Model: gpt-4o-mini

overall

| **Experiment name** | **Human annotated** | | | **Synthetic** | | |
| --- | --- | --- | --- | --- | --- | --- |
| **P** | **R** | **F0.5** | **P** | **R** | **F0.5** |
| zeroshot | 46.67 | 31.41 | 42.53 | 29.64 | 28.05 | 29.30 |
| Chain-of-Thought | 46.03 | 29.76 | 41.49 | 30.03 | 27.25 | 29.43 |
| Our system | **58.68** | **41.61** | **54.43** | 32.46 | 31.22 | 32.20 |
| mt5-base | 38.96 | 12.99 | 27.83 | **74.62** | **50.03** | **67.94** |

Integrating M1 only

| **K1 (with CoT)** | **Human annotated** | | | **Synthetic** | | |
| --- | --- | --- | --- | --- | --- | --- |
| **P** | **R** | **F0.5** | **P** | **R** | **F0.5** |
| 1 | 53.58 | 36.83 | 49.11 |  |  |  |
| 3 | 55.02 | 38.27 | 50.59 | 32.79 | 30.88 | 32.39 |
| 5 | 56.10 | 39.15 | 51.63 | 32.67 | 30.89 | 32.30 |
| 7 | 55.59 | 40.21 | 51.64 |  |  |  |

Integrating M2 only

| **K2 (with CoT)** | **Human annotated** | | | **Synthetic** | | |
| --- | --- | --- | --- | --- | --- | --- |
| **P** | **R** | **F0.5** | **P** | **R** | **F0.5** |
| 1 | 51.56 | 31.97 | 45.93 |  |  |  |
| 3 | 50.17 | 32.55 | 45.27 |  |  |  |
| 5 | 50.99 | 33.19 | 46.05 |  |  |  |
| 7 | 50.16 | 32.70 | 45.32 |  |  |  |

Delta1 k1=5, k2=2

| **Delta1** | **Initial value** | **alpha** | **Human annotated** | | | **Synthetic** | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **P** | **R** | **F0.5** | **P** | **R** | **F0.5** |
| constant | 0.3 | - | 54.82 | 38.64 | 50.58 | 32.46 | 30.91 | 32.13 |
| exp | 0.3 | 0.0035 | 57.01 | 40.72 | 52.79 | 31.66 | 31.02 | 31.53 |
| linear | 0.3 | 0.0001 | 55.36 | 40.72 | 51.65 |  |  |  |

Ablation study

| **Configuration** | | **Human annotated** | | | **Synthetic** | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **P** | **R** | **F0.5** | **P** | **R** | **F0.5** |
| Removing M2 | | 55.59 | 40.21 | 51.64 |  |  |  |
| Removing M1 | | 51.56 | 31.97 | 45.93 |  |  |  |
| Without Chain-of-Thought | | 53.58 | 36.60 | 49.03 |  |  |  |

| **k1** | **k2** | **Human annotated** | | | **Synthetic** | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **P** | **R** | **F0.5** | **P** | **R** | **F0.5** |
| 5 | 1 | 56.06 | 39.45 | 51.70 |  |  |  |
| 5 | 3 | 56.98 | 41.61 | 53.06 |  |  |  |
| 5 | 5 | 56.59 | 40.04 | 52.27 |  |  |  |
| 5 | 7 | 56.80 | 40.59 | 52.60 |  |  |  |
| 3 | 1 | 57.27 | 39.96 | 52.70 |  |  |  |
| 3 | 3 | 55.29 | 38.69 | 50.92 |  |  |  |
| 3 | 5 | 55.46 | 39.75 | 51.39 |  |  |  |
| 3 | 7 | 55.14 | 40.63 | 51.47 |  |  |  |
| 1 | 1 | 53.27 | 36.46 | 48.77 |  |  |  |
| 1 | 3 | 56.13 | 39.02 | 51.61 |  |  |  |
| 1 | 5 | 58.68 | 39.49 | 53.48 |  |  |  |
| 1 | 7 | 55.83 | 38.64 | 51.27 |  |  |  |
| 7 | 1 | **58.68** | **41.61** | **54.43** |  |  |  |
| 7 | 3 | 55.62 | 39.75 | 51.51 |  |  |  |
| 7 | 5 | 55.79 | 39.75 | 51.62 |  |  |  |
| 7 | 7 | 56.60 | 40.80 | 52.53 |  |  |  |

mT5-guj-gec

| **epochs** | **P** | **R** | **F0.5** |
| --- | --- | --- | --- |
| 4 | 37.40 | 9.96 | 24.11 |
| 8 | 40.00 | 10.82 | 25.99 |
| 9 | 39.31 | 12.34 | 27.35 |
| 10 | **38.96** | **12.99** | **27.83** |
| 12 | 37.76 | 11.69 | 26.11 |
| 13 | 39.69 | 11.26 | 26.37 |
| 14 | 36.96 | 11.04 | 25.15 |
| 15 | 38.19 | 11.90 | 26.49 |