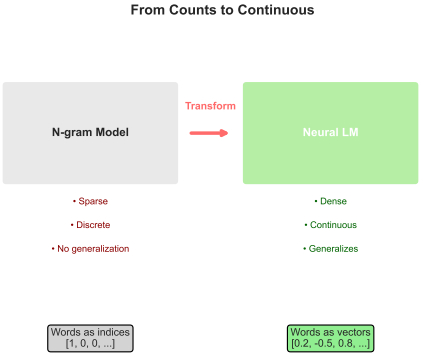
Week 2: Neural Language Models & Word Embeddings

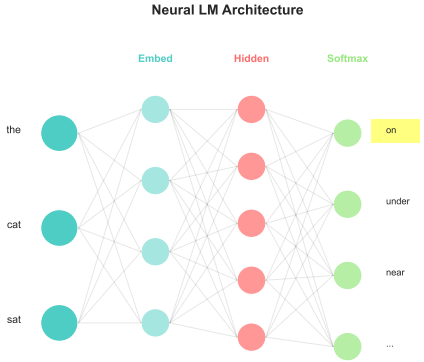
From Discrete to Continuous Representations

Natural Language Processing Course

BSc Computer Science

The Neural Revolution in Language Modeling





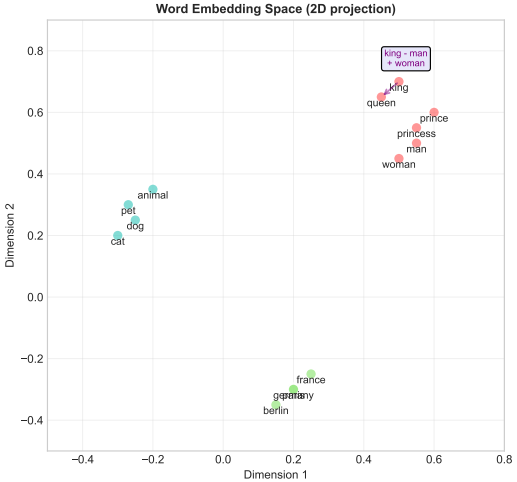
Paradigm Shift:

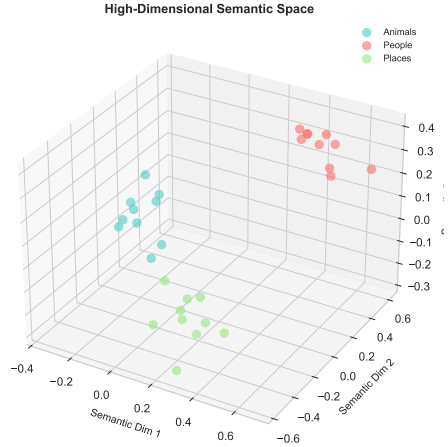
- **From Counting to Learning**
  - No more sparse matrices
  - Learn representations from data

Key Innovations:

- Distributed representations
- Backpropagation training

Word Embeddings: Semantic Space Discovery





Semantic Properties

- Similar words cluster
- Relationships as vectors

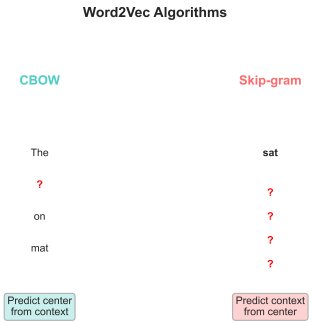
Training Methods

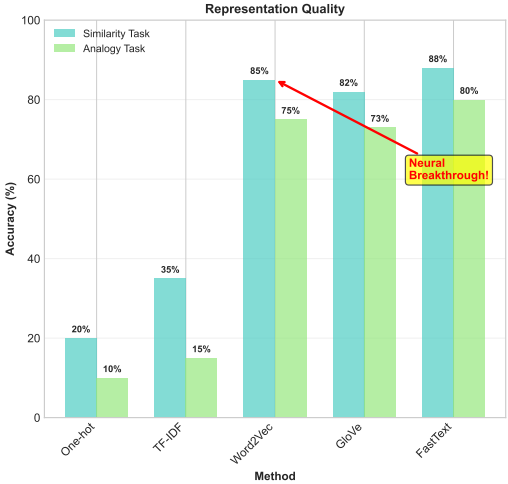
- Word2Vec (2013)
- GloVe (2014)

Dimensions

- Typical: 50-300 dims
- Each captures feature

Applications & Real-World Impact





Word2Vec Algorithms:

- **CBOW (Continuous Bag of Words)**
  - Predict center from context

Why This Matters:

- Foundation for all modern NLP
- Transfer learning enabler