

Pengenalan komponen imbuhan dan kata dasa pada isyarat kata berimbuhan dalam sibi (sistem isyarat bahasa indonesia dengan menggunakan probabilistic graphical models) = Recognition of affixes and root words in inflectional word gestures in sibi (sign language system for Indonesian language using probabilistic graphical models)

Erdefi Rakun, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20446734&lokasi=lokal>

Abstrak

ABSTRAK

SIBI merupakan bahasa isyarat resmi bagi penyandang tunarungu di Indonesia. Dalam pembentukan isyarat, SIBI mengikuti aturan tata bahasa Indonesia. Untuk membentuk isyarat kata berimbuhan, maka isyarat imbuhan awalan, akhiran dan partikel ditambahkan ke isyarat kata dasar. Karena banyak isyarat SIBI merupakan isyarat kata berimbuhan dan belum ada penelitian tentang kata tersebut, maka penelitian ini fokus pada membangun sistem penerjemah kata berimbuhan SIBI ke teks. Gerakan isyarat ditangkap oleh kamera Kinect yang menghasilkan data color, depth dan skeleton. Data Kinect ini diolah menjadi fitur yang dipakai oleh model untuk mengenali gerakan. Sistem penerjemah memerlukan teknik ekstraksi fitur, yang dapat menghasilkan sebuah feature vector set dengan ukuran yang minimal. Penelitian ini berusaha untuk dapat memisahkan isyarat imbuhan dan kata dasar pada isyarat kata berimbuhan. Dengan kemampuan ini, sistem penerjemah menghasilkan 3 feature vector set: kata dasar, awalan dan akhiran. Tanpa pemisahan, feature vector set yang harus disediakan adalah sebanyak perkalian cartesian dari ketiga feature vector set tersebut. Perkalian ketiga set ini tentunya akan menghasilkan feature vector set total yang berukuran sangat besar. Model yang dicoba pada penelitian ini adalah Conditional Random Fields, Hidden Markov Model, Long Short-Term Memory Neural Networks LSTM dan Gated Recurrent Unit. Akurasi yang terbaik yang dicapai oleh untuk LSTM 2-layer 77.04 . Keunggulan dari LSTM terletak pada inputnya yang berupa sequence-of-frames dan setiap frame direpresentasi oleh fitur lengkap, bukan fitur hasil clustering. Model sequence-of-frames lebih cocok untuk SIBI, karena gerakan isyarat SIBI memiliki long-term temporal dependencies. Error hasil prediksi banyak terjadi pada kelompok awalan dan akhiran. Hal ini karena miripnya gerakan pada isyarat-isyarat imbuhan SIBI tersebut. LSTM 2-layer yang dipakai untuk mengenali kata dasar saja memberikan akurasi yang tertinggi 95.4 .

<hr />

ABSTRACT

SIBI is the official sign language system for the Indonesian language. The formation of SIBI gestures follow Indonesian grammar rules, including inflectional words. Inflectional words are root words with prefixes, infixes, and suffixes, or a mix of the three. Inflectional gestures are made from root word gestures, with prefix, suffix and particle gestures added in the order in which they appear, all of which is unique to SIBI. This research aims to find a suitable model that can quickly and reliably perform SIBI to text translation on inflectional word gestures. The hand movement of the signer is captured by a Kinect camera. The Kinect data was then processed to yield features for the models to use recognize the gestures. Extant research have been able to translate the alphabet, root words, and numbers from SIBI to text, but none has been able to translate SIBI inflectional word gestures. In order for the translation system to work as efficiently as

possible, this research developed a new method that splits an inflectional word into three feature vector sets: root, prefix, and suffix. This ensures that a minimally descriptive feature set is used. Without using this, the feature sets would otherwise be as big as the Cartesian product of the prefixes, suffixes, and root words feature sets of the inflectional word gestures. Four types of machine learning models were tested: Conditional Random Fields, Hidden Markov Model, Long Short Term Memory Net, and Gated Recurrent Unit. The 2 layer LSTM, with an accuracy of 77.04%, has been proven to be the most suitable. This model's performance is due to the fact that it can take entire sequences as input and doesn't rely on pre-clustered per frame data. The 2 layer LSTM performed the best, being 95.4% accurate with root words. The lower accuracy with inflectional words is due to difficulties in recognizing prefix and suffix gestures.