# Experimental Results (macOS)

## Overall metrics

|  |  |  |
| --- | --- | --- |
| Accuracy (%) | Latency (ms) | Effort Saved (%) |
| 96.79 | 8003.3 | 80.0 |

## Per-test metrics

|  |  |  |  |
| --- | --- | --- | --- |
| Test case | Accuracy (%) | Latency (ms) | Effort Saved (%) |
| TC1 | 100.0 | 4497.1 | 84.9 |
| TC10 | 100.0 | 9851.7 | 70.6 |
| TC11 | 100.0 | 2682.9 | 80.0 |
| TC12 | 100.0 | 3249.3 | 83.3 |
| TC13 | 100.0 | 4802.9 | 87.5 |
| TC14 | 100.0 | 6127.1 | 75.0 |
| TC15 | 100.0 | 6012.0 | 83.3 |
| TC16 | 100.0 | 3959.1 | 71.4 |
| TC17 | 100.0 | 5969.4 | 83.3 |
| TC18 | 100.0 | 5582.3 | 81.2 |
| TC19 | 85.71 | 3545.2 | 75.0 |
| TC2 | 100.0 | 6048.3 | 79.2 |
| TC20 | 100.0 | 7041.6 | 78.6 |
| TC21 | 100.0 | 5088.3 | 83.3 |
| TC22 | 100.0 | 6068.3 | 77.3 |
| TC23 | 100.0 | 4698.4 | 76.0 |
| TC24 | 100.0 | 3243.7 | 87.5 |
| TC25 | 100.0 | 3246.1 | 86.4 |
| TC26 | 100.0 | 8312.6 | 66.3 |
| TC27 | 100.0 | 4391.0 | 81.2 |
| TC28 | 14.29 | 3068.0 | 75.0 |
| TC29 | 100.0 | 4810.1 | 75.0 |
| TC3 | 100.0 | 2896.9 | 85.7 |
| TC30 | 100.0 | 5217.7 | 78.6 |
| TC31 | 100.0 | 3580.6 | 83.3 |
| TC32 | 100.0 | 2607.6 | 80.0 |
| TC33 | 100.0 | 3716.9 | 87.5 |
| TC34 | 100.0 | 8311.7 | 65.5 |
| TC35 | 100.0 | 4257.3 | 81.2 |
| TC36 | 100.0 | 125806.9 | 69.4 |
| TC37 | 71.43 | 3459.4 | 75.0 |
| TC38 | 100.0 | 3771.6 | 81.2 |
| TC39 | 100.0 | 2997.9 | 87.5 |
| TC4 | 100.0 | 4590.1 | 86.4 |
| TC40 | 100.0 | 2878.9 | 83.3 |
| TC5 | 100.0 | 5657.3 | 84.6 |
| TC6 | 100.0 | 5877.0 | 80.0 |
| TC7 | 100.0 | 4981.4 | 83.3 |
| TC8 | 100.0 | 7231.3 | 73.2 |
| TC9 | 100.0 | 3831.9 | 85.0 |

## Robustness

Overall robustness on paraphrases: 98.75 %

## Figures







