Introduction

This work presents a flexible system for metric learning in document analysis.
- We use a Convolutional Neural Network
- Trained with a differentiable ranking loss nDCG
- The resulting embedding space distribution keeps semantic meaning.

Learning Objectives

We used the nDCG metric from information retrieval as loss function. This function is not differentiable, so we propose a smoothed version [1, 2, 3].

Datasets

We evaluated this task in two distinct datasets.

- MPS Manuscript Dataset
- XAC Newspapers Dataset

Human in the loop

A simple example of how the system could incorporate a human-in-the-loop:
- The user could give feedback of the years the model has to focus.
- The user could incorporate new data easily to improve the performance.

Embeddings Glow-Up

The resulting embeddings are more representative of its own category than the ones we get with common training loss functions.

Adaptability

Moreover, we can “specialize” the network in a certain range of categories just by increasing the values on the relevance matrix.

Examples

Example of retrieval for MPS Dataset. The model succeeds on retrieving documents from close years to the given query.

Note that the base model (left) is equally focused on all the years; we can see at first sight how the focused model (right) has made more clear differences in the early years (yellow).

Additionally, we evaluated how using the retrieval approach can lead to similar results in terms of regression. Despite we didn’t get to archive such results in MPS Regression task, it should be solved by adapting our input to the SOTA format.

Acknowledgements: This work has been partially supported by the Spanish projects RTI2018-095645-B-C21, and FCT-19-15244, and the Catalan projects 2017-SGR-1783, the Culture Department of the Generalitat de Catalunya, and the CERCA Program / Generalitat de Catalunya.

References