sama, maka kesimpulannya bisa dipercaya. Namun jika salah, JST akan menyesuaikan pembelajarannya dengan kembali ke lapisan sebelumnya untuk mengubah persamaan matematika. Ini hanya akan terjadi jika jawabannya salah. Karena proses pembelajaran yang berkelanjutan ini, JST dianggap sebagai algoritma Deep Learning. Algoritma Deep Learning adalah algoritme yang semakin bijak selama penggunaannya. Seperti yang dapat dilihat pada Gambar 1, jaringan syaraf tiruan dapat memiliki banyak tingkat kerumitan.



Gambar 2. Jaringan Syaraf tiruan menggunakan multi layer

Representasi konstruksi jaringan syaraf tiruan yang menggunakan tiga jenis lapisan yang saling berhubungan ditunjukkan pada Gambar 2. Lapisan-lapisan ini adalah sebagai berikut:

- Lapisan yang dikenal sebagai lapisan input adalah lapisan yang bertanggung jawab untuk menerima data atau informasi secara langsung dari sumber yang terletak di luar sistem. Lapisan input menerima jumlah penyesuaian yang sesuai berdasarkan input.
- Lapisan yang disebut lapisan tersembunyi adalah lapisan yang terletak di antara lapisan yang disebut lapisan masukan dan lapisan yang disebut lapisan keluaran. Lapisan tersembunyi terdiri dari neuron yang mengambil informasi dari lapisan di bawahnya, yang dikenal sebagai lapisan input.
- Lapisan yang dikenal sebagai lapisan keluaran adalah lapisan yang bertanggung jawab untuk menghasilkan keluaran akhir jaringan syaraf tiruan. Penyesuaian jumlah neuron pada lapisan keluaran ini agar sesuai dengan jumlah keluaran yang diminta oleh sistem diperlukan.

Fungsi aktivasi yang disertakan di setiap lapisan jaringan saraf tiruan. Fungsi aktivasi adalah fungsi yang menentukan apakah keluaran suatu neuron linier atau nonlinier. Fungsi ini dapat mengambil berbagai bentuk. Fungsi aktivasi berikut digunakan untuk proyek penelitian ini:

- a. Fungsi aktivasi yang dikenal sebagai fungsi aktivasi ReLU adalah fungsi yang memiliki operasi matematika langsung. Tidak ada operasi yang melibatkan pertumbuhan eksponensial, perkalian, atau pembagian, dan nilai elemen apa pun yang memiliki nilai negatif diatur ulang ke 0. Fungsi aktivasi ReLU dapat diformulasikan sedemikian rupa sehingga terkandung dalam persamaan (1).

$$f(x) = \max(0, x) \text{ atau } f(x) = \begin{cases} 0 & \text{untuk } x \leq 0 \\ x & \text{untuk } x > 0 \end{cases} \quad (1)$$

- b. Fungsi Aktivasi Softmax adalah fungsi yang digunakan dalam perhitungan probabilitas untuk menentukan klasifikasi multi-kelas dengan output kelas yang memiliki nilai probabilitas tertinggi. Nama fungsi ini berasal dari fakta bahwa ia menggunakan algoritma softmax. Nilai probabilitas yang berada di antara 0 dan 1 ditetapkan ke output yang dihasilkan oleh fungsi aktivasi softmax. Persamaan (2) mewakili versi yang dinyatakan dari fungsi aktivasi softmax.

$$(X_i) = \frac{Exp(X_i)}{\sum_{k=0}^k Exp(X_k)}, \text{ nilai } i = 0, 1, 2, \dots, k \quad (2)$$

C. Confusion Matrix

Untuk mengevaluasi penelitian kali ini, penulis menggunakan *confusion matrix* yang nantinya akan digunakan untuk menghitung *precision*, *recall*, dan *accuracy*. Dalam *confusion matrix* sendiri terdiri dari *True Positive*, *True Negative*, *False Positive*, dan *False Negative*. *True Positive* mempunyai arti jika data uji diklasifikasikan benar dan diprediksi benar. Sedangkan untuk *True Negative* mempunyai arti jika data diklasifikasikan salah tetapi diprediksi benar. *False Positive* sendiri mempunyai arti jika data uji diklasifikasikan salah dan diprediksi benar sebagai data yang salah. *False Negative* mempunyai arti jika data uji diklasifikasikan salah tapi diprediksi benar.

Mengikuti uraian yang disajikan di atas, evaluasi matriks kebingungan dapat dihitung dengan menggunakan persamaan (3) untuk presisi, persamaan (4) untuk mengingat, persamaan (5) untuk akurasi, persamaan (6) untuk spesifisitas, dan persamaan (7) untuk luas di bawah kurva.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$Akurasi = \frac{TP}{Jumlah\ data} \times 100 \tag{5}$$

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

$$AUC = \frac{Recall + Specificity}{2} \tag{7}$$

D. Kebutuhan Data

Sebuah agen virtual cerdas tentunya harus dibekali dengan sebuah *knowledge base* / pengetahuan dasar agar dapat mempelajari dan mengolehnya menjadi informasi yang dibutuhkan. *Knowledge base* / pengetahuan dasar yang dimaksudkan disini merupakan data aktual yang telah terjadi, sehingga dapat digunakan sebagai pembelajaran dalam pelatihan sebuah agen virtual. Dalam membuat sebuah agen virtual yang bertujuan dapat menanggapi respon dengan baik maka dibutuhkan sebuah kemampuan mengolah bahasa alami atau biasa disebut NLP. NLP berfungsi membantu agen virtual menerjemahkan maksud dari bahasa manusia yang digunakan oleh pelanggan. Data yang dibutuhkan untuk mengolah bahasa manusia adalah menentukan intent.

Intent digunakan untuk memahami apa yang diinginkan pengguna. Ada banyak frasa / kalimat yang memiliki arti yang sama. Setiap intent akan menampung beberapa pertanyaan dan respon umum yang biasa digunakan seorang customer service dalam melayani pelanggan. Setiap Intent kemudian akan diberi label supaya memudahkan agen virtual saat mempelajarinya. Kemudian intent tersebut disusun dalam format JSON yang nantinya akan diimport agar dapat digunakan pada pengolahan data *text preprocessing*. Sehingga nantinya siap diolah dan di *training* untuk nantinya akan di uji coba menggunakan metode JST sehingga dapat menjadi prediksi jawaban dari pertanyaan yang ditanyakan oleh *user*.

Data yang telah penulis rangkum dalam sebuah intent diperoleh dari hasil data survey pertanyaan yang sering ditanyakan pelanggan pada bisnis digital printing selama 7 hari kerja. Data yang penulis survey hanyalah 6 topik pertanyaan yang nantinya akan penulis gunakan sebagai Intent pada penelitian ini yaitu kalimat yang mengandung pertanyaan atau kata-kata yang mengandung unsur: sapaan, jam operasional, harga banner, cetak stiker, cetak kertas, dan lokasi. Berikut penulis rangkum dalam Tabel 1.

TABEL I
 TABEL HASIL SURVEY JUMLAH PERNYATAAN PELANGGAN SELAMA 7 HARI

HARI, TANGGAL	JUMLAH PERTANYAAN PADA TEMA :					
	SAPA	JAM OPERASIONAL	HARGA BANNER	CETAK STIKER	CETAK KERTAS	LOKASI
Senin, 8/7/22	15	2	12	5	3	2
Selasa, 19/7/22	22	5	10	12	6	4
Rabu, 20/7/22	20	4	18	2	1	2
Kamis, 21/7/22	25	6	15	11	6	5
Jum'at, 22/7/22	15	4	10	2	4	1
Sabtu, 23/7/22	23	5	12	10	7	1
Senin, 25/7/22	23	3	10	8	5	2

E. Formulasi Kecerdasan Buatan

Pada bagian ini dijelaskan tentang penggunaan formulasi kecerdasan buatan yang digunakan pada sistem. Kecerdasan buatan digunakan untuk prediksi respon yang ditanyakan oleh pelanggan / *user*. Berikut tahapan secara singkat formulasi kecerdasan buatan Agen virtual berbasis text pada bisnis digital printing menggunakan algoritma JST yang kemudian akan diintegrasikan ke Telegram :

1. Datasets. Data yang telah penulis rangkum dalam sebuah intent diperoleh dari hasil survey kepada pelanggan bisnis digital printing selama 7 hari. Intent tersebut disusun dalam format JSON. Yang nantinya akan diimport agar dapat digunakan pada pengolahan data *text preprocessing*.
2. Text PreProcessing. Mengimpor dataset adalah langkah pertama yang harus dilakukan sebelum memulai pemrosesan data persiapan teks. Mengimpor dataset JSON yang telah disiapkan sebelumnya agar tersedia untuk digunakan dalam

pemrosesan data preprocessing teks akan dilakukan. Data tersebut akan diolah dan diolah dalam text preprocessing menggunakan library NLTK yang merupakan aplikasi natural language processing (NLP), setelah dataset diimpor. Sebagai hasil dari prosedur penyusunan teks, diperoleh 149 token, juga dikenal sebagai kata tunggal, dan 7 label, juga dikenal sebagai klasifikasi, dalam pengolahan text preprocessing ini terdapat beberapa tahapan diantaranya :

- Tahapan case folding akan membuat semua dataset akan diubah menjadi besar atau kecil. Sehingga memudahkan data teks untuk dapat diolah lebih lanjut.
- Tahap tokenization akan mengubah sebuah kalimat atau kumpulan kata menjadi per kata.
- Tahap Stemming akan mengubah kata yang telah dicacah dalam tokenization menjadi bentuk kata dasar. Misalnya, kata "memiliki" dimulai dengan awalan "mem" dan diakhiri dengan "I". Akibatnya, susunan kata tersebut terbentuk menjadi "to have +i", dan kata dasar yang mengikuti proses stemming adalah "to have" [7].
- Setelah itu, BoW akan mengembangkan kosa kata dengan menganalisis semua data teks, yang akan menghasilkan pembentukan data kata yang unik.
- Mengukur data dari kelas atau label yang berbeda dapat dilakukan dengan bantuan *One Hot Vector*. Dengan menggunakan pendekatan ini, ia menghasilkan vektor dengan panjang yang sebanding dengan jumlah total kata yang disimpan dalam kantong kata. Jika ada kata dalam data, maka komponen vektor akan diberi nilai 1, tetapi untuk kata yang tidak ada dalam data, komponen vektor akan diberi nilai 0.

Setelah impor, data akan diproses menggunakan *library* NLTK, yang mengimplementasikan NLP. Setelah itu, informasi tersebut akan diubah menjadi matriks BoW dalam bentuk array, dan produk akhir dari transformasi tersebut adalah bilangan bulat 0 dan 1. Jika kata dari pola saat ini juga muncul dalam pola yang diminta, maka nomor dalam array akan memiliki nilai 1, jika tidak muncul dalam pola yang diperlukan, maka akan memiliki nilai 0. set. Informasi tersebut akan dipisahkan menjadi dua variabel, yang akan disebut sebagai *train_x* dan *train_y*. Dimana *train_x* akan menyimpan isi matriks sebagai token atau kata tunggal, dan *train_y* akan menyimpan isi matriks sebagai label atau kategori.

3. Metode Jaringan Syaraf Tiruan (JST)

Penelitian ini menggunakan 3 lapisan dalam proses pelatihan data untuk pembuatan *chatbot*. Lapisan pertama disebut lapisan input, dan data yang diperlukan untuk itu dikumpulkan dari langkah *preprocessing* teks yang muncul setelah tahap *preprocessing* teks.

Berdasarkan dataset JSON yang telah diolah sebelumnya, didapatkan 149 *neuron* untuk digunakan sebagai input data training. Data dihasilkan karena prosedur text preprocessing, yang menghasilkan 149 token dan tujuh kelas atau kategori. Di dalam lapisan kedua, yaitu lapisan tersembunyi, 21 *neuron* memiliki fungsi aktivasi *ReLU*. Setelah itu muncul lapisan ketiga yang juga merupakan lapisan tersembunyi dan berisi 16 *neuron* dengan fungsi aktivasi *ReLU*. Lapisan keempat adalah lapisan keluaran, terdiri dari tujuh neuron yang diaktifkan oleh fungsi aktivasi *Softmax*. Selama proses training, *Dense Layer* yang merupakan bagian dari *hidden layer* akan bertanggung jawab untuk memproses input yang terdiri dari 149 *neuron*. Fungsi Padat menambahkan Lapisan Terpasang Sepenuhnya, sebuah lapisan di mana semua aktivitas neuron sebelumnya digabungkan ke neuron pada lapisan di atasnya.

Untuk mendapatkan nilai akurasi yang sebesar mungkin dan nilai error yang serendah mungkin, maka proses training data akan dilakukan secara iteratif dan berulang. Representasi grafis dari data akan menampilkan informasi tentang keakuratan dan kerugian yang terjadi selama proses pelatihan untuk setiap zaman. Selama fase yang dikenal sebagai "pelatihan data", model yang dibangun sebelumnya akan dikompilasi menggunakan fungsi kehilangan "*Categorical Crossentropy*" dan pengoptimal SGD (*Stochastic Gradient Descent*). Untuk menyelesaikan proses pelatihan data, kami menggunakan kriteria penghentian yang didasarkan pada 300 epoch.

4. Integrasi Telegram

Agen Virtual berbasis text pada bisnis digital printing akan di implementasikan ke aplikasi Telegram menggunakan API. Berikut langkah yang digunakan untuk menghubungkan Agen virtual berbasis text ke aplikasi Telegram :

a) Membuat Bot pada BotFather Telegram

Bot Father adalah bot resmi dari telegram yang tugasnya melahirkan atau menciptakan bot baru. Caranya adalah sebagai berikut :

1. Buka Bot Father dengan mencarinya pada pencarian telegram.
2. Ketik /start untuk memunculkan menu- menu yang disediakan oleh Bot Father Telegram.
3. Pilih menu (dengan klik pada bagian menu tersebut) atau ketikkan menu /newbot.
4. Ketikkan nama untuk Bot yang akan digunakan. Disini penulis menggunakan nama panggilan bot "ArthurACM".

b) Mendapatkan Token API Telegram

Setelah membuat nama panggilan bot, Bot father akan meminta untuk menuliskan nama username yang akan digunakan. Syaratnya adalah menambahkan "Bot" pada bagian belakang username bot yang akan dibuat. Jika username belum pernah digunakan dan disetujui oleh pihak Telegram, maka akan secara otomatis dikirimkan token API Telegram yang siap digunakan.

Disini penulis membuat username "ArthurACMbot". Yang Kemudian secara otomatis akan mendapatkan token API yang siap digunakan.

c) Menginstal Library python-telegram-bot

Untuk menghubungkan bahasa pemrograman python dengan telegram sebagai interface bot, dibutuhkan sebuah library python yang bernama python-telegram-bot. Menghubungkan Agen Virtual berbasis text dengan Telegram menggunakan API Telegram yang sudah diperoleh dari Bot father.

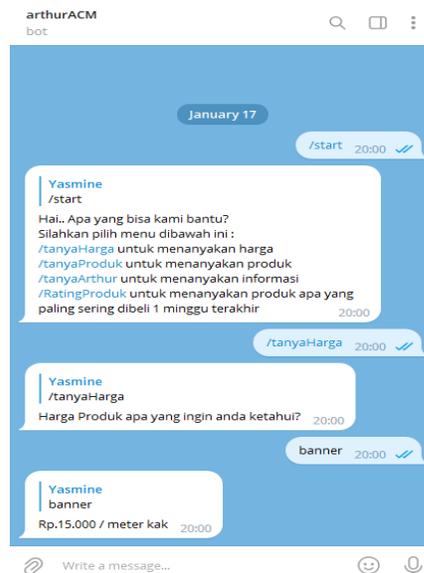
III. HASIL DAN PEMBAHASAN

Pada bagian ini akan dijelaskan hasil uji coba dari perancangan dan pembuatan agen virtual berbasis text pada bisnis digital printing menggunakan jaringan syaraf tiruan. Dalam melakukan uji coba penulis membagi menjadi tiga uji coba yaitu *black box testing*, *white box testing* dan uji penerimaan.

A. Black Box Testing

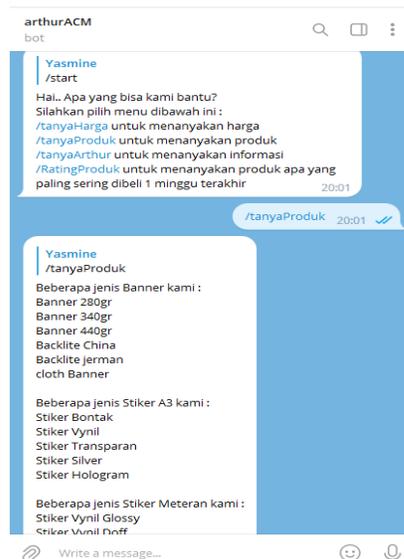
Pengujian black box merupakan salah satu cara pengujian perangkat lunak berguna untuk menentukan apakah aplikasi tersebut sudah berjalan sesuai dengan yang diharapkan atau tidak. Berikut merupakan pengujian tanya jawab oleh user dan agen virtual berbasis teks.

Menu Tanya Harga. Menu ini digunakan sebagai pilihan menanyakan harga sebuah produk. Pengujian menu ini akan dijelaskan pada Gambar 3.



Gambar 3. Menu Tanya Harga

Menu Tanya Produk. Menu ini digunakan sebagai pilihan menanyakan jenis- jenis produk pada CV. Arthur Citra Media. Pengujian menu ini akan dijelaskan pada Gambar 4.



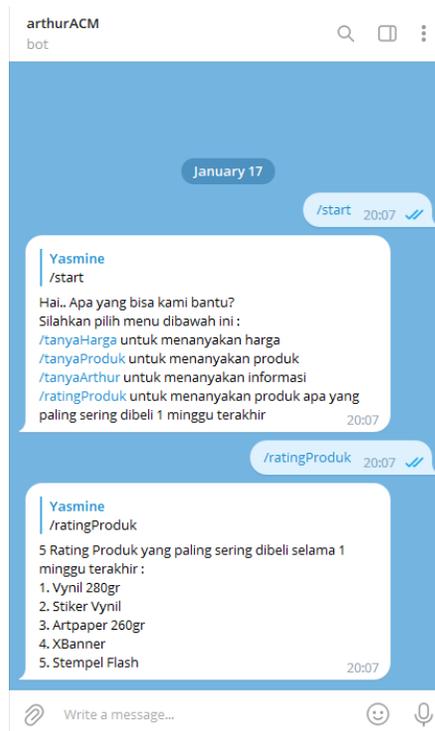
Gambar 4. Menu Tanya Produk

Menu Tanya Arthur. Menu ini digunakan sebagai opsi menanyakan Informasi seputar CV. Arthur Citra Media. Pengujian menu ini akan penulis dijelaskan pada Gambar 5. Menampilkan gambar ketika *user* memilih atau mengetik “/tanyaArthur” maka prediksi jawaban yang seharusnya keluar adalah informasi tentang CV. Arthur Citra Media, seperti lokasi, cara pemesanan, nomor telepon, dan lain sebagainya.



Gambar 5. Menu Tanya Arthur

Menu Rating Produk. Menu ini digunakan sebagai pilihan menanyakan informasi seputar peratingan produk yang sering dibeli selama 1 minggu terakhir. Pengujian menu ini akan dijelaskan pada Gambar 6. Menampilkan gambar ketika *user* memilih atau mengetik “/ratingProduk” maka prediksi jawaban yang seharusnya keluar adalah peratingan produk yang sering dibeli selama 1 minggu terakhir.



Gambar 6. Menu Rating Produk

Agen virtual ini juga dibekali kemampuan untuk tetap menjawab user dengan kalimat “Maaf, kami tidak mengerti apa yang anda maksud. Mohon perbaiki penulisan anda atau mungkin kami tidak menjual barang yang anda tanyakan.” ketika agen virtual tersebut tidak mengetahui apa yang dimaksud oleh *user*. Seperti yang akan ditampilkan pada Gambar 7. Menampilkan gambar ketika agen virtual menjawab dengan permintaan maaf apabila agen virtual tidak mengetahui apa yang ditanyakan *user*.



Gambar 7 Agen Virtual Menjawab Dengan Permintaan Maaf Apabila Agen Virtual Tidak Mengetahui Apa Yang Ditanyakan User.

B. White Box Testing

White box testing merupakan cara menguji algoritma yang digunakan dalam suatu aplikasi apakah terdapat kesalahan atau tidak, tanpa memperdulikan tampilan aplikasi. Pengujian akan dilakukan dengan menguji 10 kalimat pertanyaan atau percakapan untuk melihat apakah sistem chatbot dapat menjawab dengan sesuai atau tidak. Dari 25 kalimat pertanyaan didapatkan 25 kalimat percakapan yang berhasil menampilkan respon dengan benar dan sesuai dan terdapat 1 kalimat percakapan yang menampilkan respon tidak sesuai. Berikut ini merupakan daftar hasil pengujian dari 25 pertanyaan ditampilkan pada Tabel II.

TABEL II
 TABEL HASIL PENGUJIAN AGEN VIRTUAL BERBASIS TEXT PADA BISNIS DIGITAL PRINTING

INPUTAN USER	PREDIKSI RESPON	RESPON SEHARUSNYA	KET
Hello	Arthur disini, Apa ada yang bisa kami bantu?	Arthur disini, Apa ada yang bisa kami bantu?	TP
Hi	Arthur disini, Apa ada yang bisa kami bantu?	Arthur disini, Apa ada yang bisa kami bantu?	TP
Permisi	Arthur disini, Apa ada yang bisa kami bantu?	Arthur disini, Apa ada yang bisa kami bantu?	TP
Selamat Pagi	Arthur disini, Apa ada yang bisa kami bantu?	Arthur disini, Apa ada yang bisa kami bantu?	TP
Selamat Siang	Arthur disini, Apa ada yang bisa kami bantu?	Arthur disini, Apa ada yang bisa kami bantu?	TP
Selamat Sore	Arthur disini, Apa ada yang bisa kami bantu?	Arthur disini, Apa ada yang bisa kami bantu?	TP
Selamat Malam	Arthur disini, Apa ada yang bisa kami bantu?	Arthur disini, Apa ada yang bisa kami bantu?	TP
Lokasi toko ada dimana ya?	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	TP
Nanti saya kesana	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	TP
Apa bisa diambil gojek?	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	TP

INPUTAN USER	PREDIKSI RESPON	RESPON SEHARUSNYA	KET
Apa bisa diambil grab?	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	TP
Boleh minta alamatnya?	Kami buka Senin- Sabtu (Jam 09,00-18,00)	Kami buka Senin- Sabtu (Jam 09,00-18,00)	TP
Buka sampai jam berapa?	Kami buka Senin- Sabtu (Jam 09,00-18,00)	Kami buka Senin- Sabtu (Jam 09,00-18,00)	TP
Apa besok buka?	Kami buka Senin- Sabtu (Jam 09,00-18,00)	Kami buka Senin- Sabtu (Jam 09,00-18,00)	TP
Hari ini buka?	Kami buka Senin- Sabtu (Jam 09,00-18,00)	Kami buka Senin- Sabtu (Jam 09,00-18,00)	TP
Apa nanti buka?	Kami buka Senin- Sabtu (Jam 09,00-18,00)	Kami buka Senin- Sabtu (Jam 09,00-18,00)	TP
Banner per meter berapa ya?	15rb/ m	15rb/ m	TP
Spanduk ukuran 1 meter berapa?	15rb/ m	15rb/ m	TP
Kak, mau pesan stiker bisa?	Bisa kak	Bisa kak	TP
Mau cetak stiker Oneway	Bisa kak	Bisa kak	TP
Bisa cetak stiker meteran?	Bisa kak	Bisa kak	TP
Bisa cetak stiker A3?	Bisa kak	Bisa kak	TP
Kak, mau pesan artpaper bisa?	Bisa kak	Bisa kak	TP
Bisa cetak menu makanan?	Bisa kak	Lokasi kami berada di Jl. Bratang Binangun No. 9 Surabaya	FP
Bisa cetak HVS A3?	Bisa kak	Bisa kak	TP

Dari Tabel II maka dapat dibuat *confusion matrix multiclass* seperti Tabel III.

TABEL III
 TABEL CONFUSION MATRIX MULTICLASS TAG INTENT

DATA UJI	DATA PREDIKSI					
	SAPA	JAM OPERASIONAL	HARGA BANNER	CETAK STIKER	CETAK KERTAS	LOKASI
Greeting	7	0	0	0	0	0
JamOperasional	0	5	0	0	0	0
Lokasi	0	0	5	0	0	1
hargaBanner	0	0	0	2	0	0
cetakStiker	0	0	0	0	4	0
cetakKertas	0	0	0	0	0	1

Keterangan :

- Data Benar
- Data Salah

Sedangkan untuk akurasi dan AUC dari keseluruhan pengujian data dan testing dapat di implementasikan seperti perhitungan berikut:

$$\begin{aligned} \text{Akurasi} &= \frac{\text{TP}}{\text{Total Dataset}} \\ &= \frac{7+5+5+2+4+1}{25} \times 100\% = 96\% \end{aligned}$$

$$\begin{aligned} \text{AUC} &= \frac{\text{Recall} + \text{Specificity}}{2} \\ &= \frac{97.2\% + 99.3\%}{2} = 98.25\% \end{aligned}$$

IV. KESIMPULAN

Algoritma JST dapat diimplementasikan untuk membangun agen virtual berbasis text sehingga mampu memberikan respon yang sesuai dengan inputan pengguna dengan tingkat akurasi hasil dari uji coba sebesar 96%. Tingkat akurasi dapat dipengaruhi dengan banyaknya sampel data, dimana semakin banyak data yang dimiliki maka tingkat akurasi akan semakin tinggi. Agen Virtual ini akan di implementasikan ke aplikasi Telegram menggunakan API, sehingga dapat lebih mudah digunakan oleh pengguna utama agen virtual ini, yaitu customer.

DAFTAR PUSTAKA

- [1] Saharja, fatimah dan Siti Aisyah. 2020. *Efektifitas Digital Printing (Percetakan Digital) dalam menghasilkan Produk Cetak dan pengaruhnya terhadap konsumen. Vol.14 No.11.*
- [2] Liddy, E.D. 2001. *Natural Language Processing.* Enclopedia of Library dan Information Science, 2nd Edition, Marcel Decker Inc, NY, USA.
- [3] A. V. Vitianingsih, D. Cahyono, and A. Choiron, "Analysis and design of web-geographic information system for tropical diseases-prone areas: A case study of East Java Province, Indonesia," in 2017 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), 2017, pp. 255–260.
- [4] A. V. Vitianingsih, A. Choiron, D. Cahyono, A. Umam, and S. Suyanto, "Spatial data modeling on GIS for classification of measles-prone region using multiple attribute decision making," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 3, pp. 97–107, 2019.
- [5] Bijaksana, A. Negara, P. Safriadi, N. Studi, P. Informatika, T. Tanjungpura, U. 2017. *Prediksi Jeda Pada Ucapan Bahasa Melayu Pontianak Dengan Menggunakan Metode Shallow Parsing.*
- [6] Guzman, A. L. 2019. *Voices in and of the machine: Source orientation toward mobile virtual assistants. Computers in Human Behavior.*
- [7] Cahyono, D., Fadlil, J., Sumpeno, S., & Hariadi, M. 2008. *Temu Kembali Informasi untuk Pembangkitan Basis Pengetahuan dari Teks Bebas yang Digunakan oleh Agen Percakapan Bahasa Alami.*