Actes de CORIA-TALN 2026
Nantes, France, 29 juin au 3 juillet 2026

DiaTAC-Emo : Dialecte Tunisien - Analyse à deux niveaux pour la Classification des Émotions

Latifa Iben Nasr[1] Ismail Badache[2] Abir Masmoudi[1] Lamia Hadrich Belguith[1]
(1) MIRACL Laboratory, University of Sfax, Tunisia
(2) Aix-Marseille University, CNRS, LIS, Marseille, France
latifa.ibennasr@fsegs.usf.tn, Ismail.Badache@lis-lab.fr,
masmoudiabir@gmail.com, lamia.belguith@fsegs.usf.tn

RÉSUMÉ

Les émotions dans la parole spontanée sont authentiques mais se chevauchent souvent, en particulier dans les langues peu dotées comme le dialecte tunisien (DT), qui présente des caractéristiques linguistiques spécifiques. Nous proposons une méthode adaptative qui combine l'analyse des émotions au niveau du document et au niveau des aspects afin de mieux capturer ce chevauchement granulaire. En utilisant des transcriptions issues à la fois de modèles open source et propriétaires, nous appliquons GPT-4o pour l'augmentation des données avec une validation humaine. Gemini 1.5 Flash est utilisé pour l'extraction des aspects, soutenue par une validation humaine et une auto-validation avec injection de contexte. Pour la reconnaissance des émotions à partir de la parole (Speech Emotion Recognition : SER), des modèles préentraînés en arabe ainsi que des LLMs sont évalués ; AraBERT obtient les meilleures performances, avec une exactitude de 77 % au niveau du document et de 79,8 % au niveau des aspects. Les résultats soulignent l'importance de la qualité de la transcription, du choix des modèles et de l'analyse au niveau des aspects pour la SER dans les langues sous-représentées.

ABSTRACT

DiaTAC-Emo : Tunisian Dialect Two-level Analysis for Classification of Emotions

Emotions in spontaneous speech are genuine but often overlap, especially in low-resource languages like Tunisian Dialect (TD), which features unique linguistic traits. We propose an adaptive approach that combines document- and aspect-level emotion analysis to better capture this overlapping granularity. Using transcriptions from both open-source and proprietary models, we apply GPT-4o for data augmentation with human validation. Gemini 1.5 Flash performs aspect extraction, supported by human validation and autoranking with context injection. For Speech Emotion Recognition (SER), Arabic-pretrained models and LLMs are evaluated ; AraBERT performs best, achieving 77% accuracy at the document level and 79.8% at the aspect level. Results highlight the importance of transcription quality, model selection, and aspect-level analysis for SER in underrepresented languages.

MOTS-CLÉS : Reconnaissance des émotions, Transcription, Dialecte tunisien, Niveau du document, Niveau de l'aspect.

KEYWORDS: Emotion recognition, Transcription, Tunisian dialect, Document Level, Aspect Level.

ARTICLE ACCEPTÉ À : The 24th IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology.
URL : https://hal.science/hal-05547882

Actes des 33ème Conférence sur le Traitement Automatique des Langues Naturelles.
Volume 2 : articles déjà publiés, pages 12–12.
Cette œuvre est mise à disposition sous licence Attribution 4.0 International.