

Rancang Bangun Forum Diskusi Online untuk E-Learning dengan Opsi Partisipasi Anonim dan Sistem Moderasi Otomatis Berbasis Bidirectional Encoder Representations from Transformers (BERT) = Development of an Online Discussion Forum for E-Learning with an Anonymous Participation Option and Automatic Moderation System Based On Bidirectional Encoder Representations from Transformers (BERT)

Deandra Setyaputri, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920525674&lokasi=lokal>

Abstrak

Dalam pendidikan, partisipasi pelajar dalam kelas dapat menjadi salah satu faktor pendukung proses pembelajaran yang efektif. Demi mendukung partisipasi pelajar, penelitian ini bertujuan untuk mengembangkan sebuah forum diskusi online untuk proses e-learning dengan mengimplementasikan fitur partisipasi anonim dimana pelajar dapat mengunggah post tanpa harus menunjukkan identitas aslinya. Pilihan untuk dapat berpartisipasi secara anonim mampu meningkatkan keinginan pelajar untuk berpartisipasi dalam pembelajaran seperti melalui aksi bertanya, menjawab pertanyaan, dan berpendapat dalam kelas. Namun anonimitas yang ditawarkan dapat mengundang perilaku buruk karena berkurangnya akuntabilitas. Untuk mengatasinya, penelitian ini juga bertujuan untuk mengembangkan sistem moderasi otomatis pada forum diskusi dengan memanfaatkan model deep learning pendeteksi bahasa kasar berbasis Bidirectional Encoder Representations from Transformers atau BERT. Setiap kali pengguna ingin mengirim unggahan ke dalam forum diskusi, model pendeteksi bahasa kasar akan terlebih dahulu mengklasifikasikan teks unggahan tersebut ke dalam kelas 'abusive' jika terdapat unsur kasar, menyinggung, atau mengandung kebencian dan ke dalam kelas 'safe' jika tidak. Sistem akan secara otomatis mencegah suatu unggahan untuk terkirim jika unggahan tersebut diklasifikasikan sebagai 'abusive'. Model pendeteksi bahasa kasar tersebut dilatih dengan melakukan fine-tuning pada IndoBERT, model pre-trained Bahasa Indonesia berbasis BERT, dan IndoBERTweet yang dilatih untuk domain Twitter. Berdasarkan hasil pengujian, model dengan performa terbaik merupakan model hasil fine-tuning IndoBERTweet yang mencapai F1 Score sebesar 91,02%. Durasi waktu yang dibutuhkan oleh model untuk mengeksekusi prediksi bervariasi berdasarkan panjang input, dimana durasi bertambah seiring dengan meningkatnya jumlah karakter pada input, namun maksimum berada di kisaran 1,3 detik karena adanya batasan jumlah token input yang dapat diproses model.

.....In education, students' in-class participation can be one of the supporting factors for effective learning. In order to promote student participation, this study aims to develop an online discussion forum for e-learning that implements an anonymous participation feature where students can upload posts without having to show their real identities. The choice to be able to participate anonymously has been proven to improve students' motivation to participate in the learning process through asking and answering questions and expressing opinions in class. But the anonymity offered can be the cause of several bad behaviors due to the lack of accountability. To handle this, this research will also aim to develop an automatic moderation system for the discussion forums that uses an abusive language classifier deep learning model based on Bidirectional Encoder Representations from Transformers or BERT. Every time a user wants to upload a

post to the discussion forum, the abusive language detection model will first classify the uploaded text into the 'abusive' class if it contains abusive language or hateful content and into the 'safe' class if otherwise. The system will automatically prevent a post from being uploaded if it was classified as 'abusive'. The abusive language classifier model is trained by fine-tuning the IndoBERT model, a pre-trained Bahasa Indonesia model based on BERT, and IndoBERTweet which was trained for the Twitter domain. Based on testing results, the model with the best performance is the fine-tuned IndoBERTweet model which achieved an F1 Score of 91,02%. The duration of time required by the model to execute predictions varies based on the length of the input, where the duration increases as the number of characters in the input increases, but the maximum is around 1.2 seconds due to a limit on the number of input tokens that the model can process.