

Multi Respon Ranking Pada Percakapan Layanan Travel Berdasarkan Riwayat Obrolan

Uli Rizki

Program Pasca Sarjana Universitas Amikom Yogyakarta
Jl. Ring Road Utara, Condong Catur, Sleman, Yogyakarta, (0274) 884201 - 207
uli.rizki@students.amikom.ac.id

Abstrak

Membangun sebuah chatbot tidak lepas dari basis pengetahuan. Basis pengetahuan bisa diperoleh dari data yang telah diberi label oleh pengembang, dokumen yang telah di konversi menjadi pre-processing data, atau informasi yang diambil dari media sosial. Dalam hal ini, data yang digunakan sebagai pengetahuan adalah riwayat obrolan. Dalam riwayat obrolan tersebut tentu banyak sekali variasi jawaban dan memungkinkan satu pertanyaan menimbulkan banyak jawaban.

Untuk mengatasi multi respon tersebut maka dibuatlah pemeringkatan respon. Adanya pemeringkatan, tentu respon yang diinginkan oleh pengguna akan lebih sesuai. Tantangan dalam pemeringkatan adalah bagaimana mendapatkan inti dari sebuah pertanyaan dan menemukan pasangan tanya jawab dari data. Hal tersebut bisa diselesaikan dengan model sequence to sequence. Namun, permasalahan yang akan muncul adalah konsistensi jawaban. Adanya banyak riwayat obrolan dari layanan travel tentu menimbulkan banyak jawaban meski inti dari pertanyaan adalah sama. Untuk itu Algoritma CNN sebagai solusi untuk permasalahan tersebut. Penelitian ini menggunakan convolutional sequence to sequence yang akan diterapkan untuk respon ranking. Kami membandingkan efisiensi dari model ini. Dengan melakukan perbandingan, model ini menunjukkan hasil akurasi sebesar 86.7%.

Kata kunci: Chatbot, Sequence to sequence, CNN

1. PENDAHULUAN

Chatbot merupakan salah satu program dari kecerdasan buatan dalam simulasi percakapan interaktif kepada pengguna (manusia) melalui text, suara atau visual. Dalam mengenali dan memberikan respon yang alami layaknya percakapan manusia, chatbot menggunakan pendekatan *Natural Language Processing* (NLP). Tantangan dalam penerapan NLP diantaranya yaitu ambiguitas bahasa alami, representasi pengetahuan, tingkatan informasi, dan beragam aplikasi untuk teknologi bahasa.

Akurasi dalam menyajikan jawaban dalam sebuah percakapan robot terus dilakukan. Berbagai pendekatan dan model meningkatkan efisiensi dalam mengenali pola percakapan dan memberikan jawaban dengan cepat dari dasar pengetahuan pada sebuah percakapan manusia dan mesin. Salah satu pendekatan yang sering dilakukan adalah menyandingkan pasangan teks dengan menggunakan fitur leksikal, sintaksis dan semantik yang kemudian menghitung ukuran kesamaan antara representasi pada sebuah kalimat.

Untuk memberikan respon layaknya percakapan manusia, chatbot mengandalkan basis pengetahuan yang sudah tersusun atau masukan dari pengembangnya. Hal tersebut memberikan masalah jika membutuhkan banyak sekali pengetahuan. Kemudian, untuk meningkatkan akurasi biasanya menggunakan metode pelabelan pada setiap pola kalimat. Ini akan menjadi pekerjaan yang melelahkan jika harus memberikan label pada gudang data dan tidak efektif jika diterapkan untuk skala industri. (Bowen Wu, dkk)[1]

Pendekatan lain yang digunakan adalah menggunakan *framework sequence to sequence* atau urutan konteks percakapan (Jiwei Li, dkk, 2017)[2]. Metode ini di implementasikan ketika menggunakan data yang berasal dari percakapan group atau riwayat obrolan tertentu. Dalam percakapan sebuah group obrolan tentunya akan banyak jawaban yang berbeda tergantung dari personal penjawab. Adanya urutan (*sequence*) tentunya akan menemukan jawaban dari pertanyaan yang sama. Tetapi metode urutan ini memiliki kelemahan pada konsistensi respon. Respon yang diberikan tergantung

dari urutan yang pada suatu waktu diperbaharui sehingga dengan pertanyaan yang sama dengan waktu berbeda akan memberikan respon berbeda pula.

Permasalahan konsistensi dapat disolusikan dengan pemeringkatan respon dengan menggunakan algoritma CNN (Akshay Chaturvedi, dkk, 2018)[3]. Kemudian untuk meningkatkan relevansi sesuai dengan permintaan pengguna, ditambahkan fitur lain yaitu dengan pendekatan konten (Bingquan Liu, dkk, 2018)[4]. Pendekatan konten tersebut mengambil profil pengguna, baik dari input ke sistem maupun mengambil dari akun media sosial terkait profil pribadi pengguna. Profil pengguna yang dihubungkan dengan konten yang di posting, tanggapan, dan informasi pribadi digunakan sebagai bahan pertimbangan perspektif pengguna sehingga respon akan dipilihkan lebih tepat untuk orang tertentu.

Banyaknya penelitian terdahulu serta banyaknya referensi mengenai pemeringkatan respon *chatbot* dengan menggunakan algoritma CNN dengan beberapa model pendekatan. Namun, dari penelitian-penelitian sebelumnya mayoritas menggunakan dataset dari percakapan lebih luas dan dengan bahasan yang belum spesifik. Oleh karena itu, penelitian ini akan menggunakan data pengetahuan riwayat obrolan dalam pemeringkatan respon sistem percakapan dengan menggunakan model pendekatan *sequence to sequence* dan menggunakan algoritma CNN.

2. METODE PENELITIAN

a. Jenis, Sifat, dan Pendekatan Penelitian

1) Jenis Penelitian

Chatbot bukanlah sebuah hal baru dalam machine learning seperti saat sekarang ini, bahkan sudah banyak diimplementasikan di banyak industri layanan online. Banyak sekali penelitian untuk melatih percakapan sistem layaknya percakapan alami manusia. Penelitian dengan berbagai model dan pendekatan mampu untuk mengambil informasi tidak hanya dari basis pengetahuan yang disediakan didalam sistem chatbot sendiri tetapi sudah mengambil informasi dari luar bahkan mengambil profil pribadi berdasarkan akun media sosial dibenamkan dalam sistem chatbot itu sendiri.

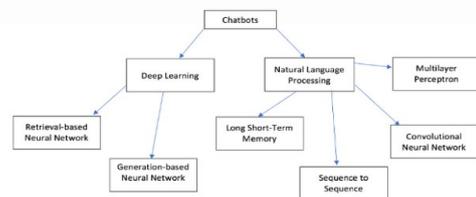
Dalam hal ini, penelitian ini dilakukan dengan melakukan eksperimen dengan cara menerapkan model untuk respon ranking chatbot dengan menggunakan riwayat obrolan sebagai basis pengetahuan chatbot tersebut.

2) Sifat Penelitian

Penelitian ini bersifat kausal yaitu model yang akan diterapkan memberikan hasil bagaimana pendekatan ini bekerja. Dan memperoleh informasi seberapa besar akurasi yang bisa ditunjukkan dengan menggunakan model yang diterapkan.

3) Pendekatan Penelitian

Dalam pembelajaran sistem terdapat banyak metode yang bisa digunakan tergantung dari permasalahan yang akan diselesaikan. Dimulai dari Natural Language Processing dan Deep Learning. Dua konsep besar tersebut bisa ditarik lagi metode yang ditunjukkan pada gambar dibawah ini.



Gambar 1. Ilustrasi Peta Konsep Chatbot

Ada terdapat banyak metode dalam membaca informasi dalam sebuah chatbot seperti terlihat pada gambar 3, namun pada penelitian ini akan dikhususkan pada penggunaan metode Sequence to Sequence dan Convolutional Neural Network. Metode seq2seq akan memproses urutan sebuah kalimat dan algoritma CNN digunakan sebagai hidden layer yang akan menghitung atau me-ranking kalimat sebagai kandidat respon yang paling berhubungan.

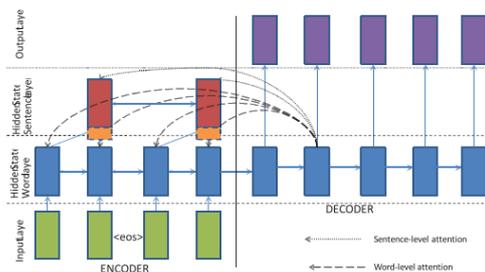
Secara garis besar langkah-langkah yang dilakukan untuk melakukan respon ranking pada sebuah sistem percakapan adalah dengan memecah sebuah kalimat berdasarkan urutan-urutan kata. Yang selanjutnya akan di filter menggunakan algoritma CNN dengan beberapa layer

tersembunyi untuk melihat atau mencari kata yang paling relevan dengan basis pengetahuan, dalam hal ini riwayat obrolan.



Gambar 2. Proses Respon ranking

Seperti yang terlihat dalam gambar 4, model re-ranking menggunakan seq2seq dan CNN akan mengadopsi dari penelitian Abstract Text Summarization with a Convolutional Seq2seq Model, Zhang dkk, MDPI 2019. Pertama sebuah kalimat dari user masuk ke input layer, disini akan dilakukan proses pemecahan kalimat. Selanjutnya, akan masuk ke Hidden State Wordaye untuk dilakukan perangkikan yang didukung dengan Hidden State Sentence Layer untuk proses filtering.



Gambar 5. Mekanisme Respon Ranking

b. Metode Pengumpulan Data

Dalam pelaksanaan penelitian mengenai respon ranking ini tentunya dibutuhkan berbagai komponen pendukung demi kelancaran penelitian. Salah satu komponen pendukung yang tidak bisa dipisahkan adalah data sebagai bahan ujicoba.

Sebagai salah satu bahan ujicoba dalam penelitian ini digunakan dataset riwayat obrolan dari layanan jasa penjualan tiket online di salah satu perusahaan startup di Yogyakarta.

Terdapat lebih dari 10000 percakapan untuk layanan penjualan tiket pesawat. Dalam riwayat obrolan tersebut merupakan obrolan dari 10 customer service dalam pelayanan dari tahun 2018.

c. Metode Analisis Data

1) Data Training

Misalkan dataset dengan komponen data N, komponen $(x^{(i)}, t^{(i)})$ terdiri dari teks sumber direpresentasikan sebagai $x^{(i)}$ dan kata kunci target M_i atau kalimat kunci yang dilambangkan sebagai

$$t^{(i)} = (t^{(i),1}, t^{(i),2}, \dots, t^{(i),M_i})$$

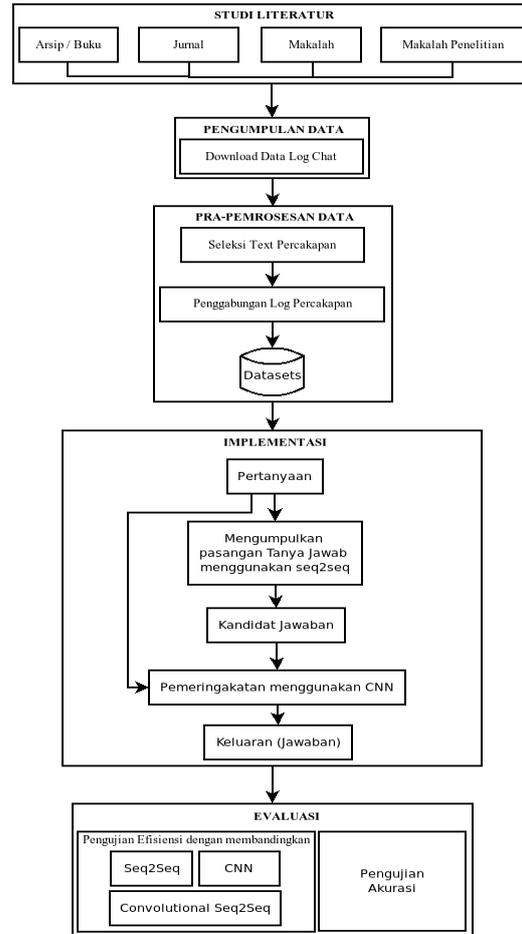
2) Urutan Kalimat

Untuk urutan input $x = (x_1, \dots, x_m)$, pertama-tama kita menyatakannya sebagai vektor dimensi rendah $u = (u_1, \dots, u_m)$, di mana $u_j \in R^d$. Sedangkan untuk embeddings posisi, pertama-tama kita memperoleh vektor untuk merekam posisi absolut suatu elemen dalam suatu urutan. Kemudian kami menggunakan lapisan embedding untuk mengubah representasi jarang dan diskrit ini menjadi embeddings kontinu sebagai $p = (p_1, \dots, p_m)$, di mana $p_j \in R^d$, sehingga memungkinkan untuk mengetahui bagian urutan mana yang sedang diproses. d adalah dimensi dari posisi embeddings dan embeddings elemen urutan input. Kami menggunakan kombinasi kedua embeddings untuk membentuk embeddings elemen input $e = (u_1 + p_1, \dots, u_m + p_m)$. Demikian pula kami juga menambahkan posisi embeddings ke embeddings elemen keluaran decoder, dan kemudian mengembalikan ke decoder.

3) Convolutional Seq2Seq Model

Kami membagi arsitektur lapisan convolutional pada encoder dan decoder yang akan menghitung status perantara melalui elemen input. Kami merepresentasikan output dari layer ke- l sebagai $z^l = (z_1^l, \dots, z_m^l)$ untuk encoder dan $h^l = (h_1^l, \dots, h_m^l)$ untuk decoder.

Setiap lapisan terdiri dari konvolusi satu dimensi dan non-linearitas. Jika decoder memiliki satu layer dengan lebar kernel menjadi k , maka outputnya h_1 akan saya kompres informasi elemen input k . Untuk memperbesar panjang elemen input, kami menumpuk blok satu sama lain, misalnya, menumpuk 6 blok dengan $k = 5$ dapat mewakili 25 elemen input. Jika diperlukan, non-linearitas memungkinkan model kami untuk menangani seluruh urutan input atau hanya beberapa elemen. Keunggulan komputasi model kami adalah melakukan komputasi paralel yang jauh lebih efisien daripada model RNN tradisional yang dikomputasi satu elemen dengan satu elemen. Seperti disebutkan dalam Bagian 1, untuk mewakili urutan dengan kata-kata, CNN hanya membutuhkan operasi $O(n/k)$, sedangkan RNN membutuhkan operasi $O(n)$. Namun, model CNN hierarkis ini tidak seefisien model CNN tradisional karena model kami harus menumpuk lapisan CNN untuk mewakili urutan yang lebih ekspresif, sementara model CNN tradisional hanya membutuhkan satu lapisan untuk menjelajahi seluruh urutan.



Gambar 3. Alur Penelitian

3. TINJAUAN PUSTAKA

Severyn, dkk melakukan penelitian mengenai arsitektur CNN untuk merangkai pasangan teks pendek dan fungsi kesamaan untuk dikaitkan dengan kata-kata yang menjadi dasar pengetahuan berdasarkan data training yang tersedia. Kelebihan metode yang digunakan mampu mengalahkan sistem state-of-art sebelumnya yaitu 3% poin absolut di Mean Average Precision (MAP) dan Mean Reciprocal Rank (MRR). Penelitian ini juga menjawab tantangan dari forum tanya jawab TREC.

Wu, dkk juga melakukan penelitian dengan menggunakan pendekatan *Sequential Matching Network* (SVM). Penelitian ini mencocokkan respon dengan setiap konteks dalam sebuah ucapan dari pengguna, dan menyaring pencocokan penting sebagai vektor dengan operasi pengumpulan. Vektor tersebut diakumulasikan dalam sebuah urutan melalui Recurrent Neural Network (RNN) yang memodelkan hubungan antar ucapan.

d. Alur Penelitian

Tahap-tahap penelitian respon ranking chatbot adalah sebagai berikut :

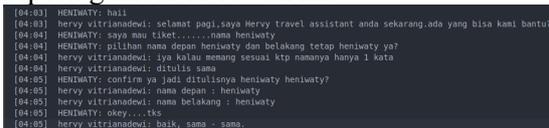
Bowen, dkk melakukan penelitian menggabungkan lapisan konvolusional dengan memahami petunjuk semantik dalam percakapan, berdasarkan representatif efektif untuk setiap kalimat. Dalam hal ini, peneliti juga menambahkan lapisan berbasis GRU untuk membangun model peringkat kandidat sehingga mengambil permodelan kalimat menggunakan CNN dan RNN sebagai permodelan urutan. Dengan permodelan tersebut mendapatkan akurasi sebesar 78,6%.

Penelitian terkait juga diimplementasikan pada tanya jawab dalam sebuah *e-commerce*. Dalam sebuah percakapan tanya jawab permasalahan yang timbul adalah mencocokkan sumber pengetahuan dengan pemrosesan bahasa alami. Untuk itu peneliti mengusulkan kerangka kerja baru yang dapat secara efektif dan efisien menyesuaikan pengetahuan yang dibagikan dari domain yang kaya sumber pengetahuan ke sumber yang kurang pengetahuan. Hasilnya, performa dari pelayanan *e-commerce* meningkat dengan $F1=0.681$ dengan waktu 80.7 per query . (Jianfei Yu, dkk, 2018)

4. HASIL DAN PEMBAHASAN

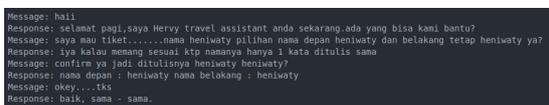
a. Membuat Datasets

Pertama, data kami ambil dari log chat dari email yang berekstensi .eml, terlihat seperti gambar di bawah ini.



Gambar 4. Log Chat dari email

Untuk melakukan penggabungan teks percakapan, penulis menggunakan skrip kode dengan bahasa python. Setelah dilakukan penggabungan maka data terlihat seperti dibawah ini.



Gambar 5. Chat yang sudah digabungkan

b. Word Vector

Untuk menghasilkan vektor kata, kami menggunakan pendekatan klasik model Word2Vec. Ide dasarnya adalah bahwa model menciptakan vektor kata dengan melihat konteks dengan kata-kata yang muncul dalam kalimat. Kata-kata dengan konteks yang sama akan ditempatkan berdekatan dalam ruang vektor.

c. Sequence to Sequence Model

Setelah melakukan pra-pemrosesan data maka masuk ke dalam implementasi yaitu dengan mengidentifikasi pertanyaan yang diajukan oleh pengguna chatbot.

Misalnya user mengajukan pertanyaan seperti “Apakah harga sudah termasuk bagasi ?”

Dalam proses ini pertanyaan tersebut akan

Apakah	harga	sudah	termasuk	bagasi	?
h1	h2	h3	h4	h5	<EOS>

dikonversi sebagai berikut:

Gambar 6. Identifikasi pertanyaan

Kemudian dengan perhitungan dengan menggunakan rumus yang telah dijelaskan pada 2.3.2 maka pertanyaan tersebut yang akan dijadikan sebagai *encoder*. Dan kemudian mesin akan menemukan pasangan pertanyaan dan jawaban. Jawaban tersebut adalah bentuk dari *decoder*. Seperti gambar

Apakah	harga	sudah	termasuk	bagasi	?	chilink	sudah	termasuk	bagasi
x1	x2	x3	x4	x5	<EOS>	y1	y2	y4	y4

dibawah ini.

Gambar 7. Bentuk encoder dan decoder seq2seq

d. Convolutional Neural Network

Setelah dilakukan pengecekan pasangan pertanyaan dan jawaban, maka kandidat respon akan diranking. Namun, dalam proses pemeringkatan, pertanyaan akan difilter lagi menggunakan algoritma CNN.

1) One-hot vector

Dalam tahap ini, kalimat atau pertanyaan akan diubah menjadi angka 1 atau 0 berdasarkan jumlah kata dalam pertanyaan yang nantinya akan menjadi sebuah matrix yang akan melalui beberapa proses perhitungan. Perubahan *one-hot vector*, dapat ditunjukkan tabel 1.

Tabel 1. One-hot Vector

	One-hot vector					
	Apakah	harga	sudah	termasuk	bagasi	?
apakah	1	0	0	0	0	0
harga	0	1	0	0	0	0
sudah	0	0	1	0	0	0
termasuk	0	0	0	1	0	0
bagasi	0	0	0	0	1	0
?	0	0	0	0	0	1

2) Convoluted Feature

Kemudian setiap vektor akan dikalikan dengan kernel yang mana akan memasang kata-kata dalam pertanyaan. Kami menggunakan contoh kernel 3x3 untuk menghitung pasangan kata yang akan diubah sebagai matrix.

Tabel 2. Kernel 3x3

Kernel 3x3		
1	0	1
0	1	0
1	0	1

Tabel 3. Convoluted Feature

Convoluted Feature			
3	0	1	0
0	3	0	1
1	0	2	0
0	1	0	3

Kemudian hasil perkalian dengan kernel ditunjukkan tabel 3.

3) Max Polling

Convoluted Feature akan di filter lagi dengan lapisan 2x2 untuk mendapatkan nilai maksimal dalam setiap lapisan. Dari tabel 3 akan diambil nilai maksimal dari 4 kolom mendapat *max-polling* seperti tabel 4.

Tabel 4. Max Polling

Max Polling	
3	1
3	2
1	3

e. Eksperimen

Dengan menggunakan dataset dari riwayat percakapan model yang diusulkan dapat meningkatkan akurasi. Jika hanya menggunakan *seq2seq* lebih sedikit akurasi karena banyaknya kandidat jawaban yang berbeda. Dengan tambahan CNN dapat meningkatkan akurasi dalam pemeringkatan respon sebuah chatbot.

Tabel 5. Hasil TestModel

Model	Akurasi
Seq2seq	68.20%
CNN	84.10%
Conv-seq2seq	86.7%

5. KESIMPULAN DAN SARAN

a. Kesimpulan

Dengan menggabungkan dua model yaitu *seq2seq* dan CNN ternyata dapat meningkatkan akurasi sebesar 86.7%. Dengan menggunakan riwayat obrolan. Jika hanya menggunakan *seq2seq* sangat memungkinkan jawaban yang berulang sesuai dengan penelitian sebelumnya.

b. Saran

Untuk kedepannya, bisa menambah dataset atau mengurangi dataset untuk menemukan kondisi terbaik ketika menggunakan model *convolusional seq2seq* dalam pemeringkatan respon sebuah mesin chatbot.

6. REFERENSI

- Chaturvedi, A., Pandit, O., & Garain, U. (2018). CNN for Text-Based Multiple Choice Question Answering. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 2, 272–277. Retrieved from <https://aclanthology.coli.uni-saarland.de/papers/P18-2044/p18-2044>
- Liu, B., Xu, Z., Sun, C., Wang, B., Wang, X., Wong, D. F., ... Zhang, M. (2018). *Content-Oriented User Modeling for Personalized Response Ranking in Chatbots*. 26(1), 122–133.
- Wu, B., Wang, B., & Xue, H. (2016). Ranking Responses Oriented to Conversational Relevance in Chat-bots. *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 652–662. Retrieved from <https://www.aclweb.org/anthology/C16/C16-1063.pdf>
- Wu, Y., Wu, W., Xing, C., Zhou, M., & Li, Z. (2016). *Sequential Matching Network: A New Architecture for Multi-turn Response Selection in Retrieval-based*

Chatbots. Retrieved from
<http://arxiv.org/abs/1612.01627>
Yu, J., Qiu, M., Jiang, J., Huang, J., Song, S.,
Chu, W., & Chen, H. (2017). *Modelling
Domain Relationships for Transfer
Learning on Retrieval-based
Question Answering Systems in E-
commerce*. 682–690. Retrieved from
<http://arxiv.org/abs/1711.08726>