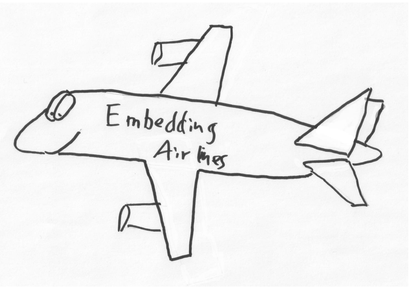
Understanding Neural Embeddings

Oliver Zeigermann @DJCordhose



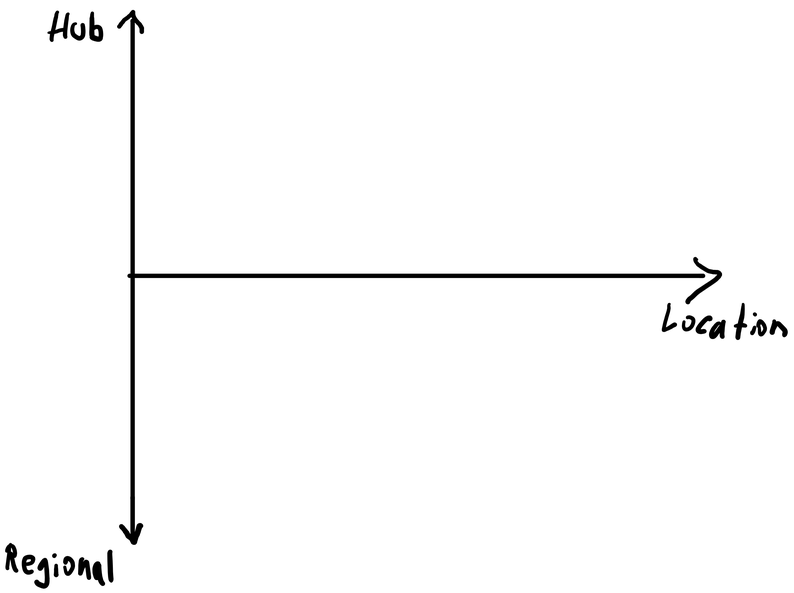
Understanding Semantic Embeddings

Basic question: How to turn categories into numbers?

- neural networks can not deal with symbols
- only numerical values can be processed
- even words and texts can be seen as categories / symbols

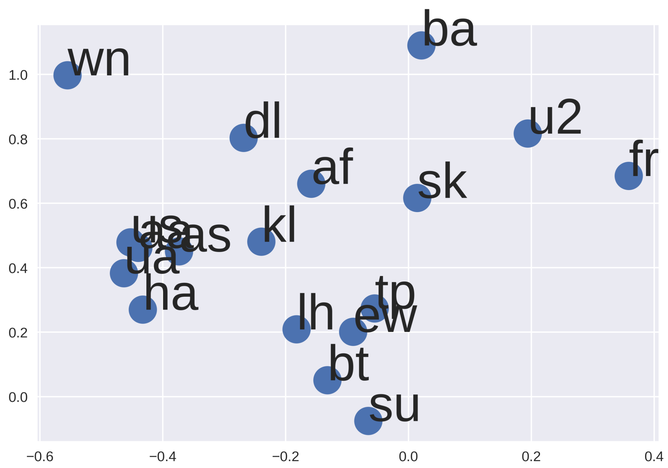
Ideal Solution: make those numbers carry the semantics of the words or symbols

Humans are good at this: Locate your favorite airport below

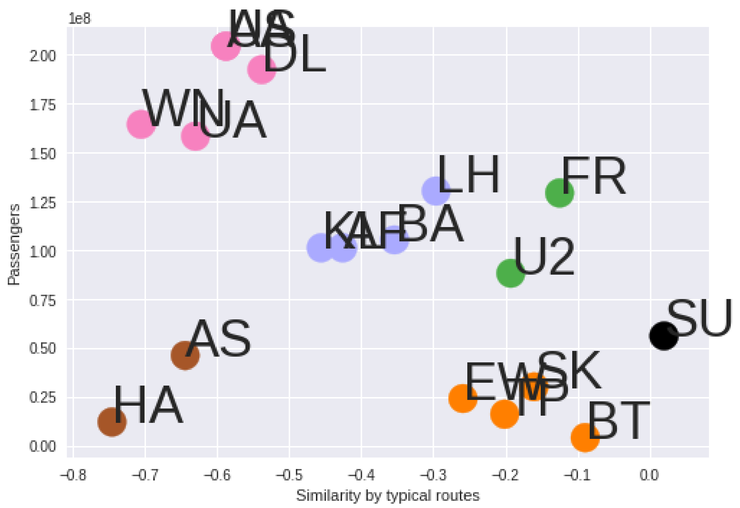


Sample Application: Airlines

Use Case 1: n-dim Embedding as input layer for further processing



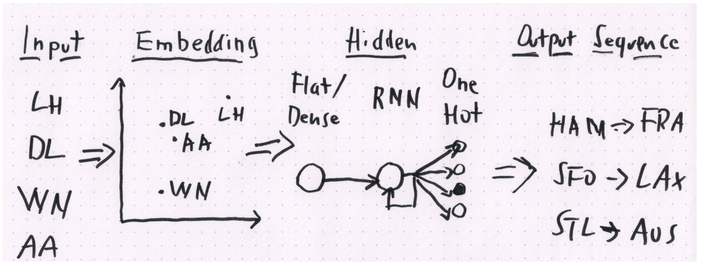
Use Case 2: 1-d Embedding vs Passengers as a plot / cluster for data visualization



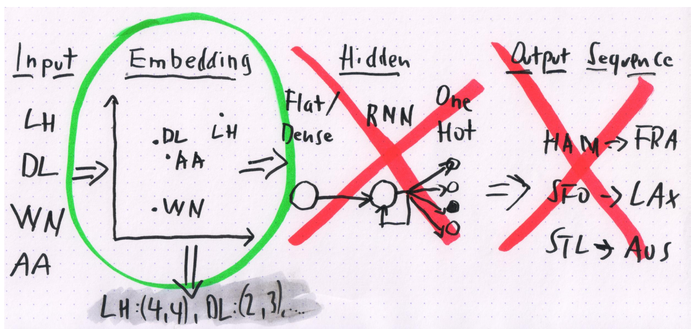
How to Embed

Basic Idea: Train a model and squeeze input through a bottle-neck. Use that bottle-neck as the semantic representation during prediction.

1. What do you want to embed? **Airlines**
2. What data do you want to use to express similarity? **Typical Routes flown by Airline**
3. Prepare data accordingly
4. Set up network and choose loss function
5. **Train** with small batch size and make sure model trains (loss goes down)



6. **Predict** categories and use Embedding layer (the bottle-neck) as new output



Link to notebook: http://bit.ly/embed-airlines