

2-

ROUGE and BLEU Scores of NLLB,IndicTrans and ChatGPT

	Rouge-1(Recall)	Rouge-1(Precision)	Rouge-1(F-Score)	Rouge-2(Recall)	Rouge-2(Precision)	Rouge-2(F-Score)	Rouge-L(Recall)	Rouge-L(Precision)	Rouge-L(F-Score)	BLEU Score
NLLB (hi to en)	0.8563	0.9592	0.9022	0.8096	0.9122	0.8549	0.8532	0.9558	0.8990	0.8715
NLLB (en to hi)	0.5565	0.6029	0.5740	0.3260	0.3547	0.3365	0.5314	0.5730	0.5470	0.5920
NLLB (hi to mr)	0.5087	0.6172	0.5547	0.3036	0.3742	0.3332	0.5006	0.6073	0.5458	0.6907
NLLB (mr to hi)	0.5015	0.5415	0.5166	0.2646	0.2842	0.2714	0.4762	0.5116	0.4894	0.5283
IndicTrans(hi to en)	0.6258	0.6749	0.6478	0.4260	0.4555	0.4391	0.5950	0.6400	0.6153	0.7517
IndicTrans(en to hi)	0.6199	0.6229	0.6196	0.3951	0.3959	0.3941	0.5961	0.5988	0.5958	0.6722
IndicTrans(hi to mr)	0.4016	0.4320	0.4129	0.1788	0.1964	0.1856	0.3803	0.4090	0.3909	0.6016
IndicTrans(mr to hi)	0.5234	0.5402	0.5279	0.2727	0.2816	0.2749	0.4896	0.5061	0.4942	0.5632
ChatGPT(hi to en)	0.4279	0.4425	0.4330	0.2782	0.2964	0.2857	0.4143	0.4284	0.4192	0.6097
ChatGPT(en to hi)	0.5430	0.5453	0.5419	0.2806	0.2812	0.2795	0.5172	0.5186	0.5158	0.6067
ChatGPT(hi to mr)	0.2405	0.2679	0.2516	0.0660	0.0723	0.0683	0.2322	0.2584	0.2427	0.4836
ChatGPT(mr to hi)	0.5036	0.5453	0.5175	0.2713	0.2960	0.2794	0.4711	0.5081	0.4832	0.5307

3 -

NLLB model demonstrates strong performance, especially for Hindi to English translation, achieving high ROUGE scores across all metrics. This indicates that NLLB produces accurate and fluent translations from Hindi to English. IndicTrans also performs well, particularly for English to Hindi translation. ChatGPT, while showing reasonable performance, tends to lag behind in certain metrics compared to NLLB and IndicTrans.

IndicTrans demonstrates competitive performance with higher ROUGE and BLEU scores for Indic to Indic translation tasks. Performs well in capturing translation quality and linguistic nuances between Indic languages like Hindi and Marathi. NLLB achieves moderate to good ROUGE scores for Indic to Indic translation (e.g., Hindi to Marathi). It also shows a decent BLEU score, indicating reasonable translation quality and fluency within Indic languages.

IndicTrans appears to be more robust for Indic to Indic translation based on the above scores, followed by NLLB, while ChatGPT may benefit from enhancements to optimise translation quality within Indic language pairs.