

Jinyuan XU

Détection Automatique des Humeurs Dépressives sur les Réseaux Sociaux Chinois

Cette thèse porte sur la détection automatique du risque dépressif à partir de textes issus des réseaux sociaux chinois. Au-delà d'une classification binaire, elle vise à proposer une analyse interprétable, traçable et éthiquement encadrée, afin de mieux comprendre l'expression de la dépression hors contexte clinique. Deux difficultés sont centrales : le manque de ressources spécialisées en chinois et l'opacité des approches génératives.

Trois enjeux structurent la recherche. Sur le plan scientifique, elle cherche à dépasser une classification uniquement corrélationnelle au moyen de catégories sémantiques adaptées au contexte chinois. Construites par l'analyse textométrique de corpus filtrés, notamment à l'aide des spécificités de Lafon, ces catégories décrivent des formes récurrentes d'expression dépressive et guident les expériences. L'enjeu sociétal tient aux normes discursives et culturelles, différentes de celles représentées dans les corpus cliniques ou anglophones. L'enjeu éthique concerne l'anonymisation, la minimisation des données, la traçabilité et la limitation des hallucinations, afin de réduire les risques d'interprétation diagnostique.

La méthode comprend quatre phases : l'établissement de modèles de référence explicables reliant catégories et décisions ; l'étude de modèles génératifs par ingénierie de prompts ; le développement et l'évaluation des TinyMentalLLMs, modèles compacts spécialisés ; et la construction de CNSocialDepress (CNSD), comprenant un jeu Gold annoté par des experts selon six dimensions et un pipeline reproductible produisant des données Silver pour la classification et l'analyse structurée.

La thèse propose ainsi une approche méthodologiquement fondée et éthiquement encadrée, ainsi qu'une ressource réutilisable, tout en maintenant une distinction claire entre analyse computationnelle et diagnostic clinique.

Mots-clés : Détection du risque dépressif ; réseaux sociaux ; textométrie ; modèles génératifs ; interprétabilité ; éthique de l'IA.

Jinyuan XU

Automatic Detection of Depressive Moods on Chinese Social Media

This thesis focuses on the automatic detection of depression risk from texts published on Chinese social media. Beyond binary classification, it aims to provide an interpretable, traceable, and ethically framed analysis in order to better understand how depression is expressed outside clinical settings. Two main challenges are addressed: the lack of specialised resources in Chinese and the opacity of generative approaches.

The research is structured around three main issues. From a scientific perspective, it seeks to move beyond purely correlational classification by developing semantic categories adapted to the Chinese context. Constructed through the textometric analysis of filtered corpora, notably using Lafon's specificity measure, these categories describe recurrent forms of depressive expression and guide the experiments. The societal issue concerns discursive and cultural norms that differ from those represented in clinical or English-language corpora. The ethical issue involves anonymisation, data minimisation, traceability, and the limitation of hallucinations in order to reduce the risks of diagnostic interpretation.

The methodology comprises four phases: the establishment of explainable baseline models linking categories to decisions; the study of generative models through prompt engineering; the development and evaluation of TinyMentalLLMs, compact specialised models; and the construction of CNSocialDepress (CNSD), which includes a Gold dataset annotated by experts according to six dimensions and a reproducible pipeline producing Silver data for classification and structured analysis.

The thesis thus proposes a methodologically grounded and ethically framed approach, together with a reusable resource, while maintaining a clear distinction between computational analysis and clinical diagnosis.

Keywords: Depression-risk detection; social media; textometry; generative models; interpretability; AI ethics.

Signature des directeurs de thèse :



Mathieu Valette, Pierre Magistry