

Implementasi Chatbot Berdasarkan Sistem Tanya Jawab dalam Program Python dengan Menggunakan Model LSTM dan Model BiLSTM = Chatbot Implementation Based on Question and Answer System in Python Program Using LSTM Model and BiLSTM Model

Prasnurzaki Anki, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20509571&lokasi=lokal>

Abstrak

Pada zaman modern ini, implementasi chatbot digunakan untuk menyimpan data yang dikumpulkan melalui sistem tanya jawab, yang dapat diterapkan dalam program Python. Data yang akan digunakan dalam program ini adalah Cornell Movie Dialog Corpus yang merupakan dataset yang berisi korpus ini berisi kumpulan percakapan fiksi kaya metadata yang besar yang diekstraksi dari skrip film. Penerapan chatbot dalam program Python, dapat menggunakan berbagai macam model, secara spesifik pada program ini akan diterapkan model LSTM, dan model BiLSTM. Penerapan chatbot dalam program Python, dapat menggunakan berbagai macam model, secara spesifik pada program ini akan diterapkan model LSTM, dan model BiLSTM. Hasil output dari program chatbot dengan penerapan model LSTM, dan BiLSTM adalah berupa akurasi, serta kumpulan data yang sesuai dengan informasi yang pengguna masukkan dalam input kotak dialog chatbot. Pemilihan model yang dapat diterapkan berdasarkan karakteristik data dapat mempengaruhi kinerja program, dengan tujuan program agar dapat menentukan tinggi atau rendahnya tingkat akurasi yang akan dihasilkan dari hasil yang diperoleh melalui sebuah program, yang dapat dijadikan faktor utama dalam menentukan model yang dipilih. Berdasarkan pertimbangan yang menjadi syarat pemilihan model dari sebuah program, pada akhirnya dipilih model LSTM, dan BiLSTM sebagai model yang akan diterapkan ke dalam program. Selain pemilihan model, berikutnya adalah menentukan metode yang digunakan dalam program, pada program ini dipilih metode greedy sebagai bentuk implementasi model LSTM dan model BiLSTM, dengan tujuan ketika dalam menjalankan program, waktu pengolahan data dapat lebih cepat, dan meningkatkan akurasi pada model yang dipilih pada program. Selain itu, atribut pendukung seperti seq2seq model, menjadi faktor penentu dalam sebuah program yang dapat berfungsi untuk memverifikasi pengolahan data apakah sesuai dengan kriteria yang dapat dijadikan sebagai pedoman dalam pengolahan data. Dalam penerapan komponen-komponen tersebut ke dalam program, seq2seq model dapat memproses kalimat input yang kemudian akan dilakukan pengolahan data tersebut menggunakan model dan struktur lain yang ada pada program, sehingga pada akhirnya dapat menghasilkan kalimat output yang berbagai macam, sebagai respon atas kalimat input yang dihasilkan dari program chatbot. Selain itu diperlukan metode evaluasi program yang dapat digunakan untuk memverifikasi apakah hasil output program sesuai dengan data yang diharapkan oleh pengguna. Berdasarkan penerapan model LSTM, dan model BiLSTM ke dalam chatbot, dapat disimpulkan bahwa dengan semua hasil uji program yang terdiri dari beragam pasangan parameter yang berbeda, maka dinyatakan Pasangan Parameter 1 (size_layer 512, num_layers 2, embedded_size 256, learning_rate 0.001, batch_size 32, epoch 20) yang berasal dari File 6 merupakan BiLSTM Chatbot dengan nilai avg accuracy 0.995217 yang menggunakan model BiLSTM menjadi pasangan parameter terbaik.

In modern times, chatbots are implemented and used to store data collected through a question and answer system which can be applied in the Python program. The data used in this program is the Cornell Movie Dialog Corpus which is a dataset containing a corpus that contains

a large collection of metadata-rich fictional conversations extracted from film scripts. The application of chatbots into the Python program can be done using various models. In this research we specifically use the LSTM and BiLSTM models. The output results from the chatbot program with the application of the LSTM and BiLSTM models are in the form of accuracy, as well as a data set that matches the information that the user enters in the chatbot dialog box input. The choice of models that will be applied is based on data that can affect program performance, with the target of the program that can determine the high or low level of accuracy that will be generated from the results obtained through the program, which is a major factor in determining the selected model.

Based on the considerations that are the required for choosing the model for the program, in the end the LSTM and the BiLSTM models are chosen and will be applied to the program. After selecting the appropriate model, the next step is to determine the method used in the program. The greedy method is chosen as a form of implementation of the LSTM and BiLSTM models that aims to decrease the data processing time of the program and make it quicker, and also increase the accuracy of the model selected for the program. In addition, supporting attributes such as the seq2seq model are a determining factor in a program that functions to verify whether data processing process matches the criteria and can be used as a guide. In applying these components to the program, the seq2seq model processes the input sentences which will then be processed using the models and other structures in the program, so that in the end it can produce various output sentences in response to the input sentences that are generated from the chatbot program. In addition, a program evaluation method is needed to verify whether the program output matches the data expected by the user. Based on the application of the LSTM dan BiLTSM models into the chatbot program, it can be concluded that between all the program test results consisting of a variety of different parameter pairs, it is stated that Parameter Pair 1 (size_layer 512, num_layers 2, embedded_size 256, learning_rate 0.001, batch_size 32, epoch 20) from File 3 is the best paramater pair of the BiLSTM Chatbot which uses the BiLTSM model, with the avg accuracy value of 0.995217.