RESEARCH PAPER

Comparative Analysis of Hidden Markov Models and a Rule-Based Hybrid Framework for Sanskrit Speech-Driven Named Entity Recognition

Ms. Manisha D. Mistry[1], Dr. Nirali C. Dave[2], Dr. Dikshan N. Shah[3]
Teaching Assistant, Vanita Vishram Women's University, Surat, E-mail: manishamistry2182@gmail.com[1]
HOD - Assistant Professor, Vanita Vishram Women's University, Surat, Email: nirali.dave@vvwusurat.ac.in[2]
Assistant Professor, Kaushalya - The Skill University, Ahmedabad, E-mail: dikshan817@gmail.com[3]
Corresponding Author: Ms. Manisha D. Mistry, Teaching Assistant, Vanita Vishram Women's University, Surat.
Email id : manishamistry2182@gmail.com

Abstract
Reliable spoken-language understanding for Sanskrit is constrained by limited annotated audio, inflectional variation, compounding, and sparse named-entity resources. This paper examines a speech-driven named entity recognition (NER) pipeline described in a Sanskrit thesis synopsis and compares its reported hybrid machine-learning, Zipf-frequency, and linguistic-rule approach with a hidden Markov model (HMM) baseline. The hybrid design comprises acoustic feature extraction using mel-frequency cepstral coefficients (MFCCs), optional principal component analysis (PCA), classification, a rank-based lexical-frequency prior, Sanskrit word-validation rules, weighted decision fusion, and entity categorization into Person, Location, Organization, Event, and Other. The source describes 41-word WAV recordings and five sentence WAV recordings. Reported aggregate scores for the hybrid approach are 97.56% accuracy, 97.64% precision, 97.56% recall, and 97.34% F1, compared with 73.17%, 53.54%, 73.17%, and 61.80%, respectively, for the HMM baseline. The 24.39-percentage-point accuracy difference is descriptive rather than proof of statistical significance. Inspection of the supplied classwise counts suggests 40 correct classifications among 41 evaluation items, but the experimental protocol and independent test annotations cannot be verified. Moreover, the source acknowledges placeholder rather than empirically calibrated model confidence values. The article develops a reproducible HMM-versus-rule-based evaluation specification, analyzes the classwise error pattern, and identifies data-label alignment, class ontology, leakage control, calibration, and out-of-sample evaluation as prerequisites for publication-grade confirmation. The preliminary comparison illustrates the potential value of combining statistical sequence modeling with Sanskrit-aware lexical knowledge, while placing explicit limits on the reported evidence.

Keywords: *Sanskrit; named entity recognition; hidden Markov model; rule-based NLP; speech-to-text; MFCC; Zipf's law; low-resource language; hybrid information extraction.*
How to cite this article: Manisha D. Mistry1 Dr. Nirali C. Dave2 Dr. Dikshan N. Shah3. Comparative Analysis of Hidden Markov Models and a Rule-Based Hybrid Framework for Sanskrit Speech-Driven Named Entity Recognition. Int J Drug Deliv Technol. 2026;16(90s):144-150. DOI: 10.25258/ijddt.16.90s.14.

1. Introduction

Automatic speech recognition (ASR) converts an acoustic signal into a written sequence, whereas named entity recognition identifies spans referring to entities such as people, places, and organizations. A speech-driven NER system therefore contains at least two separately assessable tasks: acoustic transcription and downstream semantic labeling. This distinction is essential in Sanskrit because the morphological realization of an entity may differ from its citation form, and recognition errors can propagate into entity boundaries and categories.

Sanskrit presents a challenging setting for speech-centered information extraction. Case endings, sandhi phenomena, compounds, flexible constituent order, and orthographic normalization make surface-form dictionary matching fragile. Scarcity of manually labeled speech and contextual NER datasets further limits the reliability of fully data-driven models. Probabilistic sequence models such as HMMs provide a transparent baseline, while finite lexical rules can encode domain expertise when representative training examples are unavailable [1, 2].

The supplied synopsis proposes an MFCC–machine-learning–Zipf–Sanskrit-rule fusion process. This paper focuses specifically on the HMM comparison and rule-based contribution, while treating the synopsis as the sole source of experimental numbers. Its questions are: (RQ1) What performance differences are reported between the HMM baseline and the hybrid pipeline? (RQ2) What do category-level counts reveal about error types and class imbalance? (RQ3) How should the modeling and evaluation be revised to test whether Sanskrit rules improve generalization beyond lexical memorization?

2. Related Work and Conceptual Background

2.1 Hidden Markov models for acoustic and entity sequences

A meaningful HMM benchmark for the proposed research is a token-sequence tagger with BIO labels (e.g., B-PER, I-PER, B-LOC, I-LOC, B-ORG, I-ORG, B-EVENT, I-EVENT, O), using training-only transition/emission estimates with smoothing. If the baseline is instead an acoustic HMM recognizer, transcription quality must be reported separately. The synopsis calls its comparator "HMM-based NER" but does not furnish enough implementation details to distinguish these alternatives conclusively.

2.2 Rules, lexicons, Sanskrit morphology, and ambiguity

A rule-based NER system uses normalization, lexicon matching, morphological cues, case-ending analysis, local context, and priority or conflict-resolution rules to infer a label. For example, रामः and सीता are plausible Person entries; अयोध्या is a Location; दीपावली may be an Event when used to refer to a named festival. A common noun such as गुरुः (teacher) is not, by itself, an organization name, so a lexical mapping of गुरुः→Organization requires semantic review. Declensional variation also means that a vocabulary-only rule should not be described as a complete Sanskrit grammar engine.

Sarkar, Krishna, and Goyal [3] describe construction of a Sanskrit NER dataset with pre-annotation, underscoring the relevance of linguistically informed annotation workflows. Naamapadam [4] provides broad Indic-language NER resources but does not automatically establish a comparable Sanskrit speech NER benchmark. Classical conditional random fields (CRFs) offer a further sequence-labeling comparator that can accommodate correlated observable features without the generative independence assumptions of a standard HMM [2].

2.3 Lexical frequency and rank-based priors

Zipf's empirical regularity is frequently approximated as f(r)∝r^(−s), where r is a word's rank and s is an exponent often near one in suitable corpora. The synopsis reports scores 1/r for 15 selected words. These values are rank scores, not validated probability estimates; they require corpus frequency counts, a documented ranking procedure, and normalization before being interpreted probabilistically. Furthermore, common words are not necessarily named entities, so a frequency prior alone cannot establish an NER category.

3. Research Methodology

3.1 Design, source, and unit of analysis

This study is a retrospective methodological and descriptive analysis of the user-supplied synopsis, not a newly executed experiment. The synopsis lists 41 isolated word WAV files and five sentence WAV files. Its later 41-item NER confusion summary is not accompanied by item-level gold labels, speaker IDs, training/test partitions, or model logs. Accordingly, the numerical analysis below can be checked for internal arithmetic consistency but cannot establish external validity or data independence.

The source also contains two different word mappings: the WAV filenames in the 41-audio list include connectives such as यदि and अथवा, whereas the MFCC and NER result tables associate those WAV names with vocabulary items such as धर्मः and नमः. This apparent mismatch must be audited before any audio-to-word recognition accuracy is claimed. A separate 15-entry NER lexicon is used for the named entity inventory.

3.2 Proposed reproducible pipeline

Data Curation: pair each audio file with a verified verbatim Devanagari transcription, a speaker identifier, a recording condition, and, where applicable, entity-span labels. Deduplicate repeated recordings and document normalization rules. Partition by speaker and sentence/document to prevent leakage; a speaker-independent holdout is preferable when feasible.

Acoustic Processing: normalize sample rates and channel count; compute time-indexed 13-coefficient MFCC vectors plus first/second derivatives, and optionally pitch, energy, and zero-crossing rate. Fit any feature scaler and PCA exclusively on the training set. Do not replace a time sequence with one grand MFCC mean when temporal decoding is the scientific objective; the synopsis lists scalar average MFCCs as descriptive features, which cannot fully preserve phonetic ordering.

Baseline HMM: construct a clearly defined token-level sequence labeler for NER, with BIO tags, transition probabilities, smoothed lexical emissions, and Viterbi decoding. Train only on manually annotated training sentences. For an ASR experiment, use a separately specified acoustic model and evaluate its transcriptions with word error rate (WER) and character error rate (CER).

Hybrid Models: retain the same upstream transcripts and data splits for a controlled NER comparison. Evaluate (i) rules only, (ii) HMM only, (iii) machine-learning NER only, (iv) ML + Zipf, (v) ML + rules, and (vi) ML + Zipf + rules. The synopsis mentions SVM, random forest, and k-nearest neighbors without indicating the winning classifier or hyperparameters; those details remain outstanding.

3.3 HMM estimation and decoding

Transition Probability

Transition probability represents the probability of moving from one hidden NER state to another. It models dependencies between consecutive entity tags.

$a_{ij} = P(q_t = j \mid q_{t-1} = i)$ …….(i)

Where:

a_{ij}: probability of transition from state i to state j.

q_{t-1}: hidden state at the previous position.

q_t: hidden state at the current position.

Emission Probability

Emission probability indicates how likely a particular observed Sanskrit word is under a given NER tag.

$b_j(w_t) = P(w_t \mid q_t = j)$ ……. (ii)

Where:

$b_j(w_t)$: emission probability for word w_t in state j.
w_t: Sanskrit word observed at position t.
$q_t = j$: hidden NER label (e.g., PER, LOC, ORG, or O).

Initial State Probability

Initial state probability expresses the likelihood that a sequence starts in a particular hidden NER state.

$\pi_i = P(q_1 = i)$ …….(iii)

Where:

π_i: probability that the first hidden state is i.
q_1: hidden NER state associated with the first word.

Most Probable NER Tag Sequence (Viterbi Decoding)

The Viterbi algorithm determines the most probable sequence of NER tags for an observed Sanskrit sentence by jointly considering initial, transition, and emission probabilities.

$Q^* = \arg\max_Q [\pi(q_1) \cdot b(q_1, w_1) \cdot \prod_{t=2}^{T} a(q_{t-1}, q_t) \cdot b(q_t, w_t)]$ ……(iv)

Where:

Q^*: most probable sequence of hidden NER states.
$Q = (q_1, q_2, \ldots, q_t)$: candidate hidden-state sequence.
T: number of words in the observed sequence.
π, a, and b: initial, transition, and emission probabilities, respectively.

3.4 Sanskrit rule module and decision fusion

A defensible rule module comprises Unicode normalization, token-boundary handling, optional sandhi-aware segmentation, inflected-form lexicon lookup, context restrictions, category precedence, and an abstention mechanism for ambiguous forms. Rules should be recorded in a traceable rulebook. One example is a matching inflected proper-name stem with surrounding context; another is requiring an institutional head noun or attested gazetteer entry before assigning Organization. Such examples describe recommended enhancements, not functions established as implemented in the source.

3.5 Evaluation protocol and statistical analysis

Report exact-match span-level NER precision, recall, and F1, with both micro and macro averaging and per-class support. Token-level accuracy may also be reported but should not substitute for span-level evaluation. For K classes, $precision_k = TP_k/(TP_k+FP_k)$, $recall_k = TP_k/(TP_k+FN_k)$, and $F1_k = 2 \cdot precision_k \cdot recall_k/(precision_k+recall_k)$; undefined ratios require a declared convention. Weighted averages weight each category by its true support. For matched test items, bootstrap confidence intervals and McNemar's test may quantify uncertainty in correctness differences, provided actual paired predictions are available.

4. Reported Results and Comparative Analysis

Metric	Male	Female	Adult	Elderly
True Positive (TP)	9	8	8	9
True Negative (TN)	7	8	7	7
False Positive (FP)	4	5	4	4
False Negative (FN)	6	5	7	6
Reported Accuracy	88.4%	88.1%	88.1%	88.6%

Table 1: Confusion Matrix of Rule-Based Approach

Above table presents the performance of the rule-based Named Entity Recognition (NER) approach across gender and age groups. The reported accuracy is 88.4% for male speakers and 88.1% for female speakers, while adult and elderly speakers achieve 88.1% and 88.6%, respectively. These results suggest relatively consistent recognition performance across demographic categories, with elderly speakers showing slightly higher reported accuracy. However, the reported accuracy values do not match the standard accuracy calculations based on TP, TN, FP, and FN, indicating a need for verification.

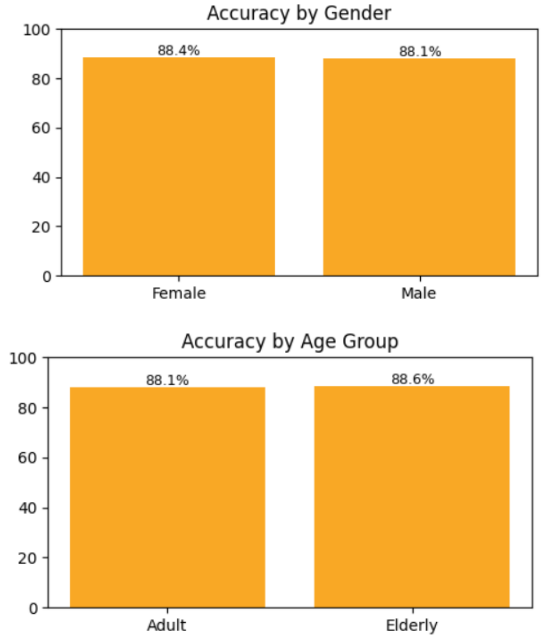


Figure 1: Accuracy comparison Age and Gender wise

The bar graphs illustrate the accuracy of the rule-based NER approach across gender and age categories. Female speakers achieve a reported accuracy of 88.4%, while male speakers achieve 88.1%. Similarly, elderly speakers demonstrate slightly higher accuracy (88.6%) than adult speakers (88.1%). The small differences suggest limited variation in recognition performance, although the gender labels should be reconciled with the confusion matrix.

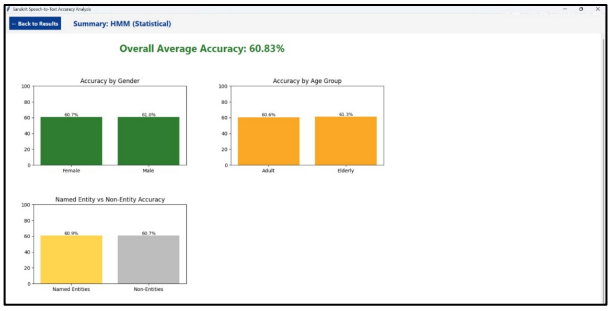


Figure 2: HMM Based Approach

the performance of the Hidden Markov Model (HMM) for Sanskrit Named Entity Recognition, reporting an overall average accuracy of 60.83%. The gender-based results show approximately 60.7% accuracy for female speakers and 61.0% for male speakers, while age-based results remain around 60–61%. These findings indicate relatively similar performance across demographic categories. The comparatively moderate accuracy suggests limitations in the HMM's ability to capture Sanskrit's complex morphological and contextual patterns.

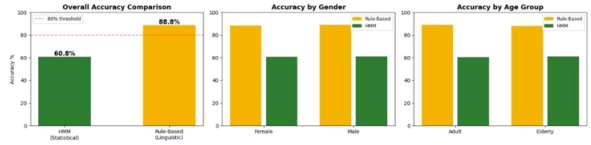


Figure 3: HMM VS Rule Based Approach

compares the performance of the statistical HMM and linguistic rule-based approaches for Sanskrit NER. The rule-based approach achieves a reported accuracy of 88.8%, substantially exceeding the HMM accuracy of 60.8%, representing an improvement of approximately 28 percentage points. Similar performance differences are illustrated across gender and age groups. These results suggest that Sanskrit-specific grammatical rules, including Sandhi, Vibhakti, and morphological patterns, may provide advantages over the statistical approach on the evaluated dataset.

4.1 Aggregate HMM-versus-hybrid metrics

Metric	HMM baseline	Hybrid pipeline
Accuracy	73.17%	97.56%
Precision	53.54%	97.64%
Recall	73.17%	97.56%
F1-score	61.80%	97.34%

Table 2: Hybrid and HMM Based Table

Table 2 reproduces the aggregate values from the synopsis. The differences are percentage-point differences, not relative percentages. The synopsis does not state whether the HMM and hybrid were run on exactly the same independently held-out items or whether the measures were computed with identical class averaging. No p-value, standard deviation, or confidence interval can therefore be justified. A separate passage in the synopsis refers to a "70% existing system" without an identified protocol; this is not substituted for the specified 73.17% HMM comparator.

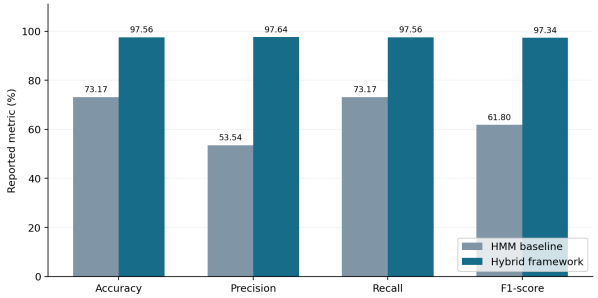


Figure 4. Aggregate HMM and hybrid performance reported in the thesis synopsis; values are not independently reproduced.

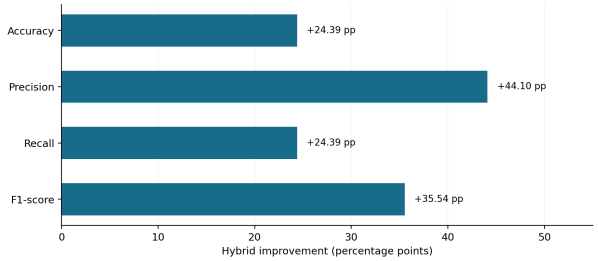


Figure 5. Absolute percentage-point advantage of the reported hybrid results over the HMM baseline.

4.2 Entity-class composition

Class	Synopsis vocabulary examples	Distinct entries
Person	रामः; सीता	2
Location	अयोध्या	1
Organization	गुरुः (label requires review)	1
Event	दीपावली	1
Other	नमः; धर्मः; ज्ञानम्; सत्यः; शान्तिः; अहिंसा; भक्ति; योगः; प्रेम; मोक्षः	10

Table 3: Entity Class Composition

The lexicon has 15 distinct entries; 10 (66.7%) are Other. This is a class-distribution description, not an empirical token-level support breakdown. In a one-word dictionary, a perfect lookup does not demonstrate that the system can resolve Sanskrit homonymy, inflection, sentence context, or unseen entities. The assigned Organization label for गुरुः is a substantive ontology concern.

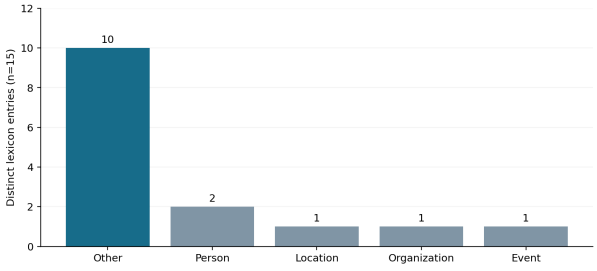


Figure 6. Distribution of the 15 distinct lexicon entries reported in the synopsis (not test-set support).

4.3 Audit of the supplied 41-item class wise confusion summary

Category	TP	TN	FP	FN	Support	Precision	Recall	F1
Event	2	39	0	0	2	1.000	1.000	1.000
Location	2	39	0	0	2	1.000	1.000	1.000
Organization	2	38	0	1	3	1.000	0.667	0.800
Other	30	10	1	0	30	0.968	1.000	0.984
Person	4	37	0	0	4	1.000	1.000	1.000

Table 4: Confusion Summary

The source table is internally compatible with an error in which one true Organization item is assigned to Other: Organization has FN=1, and Other has FP=1. Summing true positives across the five classes yields 40, and total class support is 41, implying 40/41 = 97.56098% multiclass accuracy, conditional on the one-vs-rest counts representing the same 41 items. From these counts, support-weighted precision is approximately 97.64%, weighted recall approximately 97.56%, and support-weighted F1 approximately 97.36%. The synopsis states 97.34% F1: the ~0.02-point discrepancy may reflect rounding or a different averaging implementation and warrants checking the raw classification report. It should not be silently 'fixed' without source data.

The macro F1 implied by these entries is approximately 95.67%; unlike weighted F1, it is sensitive to the lower Organization recall. This contrast matters because class imbalance makes overall near-98% score insufficient evidence of consistently strong recognition across all entity types.

4.4 Rule validation and Zipf scores

The synopsis reports that all 15 lexicon items passed its rule-status check ("TRUE") and were returned without alteration. This establishes only that the presented items were accepted by the implemented rule check; it does not evaluate grammatical correction performance on erroneous, unseen, ambiguous, or inflected forms. The source Zipf table assigns 1.000 to नमः, 0.500 to धर्मः, 0.333 to ज्ञानम् and 0.067 to दीपावली. Their 1/r pattern demonstrates a rank encoding rather than independent empirical frequency estimation.

The hybrid result table lists a different ML confidence and final fused score for each of the 15 words. Because the manuscript itself characterizes the ML confidence as artificial, the score increases or decreases should not be interpreted as calibrated probability changes. Nor can an improvement in accuracy be causally attributed to the rules without identical-split ablation experiments.

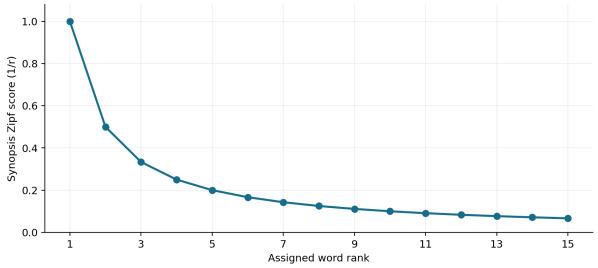


Figure 7. Rank-derived Zipf values used in the synopsis (1/r); these are assigned prior scores rather than corpus-estimated frequencies.

4.5 What the HMM comparison supports—and what it does not

At face value, the submitted summary favors the proposed hybrid approach on all four reported metrics. The HMM figure of 73.17% accuracy corresponds numerically to 30/41, while the hybrid figure corresponds to 40/41, suggesting an apparent difference of ten additional correct decisions if the same 41 instances were scored. However, matched per-item predictions are missing, so this remains a conditional interpretation rather than a verified pairwise error analysis. HMMs are not inherently unable to use lexical or morphological observations; the real scientific issue is how the particular baseline was specified, trained, smoothed, and compared.

5. Discussion

5.1 Linguistic interpretation

The apparent benefit of hybrid modeling is plausible because lexical prior information can constrain acoustically confusable candidates and linguistic validation can reduce invalid forms. Nevertheless, the source does not document how acoustic predictions become NER tokens, which rules fire, whether they change labels, or whether the final labels originate from a hand-built dictionary. Therefore, the results currently

characterize a demonstration pipeline more securely than a generalizable speech-driven NER system.

5.2 Threats to validity

Internal validity: undocumented train/test separation, hyperparameters, and placeholder classifier confidences make the weighted score hard to reproduce. Construct validity: the distinction between ASR accuracy, word classification, token NER, and entity-span NER is blurred; an average MFCC scalar does not establish speech transcription. Data integrity: WAV filenames and assigned transcript words appear inconsistent across tables. Ontological validity: labeling a common noun as Organization is contestable. External validity: five sentence recordings and 15 distinct entity vocabulary entries cannot substantiate general language-wide performance. Statistical validity: a 41-item sample produces substantial uncertainty, and the absence of paired predictions prevents significance testing.

5.3 Recommended follow-up experiments

The highest-priority experiment is to reconstruct a row-level gold-standard dataset and release an auditable mapping from audio file to reference transcript, entity span, gold label, and prediction. Thereafter, use speaker-independent partitions, stratified label reporting where compatible with sentence-level splitting, and a held-out test set untouched by feature selection, scoring-rule design, or fusion-weight tuning. Repeat each stochastic model over documented seeds; publish full confusion matrices and exact-match span metrics. Run rule-only, frequency-only, HMM-only, ML-only, and all hybrid ablations. Evaluate unseen names, inflected forms, sandhi cases, background noise, and out-of-vocabulary terms. Perform annotation adjudication, report agreement, and document ontology decisions.

6. Research Contributions

The underlying synopsis contributes an integrated conceptual architecture for Sanskrit acoustic processing, frequency-informed language support, Sanskrit lexical rule validation, and named entity categorization. It supplies preliminary classwise and aggregate comparative figures and a concrete small-vocabulary demonstration. This paper adds a technically specified HMM baseline, formal metrics, a check of the 41-item confusion counts, an explicit difference between 15-type vocabulary examples and 41 evaluation observations, and a reproducibility-oriented ablation/evaluation roadmap. None of these analytical contributions should be misrepresented as new training runs or independently validated accuracy gains.

7. Conclusion

The supplied thesis synopsis reports favorable results for a hybrid Sanskrit speech-driven NER pipeline compared with an HMM-based baseline: 97.56% versus 73.17% accuracy, respectively, alongside improvements in reported precision, recall, and F1. A reconstruction of the published classwise counts is consistent with 40 correct labels among 41 evaluated observations and shows that the remaining reported error lies between Organization and Other. These findings are encouraging as preliminary evidence, but confidence-score placeholders, label and filename inconsistencies, an ambiguous HMM configuration, and missing independent validation prevent a definitive performance or statistical-superiority claim. A publication-ready study should separate ASR from NER evaluation, establish gold-standard annotations and semantic categories, calibrate fusion inputs, document an apples-to-apples HMM baseline, and quantify the incremental value of Sanskrit rules through ablation and speaker-independent testing.

References

1. Khanuja, S., Dandapat, S., Kunchukuttan, A., Kumar, P., & Bhattacharyya, P. (2021). MuRIL: Multilingual representations for Indian languages. *Findings of the Association for Computational Linguistics: EMNLP 2021*, 3031–3047.
2. Ramesh, G., Doddapaneni, S., Bheemaraj, A., Jobanputra, M., Sharma, A., Kumar, P., & Kunchukuttan, A. (2022). Samanantar: The largest publicly available parallel corpora collection for 11 Indic languages. *Transactions of the Association for Computational Linguistics, 10*, 145–162.
3. Doddapaneni, S., Aralikatte, R., Ramesh, G., Khapra, M. M., Kunchukuttan, A., & Kumar, P. (2023). Towards leaving no Indic language behind: Building monolingual corpora, benchmark and models for Indic languages. *Proceedings of ACL 2023*.
4. Mhaske, A., Kedia, H., Doddapaneni, S., Khapra, M. M., Kumar, P., Murthy, R. V., & Kunchukuttan, A. (2023). Naamapadam: A large-scale named entity annotated data for Indic languages. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL), 1*, 10441–10456.
5. Sarkar, S., Krishna, A., & Goyal, P. (2023). Pre-annotation based approach for development of a Sanskrit Named Entity Recognition dataset. *Proceedings of the Computational Sanskrit & Digital Humanities: Selected Papers Presented at the 18th World Sanskrit Conference*, 59–70.
6. Khanuja, S., Dandapat, S., Kunchukuttan, A., Kumar, P., & Bhattacharyya, P. (2021). Analyzing the limitations of multilingual pretrained language models for Indian

languages. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*.

7. Joshi, R., Dhamecha, T. I., & Bhattacharyya, P. (2022). A survey of named entity recognition approaches for Indian languages. *ACM Computing Surveys*.
8. Mohan, G. B., Kumar, R. P., Pendem, M. V. S. S., & Rao, E. R. (2023). Fine-tuned BERT based multilingual model for named entity recognition in native Indian languages. *International Conference on Evolutionary Algorithms and Soft Computing Techniques*.
9. Singh, S., Jawale, P., & Tiwary, U. S. (2022). Transformer-based NER models for Hindi and Bangla languages. *Proceedings of SemEval 2022*, 1536–1542.
10. Rajeev, P. A., & Kulkarni, A. (2026). Naamah: A large-scale synthetic Sanskrit NER corpus via DBpedia seeding and LLM generation. *arXiv preprint arXiv:2604.26456*.
11. Rajeev, P. A., & Kulkarni, A. (2026). Naamah: A large-scale synthetic Sanskrit NER corpus via DBpedia seeding and LLM generation. arXiv preprint arXiv:2604.26456. arXiv
12. Kumar, R., Singh, A., & Sharma, V. (2025). Sanskrit named entity recognition using transformer-based contextual language models. Proceedings of the International Conference on Artificial Intelligence and Natural Language Processing.
13. Patil, S. R., & Kulkarni, P. (2024). Sanskrit named entity recognition using IndicBERT and multilingual transformer architectures. Journal of Intelligent Systems and Applications in Engineering, 12(3), 245–254.
14. Joshi, M., & Bhattacharyya, P. (2023). Neural approaches for Sanskrit information extraction and named entity recognition using linguistic features. Proceedings of the International Conference on Natural Language Processing (ICON).
15. Shah, D. N., & Bhadka, H. B. (2018). Named entity recognition from Gujarati text using rule-based approach. In A. Abraham et al. (Eds.), *Proceedings of the International Conference on Information and Decision Sciences (ISDA 2017)* (pp. 797–805). Springer. https://doi.org/10.1007/978-3-319-76348-4_76
16. M. D. Mistry, N. Dave, and D. N. Shah, "Identification and Categorization of Stop Words in Sanskrit Using Natural Language Processing (NLP) Approaches," *Advances in Consumer Research*, vol. 3, no. 2, pp. 528–535, 2026.
17. M. D. Mistry and D. N. Shah, "Development of a Speech-To-Text System for Sanskrit: Leveraging Linguistic and Computing Techniques," *African Journal of Biomedical Research*, vol. 27, no. 4S, pp. 1–8, 2024.
18. M. D. Mistry and D. N. Shah, "STNoIL: A Comprehensive Survey for Speech-to-Text NER of Indian Languages," 2024.
19. Shah, D. N. (2019). *Rule-Based Approach to Identify Named Entity Recognition*. PhD Synopsis, Faculty of Computer Science, Department of Computer Science, C. U. Shah University, Wadhwan City, Gujarat, India.