semantic-features: A User-Friendly Tool for Studying Contextual Word Embeddings in Interpretable Semantic Spaces

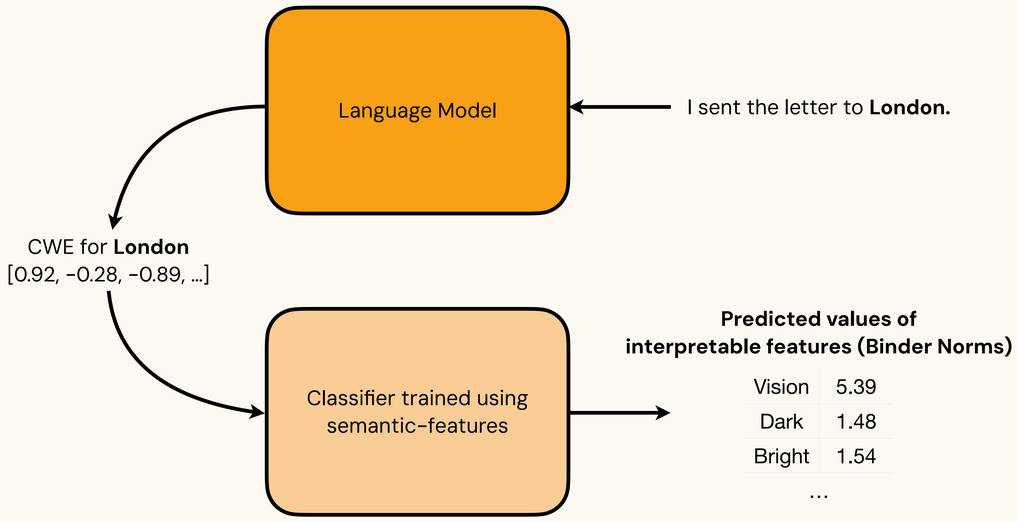
Jwalanthi Ranganathan[1], Rohan Jha[1], Kanishka Misra[1*,2], Kyle Mahowald[1]

[1] The University of Texas at Austin
[2] Toyota Technological Institute at Chicago

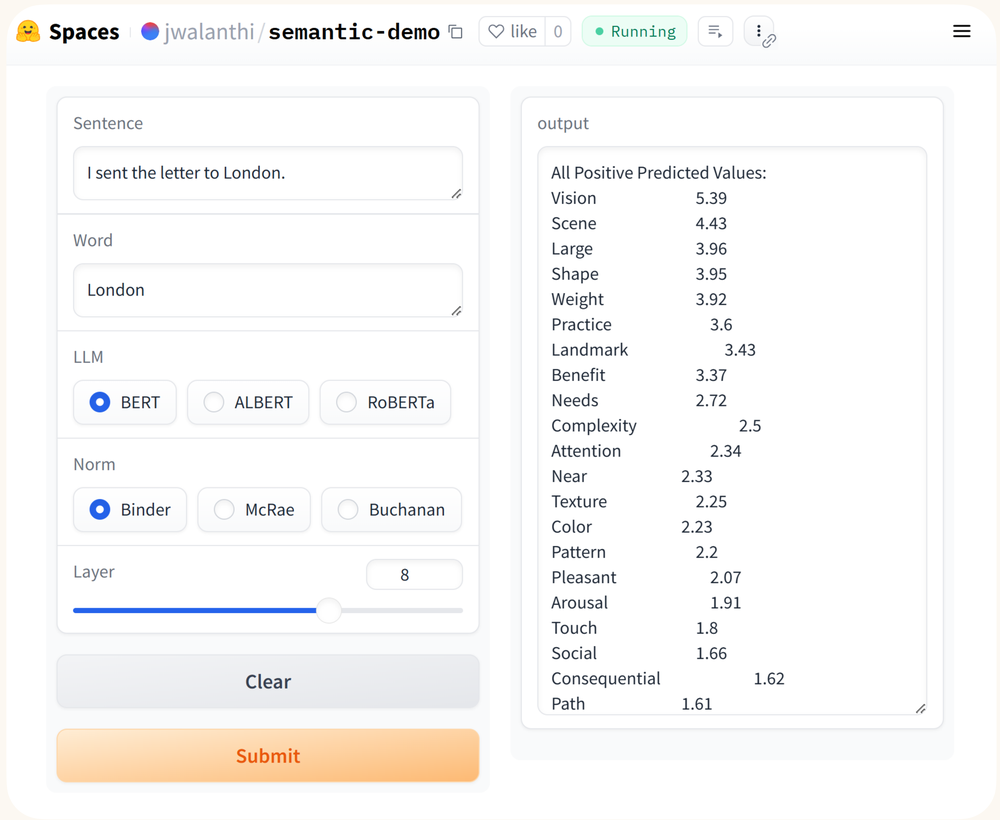
* Work partly done as a Postdoc at The University of Texas at Austin

NSF

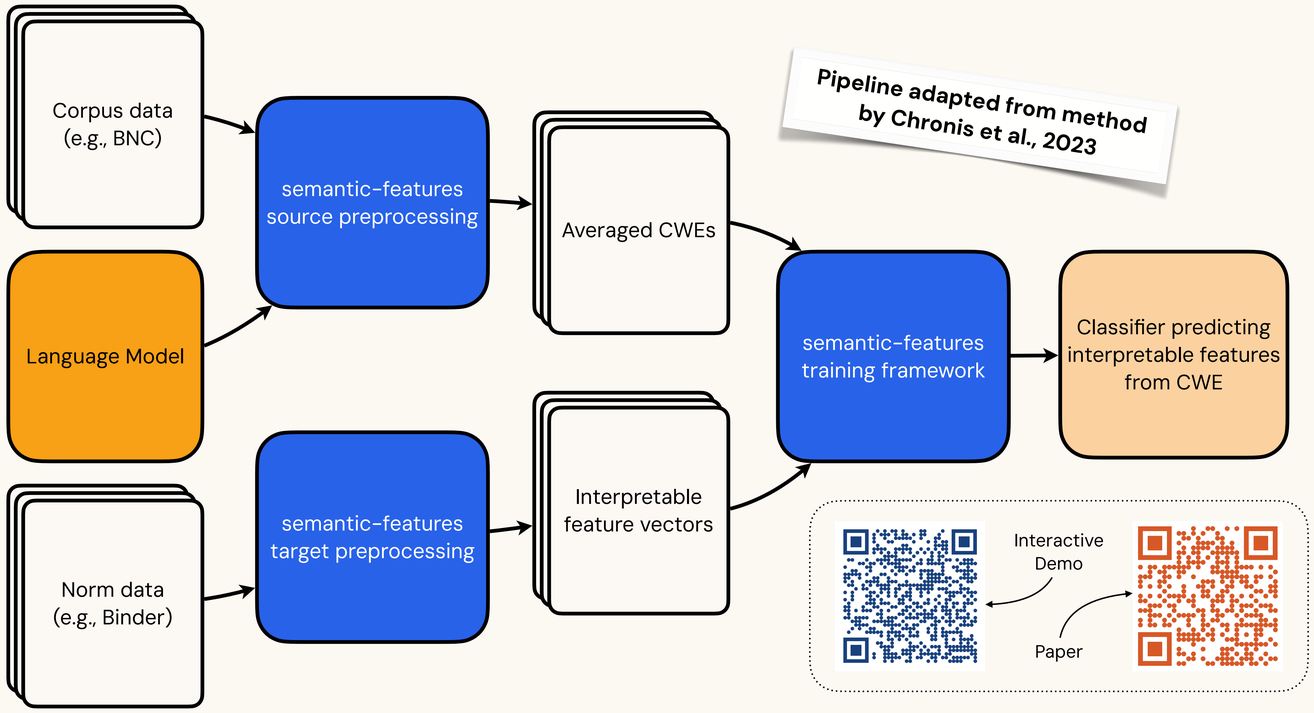
Prediction Backend



Predictions Using `semantic-demo`



Classifier Training Using `semantic-features`



Case Study: Recipient Semantics in Dative Constructions

Prepositional Object (PO): I sent the letter to **London**. → 

Double Object (DO): I sent **London** the letter. → 

(Goldberg, 1995; Hovav and Levin, 2008; Beavers, 2011)

Feature	Definition
Biomotion	showing movement like that of a living thing
Body	having human or human-like body parts
Human	having human or human-like intentions, plans, or goals
Face	having a human or human-like face
Speech	someone or something that talks
Landmark	having a fixed location, as on a map
Scene	bringing to mind a particular setting or physical location

Table 1: Feature definitions from Binder et al. (2016).

Feature	DO	PO
Biomotion	**1.19**	0.43
Body	**1.00**	0.26
Human	**0.89**	0.48
Face	**0.71**	0.19
Speech	**0.68**	0.13
Landmark	1.83	**3.43**
Scene	2.59	**4.43**

Table 2: Relevant Binder features predicted for "London" in (1) using CWEs from BERT layer 8. The PO construction lends itself more towards "location" features, and the DO more towards animate features.

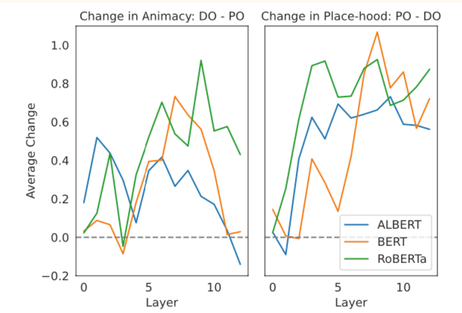


Figure 2: For each layer of an LM, we extract the CWE and project it into Binder space. **Left:** we measure the average change across the test sentences in Person features from PO to DO. The positive values indicate that recipients in the DO are found to be more animate. **Right:** we measure the average change across test sentences in Place features from DO to PO. Here, the positive values indicate that the recipients were found to be more place-like in the PO.

Correspondence: jwalanthi@utexas.edu; kanishka@ttic.edu

SCiL 2025, Eugene