

# “Si Tole” *Chatterbot* untuk Melatih Rasa Percaya Diri Menggunakan *Naive Bayes Classification*

Zamah Sari<sup>1</sup>, Moechammad Sarosa<sup>2</sup>, Suhari<sup>3</sup>

**Abstract**—Observing the behavior and character of children today, many parents are worried about their child's development. Educational observers convey the necessity of character education from an early age to instill in them with morality and noble character. Having a strong character and quality, young generation will become an ethical, tough, and superior generation. Various efforts were made to provide character education to children. One of them is to train the level of confidence of children by inviting them to speak and express opinions in the form of text/chat. In order to help that, a chat app that can help kids to tell stories in the form of text/chat was built, so the kids accustom to expressing opinions and confidently talk to others and happy to hear it. The character education raised in this application is to increase the confidence level of the child. The topics raised as conversation on the Si Tole app are about friendship, joy (food, books, movies), recreation, birthday parties, hopes, and dreams. By practicing to talk to Si Tole *Chatterbot*, it is expected that children's character development will be better. This application is suitable to be a medium of character education for adolescents (Primary School/Junior High School) to improve their confidence.

**Intisari**—Mencermati perilaku dan budi pekerti anak zaman sekarang menyebabkan banyak orang tua merasa khawatir dengan perkembangan anak-anak. Para pemerhati pendidikan menyampaikan perlunya pendidikan karakter sejak dini untuk menanamkan kepada anak-anak akhlak dan budi pekerti yang luhur. Dengan memiliki karakter yang kuat dan berkualitas, diharapkan generasi muda akan menjadi generasi yang beretika, tangguh, dan unggul. Berbagai upaya dilakukan untuk memberikan pendidikan karakter kepada anak-anak, salah satu caranya adalah melatih tingkat kepercayaan diri anak dengan mengajarkannya berbicara dan mengekspresikan pendapat dalam bentuk teks/chat. Untuk itu, dibangunlah sebuah aplikasi *chat* yang akan membantu anak-anak untuk bercerita dalam bentuk teks/chat, sehingga menjadi terbiasa untuk mengeluarkan pendapat dan percaya diri untuk berbicara dengan orang lain dan anak-anak akan senang untuk mendengarnya. Adapun pendidikan karakter yang diangkat pada aplikasi ini adalah untuk meningkatkan tingkat percaya diri anak. Topik yang diangkat sebagai bahan percakapan pada aplikasi Si Tole ini adalah tentang pertemanan, kesukaan (makanan, buku, film), rekreasi, pesta ulang tahun, harapan, dan mimpi. Dengan berlatih untuk berbicara dengan Si Tole *Chatterbot*, diharapkan perkembangan karakter anak-anak menjadi lebih baik. Aplikasi ini cocok dijadikan media pembelajaran pendidikan karakter

bagi anak-anak remaja (SD/SMP) untuk meningkatkan rasa percaya diri.

**Kata Kunci**— *chatterbot*, percaya diri, *Naive Bayes*, pendidikan karakter, remaja.

## I. PENDAHULUAN

Kualitas karakter yang tinggi di suatu masyarakat akan menumbuhkan kualitas karakter bangsa tersebut. Beberapa ahli berkeyakinan bahwa pengembangan karakter yang terbaik adalah jika dimulai sejak usia dini. Landasan karakter yang baik akan menjadi tonggak bagi perkembangan karakter mereka selanjutnya. Setelah memiliki karakter yang baik, pendidikan karakter selanjutnya diarahkan untuk menguatkan kepercayaan diri. Dalam menghadapi persaingan global diperlukan sifat percaya diri yang tinggi sehingga generasi muda berani beradu terhadap kemampuan yang dimiliki. Untuk menyampaikan pendidikan karakter kepada masyarakat tidaklah mudah. Diperlukan cara khusus dan teknik tertentu sehingga masyarakat menjadi tertarik. Dengan memperhatikan permasalahan tersebut, solusi yang ditawarkan adalah menyediakan media bantuan belajar pendidikan karakter berbasis aplikasi, yaitu berupa *chatterbot*, untuk melatih rasa percaya diri. Dengan kepercayaan diri yang tinggi, diharapkan generasi muda tidak pantang menyerah terhadap kemampuan yang dimiliki, sehingga mampu bersaing dalam dunia global.

Hasil percakapan yang diperoleh dari *log conversation* yang ada pada *server chatterbot* kemudian diolah untuk mendapatkan nilai tingkat percaya diri anak menggunakan algoritme *Naive Bayes* untuk mengklasifikasikan tingkat kepercayaan diri menjadi tidak percaya diri, percaya diri, dan sangat percaya diri.

Makalah ini merancang perangkat lunak Si Tole menggunakan skema kecerdasan buatan yang dapat dimanfaatkan untuk meningkatkan tingkat percaya diri anak, mengembangkan mesin *Artificial Linguistic Internet Computer Entity* (ALICE), dan mengintegrasikannya dengan metode *Naive Bayes* untuk mendapatkan respons dari Si Tole seperti yang diinginkan, dan mengimplementasikan Si Tole dalam perangkat lunak berbasis Android dan IOS menggunakan REST API dan Pandorabots. Penelitian ini diharapkan dapat bermanfaat bagi anak-anak remaja (SD/SMP) dalam berlatih meningkatkan rasa percaya diri.

## II. CHATTERBOT DAN ALGORITME NAIVE BAYES

### A. Chatterbot

Bagian ini membahas tentang penelitian relevan yang telah dilakukan sebelumnya. Ada beberapa penelitian yang dibahas, yang semuanya menggunakan mesin ALICE sebagai dasar penelitian dan dikembangkan sesuai dengan tujuan penelitian

<sup>1</sup> Dosen, Jurusan Informatika, Universitas Muhammadiyah Malang, Jl. Raya Tlogomas 246 Malang, 65144, Indonesia (Telp: 081945115544), e-mail: zamahsari@umm.ac.id)

<sup>2</sup> Dosen, Jurusan Teknik Elektro, Politeknik Negeri Malang, Jl. Soekarno Hatta 09 Malang, 65141, (Telp: 0341-404424, e-mail: msarosa@polinema.ac.id)

<sup>3</sup> Dosen, FKIP/PPKn, Universitas PGRI Adi Buana Surabaya, Jl. Ngagel Dadi III-B/37 Surabaya, (email: suharimau@gmail.com)

masing-masing. Kemudian berdasarkan hasil penelitian-penelitian tersebut, penelitian ini dilakukan.

*Artificial Intelligence Markup Language* (AIML) adalah bahasa bertipe *markup* yang digunakan untuk membuat logika *chatbot*. AIML merupakan kumpulan pola beserta respons yang digunakan *chatbot* untuk menelusuri jawaban dari kalimat yang dimasukkan [1]. Mesin ALICE merupakan salah satu *chatbot* yang memiliki kecerdasan buatan dengan memanfaatkan *natural language processing* dengan memanfaatkan *pattern matching* atau pencocokan pola percakapan yang ada dan disimpan dalam basis data [2].

Implementasi dua sistem yaitu ALICE dan ELIZABETH untuk membangun *chatbot* dua bahasa dengan menggunakan AIML berbasis Java menghasilkan *chatbot* yang dapat mengerti percakapan dalam bahasa Afrika ke bahasa Inggris atau sebaliknya [3]. Penelitian ini menghasilkan empat buah aplikasi, yaitu aplikasi Java berbasis sistem ALICE yang dapat diakses melalui *website*, aplikasi *desktop* menggunakan Delphi berbasis sistem ELIZABETH yang digunakan untuk mendukung evaluasi kinerja sistem ALICE, aplikasi Java yang digunakan untuk menerjemahkan kategori AIML ke dalam *subset* ELIZABETH *style rule*, dan aplikasi Java yang digunakan untuk membaca dialog dan mengubahnya ke format AIML. Permasalahan yang dihadapi adalah cara mengetahui bahasa yang digunakan oleh pengguna pada saat proses memasukkan data kemudian mengubahnya ke format AIML. Program Java yang digunakan untuk menerjemahkan transkrip dialog ke dalam representasi pengetahuan AIML memberikan implementasi dasar dari kumpulan kalimat berbasis pelatihan *chatbot* [4]. Kekurangan dari aplikasi ini adalah perangkat lunak dibangun dengan permasalahan yang masih umum, belum dikhususkan untuk permasalahan atau kondisi tertentu.

Penelitian tentang penggabungan *tutor bot* (T-Bot) dan *evaluation bot* (Q-Bot) efektif untuk membantu siswa selama proses belajar mengajar dan mendukung pengajar untuk melakukan aktivitas mengajar [5]. Penelitian ini menggunakan AIML berbasis PHP (Program E) yang dikombinasikan dengan sistem *e-learning* Claroline dan Moodle. T-bot digunakan untuk menganalisis pertanyaan siswa yang ditulis dalam bahasa natural dan memberikan jawaban sesuai dengan isi dari materi pelajaran dengan tepat. Sedangkan Q-Bot digunakan untuk melakukan supervisi dan evaluasi proses belajar siswa dengan memberikan pertanyaan yang berupa kuesioner pribadi. Pertanyaan kuesioner diberikan dalam bentuk pilihan ganda, benar/salah, dan jawaban singkat. Siswa dapat terhubung dengan *e-learning* melalui koneksi internet, kemudian siswa dapat berinteraksi dengan *bot* melalui *Bot User Interface* (BUI). Penelitian ini juga mengimplemetasikan tiga bahasa yaitu Inggris, Spanyol, dan Galisian. Kekurangan dari penelitian ini adalah pasangan masukan dan keluaran sudah ditentukan di dalam modul PHP, sehingga *chatbot* tidak dapat menciptakan jawaban teksnya sendiri.

Dalam sebuah penelitian lainnya, kemampuan emosi dan model kepribadian pada Alicebot menggunakan AIML berbasis PHP (Program E) dengan koneksi MySQL [6]. Kemampuan ini membuat Alicebot mampu membuat

keputusan sendiri berdasarkan emosi dan model kepribadian-nya. Alicebot juga mampu untuk menyatakan kesukaan atau ketidaksukaannya terhadap sesuatu. Penelitian ini menambahkan tiga modul ke dalam Alicebot, yaitu modul emosi dan kepribadian, modul konsep domain, dan modul penciptaan teks/kata. Kode-kode tambahan untuk menambahkan kemampuan emosi dan model kepribadian ditulis dalam bahasa PHP dan diintegrasikan dengan Program E. Hasil penelitian ini membuat Alicebot dapat menciptakan jawaban teksnya sendiri tanpa menggunakan *pattern* atau pola pencocokan kata [6]. Keterbatasan penelitian ini adalah aplikasi yang dibuat masih umum, belum dikhususkan untuk menangani permasalahan atau spesifikasi bidang tertentu, sehingga kosakata yang diciptakan terbatas hanya untuk percakapan umum manusia. Sedangkan pada penelitian yang lain, *chatbot* digunakan untuk menjawab pertanyaan dari lawan bicara yang jawaban dari pertanyaan-pertanyaan tersebut diambil menggunakan *web crawler* dari *website* tertentu dan membuat sebuah rangkuman dari hasil *crawling* untuk dijadikan jawaban. Pada penelitian ini, *chatbot* telah difungsikan sebagai mesin penjawab untuk hal tertentu saja, misalnya pertanyaan dan jawaban untuk hal seputar penerimaan mahasiswa baru [7].

### B. *Natural Language Processing*

Bahasa sebagai bagian yang penting dari kehidupan manusia, dalam bentuk tulis merupakan catatan dari pengetahuan yang didapat dari kehidupan manusia dari satu generasi ke generasi berikutnya, sedangkan dalam bentuk lisan merupakan sarana komunikasi antar individu dalam suatu masyarakat. Tujuan dalam bidang *natural language* ini adalah melakukan proses pembuatan model komputasi dari bahasa, sehingga dapat terjadi suatu interaksi antara manusia dengan komputer dengan perantaraan *natural language*.

Sebuah *natural language system* harus memperhatikan pengetahuan terhadap bahasa itu sendiri, baik dari kata yang digunakan, cara kata-kata tersebut digabung untuk menghasilkan suatu kalimat, arti sebuah kata, fungsi sebuah kata dalam sebuah kalimat, dan sebagainya. Akan tetapi, ada satu hal lagi yang perlu dipertimbangkan yang sangat berperan dalam bahasa, yaitu kemampuan manusia untuk mengerti dan kemampuan untuk itu diperoleh dari pengetahuan yang didapat secara terus menerus sewaktu hidup.

### C. *Algoritme Naive Bayes*

*Algoritme Naive Bayes* merupakan metode klasifikasi yang digunakan untuk memprediksi probabilitas. Algoritme ini memanfaatkan teori probabilitas yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi probabilitas di masa depan berdasarkan pengalaman di masa sebelumnya. Dua kelompok peneliti, satu oleh Pantel dan Lin, dan yang lain oleh Microsoft Research memperkenalkan metode statistik Bayesian. Tetapi yang membuat *Naive Bayes* ini populer adalah pendekatan yang dilakukan oleh Paul Graham. Banyak aplikasi menghubungkan antara atribut *set* dan variabel kelas yang *non-deterministic*. Dengan kata lain, label kelas *test record* tidak dapat diprediksi dengan peristiwa

tertentu meski atribut *set* identik dengan beberapa contoh *training*.

Algoritme *Naive Bayes* adalah metode yang mendasarkan probabilitas dan teorema Bayes dengan asumsi bahwa setiap variabel bersifat bebas (*independent*) dan mengasumsikan bahwa keberadaan sebuah fitur (*variabel*) tidak ada kaitannya dengan beradanya fitur yang lain. *Naive Bayes* adalah model penyederhanaan dari metode Bayes. *Naive Bayes* inilah yang digunakan di dalam *machine learning* sebagai metode untuk mendapatkan hipotesis untuk suatu keputusan.

#### D. Metode Rekayasa Perangkat Lunak SDLC

*System Development Life Cycle* (SDLC) merupakan proses sebuah sistem informasi dapat mendukung sebuah usaha yang dibutuhkan, desain sistem, pembangunan, dan penyampaian kepada pengguna.

### III. METODOLOGI

Langkah-langkah yang dilakukan pada penelitian ini dijelaskan melalui metode SDLC yang meliputi fase perencanaan, analisis, desain, implementasi, dan pengujian.

#### A. Perencanaan Sistem

Pada tahap ini ditetapkan spesifikasi Si Tole yang dibangun berdasarkan kebutuhan, meliputi hal-hal sebagai berikut.

1) *Subsistem Basis Data*: Subsistem basis data dalam Si Tole berfungsi sebagai sumber data informasi yang menangani masukan dari pengguna dan keluaran atau respons untuk pengguna. Spesifikasi subsistem basis data pada penelitian ini terdiri atas dua bagian yaitu sebagai berikut.

##### a. Basis data sistem ALICE.

Karena mesin ALICE yang digunakan pada penelitian ini menggunakan *interpreter* Program O, maka basis data MySQL dibutuhkan untuk menyimpan parameter masukan dan keluaran, karakteristik dari *chatterbot*, perbaikan ejaan, pengguna, data *botmaster*, dan *log* percakapan.

##### b. Basis data penilaian tingkat percaya diri.

Si Tole dibangun untuk memenuhi kebutuhan anak-anak untuk meningkatkan rasa percaya diri. Basis data ini digunakan untuk menyimpan informasi kata kunci yang digunakan untuk melakukan penilaian tingkat percaya diri dan hasilnya.

2) *Subsistem Model*: Subsistem model dalam sistem ini berfungsi untuk mengimplementasikan sistem menggunakan metode yang sudah ada pada mesin ALICE dan mengintegrasikannya dengan metode *Naive Bayes*. Spesifikasi submodel pada Si Tole terdiri atas:

- a. mesin ALICE dengan kemampuan kecerdasan buatan, dan
- b. *framework Naive Bayes* yang dibuat dengan kode PHP.

3) *Subsistem End User Interface (UI)*: Subsistem UI berfungsi untuk visualisasi perangkat lunak Si Tole *Chatterbot*, sehingga pengguna dapat melakukan interaksi dengan sistem. Spesifikasi UI berupa hal-hal sebagai berikut.

- a. Tampilan antarmuka untuk pengguna berupa halaman *chat*.
- b. Tampilan antarmuka untuk pengguna berupa halaman hasil penilaian tingkat rasa percaya diri.

#### B. Analisis

1) *Analisis Kebutuhan Perangkat Lunak*: Untuk membangun perangkat lunak Si Tole *Chatterbot* yang dapat melakukan penilaian tingkat rasa percaya diri hasil percakapan, algoritme yang digunakan adalah algoritme *Naive Bayes*. Variabel penelitian juga dibutuhkan untuk melakukan pengujian. Variabel penelitian ini terdiri atas dua buah variabel sebagai berikut.

a. Variabel bebas, yaitu variabel yang memengaruhi atau menjadi sebab perubahannya atau timbulnya variabel terikat. Variabel ini adalah masukan kata atau kalimat dari pengguna yang terdiri atas

- kata atau kalimat yang berhubungan dengan karakteristik *chatterbot*, dan
- kata atau kalimat yang berhubungan dengan topik untuk menggali tingkat kepercayaan diri dari pengguna.

b. Variabel terikat, yaitu variabel yang dipengaruhi atau yang menjadi akibat adanya variabel bebas. Variabel ini adalah respons *chatterbot* terhadap masukan pengguna yang terdiri atas

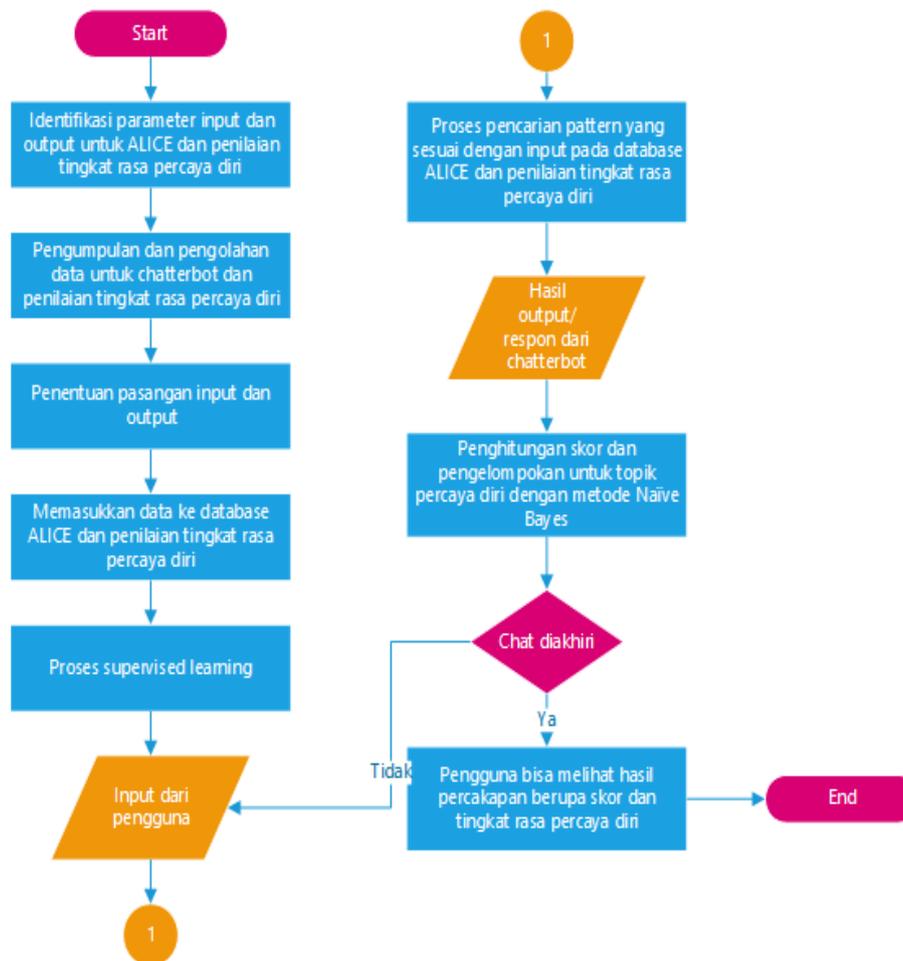
- keterkaitan topik antara masukan kata atau kalimat pengguna dengan respons kata atau kalimat yang diberikan oleh Si Tole *Chatterbot*,
- respons kata atau kalimat yang diberikan Si Tole *Chatterbot* jika informasi yang diinginkan tidak terdapat pada basis data mesin ALICE,
- respons kata atau kalimat yang diberikan Si Tole *Chatterbot* jika informasi yang diinginkan tidak terdapat pada basis data penilaian tingkat rasa percaya diri,
- respons kata atau kalimat yang diberikan Si Tole *Chatterbot* jika masukan atau kalimat tidak berhubungan dengan karakteristik *chatterbot*, penilaian tingkat rasa percaya diri, dan pengetahuan umum, dan
- waktu respons yang diberikan oleh Si Tole *Chatterbot* untuk masukan tentang *chatterbot*.

2) *Analisis Pemodelan Perangkat Lunak*: Basis data yang digunakan pada Si Tole *Chatterbot* diambil dari basis data yang dimiliki oleh mesin ALICE dan basis data penilaian tingkat rasa percaya diri. Keluaran atau respons yang dihasilkan oleh *chatterbot* merupakan integrasi dari koneksi basis data pada mesin ALICE, basis data penilaian tingkat rasa percaya diri, dan metode *Naive Bayes*.

Konsep algoritme dari sistem ini disajikan dengan diagram alir pada Gbr. 1.

#### C. Desain

1) *Subsistem Basis Data*: Desain subsistem ini memanfaatkan basis data yang sudah ada pada mesin ALICE dan basis data penilaian tingkat rasa percaya diri.



Gbr. 1 Diagram alir konsep algoritme aplikasi Si Tole Chatterbot.

2) *Subsistem Model Base*: Desain subsistem ini terdiri atas dua hal sebagai berikut.

- a. Mesin ALICE dengan kecerdasan buatan. Perancangan pada basis data mesin ALICE dikhususkan untuk membentuk karakteristik *chatterbot* sebagai Si Tole Chatterbot dengan memberikan parameter masukan dan keluaran pada tabel *bot personality*.
- b. *Framework Naive Bayes* dibuat dengan kode PHP.

3) *Subsistem User Interface*: Tampilan antarmuka untuk pengguna berupa jendela *chat* yang dibuat dengan menggunakan Android. Pada tahun kedua ini tampilan ditambahkan kepada pengguna untuk melihat hasil percakapan yang telah dilakukan.

#### IV. HASIL PENCAPAIAN

##### A. Persiapan Penelitian

Persiapan yang dilakukan dalam melaksanakan penelitian ini adalah sebagai berikut.

1. Studi pustaka terkait dengan penelitian yang digunakan sebagai referensi, dasar teori tentang beberapa algoritme yang digunakan, yaitu algoritme *Naive Bayes* dan *Phrase*

*Reinforcement*, serta *user interface* yang baru yaitu untuk *smartphone* Android [8].

2. Menentukan data yang dibutuhkan dan sumber-sumber untuk mendapatkannya.

- Data penelitian diambil dari referensi yang telah ada, dengan *chatterbot* dibuat sedemikian rupa sehingga memiliki *personality* sebagai seorang pewawancara yang melakukan tes wawancara menggunakan bahasa Inggris [9].
- Basis data untuk *knowledge base* Si Tole Chatterbot.
- Menentukan spesifikasi kebutuhan untuk merancang Si Tole Chatterbot pada *smartphone* Android.

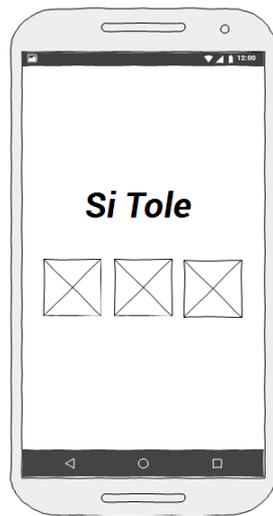
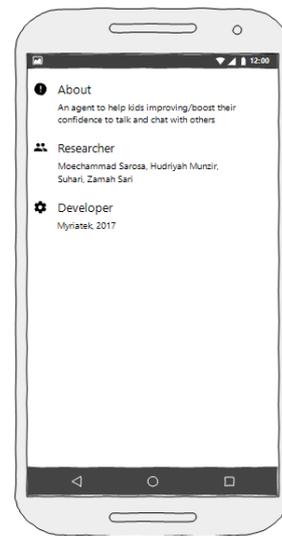
##### B. Desain Aplikasi

Aplikasi terdiri atas tiga bagian subsistem yang terdiri atas hal-hal sebagai berikut.

- 1) *Modul Android*: Modul Android berisi kode-kode Java dan Android yang dibangun pada sisi *client* untuk mengirim dan menerima *chat* yang dilakukan dengan Si Tole Chatterbot, terdiri atas lima desain *user interface* Android.

##### a. Splash Screen

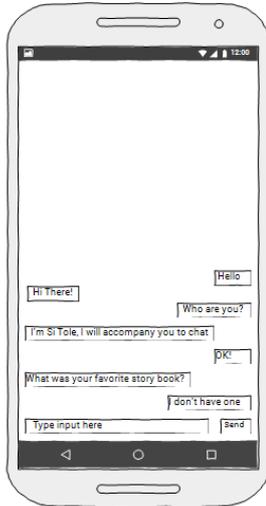
Gbr. 2 memperlihatkan tampilan awal aplikasi Si Tole Chatterbot.

Gbr. 2 Tampilan *splash screen*.Gbr. 4 Menu *About*.

b. Pada bagian dalam aplikasi sendiri terdapat beberapa tampilan, yaitu sebagai berikut.

- *Chat room*

Aplikasi Si Tole *Chatterbot* memiliki bagian utama yang digunakan untuk melakukan percakapan antara pengguna dan *bot* yang di dalamnya ditanamkan kecerdasan buatan sebagai alat untuk mendapatkan informasi yang akan digunakan untuk melakukan klasifikasi terhadap pengguna. Bagian *chatroom* ini didesain dengan satu buah *textbox* sebagai masukan, sebuah tombol untuk mengirimkan masukan teks ke *server*, dan sebuah *dialog box* yang menampilkan percakapan antara pengguna dan *bot* seperti ditunjukkan pada Gbr. 3.

Gbr. 3 *Chat room*.

- Menu *About*

Isi pada menu ini ditunjukkan pada Gbr. 4.

2) *Modul RESTful API Server*: Modul RESTful API berisi kode-kode PHP yang berfungsi untuk mengirim dan menerima data dari/ke basis data ALICE yang ada pada *web server*, kemudian diakses oleh aplikasi Android melalui API yang sudah didefinisikan pada *server*.

### C. Implementasi

Pada tahap ini desain dibawa menuju implementasi untuk algoritme *Naive Bayes*.

1) *Algoritme Klasifikasi Naive Bayes*: Teknik klasifikasi dapat dibagi menjadi lima kategori, masing-masing memiliki konsep perhitungan matematika yang berbeda. Kategori ini dibagi berdasarkan statistik, jarak, pohon keputusan, jaringan saraf tiruan, dan aturan. Masing-masing terdiri atas beberapa algoritme dan yang paling populer digunakan dari masing-masing kategori adalah *C4.5*, *Naive Bayes*, *K-Nearest Neighbors*, dan *Backpropagation Neural Network*. Algoritme *Naive Bayes* banyak digunakan karena memiliki komputasi yang lebih sederhana dalam implementasinya dibanding metode klasifikasi lainnya, meskipun mungkin dalam hal akurasi masih lebih baik jika digunakan metode *C4.5*, *K-Nearest Neighbors*, dan *Backpropagation Neural Network*. Algoritme *Naive Bayes* ini biasanya digunakan untuk aplikasi-aplikasi seperti pendeteksi *email spam*, pengelompokan kategori dari beberapa *item* secara otomatis, pendeteksi otomatis kelompok bahasa, analisis sentimen, dan lain-lain.

Dalam penelitian ini, algoritme *Naive Bayes* digunakan untuk mengelompokkan hasil percakapan menggali tingkat rasa percaya diri antara pengguna dan Si Tole *Chatterbot* yang memiliki klasifikasi: tidak percaya diri, kurang percaya diri, percaya diri, dan sangat percaya diri.

Berikut adalah sampel daftar pertanyaan yang digunakan untuk menggali tingkat rasa percaya diri anak-anak dengan topik tentang pertemanan, makanan, buku, film, musik, rekreasi, pesta ulang tahun, keinginan dan cita-cita.

- What has been your favorite moment as a \_\_\_ year old?
- What was your favorite activity this year?
- What was your favorite family event this year?
- Can you tell me a bit about your birthday party from this past year?
- What did you like about it?
- What would you like to do differently next year?

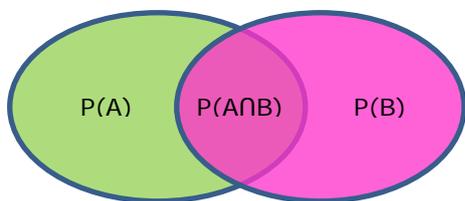
- What was your favorite story book?
- Is there a character from one of the books you (we) read that you still think about sometimes?
- What qualities about this character do you admire?
- What is something that worries you?
- Who inspires you? Why do you think that is?
- What was your favorite movie?
- What was your least favorite movie?
- What is currently your favorite food?
- What about a least favorite food?
- What food might you want to try in the next year?
- If you could only eat one thing for a whole week, what would you choose?
- What is your favorite restaurant? What do you like to eat there?
- What games do you like to play?
- What is your favorite thing to do with mom/dad/...?
- Which toys do you like to play with?
- Who is/are your best friend(s)?
- What is your favorite memory from summer break?
- Tell me something about this year that you never want to forget.
- Is there something you would like to do, or a place you would like to go in the next year? Can you tell me more about that?

Metode penentuan algoritme ini dapat dilakukan sebagai berikut.

a. Menghitung peluang dan teorema Bayes.

Misalnya diambil contoh:

- Ada sepuluh dari dua puluh percakapan yang tergolong percakapan “sangat percaya diri”.
- Ada delapan dari dua puluh percakapan yang mengandung kata-kata yang terdaftar pada *keyword* (*expect, willness, effort, interest, concern, enthusiasm, support, dan provide*).
- Lima percakapan yang mengandung kata-kata yang terdaftar pada percakapan yang “sangat percaya diri”.
- Jika ada percakapan kategori minat yang dilakukan terakhir oleh pengguna dan Si Tole *Chatterbot*, maka berapa peluang percakapan tersebut tergolong percakapan “sangat percaya diri” dan mengandung kata-kata *keyword* (*expect, willness, effort, interes, concern, enthusiasm, support, dan provide*).



Gbr. 5 Teorema Bayes.

Dengan menggunakan bentuk sederhana dari teorema Bayes, seperti pada Gbr, 5, maka dapat dihitung:

Jika  $P(A)$  = Peluang percakapan yang termasuk kategori “sangat percaya diri”.

$P(B)$  = Peluang percakapan yang mengandung kata-kata yang terdaftar pada *keyword* (*expect, willness, effort, interest, concern, enthusiasm, support, dan provide*).

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)} \tag{1}$$

sehingga

$$\begin{aligned} P(\text{sangat percaya diri}|\text{keyword}) &= \\ \frac{P(\text{sangat percaya diri}|\text{keyword}) * P(\text{sangat percaya diri})}{P(\text{keyword})} &= \\ = \frac{\frac{5}{10} * \frac{10}{20}}{\frac{8}{20}} = \frac{0.25}{0.4} &= 0.625 \end{aligned}$$

b. Pendekatan *Naive Bayes*

Dengan hasil perhitungan sebelumnya, untuk melakukan pendekatan *Naive Bayes* dapat ditambahkan permasalahan sebagai berikut.

- Ada sembilan percakapan yang mengandung kata *keywordB* (*learn, motivate, dan improve*).
- Lima percakapan tersebut dikategorikan sebagai percakapan “sangat percaya diri”.
- Berapa peluang sebuah percakapan yang dilakukan oleh pengguna yang mengandung daftar kata (*expect, willness, effort, interest, concern, enthusiasm, support, provide*) dan (*learn, motivate, improve*), yaitu *keywordA*.

Permasalahan di atas menjadi kompleks, dan tidak dapat diselesaikan menggunakan bentuk sederhana dari teorema Bayes, sehingga dibutuhkan (2).

$$\begin{aligned} P(\text{sangat berminat}|\text{keywordA, keywordB}) &= \\ \frac{P(\text{sangat PD}|\text{keywordA} \cap \text{keywordB}) * P(\text{keywordA}|\text{sangat PD}) * P(\text{sangat PD})}{P(\text{keywordA}|\text{keywordB}) * P(\text{keywordB})} &= \end{aligned} \tag{2}$$

Untuk menyederhanakan asumsi, dapat dianggap bahwa kejadian munculnya daftar kata *keywordA* tidak bergantung secara penuh terhadap kemunculan daftar kata *keywordB* pada percakapan, sehingga (2) dapat disederhanakan menjadi (3)

$$\begin{aligned} P(\text{sangat percaya diri}|\text{keywordA, keywordB}) &= \\ \frac{P(\text{keywordA}|\text{sangat PD}) * P(\text{keywordB}|\text{sangat PD}) * P(\text{sangat PD})}{P(\text{keywordA}) * P(\text{keywordB})} &= \end{aligned} \tag{3}$$

$$= \frac{\frac{5}{10} * \frac{5}{10} * \frac{10}{20}}{\frac{8}{20} * \frac{9}{20}} = \frac{250/2000}{72/400} = \frac{0.125}{0.18} = 0.694$$

Nilai peluang ini adalah kemungkinan bahwa respons percakapan yang diberikan oleh pengguna saat melakukan percakapan dengan Si Tole *Chatterbot* mengandung kata-kata yang terdaftar pada *keywordA* dan *keywordB* dan tergolong percakapan yang “sangat percaya diri”. Perhitungan di atas dilakukan untuk mengklasifikasikan percakapan termasuk percakapan "sangat percaya diri" dengan mengambil peluang dari kata-kata yang terkandung dalam percakapan tersebut.

Pendekatan *Naive Bayes* dilakukan dengan cara membuat asumsi bahwa kejadian pada satu kelompok daftar kata tidak bergantung secara penuh terhadap kemunculan kelompok daftar kata yang lain. Hal ini dilakukan untuk menyederhanakan proses perhitungan peluang.



Gbr. 6 Tampilan menu (a) *Splash Screen*, (b) *Menu drawer*.

2) *Implementasi pada Android*: Implementasi algoritme pada Android meliputi hal-hal sebagai berikut.

a. *Splash screen*

Gbr. 6(a) merupakan tampilan awal aplikasi saat dibuka.

b. *Main screen*

*Main Screen* merupakan tampilan utama saat pengguna sudah masuk ke aplikasi Si Tole Chatterbot. Di dalamnya terdapat beberapa menu yang sudah diimplementasikan.

c. *Menu drawer*

*Menu drawer* merupakan daftar menu yang tersedia pada aplikasi, seperti ditunjukkan pada Gbr. 6(b).

d. *Chat room*

*Chat room* merupakan tampilan untuk melakukan percakapan dengan Si Tole Chatterbot. Pertanyaan atau jawaban juga bisa diketikkan pada *input text box*. Pertanyaan dan jawaban yang diberikan oleh pengguna dan Si Tole Chatterbot ditampilkan pada layar percakapan seperti pada Gbr. 7(a).

e. *About*

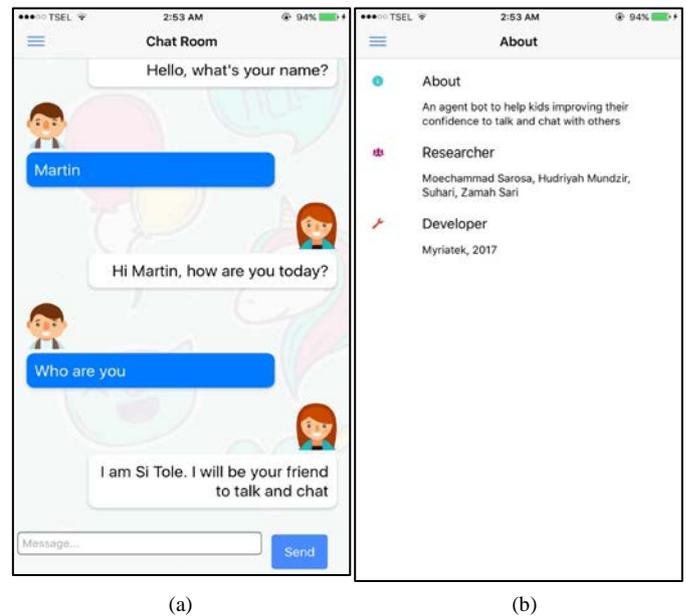
*Menu About* merupakan tampilan informasi aplikasi seperti yang ditunjukkan pada Gbr. 7(b).

D. *Testing dan Pengujian*

Pengujian terdiri atas beberapa tahapan pada pengujian sebagai berikut.

1) *Verifikasi*: Digunakan untuk melihat kesesuaian antara desain dengan hasil.

2) *Validasi*: Digunakan untuk menguji kesesuaian fungsi Si Tole Chatterbot dalam sistem ini, bisa dipenuhi atau tidak, dengan memberikan masukan dan menganalisis hasil keluaran, atau respons yang diberikan sudah sesuai dengan yang diinginkan atau tidak.



Gbr. 7 Tampilan menu (a) *Chat room*, (b) *About*.

3) *Uji Prototype*: Digunakan untuk mengetahui, sistem yang dibuat merupakan alternatif yang lebih baik dari yang ada sekarang atau tidak. Pengujian ini dilakukan dengan melibatkan responden yang melakukan percakapan dengan Si Tole Chatterbot dengan tidak diberitahu sebelumnya bahwa yang diajak berkomunikasi adalah robot. Hasil dari kuesioner yang diisi oleh responden digunakan sebagai acuan kelayakan kinerja Si Tole Chatterbot.

## V. KESIMPULAN DAN SARAN

Dari hasil pengujian yang telah dilakukan, dapat diambil beberapa kesimpulan. Implementasi aplikasi Si Tole Chatterbot pada sistem operasi Android menggunakan basis data ALICE yang ada pada *web server* dapat dilakukan dengan baik. Respons dan data yang dikirimkan melalui Android sesuai dengan masukan atau keluaran yang diinginkan. RESTful API pada *web server* yang diimplementasikan dengan menggunakan teknologi JSON, cukup cepat untuk menampilkan respons dari ALICE pada aplikasi Android.

Untuk penelitian selanjutnya, disarankan dilakukan penambahan media gambar/icon pada percakapan serta penambahan fitur masukan dan keluaran suara.

## UCAPAN TERIMA KASIH

Terima kasih disampaikan ke DRPM Direktorat Jendral Pendidikan Tinggi atas bantuan dana yang telah diberikan sehingga penelitian ini dapat terlaksana dalam program Penelitian Terapan Unggulan Perguruan Tinggi tahun anggaran 2017.

## REFERENSI

- [1] Wallace, R.S., *The elements of AIML style*, ALICE A.I. Foundation, Inc., 2003.
- [2] Wallace, R.S., *The Anatomy of A.L.I.C.E.*, Artificial Intelligence Foundation, Inc., 1996.

- [3] (2003) A.L.I.C.E AI Foundation, [Online], <http://www.alicebot.org/>, tanggal akses: 2017.
- [4] Abu Shavar, Atwell, "Using the Corpus of Spoken Afrikaans to generate an Afrikaans chatbot", *Southern African Linguistics and Applied Language Studies*, Vol. 21, No. 4, hal. 283–294, 2003.
- [5] Mikic, Burguillo, Rodríguez, Llamas, "T-BOT and Q-BOT: A Couple of AIML-based Bots for Tutoring Courses and Evaluating Student", *ASEE/IEEE Frontiers in Education Conference*, 2008, Session S3-A, hal.7-12.
- [6] Cho, K. Alvin, "Emotional and Domain Concept Enhancements to Alicebot", Master's Project, Department of Computer Science, San José State University, San Jose, USA, 2007.
- [7] Fatnuriyah, M., Sarosa, M., Santoso, P.B., "Implementasi Chatterbot Berbasis Program O Menggunakan Web Crawler dan Web Service", *Jurnal ELEKTRAN*, Vol. 2, No. 2, hal. 1-7, Des. 2012.
- [8] Sarosa, M., Junus, M., Hoesny, M.U., Sari, Z., Fatnuriyah, M., "Classification Technique of Interviewer-Bot Result Using Naïve Bayes and Phrase Reinforcement Algorithms", *International Journal of Emerging Technologies in Learning* (accepted), 2017.
- [9] Junus M., Sarosa M., Fatnuriyah M., Hoesny M.U., Sari Z., "Interviewer BOT Design to Help Student Learning English," *International Conference on Information Technology and Security (IC-ITechS)*, 2014, hal. 45-50.