

# Penerapan *Long Short-Term Memory* untuk Klasifikasi *Multi-Label* Terjemahan Al-Qur'an dalam Bahasa Indonesia

## Sitasi Artikel

Akbar, I., Faisal, M., & Chamidy, T. (2024). Penerapan *Long Short-Term Memory* untuk Klasifikasi *Multi-Label* Terjemahan Al-Qur'an dalam Bahasa Indonesia. *JOINTECS (Journal of Information Technology and Computer Science)*, 9(1), 41–54. <http://publishing-widyagama.ac.id/ejournal-v2/index.php/jointecs>

## Latar dan Tujuan

Pemahaman ayat-ayat Al-Qur'an sering kali membutuhkan klasifikasi berdasarkan tema tertentu, seperti tauhid, ibadah, akhlak, atau sejarah. Proses ini penting karena satu ayat bisa mengandung lebih dari satu tema. Penelitian ini bertujuan mengembangkan model klasifikasi *multi-label* terjemahan Al-Qur'an bahasa Indonesia menggunakan *Bi-Directional Long Short-Term Memory* (Bi-LSTM) dengan *word embedding Word2Vec*.

## Metode

Data yang digunakan berupa terjemahan resmi Kementerian Agama RI dari tiga surah, yaitu An-Nisa', Al-Maidah, dan Al-An'am dengan total 461 ayat. Data diberi label manual oleh para ahli sesuai empat kategori: Tauhid, Ibadah, Akhlak, dan Sejarah. Tahapan penelitian meliputi *preprocessing* teks (*cleansing, case folding, tokenizing*), pembagian data latih dan uji dalam empat skenario (60:40, 70:30, 80:20, 90:10), pelatihan model Bi-LSTM, serta penerapan *Word2Vec* dengan arsitektur CBOW. Evaluasi model dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *Hamming Loss*.

## Hasil/Temuan Kunci

Model dasar Bi-LSTM tanpa *word embedding* menghasilkan akurasi sekitar 50%. Setelah ditambahkan *Word2Vec* dan tuning *hyperparameter* (*epoch, batch size, learning rate*), performa meningkat signifikan dengan akurasi 70,21%, *precision* 64,31%, *recall* 61,13%, dan *Hamming Loss* 36,52%. Meski lebih baik dari metode tradisional seperti K-NN, *Naïve Bayes*, maupun SVM, hasil ini belum mencapai akurasi sempurna karena kompleksitas bahasa Al-Qur'an dan variasi penafsiran.

## Kontribusi dan Keterbatasan

Kontribusi utama penelitian ini adalah menunjukkan efektivitas Bi-LSTM dengan *Word2Vec* dalam klasifikasi multi-label teks terjemahan Al-Qur'an, yang lebih unggul dibanding metode klasik. Namun, keterbatasannya terletak pada ukuran data yang kecil, keterlibatan label manual yang terbatas, serta belum diuji pada keseluruhan teks Al-Qur'an.

## Takeaway

Pendekatan Bi-LSTM dengan *Word2Vec* terbukti menjanjikan untuk klasifikasi teks keagamaan multi-label. Untuk proyek serupa, penggunaan dataset lebih besar dan eksplorasi model modern seperti *Transformer* atau BERT bisa memberi hasil yang lebih baik.