



## हिंदी मुहावरों के भाव-विश्लेषण हेतु नियम आधारित संगणकीय प्रणाली

(A Rule Based Computational System for Sentiment Analysis of Hindi Idioms)

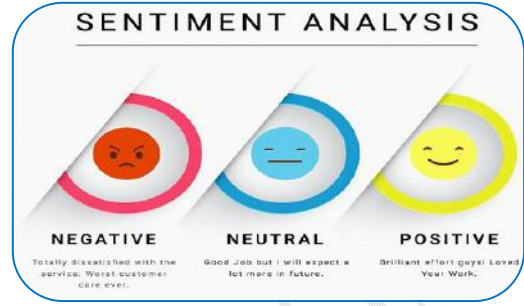
प्रितेन्द्र कुमार मालाकार<sup>1</sup>, डॉ. हर्षलता पेटकर<sup>2</sup>

<sup>1</sup>शोधार्थी, सूचना एवं भाषा अभियांत्रिकी केंद्र,

म.गां.अं.हिं.वि., वर्धा (महाराष्ट्र)

<sup>2</sup>सहायक प्रोफेसर, सूचना एवं भाषा अभियांत्रिकी केंद्र,

म.गां.अं.हिं.वि., वर्धा (महाराष्ट्र)



### सारांश (Summary):

भाव-विश्लेषण (Sentiment Analysis) एक प्रक्रिया है जिसकी सहायता से जन-अभिव्यक्तियों को किसी विषय के प्रति सकारात्मक (Positive), नकारात्मक (Negative), अथवा निष्पक्ष (Neutral) के रूप में वर्गीकृत किया जाता है ताकि उस विषय के बारे में कोई महत्वपूर्ण निर्णय लिया जा सके। हिंदी भाषा की जन-अभिव्यक्तियों के भाव-विश्लेषण संबंधी अधिकतर कार्य केवल भावयुक्त शब्दों (Sentimental Words) को ही आधार बनाकर किए गए हैं। हिंदी भाषा आधारित भाव-विश्लेषण की शुद्धता एवं गुणवत्ता में कमी का यह भी एक प्रमुख कारण है। इसी समस्या के समाधान हेतु प्रस्तुत शोध के अंतर्गत एक नियम-आधारित संगणकीय प्रणाली का निर्माण किया गया है जिसमें भावयुक्त शब्दों के साथ मुहावरों (Idioms) को भी अभिलक्षण के रूप में सम्मिलित कर हिंदी भाषा आधारित भाव-विश्लेषण की शुद्धता एवं गुणवत्ता में वृद्धि करने का प्रयास किया गया है।

**मुख्य शब्द (Keywords) :** भाव-विश्लेषण, जन-अभिव्यक्तियां, नियम-आधारित, मुहावरा, संगणकीय प्रणाली।

### प्रस्तावना:

वेब प्रौद्योगिकी के उन्नयन एवं बेहतर इनपुट टूल्स की उपलब्धता के कारण सोशल मीडिया पर संपर्क स्थापित करना तथा स्वयं की भाषा में विचार प्रकट करना और भी आसान हो गया है। आज प्रयोक्ताओं के द्वारा दैनिक जीवन से संबंधित विभिन्न विषयों के प्रति अपनी अभिव्यक्तियों को सोशल मीडिया पर नियमित रूप से व्यक्त किया जा रहा है जिसके कारण इंटरनेट पर इन अभिव्यक्तियों की संख्या में वृद्धि हुई है (पोल, 2013)। पिछले कुछ दशकों में जन-अभिव्यक्तियों की इस विशाल मात्रा ने शोधकर्ताओं का ध्यान अपनी ओर आकर्षित किया है क्योंकि इनके अध्ययन से प्राप्त महत्वपूर्ण सूचनाओं का प्रयोग शोध, राजनीतिक, सामाजिक, व्यावसायिक आदि विभिन्न क्षेत्रों में किया जा सकता है (Liu, 2012)। जन-अभिव्यक्तियों के विश्लेषण एवं वर्गीकरण हेतु विभिन्न पद्धतियों का प्रयोग किया जाता है जिसमें से भाव-विश्लेषण प्रमुख है (Pang & Lee, 2008; Arora, 2013)।

**भाव-विश्लेषण:**

भाव-विश्लेषण एक प्रक्रिया है जिसकी सहायता से जन-अभिव्यक्तियों को किसी विषय के प्रति सकारात्मक, नकारात्मक, अथवा निष्पक्ष के रूप में वर्गीकृत किया जाता है। किसी अभिव्यक्ति के भाव-विश्लेषण हेतु प्राकृतिक भाषा संसाधन (Natural Language Processing) के सिद्धांतों एवं संक्रियाओं का अनुपालन किया जाता है इसलिए भाव-विश्लेषण को प्राकृतिक भाषा संसाधन का एक अनुप्रयोग भी कहा जाता है।

**समस्या कथन:**

हिंदी भाषा की जन-अभिव्यक्तियों के भाव-विश्लेषण संबंधी अधिकतर कार्य केवल भावयुक्त शब्दों (Sentimental Words) के आधार पर ही किए गए हैं। हिंदी भाषा आधारित भाव-विश्लेषण की शुद्धता एवं गुणवत्ता में कमी का यह भी एक प्रमुख कारण है। जन-अभिव्यक्तियों के भाव-विश्लेषण हेतु मुहावरों (Idioms) का भी प्रयोग किया जाए तो हिंदी भाषा आधारित भाव-विश्लेषण की शुद्धता एवं गुणवत्ता में वृद्धि हो सकती है क्योंकि मुहावरे भी भाव की दृष्टि से सकारात्मक, नकारात्मक अथवा निष्पक्ष हो सकते हैं।

**उदाहरण**

1. दोहरे मापदंड वाले दलों (भाजपा-कांग्रेस) को अब समझना होगा। केजरीवाल ने **चौके-छक्के मार** दिए तो परेशानी हो जाएगी।
2. नोटबंदी से लोगों की **कमर टूट** गई है।

उपरोक्त वाक्यों में रेखांकित भाग मुहावरों को दर्शा रहे हैं जो भाव दृष्टि से क्रमशः **सकारात्मक** तथा **नकारात्मक** हैं। उक्त समस्या के समाधान हेतु प्रस्तुत शोध के माध्यम से एक संगणकीय प्रणाली प्रस्तावित की जा रही है जो हिंदी भाषा के वाक्यों में प्रयुक्त मुहावरों को स्वचालित रूप से पहचान कर उनका भाव-विश्लेषण कर सके।

**साहित्य पुनरवलोकन:**

इस चरण में उन शोधकार्यों का अध्ययन किया गया है जिनमें मुहावरों के भाव-विश्लेषण अथवा उनके संगणन हेतु कोई भी कार्य किया गया हो जिससे संगणकीय प्रणाली के विकास में सहायता मिल सके।

1. Agarwal & Ray, (2004) के द्वारा बंगाली भाषा के टैगरहित कार्पस से मुहावरेदार अभिव्यक्तियों के स्वतः प्रत्ययन हेतु रूपवैज्ञानिक विश्लेषण एवं सांख्यिकीय आधारित अभिगम प्रस्तावित किया गया।
2. Sharma & Goyal, (2011) के द्वारा हिंदी तथा पंजाबी के लोकोक्तियों को स्थायी (Static) एवं परिवर्तनीय (Dynamic) के रूप में विभाजित किया। स्थायी वर्ग की लोकोक्तियों को रेगुलर एक्सप्रेसंस तथा परिवर्तनीय वर्ग की लोकोक्तियों को उनके सभी संभव रूपों (inflections) के द्वारा नियंत्रित किया गया।
3. Gaule & Josan, (2012) ने मुहावरेदार अभिव्यक्तियों के स्वचालित पहचान हेतु एक नियम आधारित अभिगम (Rule based approach) प्रस्तावित किया।
4. Soliman *et al.*, (2013) ने अरबी भाषा के मुहावरों का 'SSWIL (Slang Sentimental Words and Idioms Lexicon)' नामक लेक्सीकॉन विकसित किया तथा इनके भाव-विश्लेषण हेतु SVM based classifier विधि प्रस्तावित की।
5. Bakliwal *et al.*, (2013) ने अंग्रेजी भाषा के मुहावरेदार अभिव्यक्तियों को भाव-विश्लेषण में सम्मिलित कर 57.58 % शुद्धता प्राप्त की।
6. Moreno-Ortiz *et al.*, (2013) के द्वारा स्पेनिश भाषा के बहुशब्दीय अभिव्यक्तियों के भाव-विश्लेषण हेतु 'Sentitext' नामक कोश आधारित अनुप्रयोग का निर्माण किया गया जिसमें तीन प्रमुख संसाधनों

- क्रमशः Individual Word Dictionary, Multiword Dictionary तथा Context Ruleset का प्रयोग किया गया।
7. Williams *et al.*, (2015) ने भाव-विश्लेषण के स्वचालित अभिगमों में मुहावरेदार अभिव्यक्तियों की भूमिका का अध्ययन करते हुए यह पाया कि इनके भाव-वर्गीकरण से पारंपरिक भाव-विश्लेषण के परिणाम में सुधार आ सकता है।
  8. Ibrahim *et al.*, (2015) ने अरबी भाषा के मुहावरों का एक लेक्सीकॉन 'AIPSeLEX' निर्मित किया एवं पाठ में से मुहावरों के प्रत्ययन हेतु n-gram तथा similarity measure विधि का प्रयोग किया।
  9. Borah & Sharma, (2016) ने असमिया भाषा के पदबंधों एवं मुहावरों की पहचान हेतु संदर्भ मुक्त व्याकरण (Context Free Grammar) की विधि प्रस्तावित की। इस विधि के अंतर्गत पदबंधों के निर्माण हेतु विभिन्न उत्पादन नियमों (Production Rules) को परिभाषित किया गया।

### संसाधन-निर्माण

हिंदी भाषा के भाव-विश्लेषण हेतु पूर्व-निर्मित संसाधनों (कोश, कार्पस, टैगसेट आदि) का अभाव है क्योंकि इस क्षेत्र में किए जा रहे कार्य अपनी प्रारंभिक अवस्था में हैं (Arora, 2013)। इसलिए प्रस्तावित संगणकीय प्रणाली के निर्माण हेतु निम्न संसाधनों का निर्माण किया गया है:

#### ■ कार्पस (Corpus)

यह शोध मुख्य रूप से मुहावरों पर केंद्रित है इसलिए हिंदी के 2000 ऐसे वाक्यों का कार्पस बनाया गया है जिनमें मुहावरों का प्रयोग हुआ है। संगणकीय प्रणाली के परीक्षण एवं मूल्यांकन हेतु इस कार्पस का प्रयोग किया गया है। इन वाक्यों के संकलन हेतु प्रमुख रूप से राजनीतिक एवं सामाजिक क्षेत्रों (Domains) का संदर्भ लिया है क्योंकि इन क्षेत्रों में मुहावरे अधिक संख्या में प्राप्त होते हैं।

#### ■ टैगसेट (Tagset)

इस शोध के अंतर्गत एक टैगसेट परिभाषित किया गया है जिसका प्रयोग प्रणाली के द्वारा इनपुट की गई अभिव्यक्तियों में प्रयुक्त भावयुक्त घटकों (भावयुक्त शब्दों एवं मुहावरों), उपवाक्यों एवं वाक्यों को उनकी ध्रुवणता के साथ चिन्हित करने के लिए किया गया है। इसके अतिरिक्त संयोजक अव्यय को दर्शाने हेतु भी इसमें टैग परिभाषित किया गया है। इस टैगसेट में कुल 10 टैग्स रखे गए हैं।

TAG	FULL FORM	TAG	FULL FORM
POS	POSITIVE	NM	NEGATION MARKER
NEG	NEGATIVE	IDM	IDIOM
NEU	NEUTRAL	SW	SENTIMENTAL WORD
SS	SENTENCE	POL	POLARITY
SC	CLAUSE	CONJ	CONJUNCTOR

तालिका 1: टैगसेट

#### ■ भाव-कोश

संगणकीय प्रणाली के ज्ञान-आधार के रूप एक डाटाबेस 'भाव-कोश' का निर्माण किया गया है। संगणकीय प्रणाली के द्वारा निवेशित पाठ में प्रयुक्त मुहावरों एवं भावयुक्त शब्दों की पहचान हेतु इसी डाटाबेस का प्रयोग किया जाता है। इस डाटाबेस को निम्न दो तालिकाओं के रूप में व्यवस्थित किया गया है:

1. **मुहावरा तालिका**— इस तालिका में हिंदी भाषा के 4220 मुहावरों को संग्रहित किया गया है। मैनुअल एनोटेशन विधि का प्रयोग करते हुए इस तालिका में संग्रहित मुहावरों का भाव-वर्गीकरण (सकारात्मक, नकारात्मक या निष्पक्ष के रूप में विभाजन) किया गया है।
2. **भावयुक्त शब्द तालिका**— इस तालिका में हिंदी भाषा के 2000 भावयुक्त शब्दों को संग्रहित किया गया है। भावयुक्त शब्दों के रूप भाववाचक संज्ञाओं, विशेषण एवं क्रिया-विशेषणों का संग्रहण प्रमुख रूप से किया गया है।

#### ■ नियम-समूह (Set of Rules)

प्रस्तुत शोध के अंतर्गत 09 व्याकरणिक नियमों का प्रतिपादन किया गया है। संगणकीय प्रणाली द्वारा निवेशित वाक्यों के ध्रुवणता-निर्धारण हेतु इन्हीं नियमों का अनुसरण किया गया है। प्रणाली में निवेशित वाक्यों की संरचना सरल या असरल हो सकती है। यहाँ उस वाक्य को असरल वाक्य माना गया है जिसमें उपवाक्यों का प्रयोग हुआ हो। इसलिए वाक्यों को उनकी संरचना के आधार विश्लेषित करने हेतु व्याकरणिक नियमों को दो कोटियों में विभाजित किया गया है:

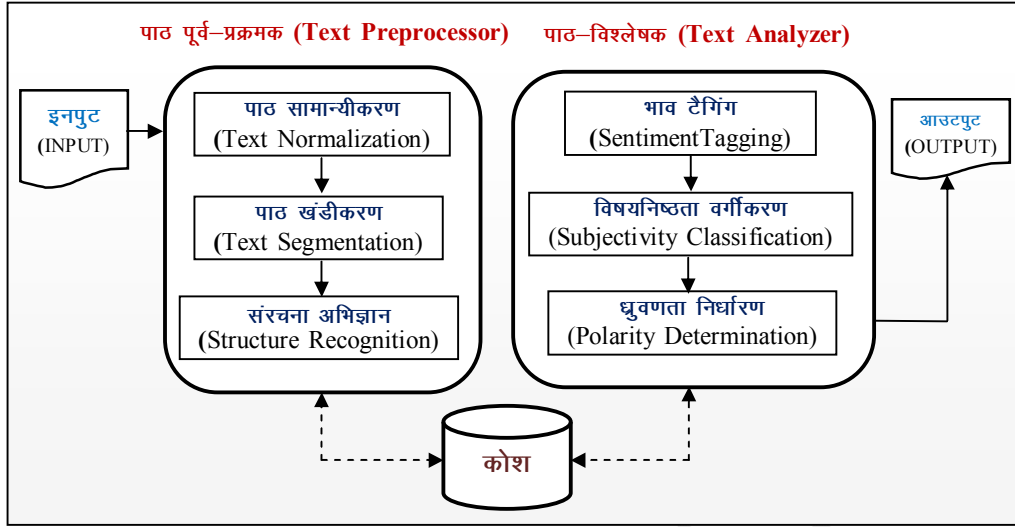
क्रं.	सरल वाक्यों की ध्रुवणता निर्धारण हेतु नियम
1.	if SS holds [IDM(POS)] Then POL(SS) = {POS};
2.	if SS holds [IDM(NEG)] Then POL(SS) = {NEG};
3.	if SS holds [IDM(NEU)] Then POL(SS) = {NEU};
4.	if SS holds [IDM(POS)] followed by [NM] Then POL(SS) = {NEG};
5.	if SS holds [IDM(NEG)] followed by [NM] Then POL(SS) = {POS};
असरल वाक्यों की ध्रुवणता निर्धारण हेतु नियम	
6.	If SS holds SC(1:POS) <CONJ> SC(2:POS) then POL(SS) = {POS};
7.	If SS holds SC(1:NEG) <CONJ> SC(2:NEG) then POL(SS) = {NEG};
8.	If SS holds SC(1:POS) <CONJ> SC(2:NEG) then POL(SS) = {NEG};
9.	If SS holds SC(1:NEG) <CONJ> SC(2:POS) then POL(SS) = {POS};

तालिका 4: भाव-विश्लेषक हेतु प्रतिपादित नियम-समूह

#### भाव-विश्लेषक: हिंदी मुहावरों के भाव-विश्लेषण हेतु संगणकीय प्रणाली

‘भाव-विश्लेषक’ एक संगणकीय प्रणाली है जिसे मुख्यतः हिंदी भाषा के मुहावरेदार अभिव्यक्तियों के भाव-विश्लेषण हेतु विकसित किया गया है। यह प्रणाली यूनिकोड आधारित देवनागरी लिपि अनूदित हिंदी पाठ पर ही कार्य करती है। सॉफ्टवेयर अभियांत्रिकी एवं प्राकृतिक भाषा संसाधन में प्रयुक्त सिद्धांतों एवं प्रविधियों को आधार बनाकर प्रणाली का निर्माण किया गया है।

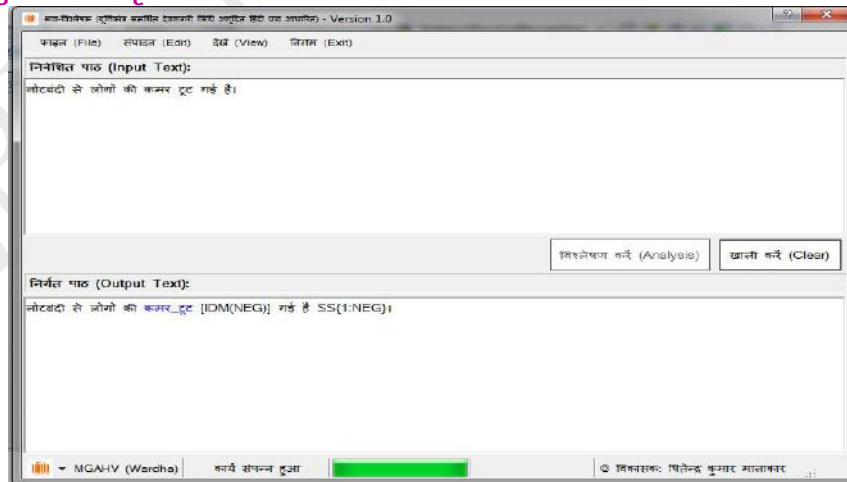
### ■ प्रणाली की संरचना एवं कार्यविधि



चित्र 1: भाव-विश्लेषक की संरचना एवं कार्यविधि

भाव-विश्लेषक में पाठ पूर्व-प्रक्रमक (Text Preprocessor) तथा पाठ-विश्लेषक (Text Analyzer) नामक दो मॉड्यूल (Modules) तथा 'कोश' नामक एक डाटाबेस का प्रयोग किया गया है। प्रत्येक मॉड्यूल के अंतर्गत अलग-अलग प्रक्रियाओं का प्रयोग किया गया है। प्रणाली के द्वारा सर्वप्रथम किसी निवेशित पाठ में उपस्थित अवांछनीय तत्वों का विलोपन किया जाता है। इसके पश्चात पाठ को वाक्य के स्तर पर खंडित किया जाता है। तत्पश्चात प्रत्येक वाक्य की संरचना का पता लगाया जाता है ताकि संरचना के अनुसार उनका अलग-अलग भाव-विश्लेषण किया जा सके। वाक्यों को उनकी संरचना के अनुसार पृथक करने के पश्चात प्रत्येक वाक्य में प्रयुक्त मुहावरों एवं भावयुक्त शब्दों की टैगिंग की जाती है। यदि वाक्य में कोई मुहावरा अथवा भावयुक्त शब्द प्राप्त होता है तो उसे विषयनिष्ठ के रूप में चिन्हित किया जाता है। इस संपूर्ण प्रक्रिया के पश्चात प्राप्त सभी विषयनिष्ठ वाक्यों की ध्रुवणता का निर्धारण किया जाता है। ध्रुवणता-निर्धारण हेतु प्रणाली द्वारा प्रतिपादित नियमों का अनुपालन किया जाता है।

### ■ प्रणाली का मुख्य अंतरापृष्ठ



चित्र 2: भाव-विश्लेषक का मुख्य अंतरापृष्ठ

उक्त चित्र में भाव-विश्लेषक के मुख्य अंतरापृष्ठ को दर्शाया गया है जिसमें सभी कंट्रोलस के नाम प्रयोक्ताओं की सुविधा की दृष्टि से हिंदी एवं अंग्रेजी में रखे गए हैं। उक्त चित्र की भाँति ही इनपुट बॉक्स में उपलब्ध पाठ का प्रणाली द्वारा भाव-विश्लेषण किया जाता है तत्पश्चात परिणाम को आउटपुट बॉक्स में प्रदर्शित किया जाता है। प्रणाली द्वारा निवेशित वाक्य में प्रयुक्त मुहावरे 'कमर टूटना' को उसकी ध्रुवणता के साथ टैग किया गया है एवं वाक्य के अंत में उसकी ध्रुवणता को दर्शाया गया है। प्रणाली द्वारा इसी प्रकार पाठ में प्रयुक्त मुहावरे की पहचान की जाती है एवं वाक्य की ध्रुवणता का निर्धारण किया जाता है।

### प्रणाली का मूल्यांकन

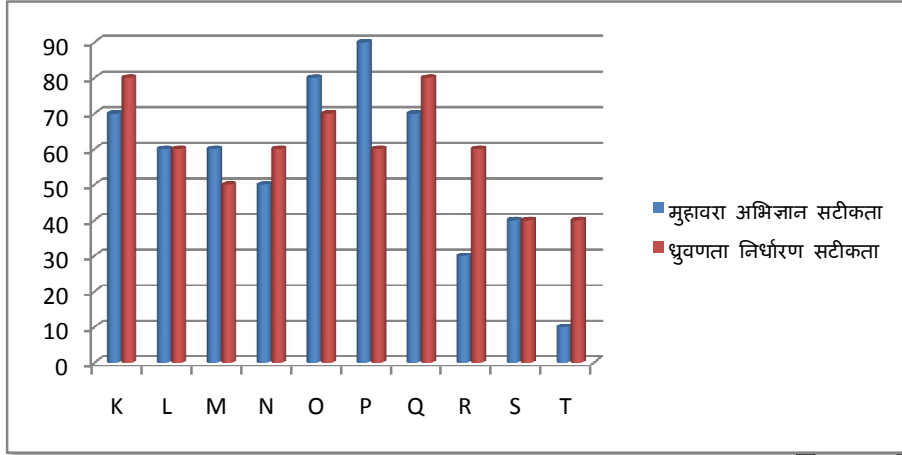
भाव-विश्लेषक के मूल्यांकन हेतु निर्मित कार्पस को परिचित (Known) एवं अपरिचित (Unknown) डाटासेट के रूप में वर्गीकृत किया गया है। परिचित डाटासेट से अभिप्राय उस डाटासेट है जिसका प्रयोग प्रणाली को विकसित करने हेतु किया गया है तथा अपरिचित डाटासेट के अंतर्गत उन वाक्यों को रखा गया है जिससे प्रणाली का परिचय नहीं हुआ है। परिचित एवं अपरिचित डाटासेट को पुनः दस अलग-अलग सेट के रूप में व्यवस्थित किया गया जिसमें से प्रत्येक सेट में 200 वाक्यों को रखा गया है जिसमें संतुलन की दृष्टि से 100 सकारात्मक तथा 100 नकारात्मक वाक्य हैं। इस प्रकार कुल 4000 वाक्यों का प्रयोग कर प्रणाली का मूल्यांकन किया गया है। प्रणाली का मूल्यांकन निम्न दो मापदंडों के आधार पर किया गया है:

1. **मुहावरा अभिज्ञान सटीकता** =  $(\sum Ri / \sum Ti) \times 100$  (जहाँ,  $Ri$  = पहचाने गए मुहावरों की कुल संख्या,  $Ti$  = कुल मुहावरों की संख्या)
2. **ध्रुवणता निर्धारण सटीकता** =  $(\text{सही निर्धारणों की संख्या} / \text{कुल वाक्य}) \times 100$

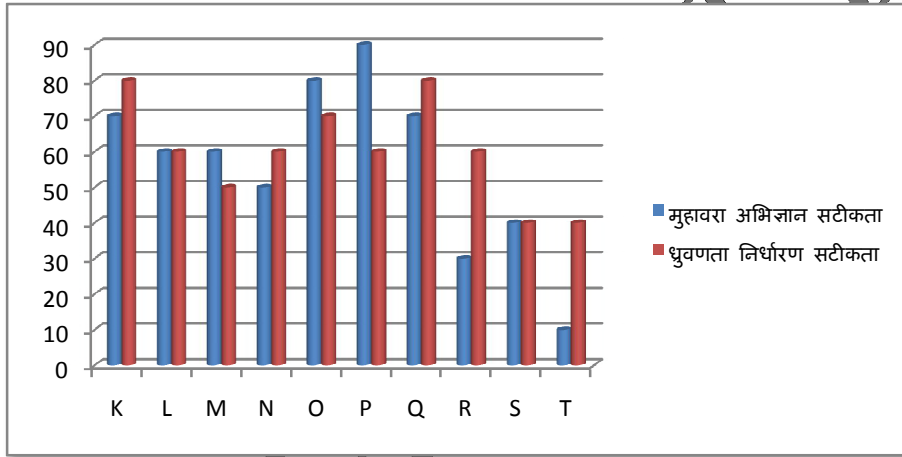
भाव-विश्लेषक के मूल्यांकन हेतु उपरोक्त कार्पस का प्रयोग किया गया है तथा प्राप्त परिणाम को निम्न तालिका के माध्यम से दर्शाया गया है:

प्रयुक्त अभिलक्षण	प्रणाली की औसत सटीकता			
	परिचित डाटासेट पर		अपरिचित डाटासेट पर	
	मुहावरा अभिज्ञान में	ध्रुवणता निर्धारण में	मुहावरा अभिज्ञान में	ध्रुवणता निर्धारण में
मुहावरा एवं भावयुक्त शब्द	61%	69%	49%	52%
भावयुक्त शब्द	-	42%	-	-

तालिका 4: भाव-विश्लेषक का मूल्यांकन



चित्र 2: परिचित डाटासेट पर भाव-विश्लेषक की सटीकता का ग्राफीय निरूपण



चित्र 3: अपरिचित डाटासेट पर भाव-विश्लेषक की सटीकता का ग्राफीय निरूपण

### निष्कर्ष

उपरोक्त तालिका से स्पष्ट है भावयुक्त शब्दों के साथ मुहावरों को भी अभिलक्षण के रूप में सम्मिलित करने पर भाव-विश्लेषक की ध्रुवणता निर्धारण सटीकता में 27% की वृद्धि हुई। प्रस्तुत शोध के माध्यम से हिंदी भाषा आधारित भाव-विश्लेषण हेतु महत्वपूर्ण संसाधनों का निर्माण एवं इसकी शुद्धता एवं गुणवत्ता को बढ़ाने का प्रयास किया गया है। भाव-विश्लेषक की क्षमता में वृद्धि हेतु डाटाबेस के आकार एवं नियमों की संख्या में वृद्धि की जा सकती है। भाव-विश्लेषक को अपरिचित कार्पस के अनुसार परिवर्तित/अनुकूलित कर बेहतर परिणाम प्राप्त किया जा सकता है।

### संदर्भ-सूची

- Agarwal, A. & Ray, B. (2004). Automatic Extraction of Multiword Expressions in Bengali: An Approach for Miserly Resource Scenarios. Proceedings of the ICON 2004, (pp. 165-172). Allied Publishers.
- Arora, P. (2013). Sentiment Analysis for Hindi Language. Hyderabad: IIIT.

- Bakliwal, A., Foster, J., Puil, v. d., O'Brien, R., Tounsi, L., & Hughes, M. (2013). Sentiment Analysis of Political Tweets: Towards an Accurate Classifier. Proceedings of the Workshop on Language in Social Media (LASM 2013), (pp- 49-58). Atlanta.
- Borah, S. K., & Sharma, U. (2016). Phrase and Idiom Identification in Assamese. 7th International conference on Intelligent Human Computer Interaction, IHCI 2015 (pp- 65- 69).
- Gaule, M., & Josan, G. S. (2012). Machine Translation of Idioms from English to Hindi. International Journal of Computational Engineering Research, (pp.50-54).
- Ibrahim, H. S., Abdou, S. M., & Gheith, M. (2015). Idioms-Proverbs Lexicon for Modern Standard Arabic and Colloquial Sentiment Analysis. International Journal of Computer Applications (0975- 8887), (pp.26-31).
- Lee, K.-J., Kim, J.-E., & Yun, B.-H. (2013). Extracting Multiword Sentiment Expressions by Using a Domain-Specific Corpus and a Seed Lexicon. ETRI Journal, (pp.838-848).
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers.
- Mittal, N., & Agarwal, B. (2013). Sentiment Analysis of Hindi Review based on Negation and Discourse. International Joint Conference on Natural Language Processing, (pp.45-50).
- Moreno-Ortiz, A., Perez-Hernandez, C., & Del-Olmo, M. A. (2013). Managing Multiword Expressions in a Lexicon-Based Sentiment Analysis System for Spanish. Proceedings of the 9th Workshop on Multiword Expressions (MWE 2013), (pp- 1-10). Atlanta, Georgia.
- Pang B. and Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval 2 (1-2), (pp.1-135).
- Sharma, M., & Goyal, V. (2011). Extracting Proverbs in Machine Translation from Hindi to Punjabi using Relational Data Approach. International Journal of Computer Science and Communication, (pp. 611-613).
- Soliman, T., Elmasry, M. A., Hedar, A.-R., & Doss, M. (2013). Mining social networks' Arabic slang comments. In Proceedings of IADIS European Conference on Data Mining 2013 (ECDM'13). Prague.
- Tsvetkov, Y., & Wintner, S. (2010). Extraction of Multi-word Expressions from Small Parallel Corpora. Coling 2010: Poster Volume, (pp. 1256-1264).
- Williams, L., Bannister, C., Arribas-Ayllon, M., Preece, A., & Spasic, I. (2015). The role of idioms in sentiment analysis. Expert Systems with Applications, (pp. 7375-7385).
- पोल, प्रशांत. (2013, जुलाई 7). कम्प्यूटर पर आईटी ने खोले भारतीय भाषाओं के रास्ते. पांचजन्य, (pp- 30-31).



**प्रितेन्द्र कुमार मालाकार**

शोधार्थी, सूचना एवं भाषा अभियांत्रिकी केंद्र, म.गां.अं.हिं.वि.,वर्धा (महाराष्ट्र)



**डॉ. हर्षलता पेटकर**

सहायक प्रोफेसर, सूचना एवं भाषा अभियांत्रिकी केंद्र, म.गां.अं.हिं.वि.,वर्धा (महाराष्ट्र)