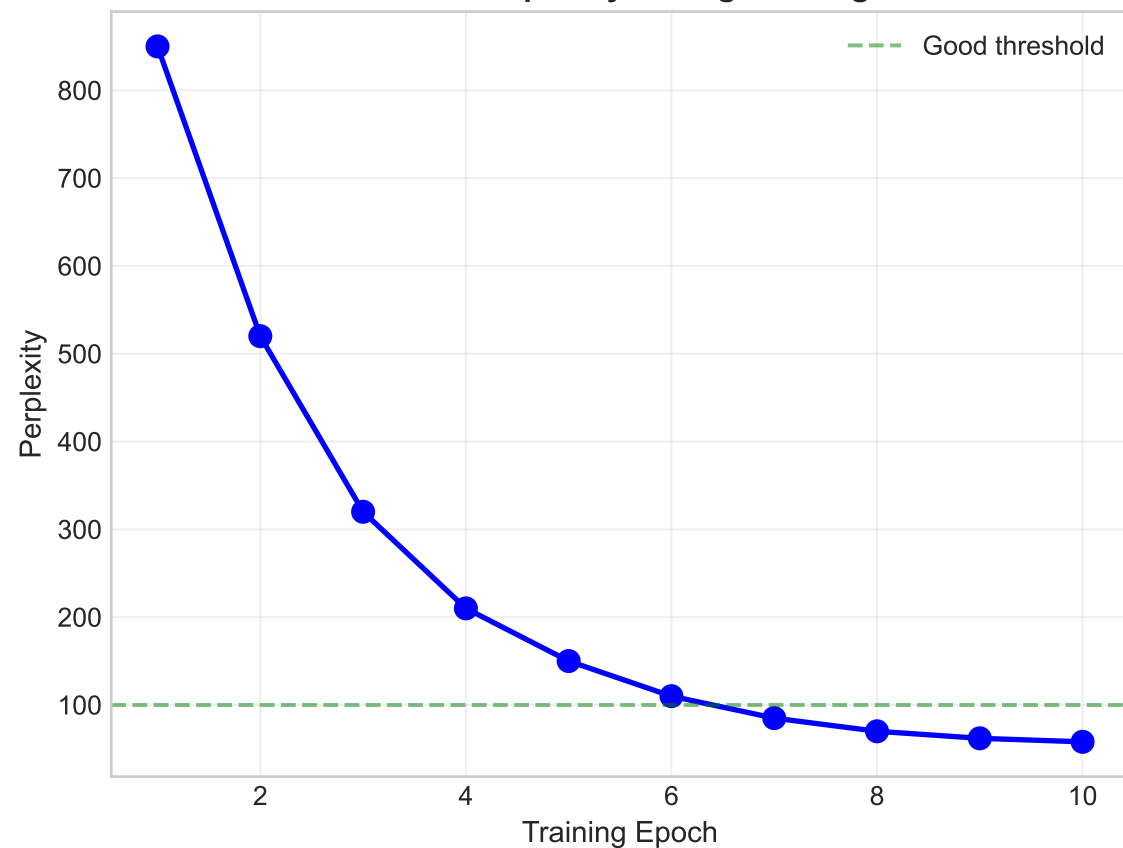
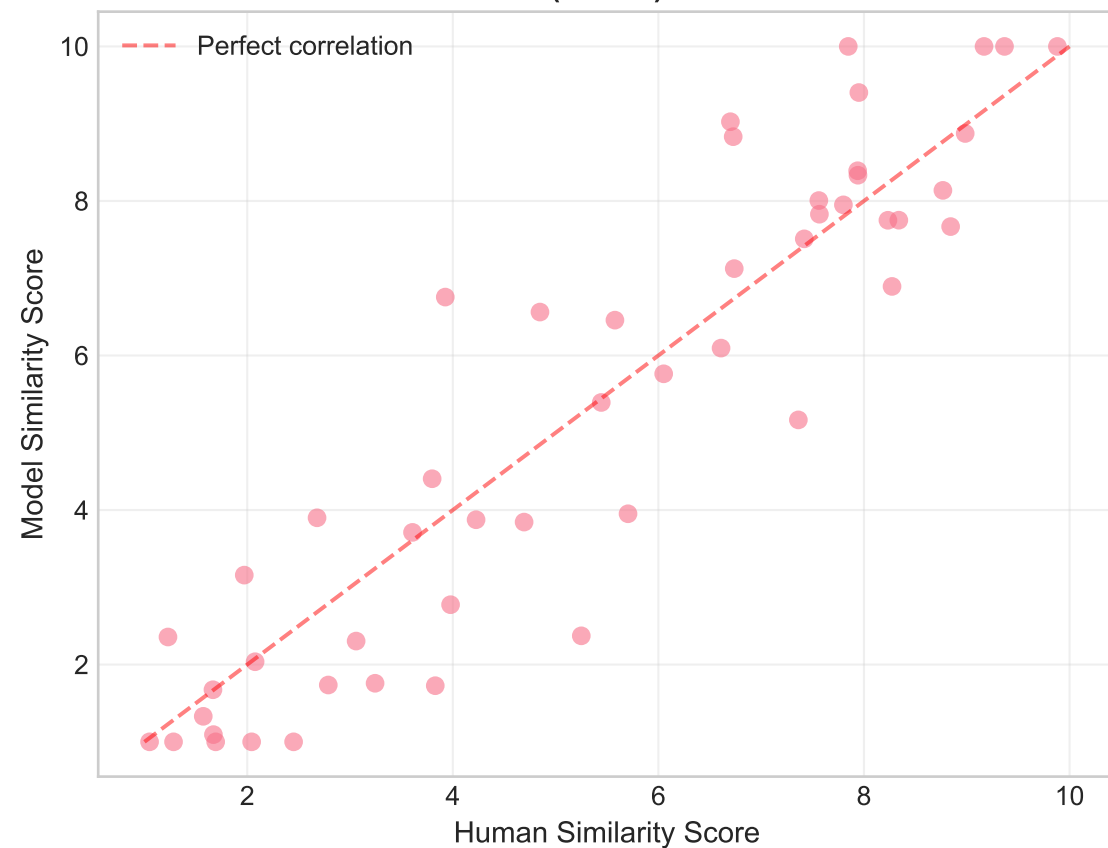


Word Embedding Evaluation Metrics

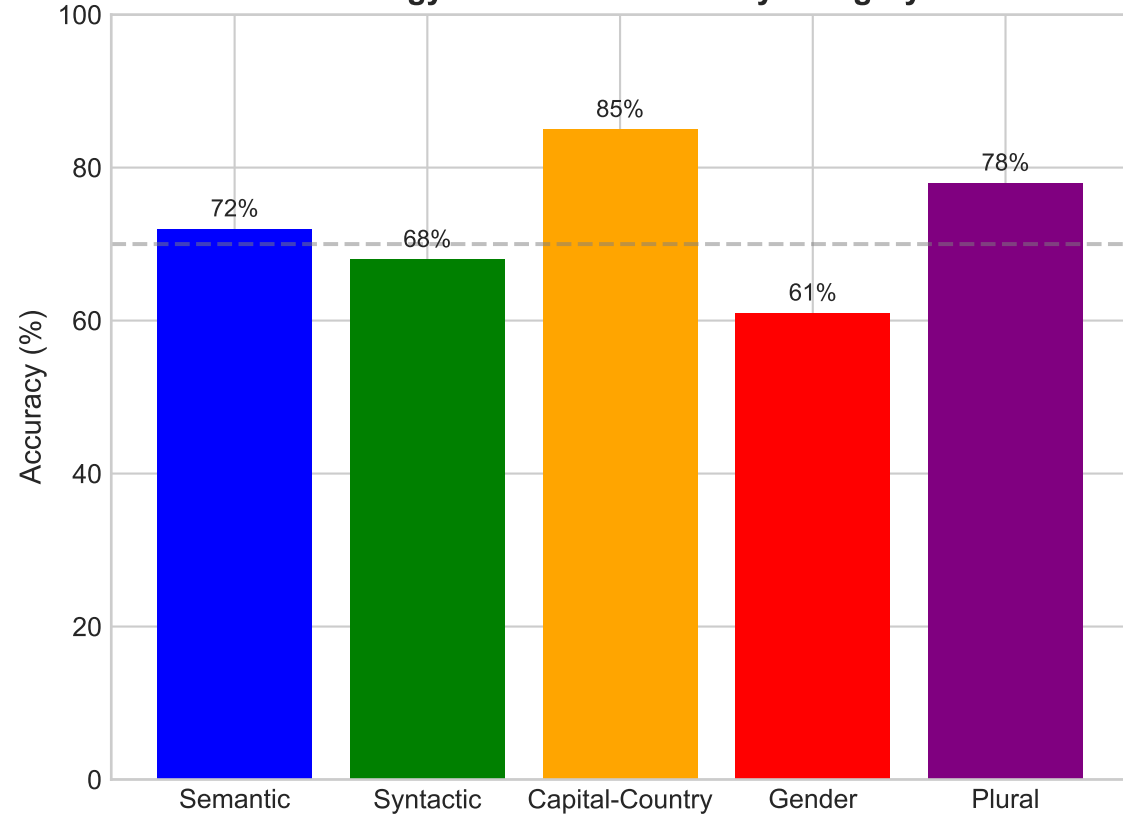
Model Perplexity During Training



Word Similarity: Human vs Model ($r=0.72$)



Analogy Task Performance by Category



Downstream Task Performance

