



■ Équipe « Decision Aid and Information Discovery » (DECIDE)

¹ CHRU de Brest / IMT Atlantique / LATIM

² IMT Atlantique / Lab-STICC / DECIDE
<https://www.labsticc.fr>

³ CHRU de Brest / IMT Atlantique /
Lab-STICC / DECIDE

Sofian BERROUIGUET¹

Romain BILLOT²

Yannis HARALAMBOUS²

Christophe LEMEY³

Philippe LENCA²

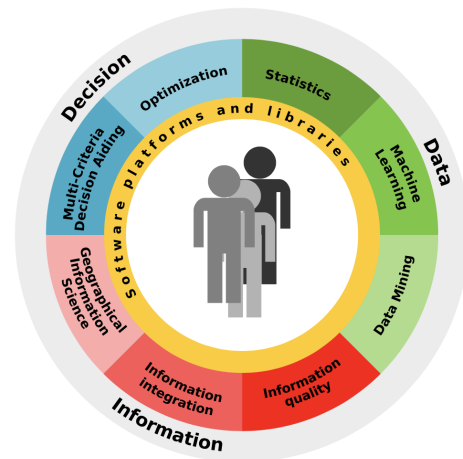
Sorin MOGA²

John PUENTES²

Introduction

L'équipe « Decision Aid and Information Discovery » (DECIDE) de l'UMR CNRS 6285 Lab-STICC mène des travaux algorithmiques et méthodologiques afin de faciliter la prise de décision dans des contextes impliquant un ou plusieurs décideurs confrontés à des données hétérogènes et complexes. Cette aide à la décision est abordée à travers le triptyque données – information – décision. Ainsi, les recherches menées puisent leurs sources dans les disciplines de la statistique, de l'apprentissage automatique et de la fouille de données (pour le point de vue données); des sciences de l'information géographique, de l'intégration et fusion d'informations et de la qualité de l'information (pour le point de vue information); et de l'optimisation et de l'aide multi-critère à la décision (pour le point de vue décision). Cette aide à la décision est soutenue à travers différentes plate-formes et logiciels, auxquels l'équipe contribue fortement, et où l'humain garde une place centrale.

Les travaux de l'équipe sont appliqués à des domaines variés, dont celui de la santé. Nous présentons brièvement ci-dessous nos principales contributions, principalement en ce qui concerne l'axe « données », dans le cadre de la santé mentale et de la détection de la maladie d'Alzheimer.



Triptyque données – information – décision

Données et santé mentale

En santé mentale, nous avons développé, avec les CHRU de Brest et Montpellier et l'*Instituto de Investigación Sanitaria de la Fundación Jiménez Díaz* de Madrid, des travaux sur la prévention du risque suicidaire et sur la détection précoce de la schizophrénie.

Les principaux résultats d'un processus complet de fouille de données sociologiques et cliniques sur un échantillon de suicidants de deux hôpitaux européens sont présentés dans [2]. Nous avons identifié trois groupes de patients similaires, à partir d'une analyse de correspondances multiples et d'une classification ascendante hiérarchique. Ensuite, les



AfIA

Association française
pour l'Intelligence Artificielle

facteurs de risque, par groupe et associés au nombre de tentatives sont, eux, identifiés à l'aide d'arbres de régression. Les résultats mettent en lumière l'apport de la fouille de données aux activités de description et d'explication ; ils mettent également en exergue la nécessité d'améliorer la qualité des données. Enfin, ces travaux seront enrichis par des évaluations écologiques [1], par le biais de données issues du milieu de vie habituel des patients, comme la qualité du sommeil [3], ou encore à travers de données liées à la représentation sociale du suicide [5].

Enfin, nous nous intéressons à la détection et à l'intervention précoce dans le cadre des troubles psychotiques débutants. À partir d'un (petit) corpus d'entretiens de bilan psychiatrique, en fusionnant des données provenant d'un analyseur syntaxique et des données chronométrées, et en y étudiant des phénomènes reliant la syntaxe du langage naturel, la prosodie et les pauses (vides ou pleines), nous avons obtenu des indicateurs de comorbidités psychiatriques permettant d'assister les psychiatres dans leurs diagnostics [4]. Pour mieux identifier les indicateurs de comorbidités dans les entretiens, nous avons créé un corpus témoin, suivant le même protocole que les entretiens psychiatriques mais avec des sujets sains. Dans le cadre d'une thèse, nous allons corrélérer les indicateurs obtenus avec des symptômes cliniques observés au cours des deux années de suivi.

Ces dernières années il y a eu, mondialement, pléthore d'articles sur l'application de l'apprentissage artificiel et du traitement automatique de la langue à la santé mentale. Ces articles diffèrent tant sur le plan des méthodes et des plate-formes que sur le plan du vocabulaire et du public visé. Pour donner une vision globale de ces efforts, nous avons étudié 327 articles et établi une synthèse de 58 d'entre eux dans [6].

Données et détection de la maladie d'Alzheimer

La maladie d'Alzheimer (MA) est un type de démence affectant des millions de personnes dans le monde. Parmi les moyens de diagnostique, l'imagerie par résonance magnétique structurale (sMRI en anglais) du cerveau peut révéler des signes de neurodégénérescence. Or, l'interprétation de ces signes tant par des experts humains que par des algorithmes est difficile, principalement à cause de la complexité de la maladie, de la variabilité des images et des différences démographiques. En outre, la fiabilité de l'analyse d'images assistée par ordinateur est directement associée à la manière de représenter l'information extraite des images.

L'équipe DECIDE, en collaboration avec le *Departamento d'Ingeniería de Sistemas y Computación* de l'*Universidad de los Andes* de Bogota, explore les approches de représentation guidées par les données des images sMRI du cerveau, focalisées sur l'analyse de la neurodégénérescence et son lien potentiel avec la MA. Des résultats encourageants suggèrent la possibilité d'identifier des images 2D représentatives d'un volume 3D, comme étant caractéristiques de l'état de la neurodégénérescence d'un sujet [8]. Deux approches de classification ont obtenu des résultats performants, en appliquant une quantification par dictionnaire non supervisé pour représenter le volume total du cerveau [7] et en définissant un modèle d'auto-encodeur supervisé qui utilise le contenu d'une seule image 2D du volume 3D du cerveau par individu [9]. Finalement, l'analyse régionale basée sur la réponse locale des autoencodeurs supervisés révèle des régions anatomiques pertinentes pour l'identification de la neurodégénérescence éventuellement associée à la MA, permettant de différencier les images pathologiques des saines [9].



Références

- [1] Sofian Berrouiguet, Romain Billot, Mark Larsen, Jorge Lopez-Castroman, Isabelle Jaussent, Michel Walter, Philippe Lenca, Enrique Baca-Garcia, and Philippe Courtet. An approach for data mining of electronic health record data for suicide risk management: Database analysis for clinical decision support. *JMIR Mental Health*, 6(5):e9766, May 2019.
- [2] Romain Billot, Sofian Berrouiguet, Mark Larsen, Michel Walter, Jorge López Castroman, Enrique Baca-Garcia, Philippe Courtet, et Philippe Lenca. Apport de la fouille de données pour la prévention du risque suicidaire. In *Conférence Internationale sur l'Extraction et la Gestion des Connaissances, Revue des Nouvelles Technologies de l'Information*, volume RNTI-E-34, pages 143–154, Paris, France, janvier 2018.
- [3] Elise Guillodo, Christophe Lemey, Mathieu Simonnet, Michel Walter, Enrique Baca-Garcia, Vincent Masetti, Sorin Moga, Mark Larsen, Hugopsy Network, Juliette Ropars, and Sofian Berrouiguet. Clinical applications of mobile health wearable-based sleep monitoring: Systematic review. *JMIR mHealth and uHealth*, 8(4), April 2020.
- [4] Yannis Haralambous, Christophe Lemey, Philippe Lenca, Romain Billot, and Deok-Hee Kim-Dufor. Using dependency syntax-based methods for automatic detection of psychiatric comorbidities. In Dimitrios Kokkinakis, Kristina Lundholm Fors, Charalambos Themistocleous, Malin Antonsson, and Marie Eckerström, editors, *Resources and processing of linguistic, para-linguistic and extra-linguistic data from people with various forms of cognitive/psychiatric/developmental impairments*, pages 142–150, Marseille, France, May 2020.
- [5] Louis Jehel, Dupont Emmanuel, Philippe Lenca, and Céline Kopp-Bigault. The ReSCU-M study: Comprehension of suicide social representation in Martinique island. Perspective for prevention and communication strategies. In *International Association for Suicide Prevention Caribbean Regional Symposium*, Trinidad, Trinidad and Tobago, May 2019.
- [6] Aziliz Le Glaz, Yannis Haralambous, Deok-Hee Kim-Dufor, Philippe Lenca, Romain Billot, Ryan Taylor, Jonathan Marsh, Jordan DeVlyder, Michel Walter, Sofian Berrouiguet, and Christophe Lemey. Machine learning and natural language processing in mental health: a systematic review. *Journal of Medical Internet Research*, 8(4), forthcoming/in press.
- [7] Ricardo Mendoza-León, Fabio A. González, Pablo Arbeláez, John Puentes, and Marcela Hernández Hoyos. Analysis of PHOW representations for Alzheimer disease classification on brain structural MRI. In *IEEE 13th International Symposium on Biomedical Imaging*, pages 24–27. IEEE, 2016.
- [8] Ricardo Mendoza-Léon, John Puentes, Fabio A. González, and Marcela Hernández Hoyos. Empirical evaluation of general-purpose image features for pathology-oriented image retrieval of Alzheimer disease cases. In *30th International Congress on Computer Assisted Radiology and Surgery, International Journal of Computer Assisted Radiology and Surgery*, volume 11, pages S39–S40. Springer, 2016.
- [9] Ricardo Mendoza-Léon, John Puentes, Luis Felipe Uriza, and Marcela Hernández Hoyos. Single-slice Alzheimer's disease classification and disease regional analysis with supervised switching autoencoders. *Computers in biology and medicine*, 116:103527, 2020.