

INTERNSHIP

Understanding users profiles by analyzing massive non structured conversations on a drug users web community.

Advisors

Raquel Urena (Associate professor in AI applied to health, SESSTIM, AMU), Benoit Favre (Full Professor in NLP, LIS, AMU)

Context

Today, digital technologies constitute promising solutions to meet the challenges facing the French healthcare system (Moquet-Anger, 2020). In the field of addictions systems for managing and analyzing massive amounts of data are already playing a major role in redefining clinical and public health research practices (Thierry et Reynaud, 2019). Nowadays, advances in AI such as social network analysis and natural language, offer novel opportunities for better understanding patterns of drug use and their associated consequences. By analyzing the massive data existing thanks to naturally occurring, self-disclosed communication on social media, it is possible to understand temporal trends of problematic drug use, market changes, social norms and cultural aspects of drug use and relevant psychological factors (Kim et al., 2017).

Objectives

In spite of storing information from almost a million of users, psychoactif.org lacks from users' structured data. Therefore, the objective of the internship is to obtain user's profiles from the textual information published and read by the user as well as its interactions in the virtual community. Then, QuanTIM researchers will carry out. Semi-structured interviews with psychoactif.org users will be carried out by SanteRCom and psychoactif.org researchers to guide the data analysis and validate the results obtained with the AI based methods.

Scientific program

This research proposal will be using the website psychoactif.org that provides drug users with an anonymous platform to exchange knowledge on drug use, practices, harm reduction strategies and prevention. Psychoactif.org is the largest French speaking web community of drug users with approximately 1 million users, 3 million page views per month and between 200 and 400 new testimonials messages created per day.

AI based analyses including, natural language processing and topic modeling to obtain a structured user profile with information about drug use, behaviors, interests and prevention measures will be carried out. textual information will be annotated and prepared by public health researcher specialists in addictions.

This is a multidisciplinary internship in which the intern will have the opportunity to deal with a real research project in which public health researchers, psychologists and data scientists are involved.

Additional information

Languages: Very good level in French required, to be able to understand the posts information.

Skills: Master-level computer science, an interest for linguistics, python programming, deep learning, Pytorch, rigor and tenacity.

Location: the internship will take place between the SESSTIM on the Faculty of Medicine and the LIS at Luminy campus of Aix-Marseille University.

Dates: Spring-summer 2022, duration 5-6 months.

Wages: regulatory internship salary (about 600 euros/month).

Contact details: To apply please send a mail with your CV and a motivation letter to raquel.urena@univ-amu.fr and benoit.favre@lis-lab.fr.

Université Aix-Marseille – Faculté de Médecine – 27, boulevard Jean Moulin 13385 Marseille Cedex 5 –
France Tél : (+33) 04 91 32 46 00 / Courriel : sesstim-u1252@inserm.fr / Site : <http://sesstim.univ-amu.fr>