MedQwen-PE: Medical Qwen for Parameter-Efficient Multilingual Patient-Centric Summarization, Question Answering and Information Extraction

Vinay Babu Ulli
Oogwai Analytics
Bangalore, India
ullivinaybabu@gmail.com

Anindita Mondal
Language Technologies Research Center
IIIT Hyderabad, India
anindita.mondal@research.iiit.ac.in

Abstract

This study addresses the Shared Task on Patient-Centric Multilingual Question Answering, which focuses on generating summaries and patient-oriented answers from multi-turn medical dialogues related to Head and Neck Cancer and Cystic Fibrosis across ten languages. The Qwen3-1.7B model is fine-tuned using QLoRA for three tasks—Summarization, Question Answering, and Information Extraction—while updating only approximately 1.6% of parameters through task-specific adapter layers. The resulting system demonstrates strong semantic fidelity, as evidenced by high BERTScore and COMET scores, particularly for Kannada, English, Telugu, and Tamil, with comparatively lower performance in Assamese, Bangla, Gujarati, and Marathi. The modular fine-tuning design enables efficient task adaptation while satisfying the constraints on model size and computational resources.

1 Introduction

In recent years, patient-centric natural language processing (NLP) has gained increasing attention for its potential to improve access to medical information and empower patients in clinical decision-making (Jerfy et al., 2024; Takale, 2024; Zhou et al., 2024; Rojas-Carabali et al., 2024). Multi-turn medical dialogues, especially for complex conditions such as Head and Neck Cancer and Cystic Fibrosis, are often difficult for non-experts to interpret, creating a need for automated systems that can generate summaries and answer patient-oriented questions. Recent advances in large language models (Singhal et al., 2023; Maity and Saikia, 2025; Meng et al., 2024) provide a robust foundation to address these challenges, offering multilingual and long-context capabilities suitable for summarization, question answering, and information extraction. The **Qwen family of models** exemplifies this evolution: **Qwen 1** (Bai et al., 2023) introduced the transformer decoder architecture with causal language modeling, while **Qwen 2** (Yang et al., 2024a) expanded scale (0.5–72 B parameters), adopted Mixture-of-Experts designs for efficiency, and demonstrated strong multilingual proficiency with extended context support of up to 128 K tokens, making it particularly effective for complex patient-centric healthcare dialogues.

The next generation, **Qwen 2.5** (Yang et al., 2024b), refined the architecture and training pipeline to push performance boundaries even further. Trained on an expanded corpus of over 18 trillion tokens and enhanced through multistage post-training with more than one million supervised samples, Qwen 2.5 achieved gains in reasoning, factual grounding, and multilingual understanding. The **Qwen3-1.7B** (Yang et al., 2025) model is a causal language model, designed primarily for generative language tasks such as text completion, summarization, question answering, and dialogue generation. As a causal model, it predicts the next token in a sequence based on all previous tokens, making it particularly effective for autoregressive text generation and understanding long-form context. The model has undergone both pretraining and post-training stages to enhance its linguistic and reasoning capabilities.

The Shared Task on Patient-Centric Question Answering focuses on multilingual health dialogue understanding, summarization, and question answering. The dataset, released as part of the NLP4Health initiative, consists of validated dialogues between patients and healthcare professionals across multiple Indian languages where each dialogue is accompanied by a structured summary and multiple patient-centric question–answer pairs. The aim is to develop models with fewer than 3 billion parameters capable of generating concise summaries of multi-turn medical dialogues and answering patient-oriented questions. The multilingual dataset is partitioned by task and each subset is used to fine-tune a separate Qwen3-1.7B in-

Second Workshop on Integrating NLP and AI for Multilingual and Patient-Centric Healthcare Communication, pages 86–92

stance via QLoRA (Dettmers et al., 2023), enabling parameter-efficient adaptation with only 1.6% trainable parameters. The resulting task-specific LoRA (Hu et al., 2022) adapters form a modular system that supports scalable extension to additional domains. System performance is evaluated using automatic metrics such as ROUGE (Lin, 2004), BLEU (Papineni et al., 2002), BERTScore (Zhang* et al., 2020) for summarization, and Exact Match and F1 score for question answering, complemented by human expert evaluation for medical correctness and clinical usefulness. Our results show strong semantic fidelity across tasks, with consistently high BERTScore and COMET (Rei et al., 2020) values, although lexical overlap remains moderate. Performance varies by language, with notably better results in Kannada, English, Telugu, and Tamil, and lower performance for Assamese, Bangla, Gujarati, and Marathi.

2 Proposed Approach

The objective of this work is to employ a language model capable of performing multilingual summarization, information extraction, and question answering. Considering the high computational and data requirements associated with training a large model from scratch, we opted to fine-tune an existing pretrained model using the task-specific dataset provided for this study. To ensure the suitability of the base model, several selection criteria were established in accordance with the shared task requirements :

1. Contain fewer than 3 billion parameters, ensuring computational efficiency and compatibility with limited hardware resources.

2. Exhibit multilingual capabilities, allowing effective processing and understanding of content across multiple languages.

3. Demonstrate strong reasoning and comprehension abilities, enabling robust performance on complex linguistic and contextual tasks.

4. Support a context window of at least 32k tokens, facilitating the handling of long documents and maintaining coherence across extended text sequences.

5. Achieve competitive performance on standard language understanding benchmarks, reflecting its generalization and robustness.

After a comprehensive evaluation of available open-source models under these constraints, we identified Qwen3-1.7B (Yang et al., 2025) as the most suitable base model for our purpose. It strikes an effective balance between model size, multilingual coverage, contextual reasoning, and computational efficiency, making it an ideal choice. With 1.7 billion parameters, the model achieves a balance between performance and efficiency, making it suitable for resource-constrained environments while maintaining strong generalization abilities. It comprises of 28 transformer layers, enabling it to capture deep hierarchical representations of text. The attention mechanism is configured with 16 query heads and 8 key-value heads, allowing the model to process complex contextual relationships across tokens efficiently. A notable feature of Qwen3-1.7B is its extended context length of 32,768 tokens, which allows it to process and reason over long documents without losing coherence. This extended context window makes it particularly suitable for tasks such as document summarization, information extraction, and long-form question answering.

2.1 System Architecture

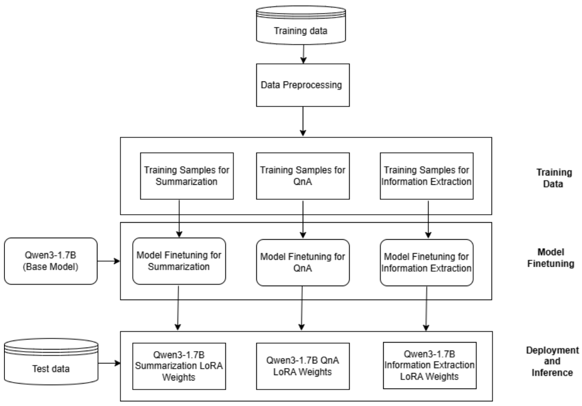

Figure 1: Training–inference pipeline for QLoRA.

As illustrated in Figure 1, the architecture comprises three main stages: 1) data preprocessing, 2) task-specific fine-tuning, and 3)deployment/inference. The process begins preprocessing the multilingual training corpus which is divided into three subsets corresponding to the target tasks: Summarization, Question Answering, and Information Extraction. Each subset is used to fine-tune a separate instance of the Qwen3-1.7B model using Supervised Fine-Tuning (SFT) with QLoRA. This design allows the system to learn task-specific representations while maintaining the efficiency

Table 1: Example dialogues which are incorrect

Language	Dialogue ID	Dialogue text
Assamese	scenario_10_984a19c41d17469cb941bc9904c637a1_IDX_05_2	I can create this long 60+ turn Assamese dialogue in JSONL, but it's best done in batches to keep it natural and accurate. Would you like me to start with Batch 1 (20 lines) now and then continue in subsequent messages?
Telugu	scenario_15_e39e255060f347e380994a0a33f6015d_IDX_04_0	Yes

of parameter-efficient fine-tuning. The QLoRA adapters are injected into the attention and MLP layers so that only a small fraction ($\approx 1.6\%$) of the total parameters are trainable. The fine-tuning phase produces three independent LoRA checkpoints—one for each task—which are subsequently used for inference on the shared-task test datasets. This modular setup facilitates easy extension to new domains by re-training only the relevant adapter weights rather than the full model.

2.2 Dataset Setup and Preprocessing

It was observed that a small portion of the training corpus contained non-conversational or incomplete dialogue structures as shown in Table 1. To improve data consistency, we applied a single-stage filtering criterion based on the number of speaker turns. Specifically, dialogues containing fewer than four occurrences of the token "speaker" were excluded, as such samples did not represent meaningful multi-turn exchanges. This preprocessing step effectively removed noisy or improperly formatted instances while preserving the linguistic diversity of the corpus. Table 2 summarizes the dataset statistics before and after preprocessing for all three tasks—Summarization, Question Answering, and Information Extraction—across ten Indic languages. Overall, the filtering step reduced the dataset size by approximately 12% for summarization, 8.6% for QA, and 10% for information extraction, resulting in 45,885, 11,463, and 25,967 high-quality samples respectively.

For the Question Answering (QnA) task, each training instance was derived from the corresponding dialogue and QnA files provided in the shared dataset. The original QnA files contained multiple question–answer pairs associated with a single dialogue. Instead of treating all questions and answers together as a single training sample, we decomposed each dialogue into multiple independent training examples—each consisting of the dialogue context, one question, and its corresponding answer. This restructuring allowed the model to learn fine-grained contextual alignment between individual questions and relevant dialogue segments. After this transformation, the QnA dataset expanded to over 164,000 training samples, substantially increasing the number of supervised examples available for fine-tuning.

2.3 Supervised Finetuning using QLoRA

To adapt the pretrained Qwen3-1.7B model to the requirements of the shared task, we employed QLoRA (Quantized Low-Rank Adaptation) (Dettmers et al., 2023) for fine-tuning. QLoRA is a parameter-efficient fine-tuning (PEFT) technique that builds upon the LoRA (Low-Rank Adaptation) (Hu et al., 2022) framework, which avoids full model retraining by freezing the pretrained weights and introducing a small number of trainable low-rank matrices within selected layers of the model. In the present work, the LoRA adapters were applied to both the attention and feed-forward (MLP) components of the transformer architecture. Specifically, the fine-tuning was performed on the following projection layers: *q_proj*, *k_proj*, *v_proj*, *o_proj*, *gate_proj*, *up_proj*, and *down_proj*. The chosen LoRA hyperparameters were as follows: *rank* = 16, *lora_alpha* = 32, and *lora_dropout* = 0.1. These settings were selected to provide a balance between adaptation flexibility and regularization, ensuring stable convergence during fine-tuning while minimizing overfitting.

The base model comprises an embedding dimension of 2048 and 28 transformer layers, resulting in a total of 1,749,017,600 parameters. Through QLoRA, only a small subset of parameters was made trainable—specifically 28,442,624 parameters, which accounts for approximately 1.63% of the total model parameters. This configuration allowed fine-tuning without significant computational overhead, while maintaining the expressive capacity of the model. The use of QLoRA thus enabled the model to adapt to task-specific data

Table 2: Dataset sizes before/after preprocessing.

Language	Summarization		QnA		Information Extraction	
	# Samples	# After Preproc.	# Samples	# After Preproc.	# Samples	# After Preproc.
Assamese	2,200	1,900	17	13	755	646
Bangla	6,153	6,037	423	418	2,503	2,461
Dogri	2,526	2,376	129	120	129	120
English	7,106	6,954	5,808	5,664	5,808	5,664
Gujarati	6,169	4,517	417	319	3,484	2,714
Hindi	6,204	6,181	1,093	1,092	5,024	5,014
Kannada	6,629	3,807	958	501	2,774	1,459
Marathi	3,624	3,547	743	726	2,390	2,343
Tamil	5,155	4,802	755	674	2,246	2,147
Telugu	6,629	5,764	2,196	1,936	3,759	3,399
Total	**52,395**	**45,885**	**12,539**	**11,463**	**28,872**	**25,967**

while remaining resource-efficient and scalable for multilingual applications.

3 Experimental Results

We evaluated the results on test set for three subtasks: Question Answering (QnA), Summarization, and Information Extraction (IE). The evaluation metrics are F1, Exact Match, ROUGE-1/2/L, BERTScore, and COMET Score, presented in Tables 3, 4, and 5 (Appendices A.1,A.2 and A.3). Exemplary outputs obtained using the models along with ground truth are available here[1].

In the QnA evaluations, the model achieved higher BERTScore and COMET scores for Kannada, English, Telugu, and Tamil, indicating strong performance in these languages. In contrast, the performance for Assamese, Bangla, Gujarati, and Marathi was considerably lower. This trend was consistent across the remaining evaluation metrics as well.

In Summarization Task, summaries are semantically aligned with the reference texts, as reflected by high BERTScore and COMET values. While the lexical overlap, measured by ROUGE-L F1, remains moderate, the COMET scores (0.6–0.7) indicate that the generated summaries maintain good semantic fidelity. Similarly, BERTScore F1 values of approximately 0.8 across languages suggest that the summaries are informative and meaningfully capture the core content.

In Information Extraction Task, the model achieves BERTScore F1 values above 0.85 for all languages, indicating strong semantic correspondence between the predicted and reference outputs. The results show a consistent pattern of strong semantic adequacy but modest lexical overlap: BERTScore/COMET are comparatively high across tasks, whereas ROUGE lag—indicative of paraphrastic correctness and formatting sensitivity in multilingual, free-form outputs. The widest dispersion appears in QnA, suggesting room for language-aware adaptation (e.g., tokenizer merges, transliteration normalization, in-language augmentation) to narrow gaps for lower-resource languages.

4 Conclusion

This work presents a parameter-efficient approach for multilingual patient-centric dialogue understanding, summarization, and question answering using the Qwen3-1.7B model. By employing QLoRA for task-specific fine-tuning, only a small fraction of model parameters (1.6%) were updated, enabling efficient adaptation under computational constraints. Experimental results demonstrate strong semantic fidelity as reflected by high BERTScore and COMET values, particularly for Kannada, English, Telugu, and Tamil, while highlighting performance gaps in lower-resource languages. The modular design of task-specific LoRA adapters allows for scalable extension to new domains without retraining the full model. Overall, this approach provides an effective and resource-efficient framework for multilingual patient-centric NLP, supporting accurate and informative dialogue summarization and question answering.

[1]https://huggingface.co/datasets/vinaybabu/NLPSharedTask-QnA-Before-After-Finetuning

5 Limitations

The Qwen3-1.7B model appears to have a stronger representation for Kannada, Tamil, Telugu, Hindi and English during pretraining relative to Assamese, Bangla, Gujarati, and Marathi. This imbalance likely contributes to the lower performance observed in the latter set of languages, particularly in the QnA task. We observed specific failure modes—repetitions, copying of question text, and irrelevant expansions particularly in Assamese, Marathi, Bangla, and Gujarati for the QnA task, with representative error cases provided in the accompanying footnote[2]. Also, we provide links [3] to the fine-tuned models for Summarization, QnA, and Information Extraction, all publicly released on Hugging Face. Our fine-tuning data is also limited to provided training data which is not enough for language understanding.

Metrics such as SummaC (Laban et al., 2022) and QAFactEval (Fabbri et al., 2022), which rely on ground truth outputs, could not be computed due to the lack of ground truth data for Question Answering (QnA), Summarization, and Information Extraction (IE) tasks within the test set. Future work includes integrating large-scale Indic corpora during fine-tuning to improve language understanding, and exploring advanced alignment methods such as Reinforcement Learning from Human Feedback(RLHF) to further refine output quality and reduce errors.

References

Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, Binyuan Hui, Luo Ji, Mei Li, Junyang Lin, Runji Lin, Dayiheng Liu, Gao Liu, Chengqiang Lu, Keming Lu, and 29 others. 2023. Qwen technical report. *Preprint*, arXiv:2309.16609.

Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. 2023. Qlora: efficient finetuning of quantized llms. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA. Curran Associates Inc.

Alexander Fabbri, Chien-Sheng Wu, Wenhao Liu, and Caiming Xiong. 2022. QAFactEval: Improved QA-based factual consistency evaluation for summarization. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2587–2601, Seattle, United States. Association for Computational Linguistics.

[2]https://huggingface.co/datasets/vinaybabu/NLPSharedTask-QnA-Error-Patterns

[3]https://huggingface.co/collections/vinaybabu/sharedtask-nlp-finetunedmodels

Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2022. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*.

Aadit Jerfy, Owen Selden, and Rajesh Balkrishnan. 2024. The growing impact of natural language processing in healthcare and public health. *Inquiry: A Journal of Medical Care Organization, Provision and Financing*, 61:469580241290095.

Philippe Laban, Tobias Schnabel, Paul N. Bennett, and Marti A. Hearst. 2022. SummaC: Re-visiting NLI-based models for inconsistency detection in summarization. *Transactions of the Association for Computational Linguistics*, 10:163–177.

Chin-Yew Lin. 2004. ROUGE: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, pages 74–81, Barcelona, Spain. Association for Computational Linguistics.

Subhankar Maity and Manob Jyoti Saikia. 2025. Large language models in healthcare and medical applications: A review. *Bioengineering*, 12(6):631.

Xiangbin Meng, Xiangyu Yan, Kuo Zhang, Da Liu, Xiaojuan Cui, Yaodong Yang, Muhan Zhang, Chunxia Cao, Jingjia Wang, Xuliang Wang, Jun Gao, Yuan-Geng-Shuo Wang, Jia ming Ji, Zifeng Qiu, Muzi Li, Cheng Qian, Tianze Guo, Shuangquan Ma, Zeying Wang, and 6 others. 2024. The application of large language models in medicine: A scoping review. *iScience*, 27(5):109713.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.

Ricardo Rei, Craig Stewart, Ana C Farinha, and Alon Lavie. 2020. COMET: A neural framework for MT evaluation. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 2685–2702, Online. Association for Computational Linguistics.

William Rojas-Carabali, Rajdeep Agrawal, Laura Gutierrez-Sinisterra, Sally L. Baxter, Carlos Cifuentes-González, Yap Chun Wei, John Abisheganaden, Palvannan Kannapiran, Sunny Wong, Bernett Lee, Alejandra de-la Torre, and Rupesh Agrawal. 2024. Natural language processing in medicine and ophthalmology: A review for the 21st-century clinician. *Asia-Pacific Journal of Ophthalmology*, 13(4):100084.

Karan Singhal, Shekoofeh Azizi, Tho Tu, and et al. 2023. Large language models encode clinical knowledge. *Nature*, 620:172–180.

Dattatray Takale. 2024. A study of natural language processing in healthcare industries.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, and 41 others. 2025. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*.

An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, and 40 others. 2024a. Qwen2 technical report. *arXiv preprint arXiv:2407.10671*.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, and 22 others. 2024b. Qwen2.5 technical report. *arXiv preprint arXiv:2412.15115*.

Tianyi Zhang*, Varsha Kishore*, Felix Wu*, Kilian Q. Weinberger, and Yoav Artzi. 2020. Bertscore: Evaluating text generation with bert. In *International Conference on Learning Representations*.

Binggui Zhou, Guanghua Yang, Zheng Shi, and Shaodan Ma. 2024. Natural language processing for smart healthcare. *IEEE Reviews in Biomedical Engineering*, 17:4–18.

A Evaluation Metrics

A.1 Question Answering

Table 3: QA metrics by language (higher is better).

Language	F1	Exact Match	ROUGE-1			ROUGE-2			ROUGE-L			BERTScore			COMET
			Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Score
Marathi	0.284	0.000	0.001	0.001	0.001	0.000	0.000	0.000	0.001	0.001	0.001	0.508	0.503	0.505	0.278
Kannada	0.492	0.000	0.012	0.010	0.011	0.000	0.000	0.000	0.012	0.010	0.011	0.855	0.835	0.845	0.475
Gujarati	0.364	0.000	0.003	0.003	0.003	0.002	0.002	0.002	0.003	0.003	0.003	0.408	0.403	0.405	0.358
English	0.623	0.000	0.329	0.390	0.343	0.064	0.076	0.067	0.194	0.229	0.202	0.857	0.857	0.857	0.703
Telugu	0.656	0.000	0.084	0.058	0.064	0.009	0.005	0.006	0.084	0.058	0.064	0.853	0.837	0.845	0.507
Tamil	0.514	0.000	0.005	0.004	0.004	0.000	0.000	0.000	0.005	0.004	0.004	0.847	0.830	0.838	0.529
Bangla	0.203	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.001	0.217	0.213	0.215	0.285
Hindi	0.462	0.000	0.007	0.007	0.007	0.000	0.000	0.000	0.007	0.007	0.007	0.656	0.651	0.653	0.370
Assamese	0.124	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.309	0.303	0.306	0.302

A.2 Summarization

Table 4: Summarization metrics by language (higher is better).

Language	F1	Exact Match	ROUGE-1			ROUGE-2			ROUGE-L			BERTScore			COMET
			Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Score
Marathi	0.562	0.000	0.491	0.343	0.357	0.110	0.075	0.079	0.227	0.171	0.162	0.795	0.782	0.788	0.640
Kannada	0.622	0.000	0.557	0.285	0.354	0.120	0.060	0.076	0.260	0.127	0.158	0.783	0.785	0.783	0.661
Gujarati	0.322	0.000	0.546	0.269	0.320	0.132	0.062	0.074	0.273	0.124	0.145	0.769	0.785	0.776	0.624
English	0.239	0.000	0.446	0.414	0.356	0.116	0.104	0.092	0.192	0.200	0.152	0.819	0.812	0.815	0.707
Telugu	0.355	0.000	0.454	0.231	0.264	0.103	0.051	0.057	0.233	0.111	0.123	0.681	0.710	0.694	0.575
Tamil	0.416	0.000	0.523	0.224	0.283	0.119	0.050	0.063	0.271	0.107	0.137	0.710	0.771	0.738	0.590
Bangla	0.277	0.000	0.462	0.276	0.300	0.101	0.061	0.066	0.224	0.141	0.141	0.778	0.781	0.778	0.638
Hindi	0.412	0.000	0.539	0.310	0.373	0.118	0.066	0.081	0.234	0.134	0.158	0.806	0.802	0.803	0.701
Assamese	0.451	0.000	0.514	0.308	0.357	0.115	0.068	0.079	0.242	0.139	0.160	0.774	0.795	0.783	0.639

A.3 Information Extraction

Table 5: Information Extraction metrics by language (higher is better).

Language	F1	Exact Match	ROUGE-1			ROUGE-2			ROUGE-L			BERTScore			COMET
			Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1	Score
Marathi	0.180	0.002	0.010	0.006	0.007	0.003	0.001	0.001	0.009	0.006	0.006	0.931	0.899	0.914	0.519
Kannada	0.279	0.041	0.145	0.096	0.103	0.036	0.014	0.018	0.138	0.092	0.098	0.856	0.851	0.852	0.546
Gujarati	0.190	0.014	0.030	0.026	0.025	0.004	0.003	0.003	0.028	0.025	0.024	0.908	0.879	0.892	0.525
English	0.258	0.039	0.130	0.088	0.096	0.046	0.028	0.030	0.120	0.083	0.090	0.886	0.867	0.876	0.535
Telugu	0.212	0.024	0.061	0.047	0.048	0.008	0.004	0.004	0.058	0.045	0.046	0.881	0.865	0.871	0.534
Tamil	0.196	0.010	0.049	0.024	0.028	0.022	0.005	0.007	0.047	0.023	0.027	0.903	0.880	0.890	0.529
Bangla	0.173	0.001	0.005	0.002	0.003	0.001	0.000	0.000	0.004	0.002	0.003	0.935	0.895	0.913	0.516
Hindi	0.252	0.033	0.122	0.078	0.084	0.037	0.015	0.017	0.113	0.073	0.078	0.868	0.854	0.860	0.531
Assamese	0.213	0.019	0.047	0.036	0.037	0.008	0.006	0.006	0.044	0.034	0.035	0.906	0.884	0.894	0.531