Appendix: Symbolic Graph Structures for Clause-level Intent Modeling

Chen et al.

Paper under-reviewed at HC@AIxIA + HYDRA 2025 Workshop

Appendix

A.1 Graph Construction Details

We summarize the construction process for all graph variants used in our experiments:

- **Dependency Graph (DEP-GCN)**: Token-level syntactic dependency graphs were extracted using `spaCy`'s dependency parser. Each clause was parsed independently to ensure alignment with sub-sentential units.
- **Semantic Role Graphs (SRL-GCN)**: Three variants were created from predicate-argument structures. The weighted variant emphasized salient roles; the predicate-centric version linked each predicate to its arguments; the anchored version centered the clause and attached all relevant semantic roles.
- **AMR Graphs (AMR-GCN)**: Clause-level AMR graphs were obtained using the SPRING model in `amrlib`. These conceptual graphs represent predicate-argument structures at an abstract level, encoding deep semantics.
- **Narrative Graphs (MSG-AMR-GCN, MSG-MLP)**: Clause-centered ego-graphs were created using predefined symbolic relations such as `next`, `elaboration`, and `contrast`. MSG-AMR-GCN integrates AMR-based features, while MSG-MLP flattens node features for non-structural input.

A.2 Additional Graph Visualizations

Figures 1–5 show additional examples of the symbolic and linguistic graph structures used in our experiments. Each figure visualizes a graph corresponding to a clause from the dataset, demonstrating the construction logic and connection patterns.

1 Training Curves for All Models

A.1 Per-Class Scores

The heatmap below shows the error distribution by true label across all test instances. Each row corresponds to a gold subcode label, annotated with the total number of occurrences. Color intensity indicates the number of incorrect predictions, providing insight into class imbalance and model-specific failure modes.

In addition, we provide raw per-class F1 scores for all 35 subcode classes across selected models in Table 1.

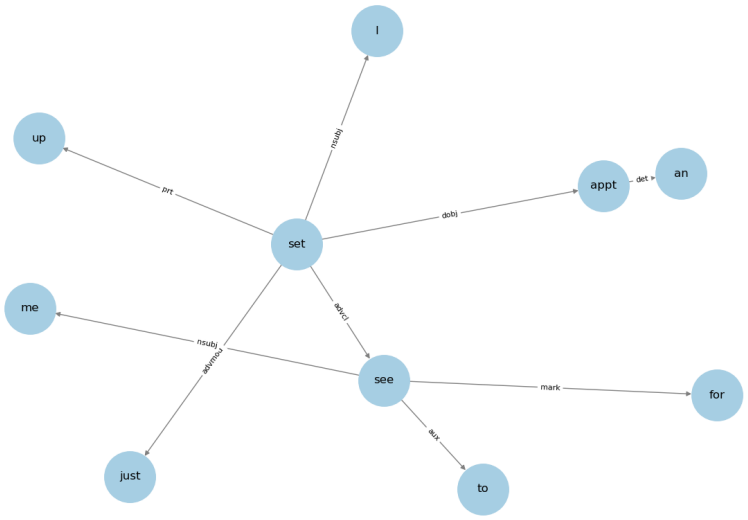


Fig. 1: Syntactic dependency graph of a clause, with grammatical edges.

Table 1: Per-class F1 scores for selected models (abbreviated sample). Full version available in supplementary material.

Subcode	MSG-AMR-GCN	SRL-GCN	DEP-GCN	Clause-MLP	Sentence-MLP
`signoff`	0.88	0.76	0.79	0.72	0.60
`schedulingAppt`	0.72	0.65	0.62	0.60	0.54
`diagnostics`	0.68	0.60	0.59	0.58	0.49
`exploreOptions`	0.33	0.29	0.24	0.22	0.21

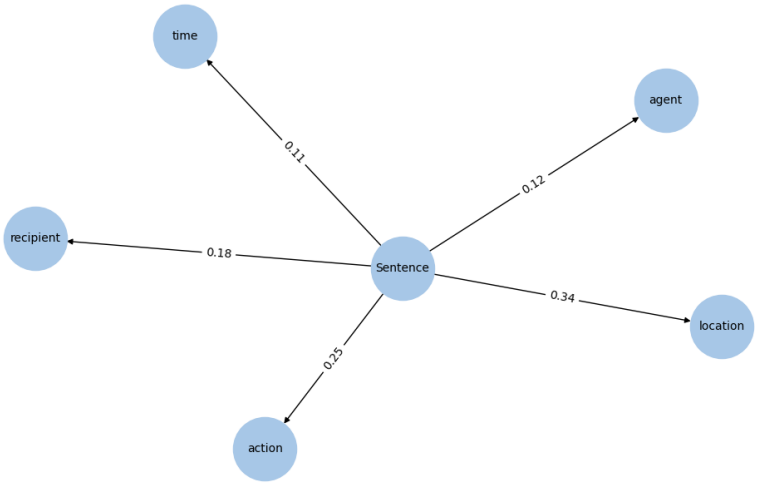


Fig. 2: SRL role-weighted focus graph with edges to salient arguments.

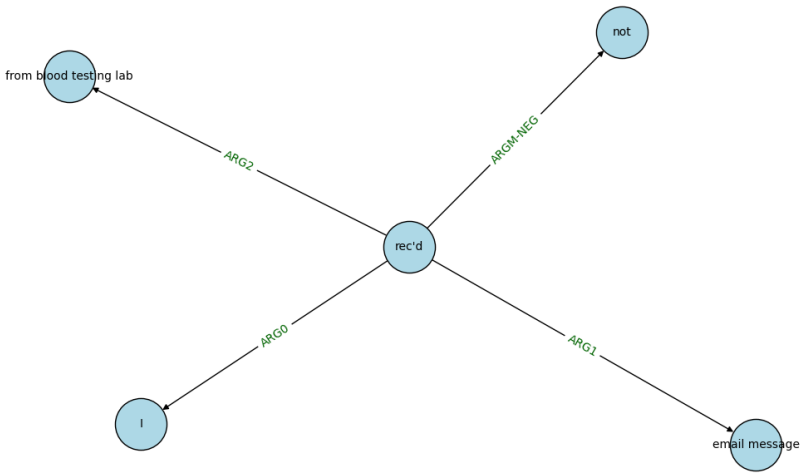


Fig. 3: Predicate-centered SRL graph using AllenNLP outputs.

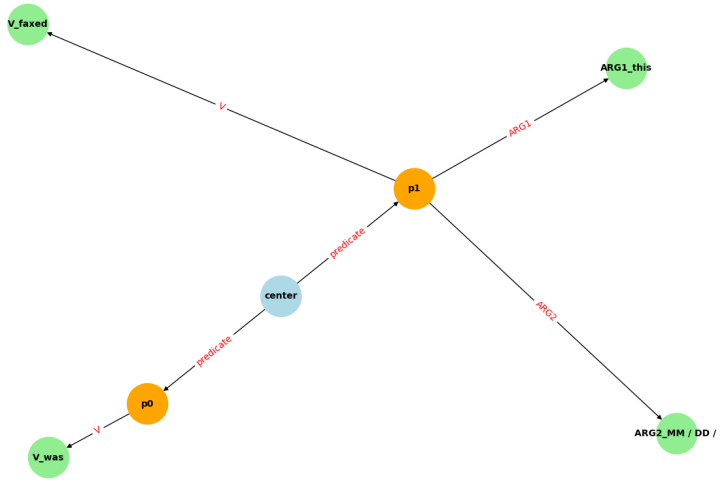


Fig. 4: Example SRL-anchored graph structure.

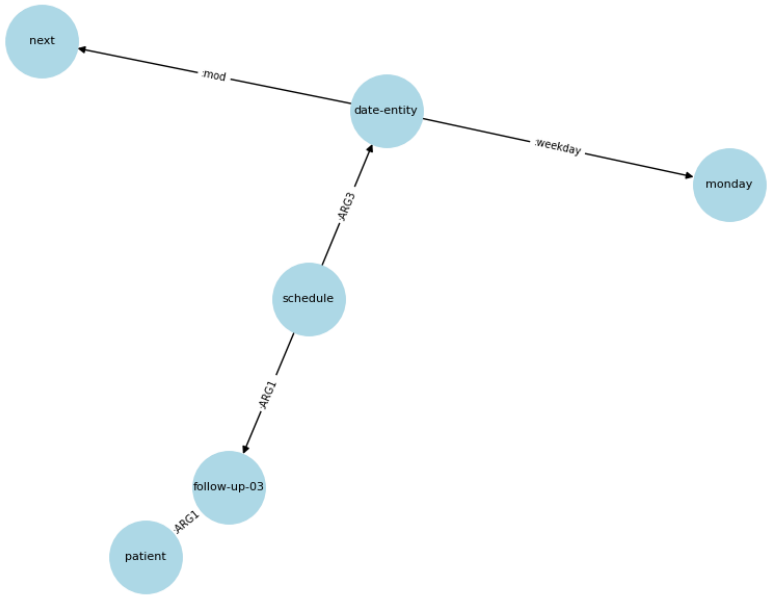


Fig. 5: Conceptual AMR graph representing a clause with abstract roles.

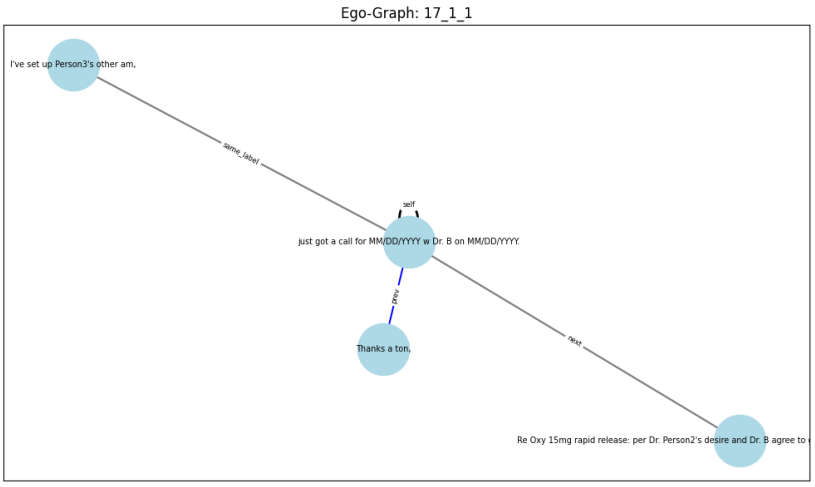


Fig. 6: Symbolic Narrative Ego-Graph

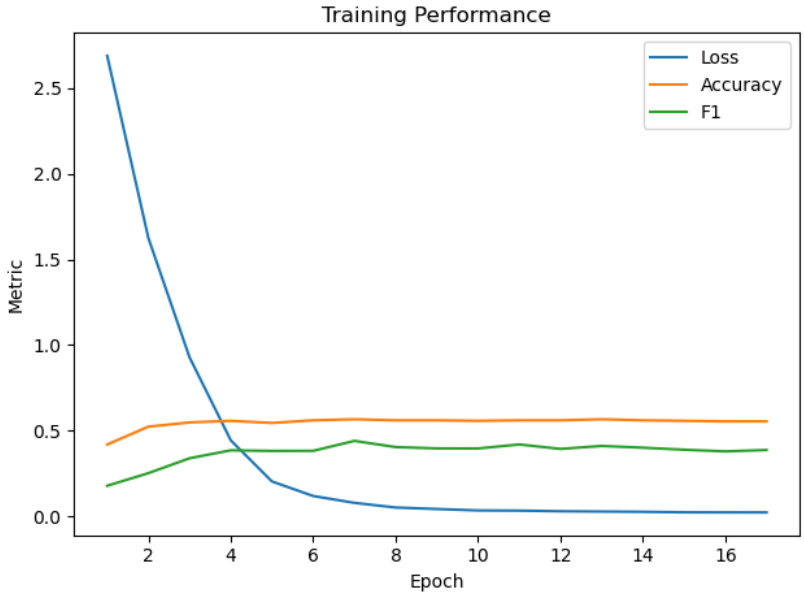


Fig. 7: Training curve for sentence-level MLP.

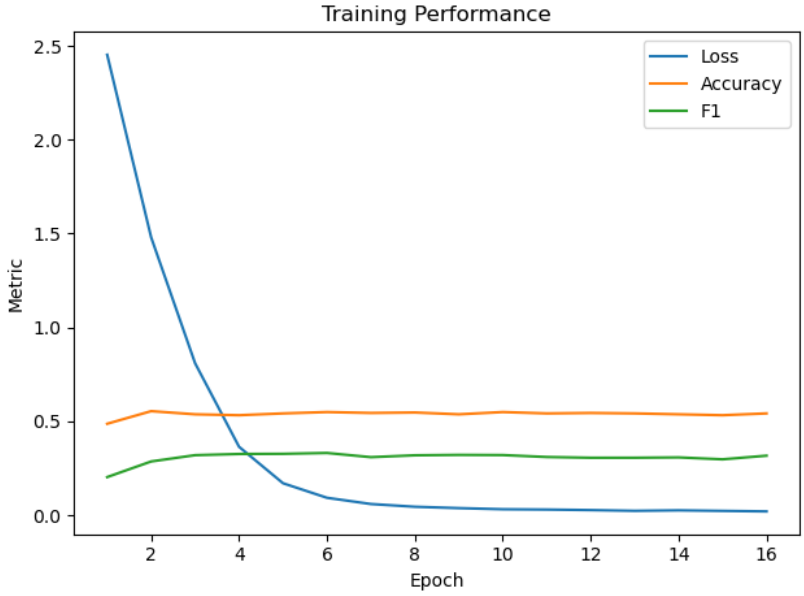


Fig. 8: Training curve for clause-level MLP.

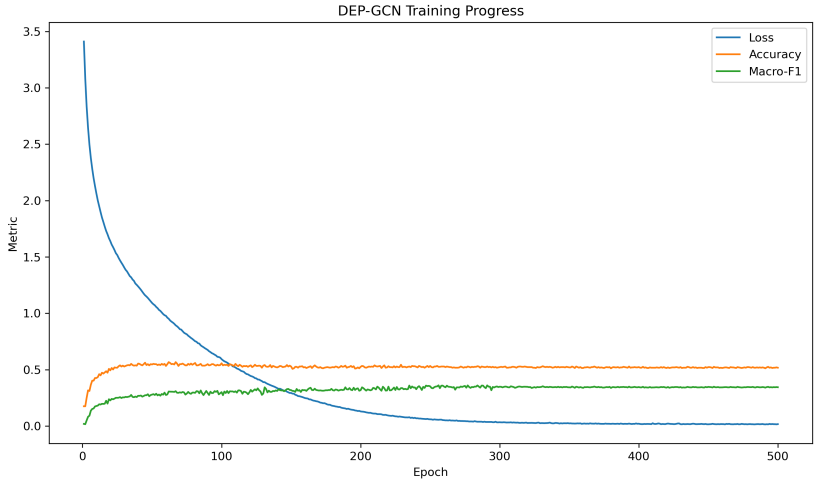


Fig. 9: Training curve for DEP-GCN.

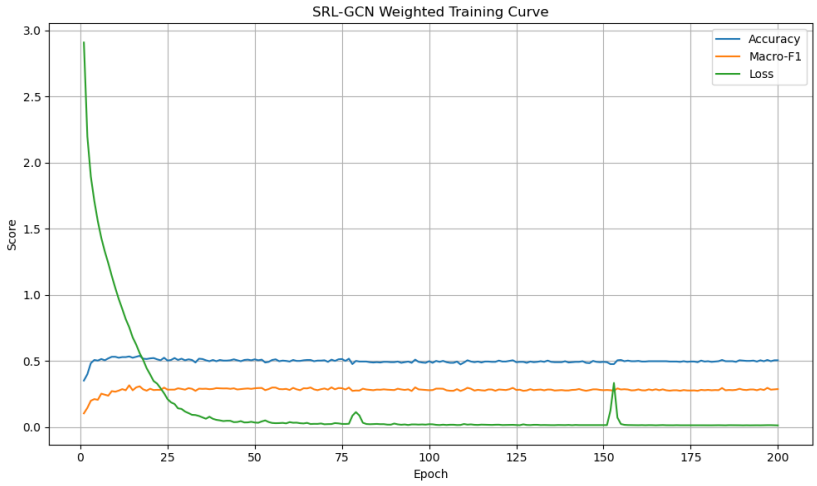


Fig. 10: Training curve for SRL-GCN-weighted.

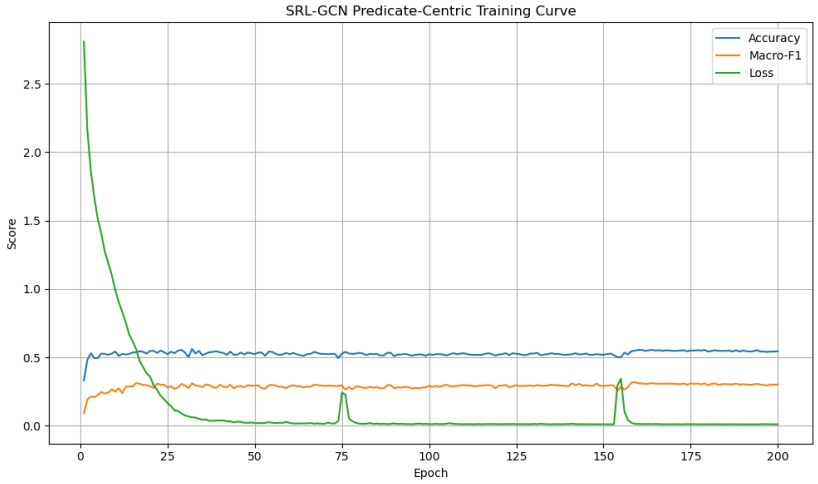


Fig. 11: Training curve for SRL-GCN-anchored.

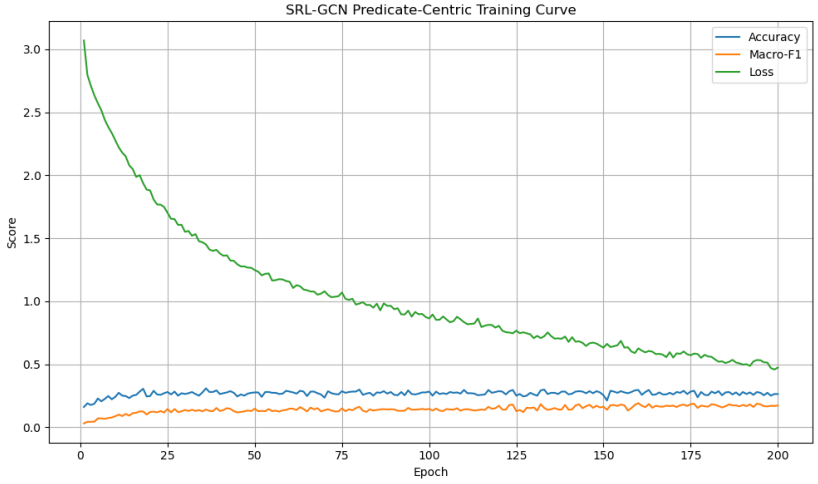


Fig. 12: Training curve for SRL-GCN-predicate.

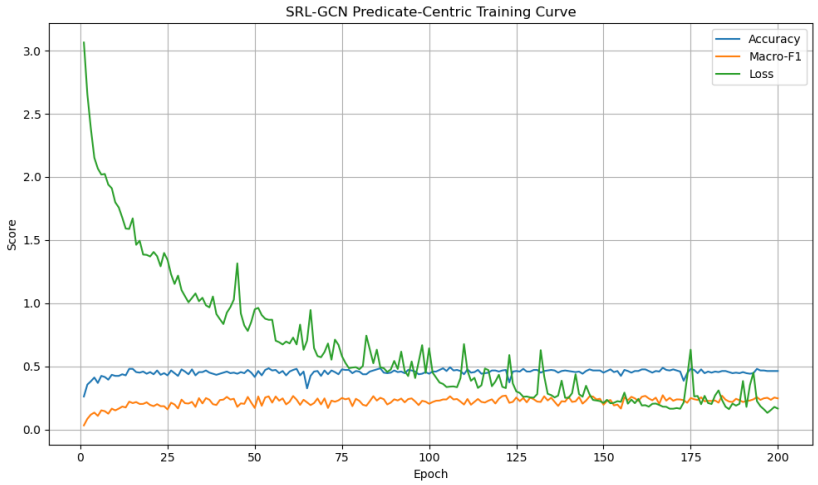


Fig. 13: Training curve for AMR-GCN.

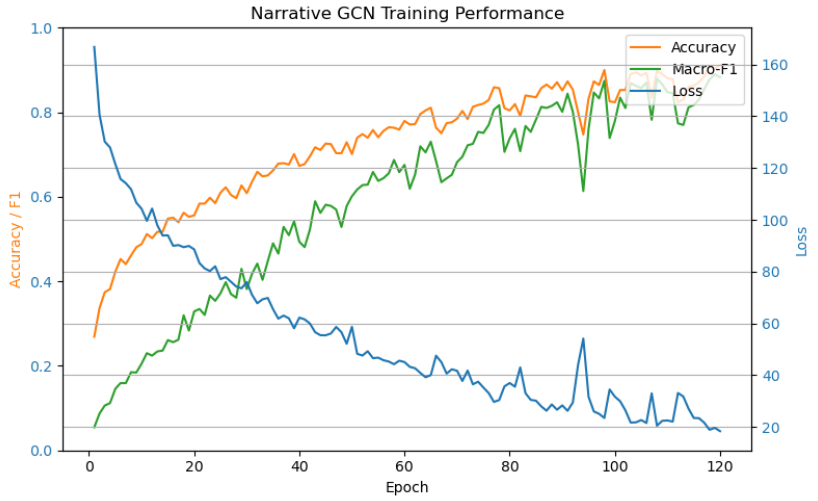


Fig. 14: Training curve for MSG-AMR-GCN.

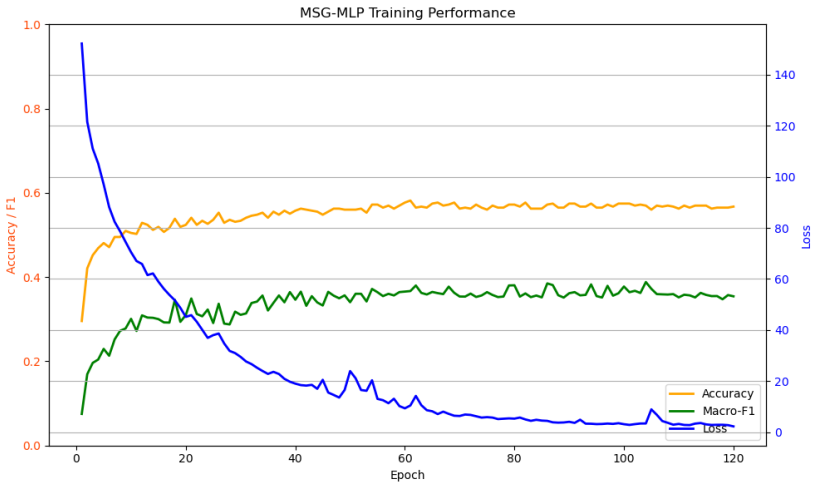
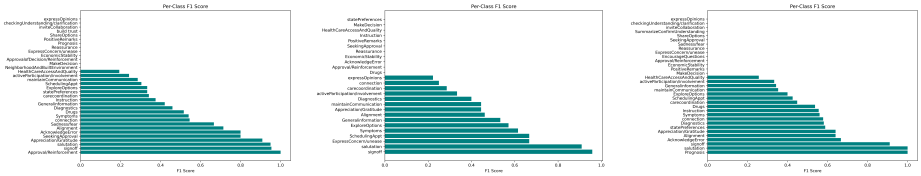


Fig. 15: Training curve for MSG-MLP.



(a) Dependency GCN (b) Narrative-AMR GCN (c) SRL-Weighted GCN

(d) Sentence MLP (e) Clause-level MLP

Fig. 16: Per-class F1 scores across five models. Each plot shows sorted F1 across 35 sub-CODE labels.

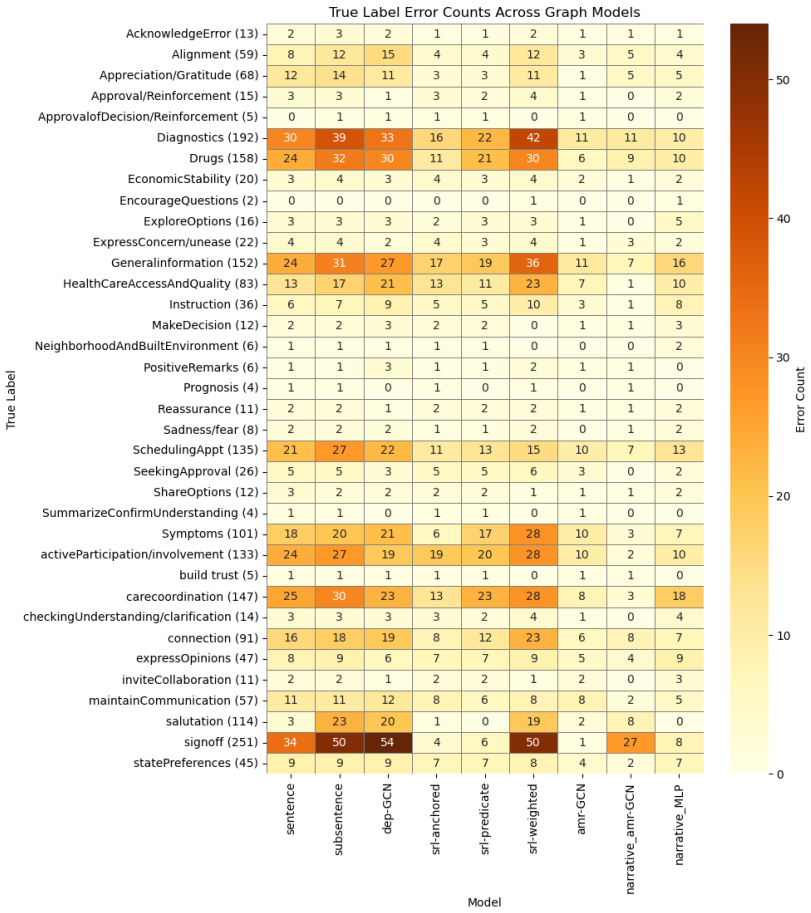


Fig. 17: Error heatmap by true label: darker shades indicate higher misclassification frequency.