|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Pearson r** | **MAE** | **MSE** | **RMSE** |
| Fresh Qwen3-0.6b | 0.0076 | 1.7388 | 3.7506 | 1.9366 |
| Fresh Qwen1.7b | 0.1906 | 0.7940 | 1.1254 | 1.0608 |
| Full finetuned Qwen3-0.6b | 0.2994 | 0.6166 | 0.7569 | 0.8700 |
| Finetuned Qwen3-0.6b with LoRA | 0.4513 | 0.6002 | 0.6908 | 0.8311 |
| Finetuned Qwen3-0.6b with GPT reasonings | 0.4820 | 0.7336 | 0.9634 | 0.9815 |
| COMET | 0.5984 | 0.8621 | 1.2633 | 1.1240 |