

Embeddings & Reinforcement Learning: The Essential Visuals

20 Charts Every Data Scientist Should Know

Prof. Dr. Jörg Osterrieder

Methods and Algorithms — MSc Data Science

Spring 2026



XKCD #1838 by Randall Munroe (CC BY-NC 2.5)

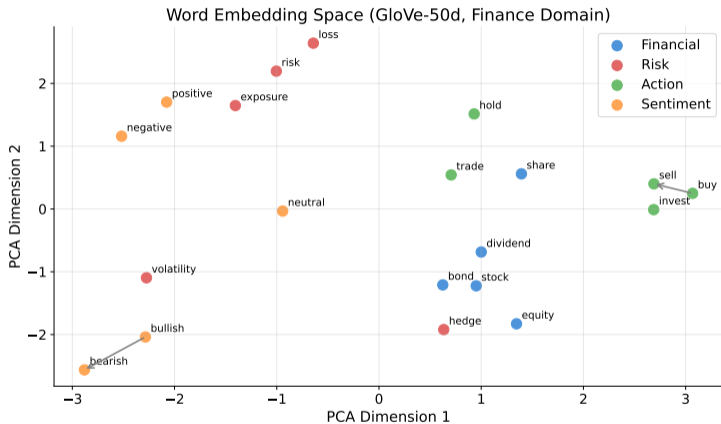
After this lecture, you will be able to:

- **Interpret** embedding visualizations to assess representation quality
- **Analyze** RL training dynamics through reward and value function charts
- **Compare** exploration strategies using cumulative regret analysis

We cover 20 essential visualizations spanning word embeddings, reinforcement learning, and advanced topics

Three parts: Word Embeddings, Reinforcement Learning, Advanced Topics

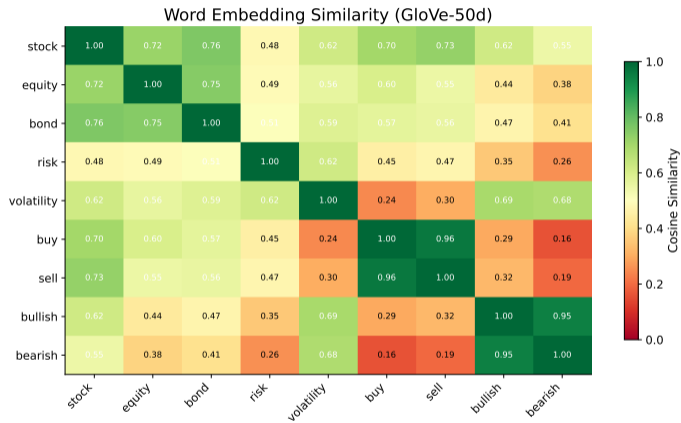
Words as Vectors: The Embedding Space



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/01_word_embedding_space

Words with similar meanings cluster together in the high-dimensional embedding space

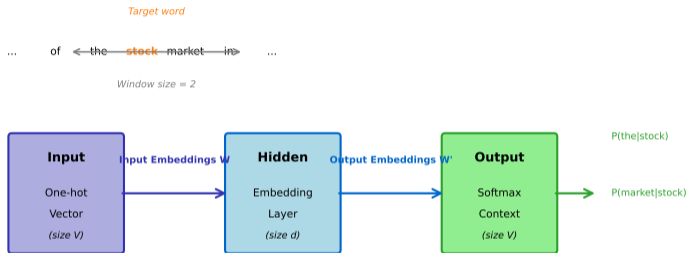
Cosine Similarity Between Word Vectors



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/02_similarity_heatmap

Cosine similarity measures the angle between word vectors—closer to 1 means more semantically similar

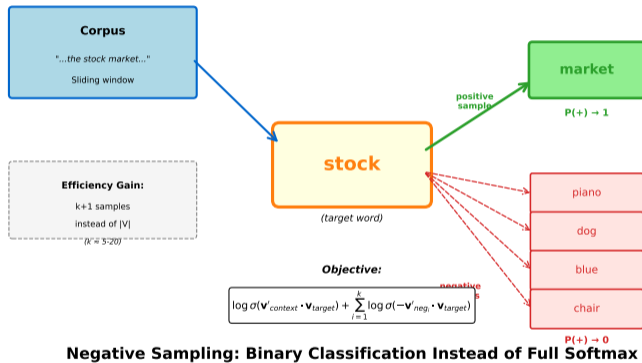
How Embeddings Are Learned: Skip-gram



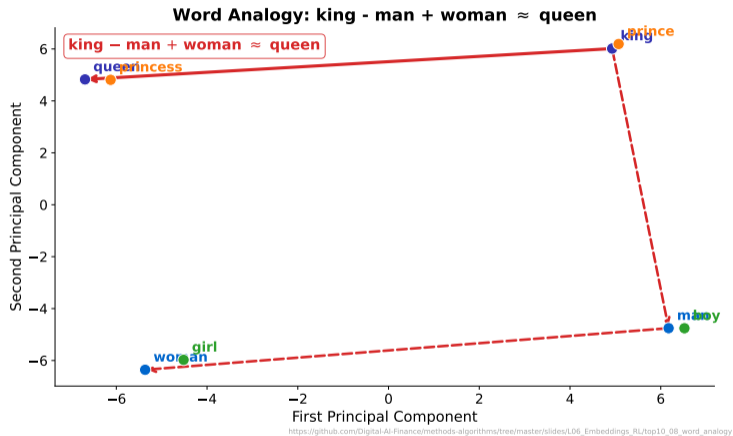
Skip-gram Architecture: Predict Context from Target

Skip-gram predicts context words from a center word, learning useful representations as a byproduct

Efficient Training: Negative Sampling

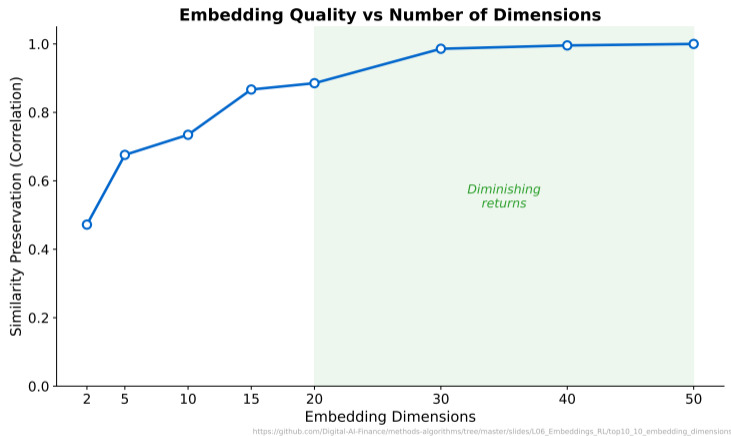


Negative sampling avoids computing the full softmax by contrasting true pairs with random negatives



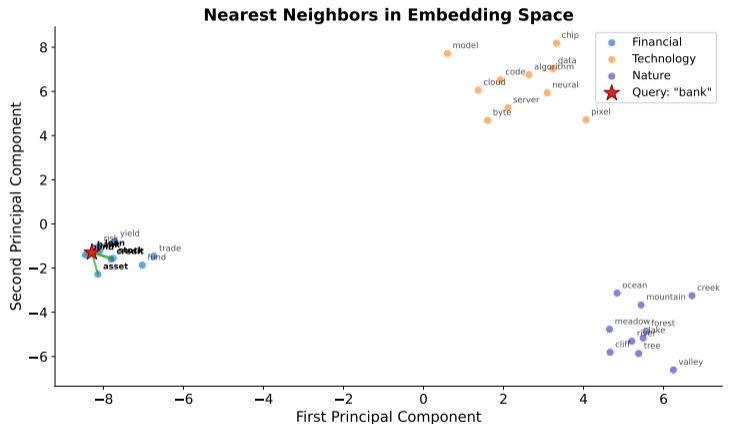
Vector arithmetic captures semantic relationships: the gender direction is consistent across word pairs

How Many Dimensions Do We Need?



Similarity structure is well-preserved beyond 20 dimensions; additional dimensions yield diminishing returns

Finding Similar Words in Embedding Space



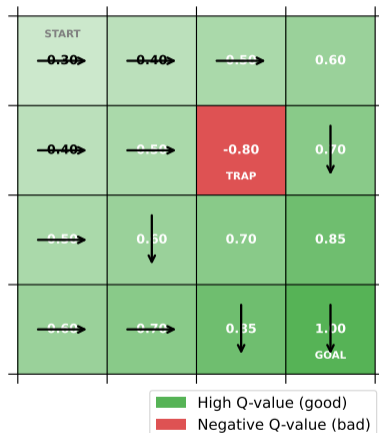
https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL_top10_13_embedding_neighbors

Nearest-neighbor queries retrieve semantically related words—the foundation of many NLP applications

- Embeddings capture **meaning** as geometric structure in vector space
- Reinforcement learning captures **behavior** as policies learned from reward signals
- Both use **learned representations** to generalize beyond training examples

Embeddings encode what words mean; RL learns what actions to take—both are fundamental to modern AI

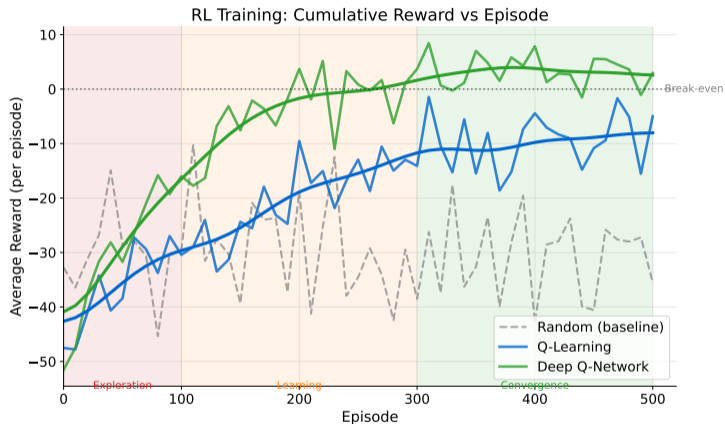
Q-Learning: Grid World with Learned Q-Values



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL04_q_learning_grid

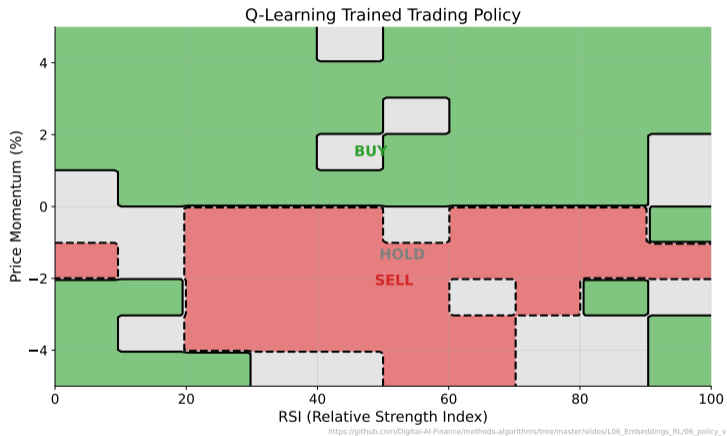
Q-values encode the expected future reward for each state-action pair in the environment

Training Progress: Reward Over Episodes



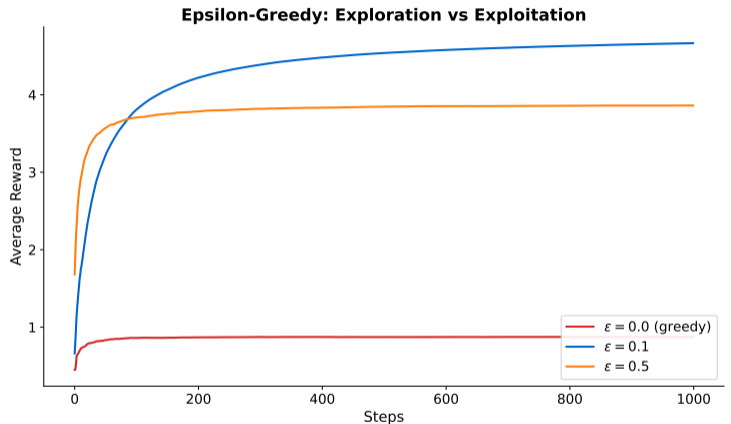
Monitoring episode reward over time reveals whether the agent is learning, plateauing, or diverging

The Learned Policy: What to Do in Each State



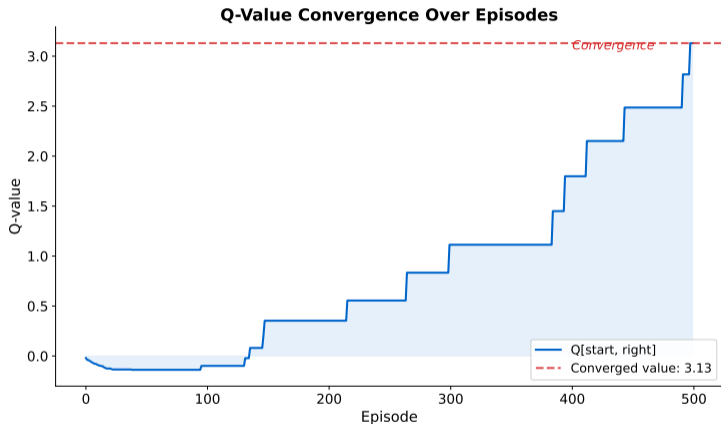
A policy maps each state to the best action—visualizing it reveals whether the agent has learned sensible behavior

Exploration vs Exploitation: The Epsilon Tradeoff



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/top10_09_epsilon_greedy

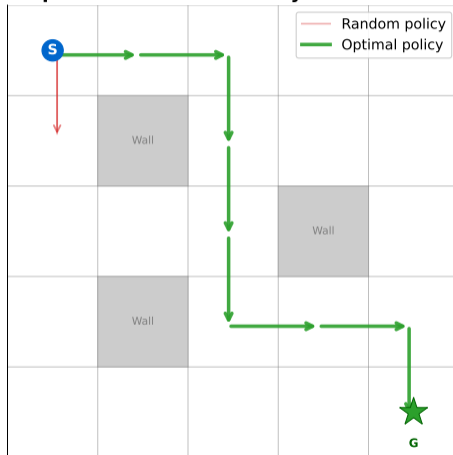
Too little exploration (greedy) gets stuck on suboptimal arms; too much exploration wastes reward on bad arms



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/top10_14_qvalue_convergence

With sufficient exploration and a decaying learning rate, Q-values provably converge to optimal values

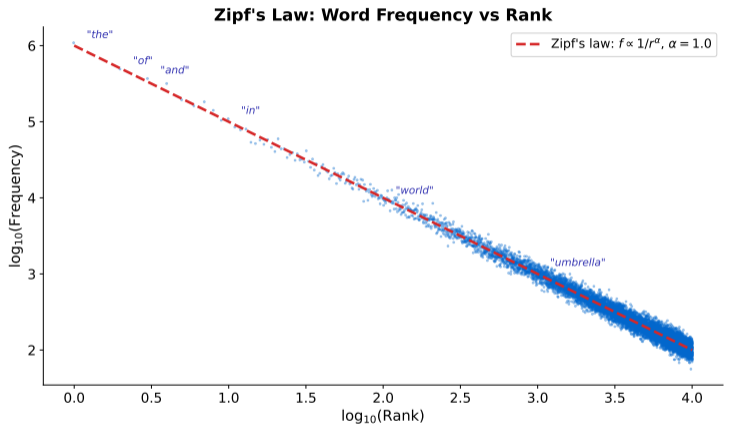
Optimal vs Random Policy in Gridworld



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/top10_19_gridworld_trajectory

After training, the learned policy finds the shortest path while random actions meander through the grid

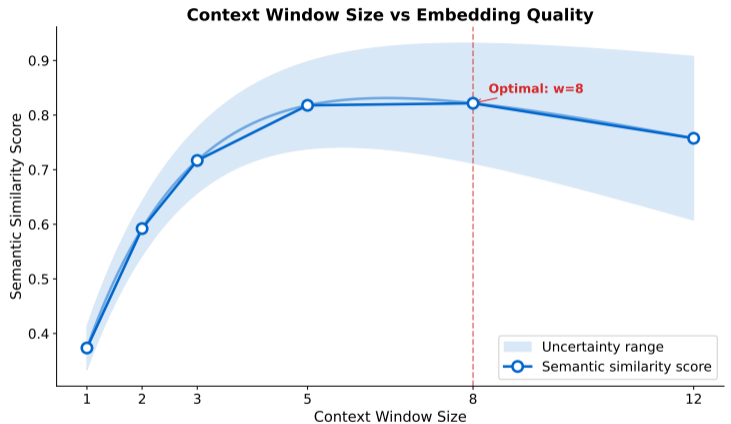
Zipf's Law: The Distribution of Words



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RU/top10_11_word_frequency_rank

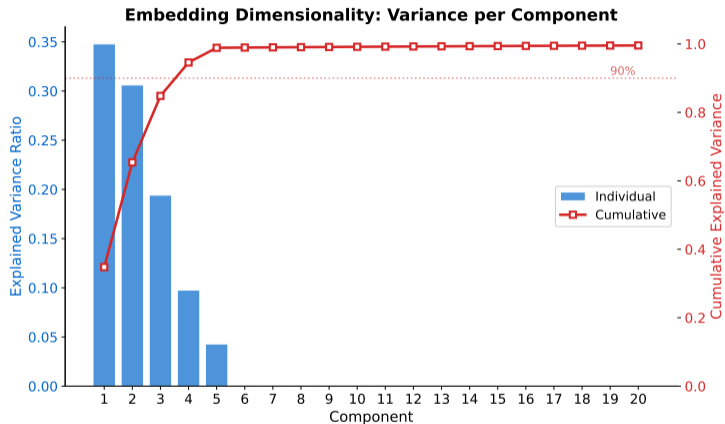
Zipf's law explains why embedding models must handle both extremely common and rare words

Context Window Size and Embedding Quality



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/top10_12_context_window

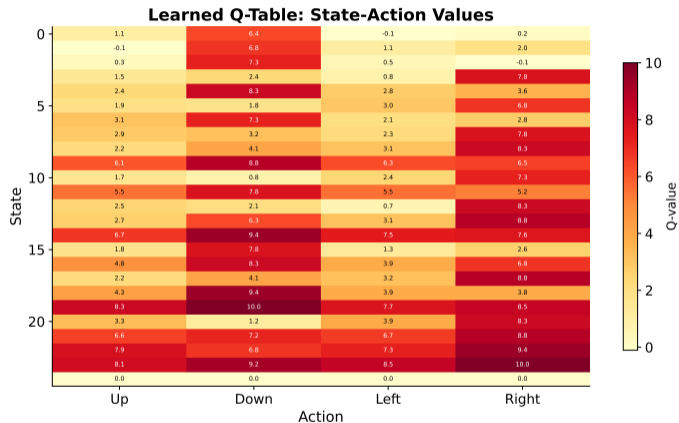
Small windows capture syntax; larger windows capture semantics—but too large introduces noise



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL_top10_18_embedding_pca_variance

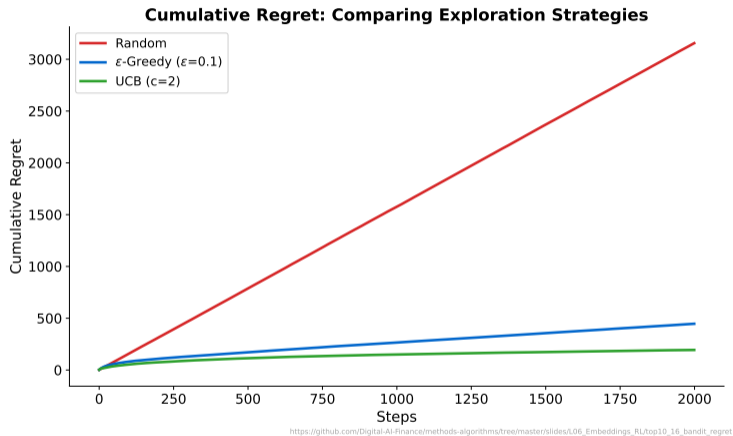
A few principal components capture most variance, confirming embeddings have low intrinsic dimensionality

The Full Q-Table: State-Action Values

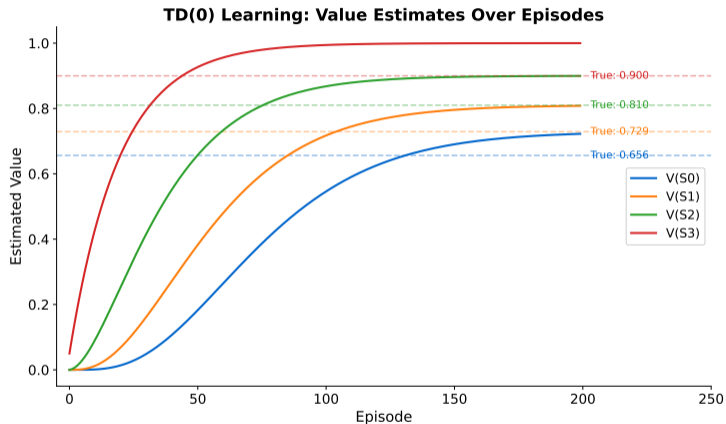


https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/top10_15_state_action_heatmap

States near the goal have higher Q-values; the optimal action per state has the highest value in its row

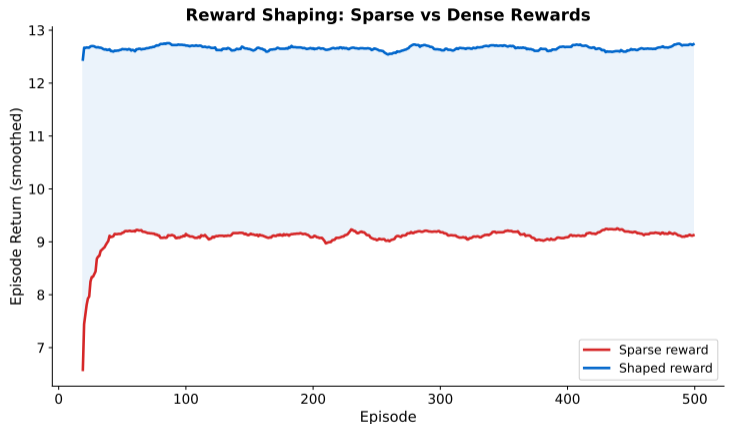


UCB achieves sublinear regret by balancing exploration and exploitation using confidence bounds



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL_top10_17_td_learning_update

TD(0) bootstraps value estimates from successor states, converging to true values without full episodes



https://github.com/Digital-AI-Finance/methods-algorithms/tree/master/slides/L06_Embeddings_RL/top10_20_reward_shaping

Shaped rewards provide learning signal throughout the episode, accelerating convergence over sparse rewards

20 Charts You Should Know

1. Embedding Space
2. Similarity Heatmap
3. Skip-gram Architecture
4. Negative Sampling
5. Word Analogy
6. Embedding Dimensions
7. Nearest Neighbors
8. Zipf's Law
9. Context Window
10. Embedding PCA Variance
11. Q-Learning Grid
12. Reward Curves
13. Policy Visualization
14. Epsilon-Greedy
15. Q-Value Convergence
16. Gridworld Trajectory
17. Q-Table Heatmap
18. Bandit Regret
19. TD Learning
20. Reward Shaping

These 20 visualizations cover embeddings, reinforcement learning foundations, and advanced analysis topics



XKCD #1838 by Randall Munroe (CC BY-NC 2.5)