CITY-GPT
Spatial mismatch with embeddings + transformers

Andrew Renninger[1,*], Elsa Arcaute[1], Neave O'Clery[1]

Applying text generation techniques to predict business configurations or visit sequences in an urban setting.

Strategy

- Predicting urban amenities from different amenity graph representations.
- Two models: **Proximity** (geographic closeness) and **Affinity** (shared visitors).
- What do the differences between these models mean?

Our model answers the following questions: What businesses tend to locate near existing amenities here? What businesses tend to attract people that come the amenities in this area?

Transforming next-word into next-business prediction allows us understand alternative configurations in cities.

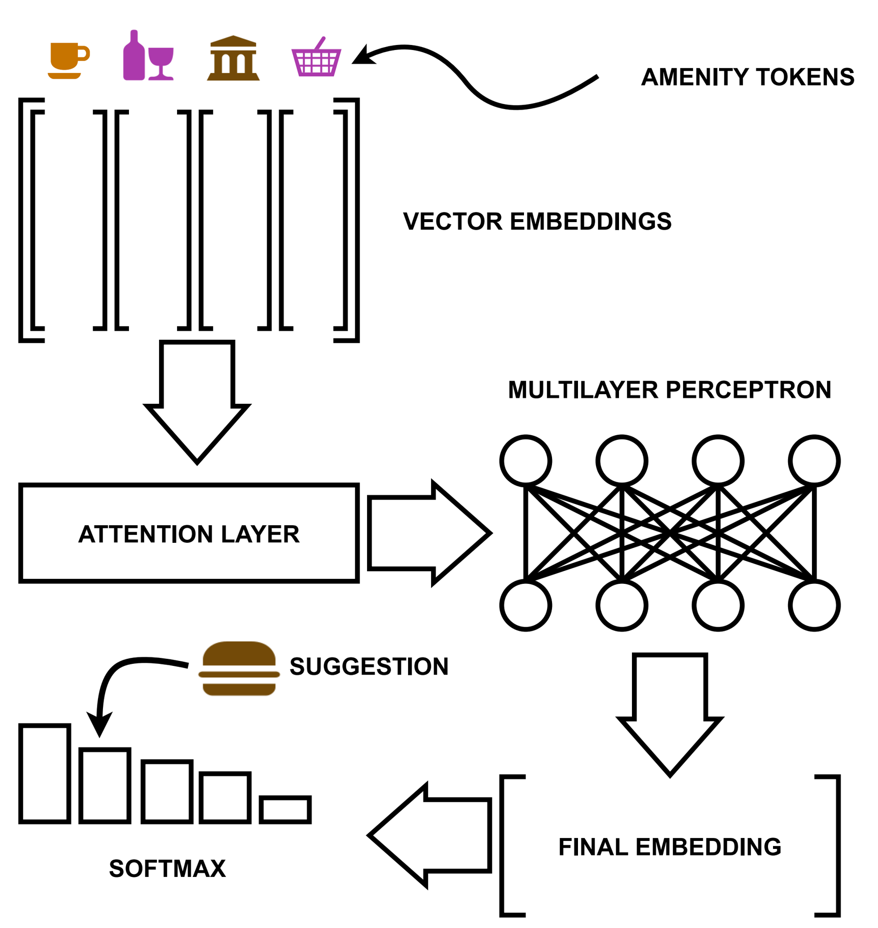

Figure 1: Diagram illustrating a neural architecture that processes batches of amenity tokens—derived from either street networks or trajectories—by converting them into vector representations from graph embeddings. These vectors are then fed into a neural network featuring an attention layer, which focuses on relevant data points. The result is a probability distribution obtained via softmax, from which the model samples to generate the final suggestions.

Why?

Our claim is that the suggestions these models make tell us important things about existing cities and amenities: comparing predicted and observed, they are measures of surprise or goodness-of-fit between amenity and location; comparing proximity-based and affinity-based predictions, they indicate spatial mismatch between urban amenities and human needs.

Know [an amenity] by the company it keeps.

GPS data from US and UK (US shown here). Nodes are brands of amenity or aggregate classes if it is not a chain (E.g. Independent Pizzeria). Weights are the number of connections—either in shared visits or shared locations—between those brands. Create graph embeddings (shown in Figs. 2A, 2B) for these representations and feed them into neural networks with (essentially) GPT-2 architecture. The models learn to predict the next point of interest given a sequence of POIs. We need "sentences" to predict missing words; For affinity we keep the existing visitation order from trajectories, for proximity we create them using random walkers along the street network.

Results

We start by examining the learned embeddings, using Jaccard similarity to quantify the similarity between each brand's nearest neighbors obtained from co-location and from co-visitation (Fig. 2C). Taking 20 neighbours, mean similarity is 0.11, indicating a fair amount of disagreement: brands that share visitors tend not to share locations.

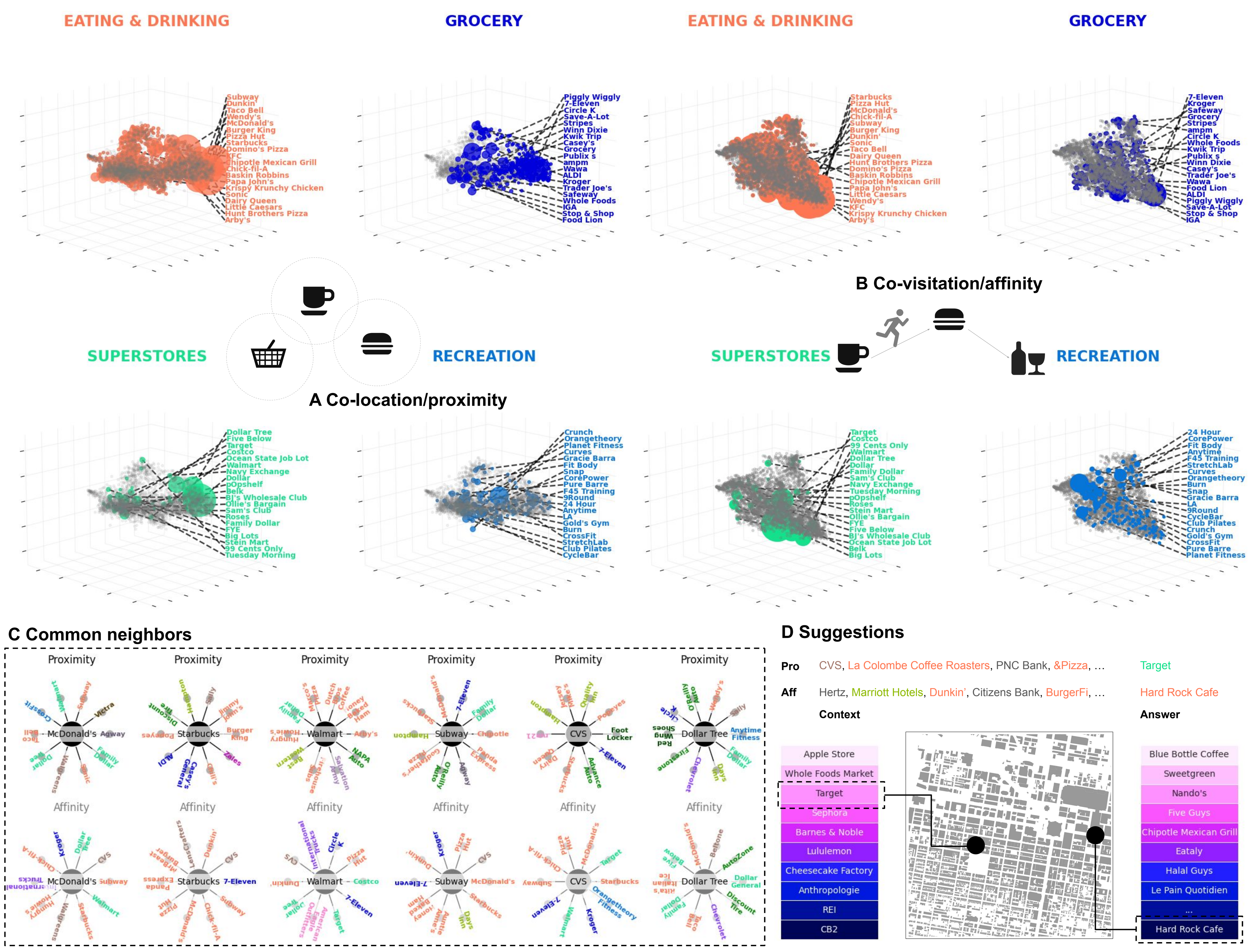

Figure 2: **A** Learned embeddings derived from the proximity network and **B** those from the affinity network. We label the 20 largest brands per category, and see that chain restaurants co-locate while grocers or recreative facilities like gyms tend not to. Dollar stores are nearby in both affinity and proximity spaces, indicating that they share visitors and locations. **C** Nearest neighbours in either spaces for select brands: McDonald's has similar compositions while Walmart does not, suggesting a mismatch. **D** Examples of predictions given a context, yielding a good fit for in one case (the observed Target is 3rd in ranked probability) and a bad fit in another (Hard Rock Cafe).

88.5%

Model accuracy on out-of-sample predictions.

Comparing predicted and observed

- Generate a "sentence" of POIs, either from random walkers or real trajectories
- Gather likely candidate POIs to complete the sentence and rank by probability
- How do the probabilities vary between proximity and affinity models?
- Where does the true next POI fall on the probability distributions?

We also test our GPT by giving limited context how often the prediction appears in a larger window. For example, **proximity** suggestions for a Target store (Fig. 2D) include predictions for an Apple Store, a Lululemon and other brands that are nearby. The **affinity** model struggles with infrequent POIs, failing to predict an observed Hard Rock Cafe.

Applications

When you transpose amenities within a city using recommendations from this model, you can create a new configuration where proximity and affinity are more correlated. This could reduce journey times and improve quality of life in cities.

Affiliations

[1] Centre for Advanced Spatial Analysis, UCL

University College London

*andrew.renninger.12@ucl.ac.uk