Forecasting and Nowcasting with Text as Data

Module 2 - Session 1: From text to signals

Renato Vassallo

April, 2026

Barcelona School of Economics

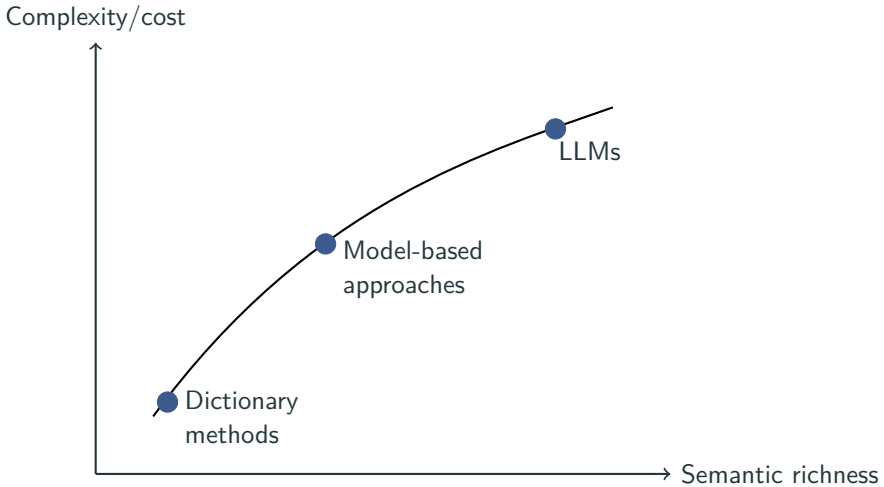
All views expressed here and any remaining errors are my own.

Introduction

Motivation

- A large share of socio-economic information is embedded in text
- Text can be systematically transformed into:
  - labels (e.g. violent vs. non-violent),
  - scores (e.g. policy uncertainty, crime sentiment),
  - signals for monitoring and forecasting.
- A rapidly growing literature exploits textual data to construct high-frequency indicators ▸ Literature
- Main challenge: mapping **unstructured language** into **structured signals**.

A simple taxonomy of text methods



Embeddings

What are embeddings?

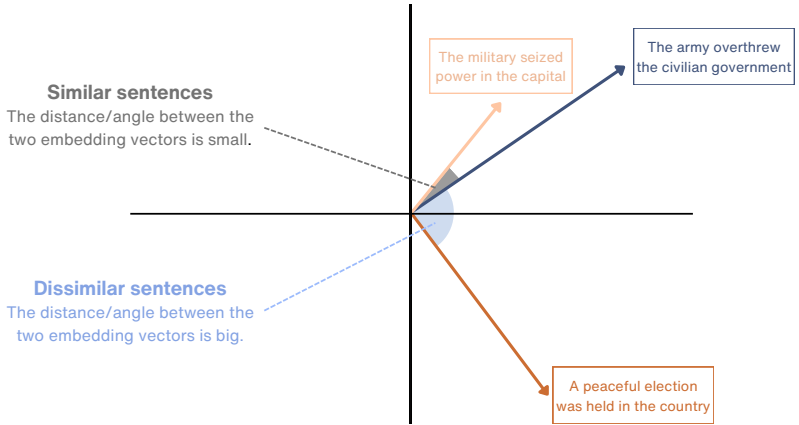


Figure 1: Illustrative representation of headlines in the embedding space. Source: Mayoral et al. (2026).

A common notation

- Let d_i denote a document, and $f_\theta : \mathcal{D} \to \mathbb{R}^k$ be a function (encoder) such that:

$$z_i = f_\theta(d_i), \qquad z_i \in \mathbb{R}^k \text{ (embedding)}$$

- A downstream task then uses this representation:

$$\hat{y}_i = g_\phi(z_i), \qquad g_\phi(\cdot) : \text{ scoring rule / classifier}$$

- Similarity measures:
  1. Dot product: $z_i' z_j$
  2. Cosine similarity: $\cos(z_i, z_j) = \frac{z_i' z_j}{\|z_i\| \, \|z_j\|}$

Contextual embeddings: BERT

- BERT (Devlin et al., 2019) is pre-trained on large corpora.
- With pre-trained BERT we can directly obtain embeddings for any text.

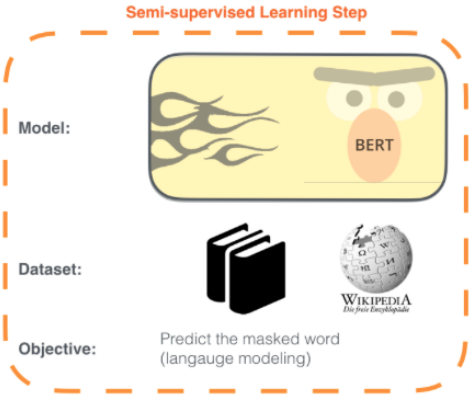


Figure 2: Source: J. Alammar (2019), *The Illustrated BERT*

Scaling up: sentence representations

- Standard BERT produces token-level embeddings:

$$z_{it} = f_\theta(w_{it} \mid w_{i1}, \ldots, w_{iT})$$

- Sentence Transformers are optimized for **semantic similarity** via contrastive objectives:

$$z_i = f_\theta(d_i) \in \mathbb{R}^k$$

An empirical illustration is provided in `session1/01_text_to_signals.ipynb`

Downstream tasks

From embeddings to downstream tasks

- Embeddings are generic; we can **fine-tune** a model to adapt to specific tasks:
  - Sentiment analysis
  - NER
  - Event detection
  - **Text classification**
  - Q&A
- The difference across methods is often not the task itself, but how f_θ and g_ϕ are constructed.
- Let's look at how to fine-tuning a pretrained transformer for text classification!

▸ Simple fine-tuning

Model choice: quality vs efficiency

- Sentence-transformer models vary along two main dimensions:
  1. **Higher-quality / specialized models**:
     - `all-mpnet-base-v2`
     - `e5-large / e5-base` (instruction-tuned embeddings)
     - domain-specific models (finance, legal, multilingual)
  2. **Lightweight models**:
     - `all-MiniLM-L6-v2`
     - `paraphrase-MiniLM`
- In practice, choice depends on **scale, latency, and task complexity**.

From representations to low-supervision decisions

- Suppose resources are limited: no large-scale fine-tuning, costly GPUs, or advanced LLM access.
- Still, pretrained representations can support powerful tasks with minimal supervision.
- We start with **zero-shot learning** (no labeled data), then move to **few-shot learning** (a small labeled set) to adapt to specific tasks.
- Key idea:

$$\text{semantic representations} \Rightarrow \text{task-specific decisions without full retraining}$$

Zero-shot learning

Zero-shot learning

- Goal: classify text into labels without task-specific training examples.
- Let $\mathcal{Y} = \{y_1, \ldots, y_K\}$ be a set of candidate labels.

$$\hat{y}_i = \arg\max_{y \in \mathcal{Y}} \text{score}(d_i, y)$$

- Very flexible
- Minimal labeling cost
- Performance depends heavily on label formulation and underlying model

Zero-shot via natural language inference

- NLI considers two sentences: a *premise* and a *hypothesis*.
- The task is to determine whether the hypothesis is true (**entailment**) or false (**contradiction**) given the premise.

Premise	Label	Hypothesis
The cat is sleeping on the couch.	Contradiction	The cat is playing outside.
The company reported a rise in profits.	Neutral	The company launched a new product.
A group of people is protesting in the street.	Entailment	There is a protest happening.

Table 1: Examples of NLI: Contradiction, Neutral, and Entailment

▸ NLI Architecture

Zero-shot via natural language inference

- For each label y, create a hypothesis:

$$h(y) = \text{“This text is about } y\text{”}$$

- Use the document as premise and evaluate entailment.

$$\hat{y}_i = \arg\max_{y \in \mathcal{Y}} \text{score}(d_i, y) \equiv \arg\max_{y \in \mathcal{Y}} P_\theta(\text{entailment} \mid d_i, h(y))$$

- Pre-trained models: `bart-large-mnli` and `roberta-large-mnli`[1]. Demo here.

An empirical illustration is provided in `session1/01_text_to_signals.ipynb`

[1] MNLI (Multi-Genre Natural Language Inference), is a dataset containing over 430k pairs of sentences labeled as **entailment**, **neutral**, or **contradiction**.

Laboratory

Lab: Time-Series Forecasting

- The literature highlights the importance of business confidence for investment dynamics.
- We will construct a **Business Confidence Index (BCI)** from news text.
- We then assess whether the BCI improves forecasts of private investment.
- Our baseline specification is an AR(1) model:

$$y_t = c + \phi y_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, 1)$$

- Forecast performance will be evaluated using a rolling-window strategy. Be careful to respect the information set available at each point in time.

An empirical illustration is provided in `session1/02_lab.ipynb`

Appendix

Literature: news-based text indices

back

Category	Representative Papers	Method
Economic Policy Uncertainty	Baker et al. (2016), Azqueta-Gavaldón (2017), Ghirelli et al. (2021), Naboka-Krell (2024)	Dictionary / Supervised
Global & Macro Uncertainty	Ahir et al. (2018), Barrett et al. (2020)	Dictionary
Economic Sentiment	Rambaccussing and Kwiatkowski (2020), Nyman et al. (2021), Consoli et al. (2022), Kalamara et al. (2022), Shapiro et al. (2022)	Supervised ML
Business Cycles / Activity	Thorsrud (2020)	Topic Model
Geopolitical / Conflict Risk	Caldara and Iacoviello (2022), Mueller and Rauh (2022), Mayoral et al. (2026)	Dictionary / Topic
Climate and Environmental Risk	Gavriilidis (2021), Ma et al. (2023)	Dictionary / Supervised
LLM-based Financial Text	Li et al. (2023), Kirtac and Germano (2024)	Large Language Models

Notes: The table summarizes representative news-based indices used in macroeconomics and finance. Methods include dictionary approaches, supervised machine learning classifiers, topic models, and large language models.

Simple fine-tuning exercise

back

- We use a simple dataset of sentences with categories: Learning, Pets, Coding.
  - `("I love machine learning", [1, 0, 0])`
  - `("Cats are cute", [0, 1, 0])`
  - `("Python is great for programming", [0, 0, 1])`

- Model architecture:
  - Encoder: $z_i = f_\theta(d_i)$
  - Linear layer: $h_i = Wz_i$
  - Softmax layer: $\hat{y}_i = \text{softmax}(h_i)$
  - Cross-entropy loss to align embeddings with target vectors

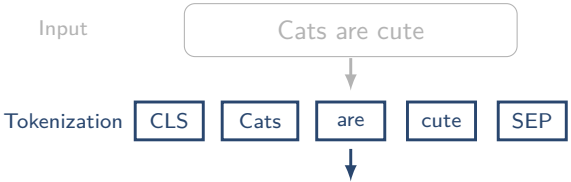
Simple fine-tuning: architecture

back

1) Start from the raw sentence

Simple fine-tuning: architecture

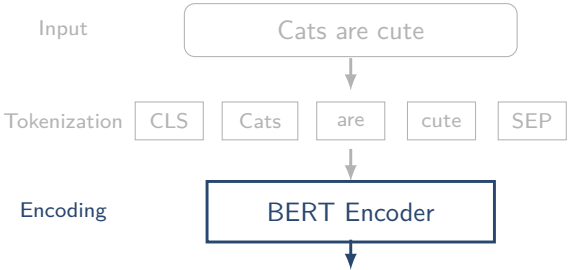
back



2) Tokenize the sentence into model-readable units

Simple fine-tuning: architecture

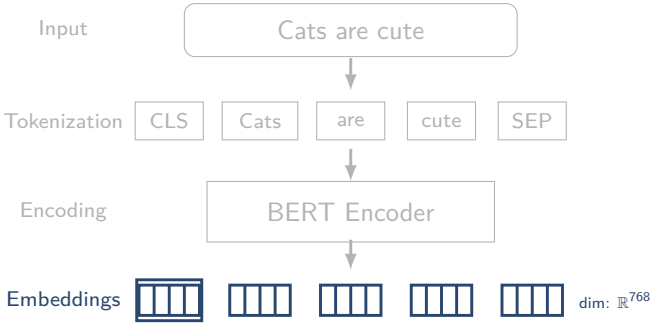
back



3) Obtain contextual representations with a pretrained encoder

Simple fine-tuning: architecture

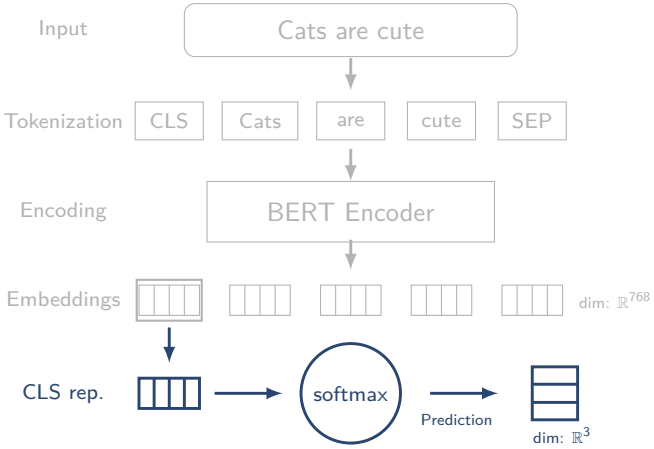
back



4) The encoder returns one high-dimensional embedding per token

Simple fine-tuning: architecture

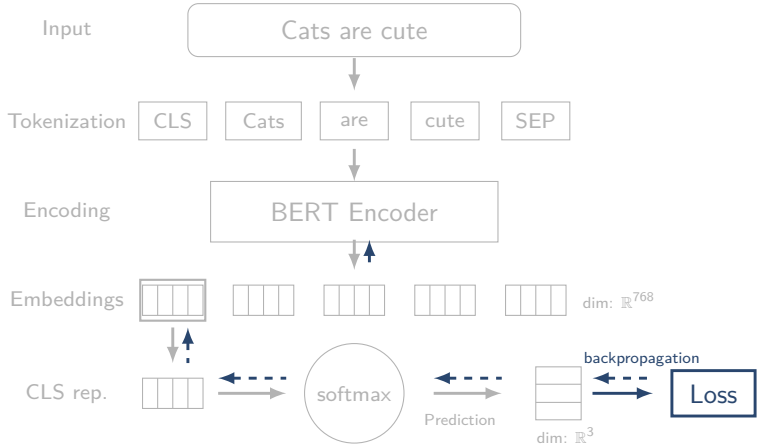
back



5) Use the [CLS] representation to generate a 3-class prediction

Simple fine-tuning: architecture

back

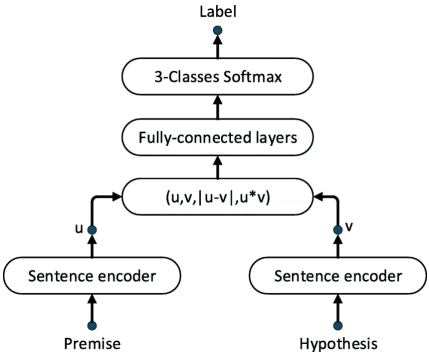


6) Finally, compute the loss and update the model through backpropagation

NLI architecture

Return

- NLI datasets are typically modeled via *sequence-pair classification*.
- The input premise and hypothesis are encoded into vectors u and v.
- They are concatenated and passed to a 3-class classifier consisting of multiple layers.



Source: Sadeghi et al. (2022).

References

Ahir, H., Bloom, N., and Furceri, D. (2018). The world uncertainty index. *Mimeo*.

Azqueta-Gavaldón, A. (2017). Developing news-based economic policy uncertainty index with unsupervised learning. *Economics Letters*, 160:109–112.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Barrett, P., Appendino, M., Nguyen, K., and de León Miranda, J. (2020). Measuring social unrest using media reports. *IMF Working Papers 2020/129*. International Monetary Fund.

Caldara, D. and Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4):1194–1225.

Consoli, S., Barbaglia, L., and Manzan, S. (2022). Fine-grained, aspect-based sentiment analysis on economic and financial lexicon. *Knowledge-Based Systems*, 247. 108781.

References ii

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. pages 4171–4186. Association for Computational Linguistics.

Gavriilidis, K. (2021). Measuring climate policy uncertainty. *Working Paper*. SSRN Electronic Journal.

Ghirelli, C., Pérez, J. J., and Urtasun, A. (2021). The spillover effects of economic policy uncertainty in latin america on the spanish economy. *Latin American Journal of Central Banking*, 2(2). 100029.

Kalamara, E., Turrell, A., Redl, C., Kapetanios, G., and Kapadia, S. (2022). Making text count: Economic forecasting using newspaper text. *Journal of Applied Econometrics*, 37(5):896–919.

Kirtac, K. and Germano, G. (2024). Sentiment trading with large language models. *Finance Research Letters*, 62(Part B). 105227.

Li, X., Chan, S., Zhu, X., Pei, Y., Ma, Z., Liu, X., and Shah, S. (2023). Are ChatGPT and GPT-4 general-purpose solvers for financial text analytics? a study on several typical tasks. *arXiv preprint: 2305.05862*.

Ma, Y.-R., Liu, Z., Ma, D., Zhai, P., Guo, K., Zhang, D., and Ji, Q. (2023). A news-based climate policy uncertainty index for china. *Scientific Data*, 10(1). 881.

References iii

Mayoral, L., Mueller, H., Rauh, C., Phillip, M., and Vassallo, R. (2026). Semantic similarity measures in newspaper text for detecting and predicting disruptive institutional events. *BSE Working Paper 1555*. Barcelona School of Economics.

Mueller, H. and Rauh, C. (2022). The hard problem of prediction for conflict prevention. *Journal of the European Economic Association*, 20(6):2440–2467.

Naboka-Krell, V. (2024). Construction and analysis of uncertainty indices based on multilingual text representations. *Economics Letters*, 237. 111653.

Nyman, R., Kapadia, S., and Tuckett, D. (2021). News and narratives in financial systems: Exploiting big data for systemic risk assessment. *Journal of Economic Dynamics and Control*, 127. 104119.

Rambaccussing, D. and Kwiatkowski, A. (2020). Forecasting with news sentiment: Evidence with uk newspapers. *International Journal of Forecasting*, 36(4):1501–1516.

Sadeghi, F., Bidgoly, A. J., and Amirkhani, H. (2022). Fake news detection on social media using a natural language inference approach. *Multimedia Tools and Applications*, 81(20):29077–29104.

Shapiro, A. H., Sudhof, M., and Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, 228(2):221–243.

References iv

Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business and Economic Statistics*, 38(2):393–409.