# Performance Metrics by Model (Median and IQR)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Cosine Similarity (Median) | Cosine Similarity (IQR) | BLEU Score (Median) | BLEU Score (IQR) |
| GPT-4o + RAG | 0.3472 | 0.2159 | 0.0046 | 0.0082 |
| GPT-4o | 0.3208 | 0.2362 | 0.0043 | 0.0087 |
| LLaMA3.1-Sonar-405b + Online | 0.3170 | 0.2384 | 0.0075 | 0.0163 |
| GPT-4o mini | 0.2974 | 0.2193 | 0.0056 | 0.0096 |
| GPT-4 Turbo | 0.2933 | 0.2061 | 0.0048 | 0.0105 |
| LLaMA3-70b | 0.2458 | 0.2038 | 0.0048 | 0.0102 |
| Gemini-1.5-Pro | 0.2317 | 0.1931 | 0.0042 | 0.0058 |
| LLaMA3.1-Nemotron-70b | 0.2286 | 0.1942 | 0.0022 | 0.0028 |
| Mixtral-8x7b | 0.2048 | 0.1922 | 0.0053 | 0.0120 |