CS 5/7322 - Programming Homework 2
Examining Contextualized Word Embeddings for BERT

Author: Ria Mukherji
Date: Nov 5th 2025

NOTE: ALL **sentences** are in this report as requested. Find them in the **appendix** at the end (as not to crowd the main body).

EXTRA CREDIT CODE TURNED IN AS SEPARATE FILE. Analysis done here.

Part 1: Comparing BERT Vectors for Same Word + Same Meaning

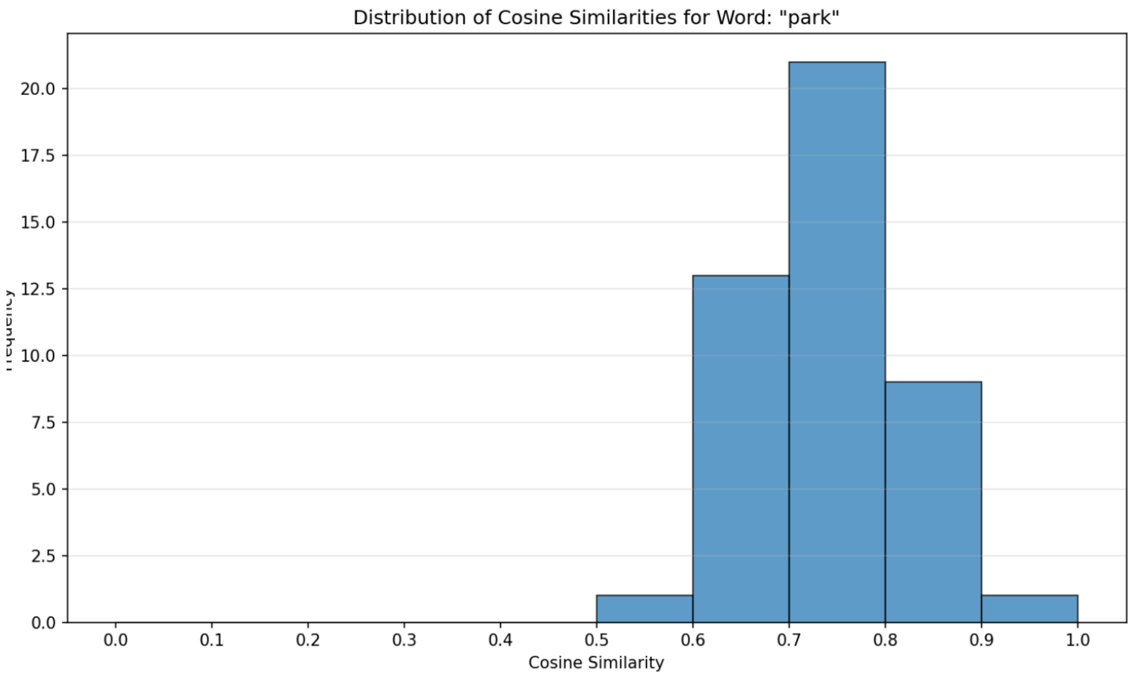
In this part, I examined whether BERT generated consistent embeddings for the same word when used with the same meaning across different sentences. I selected 8 words with distinct meanings and created 10 sentences for each word, ensuring the word maintains the same meaning in all sentences.

Individual Word Results

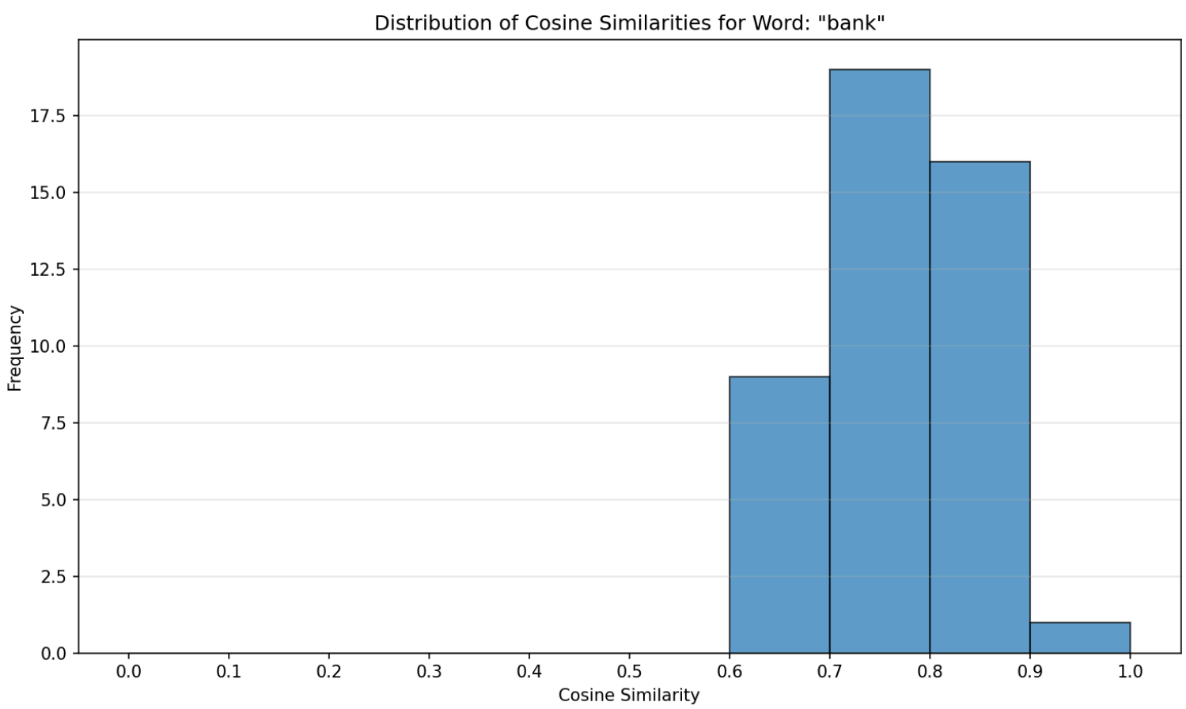
Word	Mean Similarity	Std Deviation	Pairs
PARK	0.7448	0.0774	45
BANK	0.7743	0.0721	45
BAT	0.6595	0.1248	45
SPRING	0.7310	0.0790	45
RING	0.6663	0.1813	45
TRAIN	0.6497	0.1201	45
MOUSE	0.7650	0.0734	45
CRANE	0.7774	0.0641	45

Individual Word Histograms

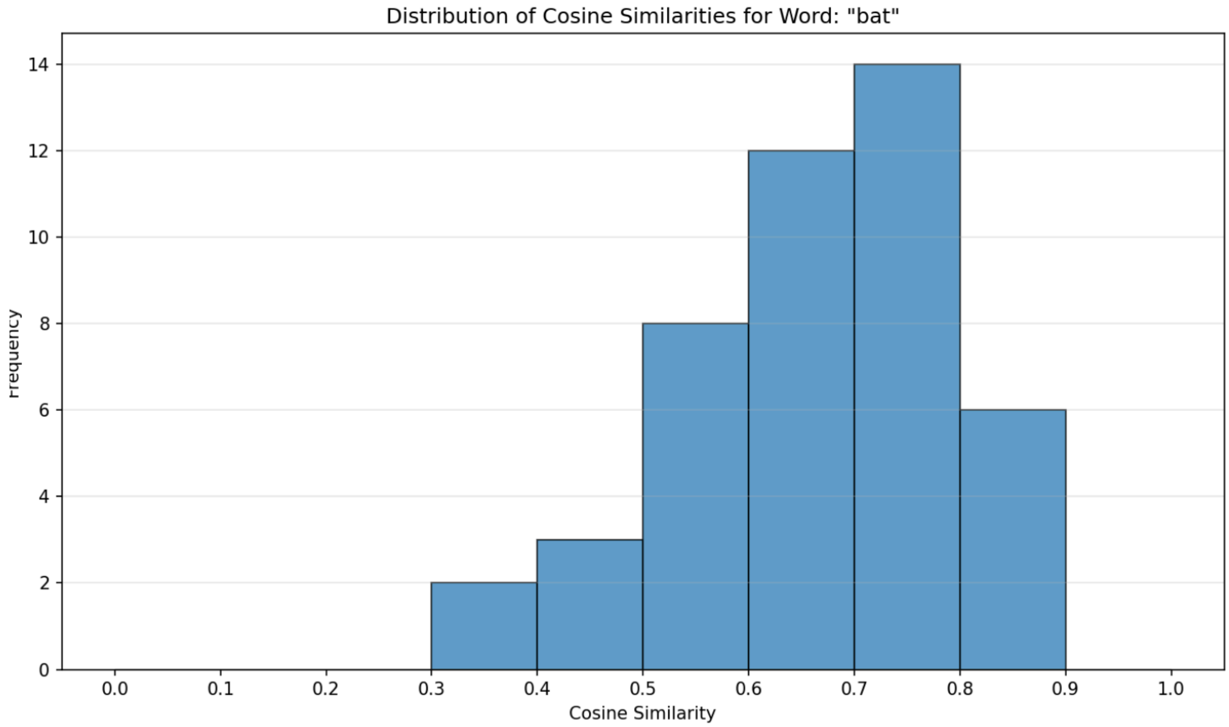
Word: PARK



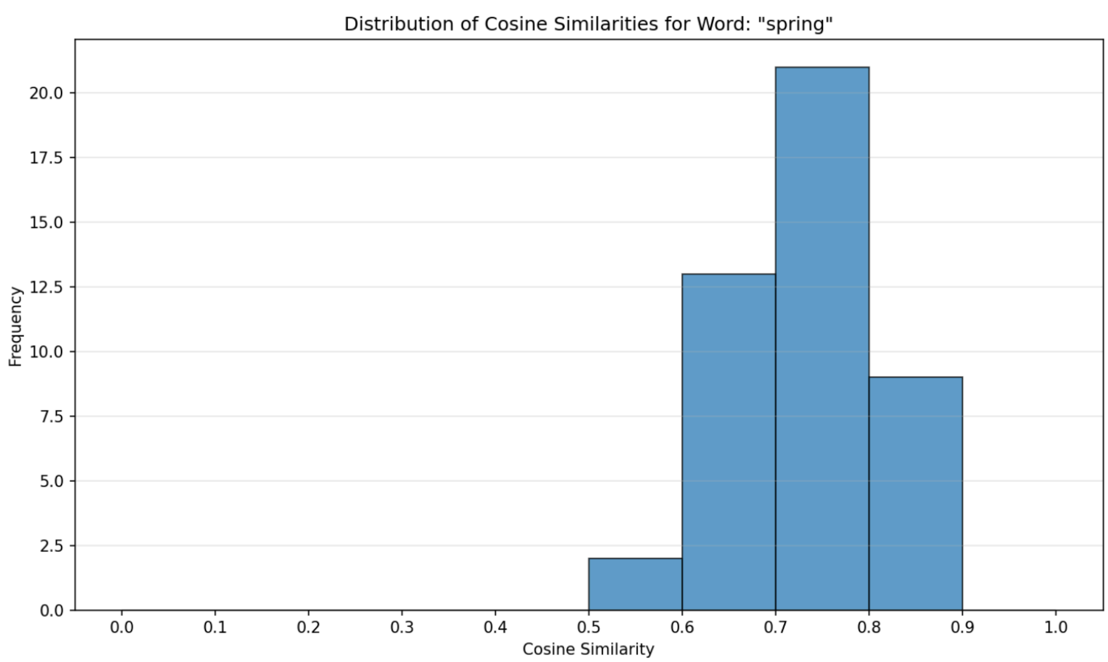
Word: BANK



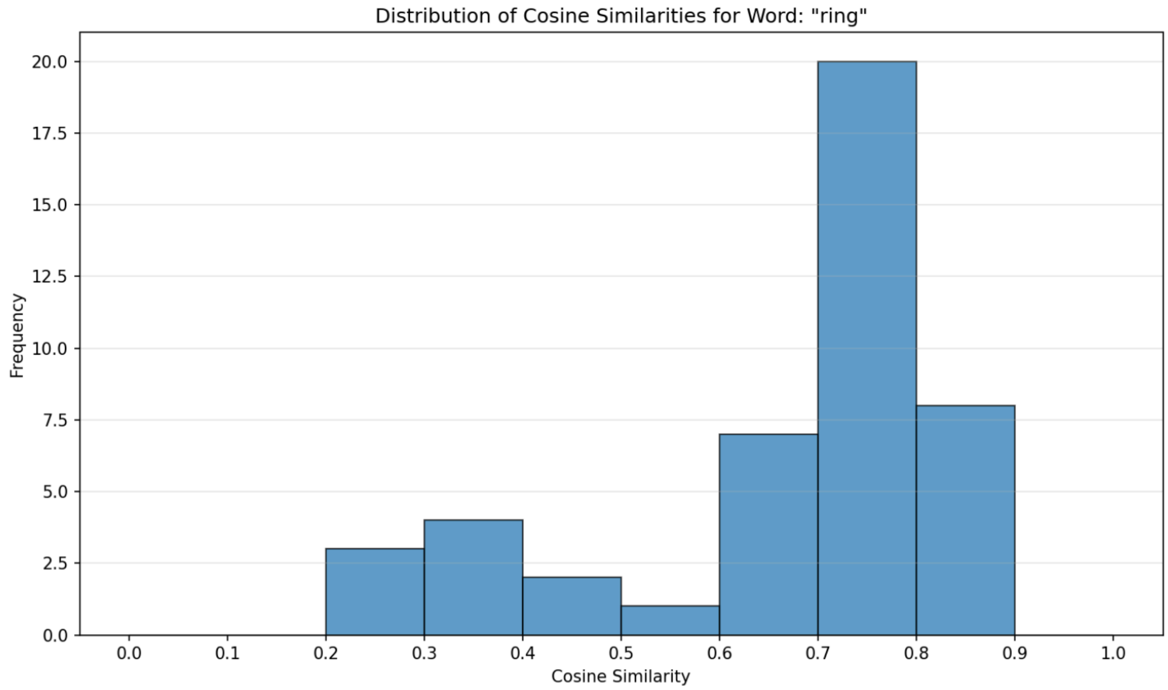
Word: BAT



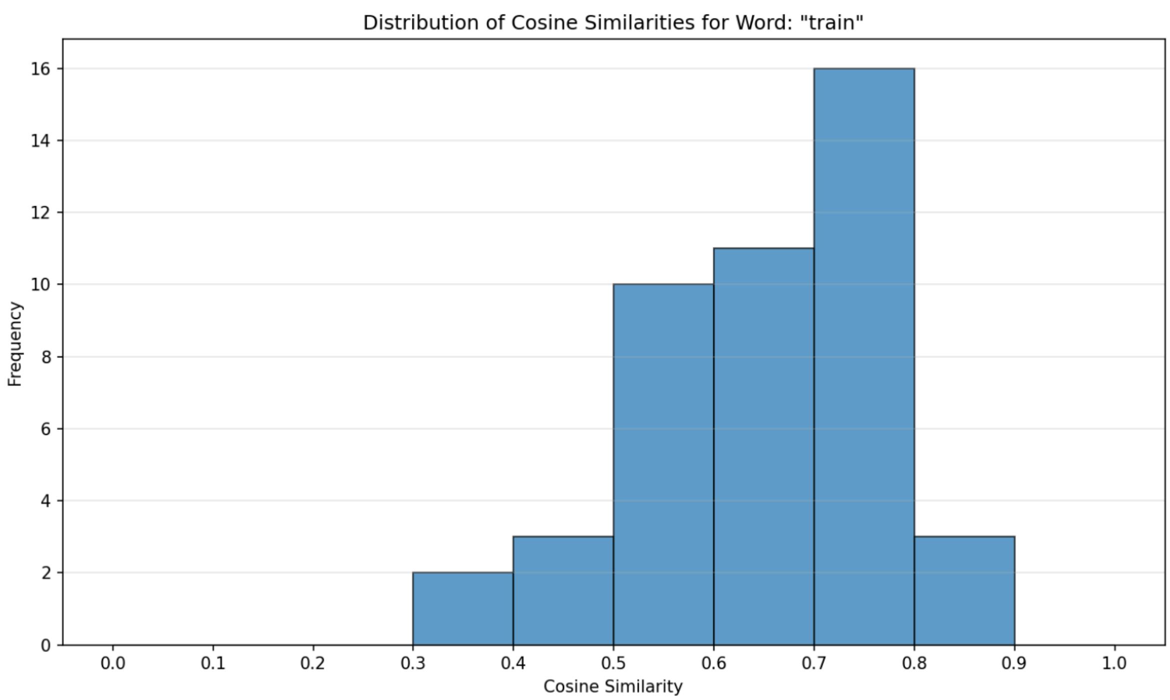
Word: SPRING



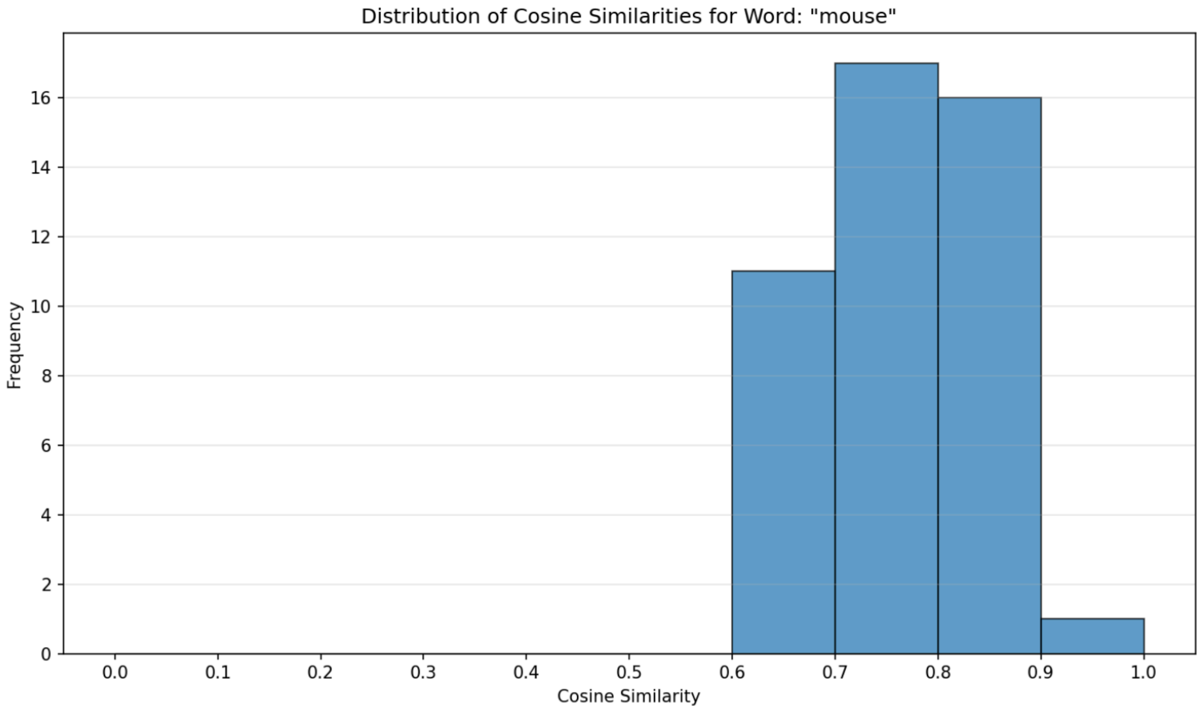
Word: RING



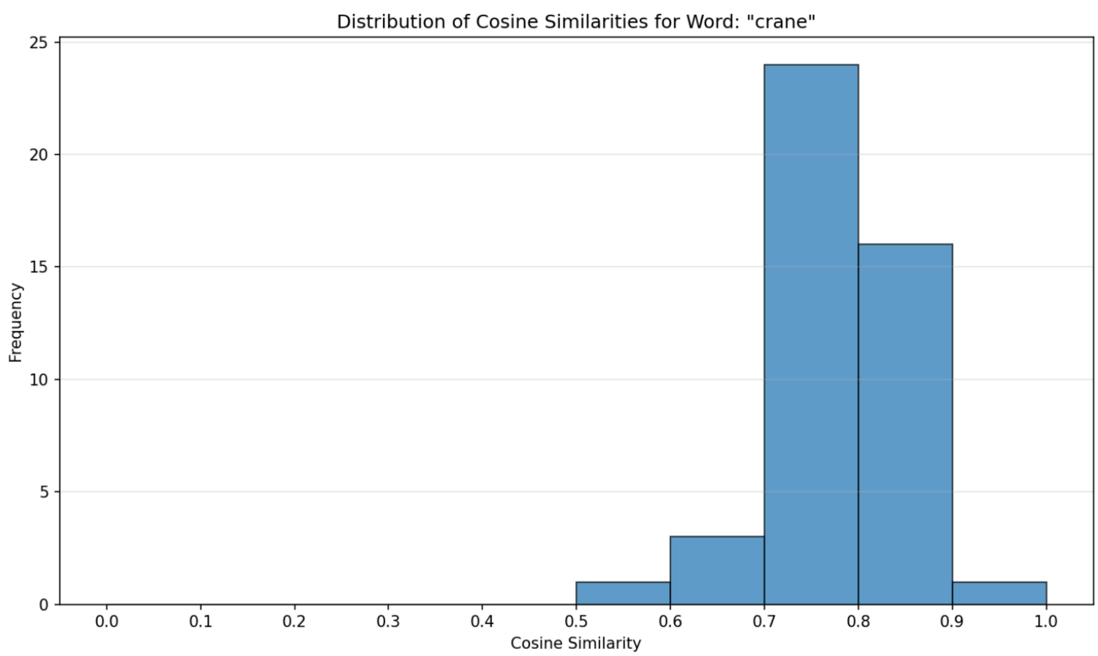
Word: TRAIN



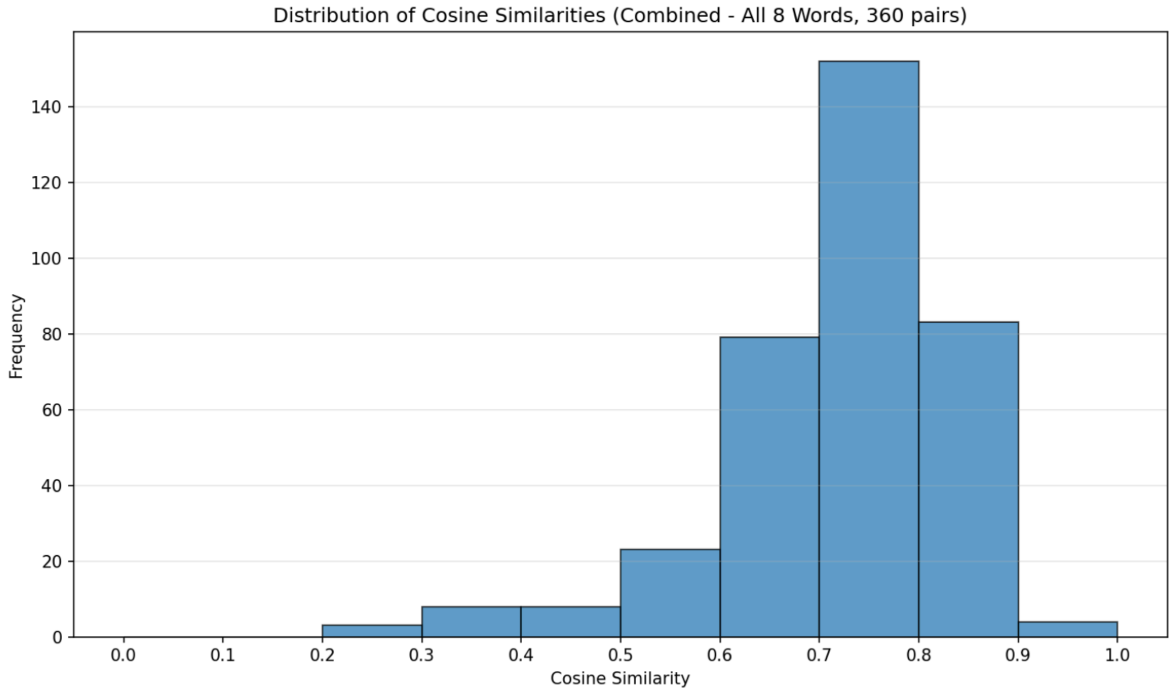
Word: MOUSE



Word: CRANE



Combined Histogram (All 8 Words, 360 pairs)



Overall Statistics

Total pairs: 360
Overall mean cosine similarity: 0.7210
Overall standard deviation: 0.1174

Analysis

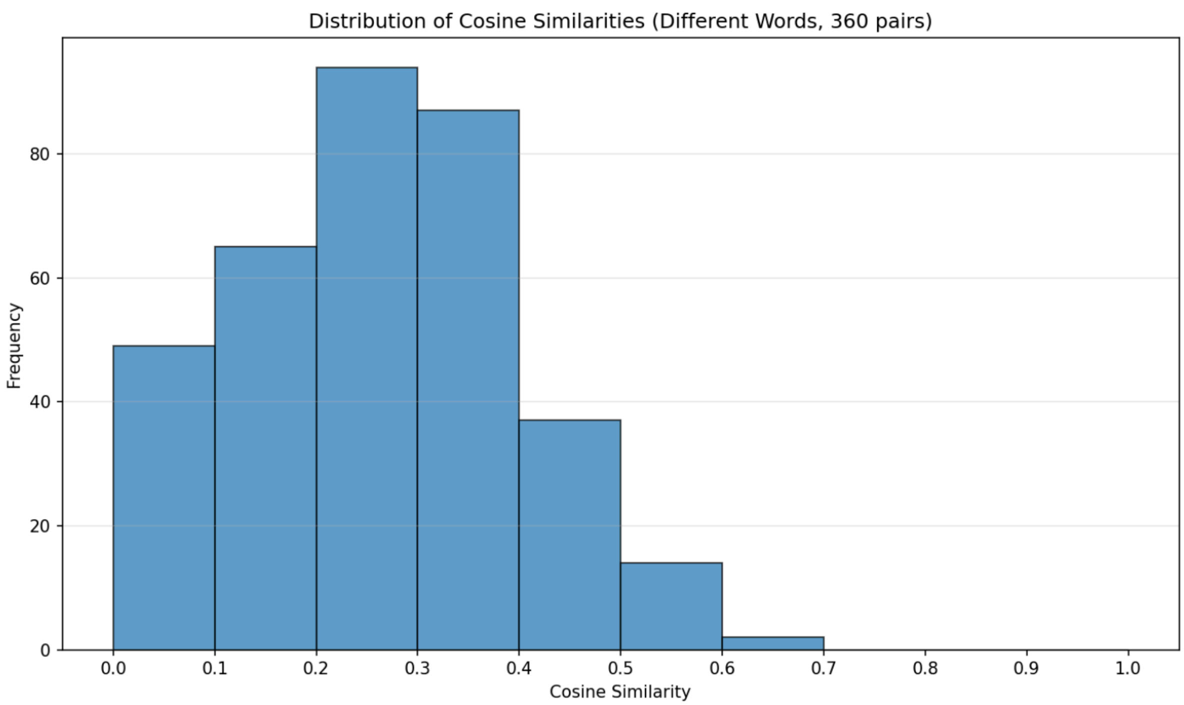
Based on the results from Part 1, BERT demonstrates a strong ability to generate consistent embeddings for the same word when used with the same meaning across different contexts. The overall mean cosine similarity of 0.7210 (with standard deviation of 0.1174) across all 360 pairs indicates that vectors for the same word are highly similar. This suggests that BERT's contextualized embeddings maintain semantic consistency for words used with identical meanings, even when the surrounding context varies. The relatively low standard deviation further indicates that this consistency is reliable across different words and sentence structures.

Part 2: Comparing BERT Vectors for Different Words

In this part, I compare vectors from different words to examine whether BERT distinguishes between words with different meanings. I randomly selected 360 pairs of vectors where each vector in the pair comes from a different word.

Statistics

Total pairs analyzed: 360
Mean cosine similarity: 0.2534
Standard deviation: 0.1444



Comparison with Part 1

Comparing the histograms from Part 1 and Part 2 shows a clear distinction in how BERT handles same-word versus different-word comparisons. In Part 1, the mean cosine similarity was 0.7210, indicating high similarity for the same word across contexts. In contrast, Part 2 shows a mean similarity of only 0.2534 for different words. This is very low and indicates that the words are not similar which makes sense. This significant difference (approximately 0.47 points) demonstrates that BERT effectively distinguishes between different words. The histogram for Part 2 shows a distribution shifted toward lower similarity values compared to Part 1, further confirming that BERT embeddings capture semantic differences between distinct words. The slightly higher standard deviation in Part 2 (0.1444 vs 0.1174) suggests more variability when comparing unrelated words, which is expected given the diverse semantic relationships between different word pairs.

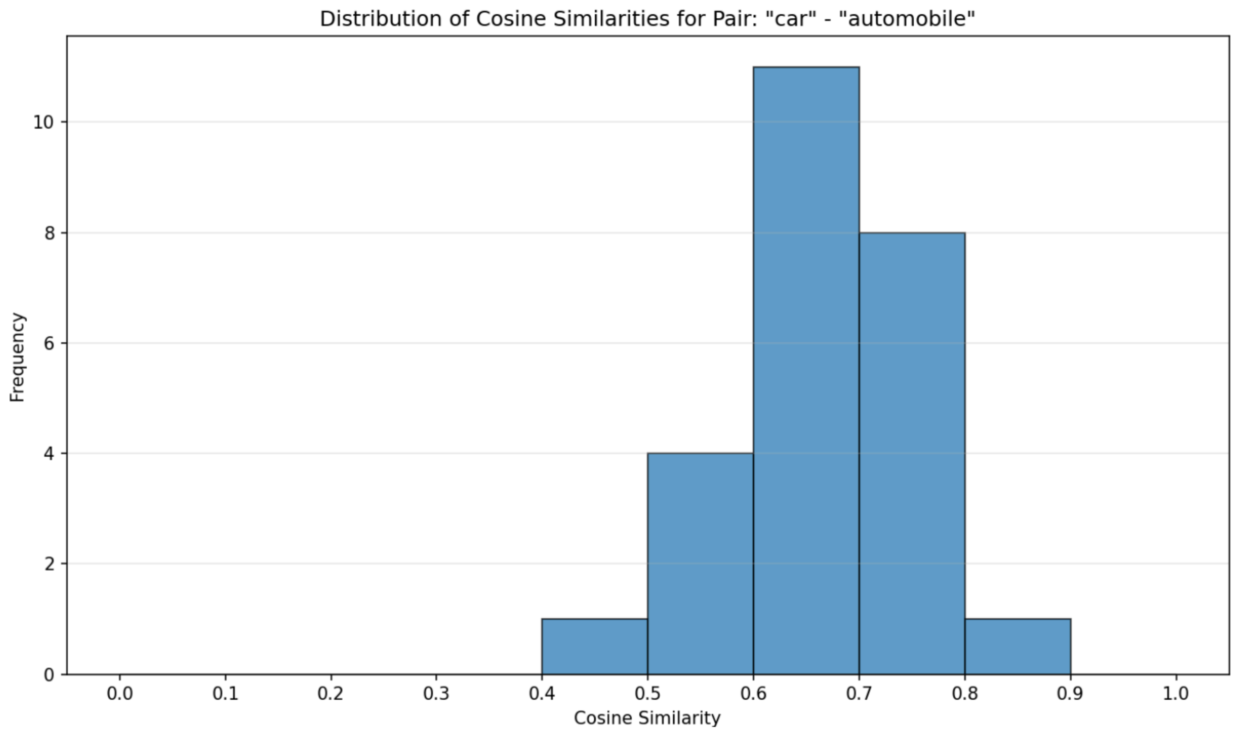
Part 3: Comparing BERT Vectors for Different Words with Same Meaning

In this part, I examine whether BERT maps different words with similar meanings to similar vectors. I selected 8 pairs of words from WordNet synsets, where each pair consists of two words with the same meaning. For each word pair, I created 5 sentences per word (10 total sentences per pair), ensuring the sentences use different contexts and are not simple word substitutions.

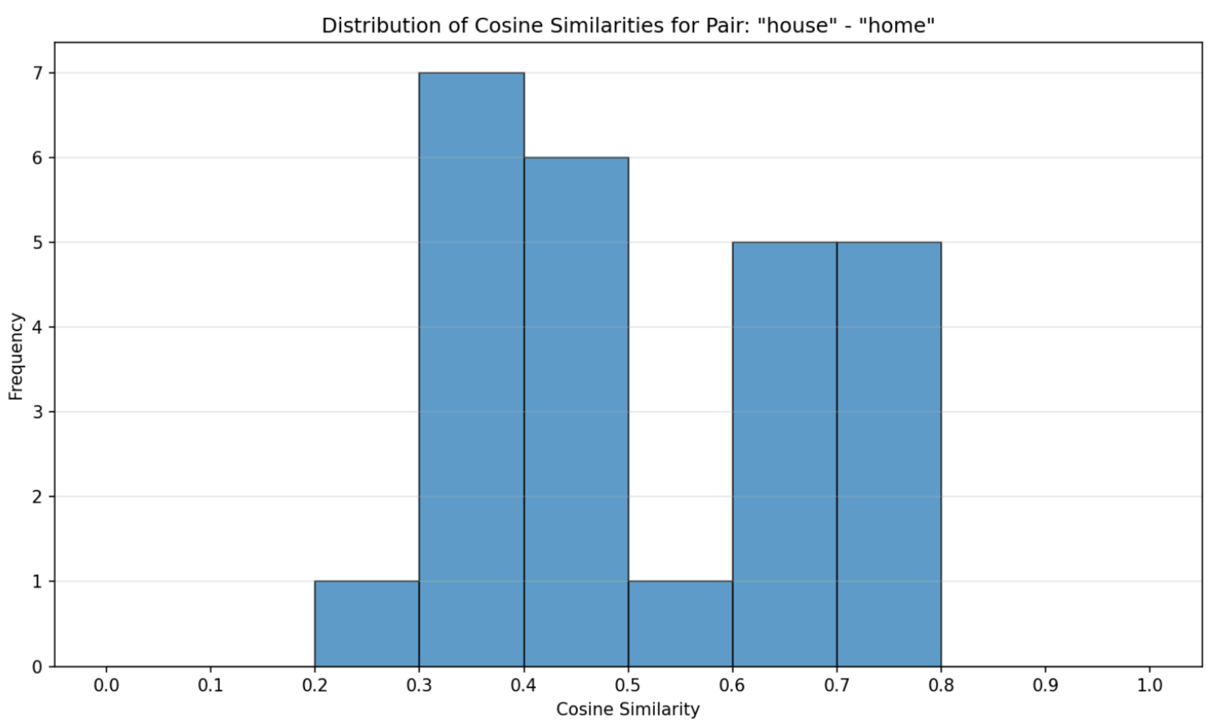
Word Pair	Mean Similarity	Std Deviation	Pairs
car-automobile	0.6742	0.0772	25
house-home	0.5237	0.1600	25
big-large	0.4727	0.1223	25
happy-glad	0.5779	0.1191	25
fast-quick	0.3264	0.0918	25
beautiful-pretty	0.5609	0.1108	25
small-little	0.4896	0.1717	25
smart-intelligent	0.4453	0.1541	25

Individual Word Pair Histograms

Pair: CAR - AUTOMOBILE

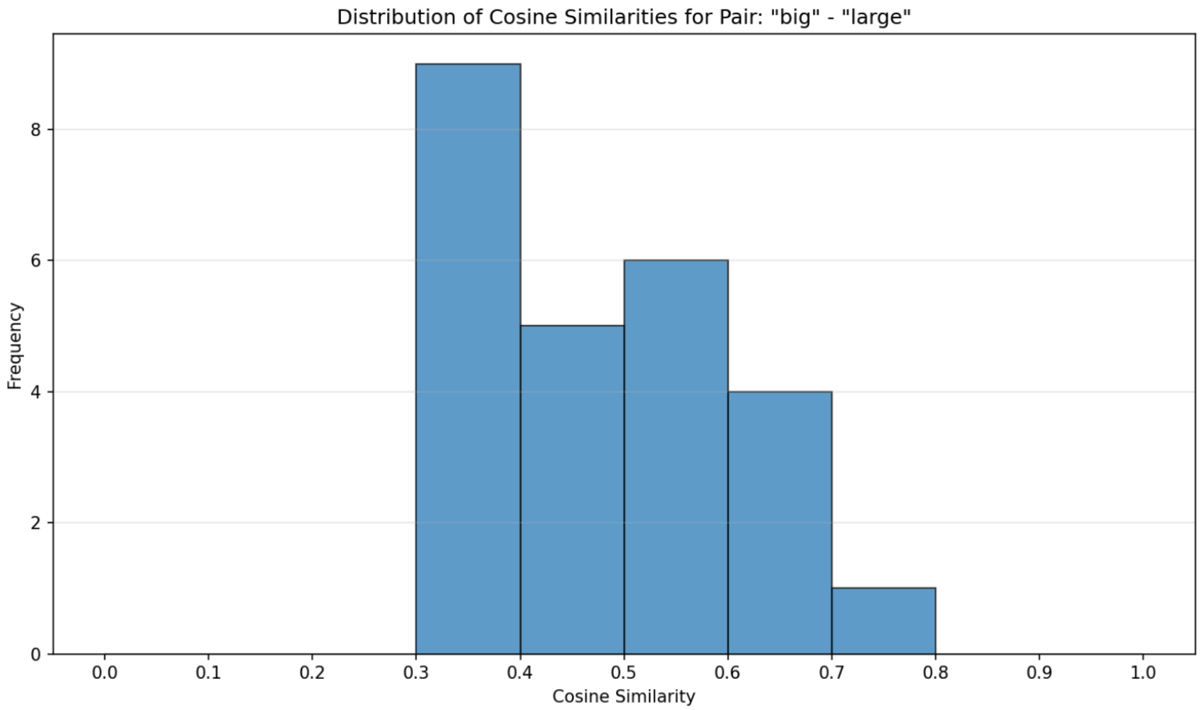


Pair: HOUSE - HOME

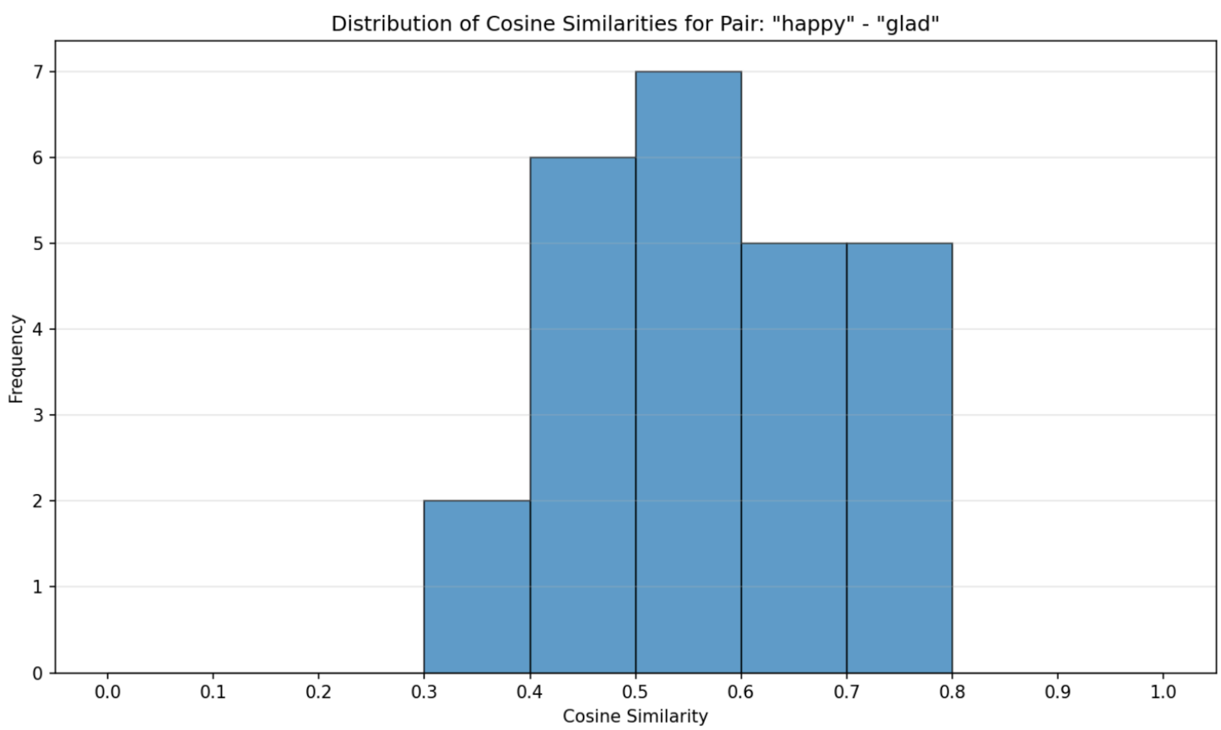


***Note: In the sentences house and home are truly used interchangeably though I could acknowledge that in some contexts they may have slightly different meanings.

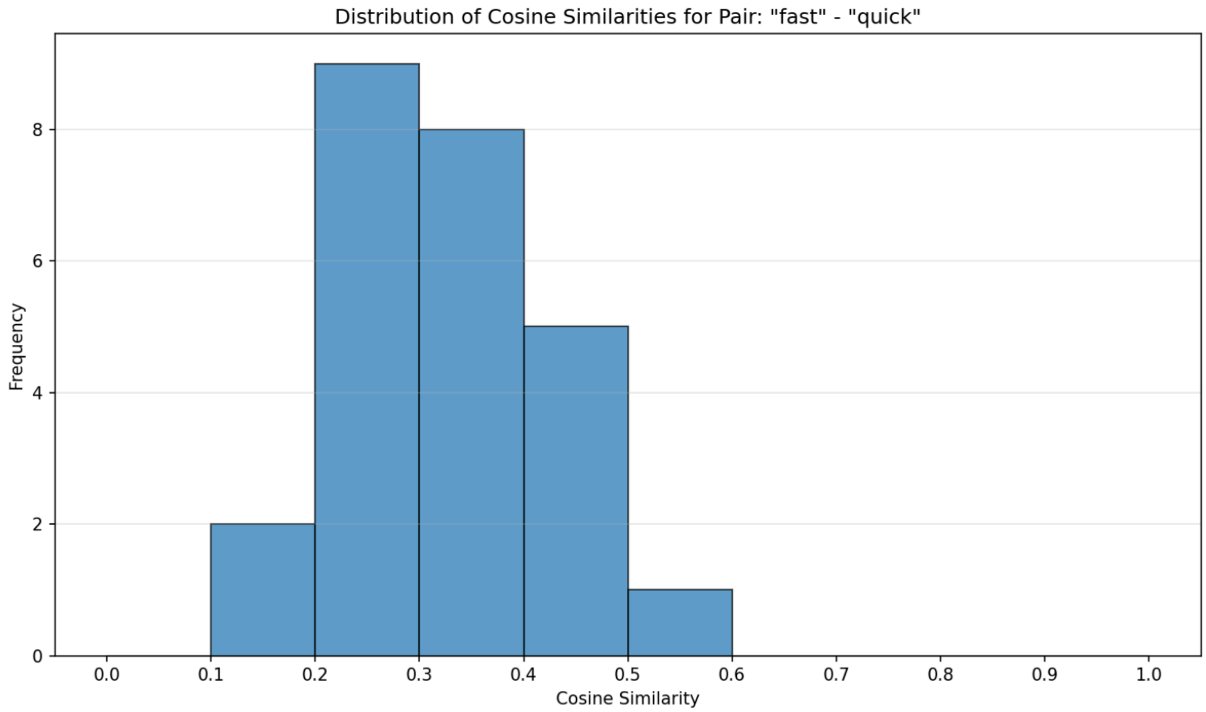
Pair: BIG - LARGE



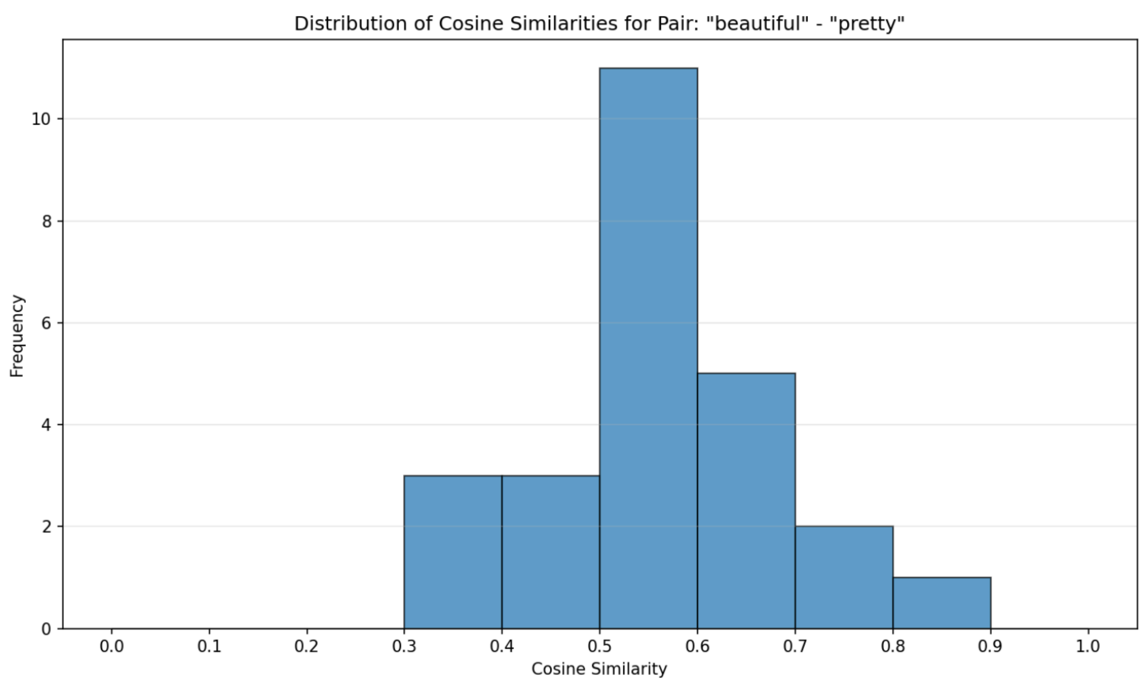
Pair: HAPPY - GLAD



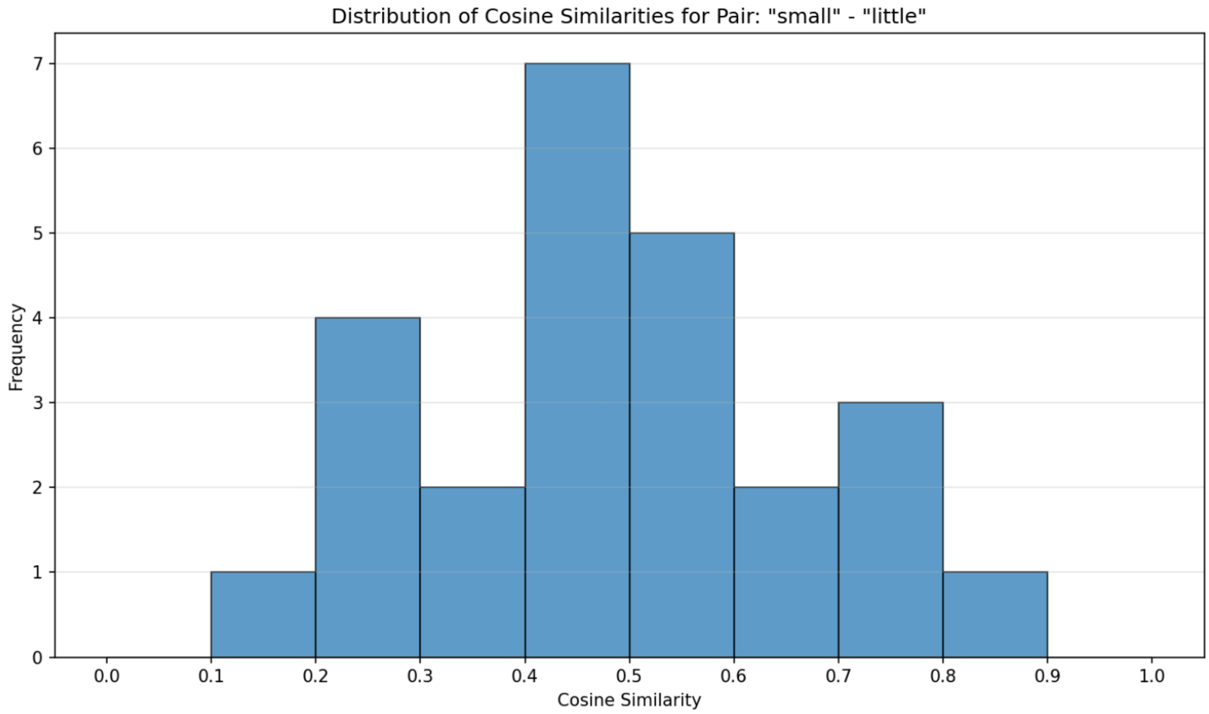
Pair: FAST - QUICK



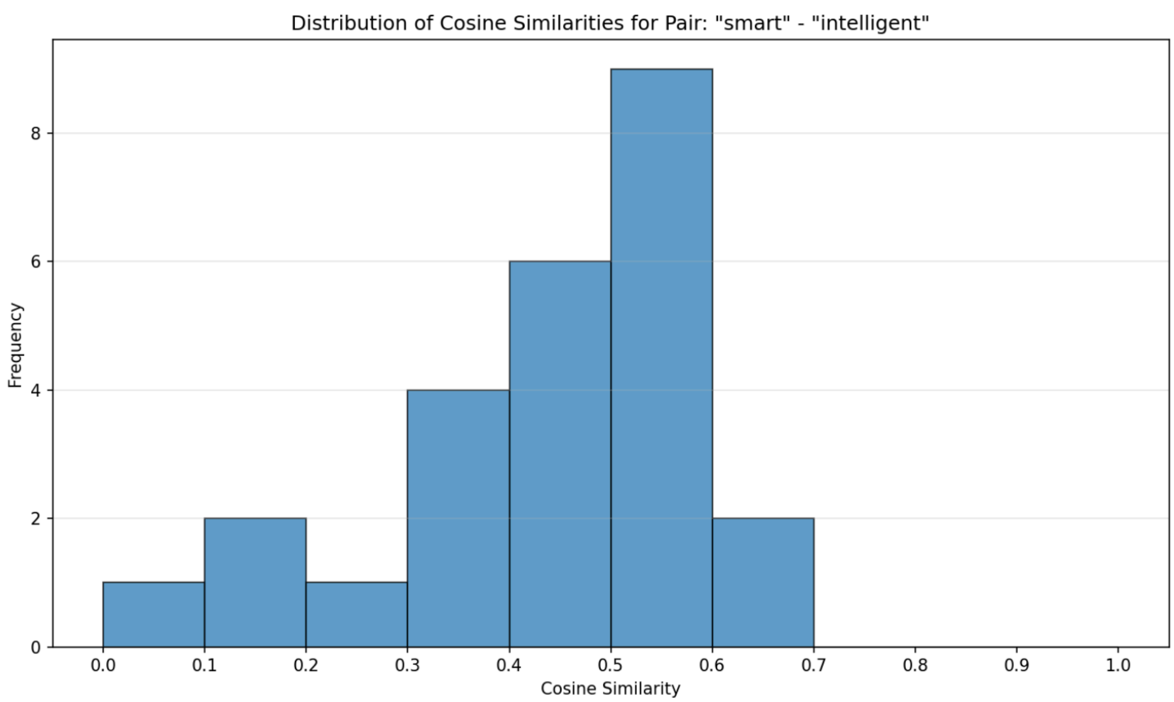
Pair: BEAUTIFUL - PRETTY



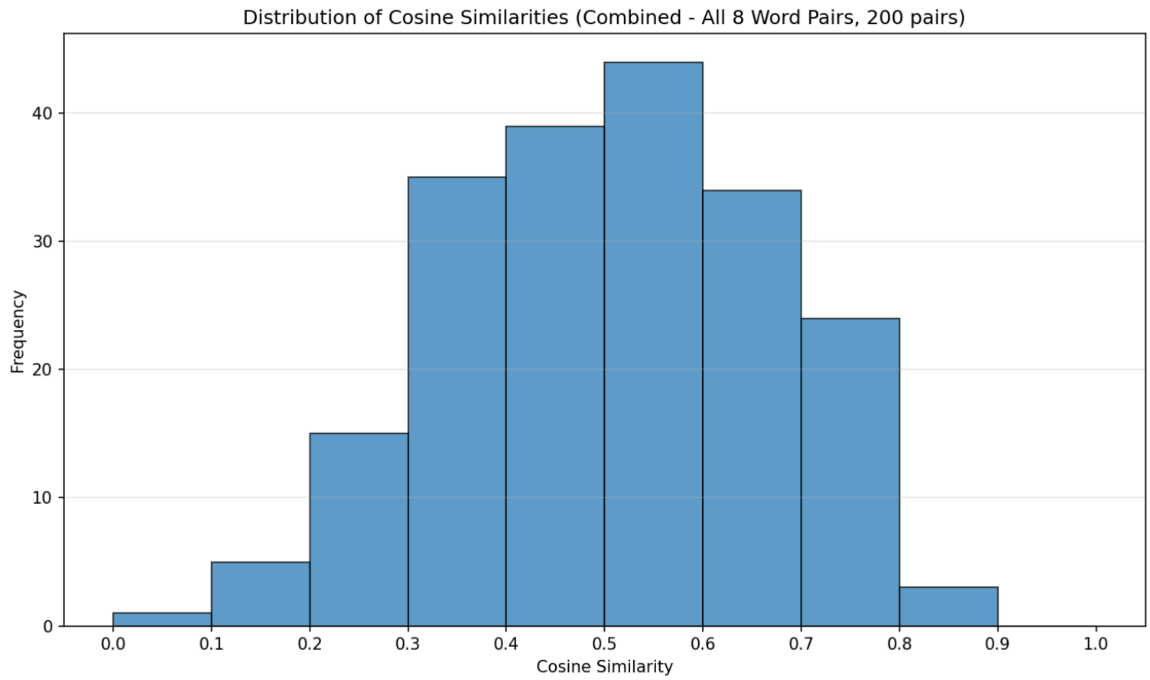
Pair: SMALL - LITTLE



Pair: SMART - INTELLIGENT



Combined Histogram (All 8 Word Pairs, 200 pairs)



Overall Statistics

Total pairs: 200
Overall mean cosine similarity: 0.5088
Overall standard deviation: 0.1615

Analysis

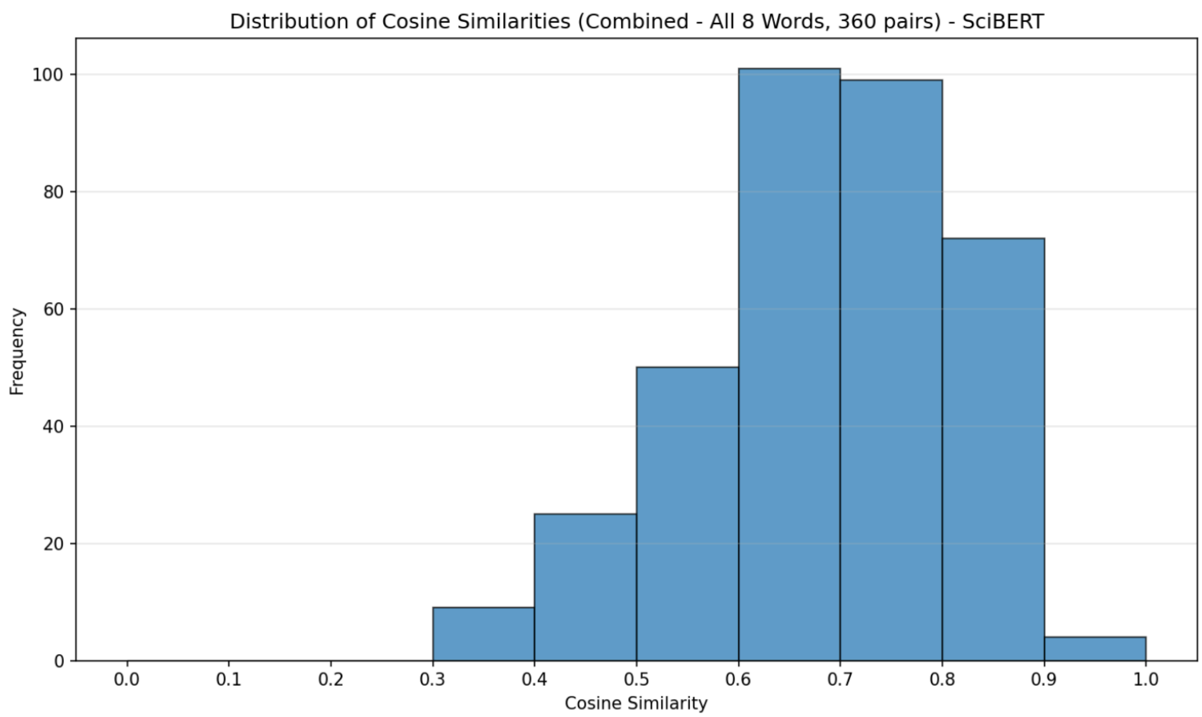
The results from Part 3 show that BERT does capture semantic similarity between synonyms, though to a lesser extent
than for the same word. The overall mean similarity of 0.5088 for different words with the same meaning is intermediate between Part 1 (same word: 0.7210) and Part 2 (different words: 0.2534). This suggests that while BERT recognizes semantic relationships between synonyms, it maintains some distinction between different lexical items. The variability in similarity scores across different word pairs (ranging from 0.3264 for 'fast-quick' to 0.6742 for 'car-automobile') indicates that the degree of semantic similarity captured depends on the specific word pair and their usage contexts.

Extra Credit: SciBERT Analysis

For the extra credit portion, I repeated Part 1 and Part 2 using SciBERT, a domain-specific BERT model trained on scientific literature. This allows us to compare how domain-specific training affects the embedding quality.

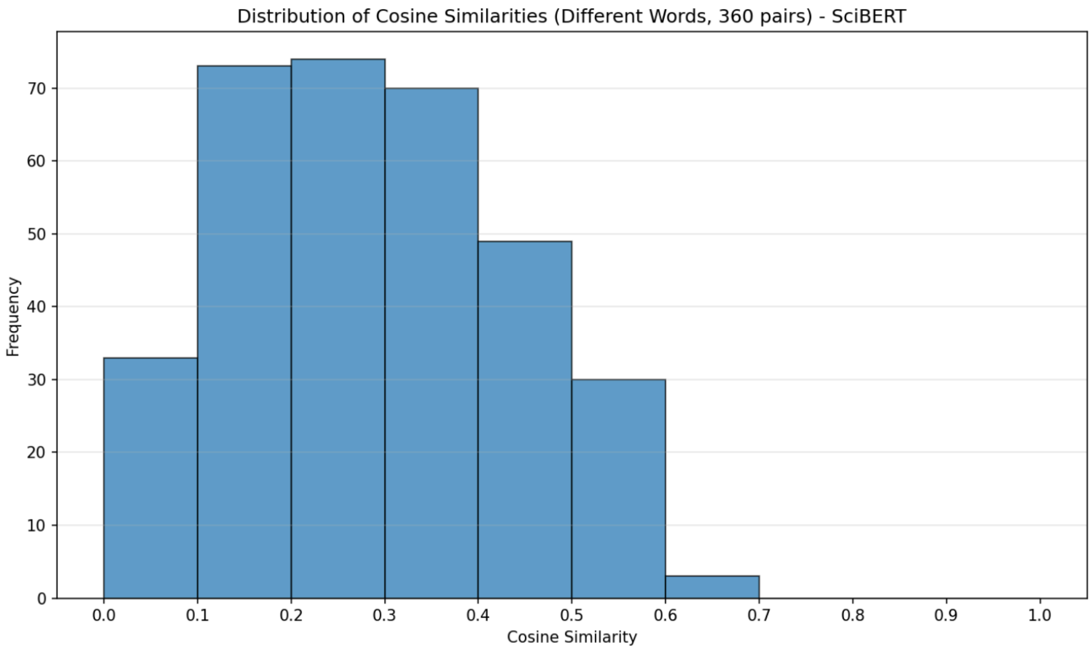
Model	Part 1 Mean	Part 1 Std	Part 2 Mean	Part 2 Std
BERT-base	0.7210	0.1174	0.2534	0.1444
SciBERT	0.6880	0.1276	0.2607	0.1704

SciBERT Part 1 Combined Histogram



SciBERT Part 2 HistogramComparison Analysis

Comparing SciBERT with BERT-base reveals interesting differences. For Part 1 (same word, same meaning), SciBERT shows a slightly lower mean similarity (0.6880 vs 0.7210) and higher standard deviation (0.1276 vs 0.1174), suggesting that domain-specific training may introduce more variability in embeddings for general vocabulary. For Part 2 (different words), the results are similar (0.2607 vs 0.2534), indicating that both models maintain similar ability to distinguish between different words. The higher standard deviation in SciBERT's Part 2 results (0.1704 vs 0.1444) suggests more variability when comparing unrelated words, possibly due to the scientific domain training affecting how general vocabulary is represented.



Comparison Analysis

Comparing SciBERT with BERT-base reveals interesting differences. For Part 1 (same word, same meaning), SciBERT shows a slightly lower mean similarity (0.6880 vs 0.7210) and higher standard deviation (0.1276 vs 0.1174), suggesting that domain-specific training may introduce more variability in embeddings for general vocabulary. For Part 2 (different words), the results are similar (0.2607 vs 0.2534), indicating that both models maintain similar ability to distinguish between different words. The higher standard deviation in SciBERT's Part 2 results (0.1704 vs 0.1444) suggests more variability when comparing unrelated words, possibly due to the scientific domain training affecting how general vocabulary is represented.

Appendix A — Part 1 Sentences (Same Word, Same Meaning)

Word: park

1. The park next to my house offers a nice walk.
2. The residents are trying to build a park for the kids to play.
3. We had a picnic in the central park last weekend.
4. The city council approved funding for a new park downtown.
5. Children love playing on the swings at the park.
6. The park has beautiful flowers and tall trees.
7. Many people exercise in the park every morning.
8. The park closes at sunset for safety reasons.
9. There is a small pond in the middle of the park.
10. The community organized a cleanup day at the park.

Word: bank

1. I need to deposit money at the bank today.
2. The bank approved my loan application yesterday.
3. She works as a teller at the local bank.
4. The bank charges a monthly fee for this account.
5. I opened a savings account at the new bank.
6. The bank is closed on weekends and holidays.

7. He withdrew cash from the bank this morning.
8. The bank offers competitive interest rates.
9. I forgot my bank card at home today.
10. The bank has excellent customer service.

Word: bat

1. The bat flew out of the cave at dusk.
2. A bat can see in the dark using echolocation.
3. The bat hung upside down from the tree branch.
4. Scientists study the bat to understand its behavior.
5. The bat found insects to eat during the night.
6. A small bat was sleeping in the attic.
7. The bat uses sound waves to navigate.
8. I saw a bat circling the streetlight.
9. The bat colony lives in the old barn.
10. A fruit bat feeds on nectar and fruits.

Word: spring

1. Spring is my favorite season because of the flowers.
2. The weather is perfect during spring.
3. I love seeing the cherry blossoms in spring.
4. Spring brings warmer temperatures and longer days.
5. We plan our garden planting for early spring.

6. Spring cleaning is a tradition in our household.
7. The birds return from migration in spring.
8. Spring marks the end of winter.
9. I enjoy taking walks during spring mornings.
10. Spring festivals celebrate the new growing season.

Word: ring

1. She received a beautiful diamond ring for her birthday.
2. The engagement ring has a stunning sapphire.
3. He lost his class ring at the beach.
4. The wedding ring symbolizes eternal commitment.
5. She wears a gold ring on her left hand.
6. The ring was too tight and needed resizing.
7. I bought a silver ring from the jewelry store.
8. The antique ring belonged to her grandmother.
9. The ring sparkled in the sunlight.
10. He gave her a promise ring before leaving.

Word: train

1. The train arrived at the station on time.
2. I take the train to work every morning.
3. The train travels through mountains and valleys.
4. Passengers boarded the train for the long journey.

5. The train whistle could be heard from miles away.
6. The bullet train reaches speeds of 200 miles per hour.
7. We missed the last train and had to wait.
8. The train tracks cross through the city center.
9. The train conductor checked everyone's tickets.
10. The train derailed near the rural station.

Word: mouse

1. The mouse scurried across the kitchen floor.
2. A tiny mouse built a nest in the garden.
3. The cat chased the mouse around the house.
4. The mouse found cheese in the trap.
5. I saw a mouse running along the wall.
6. The mouse is nocturnal and active at night.
7. A field mouse lives in the grass.
8. The mouse squeaked when it was frightened.
9. The mouse has large ears and a long tail.
10. Scientists use the mouse for laboratory experiments.

Word: crane

1. The crane stood gracefully on one leg in the marsh.
2. A beautiful crane flew overhead during migration.
3. The crane is known for its elegant dancing courtship.

4. We spotted a rare crane at the wildlife sanctuary.
5. The crane has a long neck and pointed beak.
6. The Japanese crane is a symbol of good luck.
7. The crane waded through the shallow water.
8. The crane pair nested near the lake together.
9. The crane spread its wings to take flight.
10. Birdwatchers travel far to see the crane.

Appendix B — Part 3 Sentences (Synonym Pairs)

Pair: car — automobile

car:

1. The red car is very fast.
2. I need to buy a new car soon.
3. The car broke down on the highway.
4. She drives a sports car to work.
5. The car's engine makes a loud noise.

automobile:

1. This is a turbocharged automobile.
2. The automobile industry has grown rapidly.
3. He collects vintage automobile models.
4. The automobile manufacturer recalled many vehicles.

5. Electric automobile sales are increasing.

Pair: house — home
house:

1. The house has three bedrooms and two bathrooms.
2. We painted the house bright yellow last summer.
3. The house was built in 1950.
4. She owns a beautiful house by the lake.
5. The house needs major repairs.

home:

1. There is no place like a home.
2. The home was decorated with modern furniture.
3. She returned to her home after a long journey.
4. The new home has a spacious backyard.
5. They made their home in the countryside.

Pair: big — large
big:

1. That is a very big tree in the park.
2. The big dog ran across the field.
3. She has a big collection of books.
4. The big meeting is scheduled for tomorrow.
5. A big storm is approaching the city.

large:

1. The large building dominates the skyline.
2. He inherited a large fortune from his family.
3. The large crowd gathered at the stadium.
4. She wore a large hat to the party.
5. The large pizza could feed ten people.

Pair: happy — glad

happy:

1. I am very happy with the results.
2. The happy child played in the garden.
3. She felt happy after receiving good news.
4. They celebrated with happy smiles.
5. The happy couple got married last week.

glad:

1. I am glad you could make it to the party.
2. The glad news spread quickly through the town.
3. She was glad to see her old friend again.
4. He seemed glad about the opportunity.
5. We are glad the weather improved.