Topic 3: Relation Classification

October 21, 2025

This document provides a description of the **BioRed** dataset. BioRed is based on Reddit posts, which are split into individual sentences for relation classification.

Important Note: Your Mental Health The dataset may contain sensitive or potentially distressing content. If you encounter any disturbing material during your review or experimentation, please inform your instructor immediately.

- **subreddit**: The subreddit from which the post originates.
- **text**: A single sentence extracted from a Reddit post, uniquely identified by **post_id** and **sentence_id**.
- **sentence_id**: Sentence index within a Reddit post. Each sentence in a post has a unique **sentence_id**.
- **post_title**: Title of the original Reddit post.
- **num_annotators**: Total number of unique annotators who labeled the instance.
- **annotated_labels**: Raw span-level labels from each annotator, before any aggregation. Each list entry corresponds to one annotator.
- **majority_labels**: Label(s) agreed upon by the majority of annotators (>50%), based on label presence only (not span position).
- **span_level_agreement**: measure of span-level agreement between the annotators computed using the spans and labels from **annotated_labels**, counting the number of shared points within those spans and dividing by the total unique points across annotations, e.g. 8 shared out of 10 $\rightarrow$ 0.80.
- **auto_labels**: Automatically assigned labels produced using centroid-based semantic similarity between each sentence embedding and relation prototype embeddings (built from seed examples). These labels correspond exactly to the (long, human-readable) label names defined in the *annotation_guidelines.pdf* file. A label is assigned if its similarity score exceeds a fixed threshold. **This column is not intended for model training**

— it should only be used during results analysis (e.g. to compare manual annotations with automatic labels).

- **consensus_spans**: Character-level aggregated span annotations obtained via majority voting over annotator-provided spans.
- Column **parsed_labels** If this column contains an empty list ([]), the instance is treated as having the label `no_relation` in **majority_labels**. Such "no relation" instances **must remain in the training data**, as the model should also learn to recognise when no meaningful relation is present.
- The label **unclear / uncertain** should be **removed from training**. However, such instances may still be inspected later (e.g., for explaining cases where the model struggles).
- Some instances appear multiple times because they were annotated by different annotator groups focusing on different sets of relation types (see column **group**). Duplicate instances can be detected using the pair **post_id** + **sentence_id**. These duplicates should **not be aggregated**, but it is essential to ensure all occurrences of a given instance appear in the **same data split** (to prevent train-dev-test leakage).
- The column **consensus_spans** represents the final token-level aggregated spans obtained through majority voting over the character-level overlaps of all annotators' span annotations.