NJESR/April-2024/Volume-6/Issue-4 E-ISSN-2582-5836
DOI-10.53571/NJESR.2024.6.4.41-48

हिंदी भाषा में राय जानना: एक सर्वेक्षण

किरण देवी
सहायक प्राध्यापिका (तदर्थ)
श्री लाल नाथ हिंदू महाविद्यालय, रोहतक
9053480856i@gmail.com

(Received:16March2024/Revised:8April2024/Accepted:18April2024/Published:26April2024)

अमूर्त: भारतीय भाषाई सर्वेक्षण ब्रिटिश राज के आधीन भारत का एक प्रमुख सर्वेक्षण परियोजना थी। इसके अन्तर्गत सन् १८९४ से लेकर सन् १९२८ तक भारतीय सिविल सेवा के अधिकारी जॉर्ज ग्रियर्सन के निर्देशन में कार्य हुआ। इसमें कुल ३६४ भारतीय भाषाओं एवं बोलियों का विस्तृत सर्वेक्षण किया गया। मनुष्य के जीवन में राय का बहुत महत्व है। इन मतों ने मनुष्यों को कार्यान्वित करने में सहायता की फैसले। जैसे-जैसे वेब का प्रभाव दिन-ब-दिन बढ़ता जा रहा है, वेब दस्तावेज़ों को एक नए रूप में देखा जा सकता है। मनुष्य के लिए राय का स्रोत. वेब में उपयोगकर्ताओं द्वारा उत्पन्न भारी मात्रा में जानकारी होती है!इस बड़ी मात्रा का विश्लेषण करने के लिए ब्लॉग, फ़ोरम प्रविष्टियों और सोशल नेटवर्किंग वेबसाइटों आदि के माध्यम से जानकारी के लिए एक ऐसी पद्धति विकसित करने की आवश्यकता है।जो स्वचालित रूप से उपलब्ध जानकारी को वर्गीकृत कर दे। इस डोमेन को सेंटीमेंट एनालिसिस और ओपिनियन माइनिंग कहा जाता है। राय खनन या भावना विश्लेषण एक प्राकृतिक भाषा प्रसंस्करण कार्य है जो विभिन्न पाठ्य रूपों जैसे समीक्षा, समाचार, से जानकारी प्राप्त करता है। ब्लॉग और उन्हें उनकी ध्रुवता के आधार पर सकारात्मक, नकारात्मक या तटस्थ के रूप में वर्गीकृत करें। लेकिन, आखिरी से कुछ वर्षों में वेब पर हिंदी भाषा में भारी वृद्धि देखी गई है। ओपिनियन माइनिंग में अनुसंधान अधिकतर अंग्रेजी भाषा में किया जाता है लेकिन ओपिनियन माइनिंग हिंदी में करना बहुत महत्वपूर्ण है। भाषा के साथ-साथ हिंदी में भी बड़ी मात्रा में जानकारी वेब पर उपलब्ध है। यह पेपरहिन्दी भाषा में किये गये कार्यों का एक अवलोकन देता है।

प्रमुख शब्द :- राय खनन, भावना विश्लेषण, समीक्षा, हिंदी भाषा वर्डनेट।

परिचय

हिन्दी विश्व की एक प्रमुख भाषा है और भारत की एक राजभाषा है। केन्द्रीय स्तर पर भारत में सह-आधिकारिक भाषा अंग्रेज़ी है। हिन्दी के मानकीकृत रूप को मानक हिन्दी कहा जाता है। मानक हिन्दी में संस्कृत के तत्सम तथा तद्भव शब्दों का प्रयोग अधिक है और अरबी-फ़ारसी शब्द कम हैं। हिन्दी संवैधानिक रूप से भारत की राजभाषा और भारत की सबसे

अधिक बोली और समझी जाने वाली भाषा है। हिन्दी भारत की राष्ट्रभाषा नहीं बल्कि राजभाषा (official language) है क्योंकि भारत के संविधान में राष्ट्रभाषा का कहीं उल्लेख या चर्चा नहीं है।[7][8] एथनोलॉग के अनुसार, हिन्दी विश्व की तीसरी सबसे अधिक बोली जाने वाली भाषा है।[9] विश्व आर्थिक मंच की गणना के अनुसार यह विश्व की दस शक्तिशाली भाषाओं में से एक है। आजकल अधिकांश लोग अपनी भावनाओं, अनुभवों और विचारों को साझा करना चाहते हैं। अब लोग आमतौर पर ब्लॉग, फ़ोरम, ई-न्यूज़, समीक्षा चैनल और सोशल का उपयोग करते हैं । अपने विचार और राय व्यक्त करने के लिए फेसबुक, ट्विटर जैसे नेटवर्किंग प्लेटफॉर्म। विश्ववाइड वेब जनमत एकत्र करने में महत्वपूर्ण भूमिका निभाता है। ये राय के लिए बहुत उपयोगी हैं । व्यावसायिक संगठन, क्योंकि उन्हें अपने उपयोगकर्ता की भावनाओं/राय के बारे में पता होगा । उनके उत्पाद और ग्राहकों को निर्णय लेने में भी मदद करते हैं। बड़ी मात्रा में उपयोगकर्ता सामग्रीवेब पर हर दिन डेटा उत्पन्न होता है, इस प्रकार डेटा को माइन किया जाता है और उपयोगकर्ता की भावनाओं की पहचान की जाती है,इच्छाएं, पसंद-नापसंद एक अहम काम है। भावना विश्लेषण एक प्राकृतिक भाषा प्रसंस्करण कार्य है जो अभिविन्यास खोजने से संबंधित है।किसी विषय के संबंध में पाठ के एक टुकड़े में राय [8]। यह विभिन्न पाठों से जानकारी प्राप्त करता है । समीक्षा, समाचार और ब्लॉग जैसे रूप बनाता है और उन्हें उनकी ध्रुवता के आधार पर वर्गीकृत करता है । सकारात्मक, नकारात्मक या तटस्थ. यह पाठ को व्यक्तिपरक और के स्तर पर वर्गीकृत करने पर केंद्रित है वस्तुनिष्ठ प्रकृति. विषयपरकता इंगित करती है कि पाठ में राय सामग्री शामिल है/है वस्तुनिष्ठता इंगित करती है कि पाठ बिना राय वाली सामग्री के है [2]। हिन्दी और इसकी बोलियाँ सम्पूर्ण भारत के विविध राज्यों में बोली जाती हैं। भारत और अन्य देशों में भी लोग हिन्दी बोलते, पढ़ते और लिखते हैं।[17] फ़िजी, मॉरिशस, गयाना, सूरीनाम, नेपाल और संयुक्त अरब अमीरात में भी हिन्दी या इसकी मान्य बोलियों का उपयोग करने वाले लोगों की बड़ी संख्या मौजूद है।[2] फरवरी 2019 में अबू धाबी में हिन्दी को न्यायालय की तीसरी भाषा के रूप में मान्यता मिली। 'देशी', 'भाखा' (भाषा), 'देशना वचन' (विद्यापति), 'हिन्दवी', 'दक्खिनी', 'रेख़ता', 'आर्यभाषा' (दयानन्द सरस्वती), 'हिन्दुस्तानी', 'खड़ी बोली',[21] 'भारती' आदि हिन्दी के अन्य नाम हैं, जो विभिन्न ऐतिहासिक कालखण्डों में एवं विभिन्न सन्दर्भों में प्रयुक्त हुए हैं। हिन्दी, यूरोपीय भाषा-परिवार के अन्दर आती है। ये हिन्द ईरानी शाखा की हिन्द आर्य उपशाखा के अन्तर्गत वर्गीकृत है। एथ्नोलॉग (2022, 25वां संस्करण) की रिपोर्ट के अनुसार विश्वभर में हिंदी को प्रथम और द्वितीय भाषा के रूप में बोलने वाले लोगों की संख्या के आधार पर हिंदी विश्व की तीसरी सबसे अधिक बोली जाने वाली भाषा है।

नामोत्पत्ति

हिन्दी शब्द का सम्बन्ध संस्कृत शब्द 'सिन्धु' से माना जाता है। 'सिन्धु' सिन्धु नदी को कहते थे और उसी आधार पर उसके आस-पास की भूमि को सिन्धु कहने लगे। यह सिन्धु शब्द ईरानी में जाकर 'हिन्दू ' , हिन्दी और फिर 'हिन्द' हो गया। बाद में ईरानी धीरे-धीरे भारत के अधिक भागों से परिचित होते गए और इस शब्द के अर्थ में विस्तार होता गया तथा हिन्द शब्द पूरे भारत का वाचक हो गया। इसी में ईरानी का ईक प्रत्यय लगने से (हिन्द+ईक) 'हिन्दीक' बना जिसका अर्थ है 'हिन्द का' । यूनानी शब्द 'इण्डिका' या लैटिन 'इण्डेया' या अंग्रेजी शब्द 'इण्डिया' आदि इस 'हिन्दीक' के ही दूसरे रूप हैं। हिन्दी भाषा के लिए इस शब्द का प्राचीनतम प्रयोग शरफ़ुद्दीन यज्दी के 'जफ़रनामा' (1424) में मिलता है। प्रमुख उर्दू लेखकों ने 19वीं सदी की सूचना तक अपनी भाषा को हिंदी या हिंदवी के रूप में संदर्भित करते रहे।[2-3] प्रोफ़ेसर महावीर सरन जैन ने अपने "हिन्दी एवं उर्दू का अद्वैत" शीर्षक आलेख में हिन्दी की व्युत्पत्ति पर विचार करते हुए कहा है कि ईरान की प्राचीन भाषा अवेस्ता में 'स्' ध्वनि नहीं बोली जाती थी बल्कि 'स्' को 'ह्' की तरह बोला जाता था। जैसे संस्कृत शब्द 'असुर' का अवेस्ता में सजाति समकक्ष शब्द 'अहुर' था। अफ़ग़ानिस्तान के बाद सिन्धु नदी के इस पार हिन्दुस्तान के पूरे इलाके को प्राचीन फ़ारसी साहित्य में भी 'हिन्द', 'हिन्दुश' के नामों से पुकारा गया है तथा यहाँ की किसी भी वस्तु, भाषा, विचार को विशेषण के रूप में 'हिन्दीक' कहा गया है जिसका मतलब है 'हिन्द का' या 'हिन्द से'। यही 'हिन्दीक' शब्द अरबी से होता हुआ ग्रीक में 'इण्डिके', 'इण्डिका', लैटिन में 'इण्डेया' तथा अंग्रेज़ी में 'इण्डिया' बन गया। दूसरा एक भावना के मुताबिक़, अरबी هندية हिन्दीया लफ़्ज़ के साधारण लैटिनी कृत रूप है इंडिया India. जैसे Hindiyyah (मूल अरबी)> Hindia साधारणीकृत >India लैटिनीकृत. अरबी एवं फ़ारसी साहित्य में भारत (हिन्द) में बोली जाने वाली भाषाओं के लिए 'ज़ुबान-ए-हिन्दी' पद का उपयोग हुवा है। भारत आने के बाद अरबी-फ़ारसी बोलने वालों ने 'ज़ुबान-ए-हिन्दी', 'हिन्दी ज़ुबान' अथवा 'हिन्दी' का प्रयोग दिल्ली-आगरा के चारों ओर बोली जाने वाली भाषा के अर्थ में किया।

ओपिनियन माइनिंग के घटक

ओपिनियन माइनिंग के मुख्य रूप से तीन घटक हैं, ये हैं [9]:

- राय धारक: राय धारक एक विशेष राय का धारक होता है; यह एक व्यक्ति हो सकता है या कोई संगठन जो राय रखता है. ब्लॉग और समीक्षाओं के मामले में, राय धारक वे व्यक्ति होते हैं जो ये समीक्षाएँ या ब्लॉग लिखते हैं।
- राय वस्तु: राय वस्तु वह वस्तु है जिस पर राय धारक व्यक्त कर रहा है
- राय।

- ओपिनियन ओरिएंटेशन: किसी वस्तु पर किसी राय का ओपिनियन ओरिएंटेशन यह निर्धारित करता है कि क्या किसी वस्तु के बारे में किसी राय धारक की राय सकारात्मक, नकारात्मक या तटस्थ होती है।

उदाहरण के लिए "इस मोबाइल का कै मेरा अच्छा है।" इस रिव्यू में जिस शख्स ने ये लिखा है समीक्षा राय धारक है. यहां ओपिनियन ऑब्जेक्ट मोबाइल और ओपिनियन का कैमरा है शब्द "अछा" है जो सकारात्मक रूप से उन्मुख है। सिमेंटिक ओरिएंटेशन निर्धारित करने का कार्य है [6]

राय खनन के स्तर

भावना विश्लेषण के क्षेत्र में अनुसंधान विभिन्न स्तरों पर किया जाता है जो इस प्रकार हैं [9]-

- दस्तावेज़ स्तर- दस्तावेज़ स्तर पूरे दस्तावेज़ को एकल ध्रुवता के रूप में वर्गीकृत करता है सकारात्मक, नकारात्मक या तटस्थ.
- वाक्य स्तर - वाक्य स्तर वाक्य स्तर पर दस्तावेजों का विश्लेषण करता है।
- वाक्यों का व्यक्तिगत रूप से विश्लेषण किया जाता है और सकारात्मक, नकारात्मक या तटस्थ के रूप में वर्गीकृत किया जाता है।
- पहलू स्तर- पहलू स्तर का विश्लेषण बहुत गहराई तक जाता है और इसकी पहचान करने से संबंधित है किसी दिए गए दस्तावेज़ के लिए एक वाक्य में विशेषताएँ और विशेषताओं का विश्लेषण करें और तदनुसार सकारात्मक, नकारात्मक या तटस्थ उन्हें वर्गीकृत करें।

मौजूदा अनुसंधान कार्य

ओपिनियन माइनिंग के क्षेत्र में हिंदी भाषा में थोड़ा सा काम हुआ है। बहुत ही पहला शोध हिंदी, बंगाली और मराठी भाषा में किया गया है। अमिताव दास और बंदोपाध्याय [1] ने बंगाली भाषा के लिए सेंटीवर्डनेट विकसित किया। शब्द स्तरीय शाब्दिक-स्थानान्तरण तकनीक को अंग्रेजी-बंगाली का उपयोग करके अंग्रेजी सेंटिवर्डनेट में प्रत्येक प्रविष्टि पर लागू किया गया है बंगाली सेंटिवर्डनेट प्राप्त करने के लिए शब्दकोश। उनके द्वारा 35,805 बंगाली प्रविष्टियाँ वापस कर दी गई हैं, दास और बंदोपाध्याय [2] ने एक शब्द के भाव का अनुमान लगाने के लिए चार रणनीतियाँ तैयार कीं। पहला दृष्टिकोण, उनके द्वारा एक इंटरैक्टिव गेम प्रस्तावित है जो शब्दों को एनोटेट करता है उनकी ध्रुवता. दूसरे दृष्टिकोण में, वे अंग्रेजी और भारतीय के लिए द्विभाषी शब्दकोश का उपयोग करते हैं ध्रुवीयता निर्धारित करने के लिए भाषाएँ। तीसरे दृष्टिकोण में, वे वर्डनेट का उपयोग करते हैं और उपयोग करते हैं पर्यायवाची और विपरीतार्थी संबंध, ध्रुवता निर्धारित करने के लिए। चौथे दृष्टिकोण में; सीखना ध्रुवीयता निर्धारित करने के लिए पूर्व-एनोटेटेड कॉर्पोरा से होता है।

दीपांकर दास और बंदोपाध्या[11], बंगाली में भावनात्मक अभिव्यक्तियों की पहचान उनके द्वारा की जाती है कोष. भावनात्मक अभिव्यक्तियों की पहचान भावनात्मक घटकों के आधार पर की जाती है जैसे धारक, तीव्रता और विषय। वाक्य स्तरीय व्याख्या के लिए उनके द्वारा शब्दों को छह वर्गों में वर्गीकृत किया गया भावना वर्ग और तीन प्रकार की तीव्रताओं के साथ जोशी एट अल. [3] हिंदी भाषा के लिए एक फ़ॉलबैक रणनीति का प्रस्ताव रखा। यह रणनीति तीन का अनुसरण करती है।

दृष्टिकोण: भाषा में भावना विश्लेषण, मशीनी अनुवाद और संसाधन आधारित भावना विश्लेषण. उन्होंने हिंदी सेंटीवर्डनेट (HSWN) पर आधारित एक शाब्दिक संसाधन विकसित किया। इसका अंग्रेजी प्रारूप. दो शाब्दिक संसाधनों (अंग्रेजी सेंटिवर्डनेट और अंग्रेजी-हिंदी) का उपयोग करके वर्डनेट लिंकिंग [20]) H-SWN (हिन्दी-सेंटीवर्डनेट) इनके द्वारा बनाया गया था। वर्डनेट का उपयोग करके लिंक करने पर, एचएसडब्ल्यूएन प्राप्त करने के लिए अंग्रेजी सेंटीवर्डनेट के शब्दों को समकक्ष हिंदी शब्दों से बदल दिया गया। उनके द्वारा प्राप्त अंतिम सटीकता 78.14 है। ग्राफ आधारित विधि का उपयोग करके बकलीवाल और अन्य ने शब्दकोष बनाया। वे इसका उपयोग करके यह निर्धारित करते हैं। सरल ग्राफ़ ट्रैवर्सल कि कैसे पर्यायवाची और एंटोनिम संबंध उत्पन्न करने के लिए उपयोग किया जा सकता है। व्यक्तिपरकता शब्दकोष. उनके प्रस्तावित एल्गोरिदम ने लगभग 79% सटीकता हासिल की समीक्षाओं का वर्गीकरण और मानव व्याख्या के साथ 70.4% सहमति। मुखर्जी एट अल. [6] दिखाया गया कि बैग-ऑफ-वर्ड्स मॉडल में प्रवचन मार्करों का समावेश भावना वर्गीकरण सटीकता में 2-4% सुधार होता है। बाकलीवाल एट अल. [5] एक विधि प्रस्तावित की हिंदी समीक्षाओं को सकारात्मक या नकारात्मक के रूप में वर्गीकृत करना। उन्होंने एक नया स्कोरिंग फ़ंक्शन और परीक्षण तैयार किया। दो अलग-अलग दृष्टिकोण. उन्होंने सरल एन-ग्राम और पीओएस टैग किए गए एनग्राम दृष्टिकोण के संयोजन का भी उपयोग किया।

दृष्टिकोण

ओपिनियन माइनिंग में दो प्रकार की तकनीकों का उपयोग किया जाता है:

- पर्यवेक्षित अध्ययन
- बिना पर्यवेक्षण के सीखना

पर्यवेक्षित अध्ययन

पर्यवेक्षित शिक्षण एक मशीन लर्निंग दृष्टिकोण है[6], दस्तावेज़ सेट दो प्रकार के होते हैं। इस दृष्टिकोण में आवश्यक: प्रशिक्षण सेट और परीक्षण सेट। एक प्रशिक्षण सेट ने क्लासिफायरियर को प्रशिक्षित किया।प्रशिक्षण डेटा और एक परीक्षण सेट की मदद से परीक्षण डेटा देकर क्लासिफायरियर

का परीक्षण किया जाता है। बड़ी संख्या में मशीन लर्निंग तकनीकें उपलब्ध हैं, उदाहरण के लिए, नैवे बेयस वर्गीकरण और समर्थन वेक्टर मशीनें जो राय को वर्गीकृत करती हैं।

बिना पर्यवेक्षण के सीखना

बिना पर्यवेक्षित शिक्षण, क्लस्टरिंग जैसे इनपुट के एक सेट को मॉडल करता है, लेबल के दौरान ज्ञात नहीं होता है। [6]. अनसुपरवाइज्ड लर्निंग में वर्गीकरण एक की राय की तुलना करके किया जाता है। शब्दकोषों के विरुद्ध दिया गया पाठ, जिनके भाव-मूल्य उनके प्रयोग से पहले निर्धारित होते हैं। शब्दकोषों की सहायता से दस्तावेजों के भाव-विन्यास का निर्धारण किया जाता है। भावना अभिविन्यास एक अप्रशिक्षित शिक्षा है क्योंकि राय के लिए किसी पूर्व प्रशिक्षण की आवश्यकता नहीं होती है । उदाहरण के लिए, यह पता लगाने के लिए कि दस्तावेज़ में सकारात्मक राय है या नकारात्मक राय, दस्तावेज़ में सभी राय शब्द पहले निकाले जाते हैं और उनकी ध्रुवता निर्धारित की जाती है। शब्दकोष की मदद से जिसमें ध्रुवता के साथ राय वाले शब्द शामिल हैं। जब ध्रुवता सभी राय शब्दों का निर्धारण किया जाता है, फिर यह जांचा जाता है कि दस्तावेज़ में और भी शब्द हैं या नहीं सकारात्मक शब्द या अधिक नकारात्मक शब्द. यदि दस्तावेज़ में अधिक सकारात्मक शब्द हैं। नकारात्मक शब्द, दस्तावेज़ का भावना अभिविन्यास सकारात्मक है अन्यथा यह नकारात्मक है। पार्ट-ऑफ-स्पीच (पीओएस) टैग का उपयोग राय वाले शब्दों को निर्धारित करने के लिए किया जाता है। भाषण का एक भाग टैगर (पीओएस टैगर) एक सॉफ्टवेयर है जो इनपुट टेक्स्ट को पढ़ता है और प्रत्येक शब्द के लिए भाषण के कुछ हिस्सों को निर्दिष्ट करता है।[8]

निष्कर्ष

राय खनन एक उभरता हुआ अनुसंधान क्षेत्र है और यह बहुत महत्वपूर्ण है क्योंकि मनुष्य हैं। आजकल काफी हद तक वेब पर निर्भर हैं। हिंदी भाषा में उपयोगकर्ता-जनित सामग्री में वृद्धि विभिन्न शैलियों-समाचार, संस्कृति, कला, खेल आदि में डेटा का पता लगाने और खनन करने का अवसर उपलब्ध है। उपभोक्ताओं को बेहतर सेवाएँ और सुविधाएँ प्रभावी ढंग से प्रदान करना। राय खनन बड़ा है। शॉपिंग जैसे एप्लिकेशन क्षेत्र, जहां Flipkart.com, amazon.com आदि जैसी वेबसाइटें अनुमति देती हैं।ग्राहक अपनी वेबसाइटों पर अपनी राय व्यक्त कर सकते हैं जिससे अन्य ग्राहकों को निर्णय लेने में मदद मिलती है। उत्पाद खरीदना है या नहीं. मनोरंजन, जहां लोग आसानी से आलोचकों को देख सकें और दर्शकों की उनकी पसंदीदा फिल्मों और शो की ऑनलाइन समीक्षाएँ। मार्केटिंग, अब हर संगठन और निजी कंपनियाँ अपने ग्राहकों को उनके उत्पादों से संबंधित समीक्षाएँ लिखने की अनुमति देती हैं। ऐसी वेबसाइटें जो सर्वेक्षण करने की आवश्यकता को समाप्त कर देती हैं।

ओपिनियन माइनिंग में बड़ी मात्रा में काम अंग्रेजी भाषा में किया गया है, क्योंकि अंग्रेजी एक है। वैश्विक भाषा, लेकिन अन्य भाषाओं में भी ओपिनियन माइनिंग करने की आवश्यकता है। बड़ा वेब पर अन्य भाषाओं की ढेर सारी सामग्री उपलब्ध है जिसका उपयोग करने की आवश्यकता है। राय निर्धारित करें. हिंदी भारत की राष्ट्रीय भाषा है, बड़ी मात्रा में हिंदी सामग्री है। वेब पर उपलब्ध है. गूगल तथा याहू सर्च इंजन भी हिन्दी में सेवाएँ प्रदान करते हैं।भाषा; इन सर्च इंजनों पर कोई भी जानकारी हिंदी में लिख और खोज सकता है। राय हिंदी भाषा में माइनिंग आवश्यक है। पिछले कुछ वर्षों से शोधकर्ताओं ने प्रदर्शन किया है। हिंदी भाषा में ओपिनियन माइनिंग। इस पेपर में हिंदी आधारित ओपिनियन माइनिंग का एक सिंहावलोकन है। दिया है, जो हिंदी भाषा में किए गए मौजूदा शोधों के आधार पर दिया गया है। हिंदी भाषा में खनन करना कोई आसान काम नहीं है, क्योंकि भारतीय भाषाओं की प्रकृति अलग-अलग होती है। लिपि, प्रतिनिधित्व स्तर और भाषाई विशेषताओं आदि के संदर्भ में बहुत कुछ भारतीय भाषाओं के व्यवहार को समझें, हिंदी भाषा के लिए राय का खनन बड़े पैमाने पर काम करने की जरूरत है।

हवाला

[1]."नगर राजभाषा कार्यान्वयन समितियों की सूची". मूल से 14 अगस्त 2020 को पुरालेखित. अभिगमन तिथि 22 अप्रैल 2020.

[2]."प्रचार समिति वर्धा : एक विहंगावलोकन". मूल से 1 अक्तूबर 2022 को पुरालेखित. अभिगमन तिथि 1 अक्तूबर 2022.

[3].शाहदे, ए.के., वाल्से, के.एच., और ठाकरे, वी.एम. (2022)। बहुभाषी राय खनन पर एक व्यापक सर्वेक्षण। मोबाइल कंप्यूटिंग और सतत सूचना विज्ञान: आईसीएमसीएसआई 2022 की कार्यवाही , 43-55।

[4].सोनी, वी.के., और सेलोट, एस. (2022)। हिंदी भाषा के लिए भावना विश्लेषण के क्षेत्र में गहन शिक्षण तकनीकों का सर्वेक्षण। आई-मैनेजर जर्नल ऑन कंप्यूटर साइंस , 10 (1)।

[5].थोरात, एम., और गाइड, एन.एफ. (2022)। हिंदी भाषा के लिए भावना विश्लेषण पर समीक्षा पत्र। ग्रेंज इंटरनेशनल जर्नल ऑफ इंजीनियरिंग एंड टेक्नोलॉजी (जीआईजेईटी) , 8 (1)।

[6].शाहदे, ए.के., वाल्से, के.एच., और ठाकरे, वी.एम. (2022, फरवरी)। एक नया डीप लर्निंग दृष्टिकोण आधारित बहुभाषी राय खनन। 2022 में इलेक्ट्रिकल, इलेक्ट्रॉनिक्स, सूचना और संचार प्रौद्योगिकी (ICEEICT) पर पहला अंतर्राष्ट्रीय सम्मेलन (पृष्ठ 1-4)।

[7].शाहदे, ए.के., वाल्से, के.एच., ठाकरे, वी.एम., और अतीक, एम. (2023)। सोशल मीडिया प्रवचनों के लिए बहुभाषी राय खनन: एडम ऑप्टिमाइज़र के साथ डीप लर्निंग आधारित

हाइब्रिड फ़ाइन-ट्यून्ड स्मिथ एल्गोरिदम का उपयोग करने वाला एक दृष्टिकोण। इंटरनेशनल जर्नल ऑफ़ इन्फ़ॉर्मेशन मैनेजमेंट डेटा इनसाइट्स , 3 (2), 100182।

[8].कुमारी, आर., साह, डी.के., सेंगिज़, के., इवकोविक, एन., और बालाजी, पी. (2023)। मेडिकल रिव्यू सेंटीमेंट एनालिसिस के लिए एशियाई हिंदी भाषाओं के लिए स्वचालित ग्राफ़ निर्माण और विभिन्न प्रकार के LSTM की खोज। एशियाई और कम संसाधन वाली भाषा सूचना प्रसंस्करण पर ACM लेनदेन ।

[9].शाहदे, ए., वाल्से, डी.के., और वी.एम., टी. (2022, मई)। बहुभाषी राय खनन में उपयोग किए जाने वाले डेटासेट पर एक अध्ययन। इनोवेटिव कंप्यूटिंग और संचार (ICICC) पर अंतर्राष्ट्रीय सम्मेलन की कार्यवाही में ।

[10]. यादव, पी., कौर, आर., और कुमार, ए. (2023, अगस्त)। हिंदी भाषा में सोशल मीडिया सेंटीमेंटल एनालिसिस पर एक अध्ययन। 2023 इंटरनेशनल कॉन्फ्रेंस ऑन सर्किट पावर एंड कंप्यूटिंग टेक्नोलॉजीज (ICCPCT) में (पृष्ठ 1395-1400)।