

## INTISARI

### Klasifikasi Artikel Berita Berbahasa Indonesia Menggunakan *Convolutional Neural Network*

Oleh

AR RAZI

14/371918/PPA/04617

Perkembangan teknologi informasi dan komunikasi saat ini mengakibatkan penyebaran informasi dalam bentuk berita digital telah mengalami pertumbuhan yang sangat pesat. Aliran informasi terutama berita dalam bentuk digital terus menerus bertambah dalam setiap hari sehingga *volume* berita terus bertambah banyak dengan cepat. Mengelola dokumen berita dengan dengan *volume* yang sangat besar tidak mudah dilakukan karena membutuhkan waktu yang sangat lama, sehingga dibutuhkan suatu sistem otomatisasi yang dapat digunakan untuk mengelola dan mengelompokkan dokumen berita tersebut. Pengelompokan dokumen berita dibutuhkan untuk mempermudah dan mempercepat dalam pencarian suatu informasi.

Penelitian ini bertujuan untuk mengelompokkan artikel berita berbahasa Indonesia menggunakan lima kelas yaitu kelas Entertainment, Kesehatan, Olahraga, Teknologi dan Ekonomi dengan mengimplementasikan salah satu metode machine learning yang dapat digunakan untuk klasifikasi dokumen text yaitu CNN (*Convolutional Neural Network*). Sebelum dilakukan klasifikasi, terlebih dahulu dilakukan pengubahan kata-kata ke dalam bentuk vektor dengan menggunakan *word2vec*. sehingga hasil dari pengubahan tersebut dapat diinputkan ke dalam CNN.

Hasil pengujian pada sistem yang dibangun memperlihatkan bahwa kombinasi metode *Convolutional Neural Network* dan *word2vec* memberikan hasil akurasi yang lebih baik daripada metode *naive bayes*, dengan nilai akurasi sebesar 96.70 %, dan presisi, *recall* serta *f-measure* mencapai 96,60 %.

**Kata kunci:** klasifikasi, berita, *word2vec*, *Continuous bag-of-word*, *skip-gram*, *convolutional neural Network*

## ABSTRACT

### **Classification Of Indonesian News Articles Using Convolutional Neural Network**

by

AR RAZI

14/371918/PPA/04617

The current development of information and communication technology resulted in the dissemination of information in the form of digital news has experienced a very rapid growth. The flow of information, especially news in digital form continues to grow in every day so that the volume of news continues to multiply rapidly. Managing news documents with a very large volume is not easy to do because it takes a very long time, so it takes an automation system that can be used to manage and classify the news documents. Grouping of news documents is needed to simplify and speed up the search for information.

This research aimed to categorization Indonesian language news articles using the five classes namely class Entertainment, Health, Sports, Technology and Economics by implementing a machine learning method that can be used for classification of text documents that CNN (Convolutional Neural Network). Before the classification, first performed the conversion of words into a vector form using word2vec. so the result of that change can be entered into CNN.

The test results of system that is built to show that a combination of methods Convolutional Neural Network and word2vec provide accuracy results better than Naive Bayes method, the value of accuracy of 96.70 % and precision, recall, f-measure reach 96,60 %.

**Keyword:** categorization, news, word2vec, Continuous bag-of-word, skip-gram, convolutional neural Network