



Extraction of medical terms from non-structured (textual) data from online news and social media – *Workshop 5 - SVEPM 2021*

Part 3: Classification and animal disease surveillance systems

KEY-WORDS: Text Mining, Information Retrieval, Health Surveillance

ABSTRACT

In the context of this workshop, 3 talks will be done. First, we will present an overview of NLP (Natural Language Processing) approaches in order to mine textual data [PART 1]. The second part will present two NLP tasks dedicated to Health domain: (i) acquisition of tweets, (ii) terminology extraction in (social) media [PART 2]. The last part of this workshop is dedicated to textual classification issues based on machine learning techniques with Weka [PART 3]. A part of this presentation will use a public dataset: <https://dataverse.cirad.fr/dataset.xhtml?persistentId=doi:10.18167/DVN1/POIZMA>

Contact:

mathieu.roche@cirad.fr

CIRAD, TETIS, F-34398 Montpellier, France

TETIS, Univ. Montpellier, AgroParisTech, CIRAD, CNRS, INRAE, Montpellier, France