

Proposing an NLP-Based Model for Multi-Class Classification of Holland Personality

Abstrak

Menyesuaikan peran pekerjaan dengan kepribadian individu merupakan isu krusial dalam lingkungan kerja yang dinamis saat ini, di mana ketidakcocokan dapat mengganggu kinerja dan menyebabkan inefisiensi. Banyak karyawan menemukan diri mereka berada dalam peran yang tidak sesuai dengan sifat bawaan mereka, yang mengakibatkan ketidakpuasan, kurangnya keterlibatan, dan tingkat turnover yang tinggi. Meskipun ada kemajuan, organisasi sering kali mengandalkan penilaian tradisional seperti survei dan kuesioner, metode yang dapat mahal dan memakan waktu. Studi ini memperkenalkan solusi yang skalabel dan berbasis data menggunakan analisis teks untuk mengatasi kesesuaian kepribadian dengan peran pekerjaan. Kami mengevaluasi tiga model NLP terdepan—RoBERTa, BERT, dan DistilBERT—yang bertujuan untuk mengklasifikasikan tipe kepribadian Holland melalui profil LinkedIn. Metodologi kami mencakup pengumpulan data, prapemrosesan, penandaan otomatis menggunakan GPT 4.0 mini, dan penyempurnaan model di Google Colab untuk klasifikasi multikelas. Kami mengukur kinerja menggunakan akurasi, presisi, recall, dan F1-score dengan rata-rata makro untuk evaluasi yang seimbang. Hasil menunjukkan bahwa DistilBERT, ketika disesuaikan dengan hiperparameter optimal (tingkat pembelajaran $5e-5$, ukuran batch 16, dan 6 epoch), melampaui model lain, mencapai akurasi 90,77%, dengan presisi, recall, dan F1-score sebesar 0,91. Temuan ini menunjukkan kemampuan DistilBERT dalam mengidentifikasi ciri kepribadian, menandakan pergeseran dari metode tradisional ke solusi yang lebih efisien. Penelitian ini memberikan wawasan bagi perekrutan talenta dan optimasi tenaga kerja, membantu organisasi memahami ciri kepribadian untuk meningkatkan keputusan perekrutan. Studi masa depan dapat memperkuat temuan ini dengan menggunakan dataset yang lebih luas dan mengeksplorasi model alternatif untuk meningkatkan presisi klasifikasi.