General Method For Document Date Estimation

Adrià Molina Lluis Gomez Oriol Ramos Terrades Josep Lladós

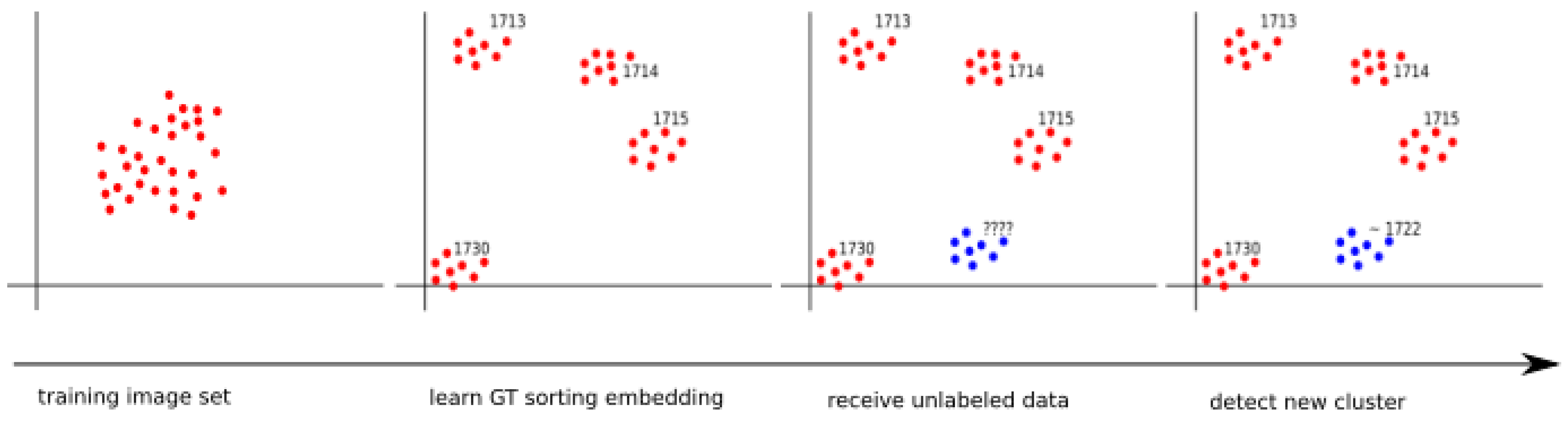
Computer Vision Center and Computer Science Department,
Universitat Autònoma de Barcelona, Catalunya

DAS 2022
May 22-25, La Rochelle

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].

DCG

Measures the retrieval performance for a graded **relevance scale**.

$$\mathrm{DCG}_q = \sum_{n=1}^{|\Omega_q|} \frac{r(n)}{\log_2(n+1)},$$

smooth-DCG

Is differentiable by using an smooth **indicator function**.

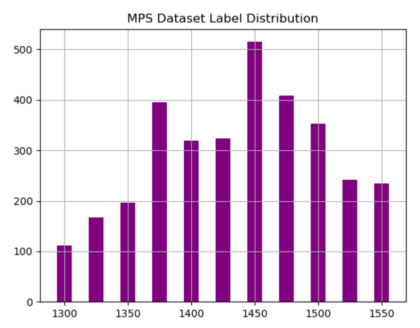
$$\mathrm{DCG}_q \approx \sum_{i\in\Omega_q} \frac{r(i)}{\log_2\left(2+\sum_{j\in\Omega_q, j\neq i} \mathcal{G}(D_{ij};\tau)\right)}$$

Datasets

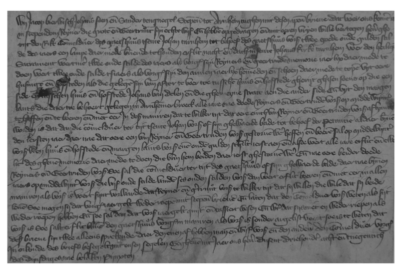
We evaluated this task in two distinct datasets.

MPS Manuscript Dataset

Histogram Distribution



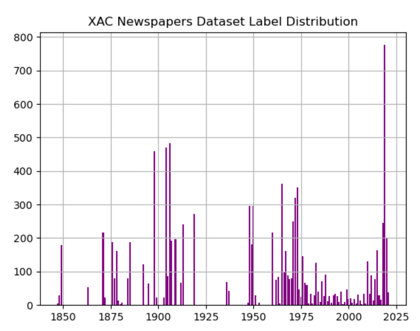
Sample Image



XAC Newspapers Dataset

Sample Image

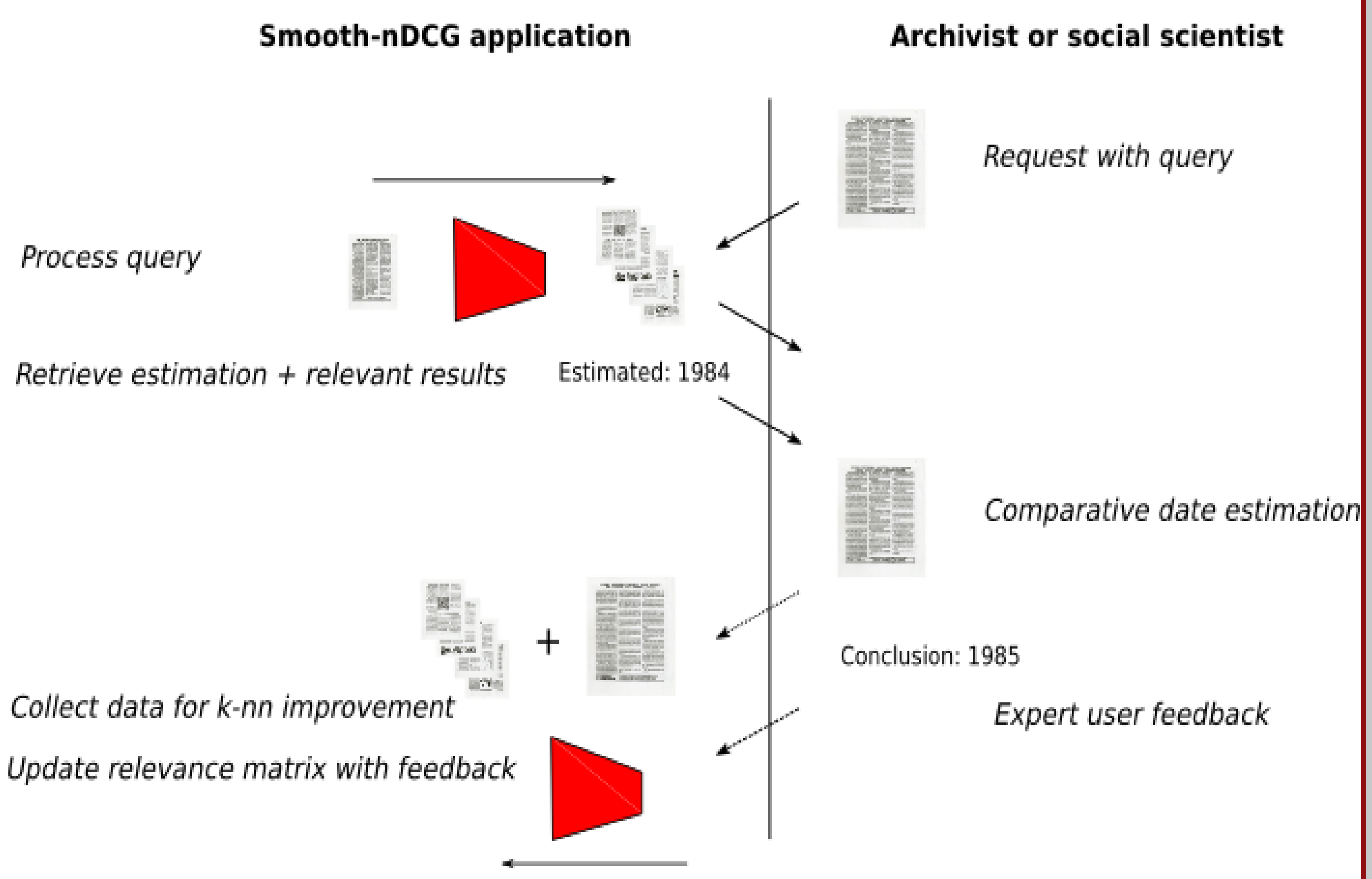
Histogram Distribution



Human in the loop

A simple example of how the system could incorporate a human-in-the-loop:

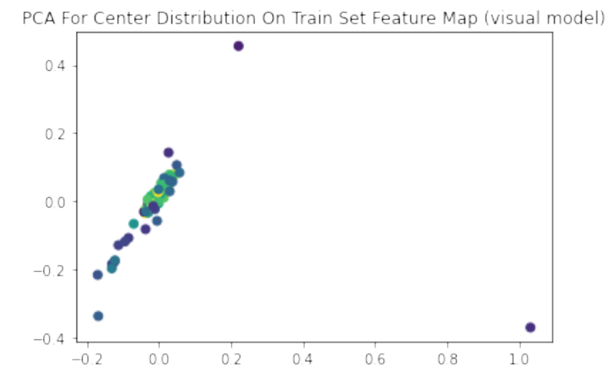
- The user could give feeback of the years the model has to focus.
- The user could incorporate new data easily to improve the performance.



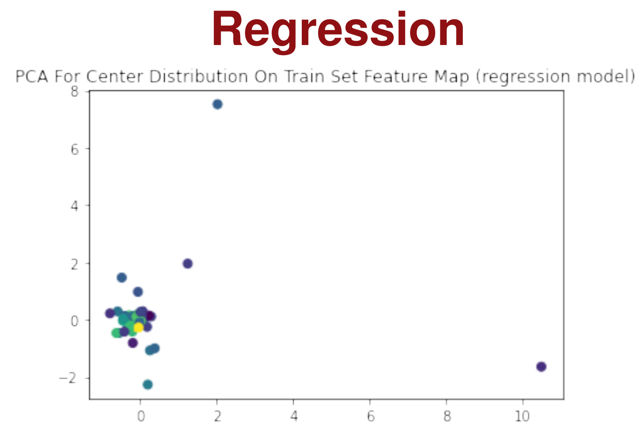
Embeddings Glow-Up

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

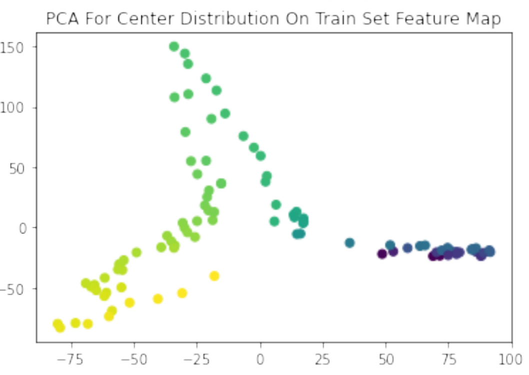
Classification



Regression

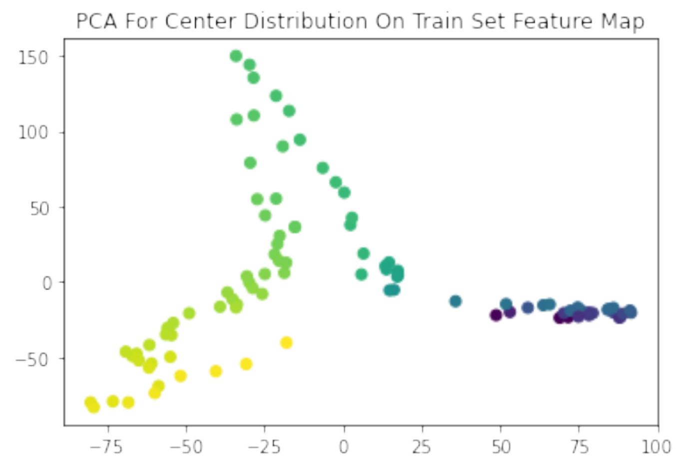


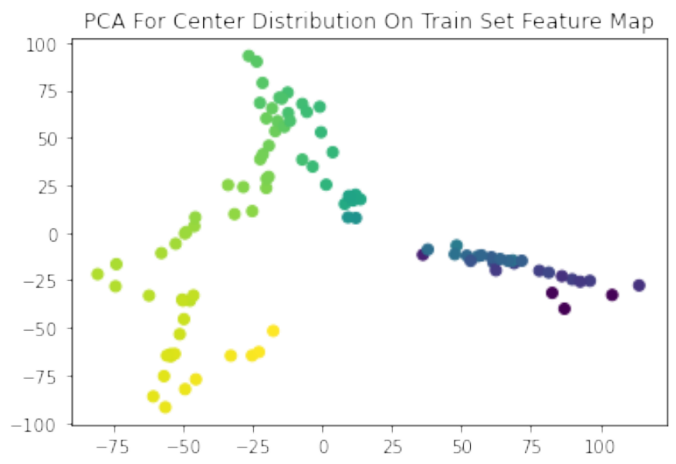
nDCG



Adaptability

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

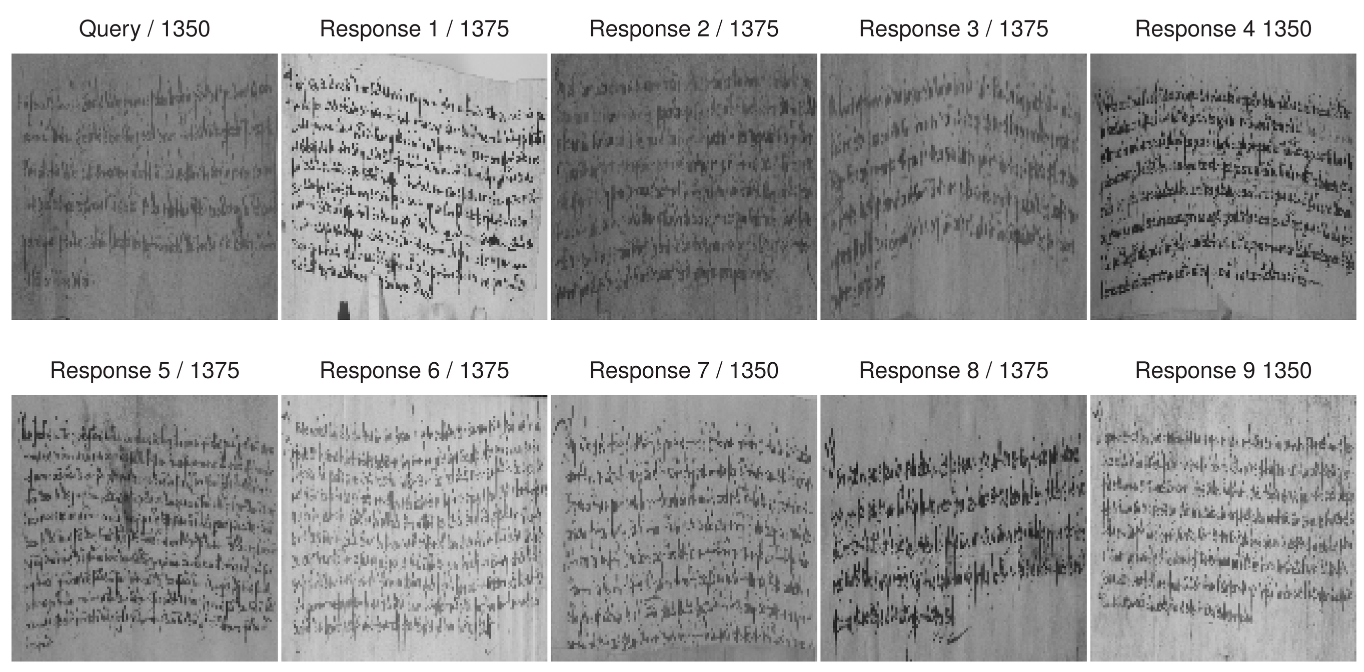




Note that the base model (left) is equally focused on all the years; we can see at first sight how the focused model (right) has made more clear differences in the early years (yellow).

Examples

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



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.

Baseline	MAE	mAP
Regression Baseline (Inception v3)	3.5	0.24
Smooth-nDCG Newspaper Date Estimation (Inception v3)	2.9	0.49
InceptionResnetV2 (MPS)	3.01	-
Smooth-nDCG Manuscript Retrieval (MPS)	23.8	0.43

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

[1] A. Brown, W. Xie, V. Kalogeiton, and A. Zisserman, "Smooth-ap: Smoothing the path towards large-scale image retrieval," in *Proceedings of the European Conference on Computer Vision*, pp. 677–694, 2020.

[2] A. Molina, P. Riba, L. Gomez, O. Ramos-Terrades, and J. Lladós, "Date estimation in the wild of scanned historical photos: An image retrieval approach," in *International Conference on Document Analysis and Recognition*, pp. 306–320, Springer, 2021.

[3] P. Riba, A. Molina, L. Gomez, O. Ramos-Terrades, and J. Lladós, "Learning to rank words: Optimizing ranking metrics for word spotting," in *International Conference on Document Analysis and Recognition*, pp. 381–395, Springer, 2021.