



Algoritma *Artificial Neural Network* pada *Text-based Chatbot Frequently Asked Question (FAQ) Web* Kuliah Universitas Nasional

Feri Mustakim ^{*1}, Fauziah ², Nur Hayati ³

^{1,2,3} Fakultas Teknologi Komunikasi dan Informatika, Universitas Nasional

article info

Article history:

Received 27 December 2020

Received in revised form

4 February 2021

Accepted 13 February 2021

Available *online* October 2021

DOI:

<https://doi.org/10.35870/jtik.v5i4.261>

Keywords:

Chatbot, Artificial Neural Network, ANN, Text-based.

Kata Kunci:

Chatbot, Artificial Neural Network, ANN, Berbasis teks.

abstract

The technology development increase the number of automation system in industry. One of them is *Chatbot* application in education industry. This automation technology is able to lessen university's service in order to facilitate the students' need of information whenever and wherever they are. Lack of student literacy regarding the functions and use of the *web* in conducting online lectures causes the same number of questions repeatedly to the university, which are actually frequently asked questions that have been written in a list of frequently asked questions (FAQ), such as: assignment submission, forget *passwords*, lectures online, *video conference* lectures and lecture *web* applications on android. *Chatbot* will automatically answer students' question in university *web* page by providing information and suggesting a proper answer suit to the question. This research will develop *Chatbot* type based on text by applying Artificial Neural Network (ANN) algorithm. The applied data set while conducting the *Chatbot* coaching is the questions data which frequently being asked (FAQ) in the study *web*, 25 questions with its answer which is divided into 16 labels or classes. The testing is conducted by using 110 different conversations from the *dataset* but have the same intention. From those 110 conversation, the *Chatbot* succeed in answering 107 questions precisely and made 3 wrong conversation. The testing result shows a good result by having 97,27% accuracy and 2,72% error.

abstract

Kemajuan teknologi membuat banyak otomatisasi dalam dunia industri, Salah satunya penerapan *Chatbot* pada industri pendidikan. Dengan teknologi otomatisasi ini, Memudahkan universitas untuk melayani mahasiswa kapanpun waktunya dan dimanapun tempatnya. Kurangnya literasi mahasiswa terhadap fungsi dan penggunaan *web* kuliah dalam melaksanakan kuliah online menyebabkan banyaknya pertanyaan yang sama secara berulang kepada pihak universitas yang sebenarnya pertanyaan yang sering ditanyakan sudah ditulis dalam daftar pertanyaan yang sering muncul (FAQ), seperti: pengumpulan tugas, lupa *password*, kuliah online, kuliah *video conference* dan aplikasi *web* kuliah di android. Dengan menggunakan *Chatbot* secara otomatis akan menjawab pertanyaan mahasiswa dilaman *web* kuliah dengan memberikan informasi dan mengarahkan sesuai dengan jawaban pertanyaan. Dalam penelitian ini akan dikembangkan model *Chatbot* berbasis teks dengan menggunakan algoritma Artificial neural network (ANN). *Dataset* yang digunakan untuk melakukan pelatihan *Chatbot* merupakan data pertanyaan sering timbul (FAQ) didalam *web* kuliah yaitu 25 pertanyaan beserta jawaban yang dibagi kedalam 16 label atau kelas. Pengujian dilakukan dengan menggunakan 110 percakapan yang berbeda dengan *dataset* tetapi mempunyai maksud yang sama. Dari 110 percakapan *Chatbot* berhasil menjawab dengan tepat sebanyak 107 percakapan dan jawaban salah sebanyak 3 percakapan. Hasil pengujian menunjukkan hasil yang baik yaitu mempunyai akurasi 97,27 % dan kesalahan 2,72 %.

*Corresponding author. Email: ferimustakim@gmail.com¹.

© E-ISSN: 2580-1643.

Copyright © 2021. Published by Lembaga Informasi dan Riset (KITA INFO dan RISET), Lembaga KITA (<http://creativecommons.org/licenses/by/4.0/>).

1. Latar Belakang

Kurangnya literasi mahasiswa terhadap fungsi dan penggunaan *web* kuliah dalam melaksanakan kuliah online menyebabkan banyaknya pertanyaan yang sama secara berulang kepada pihak universitas yang sebenarnya pertanyaan yang sering ditanyakan sudah ditulis dalam daftar pertanyaan yang sering muncul (Faq), seperti: pengumpulan tugas, lupa *password*, kuliah online, kuliah *video conference* dan aplikasi *web* kuliah di android. Adapun tujuan yang ingin dicapai dalam penelitian adalah pengembangan model *Chatbot* dengan algoritma *artificial neural network* (ANN) untuk menjawab pertanyaan yang sering ditanyakan tentang penggunaan *web* kuliah Universitas.

Chatbot adalah suatu program kecerdasan buatan yang berbentuk simulasi percakapan interaktif antara mesin dengan manusia melalui teks, suara dan visual atau gambar. Dalam mengenali dan memberikan respon layaknya seperti percakapan manusia, *Chatbot* sangat bergantung dengan data pengetahuan yang sudah dibuat atau masukan dari pengembang sistem. Untuk meningkatkan akurasi sistem dapat menggunakan cara pemberian label pada setiap pola kalimat pada *dataset*. (Bowen Wu, dkk) [1].

Bowen dkk, dalam penelitiannya melakukan penggabungan antara lapisan konvolusional dengan memahami petunjuk semantik dalam percakapan, sehingga dapat memodelkan semantik dari urutan ucapan berdasarkan representasi efektif untuk setiap kalimat. Penelitian ini lebih menekankan ke bagian kata-kata penting dalam percakapan, sehingga representasi konteks cenderung lebih bermakna dan dapat meningkatkan kandidat jawaban. Dalam penelitian algoritma yang digunakan adalah CNN dan RNN untuk mengembangkan model peringkat kandidat jawaban. Pada pengembangan model ini didapat nilai akurasi yang cukup baik yaitu 78,6% [1].

Penelitian dengan algoritma *sequential matching network* (SMN) dilakukan oleh Wu dkk. SMN pertama digunakan untuk mencocokkan respon dengan setiap kalimat dalam konteks pada berbagai tingkat perincian, dan menyaring informasi penting yang cocok dari setiap pasangan sebagai vektor dengan operasi konvolusi dan pooling. Dengan menggunakan recurrent neural network (RNN)

vektor diakumulasikan dalam model urutan hubungan antar ucapan [2].

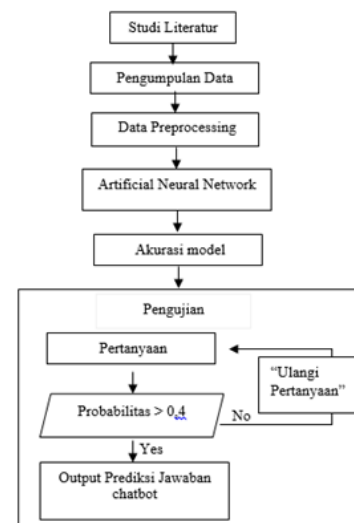
Yogi Wisesa Candra dan Suyanto juga melakukan penelitian *Chatbot* dengan model *sequence to sequence*, *dataset* yang digunakan merupakan pesan percakapan whatsapp panitia penerimaan mahasiswa dengan calon mahasiswa. Hasil penelitian menghasilkan model dengan skor BLUE yang cukup baik 41,6 % [3].

Uli Rizki melakukan penelitian tentang multi respon jawaban dalam sebuah *Chatbot*. Peneliti menggunakan perangkikan terhadap respon untuk mengatasi masalah multi respon tersebut. Peneliti menggunakan dua algoritma yaitu *sequence to sequence* dan CNN. Algoritma *sequence to sequence* digunakan untuk perangkikan pasangan antara pertanyaan dan jawaban, sedangkan algoritma CNN untuk menjaga konsistensi jawaban. Dari penelitian ini didapatkan hasil akurasi dengan nilai 86,7% [4].

Berdasarkan referensi hasil penelitian yang disebutkan sebelumnya diatas peneliti menjadikan sebagai acuan dalam penelitian ini. Fokus penelitian yaitu bagaimana mengembangkan sebuah *Chatbot* dengan algoritma artificial neural network (ANN). Penelitian ini akan menggunakan *dataset* pertanyaan yang sering ditanyakan oleh mahasiswa (Faq) tentang penggunaan *web* kuliah yang diambil dari halaman *web* kuliah Universitas dan diubah dalam bentuk file JSON.

2. Metode Penelitian

Kerangka Penelitian



Gambar 1. Alur Penelitian

Pada gambar 1 merupakan alur penelitian yang telah dilakukan dengan melalui proses studi literatur, pengumpulan data, *data preprocessing*, model *artificial neural network*, akurasi model, dan pengujian. Adapun penjelasan dari alur penelitian yang telah dilakukan antara lain:

1) Studi Literatur

Studi literatur yang dilakukan penulis adalah membaca sumber-sumber tertulis seperti jurnal ilmiah dari penelitian terdahulu dan referensi lainnya yang berguna sebagai dasar acuan melakukan penelitian yang dilakukan sekarang.

2) Metode Pengumpulan Data

Dalam penelitian ini, data yang digunakan adalah kumpulan data dari pertanyaan yang sering timbul atau ditanyakan oleh mahasiswa di *web* kuliah tentang penggunaan aplikasi *web* kuliah.

3) Data Pre-processing

Data pre-processing disini adalah proses pengolahan *text* sebelum digunakan kedalam algoritma ANN. Proses pengolahan *text* menggunakan teknik dari *natural language processing* (NLP) atau disebut dengan pengolahan bahasa alami manusia. Teknik ini menerjemahkan Bahasa manusia menjadi bahasa yang dimengerti oleh komputer. Teknik yang digunakan meliputi:

a. Case Folding

Mengubah semua data yang digunakan sebagai masukan (*corpus* atau *dataset*) menjadi huruf besar atau kecil. Ini akan menghindari kesalahan penafsiran kata yang salah jika dieja dengan huruf besar atau kecil.

b. Tokenization

Mengubah suatu kalimat (kumpulan kata) menjadi perkata.



Gambar 2. Mengubah Kalimat menjadi perkata

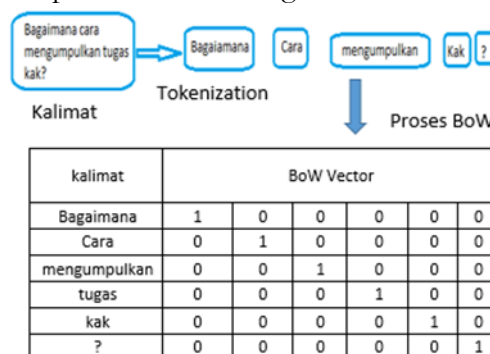
c. Stemming

Merupakan proses mengembalikan kata-kata menjadi bentuk kata dasar. Proses *stemming* membantu mengurangi kantong kata-kata dengan mengaitkan kata-kata yang mirip dengan kata-kata dasar yang sesuai. Contoh: mengumpulkan menjadi kumpul,

bagaimanakah menjadi bagaimana, perbedaan menjadi berbeda, dan lain-lain.

d. Bag of Words (BoW)

BoW merupakan pemodelan yang mempelajari sebuah kosakata dari seluruh data teks atau *corpus* sehingga terbentuk data kata yang unik. Kemudian memodelkan setiap data dengan menghitung jumlah kemunculan setiap kata (Deepu, Pethuru, & Rajaraajeswari, 2016). Berikut contoh teks yang direpresentasikan sebagai BoW.



Gambar 3. Proses penghitungan frekuensi kata dalam data.

e. One-Hot Vector

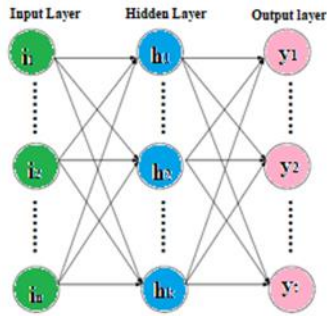
Metode *one-hot vector* berguna untuk mengukur data dari kelas/label. Dalam metode ini menghasilkan vektor dengan panjang sama dengan jumlah kata yang terdapat dalam *bag of words*. Jika suatu kata termasuk didalam data maka komponen vektor akan diberi nilai 1 dan nilai 0 untuk yang tidak terdapat dalam data.

Tabel 1. One-hot vector

Kalimat	One Hot Vector					
	Bagaimana	Cara	kumpul	tugas	kak	?
Bagaimana	1	0	0	0	0	0
Cara	0	1	0	0	0	0
mengumpulkan	0	0	1	0	0	0
tugas	0	0	0	1	0	0
kak	0	0	0	0	1	0
?	0	0	0	0	0	1

4) Artificial Neural Network (ANN)

Artificial neural network adalah suatu model yang tersusun dari neuron buatan dan saling terhubung yang mempunyai karakteristik menyerupai jaringan syaraf biologis otak manusia yang cara kerjanya dengan diberikan rangsangan / *input*, melakukan proses, kemudian menghasilkan *output*.



Gambar 4. *Artificial neural network* dengan menggunakan *multilayer*

Dari gambar 4 merupakan struktur dari *artificial neural network* dengan menggunakan 3 jenis layer yang saling terhubung, yaitu:

a. *Input Layers*

Input layer adalah lapisan yang bertugas menerima input / masukan langsung dari luar sistem. *Input layer* disesuaikan dengan jumlah input.

b. *Hidden Layer*

Hidden layer adalah lapisan yang terletak diantara input layer dan output layer. *Hidden layer* terdiri dari *neuron – neuron* yang menerima data dari *input layer*.

c. *Output Layer*

Output layer adalah lapisan yang menghasilkan output akhir dari jaringan *artificial neural network*. Pada *layer output* ini jumlah *neuron* harus disesuaikan dengan jumlah *output* yang diinginkan oleh sistem.

Setiap *layer* pada *artificial neural network* memiliki fungsi aktivasi. Fungsi aktivasi merupakan fungsi yang menentukan keluaran suatu *neuron* yang berbentuk *linear* atau *nonlinear*. Dalam penelitian ini fungsi aktivasi yang dipakai adalah sebagai berikut:

a. Fungsi aktivasi Relu

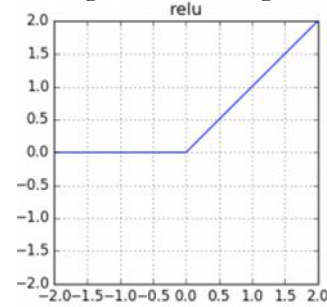
Fungsi *Rectified Linear Unit* (ReLU) merupakan fungsi aktivasi yang memiliki perhitungan sederhana. Jika ada elemen yang bernilai negatif maka nilai diubah menjadi 0, tidak ada operasi eksponensial, perkalian atau pembagian. Fungsi aktivasi relu dapat di rumuskan sebagai berikut (1):

$$f(x) = \max(0, x) \quad (1)$$

atau

$$f(x) = \begin{cases} 0 & \text{untuk } x \leq 0 \\ x & \text{untuk } x > 0 \end{cases}$$

Berikut ini bentuk grafik dari fungsi aktivasi relu :



Gambar 5. Grafik ReLU

Berdasarkan gambar 5 grafik relu menunjukkan apabila nilai masukan bernilai negatif maka dengan fungsi aktivasi relu akan mengubah nilainya menjadi nol.

b. Fungsi aktivasi Softmax

Fungsi aktivasi softmax adalah fungsi yang berguna dalam menghitung probabilitas untuk menentukan klasifikasi multi kelas dengan output kelas yang memiliki nilai probabilitas yang paling tinggi. Fungsi aktivasi softmax menghasilkan output yang memiliki nilai antara 0 sampai 1. Fungsi aktivasi softmax dapat dirumuskan seperti berikut (2):

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

Pada penelitian pengembangan model *Chatbot* dengan *artificial neural network* terdapat juga beberapa parameter yang digunakan dalam pelatihan data, diantaranya:

a. *Epoch*

Epoch adalah satu siklus pengulangan dari proses belajar algoritma *artificial neural network* pada seluruh data pelatihan. Satu *epoch* dapat diartikan bahwa *artificial neural network* telah melakukan proses belajar dari data pelatihan secara keseluruhan. Dalam *artificial neural network*, proses pembelajaran yang dilakukan secara berulang bertujuan untuk mencapai nilai *error* dan akurasi yang sebaik mungkin.

b. *Batch Size*

Batch Size merupakan jumlah sampel data yang disebar dalam proses pelatihan oleh *Neural Network*.

c. *RMSProp*

Root Mean Square Propagation (RMSProp)

merupakan algoritma pengoptimal yang mengatur berapa nilai dari *learning rate* berdasarkan besaran nilai rata-rata dari *weight*. RMSProp dalam menentukan nilai rata-rata dari *weight* menggunakan nilai pertama dari *gradient* (Duchi, Hazan, Elad, Singer, & Yoram, 2011).

- d. *Mean Squared error* (MSE)
Mean Squared Error (MSE) merupakan salah satu metode yang dapat mengevaluasi metode peramalan. Masing-masing kesalahan atau sisa dikuadratkan. Kemudian dijumlahkan dan hasilnya akan ditambahkan lagi dengan jumlah observasi. Metode ini dapat mengatur kesalahan peramalan yang besar karena kesalahan-kesalahan itu dikuadratkan.

- 5) Akurasi Model
Cara mendapatkan akurasi model yang bagus penulis melakukan proses pelatihan data secara berulang kali secara bertahap sehingga didapatkan hasil nilai kesalahan yang semakin kecil dan nilai akurasi yang paling besar.

- 6) Pengujian
Proses pengujian yang dilakukan penulis adalah mencocokkan apakah *output* yang di keluarkan sudah sesuai dengan yang diharapkan

3. Hasil dan Pembahasan

Dataset

Dataset dibuat dengan mengambil daftar pertanyaan yang sering muncul (Faq) di *web* kuliah.



Gambar 6. Daftar pertanyaan yang sering timbul.

Kemudian data tersebut diubah kedalam bentuk file *.json*. Berikut tampilan file dalam bentuk *.JSON*.



Gambar 7. Tampilan file dalam bentuk *.json*

Dataset dibuat secara manual kedalam bentuk file *json* dengan rancangan sebagai berikut:

- a. Intents: kumpulan semua data input dan output yang digunakan untuk melatih *Chatbot*
- b. Patterns: berisi data pola *input* yang diinginkan pengguna.
- c. Responses: berisi data pola *output* yang dikirimkan *Chatbot* kepada pengguna
- d. Tag: mengelompokkan data teks yang serupa dan menggunakan yang sama sebagai keluaran yang ditargetkan untuk melatih jaringan neural.

Data Preprocessing

Sebelum melakukan data *preprocessing dataset* harus diimport terlebih dahulu, dengan cara sebagai berikut di gambar 8:

```
import json
with open('intents.json') as file:
    data = json.load(file)
```

Gambar 8. Import *dataset* yang akan digunakan

Setelah diimport selanjutnya dilakukan ekstrak *dataset* untuk dilakukan *stemming*, pelabelan, pembuatan token, dan daftar pertanyaan semua dalam bentuk array. Hasil dari ekstrak *dataset* didapatkan 47 token atau kata unik, 16 labels atau nama kelompok pertanyaan beserta jawaban, 25 daftar pertanyaan.

47 Token atau kata unik ['antara', 'apa', 'apakah', 'aplikasi', 'app', 'assalamualikum', 'bagaimana', 'beda', 'cara', 'confere nce', 'dengan', 'di', 'file', 'guna', 'hai', 'hallo', 'install', 'itu', 'jika', 'jumpa', 'kak', 'kasih', 'kata', 'kuliah', 'kum pul', 'lupa', 'maksud', 'malam', 'mana', 'menginstall', 'mengupload', 'mobile', 'online', 'pagi', 'password', 'private', 'sampa i', 'sandi', 'saya', 'selamat', 'siang', 'sore', 'tanya', 'terima', 'tugas', 'web', 'yang']
16 Kelompok pertanyaan ['halo', 'install', 'kuliah online', 'malam', 'mobile app', 'pagi', 'password', 'perbedaan', 'private f ile', 'salam', 'sampai jumpa', 'siang', 'sore', 'terima kasih', 'tugas', 'web conference']
25 Daftar Pertanyaan [['Assalamualikum', 'Kak'], ['Hai', 'kak'], ['Hallo', 'kak'], ['selamat', 'pagi', 'kak'], ['selamat', 'si ang', 'kak'], ['selamat', 'sore', 'kak'], ['selamat', 'malam', 'kak'], ['selamat', 'kasih', 'kak'], ['Sampai', 'jumpa', 'kak'], ['Dimana', 'saya', 'bertanya', 'jika', 'lupa', 'password'], ['Lupa', 'kata', 'sandi'], ['Bagaimana', 'cara', 'mengupload', 'tug as', 'di', 'web', 'kuliah', '?'], ['Bagaimana', 'cara', 'menggumpulkan', 'tugas', 'di', 'web', 'kuliah', '?'], ['Apakah', 'yan g', 'dinaksud', 'dengan', 'web', 'conference'], ['Bagaimana', 'cara', 'penggunaan', 'web', 'conference'], ['apa', 'itu', 'perku liahan', 'online'], ['apa', 'yang', 'dinaksud', 'dengan', 'kuliah', 'online', '?'], ['apa', 'perbedaan', 'antara', 'web', 'conf erence', 'dengan', 'web', 'kuliah'], ['apa', 'kegunaan', 'private', 'file', '?'], ['apa', 'itu', 'mobile', 'app', '?'], ['apa', 'itu', 'aplikasi', 'mobile'], ['Bagaimana', 'cara', 'menginstall', 'mobile', 'app', '?'], ['bagaimana', 'cara', 'penggunaan', 'mobile', 'app', '?'], ['cara', 'install', 'aplikasi', 'mobile'], ['cara', 'menggunakan', 'aplikasi', 'mobile']]

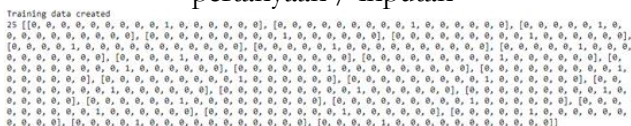
Gambar 9. Hasil ekstrak *dataset*

Training Dataset

Proses *training* atau pelatihan dilakukan dengan cara memberikan *input* dan *output*, kemudian mengubah / mengkonversi semua data *input* dan *output* menjadi angka berbentuk array. Hasil dari konversi tersebut panjang array akan sama dengan ukuran banyak kata unik, angka 1 akan ditetapkan ketika sebuah kata dari pola saat ini berada di pattern yang ditetapkan. Seperti gambar 10 dan 11.



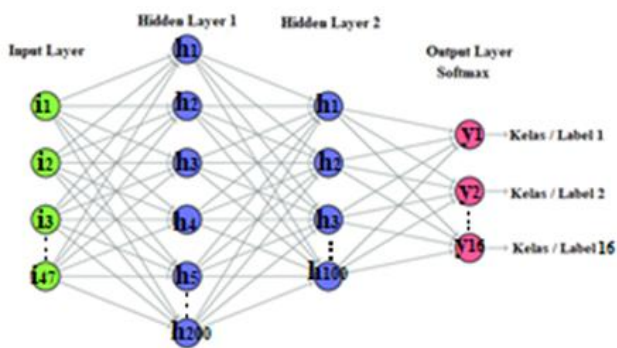
Gambar 10. Konversi pelatihan data untuk daftar pertanyaan / inputan



Gambar 11. Konversi pelatihan data untuk daftar label / output

Model Artificial Neural Network (ANN)

Dalam pengembangan model *Chatbot* dengan algoritma ANN memiliki 4 *layers*, pertama *input layer* yang memiliki *neuron* sama dengan jumlah *input* 47 *neuron*, *layer* kedua merupakan *hidden layer* yang memiliki 200 *neuron*, fungsi aktivasi ReLu dan *batch size* 30, yang ketiga juga merupakan *hidden layer* memiliki 100 *neuron* dan menggunakan fungsi aktivasi ReLu, dan yang keempat merupakan *output layer* yang memiliki jumlah *neuron* sama dengan jumlah *output* yaitu 16 dan menggunakan fungsi aktivasi softmax.



Gambar 12. Arsitektur ANN dengan 4 layer (Sumber gambar: Dexlab)

Proses pelatihan data dilakukan secara berulang kali secara bertahap sehingga didapatkan hasil nilai kesalahan yang semakin kecil dan nilai akurasi yang paling besar. Berikut tabel tahapan pelatihan dengan nilai *epoch* dan *neuron* yang berbeda.

Tabel 2: Hasil pelatihan data dengan nilai *epoch* dan jumlah *neuron* berbeda.

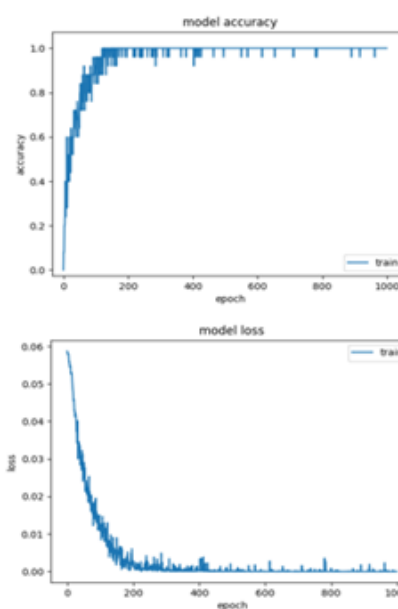
Epoch	Hidden layer neuron 100 dan 50		Hiddn layer neuron 200 dan 100	
	Loss	Acc	Loss	Acc
200	0.0097	0.9200	3.9650e-04	1.0000
400	0.0010	1.0000	6.9595e-07	1.0000
600	9.7474e-04	1.0000	1.3115e-06	1.0000
800	8.1390e-06	1.0000	1.9222e-09	1.0000
1000	0.0041	0.9200	1.7395e-06	1.0000

Pelatihan data dilakukan dengan meningkatkan epoch dari 200 menjadi 1000 dan jumlah *neuron* di *hidden layer* 1 dan *hidden layer* 2 dari 100 dan 50 menjadi 200 dan 100. Hasil yang didapatkan dari pelatihan dengan *epoch* sebanyak 1000 menunjukkan model lebih baik dengan menggunakan *hidden layer* 200 dan 100.

```

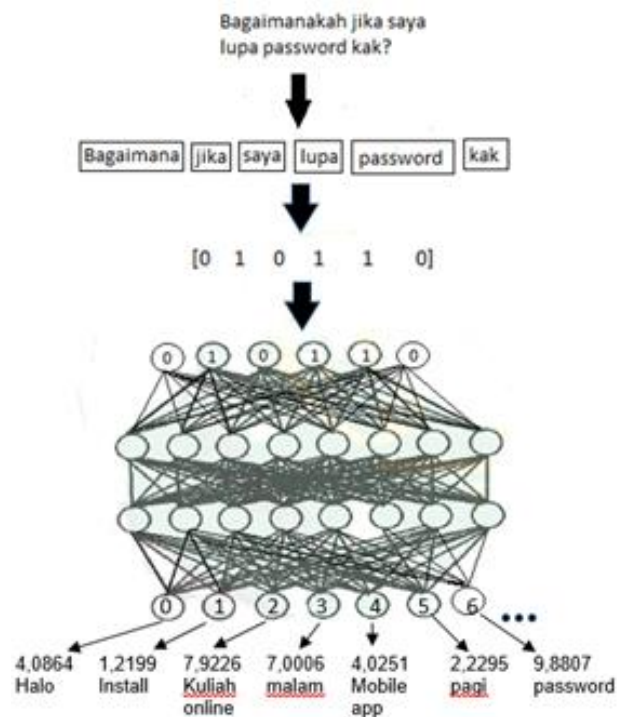
model.compile(optimizer="rmsprop", loss="mse", metrics=["accuracy"])
#model.fit(training, output, epochs=200, batch_size=32)
model.fit(numpy.array(train_x), numpy.array(train_y), epochs=1000, batch_size=30)
model.save("model.h5")
print("saved model to disk")
Epoch 992/1000
1/1 [-----] - 0s 2ms/step - loss: 4.7562e-11 - accuracy: 1.0000
Epoch 993/1000
1/1 [-----] - 0s 3ms/step - loss: 6.8394e-10 - accuracy: 1.0000
Epoch 994/1000
1/1 [-----] - 0s 4ms/step - loss: 1.9181e-07 - accuracy: 1.0000
Epoch 995/1000
1/1 [-----] - 0s 2ms/step - loss: 9.6482e-13 - accuracy: 1.0000
Epoch 996/1000
1/1 [-----] - 0s 3ms/step - loss: 7.4346e-07 - accuracy: 1.0000
Epoch 997/1000
1/1 [-----] - 0s 3ms/step - loss: 3.3291e-11 - accuracy: 1.0000
Epoch 998/1000
1/1 [-----] - 0s 2ms/step - loss: 2.2078e-08 - accuracy: 1.0000
Epoch 999/1000
1/1 [-----] - 0s 2ms/step - loss: 4.9287e-10 - accuracy: 1.0000
Epoch 1000/1000
1/1 [-----] - 0s 2ms/step - loss: 1.7395e-06 - accuracy: 1.0000
saved model to disk
    
```

Gambar 13. Hasil *compile* model ANN dengan nilai *epoch* 1000.



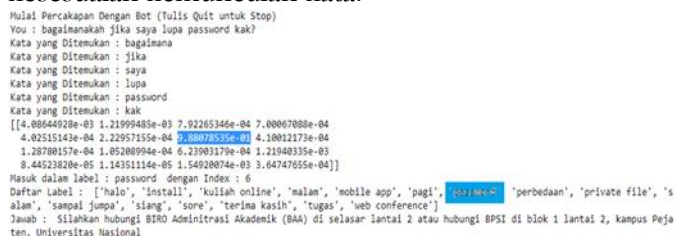
Gambar 14. Grafik *accuracy* dan *loss* model ANN.

Tujuan dari *artificial neural network* yang ada dalam sistem ini adalah untuk dapat mengklasifikasikan pertanyaan yang dimasukkan oleh pengguna kedalam sebuah label yang sesuai dengan nilai *probabilitasnya*. Sehingga didapat jawaban mana yang harus dipilih berdasarkan label yang sudah terpilih. Pada gambar 15 merupakan proses pelatihan dari *artificial neural network* mulai dari kalimat pertanyaan, proses *preprocessing* data, pelatihan model, *output* yang merupakan *label* dan nilai hasil *probabilitas* setiap *label*.



Gambar 15. Proses pelatihan model *artificial neural network*

Berikut **gambar 16** contoh hasil dari sistem yang dibuat. Dimana dari kalimat pertanyaan dilakukan proses pengolahan teks, kemudian hasil pengelolaan teks dicocokkan dengan kata unik yang ada di dalam *bag of word*, setelah itu dihitung nilai probabilitas kesesuaian kemunculan kata.



Gambar 16. Proses pengambilan jawaban dari pertanyaan berdasarkan nilai probabilitas yang paling besar.

Dari gambar 16 dapat dilihat bahwa pengambilan jawaban berdasarkan nilai probabilitas yang paling besar antara pertanyaan dengan label yang sudah dilakukan proses pelatihan. Dimana user bertanya tentang “*bagaimana jika lupa password?*” Sistem melakukan proses dan menghasilkan nilai *probabilitas* yang paling besar di label / kelas *password* dengan nilai *probabilitas* 0,988 yang merupakan nilai paling besar diantara nilai probabilitas 16 label lainnya. Dari label yang sudah didapat kemudian dipilih untuk ditampilkan output jawaban yang sesuai. Dari label *password* mempunyai kandidat jawaban yang ditampilkan adalah “Silahkan hubungi BIRO Adminitrasi Akademik (BAA) di selasar lantai 2 atau hubungi BPSI di blok 1 lantai 2, kampus Pejaten, Universitas Nasional”.

Pengujian Sistem

Pengujian sistem dilakukan langsung dengan cara membuat *Graphic User Interface* (GUI) *Chatbot* kemudian menginputkan teks berupa pertanyaan ke dalam sistem GUI yang sudah dibuat, gambar tampilan GUI ditampilkan dalam gambar 17. Kemudian mencocokkan apakah *output* yang di keluarkan sesuai dengan yang terdaftar pada *faq web* kuliah. Pengujian menggunakan 110 percakapan yang berbeda dengan dengan data yang dilatih tetapi mempunyai kesamaan maksud. Model diberi nilai *threshold*(nilai ambang) antara 0,1 sampai 0,5. Selain nilai *threshold* model diuji dengan memberikan nilai batas *probabilitas* label yang dipilih sebagai jawaban, nilai batas *probabilitas* yang diuji 0,1 sampai 0,9. Apabila nilai *probabilitas* dari model dibawah yang ditetapkan maka sistem *Chatbot* akan memberikan jawaban untuk mengulang pertanyaan. Tabel hasil akurasi pengujian dengan nilai *threshold* dan batas *probabilitas* dapat dilihat di **tabel 3**.

Tabel 3: Hasil akurasi pengujian dengan nilai *threshold* 0,1 – 0,5 dan nilai *probabilitas* 0,1 – 0,9

Batas Probabilitas	Nilai Threshold				
	0,1	0,2	0,3	0,4	0,5
0,1	97,27	96,36	96,36	96,36	95,54
0,2	97,27	96,36	96,36	96,36	95,54
0,3	97,27	96,36	96,36	96,36	95,54
0,4	97,27	96,36	96,36	96,36	95,54
0,5	97,27	96,36	96,36	96,36	95,54
0,6	97,27	96,36	96,36	96,36	95,54
0,7	97,27	96,36	96,36	96,36	95,54
0,8	97,27	96,36	96,36	96,36	95,54
0,9	97,27	96,36	96,36	96,36	95,54

Berdasarkan tabel 3 dapat dilihat bahwa pengujian dengan menggunakan nilai *threshold* 0,1 sampai 0,5 dan nilai batas *probabilitas* 0,1 sampai 0,9 didapatkan hasil yang lebih baik jika sistem *Chatbot* menggunakan nilai *threshold* 0,1 dan batas nilai *probabilitas* yang dapat dipilih sebagai kandidat jawaban adalah diatas 0,4. Apabila nilai *probabilitas* antara pertanyaan yang diinputkan dengan model *Chatbot* dibawah 0,4, sistem akan memberikan tanggapan untuk mengulangi pertanyaan yang lebih jelas.

Pengujian sistem *Chatbot* dengan nilai *threshold* 0,1 dan nilai batas *probabilitas* 0,4 menunjukkan dari 110 percakapan *Chatbot* berhasil menjawab dengan tepat sebanyak 107 percakapan dan jawaban salah sebanyak 3 percakapan. Hasil pengujian pengembangan model *Chatbot* berbasis teks menunjukkan akurasi yang cukup baik sebesar 97.27% dan kesalahan 2.72%. Cara menghitung nilai akurasi dan kesalahan dapat dirumuskan sebagai berikut (3), (4):

$$\text{Akurasi} = \frac{\text{Jumlah pengujian yang benar}}{\text{Jumlah seluruh pengujian}} \times 100 \% \quad (3)$$

$$\text{Akurasi} = \frac{107}{110} \times 100 \%$$

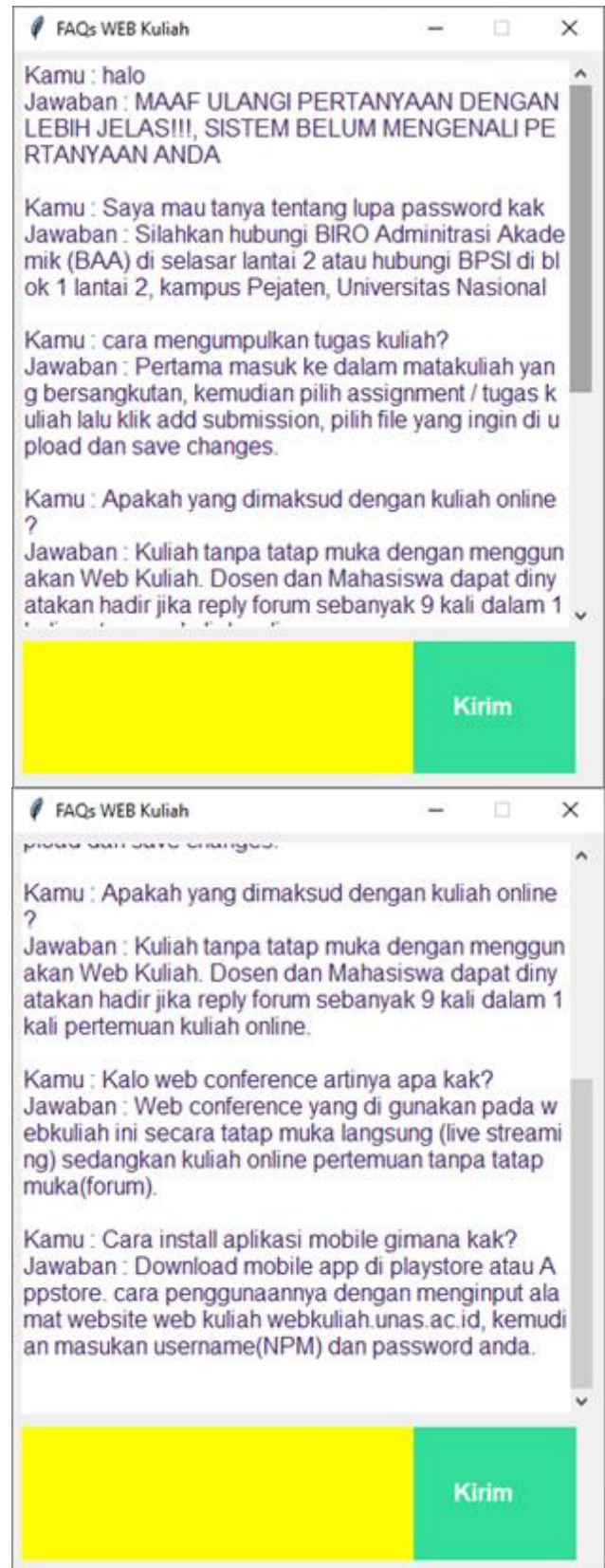
$$\text{Akurasi} = 97,27 \%$$

$$\text{Kesalahan} = \frac{\text{Jumlah pengujian yang salah}}{\text{Jumlah seluruh pengujian}} \times 100 \% \quad (4)$$

$$\text{Kesalahan} = \frac{3}{110} \times 100 \%$$

$$\text{Kesalahan} = 2,72 \%$$

Nilai akurasi yang didapat sesuai dengan perhitungan sistem yang telah dilakukan, dapat dilihat pada tabel 3, yaitu dengan nilai *threshold* 0,1 dan nilai batas *probabilitas* 0,4 menghasilkan nilai akurasi 97,27%. Berikut ini merupakan contoh daftar hasil pengujian dengan 5 pertanyaan dan gambar GUI yang ditampilkan dalam gambar 17 dan tabel 4.



Gambar 17. Hasil GUI pengujian *Chatbot*

Tabel 4: Daftar contoh pertanyaan dan jawaban dari pengujian sistem

No	Pertanyaan	Nilai Proba bilitas	Label /kelas	Jawaban	Validasi
1	Saya mau tanya tentang lupa password kak	1,0	password	Silahkan hubungi BIRO Administrasi Akademik (BAA) di selasar lantai 2 atau hubungi BPSI di blok 1 lantai 2, kampus Pejaten, Universitas Nasional Pertama masuk ke dalam matakuliah yang bersangkutan, kemudian pilih assignment / tugas kuliah lalu klik add submission, pilih file yang ingin di upload dan save changes.	Benar
2	Cara mengumpulkan tugas kuliah?	0,99	tugas	Kuliah tanpa tatap muka dengan menggunakan Web Kuliah. Dosen dan Mahasiswa dapat dinyatakan hadir jika reply forum sebanyak 9 kali dalam 1 kali pertemuan kuliah online.	Benar
3	Apakah yang dimaksud dengan kuliah online?	0,99	Kuliah online	Download mobile app di playstore atau Appstore. cara penggunaannya dengan menginput alamat website web kuliah webkuliah.unas.ac.id, kemudian masukan username(NPM) dan password anda	Benar
4	Cara install aplikasi mobile gimana kak?	1,0	Install	Maaf ulangi pertanyaan dengan lebih jelas!!!, sistem belum mengenali pertanyaan anda	Benar
5	Halo	0,11	Password		Benar

4. Kesimpulan dan Saran

Pengembangan *Chatbot* menggunakan artificial neural network dilakukan dengan melakukan pelatihan model *Chatbot* dengan pengulangan sampai 1000 epoch untuk 200 neuron di hidden layer 1 dan 100 neuron di hidden layer 2. Pelatihan menggunakan data dari pertanyaan yang sering timbul (Faq) di *web* kuliah sebanyak 25 percakapan. Pengujian menggunakan 110 percakapan yang berbeda dengan data yang dilatih tetapi mempunyai kesamaan maksud. Dari pengujian model *Chatbot* berbasis teks menghasilkan *Chatbot* dengan akurasi yang baik yaitu sebesar 97,27% dan kesalahan 2,72%.

Hasil yang didapat dari pengembangan model *Chatbot* memudahkan mahasiswa dan pihak universitas dalam pelaksanaan kuliah online. Universitas akan lebih cepat dalam menjawab kendala yang sudah sering ditanyakan oleh mahasiswa dan mahasiswa tidak perlu menunggu lama jawaban dari universitas terkait kendala penggunaan *web* kuliah yang dialami.

5. Daftar Pustaka

- [1] Wu, B., Wang, B. and Xue, H., 2016, December. Ranking responses oriented to conversational relevance in chat-bots. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers (pp. 652-662).
- [2] Wu, Y., Wu, W., Xing, C., Zhou, M. and Li, Z., 2016. Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots. arXiv preprint arXiv:1612.01627.
- [3] Chandra, Y.W. and Suyanto, S., 2019. Indonesian chatbot of university admission using a question answering system based on sequence-to-sequence model. Procedia Computer Science, 157, pp.367-374.
- [4] Rizki, U., 2019. Multi Respon Ranking Pada Percakapan Layanan Travel Berdasarkan Riwayat Obrolan. Jurnal Informa, 5(3), pp.73-79.
- [5] Bhagwat, V.A., 2018. Deep Learning for Chatbots., San Jose State University, 2018.
- [6] Jwala, K., Sirisha, G.N.V.G. and Raju, G.P., 2019. Developing a Chatbot using Machine Learning. International Journal of Recent Technology and Engineering (IJRTE), 8(1S3), pp.89-92.
- [7] Imamah, F. and Dores, A., 2018. Aplikasi chatbot (milki bot) yang terintegrasi dengan web CMS untuk customer service pada UKM MINSU. Jurnal Cendikia, 16(2 Oktober), pp.100-106.
- [8] Trivedi, A., Gor, V. and Thakkar, Z., 2019. Chatbot generation and integration: A review. International Journal of Advance Research, Ideas and Innovations in Technology, 5(2), pp.1308-1311.
- [9] Nagarajan, S.K., 2019. Server-Less Rule-Based Chatbot Using Deep Neural Network.

- [10] Wibawa, M.S., 2017. Pengaruh Fungsi Aktivasi, Optimisasi dan Jumlah Epoch Terhadap Performa Jaringan Saraf Tiruan. *Jurnal Sistem dan Informatika (JSI)*, 11(2), pp.167-174.