

Perbandingan convolutional neural network dan multihead attention dengan recurrent neural network sebagai multiclass keyword spotting pada edge device = Comparison of convolutional neural network and multihead attention with recurrent neural network as multiclass keyword spotting on edge devices.

Anandwi Ghurran Muhajjaln Arreto, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20516390&lokasi=lokal>

Abstrak

Artificial Intelligence (AI) telah berkembang sangat pesat sehingga sudah sering terlihat dan digunakan secara umum oleh masyarakat. Salah satu jenis AI yang sering digunakan adalah speech recognition terutama keyword spotting yang disebabkan karena pandemi COVID-19. Implementasi keyword spotting dapat diterapkan pada lift sebagai sistem navigasi agar para pengguna lift tidak perlu melakukan kontak pada tombol, melainkan dapat menggerakkan lift hanya dengan mengucapkan lantai yang dituju. Metode untuk melakukan implementasi keyword spotting pada sistem lift dapat dilakukan dengan banyak metode, namun pada skripsi ini, metode yang diujikan adalah CNN (Convolutional Neural Network) dan MHAtt RNN (Multihead Attention Recurrent Neural Network). Penelitian yang dilakukan memiliki batasan untuk setiap metode agar dapat melakukan klasifikasi enam keyword dan melihat performa kedua metode dalam berbagai skenario yang dapat terjadi dalam lift. Dalam pembentukan model dari MHAtt RNN, dapat diketahui bahwa model memiliki performa terbaik ketika dibentuk dengan jumlah head untuk attention sebesar 8 dan LSTM dengan jumlah unit sebanyak 32. Pelatihan pada model dilakukan menggunakan optimizer Adam dengan learning rate sebesar 0.001 dan decay 0.005 agar pelatihan dapat menghasilkan model yang paling baik. Setelah melakukan pengujian pada berbagai skenario yang dapat terjadi di dalam sebuah lift, didapatkan hasil bahwa secara keseluruhan model CNN memiliki performa yang lebih baik dibandingkan model MHAtt RNN karena memiliki nilai F1-score dan precision yang lebih tinggi.

.....Artificial Intelligence (AI) has grown so rapidly that it has often been seen and used in general by the public. One type of AI that is often used is speech recognition, especially keyword spotting caused by the COVID-19 pandemic. The implementation of keyword spotting can be applied to elevators as a navigation system so that elevator users do not need to make contact with buttons but can move the elevator just by saying the intended floor. There are many methods to implement keyword spotting in elevator systems, but in this thesis, the methods tested are CNN (Convolutional Neural Network) and MHAtt RNN (Multihead Attention Recurrent Neural Network). The research conducted has limitations for each method in order to be able to classify six keywords and see the performance of both methods in various scenarios that can occur in an elevator. In forming the model from MHAtt RNN, it can be seen that the model has the best performance when it is formed with the number of heads for attention of 8 and the LSTM with the number of units of 32. The training on the model is carried out using the Adam optimizer with a learning rate of 0.001 and a decay of 0.005 so that the training can produce the best models. After testing on various scenarios that can occur in an elevator, the results show that the CNN model overall has better performance than the MHAtt RNN model because it has a higher F1-score and precision.