



Using diplomacy for... Machine Learning?

Applying AI to archival problems

Corinne Rogers

TAATU

June 14, 2022

The problem

- Automatic analysis of digitized ancient parchments has become an important field of research for
 - archival description (archival domain)
 - image and pattern recognition (AI domain)
- Example: the State Archives of Milan holds <130,000 parchments covering 1,000 years of history
 - how can these parchments be analyzed and made available?



AI & Machine Learning

- In the simplest terms, machine learning is a branch of AI that uses data and algorithms to imitate human learning for analytical and predictive purposes
- Machine Learning (ML) > Deep Learning (DL) > Neural Networks
- Applications include image recognition, natural language processing, speech recognition
- **Patterns!**



Diplomatics

- At the core of diplomatics lies the idea that all records can be analyzed, understood, and evaluated in terms of a system of formal elements, conditioned by time and place
- Diplomatics conceives documents as a system of physical (external) and intellectual (internal) elements of form
- These formal elements can be examined separately from content
- **Patterns!**



Diplomatics + DL = AI in Archives

- DL is being applied to digitized notarial parchments from the Middle Ages in Milan*
- The diplomatic element of form being used is the *signum* or Notarial sign - a unique drawn mark identifying each notary



* This InterPARES Trust AI study is led by Benedetto Luigi Compagnoni - Director, Archivio di stato di Milano, and Emanuele Frontoni, Professor, Computer Science and Computer Vision, University of Macerata



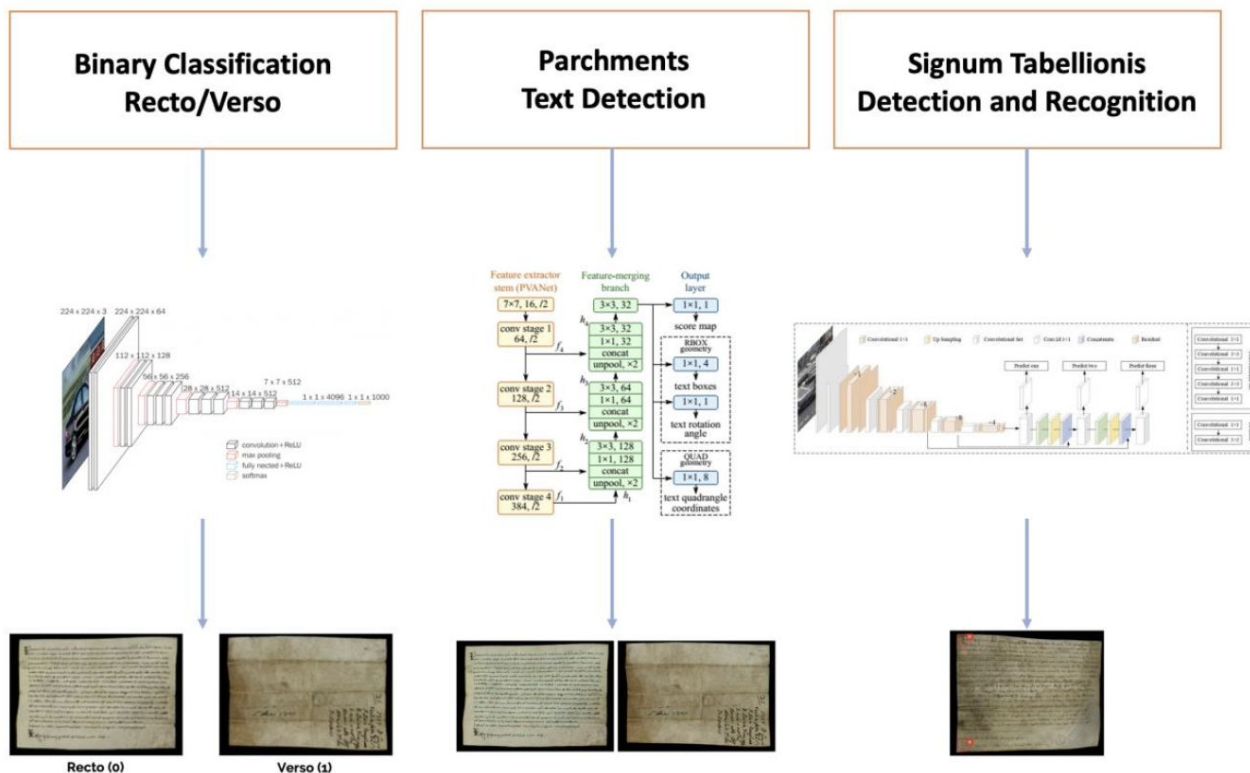
Diplomatics + DL = AI in Archives

- The aim of this project is to use AI to identify and build a complete list of Milanese notaries from the 12th and 13th centuries, and generate a list of all the surviving documents related to each notary



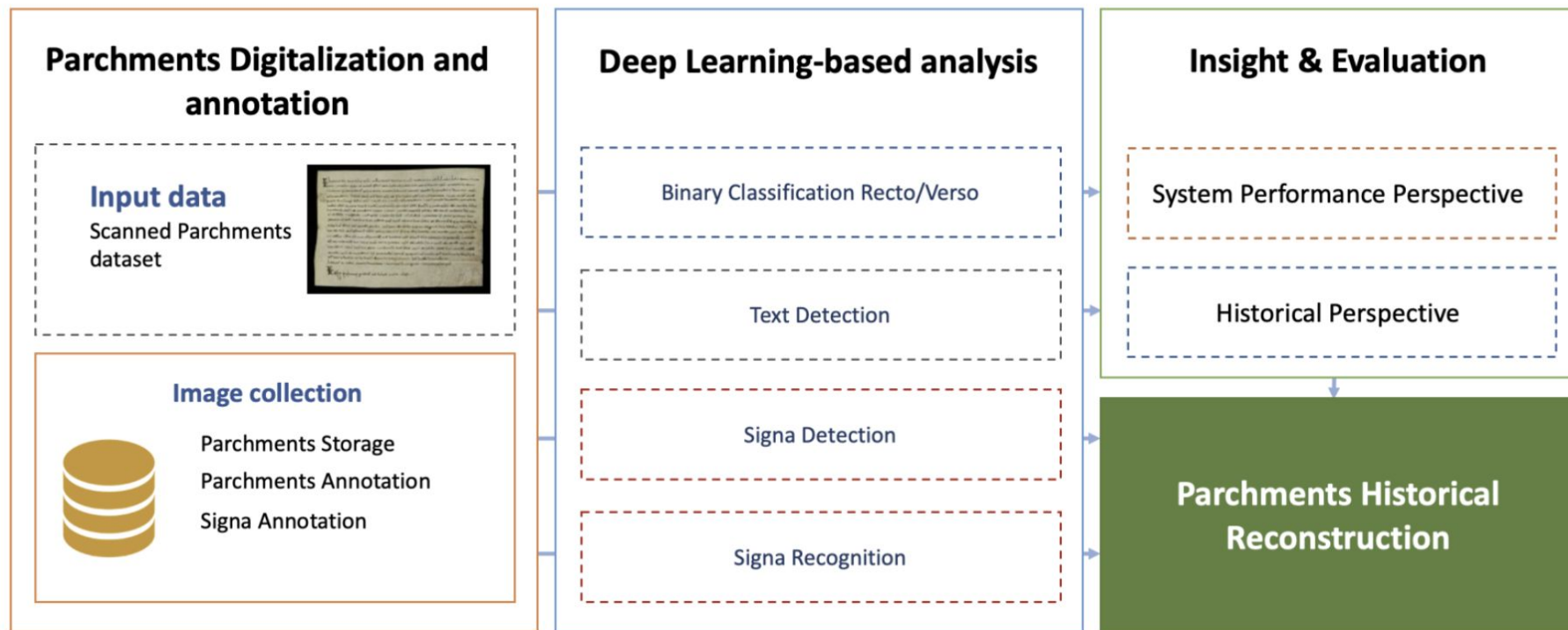
Diplomatics + DL = AI in Archives

Deep Learning Pipeline



Diplomatics + DL = AI in Archives

PergaNet



Diplomatics + DL = AI in Archives

- There are many other applications possible:
- Recognize the specific system of writing of each notary
- Recognize annotations
- Retrace previous archival arrangements
- Recognize recurring images and other formal elements
- Identify common patterns in manuscript maps or drawings
- Many others...



Where to find out more

- <https://interparestrustai.org>
- a presentation that describes more about the Milan project (and other projects in ITrust AI)
https://www.youtube.com/watch?v=Hup_d0uJyV8
- a great intro to deep learning from a non-technical perspective
<https://mumageed.medium.com/how-deep-learning-actually-learns-about-language-b3f95d752acd>

Thank you!

InterPARES
TrustAI

