

INTISIARI

AUTHOR OBFUSCATION UNTUK BAHASA INDONESIA MENGGUNAKAN SYNONYM REPLACEMENT WORD EMBEDDING

Oleh

Adi Nursetya Pratama

17/409422/PA/17729

Author obfuscation merupakan suatu cara untuk melakukan modifikasi dokumen dengan mengubah *writing style* dari dokumen. *Author obfuscation* merupakan salah satu cara untuk menjaga anonimitas author terhadap *authorship attribution*. *Author attribution* merupakan proses untuk mengidentifikasi suatu penulis dari sebuah dokumen yang diberikan, berdasarkan kumpulan dokumen dari penulis yang sudah diketahui. Hal ini tentunya merupakan ancaman bagi kebebasan berpendapat dan privasi. Untuk melawan ancaman tersebut, metode *author obfuscation* diusulkan untuk memodifikasi suatu teks supaya penulisnya sulit diidentifikasi tanpa mengaburkan topik utamanya

Pada penelitian ini, model *author obfuscation* dibuat berbasis *word embedding* untuk memodifikasi artikel berita berbahasa Indonesia agar penulis aslinya tidak teridentifikasi secara tepat oleh model *authorship attribution* dengan tetap menjaga semantik artikel yang dimodifikasi sama dengan aslinya. Pada model ini, setiap artikel akan mengalami prapemrosesan berupa tokenisasi dan PoS *tagging*. Kemudian, akan dideteksi kata yang dijadikan sebagai kata target untuk diganti dengan sinonim yang dihasilkan dari *word embedding* berdasar nilai *cosine similarity* terhadap kata target. Adapun *word embedding* yang digunakan dalam penelitian ini adalah Word2Vec, Glove, dan FastText. Selanjutnya, susunan kata dan kalimat digabungkan kembali menjadi artikel utuh untuk dilanjutkan proses evaluasi.

Model tersebut dievaluasi berdasarkan aspek *safety*, *soundness*, dan *sensibleness*. Dari aspek *safety* model FastText mendapat hasil paling baik karena dapat menurunkan akurasi model *authorship attribution* sebesar 0,1150. Untuk aspek *soundness* model FastText mendapat hasil paling baik dengan kemiripan artikel hasil obfuskasi dengan artikel asli sebesar 0,9935. Namun, untuk aspek *sensibleness* yang dievaluasi secara manual, model Word2Vec yang mendapat hasil paling baik sebesar 2,756 dari skala 1-5. Dari hasil evaluasi ketiga aspek, model FastText yang paling baik meskipun dari segi tata bahasa dan pemilihan diksi masih belum optimal.

Kata kunci : *author obfuscation*, *author attribution*, *word embedding*, *Word2Vec*, *Glove*, *FastText*

ABSTRACT

AUTHOR OBFUSCATION FOR INDONESIAN USING SYNONYM REPLACEMENT WORD EMBEDDING

By

Adi Nursetya Pratama

17/409422/PA/17729

Author obfuscation is a way to modify a document by changing the writing style of the document. Author obfuscation is one way to maintain author anonymity against authorship attribution. Author attribution is the process of identifying an author of a given document, based on a document set of known authors. This is certainly a threat to freedom of expression and privacy. To counter this threat, the author obfuscation method is proposed to modify a text so that the author is difficult to identify without obscuring the main topic.

In this research, the author obfuscation model was made based on word embedding to modify Indonesian news articles so that the original author was not correctly identified by the authorship attribution model while maintaining the semantics of the modified article the same as the original. In this model, each article will undergo preprocessing in the form of tokenization and PoS tagging. Then, the word that becomes the target word will be detected to be replaced with a synonym generated from the word embedding based on the cosine similarity value to the target word. The word embedding used in this research is Word2Vec, Glove, and FastText. Furthermore, the arrangement of words and sentences is recombined into a complete article to be continued with the evaluation process.

The model is evaluated based on aspects of safety, soundness, and sensibleness. From the safety aspect, the FastText model gets the best results because it can reduce the accuracy of the authorship attribution model by 0.1150. For the soundness aspect, the FastText model got the best results with the similarity of the obfuscated article with the original article of 0.9935. However, for the sensibleness aspect that was evaluated manually, the Word2Vec model that got the best results was 2,756 on a scale of 1-5. From the evaluation results of the three aspects, the FastText model is the best, although in terms of grammar and diction selection it is still not optimal.

Keyword : *author obfuscation, author attribution, word embedding, Word2Vec, Glove, FastText*