

Federated multi-label classification applied to International Classification of Diseases

IXA Group

Maite Oronoz, Alicia Pérez, Arantza Casillas, Olatz Pérez de Viñaspre

Contact: maite.oronoz@ehu.eus

1 Description

International Classification of Diseases is an international means of encoding diagnoses. This is a relevant way of assessing statistics about the spread of diseases world-wide and get alarms about morbidity. These codes are also used from insurance companies for billing.

Within the context of clinical text mining a real issue rests on data-sharing due to confidentiality regulations. While **Deep Neural Networks** are strong, big data are required, however, getting health records from different hospitals is barely feasible. To cope with this issue, we resort to **federated** learning. This enables to infer a model from each hospital and combine the models' parameters without the need of sharing data.

The aim of this project is to assess the performance of individual models learned from different cohorts or hospitals and compare to a federated model. Does the combination gain performance?

2 Goals

The student will apply different techniques and relatedness metrics in order to build a prototype able to identify relations not explicitly expressed in documents. The key objectives are:

- Cope with extreme multi-label classification.



- Evaluate and compare the base-models with respect to a federated approach.

3 Requirements

- Good programming skills.
- The master dissertation can be written in English, Spanish or Basque.

4 Framework

Python

5 Learning outcomes

The student will acquire background in information extraction reinforcing the following areas:

- deep learning applied to information extraction
- multi-label classification
- federated approaches
- submit the results to a research paper

6 Tasks and plan

- 3 weeks: study literature, test approaches
- 9 weeks: experimentation and evaluation
- 4 weeks: write and submit a research article
- 4 weeks: write down the thesis