LLMs vs Human Annotations Schema Matches

% of responses

matching_df mismatching_df