

ABSTRACT

Utilization of this system is also growing, as well as learning tools. Speech Recognition accuracy is influenced by the algorithm used in recognizing sound. The aim of this research is to improve Speech Recognition accuracy by adding the K-Nearest Neighbor method or Dynamic Time Warping algorithm, determine the optimal k value used and calculate the comparison of Speech Recognition accuracy level with DTW algorithm only and DTW added K-NN method.

This research is an experimental type of research, conducted to determine the effect of giving a treatment or treatment of research subjects. This research is an experimental type because the object under study is processed through an experiment using the DTW algorithm plus the K-NN method for the process of introducing English vocabulary and seeing the results of its accuracy.

Previous Speech Recognition research using DTW algorithm produced a fairly large accuracy of 95.85%, but still has a weakness that is the lack of training data in the dataset that can affect the results of the matching process and the accuracy value. The disadvantage of DTW is that the greater number of dictionaries ensures that an increase in the average successful speech recognition process results in longer processing times but higher accuracy, so you have to choose between accuracy or speed. Based on this, the classification algorithm is included, K-NN, because one of the main functions of classification is for the introduction / classification / prediction of an object so that it is suitable to be used as the final validation in a Speech Recognition system.

K-NN algorithm works based on a query or nearest neighbor by entering the value 'k'. The value of k entered is very influential on the results of system accuracy. The smaller the value of k entered will make the classification results more rigid because it takes into account some of the closest neighbors. Meanwhile, if too many k values are entered which will result in a vague classification meaning too many and too far the neighbors are calculated. In this study the k values included are 5, 10, 11, 20, 25, 50 and 75. From the several k values that have been tested, the results can be concluded that the value of k = 5 is the most optimal value because it has the highest level of accuracy. that is 99% compared to the other k values. The accuracy rate is 3.15% higher compared to previous studies that only use the MFCC and DTW Algorithm without any Classification.

Keywords: Dynamic Time Warping, K-Nearest Neighbor, Speech Recognition, Word Recognition.

ABSTRAK

Seiring dengan perkembangan teknologi, kebutuhan akan adanya suatu sistem yang mampu mengenali, menganalisis dan mengidentifikasi suatu sinyal suara pun semakin tinggi. Pemanfaatan sistem ini juga semakin berkembang, seperti halnya sebagai sarana pembelajaran. Tingkat akurasi *Speech Recognition* dipengaruhi oleh algoritma yang digunakan dalam mengenali suara. Tujuan yang ingin dicapai pada penelitian ini adalah untuk meningkatkan akurasi *Speech Recognition* dengan melakukan penambahan metode atau algoritma K-Nearest Neighbour pada algoritma Dynamic Time Warping, menentukan nilai k optimal yang digunakan dan menghitung perbandingan tingkat akurasi *Speech Recognition* dengan algoritma DTW saja dan DTW yang ditambah metode K-NN.

Penelitian ini merupakan penelitian jenis eksperimental, dilakukan untuk mengetahui pengaruh pemberian suatu treatment atau perlakuan terhadap subjek penelitian. Penelitian ini berjenis eksperimental karena objek yang diteliti diproses melalui suatu eksperimen yang menggunakan algoritma DTW ditambah metode K-NN untuk proses pengenalan kosa kata bahasa Inggris dan melihat hasil akurasinya.

Penelitian *Speech Recognition* sebelumnya dengan algoritma DTW menghasilkan akurasi yang cukup besar yaitu 95,85%, namun masih memiliki kelemahan yaitu kurangnya jumlah data pelatihan (*data training*) dalam dataset yang dapat mempengaruhi hasil dari proses pencocokan dan nilai akurasi. Kelemahan dari DTW yaitu makin banyak jumlah kamus yang memastikan bahwa kenaikan rata-rata sukses proses pengenalan suara mengakibatkan waktu pemrosesan yang makin lama namun ketepatan semakin tinggi, jadi harus memilih antara ketepatan atau kecepatan. Berdasarkan hal tersebut maka dimasukan algoritma klasifikasi yaitu K-NN, karena salah satu fungsi utama dari klasifikasi yaitu untuk pengenalan / klasifikasi / prediksi pada suatu objek sehingga cocok digunakan sebagai validasi akhir dalam suatu sistem *Speech Recognition*.

Algoritma K-NN bekerja berdasarkan query atau tetangga terdekat yaitu dengan memasukan nilai 'k'. Nilai k yang dimasukan sangat berpengaruh terhadap hasil akurasi sistem. Semakin kecil nilai k yang dimasukan akan membuat hasil klasifikasi semakin kaku karena memperhitungkan beberapa tetangga terdekat saja. Sedangkan jika terlalu banyak nilai k yang dimasukan akan menghasilkan klasifikasi yang samar artinya terlalu banyak dan terlalu jauh tetangga yang di hitung. Pada penelitian ini nilai k yang dimasukan yaitu 5, 10, 11, 20, 25, 50 dan 75. Dari beberapa nilai k yang telah diuji, hasilnya dapat disimpulkan bahwa nilai $k = 5$ adalah nilai k yang paling optimal karena memiliki tingkat ketepatan paling tinggi yaitu sebesar 99% dibandingkan dengan nilai k yang lainnya. Tingkat akurasi tersebut 3,15% lebih tinggi dibandingkan dengan penelitian sebelumnya yang hanya menggunakan Algoritma MFCC dan DTW tanpa adanya Klasifikasi.

Kata Kunci: Dynamic Time Warping, K-Nearest Neighbour, Pengenalan Suara, Pengenalan Kata.