

Towards Natural Language Processing with Figures of Speech in Hindi Poetry

Milind Kumar Audichya¹

Computer Science Department
Gujarat Technological University, Ahmedabad, India

Jatinderkumar R. Saini^{2*}

Symbiosis Institute of Computer Studies and Research
Symbiosis International (Deemed University), Pune, India

Abstract—Poems have always been an excellent way of expressing emotions in any language. In particular, Hindi poetry is having versatile popularity among native and non-native speakers all over the world. A typical poem in Hindi is characterized by meter (“Chhand”), emotion (“Rasa”), and figure of speech (“Alankaar”). The present research work is the first of its kind in Hindi Natural Language Processing (NLP), which touches on the area of Hindi figure of speech. The authors have created a systematic hierarchical structure of Hindi “Alankaar” types and sub-types and attempted and extended the work to identify a few. A taxonomical list of 58 Hindi figures of speech is presented along with their nearest mapping to English equivalents. On the sidelines, the paper also presents the distinct rules for each type and sub-type needed for the classification task of NLP. The authors achieved 97% efficiency in reporting the first results with an average execution time of 0.002 seconds.

Keywords—“Alankaar”; figure of speech; Hindi; Natural Language Processing (NLP); poetry

I. INTRODUCTION

Poetry refers to the poetic creation of a person and consists of a series of verses. Different types of poetry have been written from as early as the 25th century BCE [1]. Different types of rules and regulations are followed for writing poetry in different language scripts, yet maintaining grammar [2]. Hindi is one of the prevalent languages of the world. To some extent, Hindi is a majorly spoken language for communication in India and is written using the Devanagari script. It is used along with English as the official language of the Government of India [3].

Many well-known writers have done many poetic creations in the Hindi language, and every day many writers are writing some new poems. In Hindi poetry, “Rasa” (i.e., “emotion”), “Chhand” (i.e., verse meter), and “Alankaar” (i.e., the figure of speech) are essentials of the poetic composition [4]. Though little progress can be seen in the “Rasa” and verse-related research works, it is almost absent when it comes to the figure of speech. The figure of speech known as “Alankaar” in Hindi is capable enough to make any poem’s creation magical through its presence.

Our contribution through the present research work includes:

- Detailed exploration of Hindi “Alankaar”,
- The standardization of taxonomical classification structure for different types of “Alankaar”, and

- The specific methodology for identification of the three trendy Hindi “Alankaar”.

The former could be well exploited for Natural Language Processing (NLP) of Hindi language, particularly for the classification task.

Notably, all of these are worked upon and reported for the first time in the scientific literature. The structure of the remainder of the paper includes literature review, description of rules and creation of “Alankaar” hierarchy, “Alankaar” identification, and results. The paper ends with the conclusions derived from the work and some pointers to future work.

II. LITERATURE REVIEW

An extensive literature review was carried out for this research, in which we tried to dig up the research items, books, blogs, and online portals for the different kinds of information retrieval and to know the current state of the research progress in this specific segment. Research works for the internationally well-known languages that can be seen concerning poetry and related nearby segments such as emotion detection, text classification, and identification in different languages such as Arabic, Chinese, English, and Persian [5-8]. Some research works related to poetry were found for Indian regional languages like Hindi, Marathi, and Punjabi [9-11].

Saini and Kaur [13] worked for Punjabi poems annotated corpus for emotion detection based on the nine different types of emotions (“Rasa”). Pal and Patel [12] introduced a model for the classification of Hindi poems. Audichya and Saini [14] worked for the unified rule-based technique for automatic metadata generation based on different meter rules in Hindi poetry. Kushwah and Joshi [15] researched the detection of a specific type of verse meter named “Rola”. Bafna and Saini [16-19] also worked using eager machine learning and concept learning algorithms to classify Hindi verses.

After founding and exploring this much, authors can powerfully convey that the figure of speech as known as “Alankaar” in Hindi is an untouched portion of the research works related to Hindi or any other Indian regional level languages. The main reason behind less work in this area is because it is tedious to deal with, and no such initial research works had been done or carried out so far. To fulfill this gap, we have presented a path to work further in Hindi NLP from the perspective of “Alankaar”. One of the goals of this research work is to organize and manage all available information related to the figure of speech in the Hindi language after

*Corresponding Author

proper collection, verification, and validation for a better research approach so that one is not needed to deal with insufficient knowledge or contradictory information in upcoming times.

After the detailed and in-depth literature review, it was observed that there is a lack of identification mechanism which can detect and identify the different “Alankaar” in Hindi. The vast Hindi content data can be sorted in an organized manner with the generated metadata based on the detection. The metadata can help populate better search results instead of regular keywords-based searching. Apart from that, this research work can help digital libraries to manage the content based on the different types of Hindi “Alankaar”. With the perspective of computational logistics, this research work can help analyze the write-ups based on the different types of the Hindi “Alankaar”. So as authors felt, it can be a valuable and necessary novel work for current and upcoming times, which can be understood with already emphasized points, and there can be many more uses scenarios also. These all aspects were the actual motivation to carry out this research work.

III. “ALANKAAR”: THE HINDI FIGURE OF SPEECH

The figure of speech which is scripted as “अलंकार” (“Alankaar”) in the Devanagari script of the Hindi language, is an essential part of the creation of the poem. “Alankaar” means “ornament”, and just as the beauty of a person is adorned with ornamentation, in the same way, the grace of poetry is ornamented. Something which embellishes the poem is known as the figure of speech.

To identify and detect the different types of “Alankaar” we first need standard types and rules, which are not available systematically, and where ever is available, it is either missing some information or have some fewer details [20-21]. The first and significant time-consuming task of this research work was to go through the various sources to collect, verify, and systematically structure hierarchical classes for better research.

The rules and all their relevant information collection were carried out through different places such as educational materials, websites, blogs, and portals [22-24]. After this process, we have structured everything in such a way with Hindi experts’ opinions from academia to be helpful for the upcoming research works.

A. Types of “Alankaar”

Mainly as per the characteristics, the types of “Alankaar” are divided into three streams which are as follows:

- 1) “Shabd Alankaar” (“शब्दालंकार”)
- 2) “Arth Alankaar” (“अर्थालंकार”)
- 3) “Ubhay Alankaar” (“उभयालंकार”)

Each has its own set of rules and further subtypes, and even sub-sub types [25]. We discuss each of them quickly in the subsequent sections.

1) “Shabd Alankaar” (“शब्दालंकार”): “Shabd Alankaar” is the first type of the Hindi figure of speech, those which embellish the poems through the words in the figure of speech, that is, by putting a particular word in a poem, the

beauty comes, and if the beauty is lost when using a synonym, are called the “Shabd Alankaar”. Although not all the “Alankaars” that come in this category are purely based on words, the primary focus in those “Alankaars” is on the words, which is why they are added into this category.

“Shabd Alankaar” have further sub and sub-sub classes as per their types mentioned as follows: Each has its own set of rules and further subtypes, and even sub-sub types [25]. We discuss each of them quickly in the subsequent sections.

1.1 Alliteration (“अनुप्रास अलंकार”)

- 1.1.1 ‘Chekanupras’ (“छेकानुप्रास अलंकार”)
- 1.1.2 ‘Vrutyanupras’ (“वृत्यानप्रास अलंकार”)
- 1.1.3 ‘Latanupras’ (“लाटानुप्रास अलंकार”)
- 1.1.4 ‘Antyanpras’ (“अन्तत्यानप्रास अलंकार”)
- 1.1.5 ‘Shrtyanpras’ (“श्रत्यानप्रास अलंकार”)

1.2 ‘Yamak’ (“यमक अलंकार”)

1.3 ‘Punrukti’ (“पुनरुक्ति अलंकार”)

1.4 ‘Vipsa’ (“विप्सा अलंकार”)

1.5 ‘Vakrokti’ (“वक्रोक्ति अलंकार”)

- 1.5.1 ‘Kaku Vakrokti’ (“काकु वक्रोक्ति अलंकार”)
- 1.5.2 ‘Shelsh Vakrokti’ (“श्लेष वक्रोक्ति अलंकार”)

1.6 Pun or Irony (“श्लेष अलंकार”)

- 1.6.1 ‘Abhang Shlesh’ (“अभंग श्लेष अलंकार”)
- 1.6.2 ‘Sabhang Shlesh’ (“सभंग श्लेष अलंकार”)

2) “Arth Alankaar” (“अर्थालंकार”): “Arth Alankaar” is mainly related to the meaning of the words, so those which embellish the poems through the meaning of the words in the figure of speech, That is, by putting a particular word in a poem and due to the meaning of that word the miracles occur in poetry, are called the “Arth Alankaar”. Although not all the “Alankaar” which comes in this category are purely based on the meaning of words, yet the primary focus in those “Alankaars” is on the meaning of the words that’s why they are added into this category.

“Arth Alankaar” are having further sub and sub-sub classes as per their types which are mentioned as following:

1.1 Simile (“उपमा अलंकार”)

- 1.1.1 ‘Purnopama’ (“पूर्णोपमा अलंकार”)
- 1.1.2 ‘Luptopama’ (“लुप्तोपमा अलंकार”)

1.2 Metaphor (“रूपक अलंकार”)

- 1.2.1 ‘Sam Rupak’ (“सम रूपक अलंकार”)
- 1.2.2 ‘Adhik Rupak’ (“अधिक रूपक अलंकार”)
- 1.2.3 ‘Nyun Rupak’ (“न्यून रूपक अलंकार”)

1.3 Poetic Fancy (“उत्प्रेक्षा अलंकार”)

- 1.3.1 ‘Vastupreksha’ (“वस्तुप्रेक्षा अलंकार”)
- 1.3.2 ‘Hetupreksha’ (“हेतुप्रेक्षा अलंकार”)
- 1.3.3 ‘Falotpreksha’ (“फलोत्प्रेक्षा अलंकार”)

1.4 Exemplification (“द्रष्टान्ति अलंकार/दृष्टान्त”)

1.5 Doubt (“संदेह अलंकार”)

1.6 Hyperbole (“अतिशयोक्ति अलंकार”)

1.7 ‘Upmeyopma’ (“उपमेयोपमा अलंकार”)

- 1.8 *Converse* (“प्रतीप अलंकार”)
- 1.9 *Self Comparison* (“अनन्वय अलंकार”)
- 1.10 *Error* (“भ्रांतिमान अलंकार”)
- 1.11 *Illuminator* (“दीपक अलंकार”)
- 1.12 *Concealment* (“अपहृति अलंकार”)
- 1.13 *‘Vyatirek’* (“व्यतिरेक अलंकार”)
- 1.14 *Peculiar Causation* (“विभावना अलंकार”)
- 1.15 *Peculiar Allegation* (“विशेषोक्ति अलंकार”)
- 1.16 *Corroboration* (“अर्थान्तरन्यास अलंकार”)
- 1.17 *‘Ullek’* (“उल्लेख अलंकार”)
- 1.18 *Contradiction* (“विरोधाभास अलंकार”)
- 1.19 *Disconnection* (“असंगति अलंकार”)
- 1.20 *Personification* (“मानवीकरण अलंकार”)
- 1.21 *‘Anantyo’kti’* (“अन्तयोक्ति अलंकार”)
- 1.22 *Poetical Reason* (“काव्यलिंग अलंकार”)
- 1.23 *Natural Description* (“स्वभावोती अलंकार”)
- 1.24 *Typical Comparison* (“प्रतिवस्तूपमा”)
- 1.25 *Chain of Similes* (“मालोपमा”)
- 1.26 *Equal Pairing* (“तुल्योयिता”)
- 1.27 *Illustration* (“निदर्शना”)
- 1.28 *Speech of Brevity* (“समासोक्ति”)
- 1.29 *Indirect Dissection* (“अप्रस्तुतप्रर्शासा”)
- 1.30 *Special Mention* (“परिसंख्या”)

| Types of ‘ALANKAAR’ - The Hindi Figure of Speech (“अलंकार के भेद”) | |
|--|---|
| 1. ‘Shabd Alankaar’ (“शब्दालंकार”) | 2. ‘Arth Alankaar’ (“अर्थालंकार”) |
| 1.1 Alliteration (“अनुप्रास अलंकार”) | 2.1 Simile (“उपमा अलंकार”) |
| 1.1.1 ‘Chekanupras’ (“छेकनुप्रास अलंकार”) | 2.1.1 ‘Purnopama’ (“पूर्णोपमा अलंकार”) |
| 1.1.2 ‘Vrutyanupras’ (“वृत्तानुप्रास अलंकार”) | 2.1.2 ‘Luptopama’ (“लुप्तोपमा अलंकार”) |
| 1.1.3 ‘Latanupras’ (“लटानुप्रास अलंकार”) | 2.2 Metaphor (“रूपक अलंकार”) |
| 1.1.4 ‘Antyanpras’ (“अन्तयानुप्रास अलंकार”) | 2.2.1 ‘Sam Rupak’ (“सम रूपक अलंकार”) |
| 1.1.5 ‘Shrtyanpras’ (“श्रयानुप्रास अलंकार”) | 2.2.2 ‘Adhik Rupak’ (“अधिक रूपक अलंकार”) |
| 1.2 ‘Yamak’ (“यमक अलंकार”) | 2.2.3 ‘Nyun Rupak’ (“न्यून रूपक अलंकार”) |
| 1.3 ‘Punrukti’ (“पुनरुक्ति अलंकार”) | 2.3 Poetic Fancy (“उपेक्षा अलंकार”) |
| 1.4 ‘Vipsa’ (“विप्सा अलंकार”) | 2.3.1 ‘Vastupreksha’ (“वस्तुप्रेक्षा अलंकार”) |
| 1.5 ‘Vakrokti’ (“वक्रोक्ति अलंकार”) | 2.3.2 ‘Hetupreksha’ (“हेतुप्रेक्षा अलंकार”) |
| 1.5.1 ‘Kaku Vakrokti’ (“ककु वक्रोक्ति अलंकार”) | 2.3.3 ‘Falotpreksha’ (“फलतोपेक्षा अलंकार”) |
| 1.5.2 ‘Shesh Vakrokti’ (“शेष वक्रोक्ति अलंकार”) | 2.4 Exemplification (“प्रदर्शन अलंकार/दृष्टान्त”) |
| 1.6 Pan or Irony (“श्लेष अलंकार”) | 2.5 Doubt (“संशय अलंकार”) |
| 1.6.1 ‘Abhang Shlesh’ (“अभंग श्लेष अलंकार”) | 2.6 Hyperbole (“अतिव्योक्ति अलंकार”) |
| 1.6.2 ‘Sabhang Shlesh’ (“सभंग श्लेष अलंकार”) | 2.7 ‘Upmeyopma’ (“उपमेयोपमा अलंकार”) |
| 3. ‘Ubhay Alankaar’ (“उभयालंकार”) | 2.8 Converse (“प्रतीप अलंकार”) |
| 3.1 Combination of Figures of Speech (“संयुक्ति”) | 2.9 Self Comparison (“अनन्वय अलंकार”) |
| 3.2 The fusion of Figures of Speech (“संकर”) | 2.10 Error (“भ्रंतिमान अलंकार”) |
| | 2.11 Illuminator (“दीपक अलंकार”) |
| | 2.12 Concealment (“अपहृति अलंकार”) |
| | 2.13 ‘Vyatirek’ (“व्यतिरेक अलंकार”) |
| | 2.14 Peculiar Causation (“विभावना अलंकार”) |
| | 2.15 Peculiar Allegation (“विशेषोक्ति अलंकार”) |
| | 2.16 Corroboration (“अर्थान्तरन्यास अलंकार”) |
| | 2.17 ‘Ullek’ (“उल्लेख अलंकार”) |
| | 2.18 Contradiction (“विरोधाभास अलंकार”) |
| | 2.19 Disconnection (“असंगति अलंकार”) |
| | 2.20 Personification (“मानवीकरण अलंकार”) |
| | 2.21 ‘Anantyo’kti’ (“अन्तयोक्ति अलंकार”) |
| | 2.22 Poetical Reason (“काव्यलिंग अलंकार”) |
| | 2.23 Natural Description (“स्वभावोती अलंकार”) |
| | 2.24 Typical Comparison (“प्रतिवस्तूपमा”) |
| | 2.25 Chain of Similes (“मालोपमा”) |
| | 2.26 Equal Pairing (“तुल्योयिता”) |
| | 2.27 Illustration (“निदर्शना”) |
| | 2.28 Speech of Brevity (“समासोक्ति”) |
| | 2.29 Indirect Dissection (“अप्रस्तुतप्रर्शासा”) |
| | 2.30 Special Mention (“परिसंख्या”) |

Fig. 1. Systematically Structured Graphical Representation of Hindi Alankaars.

B. Selecting a Template

Research works are always challenging, and that’s what the beauty of research is, but when it comes to the research with the figure of speech in Hindi, it is very tedious and challenging. That’s the only reason this segment was still untouched, and no such initial research was found. While carrying out this research work, the following challenges were faced.

1) *No previous research works*: Initial level research work requires some extra efforts as we discussed already that no previous research work or articles had been found so far, so one needs to create their path or way to work to accomplish the research-related tasks and it requires massive efforts because there is no dataset, algorithms or implementation strategy is existing.

2) *Missing and conflicting information*: Information Collection, verification, and systematic arrangement are some of the initial tasks of any research work, and this is more important when dealing with a purely new segment where no such past research works or articles can be seen. The authors came across different sources in this collection and validation process where either some types were missing or having incomplete information.

3) *Context-based meaning*: To deal with Hindi words’ meanings, one can integrate with the existing Hindi wordnet or other libraries, but the context-based meaning is required, which is missing, or still, some research works are going on in the same segment and research in its own [26].

TABLE I. ALANKAARS COMPARATIVE SCENARIO AND STATS

| Alankaars Comparative Scenario | Count |
|--|-------|
| Total Alankaars for which English equivalent is found: | 30 |
| Total Alankaars for which English equivalent is not found: | 28 |
| Total Alankaars in Hindi: | 58 |

4) *Homonymy and polysemy*: “Alankaars” are all about the words and their meanings, here a single word can have multiple and can be used to express different things, which are polysemy, and similar words that are either spelled similar or sound the same but have different meanings are homonymy. That is another challenge level, which is still a vast issue and essential for this research work, too [27].

5) *Multiple “Alankars” detection*: As per the nature and characteristics of “Alankaar”, there can be multiple “Alankaars” in the same poem lines or even in a part of the poem, comparatively in “Rasa” and “Chhand” usually it has been observed that mostly there will be only a single type of “Rasa” or “Chhand” will be there in a part of the poem.

6) *Unavailability of datasets for experiments*: To carry out any research work, one will always need a dataset, as there is no such research work done in this specific problem segment, and in other poem related Hindi research also works dataset is a challenge because there is no such ready dataset or open-source datasets are available. To deal with such things, one has to follow one and the only thing that makes the dataset by self, and again it requires some additional effort and time.

Despite all of the listed challenges, we followed the approach of focusing on the best optimum problem-solving methods, and the same is discussed in the following “Alankaar” Identification section.

IV. “ALANKAAR” IDENTIFICATION

To Identify and detect the “Alankaars” used in Hindi poetry based on the different rules of “Alankaars”, we tried to implement the viral, trendy, and three primarily used “Alankaars” out of the all mentioned 58 different types.

For example, to identify and detect these two “Alankaars”, namely “Anupras” (i.e., Alliteration) “Punrukti”, we need to know the respective appropriate rules of both types. If we consider “Anupras”, the rule says that when a specific character occurs repeatedly, there is “Anupras”. If we talk about the “Punrukti”, a word that occurs twice consecutively, then there is “Punrukti”. Let us understand with the following example which fits for both the types:

‘ढुमुकि - ढुमुकि रुनझुन धुनि - सुनि,
कनक अजिर शिशु डोलता’

In this example, the Unicode Standard [28] Unicode Transformation Format - 8 (UTF-8) based text is accepted as input, and if we observe closely, we can see that the character ‘क’, ‘न’ and more occurs more than once, repeated and again and again so “Anupras” is here. Also, there is “ढुमुकि” word which is occurring twice consecutively it is fulfilling “Punrukti” rule. This is how one can understand this concept, but to make a computer computationally understand the same, we need to follow some systematic process so we have designed in such a way that in the case in near future we need to add some more “Alankaar” implementation we can do that very quickly.

The simplest way to understand the implementation methodology is as follows:

Step 1: Start.

Step 2: Input the data in “UTF-8” format.

Step 3: Cleaning and Preprocessing operations.

Step 4: Perform Character Count and Word Count.

Step 5: Send the data to check the “Alankaar”.

Step 6: Check “Alankaar” in “Shabd Alankaar” where it will further pass on the sub-type functions, and if any type gets detected, it will be added to the output result buffer.

Step 7: Check “Alankaar” in “Arth Alankaar” where it will further pass on the subtype functions and if any type gets detected, it will be added to the output result buffer.

Step 8: Check “Alankaar” in “Ubhay Alankaar” where it will further pass on the sub-type functions and if any type gets detected, it will be added to the output result buffer.

Step 9: Return appropriate output by merging all the output in the buffer.

Step 10: Stop.

With this methodology, many “Alankaars” can be easily covered as soon as the modeling of the specific “Alankaar” rules is done in the implementation script’s functional modules. Let’s have a look at the pseudo-code for this implementation.

```
if isshabdAlankaar():  
    outputformator(output)  
elif isarthAlankaar():  
    outputformator(output)  
elif isubhyaAlankaar():  
    outputformator(output)
```

If one need to check for specific “Alankaar” functional code goes as follow:

```
def isanupras():  
    global Input,output  
    anupras = False  
    count = character_count(Input)  
    for key, value in count.items():  
        if value >= 2:  
            anupras = True
```

The function gets called while checking “Shabd Alankaar”. As same as the “isanupras()” other alankars methods also gets executed automatically while checking different classes of “Alankaar”. Along with that, by keeping the computational perspective in mind, if required, the appropriate position of the detected “Alankaar” can also be populated along with the final metadata.

V. RESULTS

In this research work, we started from scratch and as a final result, we have systematically sorted and arranged standard hierarchical data of “Alankaars”. Apart from that, the authors were also able to execute binary classification for the three “Alankaars” successfully. From an implementation perspective, the authors have already implemented “Anupras”, “Punrukti” and “Yamak” “Alankaars”. The same example is used to explain the result, which was used to discuss in the

“Alankaar” Identification section. The code for the output depicted in Fig. 1 was written using Python version 3.9 and executed on MacBook Air 13-inch, 2017 system with macOS Big Sur Version 11.1 having 8 GB 1600 MHz DDR3 Memory along with 1.8 GHz Dual-Core Intel Core i5 processor.

```
कृपया इनपुट उपलब्ध करवाए (इनपुट के बाद Ctrl+D दबाए) :  
डुमुकि - डुमुकि रुनरुक्ति धुनि - सुनि ,  
कनक अजिर शिशु डोलत।  
^D  
Type : शब्दालंकार  
Sub Type : अनुप्रास अलंकार  
Type : शब्दालंकार  
Sub Type : पुनरुक्ति अलंकार  
Time Duration (In Seconds) : 0.00023818016052246094
```

Fig. 2. Output of the Automatic “Alankaar” Identification.

Fig. 2 shows that it took input and processed the same as discussed in section IV, and on completion, it returns that the input is consists of two “Alankaars” which comes under the primary type “Shabd Alankaar” and their subtypes are “Anupras” and “Punrukti”. Also, one thing to notice here is that the whole process took just 0.002 seconds, which is a rapid execution time. Apart from these two, we have also incorporated the “Yamak Alankaar” identification which works quite well, but it does not work in some scenarios as we need to make it work better using the integration of the wordnet for the meanings of the words for the comparison of different words.

Table II shows the result related stats of this research work carried out after the working model’s design and implementation integration. There is no training mechanism based on data in this research study, so whatever data inputted for the tests were genuinely on unseen data only. The test was carried out on the 78 different UTF-8 based input, and based on the results, we were finally able to achieve overall 97.00% accuracy in 0.002 second average execution time.

TABLE II. ALANKAARS IDENTIFICATION ACCURACY AND EXECUTION TIME

| Alankaar | Accuracy % | Execution Time (Seconds) |
|----------------|----------------|--------------------------|
| “Anupras” | 100.00% | 0.002 Second |
| “Punrukti” | 100.00% | 0.002 Second |
| “Yamak” | 091.00% | 0.002 Second |
| Total | 291.00% | 0.006 Second |
| Maximum | 100.00% | 0.002 Second |
| Minimum | 091.00% | 0.002 Second |
| Average | 097.00% | 0.002 Second |

VI. CONCLUSION AND FUTURE WORK

This research work consists of a lot of time-consuming efforts, which concludes several things. Information related to “Alankaar” is now having some research sort of systematic arrangement which can be used to take it further. Dealing with “Alankaar” is not an easy task. “Alankaars” are itself a world

inside them, and it is not possible to cover all the “Alankaar” identification or detection in one single research work.

Nevertheless, the authors introduced how “Alankaar” identification and detection can be made automatically. Three well-known “Alankaars” are currently modeled, which are “Anupras”, “Punrukti” and “Yamak”. Wordnet integration with advanced grammatical aspects can strengthen this identification system more effectively in upcoming times. In further research works, we will be adding more types. The authors will also be looking for solutions to the challenges that were not able to overcome currently, such as the context-based meaning of Hindi words and other ambiguity-related issues.

REFERENCES

- [1] Wikipedia Contributors, ‘Poetry,’ Wikipedia, 10-Mar-2019. [Online]. Available: <https://en.wikipedia.org/wiki/Poetry>.
- [2] ‘7 Fundamental Rules of Poetry,’ 7 Fundamental Rules of Poetry: Grammarly Blog, 06-May-2015. [Online]. Available: <https://www.grammarly.com/blog/7-fundamental-rules-of-poetry/>. [Accessed: 04-Dec-2020].
- [3] Wikipedia Contributors, ‘Hindi,’ Wikipedia, 04-May-2019. [Online]. Available: <https://en.wikipedia.org/wiki/Hindi>.
- [4] Contributors to Wikimedia projects, ‘रस,’ Wikipedia.org, 29-Nov-2006. [Online]. Available: <https://hi.wikipedia.org/s/a4j> [Accessed: 04-Dec-2020].
- [5] Z. He, W. Liang, L. Li and Y. Tian, ‘SVM-Based Classification Method for Poetry Style,’ 2007 International Conference on Machine Learning and Cybernetics, Hong Kong, 2007, pp. 2936-2940, doi: 10.1109/ICMLC.2007.4370650.
- [6] S. Hamidi, F. Razzazi and M. P. Ghaemmaghami, ‘Automatic meter classification in Persian poeties using support vector machines,’ 2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, 2009, pp. 563-567, doi: 10.1109/ISSPIT.2009.5407514.
- [7] N. Ghneim and O. Alsharif, ‘Emotion Classification in Arabic Poetry Using Machine Learning building a model for gender, dialect and age prediction from Arabic tweets View project Interactive Arabic Dictionary View project Emotion Classification in Arabic Poetry using Machine Learning,’ International Journal of Computer Applications, vol. 65, no. 16, pp. 975–8887, 2013.
- [8] A. Abbasi, H. Chen, and A. Salem, ‘Sentiment analysis in multiple languages,’ ACM Transactions on Information Systems, vol. 26, no. 3, pp. 1–34, Jun. 2008, doi: 10.1145/1361684.1361685.
- [9] J. Kaur and J.R.Saini, ‘Designing Punjabi Poetry Classifiers Using Machine Learning and Different Textual Features,’ The International Arab Journal of Information Technology, pp. 38–44, Jan. 2019, doi: 10.34028/iajit/17/1/5.
- [10] B.K. Joshi and K.K. Kushwah, ‘A Novel Approach to Automatic Detection of Chaupai Chhand in Hindi Poems,’ 2018 International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, Uttar Pradesh, India, 2018, pp. 223-228, doi: 10.1109/GUCON.2018.8675052.
- [11] P.B. Bafna and J.R.Saini, ‘An Application of Zipf’s Law for Prose and Verse Corpora Neutrality for Hindi and Marathi Languages,’ International Journal of Advanced Computer Science and Applications, vol. 11, no. 3, 2020, doi: 10.14569/ijacsa.2020.0110331.
- [12] K. Pal and B.V. Patel, ‘Model for Classification of Poems in Hindi Language Based on Rasa,’ Smart Systems and IoT: Innovations in Computing, pp. 655–661, Oct. 2019, doi: 10.1007/978-981-13-8406-6_62.
- [13] J.R.Saini and J.Kaur, ‘Kāvi: An Annotated Corpus of Punjabi Poetry with Emotion Detection Based on ‘NavRasa’, Procedia Computer Science, vol. 167, pp. 1220–1229, 2020, doi: 10.1016/j.procs.2020.03.436.
- [14] M.K. Audichya and J.R. Saini, ‘Computational linguistic prosody rule-based unified technique for automatic metadata generation for Hindi poetry,’ 2019 1st International Conference on Advances in Information

- Technology (ICAIT), Chikmagalur, India, 2019, pp. 436-442, doi: 10.1109/ICAIT47043.2019.8987239.
- [15] K.K. Kushwah and B.K. Joshi, 'Rola: An Equi-Matrik Chhand of Hindi Poems.', International Journal of Computer Science and Information Security (IJSIS), Vol. 15, No. 3, March 2017.
- [16] P.B. Bafna and J.R. Saini, 'On Exhaustive Evaluation of Eager Machine Learning Algorithms for Classification of Hindi Verses,' International Journal of Advanced Computer Science and Applications, vol. 11, no. 2, 2020, doi: 10.14569/ijacsa.2020.0110224.
- [17] P.B. Bafna and J.R. Saini, 'Hindi Verse Class Predictor using Concept Learning Algorithms,' 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, 2020, pp. 318-322, doi: 10.1109/ICIMIA48430.2020.9074850.
- [18] P.B. Bafna and J.R. Saini, 'BaSa: A Technique to Identify Context based Common Tokens for Hindi Verses and Proses,' 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154124.
- [19] P.B. Bafna and J.R. Saini, 'Hindi Poetry Classification using Eager Supervised Machine Learning Algorithms,' 2020 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2020, pp. 175-178, doi: 10.1109/ESCI48226.2020.9167632.
- [20] 'भारतीय काव्यशास्त्र (दिवि)/अलंकार - विकिपुस्तक,' Wikibooks.org, 2020. [Online]. Available: [https://hi.wikibooks.org/wiki/%E0%A4%AD%E0%A4%BE%E0%A4%B0%E0%A4%A4%E0%A5%80%E0%A4%AF%E0%A4%95%E0%A4%BE%E0%A4%B5%E0%A5%8D%E0%A4%B6%E0%A4%BE%E0%A4%B8%E0%A5%8D%E0%A4%A4%E0%A5%8D%E0%A4%B0_\(%E0%A4%A6%E0%A4%BF%E0%A4%B5%E0%A4%BF\)/%E0%A4%85%E0%A4%B2%E0%A4%82%E0%A4%95%E0%A4%BE%E0%A4%B0](https://hi.wikibooks.org/wiki/%E0%A4%AD%E0%A4%BE%E0%A4%B0%E0%A4%A4%E0%A5%80%E0%A4%AF%E0%A4%95%E0%A4%BE%E0%A4%B5%E0%A5%8D%E0%A4%B6%E0%A4%BE%E0%A4%B8%E0%A5%8D%E0%A4%A4%E0%A5%8D%E0%A4%B0_(%E0%A4%A6%E0%A4%BF%E0%A4%B5%E0%A4%BF)/%E0%A4%85%E0%A4%B2%E0%A4%82%E0%A4%95%E0%A4%BE%E0%A4%B0). [Accessed: 07-Dec-2020].
- [21] Contributors to Wikimedia projects, 'अलंकार,' Wikipedia.org, 19-Mar-2007. [Online]. Available: <https://hi.wikipedia.org/s/9ll> [Accessed: 07-Dec-2020].
- [22] 'अलंकार - अलंकार की परिभाषा, भेद, उदाहरण, Hindi Grammar.' [Online]. Available: <https://www.mycoaching.in/2018/09/Alankaar.html>. [Accessed: 07-Dec-2020].
- [23] 'अलंकार - भारतकोश, ज्ञान का हिन्दी महासागर,' Bharatdiscovery.org, 2020. [Online]. Available: <https://bharatdiscovery.org/india/%E0%A4%85%E0%A4%B2%E0%A4%82%E0%A4%95%E0%A4%BE%E0%A4%B0>.
- [24] 'अलंकार की परिभाषा, भेद, प्रकार और उदाहरण सहित पूरी जानकारी,' Hindi vibhag, 26-Aug-2017. [Online]. Available: <https://www.hindivibhag.com/%E0%A4%85%E0%A4%B2%E0%A4%82%E0%A4%95%E0%A4%BE%E0%A4%B0/>. [Accessed: 07-Dec-2020].
- [25] 'Alankaar'(Figure of speech)(अलंकार) Hindi Grammar,' hindigrammar.in. [Online]. Available: <http://hindigrammar.in/Alankaar%2.html>. [Accessed: 07-Dec-2020].
- [26] S. Rajendran and S. Arulmozi, 'Augmenting Indo-wordnet with Context' [Online]. Available: http://www.cfilt.iitb.ac.in/wordnet/webhwn/IndoWordnetPapers/13_iwn_Augmenting%20Indo-wordnet%20with%20Context.pdf. [Accessed: 07-Dec-2020].
- [27] N. Dash, 'Polysemy and Homonymy: A Conceptual Labyrinth' [Online]. Available: http://www.cfilt.iitb.ac.in/wordnet/webhwn/IndoWordnetPapers/08_iwn_Polysemy%20and%20Homonymy.pdf. [Accessed: 07-Dec-2020].
- [28] 'The Unicode Standard, Version 12.0 - Devanagari.' [Online]. Available: <https://www.unicode.org/charts/PDF/U0900.pdf>.