

Analisis Konten Otomatis Level Kehadiran Kognitif pada Transkrip Teks Diskusi Asinkron Berbahasa Indonesia Menggunakan BERT Language Model = Automated Content Analysis of Cognitive Presence Levels in Indonesian-Language Asynchronous Discussion Transcripts Using BERT Language Model

Mubarik Ahmad, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920546050&lokasi=lokal>

Abstrak

Forum diskusi asinkron adalah salah satu media pembelajaran kolaboratif daring yang mampu mendorong pemikiran kritis, pertukaran gagasan, dan pembentukan pengetahuan. Analisis konten merupakan metode ilmiah yang dapat digunakan untuk mengidentifikasi keterampilan berpikir kritis dari transkrip pada forum diskusi asinkron. Metode analisis konten konvensional membutuhkan tahapan pengodean manual yang membutuhkan banyak waktu dan tenaga. Hal ini dapat mengakibatkan pengajar terlambat dalam memberikan intervensi instruksional karena informasi keterampilan berpikir kritis tidak dapat diperoleh secara cepat.

Penelitian ini mengacu pada kerangka kerja Community of Inquiry (CoI) di mana keterampilan berpikir kritis dioperasionalkan melalui empat level dalam kehadiran kognitif yaitu pemantik diskusi, eksplorasi, integrasi, dan resolusi. Tujuan penelitian adalah mengembangkan model klasifikasi berbasis machine learning yang mampu menganalisis secara otomatis kehadiran kognitif pada transkrip diskusi berbahasa Indonesia. Desain penelitian menggunakan metode campuran kuantitatif dan kualitatif. Data eksperimen berjumlah 1.200 pesan diskusi dari mata kuliah Aljabar Linear di lingkungan pembelajaran bauran. Hasil penelitian menunjukkan bahwa kesiapan mahasiswa dalam mengelola pembelajaran dan lingkungan e-learning berpengaruh signifikan terhadap pengembangan kehadiran sosial dan kehadiran kognitif. Dataset level kehadiran kognitif pada transkrip diskusi asinkron dibangun dengan metode analisis konten yang reliabel kategori hampir sempurna (Cohen's kappa = 0,88). Eksperimen pengembangan model analisis kehadiran kognitif menggunakan sepuluh basis algoritma yaitu XGBoost, Random Forest, Support Vector Machine, Logistic Regression, Naive Bayes, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), IndoBERT-base, IndoBERT-large dan XLM-RoBERTa. Model berbasis IndoBERT-large memiliki performa terbaik dengan akurasi sebesar 0,825. Prototipe sistem Cognipresa (cognitive presence analytics) telah dikembangkan untuk memfasilitasi pengajar dengan menganalisis kehadiran kognitif mahasiswa dalam diskusi secara otomatis. Evaluasi sistem menunjukkan hasil yang menjanjikan dari sisi usability dengan nilai System Usability Scale (SUS) sebesar 80,83.

.....The asynchronous discussion forum serves as a collaborative online learning platform capable of stimulating critical thinking, exchanging ideas, and shaping knowledge. Content analysis is a scientific method that can be employed to identify critical thinking skills from transcripts in asynchronous discussion forums. Conventional content analysis methods entail manual encoding stages, which consume a significant amount of time and effort. This may lead to instructors being delayed in providing instructional interventions due to the inability to swiftly obtain information on critical thinking skills.

This study references the Community of Inquiry (CoI) framework, where critical thinking skills are operationalized through four levels of cognitive presence: triggering event, exploration, integration, and

resolution. The research's objective is to develop a machine learning-based classification model capable of automatically analyzing cognitive presence in Indonesian-language discussion transcripts. The research design incorporates both quantitative and qualitative methods. The experimental data consists of 1,200 discussion messages from the Linear Algebra course in a blended learning environment.

The research findings indicate that students' preparedness in managing learning and e-learning environment significantly influences the development of social presence and cognitive presence. The dataset for cognitive presence at the transcript of asynchronous discussions was constructed using a content analysis method with a reliably almost perfect category (Cohen's kappa = 0.88). An experimental development of the cognitive presence analysis model was conducted using ten algorithmic bases, namely XGBoost, Random Forest, Support Vector Machine, Logistic Regression, Naive Bayes, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), IndoBERT-base, IndoBERT-large, and XLM- RoBERTa. The IndoBERT-large-based model demonstrated the best performance with an accuracy of 0.825. A prototype system called Cognipresa (cognitive presence analytics) has been developed to facilitate educators in automatically analyzing students' cognitive presence in discussions. The system evaluation indicates promising results in terms of usability, with a System Usability Scale (SUS) score of 80.83.