UNIVERSITÀ DEGLI STUDI DI MILANO

MSC IN DATA SCIENCE AND ECONOMICS

Mining arguments in ECB speeches

Author: Luca Cerabone (matr. 960070) - **Supervisor:** Prof. Alfio Ferrara

Contents

The problem - **Argumentation**

Computer replace man power in jobs

With the current trend today of the ever-changing technology, it cannot be denied that several jobs are replaceable. On the other hand, **there are still jobs which require the expertise of a human employee** especially when *it entails decision making situations*. **One example of a job that got replaced by the availability of technology or computer to the public is the reservation and ticketing officers of travel agencies and airlines.** *You can now see the availability of your intended travel dates online, make the reservation, know the price of the ticket you want to purchase and pay conveniently with your credit card and you're all done. You even have the choice of printing your own ticket or just pick it up at the airport at the day of your departure.* But of course, *there would still be several customers that prefer to talk with an agent in completing their purchase of a ticket.* They may have several reasons to it but then maybe *they are just comfortable dealing with a human being than a computer.* **Jobs that the most computer can do is assist, would be that of a medical doctor.** It's true that *technology and computers do make their jobs easier* but *it cannot definitely replace them. It can have less work for them*, just for instance, *instead of opening up a patient to get a gallstone, they would just use an instrument that would make a point incision and guide it through inside the body to do the job for them.* With this, i still believe that ***technology and computers cannot eventually replace the workforce and drive employees to be jobless.*** *We still need the personal touch of a human in each and everything that technology or computer could offer us.*

Literature review

2009

Palau, R. M. and Moens, M. F. *Argumentation mining: the detection, classification and structure of arguments in text.*

2010

Palau, R. M. and Moens, M.F. *Argumentation mining. Artificial Intelligence and Law.*

2016

Lippi, M. and Torroni, P. *Argumentation mining: State of the art and emerging trends.*

2018

Devlin, J., Chang, M.W., Lee, K., and Toutanova, K. *Bert: Pretraining of deep bidirectional transformers for language understanding.*

2019

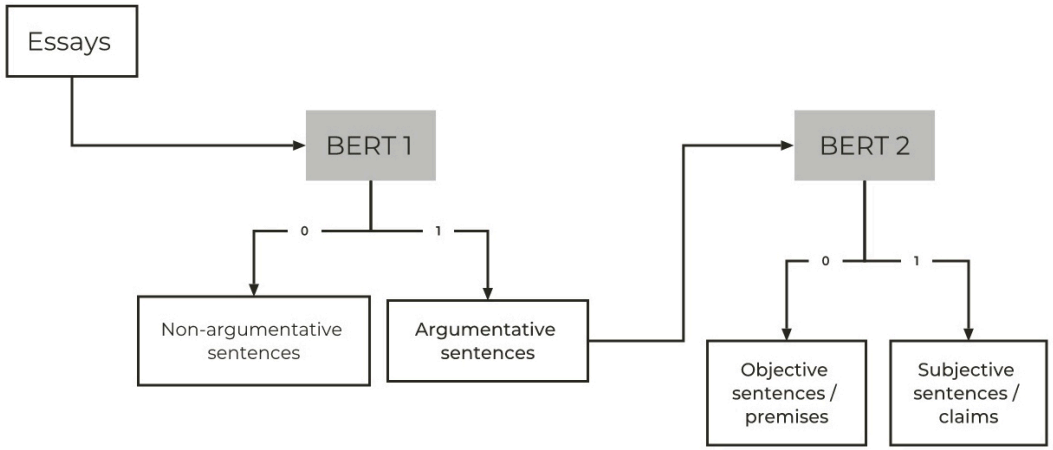
Reimers, N., Schiller, B., Beck, T., Daxenberger, J., Stab, C., and Gurevych, I. *Classification and clustering of arguments with contextualized word embeddings.*

The approach

Argument = a span of text expressing evidence or reasoning supporting or attacking a given **topic**.

Claim = the central component of an argument, it could be true or false but it's controversial;

Premise = reason justifying the claim or proving it wrong.



Argumentative sentences detection

- Split a document in sentences
- Classify each sentence as argumentative or not for the topic discussed in the document

Objective sentence detection

- Take argumentative sentences from step 1
- Classify them as objective or subjective

Predictions combination

- The **claim** expresses the views of the speaker/writer, we consider it as **subjective**
- The **premises** are facts or objective reasons that allow the speaker to justify his position, we consider them as **objective**

BERT

Bidirectional Encoder Representations from Transformers

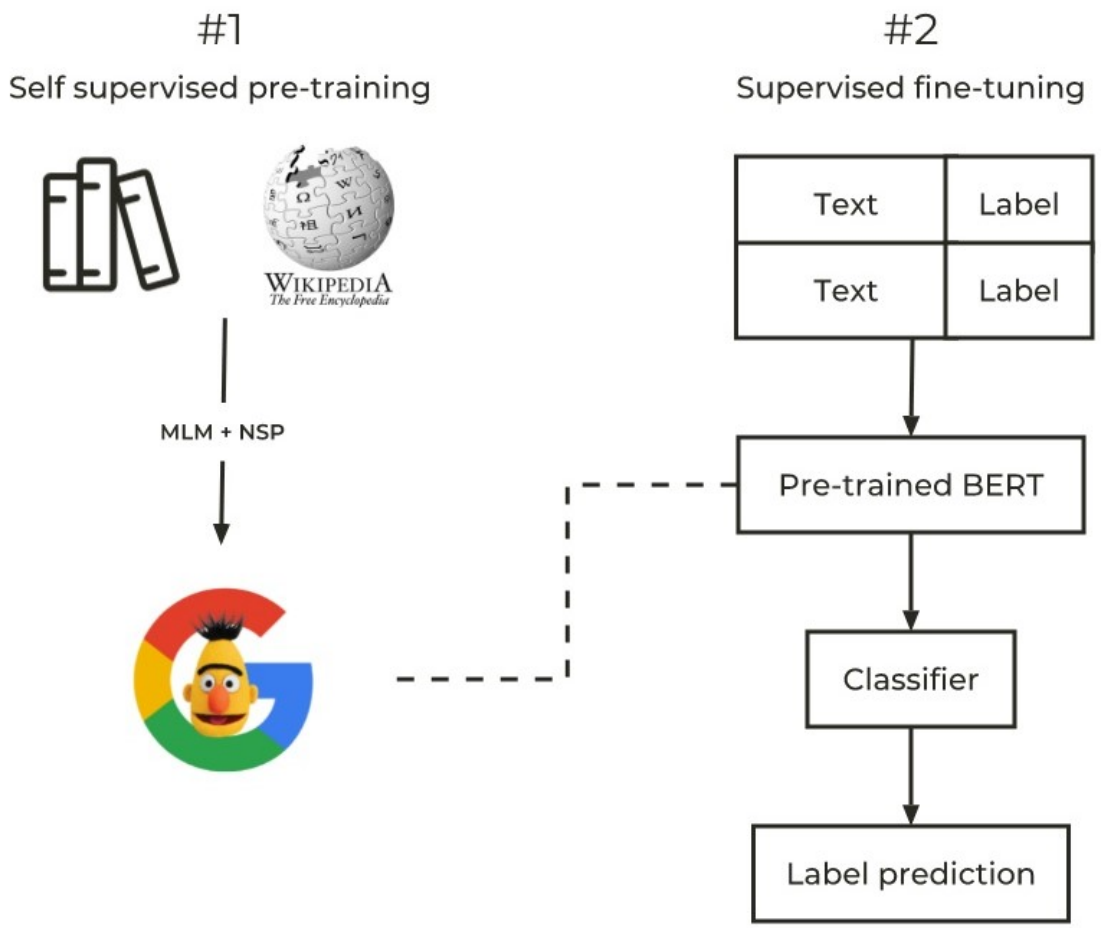
BERT [Devlin et al., 2018] encodes semantics and syntactic information about the language. It uses a deep Transformer network [Vaswani et al., 2017] with 12 or 24 layers to derive word representations. It was pre-trained for masked language modeling and next sentence prediction on two datasets: the BooksCorpus (800M words) and English Wikipedia (2,500M words).

Contextualized word embeddings

For each word in a sentence, BERT generates a vector representation that is based on the other words in that sentence.

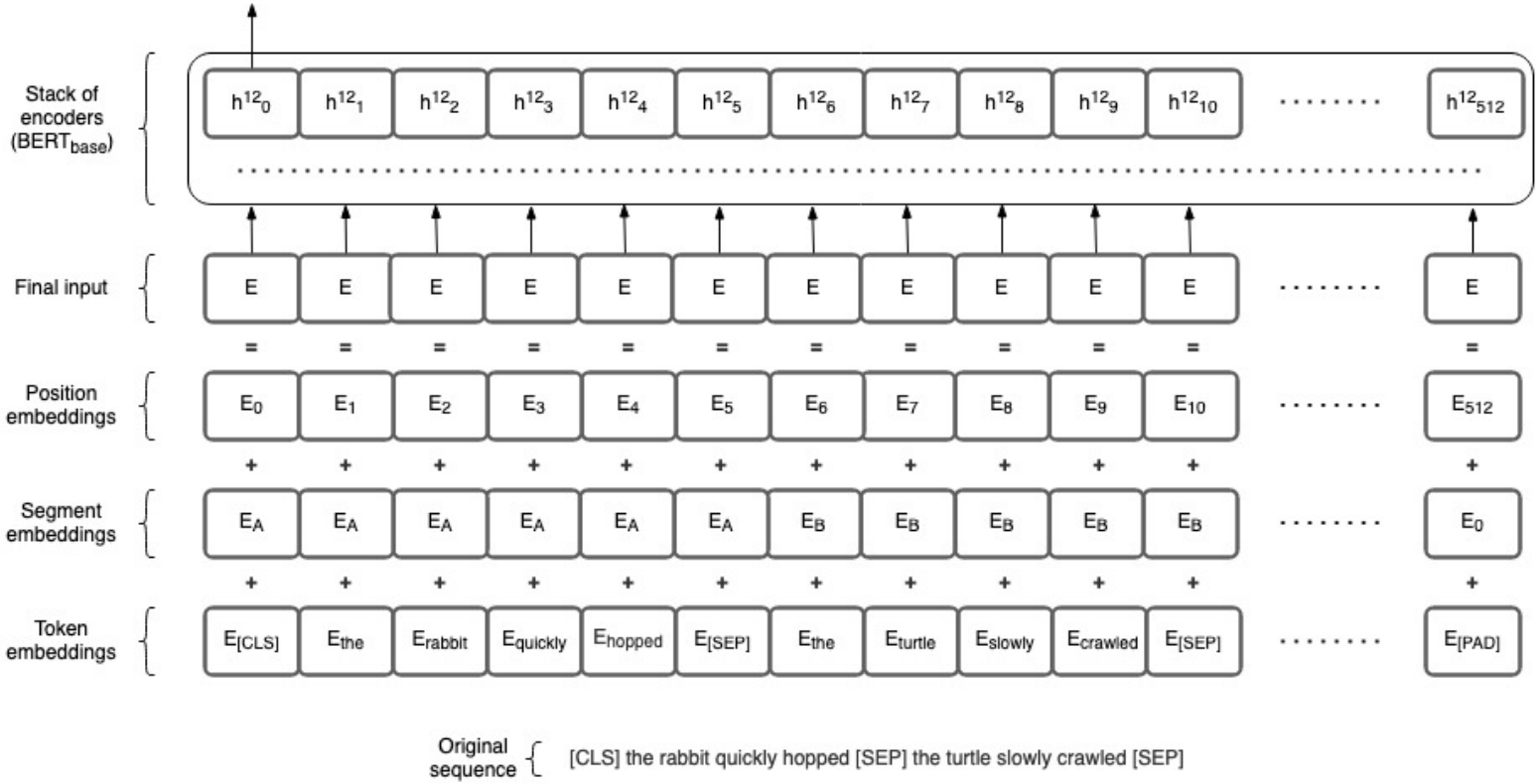
Transfer learning

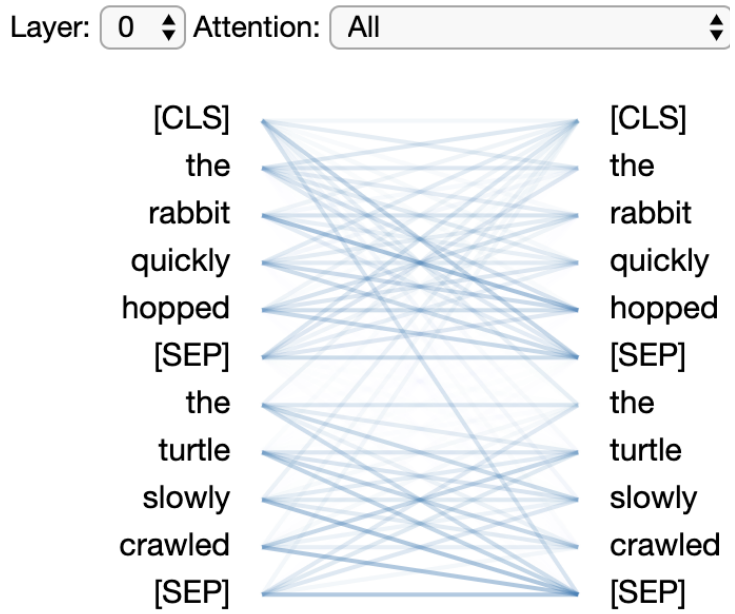
The model is pre-trained on some source task and fine-tuned on some target task. Its previous knowledge translates in a more efficient final model.



The attention mechanism

The input vector for BERT is given by the sum of token, segment and position embeddings.
Through attention a language model assigns weights to input features based on their importance for the final task.
BERT uses attention to form connections between words, allowing it to learn a variety of rich lexical relationships.





BERT 1 - training

UKP Sentential Argument Mining Corpus (25,492 sentences, 8 topics, 3 labels) [Stab et al., 2018].

The BERT output is passed to a dense layer with 512 neurons, a dropout layer is applied to mitigate overfitting and a final dense layer with one neuron classifies the sentences as non argumentative or argumentative.

Topic	Sentence	Label
Minimum wage	We should abolish all Federal wage standards and allow states and localities to set their own minimums.	No argument
Nuclear energy	It has been determined that the amount of greenhouse gases have decreased by almost half because of the prevalence in the utilization of nuclear power.	Argument
Abortion	Also, in that decision, he accepted scientific arguments that many women do come to regret having had abortions, and thus the procedure may cause psychological harm.	Argument

BERT 1 - results

Averaged scores obtained by training the model on 7 topics and testing it on the 8th for all the 8 topics

	Precision	Recall	F1 Score
bilstm [Stab et al., 2018]	0.73	0.38	0.61
biclstm [Stab et al., 2018]	0.62	0.70	0.64
mtl+biclstm+dip2016 [Stab et al., 2018]	0.65	0.67	0.67
BERT 1	**0.78**	**0.76**	**0.79**

BERT 2 - training

Training data: SUBJ dataset (10,000 sentences, 2 labels) [Pang and Lee, 2004].

The BERT output is passed to a dense layer with 256 neurons, a dropout layer is applied to mitigate overfitting and a final dense layer with one neuron classifies the sentences as objective or subjective.

Sentence	Label
The movie begins in the past where a young boy named Sam attempts to save Celebi from a hunter.	Objective
Watching the film is like reading a times portrait of grief that keeps shifting focus to the journalist who wrote it.	Subjective

BERT 2 - results

Score obtained by training on 80% of data and testing on 20%

	Accuracy
CNN [Mu et al., 2017]	0.8733
Vanilla RNN [Mu et al., 2017]	0.8760
GRU-RNN [Mu et al., 2017]	0.9185
CNN-LSTM [Khodak et al., 2018]	0.9360
BERT 2	**0.9575**

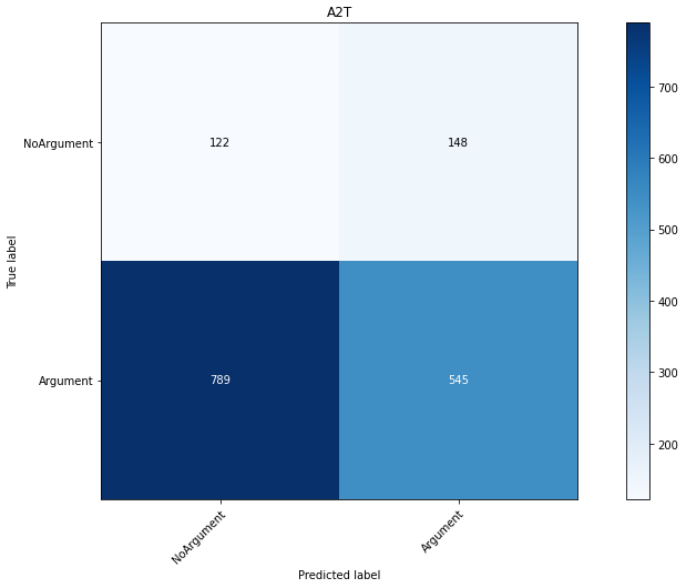
Pipeline test

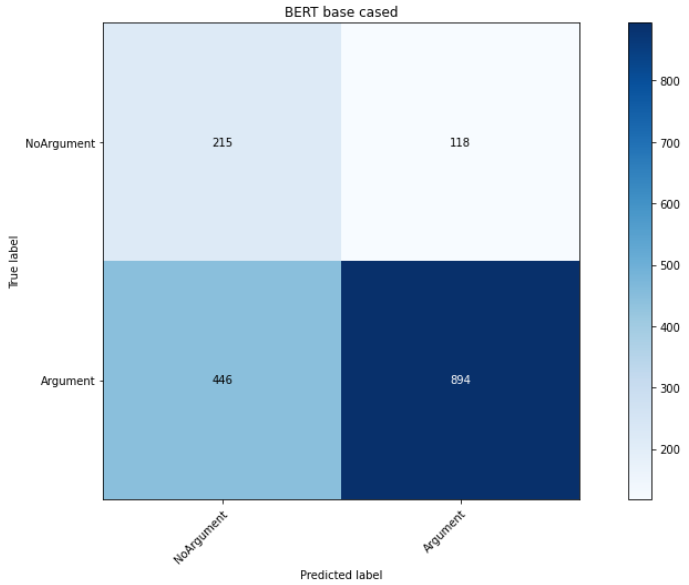
90 essays dataset (1,673 sentences) [Stab and Gurevych, 2014].

Essays were split in sentences each annotated as ”Argument” if it contained a major claim, claim or a premise or ”No Argument” otherwise. Each sentence was also annotated according to the argument component it contained (”Premise” or ”Claim”).

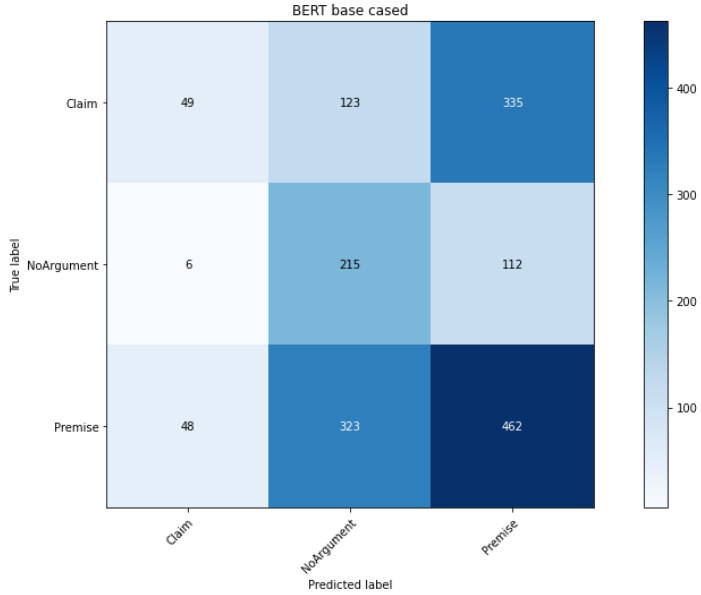
Topic	Sentence	Label0	Label1
Why you should exercise	The number one reason why people should exercise is because it will keep you healthy.	Argument	Claim
Living and studying overseas	It is every student’s desire to study at a good university and experience a new environment.	No argument	No Argument
More people are migrating to other countries than ever before	These feedbacks, in turn, help raise one’s pride of their cultures and help people understand each other more.	Argument	Premise

A2T [Ferrara et al., 2017] vs BERT - argument identification





BERT - argument component identification



Case study - ECB Board Members speeches

“Since I’ve become a central banker, I’ve learned to mumble with great incoherence. If I seem unduly clear to you, you must have misunderstood what I said.”

Alan Greenspan, former FED chair, **1987**

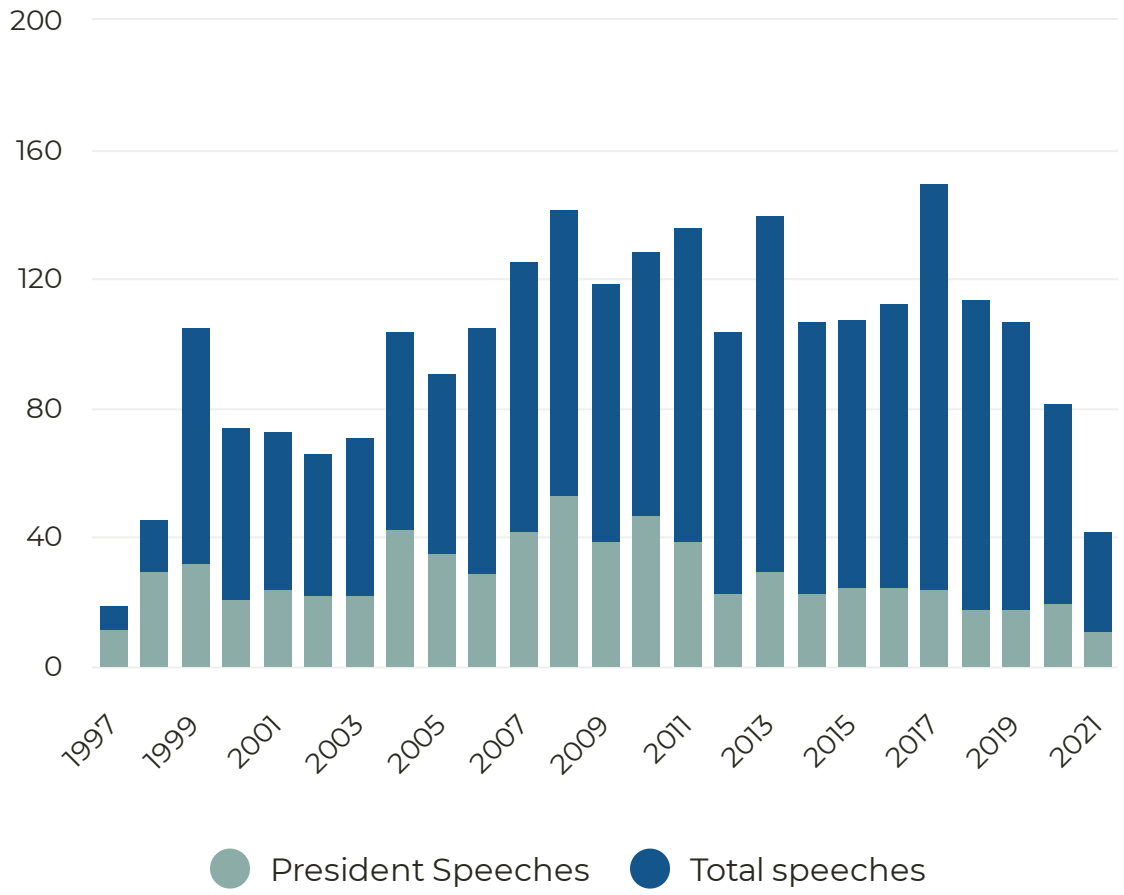
“Today, central bank communication is at the heart of monetary policy. It is actually a monetary policy tool in itself.”

Mario Draghi, former ECB President, **2014**

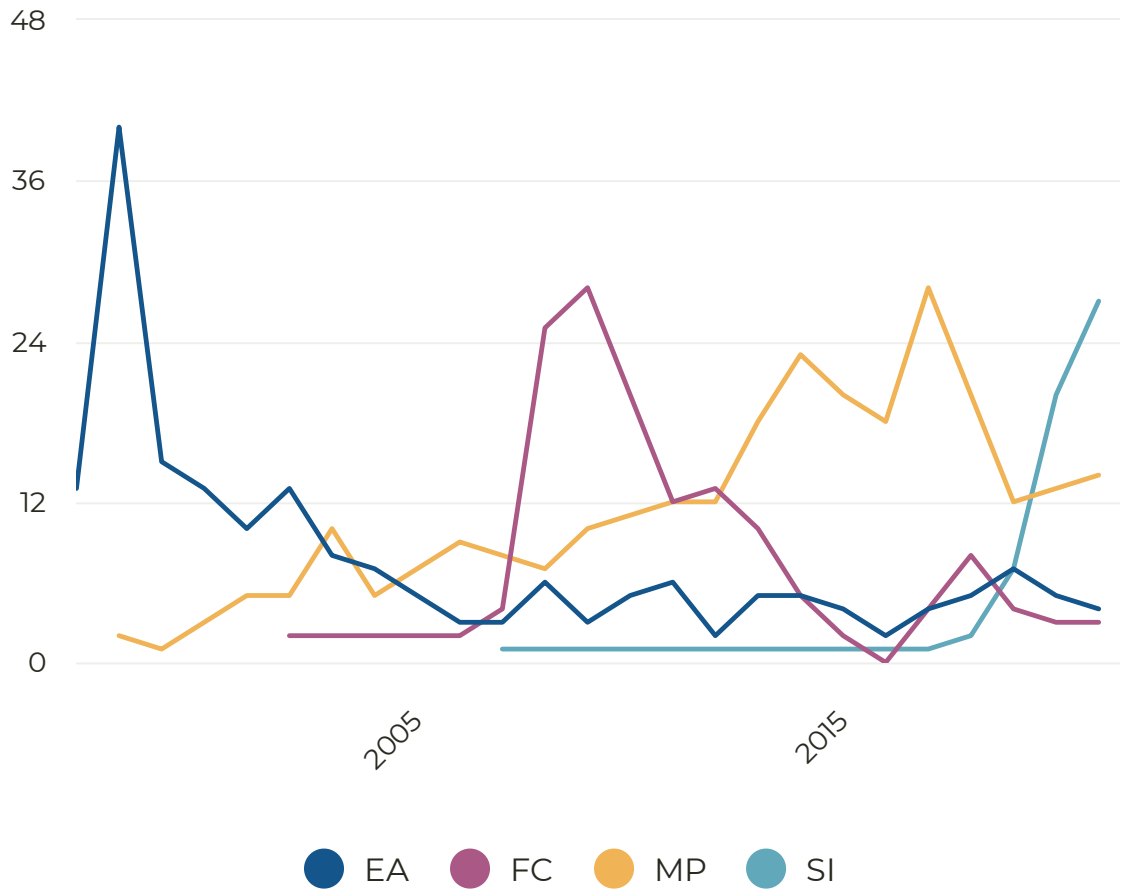
Exploratory data analysis

AVG speech time = 21 minutes

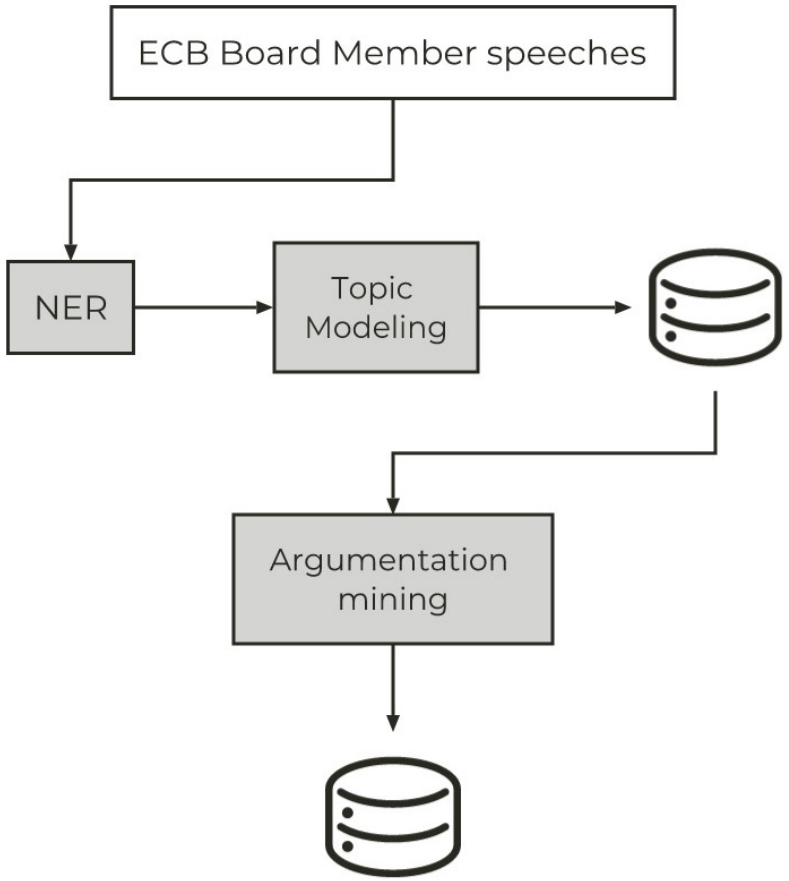
Number of speeches per year



Dynamic Topic Modeling



Extracting structured information from an unstructured set of textual documents



Named Entity Recognition

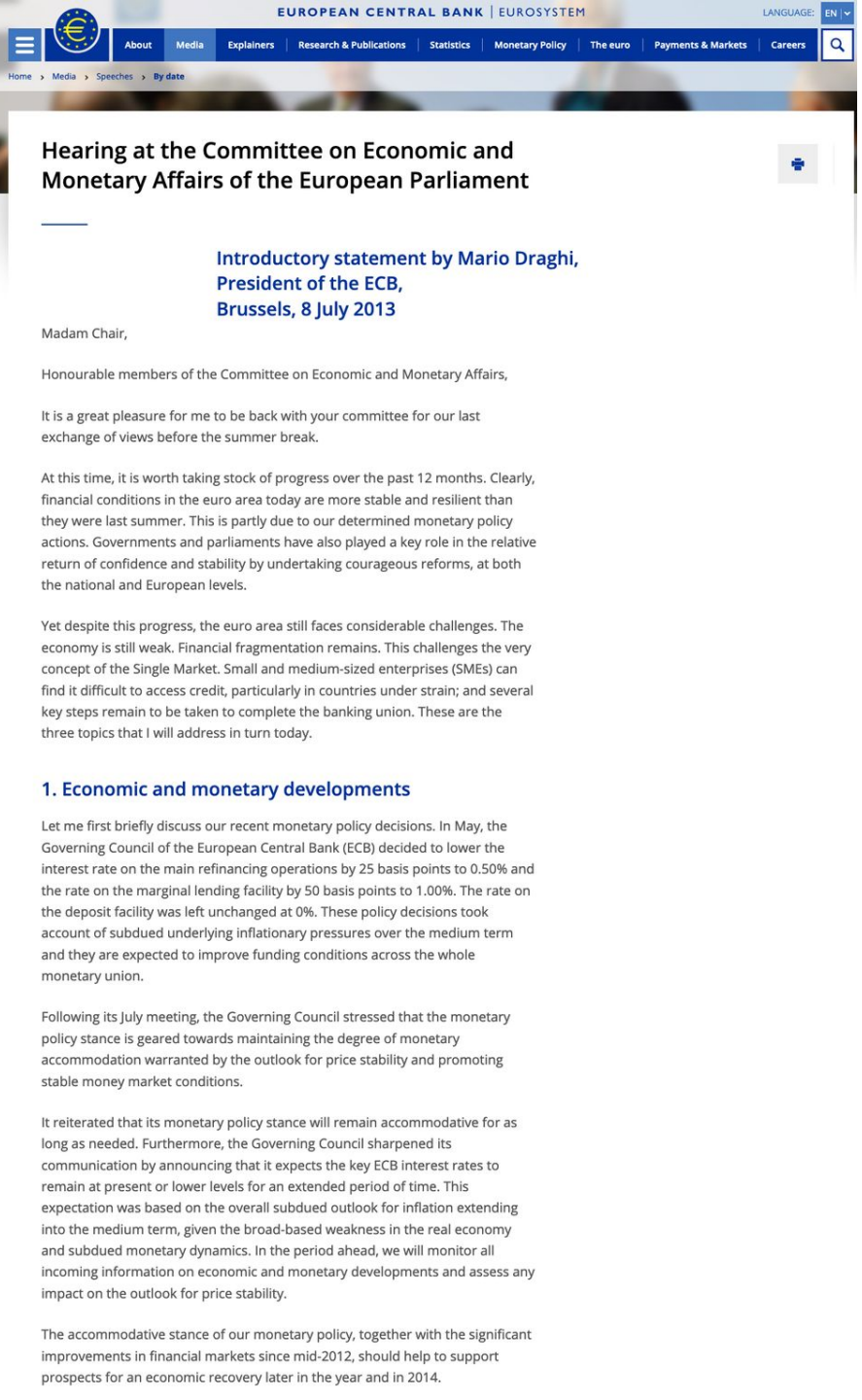
Associate each speech with the named entities it contains (e.g: speakers, locations, conferences)

Topic Modeling

Associate each speech with the topic it discusses and construct a first database associating speeches with topics and named entities

Argumentation Mining

Select the speeches you are interested in and associate them with short sets of argumentative sentences to construct a final index



An example

Hearing at the Committee on Economic and Monetary Affairs of the European Parliament, Mario Draghi, Brussels, July 8, 2013

Non argumentative

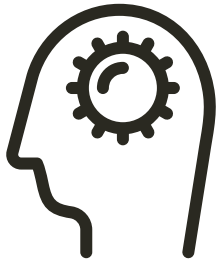
«Let me first briefly discuss our recent monetary policy decisions. In May, the Governing Council of the European Central Bank (ECB) decided to lower the interest rate on the main refinancing operations by 25 basis points to 0.50% and the rate on the marginal lending facility by 50 basis points to 1.00%. The rate on the deposit facility was left unchanged at 0%.»

Premise

«These policy decisions took account of subdued underlying inflationary pressures over the medium term and they are expected to improve funding conditions across the whole monetary union.»

Conclusions and next steps

- Domain-agnostic argumentation mining model to extract structured information from textual documents
- New SOTA performances in cross-domain argument detection
- New SOTA performances in cross-domain argument component detection
- Monetary policy decisions prediction combining standard quantitative data and new textual data
- Explore other language models (e.g: RoBERTa)
- Explore other domains (e.g: political debates)



Argumentation mining helps us defining **what** are people's opinions, but more importantly **why** people hold those particular opinions.

Thank you.