

ANALISIS METODE *NORMALIZED COMPRESSION DISTANCE* (NCD) DALAM DETEKSI PERNYATAAN FAKTA

Nur Muhammad Al Akbari
10/302577/PA/13437

INTISARI

Banyaknya berita palsu yang beredar akan memberikan dampak negatif dalam masyarakat. Hal ini terlihat dari kecenderungan masyarakat yang sangat suka mendapatkan informasi khususnya secara *online*. Seiring dengan hal tersebut berkembang berbagai macam situs berita *online* yang sangat membantu dalam mendapatkan informasi secara aktual. Informasi yang ada pun tidak hanya berasal dari satu aspek saja, melainkan dari banyak aspek seperti politik, ekonomi, budaya, kesehatan dan masih banyak lagi lainnya. Semua aspek tersebut ada aspek yang jika telaah lebih lanjut, merupakan aspek penting dalam kehidupan, yaitu aspek kesehatan. Tentunya informasi dalam aspek kesehatan ini tidak boleh bersifat palsu atau tidak benar, melainkan harus bersifat fakta agar tidak berdampak buruk pada kehidupan sehari-hari.

Pada penelitian ini, akan diterapkan identifikasi dan pengklasifikasian berita kesehatan tersebut sehingga dapat diketahui mana yang bersifat fakta dan mana yang bersifat *hoax*. Metode yang dipakai adalah metode *Normalized Compression Distance* (NCD). Metode tersebut akan memproses file berita yang dijadikan sumber dan file berita input untuk dicari nilai similaritas nya. Proses kompresi kedua file tadi nantinya akan menghasilkan sebuah nilai antara 0-1 yang kemudian diolah selanjutnya agar dapat menggolongkan berita tersebut ke dalam berita fakta maupun *hoax*. Hasil dari data yang sudah diproses tadi akan dilakukan pengujian dengan perhitungan *precision*, *recall* dan *accuracy* untuk melihat apakah metode efektif jika diterapkan pada kasus ini.

Hasil penelitian ini menunjukkan bahwa metode *Normalized Compression Distance* (NCD) kurang tepat jika digunakan untuk pendeteksian *hoax* dalam sebuah berita karena tidak memenuhi beberapa syarat uji secara sistematis dan logika. Hal ini dikarenakan metode *Normalized Compression Distance* (NCD) hanya mencari nilai similaritas dalam struktur file yang di proses dan tidak relevan jika digunakan untuk membandingkan file input yang masuk dan file sumber guna penentuan berita fakta ataupun *hoax*.

Kata kunci: *Normalized Compression Distance*, *Hoax Detection*, Berita Kesehatan

***NORMALIZED COMPRESSION DISTANCE METHOD ANALYSIS FOR
FACT DETECTION***

Nur Muhammad Al Akbari
10/302577/PA/13437

ABSTRACT

A lot of false news spreading in the community can cause negative effects. It is related to the community's nowadays trend of getting news from the internet. That is why there are a lot of news situs in Indonesia. This news situs covers wide varieties of news such as politics, economics, cultural, and medical-related news. Medical-related news can be crucial to one's life since it is referred a lot. That is why this kind of news can't happen to be false so there won't be any damage done to the readers.

The Identification and classification of false news from medical-related field will be done in this research. Normalized Compression Distance (NCD) will be used for those purpose. This method will proceed the news input and look for the similarity that will be measured from 0-1 scale then it can be determined that the news whether or not is a false news. From the processed data, the precision, recall dan accuracy also being determined to know whether or not this method is suitable to do such analysis.

The result showed that the Normalized Compression Distance is likely to be not accurate to do an analysis for false news since it doesn't fulfill the sistematic and logical criteria. This method can only determine the similarities of the news input that can't be used for the determination of whether or not the news is likely to be false.

Keywords: Normalized Compression Distance, Hoax Detection, Medical-related news