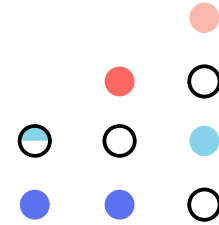


IMPLEMENTASI MODEL LSTM PADA CHATBOT INFORMASI KERAGAMAN BUDAYA INDONESIA MENGGUNAKAN BLACK BOX TESTING

+ Oleh:

Muhammad Tegar Lazuardi (Jupyter XXI)
Mochamad Ali Fajar (Cordoba)
Dwi Alda Agustini (Cordoba)
Luki Iswanto (Cordoba)
Ajeng Dyanti Putri (Leibniz)



PENDAHULUAN

- Manusia merupakan makhluk yang tidak lepas dari kebudayaan karena kebudayaan ada dan lestari jika manusia ikut menerapkannya dalam kehidupan sehari-hari. Oleh karenanya, diperlukan sarana untuk mempermudah masyarakat untuk mengakses informasi mengenai kebudayaan. Salah satunya adalah dengan memanfaatkan AI (Artificial Intelligence) berupa aplikasi chatbot.
- Dari pengembangan media informasi dan pembelajaran berbasis chatbot ini diharapkan dapat meningkatkan kesadaran masyarakat khususnya generasi muda bangsa Indonesia dengan memberikan kemudahan dalam hal pengenalan dan pemahaman mengenai kebudayaan Indonesia.
- Berdasarkan permasalahan yang telah diuraikan, pada penelitian ini akan dibuat sebuah chatbot menggunakan algoritma Long Short-Term Memory (LSTM) serta menggunakan pengujian Blackbox untuk melakukan training dan testing model.

TINJAUAN PUSTAKA

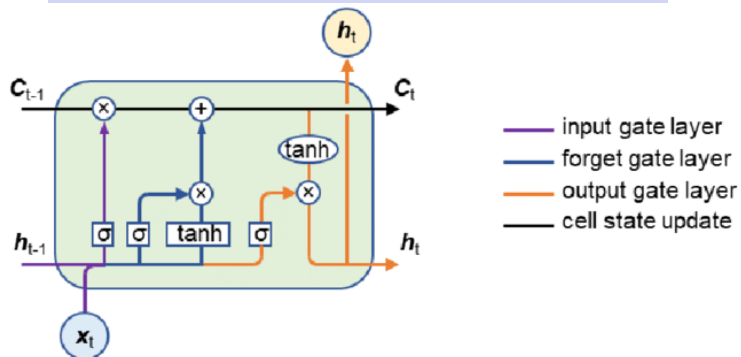
1. *Natural Language Processing (NLP)*

NLP adalah teknik pemodelan yang mengkaji struktur suatu Bahasa. Peran NLP disini adalah untuk melihat struktur data yang diproses agar dapat memberikan jawaban dalam bentuk Bahasa.

2. *Recurrent Neural Network (RNN)*

RNN merupakan arsitektur jaringan saraf tiruan atau biasa disebut jaringan saraf berulang. Model ini bekerja dengan menggunakan nilai neuron pada hidden layer secara berulang sebagai masukan.

Alur perulangan 4 layer LSTM



3. Long Short-Term Memory (LSTM)

LSTM adalah salah satu jenis dari *Machine Learning* yang merupakan bagian dari model RNN. Perbedaannya terletak pada penambahan sel memori. Sel memori membantu menyimpan informasi untuk waktu yang lama. Secara umum LSTM terdiri dari *memory cell*, *input gate*, *output gate*, dan *forget gate*. Adapun alur keempat layer tersebut dapat dilihat pada gambar disamping.

a Memory Cell

Mengambil *input* dan menyimpannya dalam beberapa waktu.

b Input Gate

Mengontrol sejauh mana nilai baru akan berjalan ke dalam *cell*

c Forget Gate

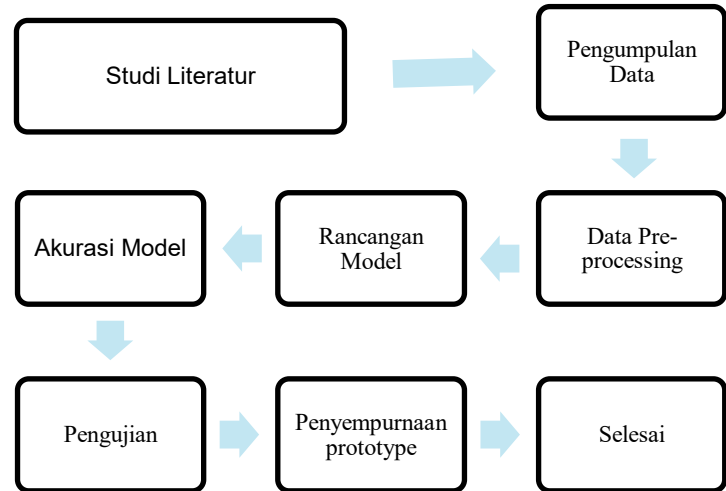
Mengontrol sejauh mana nilai tetap di dalam *cell*

d Output Gate

Mengontrol sejauh mana nilai dalam *cell* digunakan untuk menghitung aktivasi keluaran dari unit LSTM.

METODE PENELITIAN

Pada penelitian ini, objek penelitian yang digunakan adalah Google Colaboratory. Pelayanan yang diberikan google colab ialah fasilitas bahasa pemrograman python dari Google. Fasilitas ini memungkinkan para pengguna Google Colab untuk berkolaborasi dalam membuat program dengan python. Adapun alur dari penelitian yang akan dilakukan adalah sebagai berikut:



METODE PENELITIAN

1. Pengumpulan Data

Tahap pertama yang dilakukan dalam penelitian ini adalah pengumpulan kebutuhan dataset. Data yang digunakan adalah kumpulan dari pertanyaan-pertanyaan terkait kebudayaan Indonesia

2. Data Pre-processing

Proses pengolahan text menggunakan teknik dari natural language processing (NLP) atau disebut dengan pengolahan bahasa alami manusia.

3. Rancangan Model

Pada penelitian ini, pemodelan yang digunakan untuk membuat *chatbot* ialah *Long Short-Term Memory*. Pelatihan yang dilakukan menggunakan parameter *epoch*.

4. Akurasi Model

Untuk mendapatkan akurasi model yg bagus penulis melakukan proses pelatihan data secara berulang kali secara bertahap sehingga didapatkan hasil nilai kesalahan yg semakin kecil dan nilai akurasi yang paling besar.

METODE PENELITIAN

5. Pengujian Sistem

Proses pengujian yang dilakukan penulis adalah mencocokkan apakah output yang di keluarkan sudah sesuai dengan yang diharapkan menggunakan *blackbox testing*.

6. Penyempurnaan Prototipe

Tahapan ini dilakukan untuk merubah data sehingga dapat menyempurnakan prototype yang dibangun agar menghasilkan prototype yang sesuai dengan kebutuhan.

PEMBAHASAN

DATASET

Dataset dibuat dengan mengambil daftar pertanyaan terkait kebudayaan Indonesia. Ada 780 pertanyaan yang sudah dikumpulkan. Data tersebut kemudian diubah kedalam bentuk file .json. Dataset dibuat secara manual kedalam bentuk file .json dengan rancangan sebagai berikut:

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

PEMBAHASAN

AKURASI DAN PEMODELAN LSTM

Pelatihan data dilakukan dengan meningkatkan epoch dari 200, 300, 400 dan 500. Berikut tabel perbandingan nilai akurasi yang didapatkan pada setiap epoch yang digunakan.

No	Epoch	Loss	Akurasi
1	200	0.1132	0.9177
2	300	0.1070	0.9191
3	400	0.0970	0.9218
4	500	0.0964	0.9246

Hasil yang didapatkan dari pelatihan dengan *epoch* sebanyak 500 menunjukkan model yang lebih baik dengan tingkat akurasi paling besar 0.9246 dan *loss* 0.0964 pada *epoch* 500.

PENGUJIAN SISTEM

Pengujian sistem dilakukan dengan penguji blackbox menggunakan 31 pertanyaan. Hasil pengujian pengembangan model *chatbot* berbasis teks menunjukkan akurasi yang cukup baik sebesar 93.54%.

05

KESIMPULAN

- Pada penelitian dengan model LSTM menghasilkan tingkat *accuracy* 0.9246 dan *loss* 0.0964 pada 500 *epoch*.
- Dengan model LSTM, chatbot mampu memprediksi dan memberikan jawaban sesuai pertanyaan pengguna.
- Hasil pengujian menggunakan 31 pertanyaan, menunjukkan hasil yang cukup baik dengan total 29 jawaban benar, dan 2 jawaban salah. Akurasi dari pengujian menggunakan *blackbox testing* adalah 93.54% dan kesalahan 6.46%.
- Aplikasi *chatbot* memudahkan masyarakat khususnya generasi muda untuk mempelajari budaya bangsa Indonesia secara online dimana pun dan kapan pun.

SARAN

- Saran untuk pengembangan penelitian selanjutnya, untuk meminimalisir kesalahan ejaan kata dan Bahasa yang tidak baku dari pertanyaan pengguna, penulis menyarankan untuk meningkatkan "*Sense of Human*".
- Diharapkan pada masyarakat agar mau memaksimalkan fasilitas yang ada seperti *chatbot* karena dengan mudahnya dapat kita gunakan dimana pun dan kapan pun untuk mencari informasi.

Budi Cahyo Suryo, Putro S., et al. "Improved Time Series Prediction Using LSTM Neural Network for Smart Agriculture Application." *2019 5th International Conference on Science and Technology (ICST)*, IEEE, 2019, pp. 1–4. DOI.org (Crossref), <https://doi.org/10.1109/ICST47872.2019.9166401>.

Fatonah, Fatonah. "BELAJAR DARI KARAKTERISTIK BANGSA JEPANG DALAM MENGHARGAI KEBUDAYAAN." *Tsaqofah Dan Tarikh: Jurnal Kebudayaan Dan Sejarah Islam*, vol. 2, no. 2, Dec. 2017, p. 119. DOI.org (Crossref), <https://doi.org/10.29300/ttjksi.v2i2.711>.

Jawed, Soyiba, et al. "EEG Visual and Non- Visual Learner Classification Using LSTM Recurrent Neural Networks." *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, IEEE, 2018, pp. 467–71. DOI.org (Crossref), <https://doi.org/10.1109/IECBES.2018.8626711>.

Yadav, Anita, et al. "Optimizing LSTM for Time Series Prediction in Indian Stock Market." *Procedia Computer Science*, vol. 167, 2020, pp. 2091–100. DOI.org (Crossref), <https://doi.org/10.1016/j.procs.2020.03.257>.

https://slidesgo.com/

Thanks!



Please keep this
slide for attribution



CREDITS: This presentation
template was created by **Slidesgo**,
including icons by **Flaticon**,
and infographics & images by **Freepik**

