

ABSTRACT

IMPLEMENTASI *RECURRENT NEURAL NETWORK* PADA *MULTICLASS TEXT CLASSIFICATION* JUDUL BERITA

Oleh
Bagus Bramantyo
19312039

Currently information can be obtained in a relatively easy way with the existence of internet technology, one of which is by accessing online news portals. Online news portals can make it easier to obtain information because they are on demand so they can display the desired information such as political, economic, health, and sports events. With the dissemination of information on a very large scale, sometimes we find it difficult to find news that matches the desired content, this happens because there is a lot of news content that must be categorized by the author manually so that errors can occur which result in news being mixed up between one topic and another. other. This research was conducted to determine the performance of the Deep Learning method using a Recurrent Neural Network (RNN) in carrying out multi-classification tasks on news headlines on the topics of Economics, Health, Sports and Politics. The news titles used as training data and test data were obtained from news portals using the Web Scraping method taken from March 2021 to March 2023, a total of 3004 news titles with a ratio of 80% used as training data and 20% as test data. Next, the Text Preprocessing process will be carried out which aims to prepare data before the model training process. The stages include case folding, tokenization, and stopwords removal. Next, a feature extraction process will be carried out using the TF-IDF method so that the training data can be inputted into the model for training. The results of model performance testing using the Confusion Matrix method obtained an accuracy value of up to 97%, so that the RNN method can be used for multi-classification tasks on news headlines quite well and can be applied to news title classification systems.

Keywords: *Confusion Matrix, Deep Learning, Multiclassification, News Title, Recurrent Neural Network, Text Preprocessing, TF-IDF, Web Scraping*

INTISARI

IMPLEMENTASI *RECURRENT NEURAL NETWORK* PADA *MULTICLASS TEXT CLASSIFICATION* JUDUL BERITA

Oleh
Bagus Bramantyo
19312039

Saat ini informasi dapat diperoleh dengan cara yang relatif mudah dengan adanya teknologi internet, salah satunya yaitu dengan mengakses portal berita online. Portal berita online dapat mempermudah dalam memperoleh informasi karena bersifat *on demand* sehingga dapat menampilkan informasi yang diinginkan seperti peristiwa politik, ekonomi, kesehatan, hingga olahraga. Dengan penyebaran informasi dalam skala yang sangat besar terkadang kita merasa kesulitan dalam mencari berita yang sesuai dengan konten yang diinginkan, hal ini terjadi karena banyaknya konten berita yang harus dikategorikan oleh author secara manual sehingga dapat terjadi kesalahan yang mengakibatkan berita tercampur antara topik satu dengan topik yang lainnya. Penelitian ini dilakukan untuk mengetahui kinerja dari metode *Deep Learning* menggunakan *Recurrent Neural Network* (RNN) dalam melakukan tugas multiklasifikasi pada judul berita dengan topik Ekonomi, Kesehatan, Olahraga, dan Politik. Judul berita yang digunakan sebagai data latih dan data uji diperoleh dari portal berita menggunakan metode *Web Scraping* yang diambil dari bulan maret 2021 sampai dengan bulan maret 2023 yang berjumlah 3004 judul berita dengan perbandingan 80% digunakan sebagai data latih dan 20% sebagai data uji. Selanjutnya akan dilakukan proses *Text Preprocessing* yang bertujuan untuk mempersiapkan data sebelum proses pelatihan model. Tahapannya yaitu meliputi case folding, tokenization, dan stopwords removal. Selanjutnya akan dilakukan proses ekstraksi fitur menggunakan metode *TF-IDF* agar data latih dapat diinputkan kedalam model untuk dilakukan pelatihan. Hasil dari pengujian kinerja model menggunakan metode *Confusion Matrix* diperoleh nilai akurasi mencapai 97%, sehingga metode RNN dapat digunakan untuk tugas multiklasifikasi pada judul berita dengan cukup baik dan dapat diterapkan pada sistem pengklasifikasian judul berita.

Kata kunci: *Confusion Matrix, Deep Learning, Judul Berita, Multiklasifikasi, Recurrent Neural Network, Text Preprocessing, TF-IDF, Web Scraping*