

L'algorithme qui détecte la dépression dans vos messages

NLP, analyse comportementale digitale et données pour la santé mentale

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | Ingénieur en Travaux Statistiques
tpineldahoui@gmail.com | github.com/Baudelaire12

2 Mars 2026

Série LinkedIn STEM au Quotidien — Jour 6/40

Plan de la présentation

- 1 Mise en contexte
- 2 Concepts clés
- 3 Développement mathématique
- 4 Cas pratique
- 5 Conclusion
- 6 Références

La santé mentale : une crise mondiale et silencieuse

Chiffre clé

Selon l'OMS (2022), **1 personne sur 8** dans le monde souffre d'un trouble mental. Moins de **2%** du budget de santé mondial est alloué à la santé mentale. En Afrique, on compte **1 psychiatre pour 1,4 million d'habitants** (OMS, 2021).

- En moyenne, **11 ans** s'écoulent entre les premiers symptômes de dépression et la première consultation psychiatrique (Wang et al., 2007).
- Des études montrent que le langage écrit des personnes dépressives diffère statistiquement du langage des personnes en bonne santé (Rude et al., 2004).
- Twitter et Reddit sont devenus des mines de données pour la recherche sur le suicide et la dépression (De Choudhury et al., 2013).

Question centrale

Peut-on détecter des signaux précoces de détresse psychologique à partir de données numériques comportementales — sans intrusion dans la vie privée ?

Phenotypage digital : le comportement comme signal clinique

Définition

Le **phénotypage digital** désigne l'utilisation de données comportementales numériques passives (utilisation du smartphone, langage, mobilité) pour inférer des états mentaux ou des traits psychologiques.

Sources de données comportementales :

Source	Signal capté	Indicateur mental
Réseaux sociaux	Fréquence, contenu, horaires	Dépression, anxiété
Accéléromètre	Activité physique, sommeil	Épisode maniaque/dépressif
GPS	Mobilité réduite, isolement	Dépression sévère
Appels/SMS	Fréquence de communication	Isolement social
Micro (passif)	Prosodie vocale	Schizophrénie, dépression
Rythme de frappe	Variabilité, erreurs	Anxiété, burn-out

Analyse de sentiment et modèles de langage

Représentation TF-IDF (Term Frequency-Inverse Document Frequency) :

$$\text{TF-IDF}(t, d, D) = \text{TF}(t, d) \times \log \frac{|D|}{|\{d' \in D : t \in d'\}|}$$

Modèle BERT (Bidirectional Encoder Representations from Transformers) — représentation contextuelle :

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V$$

où Q , K , V sont les matrices de requêtes, clés et valeurs, d_k la dimension des clés.

LIWC (Linguistic Inquiry and Word Count) — dictionnaire psycholinguistique :

$$\text{Score}_{\text{catégorie}} = \frac{\text{Nombre de mots de la catégorie}}{\text{Nombre total de mots}} \times 100$$

Les personnes dépressives utilisent significativement plus de pronoms à la première personne du singulier (*je*, *moi*, *mon*) — OR = 1.35 (Rude et al., 2004).

Pipeline de détection de dépression — NLP

Pipeline complet d'un modèle de détection de détresse :

- ① **Collecte** : posts Reddit (r/depression, r/SuicideWatch) vs r/casual — (Yates et al., 2017)
- ② **Prétraitement** : tokenisation, lemmatisation, suppression des stopwords
- ③ **Vectorisation** : BERT fine-tuné sur données cliniques (Mental-BERT, Ji et al., 2021)
- ④ **Feature engineering** :
 - Scores LIWC : catégories affect négatif, mort, cognition
 - Fréquence de publication (rythme circadien perturbé)
 - Sentiments sur l'axe valence-arousal-dominance (Russell, 1980)
- ⑤ **Classification** : Random Forest / XGBoost / BERT fine-tuné
- ⑥ **Évaluation** : Sensibilité, Spécificité, AUC-ROC, F1-score

Résultat

Mental-BERT (Ji et al., 2021) atteint une AUC de **0.93** sur la détection de dépression à partir de textes Reddit.

Cas pratique : l'étude de De Choudhury et al. (2013)

Étude pionnière

De Choudhury, Gamon, Counts & Horvitz (2013) ont analysé les tweets de **476 utilisateurs** cliniquement diagnostiqués dépressifs vs un groupe contrôle.

Résultats clés :

Caractéristique Twitter	Dépressifs	Contrôle
Heure de publication (nuit)	Plus élevée	Moindre
Émotion négative (LIWC)	+34%	Référence
Engagement social	-27%	Référence
Références à soi ("je")	Plus fréquentes	Moindres
AUC du modèle prédictif	0.87	

Enjeu éthique majeur : Comment détecter sans stigmatiser ? Comment recueillir le consentement éclairé pour l'analyse de données publiques à des fins cliniques ?

Synthèse

Le NLP et le phénotypage digital ouvrent des voies prometteuses pour la détection précoce des troubles mentaux, particulièrement dans des régions où l'accès aux psychiatres est limité. Mais ces outils soulèvent des questions éthiques fondamentales sur la vie privée, le consentement et la stigmatisation.

Perspectives :

- Développement d'outils NLP en langues africaines (fon, haoussa, swahili) pour dépistage local
- Applications mobiles de suivi de l'humeur intégrant le NLP passif
- Cadres réglementaires clairs pour l'utilisation éthique des données comportementales
- Chatbots thérapeutiques validés cliniquement (Woebot — Fitzpatrick et al., 2017)

« Les données numériques ne remplaceront jamais l'empathie clinique — mais elles peuvent ouvrir la porte à ceux qui n'osent pas frapper. »

— Adapté de Munmun De Choudhury, 2013

Références bibliographiques

-  De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *ICWSM*, 7(1), 128–137.
-  Ji, S., et al. (2021). MentalBERT : Publicly available pretrained language models for mental healthcare. *INTERSPEECH 2022*.
-  Rude, S., Gortner, E.-M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition & Emotion*, 18(8), 1121–1133.
-  World Health Organization. (2022). *World Mental Health Report : Transforming Mental Health for All*. WHO Press.
-  Wang, P. S., et al. (2007). Delay and failure in treatment seeking after first onset of mental disorders in WHO World Mental Health Survey Initiative. *World Psychiatry*, 6(3), 177–185.
-  Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression via a conversational agent (Woebot). *JMIR Mental Health*, 4(2), e19.
-  Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT : Pre-training of deep bidirectional transformers. *NAACL-HLT 2019*, 4171–4186.

Merci pour votre attention

Pinel Baudelaire T. DAHOUI

Data Scientist & IA | Ingénieur en Travaux Statistiques

`tpineldahoui@gmail.com`

`github.com/Baudelaire12`

LinkedIn : @Pinel Baudelaire DAHOUI

Série STEM au Quotidien — 40 jours