

ABSTRAK

Perkembangan komunikasi digital mendorong meningkatnya interaksi manusia melalui media daring. Hal tersebut memicu dinamika hubungan interpersonal secara lebih mendalam. Salah satu aspek penting dalam interaksi tersebut adalah kedekatan interpersonal yang berkembang melalui proses pendalaman pengungkapan diri. Berdasarkan teori penetrasi sosial, kedekatan interpersonal memiliki struktur bertingkat yang bersifat ordinal. Namun, penelitian sebelumnya masih terbatas pada analisis pengungkapan diri tanpa mengkaji tingkat kedalaman pengungkapan diri sebagai proksi kedekatan interpersonal secara otomatis menggunakan pendekatan *deep learning*.

Penelitian ini bertujuan mengembangkan dan membandingkan model klasifikasi tingkat kedalaman pengungkapan diri sebagai proksi kedekatan interpersonal pada teks berbahasa Indonesia menggunakan pendekatan berbasis *Transformer* dan *ordinal regression*. *Dataset* dikembangkan secara mandiri dan terdiri dari 1.890 kalimat yang dibuat secara manual dan dengan bantuan kecerdasan buatan. Data divalidasi oleh dua ahli psikologi sosial dan melalui *agreement-based filtering* diperoleh 1.739 data berlabel akhir. Penelitian membandingkan lima pendekatan, yaitu TF-IDF dengan Linear SVM sebagai *baseline* tradisional, IndoBERT *Frozen Classifier*, IndoBERT dengan *Softmax* dan *Cross Entropy Loss*, IndoBERT dengan *Softmax* dan *Weighted Kappa Loss*, serta IndoBERT dengan CORAL dan *Binary Cross Entropy Loss* sebagai pendekatan *ordinal regression*. Eksperimen model berbasis IndoBERT dilakukan menggunakan lima *random seed* untuk mengevaluasi *robustness* model. Evaluasi dilakukan menggunakan *F1-score*, MAE, RMSE, QWK, *Spearman rank correlation*, dan *error distance table* dengan pembobotan linear, eksponensial, dan logaritmik, serta uji signifikansi menggunakan *Wilcoxon Signed-Rank Test*.

Hasil penelitian menunjukkan bahwa pendekatan IndoBERT dengan CORAL-3 dan *Binary Cross Entropy Loss* dengan *learning rate* CORAL sebesar 8×10^{-3} memberikan kinerja terbaik. Berdasarkan evaluasi *Best Loss* pada lima *random seed*, model tersebut memperoleh nilai MAE sebesar $0,0529 \pm 0,0026$, RMSE sebesar $0,2622 \pm 0,0158$, QWK sebesar $0,9712 \pm 0,0032$, dan *Spearman* sebesar $0,9710 \pm 0,0038$. Model CORAL-3 juga menghasilkan total *error distance* terendah pada pembobotan linear dan eksponensial, masing-masing sebesar $18,40 \pm 0,89$ dan $57,91 \pm 5,67$. Hasil uji *Wilcoxon Signed-Rank Test* menunjukkan CORAL-3 memiliki keunggulan signifikan dibandingkan TF-IDF dan SVM dan IndoBERT *Frozen Classifier* pada beberapa kondisi pengujian, sedangkan terhadap model *Softmax* tidak signifikan. Hasil penelitian menegaskan bahwa pemodelan ordinal dengan CORAL pada IndoBERT menghasilkan klasifikasi tingkat kedalaman pengungkapan diri yang akurat, stabil, dan lebih mempertimbangkan jarak antar kelas dibandingkan pendekatan klasifikasi konvensional.

Kata kunci : Kedekatan Interpersonal, Pengungkapan Diri, Klasifikasi Teks, *Ordinal Regression*, IndoBERT

ABSTRACT

The rapid development of digital communication has led to an increase in human interactions through online media, fostering more complex interpersonal relationship dynamics. One important aspect of these interactions is interpersonal closeness, which develops through progressively deeper self-disclosure. According to Social Penetration Theory, interpersonal closeness follows a hierarchical structure with an inherent ordinal nature. However, previous studies have primarily focused on self-disclosure analysis without investigating the level of self-disclosure depth as a proxy for interpersonal closeness using automated deep learning approaches.

This study aims to develop and compare models for classifying the depth of self-disclosure as a proxy for interpersonal closeness in Indonesian-language text using Transformer-based and ordinal regression approaches. The dataset was independently developed and consists of 1,890 sentences generated manually and with the assistance of artificial intelligence. The data were validated by two social psychology experts, and agreement-based filtering resulted in 1,739 final labeled data. The study compares five approaches: TF-IDF with Linear SVM as the traditional baseline, IndoBERT Frozen Classifier, IndoBERT with Softmax and Cross Entropy Loss, IndoBERT with Softmax and Weighted Kappa Loss, and IndoBERT with CORAL and Binary Cross Entropy Loss as the ordinal regression approach. The IndoBERT-based models were experimentally evaluated using five random seeds to assess model robustness. Evaluation was conducted using F1-score, MAE, RMSE, QWK, Spearman rank correlation, and an error distance table with linear, exponential, and logarithmic weighting, as well as a significance test using the Wilcoxon Signed-Rank Test.

The results show that the IndoBERT approach with CORAL-3 and Binary Cross Entropy Loss, using a CORAL learning rate of 8×10^{-3} , achieved the best performance. Based on the Best Loss evaluation across five random seeds, the model achieved an MAE of 0.0529 ± 0.0026 , RMSE of 0.2622 ± 0.0158 , QWK of 0.9712 ± 0.0032 , and Spearman of 0.9710 ± 0.0038 . The CORAL-3 model also achieved the lowest total error distance under linear and exponential weighting, with values of 18.40 ± 0.89 and 57.91 ± 5.67 , respectively. The results of the Wilcoxon Signed-Rank Test indicate that CORAL-3 demonstrated a significant advantage over TF-IDF with SVM and the IndoBERT Frozen Classifier under several testing conditions, whereas its differences from the Softmax-based models were not statistically significant. These findings confirm that modeling ordinal structure using CORAL with IndoBERT can produce accurate and stable classification of self-disclosure depth while better accounting for the distance between classes compared with conventional classification approaches.

Keywords : Interpersonal Closeness, Self-Disclosure, Text Classification, IndoBERT, Ordinal Regression