

## INTISARI

### **ANALISIS KLASIFIKASI TUTUR MENGGUNAKAN *CONVOLUTIONAL NEURAL NETWORK, DEEP LONG SHORT TERM MEMORY, DAN REGRESI LOGISTIK MULTINOMIAL***

Oleh :

Muhammad Allafa Budi Pratama

17/412743/PA/18062

Kesalahan pada *auto-generated subtitle* Youtube menggunakan Bahasa Indonesia masih sering terjadi dan merugikan penikmat Youtube yang mengalami gangguan pendengaran sehingga pembelajaran mengenai klasifikasi tutur dilakukan. Penelitian ini bertujuan mempelajari klasifikasi tutur dengan menggunakan data suara pengucapan kata barat, selatan, timur, dan utara. Penelitian ini melalui dua tahapan besar yaitu ekstraksi ciri dan analisis klasifikasi. Ekstraksi ciri data suara yang digunakan yaitu ekstraksi ciri *Mel Frequency Cepstrum Coefficient* (MFCC). Analisis klasifikasi yang digunakan yaitu *Convolutional Neural Network* (CNN), *Deep Long Short Term Memory* (DLSTM), dan regresi logistik multinomial. Ketiga metode klasifikasi ini akan dibandingkan performa modelnya berdasarkan nilai akurasi dan *loss function* dengan tiap metode dilatih dengan 3 skenario pembagian data *training*, *validation*, dan *testing*.

Pada analisis CNN, dibangun 597.316 parameter untuk dilatih dan menghasilkan klasifikasi terbaik yaitu pada skenario pembagian jumlah data *training*, data *validation*, dan data *testing* sebesar 64:16:20 dengan nilai akurasi sebesar 97,31% pada data *training*, 91,11% pada data *validation*, dan 99,07% pada data *testing*. Pada analisis DLSTM, dibangun 215.644 parameter untuk dilatih dan menghasilkan klasifikasi terbaik yaitu pada skenario pembagian jumlah data *training*, data *validation*, dan data *testing* sebesar 64:16:20 dengan nilai akurasi sebesar 91,31% pada data *training*, 84,63% pada data *validation*, dan 98,24% pada data *testing*. Pada analisis regresi logistik multinomial, dibentuk 3 model berdasarkan skenario dan menghasilkan klasifikasi terbaik yaitu pada skenario pembagian jumlah data *training*, data *validation*, dan data *testing* sebesar 64:16:20 dengan nilai akurasi sebesar 27,06% pada data *training*, 24,03% pada data *validation*, dan 24,92% pada data *testing*. Didapatkan kesimpulan hasil klasifikasi menggunakan metode CNN lebih baik jika dibandingkan dengan metode DLSTM, terbukti dari plot akurasi dan *loss function* yang menunjukkan hasil klasifikasi DLSTM kurang stabil pada skenario pembagian jumlah data *training*, data *validation*, dan data *testing* sebesar 64:16:20.

**Kata Kunci :** *Speech Classification, Convolutional Neural Network, Deep Long Short Term Memory, Regresi logistik multinomial, MFCC.*

## **ABSTRACT**

### ***SPEECH CLASSIFICATION ANALYSIS USING CONVOLUTIONAL NEURAL NETWORK, DEEP LONG SHORT TERM MEMORY, AND MULTINOMIAL LOGISTIC REGRESSION***

By :

Muhammad Allafa Budi Pratama

17/412743/PA/18062

*Errors in auto-generated Youtube subtitles using Indonesian still often occur and are detrimental to Youtube viewers who have hearing loss, so learning about speech classification is carried out. This study aims to study speech classification by using voice data for the pronunciation of the words west, south, east, and north. This research went through two major stages, namely feature extraction and classification analysis. Feature extraction of voice data used is feature extraction of Mel Frequency Cepstrum Coefficient (MFCC). The classification analysis used is Convolutional Neural Network (CNN), Deep Long Short Term Memory (DLSTM), and multinomial logistic regression. The three classification methods will compare the performance of the model based on the value of accuracy and loss function with each method being trained with 3 scenarios for sharing data, training, validation, and testing.*

*In the CNN analysis, 597,316 parameters were built to be trained and resulted in the best classification, namely in the scenario of dividing the amount of training data, validation data, and testing data of 64:16:20 with an accuracy value of 97.31% on training data, 91.11% on data validation, and 99.07% on data testing. In the DLSTM analysis, 215,644 parameters were built to be trained and produced the best classification, namely in the scenario of dividing the amount of training data, validation data, and testing data of 64:16:20 with an accuracy value of 91.31% on training data, 84.63% on data validation, and 98.24% on data testing. In multinomial logistic regression analysis, 3 models were formed based on scenarios and resulted in the best classification, namely in the scenario of dividing the amount of training data, data validation, and data testing of 64:16:20 with an accuracy value of 27.06% on training data, 24.03 % on data validation, and 24.92% on data testing. It is concluded that the classification results using the CNN method are better than the DLSTM method, as evidenced by the accuracy and loss function plots which show that the DLSTM classification results are less stable in the scenario of dividing the amount of training data, data validation, and data testing of 64:16:20.*

**Key Words :** *Speech Classification, Convolutional Neural Network, Deep Long Short Term Memory, Multinomial logistic regression, MFCC.*