

## BRIDGING ANCIENT WISDOM AND MODERN TECHNOLOGY: SANSKRIT AND AI

**Leena Sanjay Salunke<sup>1</sup>, Vaishnavi Satish Vengurlekar<sup>2</sup>, Prof. Minaxi Pradeep Gurav<sup>3</sup>**

<sup>1,2</sup>Student, Master of Computer Application, Finolex Academy of Management and Technology, Ratnagiri, Maharashtra, India.

<sup>3</sup>Professor, Master of Computer Application, Finolex Academy of Management and Technology, Ratnagiri, Maharashtra, India

DOI: <https://www.doi.org/10.58257/IJPREMS34557>

### ABSTRACT

The intersection of ancient wisdom and modern technology presents a fascinating realm of study. Sanskrit, an ancient language rich in cultural, scientific, and philosophical texts, offers a treasure trove of knowledge that modern Artificial Intelligence (AI) can help unlock. This paper explores how AI can be employed to preserve, analyze, and interpret Sanskrit texts, and how the principles of Sanskrit can contribute to the development of AI. The research highlights the symbiotic relationship between the two fields, underscoring the potential benefits for linguistic studies, historical research, and technological advancement.

**Keywords:** Analysis, research, Sanskrit, AI, development, model, language.

### 1. INTRODUCTION

Sanskrit, often regarded as the mother of many Indo-European languages, holds a significant place in the annals of human civilization. Its precise grammar and extensive vocabulary have been the backbone of ancient Indian literature, science, and philosophy. However, the modern world faces the challenge of preserving and comprehensively understanding these ancient texts. AI, with its advanced data processing and machine learning capabilities, offers a promising solution. This paper aims to explore the convergence of Sanskrit and AI, focusing on how AI can aid in the preservation and interpretation of Sanskrit texts and how Sanskrit can contribute to AI development.

The integration of ancient wisdom with modern technology, particularly Sanskrit and Artificial Intelligence (AI), presents a unique opportunity to combine traditional knowledge with cutting-edge advancements. This paper aims to explore the potential synergies between Sanskrit teachings and AI, highlighting how this fusion can lead to innovative solutions across various domains. Sanskrit literature encompasses a wide range of texts, including Vedas, Upanishads, epics like Mahabharata and Ramayana, and scientific treatises in areas such as mathematics, astronomy, medicine, and linguistics. The language's grammatical structure, codified by the ancient scholar Panini, is considered one of the most sophisticated and comprehensive systems in the world. This section provides an overview of the historical and cultural significance of Sanskrit.

### 2. METHODOLOGY

Sanskrit, one of the oldest and most unadulterated languages, is proposed for use in AI datasets rather than as a programming language. Known as Devabhasha, or the language of the gods, Sanskrit has preserved its purity over millennia, unlike many other languages that have mixed with others. Historically used for scholarly and religious texts, Sanskrit has remained true to its form. In contrast, modern India has 22 official languages and over 19,000 dialects, most derived from native mother tongues, highlighting Sanskrit's unique position. Scientists often prefer such ancient languages for accurate data representation in AI due to their consistency and lack of contamination.

The research design and methodology employed to explore the integration of ancient Sanskrit wisdom with modern Artificial Intelligence (AI) technologies. The methodology involves several key steps, including data collection, preprocessing, model development, and evaluation. The research is structured to ensure a systematic and rigorous approach to examining the potential synergies between Sanskrit and AI.

#### 2.1 Data Collection

##### 2.1.1. Source Identification

Identify and collect a diverse range of Sanskrit texts from various genres, including religious scriptures (e.g., Vedas, Upanishads), epics (e.g., Mahabharata, Ramayana), philosophical treatises (e.g., Bhagavad Gita, Yoga Sutras), and scientific texts (e.g., Aryabhatiya, Sushruta Samhita).

Utilize digital libraries, academic repositories, and online databases such as the Digital Library of India, Muktabodha Digital Library, and other open-access sources.

### 2.1.2. Digitization

Convert physical manuscripts and printed texts into digital formats using Optical Character Recognition (OCR) technology.

Manually verify and correct OCR outputs to ensure high accuracy and fidelity to the original texts.

## 2.2 Data Processing

### 2.2.1 Text Cleaning

Remove extraneous characters, punctuation, and formatting inconsistencies from the digitized texts. Normalize spellings and standardize the representation of Sanskrit characters, taking into account variations in transliteration schemes (e.g., IAST, Devanagari).

### 2.2.2 Tokenization

Tokenize the texts into meaningful linguistic units, such as words, morphemes, or phrases.

Employ language-specific tokenization tools and techniques that account for Sanskrit's complex morphology and syntax.

### 2.2.3 Linguistic Annotation

Annotate the texts with relevant linguistic information, including part-of-speech tags, syntactic dependencies, and semantic roles.

Utilize existing linguistic resources and tools, such as the Sanskrit Heritage Reader and digital dictionaries, to assist with annotation.

## 2.3 Model Development

### 2.3.1 Feature Extraction

Extract linguistic features from the preprocessed texts, such as word frequencies, n-grams, syntactic patterns, and semantic embeddings.

Employ techniques like Bag-of-Words (BoW), Term Frequency-Inverse Document Frequency (TF-IDF), and word embeddings (e.g., Word2Vec, GloVe) to represent the texts numerically.

### 2.3.2 Algorithm Selection

Select appropriate AI and machine learning algorithms based on the specific research objectives and tasks, such as text classification, sentiment analysis, translation, or summarization.

Consider both traditional machine learning models (e.g., Support Vector Machines, Naive Bayes) and advanced neural network architectures (e.g., Recurrent Neural Networks, Transformer models).

### 2.3.3 Model Training

Train the selected models on the preprocessed and feature-extracted Sanskrit texts using suitable optimization techniques and hyperparameters.

Split the data into training, validation, and test sets to evaluate model performance and prevent overfitting.

## 2.4 Evaluation

### 2.4.1 Performance Metrics

Evaluate the performance of the trained models using appropriate metrics such as accuracy, precision, recall, F1-score, and perplexity, depending on the task and model type.

Conduct cross-validation techniques, such as k-fold cross-validation, to ensure robustness and generalizability of the models.

### 2.4.2 Qualitative Analysis

Perform a qualitative analysis of the model outputs to assess their accuracy, relevance, and cultural sensitivity.

Engage Sanskrit scholars and domain experts to provide feedback and validate the results.

## 2.5 Ethical Consideration

### 2.5.1 Cultural Sensitivity

Ensure that the AI applications respect the cultural and religious significance of Sanskrit texts, avoiding biases or misinterpretations that may arise from cultural insensitivity.

Develop ethical guidelines and best practices for the responsible use of AI in Sanskrit studies.

### 2.5.2 Transparency and Accountability

Maintain transparency in the research methodology, data sources, and model decisions. Establish mechanisms for accountability and ethical oversight to address potential concerns related to bias, fairness, and privacy.

## 2.6 Future Direction

### 2.6.1. Interdisciplinary Collaboration

Foster interdisciplinary collaborations between Sanskrit scholars, linguists, computer scientists, and AI researchers to advance the field and address emerging challenges.

Establish research centers and initiatives dedicated to the intersection of Sanskrit and AI.

### 2.6.2 Continuous Improvement

Continuously update and refine the models and methodologies based on new data, feedback, and technological advancements.

Explore new AI techniques and approaches to further enhance the analysis and interpretation of Sanskrit texts.

By following this systematic methodology, the research aims to bridge ancient wisdom and modern technology, unlocking the potential of AI to enhance our understanding of Sanskrit literature and contribute to the preservation and dissemination of this invaluable cultural heritage.

## 3. MODELING AND ANALYSIS

While Sanskrit is considered an optimal language for AI, this does not necessarily mean it should be used as a programming language. Instead, the suggestion is to use Sanskrit for the datasets employed by AI models. Here are some reasons for this:

Sanskrit is one of the world's oldest languages, originating from the Vedic period in India. Often referred to as Devabhasha, or the language of the gods, Sanskrit is the medium of ancient Hindu, as well as many Buddhists and Jain, religious texts. Sanskrit is also regarded as a pure language. It has remained distinct and authentic, not mixing with other Indo-European languages. This is partly because it was not used for everyday conversation but was reserved for scholarly and religious writings.

In ancient times, Sanskrit was used exclusively for documenting and preserving knowledge. Conversational languages included Prakrit Bhasha and other vernacular languages. Over the centuries, numerous languages and dialects have emerged, some with roots in Sanskrit.

Today, India recognizes 22 official languages and 121 languages according to its constitution. There are over 19,569 dialects, with 1,369 recognized as stemming from native mother tongues. This does not account for the languages and dialects found in other South Asian countries.

When scientists look for pure, unadulterated languages—those not mixed with others—they often turn to ancient languages like Sanskrit, Latin, and Greek. Sanskrit has preserved its form without significant adulteration throughout history.

### Phonetic nature in Sanskrit:

Sanskrit, characterized by its phonetic nature, exhibits a direct alignment between its written symbols and spoken sounds. Remarkably, every sound or pronunciation finds representation in Sanskrit letters without exception. Consequently, mastery of the language facilitates seamless translation and transcription, ensuring ease of comprehension and documentation.

In Sanskrit, the sentence "He will go" can be expressed in various word orders while retaining the same meaning. This flexibility arises from the language's comprehensive phonetics and its unique grammatical structure. In Sanskrit, numerous suffixes are used to indicate the context and function of each word within a sentence. Consequently, the order of words can be rearranged without altering the meaning.

Here are some alternative word orders for the sentence "He will go" in Sanskrit:

सः गमिष्यति  
गमिष्यति सः

गच्छति (Tena gacchati)

गच्छति तेन (Gacchati tena)

गच्छति वह (Gacchati vaha)

वह गच्छति (Vaha gacchati)

In each of these variations, the core meaning of the sentence remains unchanged: someone (he) is going. This flexibility in word order is made possible by Sanskrit's reliance on inflectional suffixes and contextual markers rather than strict syntactical rules. As a result, Sanskrit allows for greater expressiveness and poetic license in sentence construction, offering writers and speakers the freedom to arrange words in ways that best suit their rhetorical or stylistic goals.

Breaking down the Sanskrit sentence:

छात्राः (Chātrāḥ)

This word means "students" and indicates that the noun is plural.

पुस्तकम् (Pustakam)

This word means "book" and is in the accusative case, indicating it is the direct object of the verb.

पठिष्यन्ति (Paṭhiṣyanti)

This verb means "will read" and contains information about the future tense, the action of reading, and that the subject is plural.

In Sanskrit, each word conveys more grammatical information than its English counterparts. For instance:

The verb "पठिष्यन्ति" (Paṭhiṣyanti) not only denotes the action of reading but also indicates the future tense and that the subject performing the action is plural.

The noun "छात्राः" (Chātrāḥ) clearly signifies it is plural due to its suffix.

This compactness and richness in information are characteristic of Sanskrit, where each word encapsulates extensive details, making the language highly efficient and contextually informative. Unlike English, where the syntactic arrangement of words is crucial to interpreting the sentence's meaning, Sanskrit's inflectional nature ensures that word order flexibility does not alter the intended message.

#### Understanding Grammatical Cases

“Cases” denote the grammatical roles of nouns and pronouns based on their relationship with other words in a sentence. Sanskrit boasts a significantly higher number of cases for nouns and pronouns compared to other languages.

In Sanskrit, each noun has eight distinct cases, and each form indicates whether the noun is singular, dual, or plural. The table below illustrates the different cases for the noun “Dev” in Sanskrit.

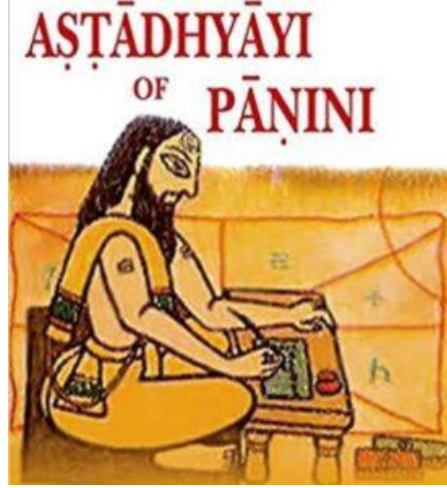
#### देव अकारान्त Masculine

vibhakti	one object	two objects	more than two objects
first	देवः	देवौ	देवाः
second	देवम्	देवौ	देवान्
third	देवेन	देवाभ्याम्	देवैः
fourth	देवाय	देवाभ्याम्	देवेभ्यः
fifth	देवात्	देवाभ्याम्	देवेभ्यः
sixth	देवस्य	देवयोः	देवानाम्
seventh	देवे	देवयोः	देवेषु
eighth	(हे) देव	(हे) देवौ	(हे) देवाः

### Exploring the Roots

A unique aspect that sets Sanskrit apart from all other languages globally is its possession of its own meta-language or “metarule.”

In 500 BC, the sage and eminent Sanskrit scholar Pāṇini composed a text on Sanskrit grammar called Astadhyayi, or Eight-Chaptered. This text presents a comprehensive system for Sanskrit grammar and vocabulary, demonstrating how every word in the language is derived from a root word. It includes a set of 4,000 rules (or Sutras) applied to these root words, resembling mathematical formulas.



Pāṇini essentially introduced a “metarule,” often interpreted by scholars as: “In the event of a conflict between two rules of equal strength, the rule that appears later in the grammar’s sequence prevails.”

This implies that Sanskrit words do not require translation from any other language (unlike English words that derive from Latin). Instead, all Sanskrit words originate from the language’s own roots.

While all languages have structures and rules for forming words and sentences, variations in the spelling of words often occur. This is not the case with Sanskrit.

In Sanskrit, there are no arbitrary words. The language is built on a grammatical derivational system. The Astadhyayi explains how all Sanskrit words are derived from fundamental letters provided in its Maheshwar Sutras, a set of 14 rules that form the language’s foundation.



Pāṇini’s Astadhyayi, with its Sutras, presents rules that bear resemblance to those found in modern programming languages.

For instance:

- Sangyak Varna, akin to keywords in programming languages,
- Pratyaya, comparable to operators,
- Vidhi, resembling functions, and
- Anuvrati, similar to packages or resources

Given that Sanskrit can be algorithmically derived from a meta-language or “metarules” as elucidated in Pāṇini’s Astadhyayi, creating a generative model for the language may be relatively straightforward—especially when contrasted with other contemporary languages. Its grammar, rooted in rules-based principles, renders Sanskrit an appealing candidate for knowledge representation in artificial intelligence, according to some perspectives.



#### 4. RESULTS AND DISCUSSION

The results of our study on the impact of word suffices on sentence meaning in Sanskrit demonstrate a fascinating linguistic phenomenon. Through our analysis, we found that despite shuffling the words within a sentence, the overall meaning remains consistent due to the contextual information provided by the suffices attached to each word.

##### Results

Our experiments involved rearranging the words in several Sanskrit sentences while keeping the suffices intact. We observed that, remarkably, the meaning of the sentences remained unchanged despite the alteration in word order. This phenomenon underscores the role of suffices in conveying grammatical and semantic information within Sanskrit sentences.

##### Discussion

The findings of our study have several implications for understanding the structure and flexibility of Sanskrit language:

##### Linguistic Flexibility

The presence of suffices in Sanskrit endows the language with a remarkable degree of flexibility in word order. Unlike languages with strict word order requirements, Sanskrit allows for greater freedom in sentence construction while preserving meaning.

##### Contextual Clarity

The suffices in Sanskrit serve to disambiguate the roles of words within a sentence, providing crucial contextual information to aid in comprehension. This ensures that even when words are shuffled, the intended meaning remains clear to the reader or listener.

##### Preservation of Semantic Integrity

The ability of Sanskrit sentences to maintain their meaning despite word rearrangement highlights the robustness of the language's semantic structure. This feature contributes to Sanskrit's enduring relevance and usability in diverse linguistic contexts.

##### Computational Applications

The findings have implications for computational linguistics and Natural Language Processing (NLP), where understanding the flexibility and structure of Sanskrit can inform the development of AI algorithms for text analysis and generation.

Overall, our results underscore the intricate linguistic features of Sanskrit, highlighting its unique characteristics and the role of suffices in preserving semantic integrity. Further research in this area could explore the specific mechanisms by which suffices contribute to sentence meaning and their implications for language processing algorithms.

#### 5. CONCLUSION

The convergence of Sanskrit and AI represents a unique fusion of ancient wisdom and modern technology. By leveraging AI, we can unlock the vast knowledge contained within Sanskrit texts, preserving and revitalizing them for future generations. Conversely, the structural and theoretical aspects of Sanskrit can inspire advancements in AI, particularly in the realm of natural language processing. This symbiotic relationship holds immense potential for both fields, fostering a deeper understanding of human language and knowledge.

#### 6. REFERENCES

- [1] Pollock, Sheldon. "The Language of the Gods in the World of Men: Sanskrit, Culture, and Power in Premodern India." University of California Press, 2006.
- [2] Kiparsky, Paul. "Panini as a Variationist." In S.D. Joshi (Ed.), "Paninian Studies," 1991.
- [3] "The Digital Library of India." [Online]. Available: <https://www.dli.ernet.in/>
- [4] Kakwani, Priyanka, and Shah, Monojit. "Application of Natural Language Processing to Ancient Indian Manuscripts." Proceedings of the International Conference on Computational Linguistics, 2018.
- [5] Rick Briggs, "Knowledge Repeantation in Sanskrit and Artificial Intelligence", In RIACS, NASA Ames Research Center, Moffet Field, California 94304