



# Day 8: Word2Vec & GloVe – Learning Word Embedding Techniques

This presentation explores Word2Vec and GloVe. Learn about how machines comprehend word meanings. These techniques transform words into numerical vectors. They capture semantic relationships. This allows machines to process and understand human language. Let's dive into the world of word embeddings!

# Introduction to Word Embeddings

## Definition

Transforming words into numerical vectors. This captures semantic relationships between words.

## Purpose

Enables machines to process and understand human language effectively.

## How it Works

Each word is mapped to a high-dimensional vector. The vector represents its contextual meaning.

## The image is a full-page abstract graphic. It features a dense, interconnected network of stylized, elongated shapes that resemble molecular structures or a complex web. The shapes are primarily in shades of purple and blue, with some lighter, almost white, lines connecting them. The text within these shapes is mirrored and rotated, creating a complex, non-representational pattern. The overall effect is one of a highly detailed, futuristic, or scientific structure.

Words appearing in similar contexts tend to have similar meanings.

Meaning is derived from the words that surround a given word.

"Cat" and "Dog" appear in similar contexts. This indicates related meanings.



# Overview of Word2Vec



## Developed By

Tomas Mikolov et al. at Google in 2013.



## Approach

Predicts a word based on its context (CBOW). Or, it predicts context words from a target word (Skip-Gram).



## Key Feature

Focuses on local context to learn word relationships.



# Overview of GloVe

1

Developed By

Researchers at Stanford University created GloVe.

2

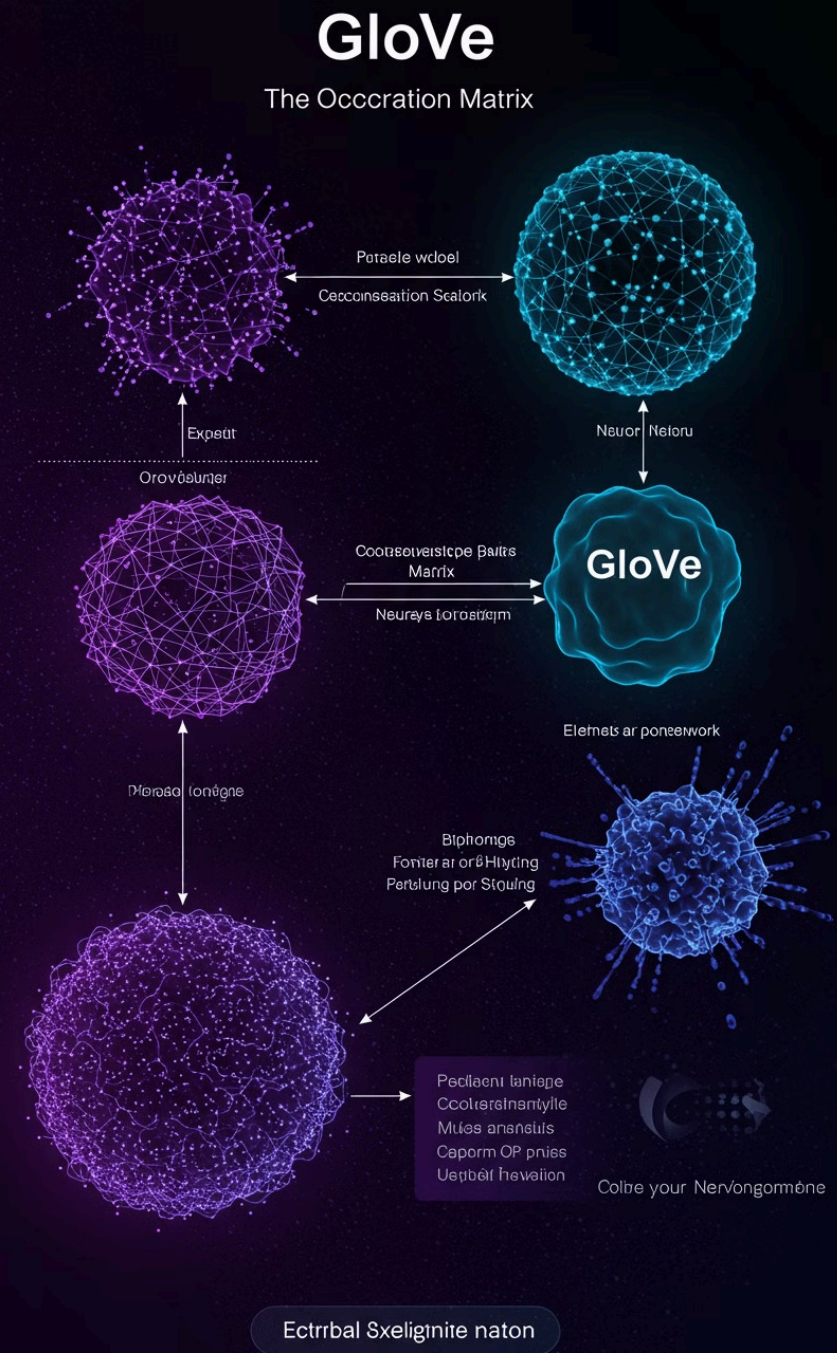
Approach

It utilizes global word co-occurrence statistics. This captures the overall meaning of words in context.

3

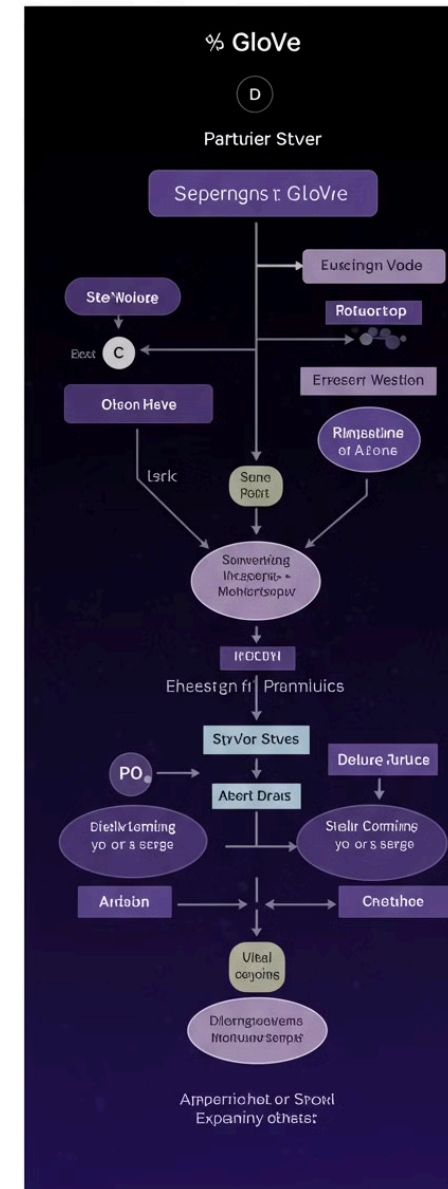
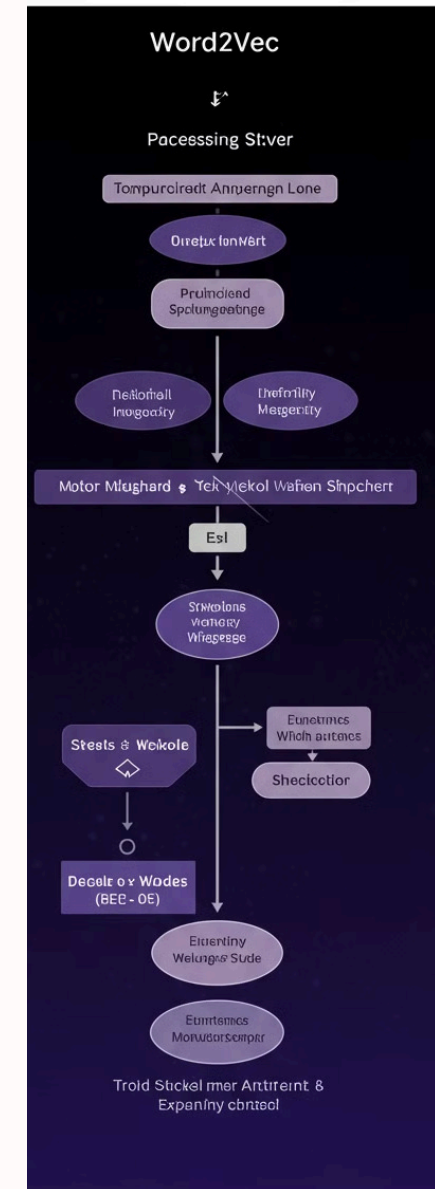
Key Feature

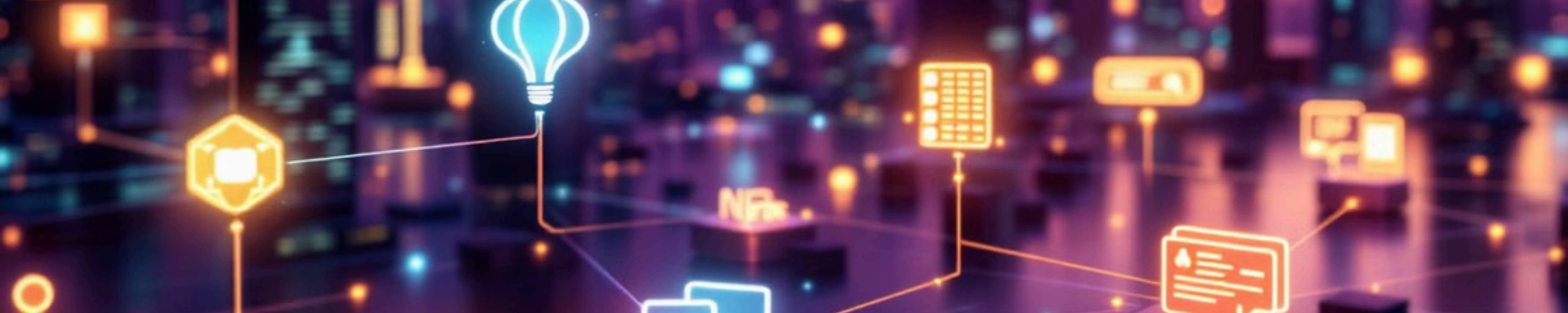
It balances local context with global statistical information for accuracy.



# Key Differences: Word2Vec vs. GloVe

Characteristic	Word2Vec	GloVe
Training Method	Predictive, focuses on local context	Counts-based, leverages global co-occurrence
Data Efficiency	Efficient with large datasets	Effective with smaller, well-structured data





# Practical Applications

1

## Sentiment Analysis

Determine the emotional tone behind a text.

2

## Machine Translation

Translate text from one language to another.

3

## Information Retrieval

Find relevant documents based on a query.

# Training with Gensim

## Preprocess Text Data

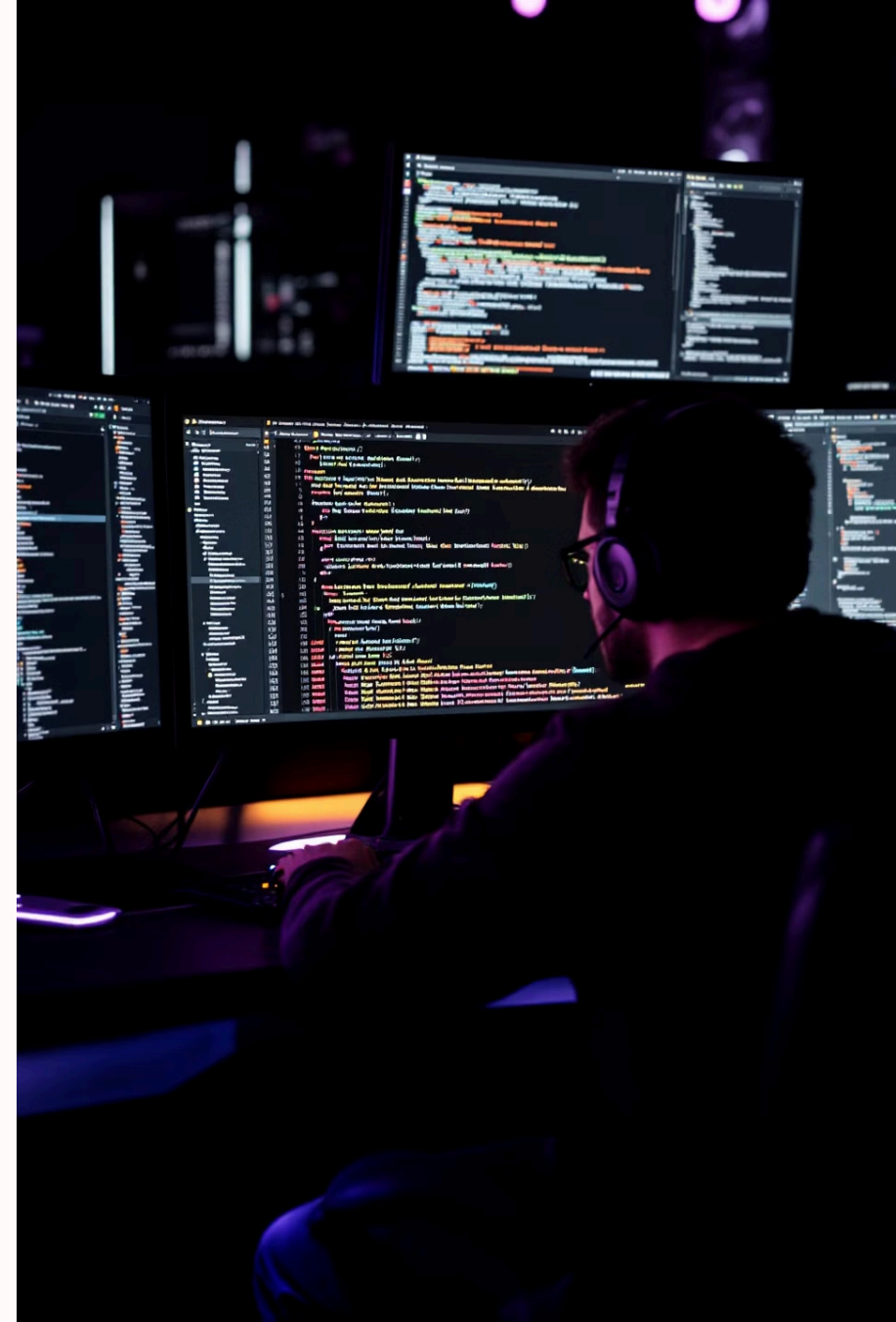
Clean and format text for training.

## Initialize and Train

Set parameters and train the Word2Vec model.

## Explore Embeddings

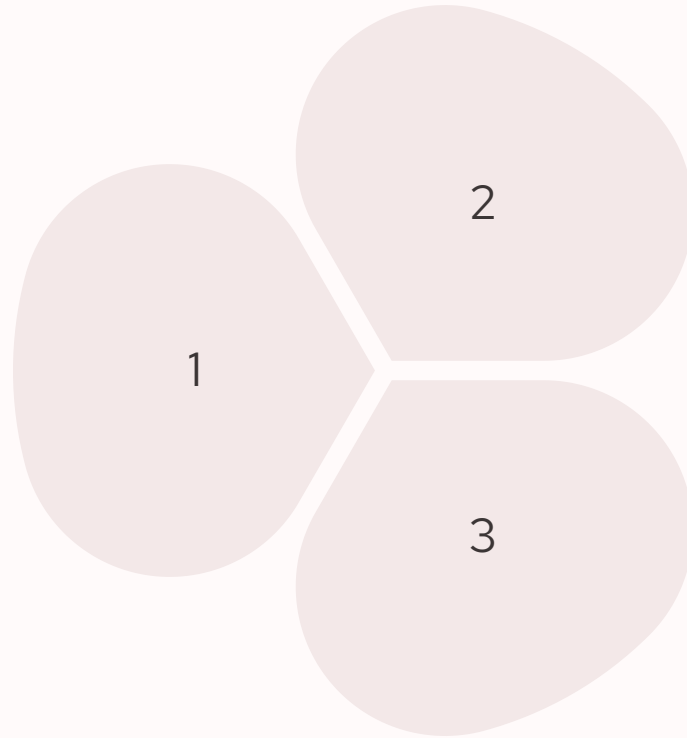
Analyze and use the learned word vectors.





# Conclusion & Key Takeaways

Numerical Representations  
Word embeddings transform text.



Foundational Techniques  
Word2Vec and GloVe are key.

Enhance Models  
Understanding these methods is crucial.