

Penggunaan text mining pada berita daring dan sosial media LAPOR! untuk penggalan informasi kedaulatan pangan di Indonesia = Text mining on online news and social media LAPOR! to extract information on food sovereignty in Indonesia

Yosia Rimbo Deantama, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20482635&lokasi=lokal>

Abstrak

ABSTRAK

Pangan merupakan hak asasi manusia yang harus senantiasa terpenuhi oleh masyarakat dengan daya beli yang sesuai dan mempunyai kualitas pangan yang tinggi dan aman. Hal tersebut mendorong kedaulatan pangan suatu negara, yang secara mandiri memenuhi kebutuhan pangan masyarakatnya berdasarkan sistem pangan yang adil bagi seluruh masyarakat. Peraturan Pemerintah Republik Indonesia Nomor 17 Tahun 2015 yang mewajibkan adanya sistem informasi tentang pangan dan gizi dan teori evolusi *e-government 3.0*. Oleh karena itu salah satu solusi yang mendukung peraturan tersebut dan pendekatan *e-government 3.0* adalah dengan pendekatan *text mining*. Penelitian ini mengolah data dari LAPOR! dan berita daring mengenai kedaulatan pangan untuk mengekstrak informasi dan menemukan pola-pola yang akan menghasilkan informasi tentang kedaulatan pangan di Indonesia sehingga dapat membantu pengambilan keputusan yang berdasar pada data melalui representasi visualisasi berbasis *web*. Jenis analisis informasi yang digunakan adalah Klasifikasi Dokumen untuk penyaringan dokumen, *Named Entity Recognition* yang digunakan untuk mengetahui entitas lokasi dan komoditas pangan dari data tekstual, dan *Topic Modelling* untuk menemukan topik dari sekumpulan teks dokumen berita dan aduan LAPOR!. Algoritma yang dipakai dalam penelitian ini adalah *Conditional Random Fields* dan *Conditional Markov Model* untuk implementasi *Named Entity Recognition*. *Latent Dirichlet Allocation* dan *Non-Negative Matrix Factorization* untuk implementasi *Topic Modelling*. Selain itu *Naïve Bayes*, *Support Vector Machine*, dan *Logistic Regression* digunakan untuk klasifikasi dokumen. Sedangkan pemilihan model ini menggunakan *Conditional Random Field* dengan nilai *F1-score* pada entitas lokasi sebesar 83.85 dan entitas komoditas pangan sebesar 90.98 yang digunakan pada data berita daring, pada data aduan LAPOR!, entitas lokasi menggunakan *Conditional Markov Model* dengan nilai *F1-Score* sebesar 60.35 dan entitas komoditas pangan sebesar 89.74. Pada klasifikasi dokumen, model *Support Vector Machine* dengan fitur *unigram* memiliki nilai presisi sebesar 92.00. Pada *Topic Modelling*, model *Non-Negative Matrix Factorization* memiliki nilai *coherence* yang lebih tinggi daripada *Latent Dirichlet Allocation* pada tiga eksperimen dengan *dataset* yang berbeda. Di samping itu, dilakukan visualisasi tentang kedaulatan pangan berdasarkan pengolahan data tersebut di atas untuk memudahkan pengambilan kebijakan oleh pimpinan seperti Tim Ahli di Kantor Staf Presiden.

ABSTRACT

Food is a human right that must always be fulfilled by the society with the appropriate purchasing power

and high and safe food quality. This encourages food sovereignty of a country, which independently meets the food needs of its people based on a food system that is fair to the entire community. Peraturan Pemerintah Republik Indonesia Nomor 17 Tahun 2015 requires an information system on food and nutrition and the theory of e-government 3.0 evolution. Therefore, one solution that supports these regulations and the e-government 3.0 approach is the text mining approach. This research processes data from LAPOR! and online news on food sovereignty to extract information and find patterns that will produce information on food sovereignty in Indonesia so that it can assist decision-making based on data through web-based visualization representation. The type of information analysis used is Document Classification for document filtering, Named Entity Recognition which is used to find out location entities and food commodities from textual data, and Topic Modeling to find topics from a collection of text news documents and complaints LAPOR !. The algorithm used in this study is Conditional Random Fields and Conditional Markov Models for the implementation of Named Entity Recognition. Latent Dirichlet Allocation and Non-Negative Matrix Factorization for the implementation of Topic Modeling. In addition Naïve Bayes, Support Vector Machine, and Logistic Regression are used for document classification. Whereas the selection of this model uses Conditional Random Field with an F1-score value for location entities of 83.85 and a food commodity entity of 90.98 used in online news data. In the LAPOR! Complaint data, the location entity uses Conditional Markov Model with an F1-Score value of 60.35 and food commodity entities amounting to 89.74. In classifying documents, the Support Vector Machine model with unigram features has a precision value of 92.00. In Topic Modeling, the Non-Negative Matrix Factorization model has a higher coherence value than the Latent Dirichlet Allocation in three experiments with different datasets. In addition, visualization of food sovereignty is based on the processing of the data above to facilitate policy making by leaders such as the Expert Team at the Kantor Staf Presiden.