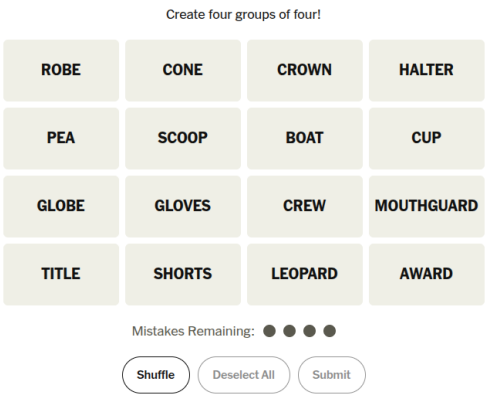
Explicit Semantic Modeling for Solving Word Association Puzzles

Aayush Pandey 2025201058
Srushti Pekamwar 2025201066
Suparshwa Patil 2025201090
Anurag Kacholiya 2025202025

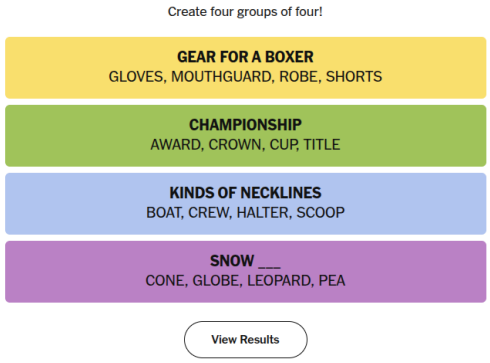
Abstract

Word connection puzzles such as *New York Times Connections* require grouping a fixed set of words into subsets based on shared semantic, contextual, or structural relations. While large language models and embedding-based clustering methods show moderate success, prior analyses demonstrate that they fail systematically on puzzles involving polysemy, distractors, and form-based categories, largely due to their reliance on surface-level similarity and lack of explicit relational reasoning. Cognitive and psycholinguistic research suggests that human word grouping relies on structured semantic networks and constraint satisfaction rather than independent similarity judgments. This project investigates whether explicit semantic knowledge, combined with neural scoring and global optimization, can better model the reasoning required for such puzzles.

We propose a knowledge-guided neural framework that treats puzzle solving as a **group validity scoring problem** rather than a word-level classification task. Given a puzzle, the model evaluates candidate groups of four words using a hybrid representation that integrates semantic embeddings, curated knowledge graph relations (from WordNet, ConceptNet, WikiData and BabelNet), and form-based features capturing patterns such as shared affixes or lexical constructions. Training data is constructed by enumerating all candidate groups within each puzzle and supervising the model using gold groups, near-miss groups, and hard negatives. At inference time, a global optimization step selects a partition of groups that maximizes total validity while satisfying puzzle constraints. The project is structured in three parts: (i) establishing empirical baselines using embedding-based clustering, (ii) developing the knowledge-guided group scoring solver, and (iii) exploring controlled puzzle generation for stress-testing and analysis (Optional). This framework aims to improve robustness on adversarial categories while offering interpretable, human-aligned reasoning.



(fig 1.1) Example Puzzle.



(fig 1.2) Solution.

.