0: Remaining 0 -1: Non-systematic-review 2: Human-non-RCT-non-drug-intervention - 703: Human-non-RCT-drug-intervention 4: Human-case-report 5: Animal-other 2 -6: Animal-drug-intervention

Confusion Matrix for Model gpt-3.5-turbo and Prompt P2\_HIERARCHY

8 -

9 -

10 - 3

11 - 0

12 - 0

13 -

**Predicted Labels** 

- 607: Human-systematic-review 3 -8: In-vitro-study 9: Human-RCT-non-drug-intervention 

4 -10: Animal-non-drug-intervention - 5011: Human-RCT-drug-intervention

5 -12: Clinical-study-protocol 13: Human-RCT-non-intervention

6 -

- 40

**True Labels** 

- 30

- 20

- 10

- 0