Word Embeddings: Word2Vec & GloVe

> Dense vector representations that capture semantic relationships

🔤 Word2Vec

Learns embeddings from local context

CBOW | Skip-gram

🌍 GloVe

Learns from global word co-occurrence statistics

Key Properties

- 🎯 Similar words have similar vectors
- 📦 Much lower dimensions than vocabulary
- 🔄 Transfer learning to downstream tasks
- 🧮 Captures semantic analogies

🎓 Classic Analogy Example

```
king - man + woman ≈ queen
```

⚙️ Typical Dimensions: **100-300** (vs. thousands in vocabulary size)

📚 Word2Vec Training Process (Window Size = 2)

CBOW (Continuous Bag of Words)

"The quick brown fox jumps over "

Step 1: One-Hot Encoding ↓

Context Words (Input):

```
quick: [0,1,0,0,0,...] (vocab size)
brown: [0,0,1,0,0,...]
jumps: [0,0,0,1,0,...]
over: [0,0,0,0,1,...]
```

Input Layer (One-Hot)
↓
Hidden Layer (Embeddings)
↓
Output Layer (Softmax)

Target (Output):

```
fox: [0,0,0,0,0,1,0,...] (one-hot)
```

🎯 Predict center word from averaged context embeddings

Skip-gram

"The quick brown fox jumps over"

Step 1: One-Hot Encoding ↓

Center Word (Input):

```
fox: [0,0,0,0,0,1,0,...] (vocab size)
```

Input Layer (One-Hot)
↓
Hidden Layer (Embeddings)
↓
Output Layer (Softmax) ×4

Target (4 Context Words):

```
quick: [0,1,0,0,0,...]
brown: [0,0,1,0,0,...]
jumps: [0,0,0,1,0,...]
over: [0,0,0,0,1,...]
```

🎯 Predict each context word from center word embedding