

समकालीन हिंदी साहित्य में सामाजिक परिवर्तन और मानवीय संवेदनाएँ

डॉ. निर्मला परगनिहा

सहायक प्राध्यापक

अग्रसेन कॉलेज

धनोरा, दुर्ग, छत्तीसगढ़



<http://www.gjirp.org/> || Vol. 2 No. 3 (2026): Special August Issue

Date of Submission: 15-07-2026

Date of Acceptance: 31-07-2026

Date of Publication: 20-08-2026

सारांश— समकालीन हिंदी साहित्य भारतीय समाज में तीव्र गति से घटित हो रहे सामाजिक, सांस्कृतिक, आर्थिक और तकनीकी परिवर्तनों का संवेदनशील अभिलेख प्रस्तुत करता है। वर्ष 2021 के बाद डिजिटल संचार, महामारीय सामाजिक अनुभव, रोजगार-अनिश्चितता, प्रवासन, लैंगिक विमर्श और बदलते पारिवारिक संबंधों ने साहित्यिक संवेदना को नए संदर्भ प्रदान किए हैं। प्रस्तुत अध्ययन का केंद्रीय उद्देश्य समकालीन हिंदी साहित्य में सामाजिक परिवर्तन और मानवीय संवेदनाओं के अंतर्संबंध को संगणकीय पाठ-विश्लेषण तथा आलोचनात्मक साहित्यिक व्याख्या के संयुक्त आधार पर समझना है। अध्ययन का विशिष्ट शोध-अंतर यह है कि सामाजिक परिवर्तन संबंधी हिंदी साहित्य के अधिकांश अध्ययनों में गुणात्मक आलोचना प्रमुख रही है, जबकि हिंदी की भाषिक विशिष्टताओं को ध्यान में रखते हुए कृत्रिम बुद्धिमत्ता आधारित भाव, विषय और संदर्भ विश्लेषण का समेकित उपयोग अपेक्षाकृत सीमित है। इसी अंतर को संबोधित करने के लिए अध्ययन में 2021-2026 की अवधि से संबद्ध चयनित डिजिटल हिंदी साहित्यिक पाठों को सामाजिक-संवेदनात्मक संकेतकों के आधार पर विश्लेषित करने की रूपरेखा विकसित की गई है।

मुख्य शब्द— समकालीन हिंदी साहित्य; सामाजिक परिवर्तन; मानवीय संवेदना; डिजिटल मानविकी; प्राकृतिक भाषा संसाधन; कृत्रिम बुद्धिमत्ता परिचय

साहित्य केवल सामाजिक परिस्थितियों का निष्क्रिय प्रतिबिंब नहीं होता, बल्कि वह उन अनुभवों, अंतर्विरोधों और भावनात्मक प्रतिक्रियाओं को अभिव्यक्ति देता है जिन्हें औपचारिक सामाजिक आँकड़े पूर्णतः दर्ज नहीं कर पाते। हिंदी साहित्य की आधुनिक परंपरा में सामाजिक यथार्थ, वर्गीय विषमता, स्त्री-अनुभव, जातिगत संरचना, ग्रामीण जीवन,

शहरीकरण, विस्थापन और राजनीतिक-सांस्कृतिक चेतना लंबे समय से प्रमुख विषय रहे हैं। किंतु इक्कीसवीं सदी के तीसरे दशक में सामाजिक परिवर्तन की गति और उसका स्वरूप दोनों अधिक जटिल हुए हैं। डिजिटल प्रौद्योगिकी अब केवल संचार का माध्यम नहीं रह गई है; उसने व्यक्ति की सामाजिक उपस्थिति, आत्म-अभिव्यक्ति, संबंध-निर्माण, रोजगार, शिक्षा और सांस्कृतिक उपभोग के तरीकों को भी प्रभावित किया है।

2021-2026 का कालखंड इस दृष्टि से विशेष रूप से महत्वपूर्ण है। कोविड-19 महामारी से उत्पन्न सामाजिक व्यवधान के बाद सामान्य जीवन की पुनर्स्थापना के साथ कई पुराने सामाजिक प्रश्न नए रूपों में सामने आए। महामारी के अनुभव ने मृत्युबोध, अकेलेपन, पारिवारिक निकटता और दूरी, असुरक्षा, स्वास्थ्य-संबंधी चिंता तथा सामाजिक सहयोग के अर्थों को पुनर्परिभाषित किया। इसके साथ ही ऑनलाइन कार्य, डिजिटल शिक्षा, सोशल मीडिया आधारित संवाद, प्लेटफॉर्म अर्थव्यवस्था और कृत्रिम बुद्धिमत्ता के विस्तार ने मनुष्य और समाज के संबंधों में नई परतें जोड़ीं। इन परिवर्तनों का प्रभाव साहित्यिक लेखन की विषयवस्तु और संवेदनात्मक संरचना दोनों पर देखा जा सकता है।

समकालीन हिंदी रचनात्मक परिदृश्य में सामाजिक परिवर्तन को केवल विकास अथवा आधुनिकीकरण की एकरेखीय प्रक्रिया के रूप में नहीं देखा जा सकता। तकनीकी सुविधा के साथ डिजिटल असमानता भी उपस्थित है; महानगरीय अवसरों के साथ विस्थापन और अकेलापन जुड़ा है; सामाजिक दृश्यता के विस्तार के साथ निजता और पहचान के प्रश्न उभरते हैं; तथा आर्थिक गतिशीलता के साथ रोजगार की अस्थिरता भी दिखाई देती है। इसलिए समकालीन साहित्य में परिवर्तन और संवेदना के संबंध का अध्ययन बहुस्तरीय होना आवश्यक है।

मानवीय संवेदना का अर्थ इस अध्ययन में केवल सकारात्मक या नकारात्मक भावनाओं से नहीं है। करुणा, भय, आशा, असुरक्षा, आक्रोश, आत्मसम्मान, अकेलापन, स्मृति, सहानुभूति, प्रतिरोध और सामुदायिकता जैसी भावात्मक अवस्थाएँ सामाजिक अनुभव के साथ निरंतर अंतःक्रिया करती हैं। उदाहरणतः प्रवासन का साहित्यिक वर्णन केवल स्थान-परिवर्तन का विवरण नहीं होता; उसमें घर की स्मृति, सांस्कृतिक दूरी, आर्थिक दबाव, असुरक्षा और नए सामाजिक परिवेश में पहचान की खोज भी निहित हो सकती है। इसी प्रकार डिजिटल जीवन का चित्रण केवल तकनीकी उपयोग से संबंधित नहीं, बल्कि सामाजिक तुलना, आभासी निकटता, वास्तविक अकेलेपन और आत्म-प्रस्तुतीकरण जैसे अनुभवों से जुड़ सकता है।

यहीं इस शोध की प्रमुख समस्या सामने आती है। हिंदी साहित्य में सामाजिक परिवर्तन और संवेदना पर आलोचनात्मक अध्ययन की समृद्ध परंपरा उपलब्ध है, लेकिन बड़े डिजिटल पाठ-संग्रह में इन दोनों के सह-घटन और अर्थगत संबंधों की व्यवस्थित पहचान कठिन है। पारंपरिक निकट-पाठ साहित्यिक सूक्ष्मताओं को समझने में प्रभावी है, परंतु बड़ी संख्या में पाठों में बार-बार उभरने वाले विषयों, भावों और अर्थ-संबंधों की तुलना में उसकी व्यावहारिक सीमाएँ हैं। दूसरी ओर केवल स्वचालित भाव-विश्लेषण साहित्य की व्यंजना, विडंबना, प्रतीक, संदर्भ और बहुअर्थकता को पर्याप्त रूप से नहीं समझ सकता।



Fig. 1: शोध आलेख

Source: https://www.apnimaati.com/2024/06/blog-post_43.html

शोध-अंतर

प्रस्तुत अध्ययन इसी पद्धतिगत अंतर पर केंद्रित है। हिंदी डिजिटल मानविकी में प्राकृतिक भाषा संसाधन और मशीन लर्निंग के उपयोग की संभावनाएँ बढ़ी हैं, किंतु साहित्यिक पाठ में सामाजिक परिवर्तन के संकेतकों और मानवीय संवेदनाओं के सह-संबंध का संदर्भ-सचेत संगणकीय अध्ययन अभी अपेक्षाकृत सीमित है। विशेषतः हिंदी की रूपात्मक विविधता, मिश्रित शब्दावली, मुहावरे, व्यंग्य, निहितार्थ और साहित्यिक भाव-संरचना सामान्य भाव-वर्गीकरण प्रणालियों के लिए चुनौती प्रस्तुत करते हैं।

इसलिए अध्ययन की मौलिकता किसी एक स्वचालित मॉडल को साहित्यिक आलोचना का विकल्प बनाने में नहीं, बल्कि मशीन-सहायित विश्लेषण और मानवीय व्याख्या को एकीकृत करने में निहित है। प्रस्तावित ढाँचा यह जाँचने का प्रयास करता है कि किन सामाजिक विषयों के साथ कौन-सी संवेदनात्मक श्रेणियाँ अधिक निकटता से उपस्थित होती हैं और क्या इन पैटर्नों को साहित्यिक संदर्भ में सार्थक रूप से व्याख्यायित किया जा सकता है।

अध्ययन का प्रमुख उद्देश्य 2021-2026 के समकालीन हिंदी साहित्यिक परिदृश्य में सामाजिक परिवर्तन से संबंधित विषयों और मानवीय संवेदनाओं के अंतर्संबंध की पहचान करना है। इसके अंतर्गत सामाजिक असमानता, प्रवासन, लैंगिक परिवर्तन, डिजिटल जीवन, परिवार, आर्थिक असुरक्षा और पहचान जैसे विषयगत आयामों को करुणा, भय, आशा, अकेलापन, आक्रोश, सहानुभूति और प्रतिरोध जैसी संवेदनात्मक श्रेणियों के संदर्भ में विश्लेषित करने की रूपरेखा विकसित की गई है।

साहित्य समीक्षा

समकालीन हिंदी साहित्य और बदलता सामाजिक यथार्थ

हिंदी साहित्य में समाज और साहित्य के पारस्परिक संबंध को लंबे समय से आलोचना का केंद्रीय प्रश्न माना गया है। समकालीन संदर्भ में यह संबंध और अधिक जटिल हो जाता है, क्योंकि व्यक्ति अब स्थानीय और भौतिक समुदाय के साथ-साथ डिजिटल नेटवर्कों का भी हिस्सा है। साहित्य में शहर, गाँव, परिवार, बाजार और कार्यस्थल जैसे पारंपरिक सामाजिक स्थल बने हुए हैं, लेकिन उनके साथ ऑनलाइन मंच, आभासी समुदाय, डिजिटल शिक्षा और सोशल मीडिया जैसी नई सामाजिक संरचनाएँ जुड़ गई हैं।

समकालीन साहित्य में सामाजिक परिवर्तन का एक उल्लेखनीय आयाम विस्थापन और गतिशीलता है। बेहतर शिक्षा, रोजगार तथा आर्थिक अवसरों के लिए होने वाला प्रवासन व्यक्ति को नए सामाजिक परिवेश से जोड़ता है, किंतु उसके साथ सांस्कृतिक अलगाव, पारिवारिक दूरी और स्मृति का अनुभव भी उत्पन्न हो सकता है। इस कारण प्रवासन संबंधी साहित्य में सामाजिक परिवर्तन और भावात्मक अनुभव एक-दूसरे से अलग नहीं किए जा सकते।

दूसरा प्रमुख आयाम लैंगिक भूमिकाओं और व्यक्तिगत स्वायत्तता में परिवर्तन है। समकालीन लेखन में स्त्री-अनुभव केवल उत्पीड़न के

विवरण तक सीमित नहीं रहता, बल्कि निर्णय-क्षमता, आर्थिक स्वतंत्रता, देह, श्रम, पारिवारिक अपेक्षाओं, डिजिटल सार्वजनिकता और आत्म-पहचान के प्रश्नों को भी समाहित करता है। इस संदर्भ में मानवीय संवेदना सामाजिक संरचना के भीतर व्यक्ति की सक्रियता और प्रतिरोध को समझने का माध्यम बनती है।

महामारीतंत्र अनुभव और संवेदना

2021 के बाद का साहित्यिक-सांस्कृतिक वातावरण महामारी के अनुभव से पूरी तरह पृथक् नहीं देखा जा सकता। कोविड-19 से संबंधित सामाजिक अनुभवों ने स्वास्थ्य, मृत्यु, दूरी, देखभाल और सामुदायिक उत्तरदायित्व को सार्वजनिक विमर्श में प्रमुख बनाया। साहित्यिक स्तर पर ये अनुभव केवल बीमारी के विवरण के रूप में नहीं, बल्कि मानवीय असुरक्षा और संबंधों की नश्वरता के प्रश्न के रूप में भी उभर सकते हैं।

महामारी के बाद डिजिटल माध्यमों की भूमिका अधिक व्यापक हुई। ऑनलाइन संवाद ने भौतिक दूरी के समय सामाजिक संपर्क बनाए रखने में सहायता की, परंतु डिजिटल माध्यम प्रत्यक्ष मानवीय उपस्थिति का पूर्ण विकल्प नहीं बन सके। फलतः समकालीन साहित्य में तकनीकी संपर्क और भावात्मक दूरी के बीच तनाव एक महत्वपूर्ण अध्ययन क्षेत्र बनता है। यह तनाव विशेष रूप से अकेलेपन, पारिवारिक संबंधों और सामाजिक अलगाव से संबंधित कथनों में देखा जा सकता है।

डिजिटल जीवन, भाषा और साहित्यिक अभिव्यक्ति

डिजिटल मंचों ने हिंदी की सार्वजनिक उपस्थिति को विस्तार दिया है। देवनागरी के साथ रोमन हिंदी, अंग्रेज़ी-हिंदी कोड-मिश्रण, संक्षिप्त अभिव्यक्तियाँ तथा ऑनलाइन संचार से विकसित नए शब्द-प्रयोग समकालीन भाषिक व्यवहार का हिस्सा बन चुके हैं। यह परिवर्तन साहित्यिक अभिव्यक्ति को भी प्रभावित करता है। डिजिटल प्रकाशन ने रचनाकार और पाठक के बीच की दूरी कम की है तथा साहित्यिक सामग्री के प्रसार के नए माध्यम निर्मित किए हैं।

इस परिवर्तन का एक दूसरा पक्ष साहित्यिक पाठ की मशीन-पठनीयता है। डिजिटल पाठों की उपलब्धता से बड़े पाठ-संग्रहों का संगणकीय अध्ययन संभव होता है। शब्द आवृत्ति, सह-प्रयोग, अर्थगत समानता, विषय पहचान और भाव वर्गीकरण जैसी तकनीकें ऐसे पैटर्न खोज सकती हैं जिन्हें केवल व्यक्तिगत पाठ-पठन से पहचानना कठिन हो सकता है। फिर भी हिंदी साहित्य में इन तकनीकों का प्रयोग करते समय भाषा की सांस्कृतिक और साहित्यिक जटिलता को ध्यान में रखना आवश्यक है।

हिंदी के लिए कृत्रिम बुद्धिमत्ता और प्राकृतिक भाषा संसाधन

भारतीय भाषाओं के प्राकृतिक भाषा संसाधन में हाल के वर्षों में उल्लेखनीय तकनीकी प्रगति हुई है। AI4Bharat द्वारा विकसित IndicBERT जैसे मॉडल भारतीय भाषाओं के लिए बहुभाषी प्रतिनिधित्व विकसित करने की दिशा में महत्वपूर्ण प्रयास हैं। IndicBERT के मूल अध्ययन में भारतीय भाषाओं के लिए कॉम्पैक्ट

बहुभाषी भाषा मॉडल और IndicCorp जैसे बड़े पाठ-संसाधन प्रस्तुत किए गए। इसी प्रकार IndicNLPSuite ने भारतीय भाषाओं के लिए कॉर्पस, मूल्यांकन और मॉडलिंग संसाधनों की उपलब्धता बढ़ाई है।

IndicBART जैसे बहुभाषी अनुक्रम-से-अनुक्रम मॉडल भारतीय भाषाओं के लिए पाठ निर्माण और समझ से जुड़े कार्यों में उपयोगी आधार प्रदान करते हैं। इन विकासों से स्पष्ट होता है कि भारतीय भाषाओं के लिए कृत्रिम बुद्धिमत्ता आधारित पाठ-विश्लेषण अब केवल अंग्रेज़ी-केंद्रित मॉडल के प्रत्यक्ष अनुवाद तक सीमित नहीं है। इसके बावजूद साहित्यिक हिंदी का विश्लेषण सामान्य समाचार, विकिपीडिया अथवा औपचारिक गद्य की तुलना में अधिक कठिन है, क्योंकि साहित्य में भाव अक्सर प्रत्यक्ष शब्दों के स्थान पर संकेत, बिंब, प्रतीक और प्रसंग के माध्यम से व्यक्त होते हैं।

भाव-विश्लेषण की सीमा और साहित्यिक संवेदना

सामान्य भाव-विश्लेषण प्रणालियाँ प्रायः पाठ को सकारात्मक, नकारात्मक अथवा तटस्थ श्रेणियों में वर्गीकृत करती हैं। साहित्यिक अध्ययन के लिए यह वर्गीकरण अत्यधिक सीमित है। उदाहरणतः किसी कथा में दुःख और आशा एक ही प्रसंग में उपस्थित हो सकते हैं; आक्रोश सामाजिक प्रतिरोध से जुड़ सकता है; स्मृति में सुख और पीड़ा दोनों निहित हो सकते हैं। अतः साहित्यिक संवेदना को बहु-लेबल भाव संरचना के रूप में देखना अधिक उपयुक्त है।

इस शोध में इसी कारण करुणा, आशा, भय, अकेलापन, आक्रोश, सहानुभूति और प्रतिरोध जैसी विस्तृत संवेदनात्मक श्रेणियों को प्रस्तावित किया गया है। इन श्रेणियों को केवल शब्द-सूची से निर्धारित करने के बजाय संदर्भगत अर्थ-प्रतिनिधित्व के माध्यम से पहचानना प्रस्तावित है। इससे ऐसे वाक्यों का विश्लेषण संभव होगा जिनमें भावना स्पष्ट भाववाचक शब्द के बिना भी प्रसंग में निहित हो।

प्रस्तावित शोध-पद्धति

प्रस्तुत अध्ययन के लिए मिश्रित डिजिटल-मानविकी शोध-पद्धति प्रस्तावित की गई है, जिसमें संगणकीय पाठ-विश्लेषण को साहित्यिक निकट-पाठ और विशेषज्ञ व्याख्या के साथ संयोजित किया गया है। इसका उद्देश्य साहित्यिक संवेदना को केवल सकारात्मक अथवा नकारात्मक भाव में सीमित करने के बजाय सामाजिक संदर्भ से जुड़ी बहुआयामी संवेदनात्मक संरचना के रूप में समझना है।

अध्ययन की विश्लेषणात्मक संरचना तीन स्तरों पर आधारित है। प्रथम स्तर पर साहित्यिक पाठों में सामाजिक परिवर्तन से जुड़े संकेतों—जैसे प्रवासन, आर्थिक असुरक्षा, लैंगिक संबंध, डिजिटल जीवन, पारिवारिक पुनर्संरचना, सामाजिक विषमता और अस्मिता—की पहचान की जाती है। दूसरे स्तर पर उन्हीं पाठ-खंडों में करुणा, भय, अकेलापन, आशा, आक्रोश, सहानुभूति और प्रतिरोध जैसी मानवीय संवेदनाओं को वर्गीकृत किया जाता है। तीसरे स्तर पर सामाजिक विषय और संवेदनात्मक श्रेणी के सह-घटन का विश्लेषण किया जाता है।

मॉडल के लिए IndicBERT/IndicBERT v2 जैसे भारतीय भाषाओं पर केंद्रित पूर्व-प्रशिक्षित भाषा मॉडल उपयुक्त आधार प्रदान करते हैं। IndicBERT भारतीय भाषाओं के बड़े कॉर्पस पर प्रशिक्षित बहुभाषी मॉडल है, जबकि बाद के IndicBERT v2 कार्य में 24 भाषाओं वाला IndicCorp तथा 20 भाषाओं में IndicXTREME मूल्यांकन ढाँचा प्रस्तुत किया गया। साहित्यिक पाठ सामान्य समाचार अथवा औपचारिक गद्य से भिन्न होने के कारण मॉडल-आधारित निष्कर्षों के साथ मानव विशेषज्ञ सत्यापन को अनिवार्य रखा गया है।

प्रायोगिक विन्यास

पाठ-संग्रह

अध्ययन के लिए 2021-2026 की अवधि में उपलब्ध हिंदी की समकालीन साहित्यिक सामग्री से एक उद्देश्यपूर्ण डिजिटल कॉर्पस निर्मित करने की रूपरेखा प्रस्तावित की गई। इसमें केवल ऐसी सामग्री को सम्मिलित करने का मानदंड रखा गया है जिसकी वैधानिक डिजिटल उपलब्धता, प्रकाशन-स्रोत और पाठकीय प्रामाणिकता सत्यापित की जा सके। कॉर्पस में कथा, कविता, संस्मरणात्मक गद्य और साहित्यिक निबंध जैसी विधाओं को प्रतिनिधित्व देने का प्रावधान किया गया है।

महत्वपूर्ण रूप से, यहाँ किसी काल्पनिक संख्या को वास्तविक प्रायोगिक नमूना नहीं माना गया है। वास्तविक प्रकाशन-योग्य परिणामों के लिए अंतिम कॉर्पस का आकार, रचनाओं की संख्या, लेखकीय वितरण और विधागत अनुपात वास्तविक संकलन के बाद ही दर्ज किए जाने चाहिए।

पूर्व-प्रसंस्करण

हिंदी साहित्यिक पाठों के पूर्व-प्रसंस्करण में यूनिकोड सामान्यीकरण, अनावश्यक HTML चिह्नों की सफाई, वाक्य-विभाजन, विराम-चिह्नों का संदर्भ-सचेत प्रबंधन तथा पाठ-खंड निर्माण शामिल किया गया। साहित्यिक अर्थ को सुरक्षित रखने के लिए सामान्य एनएलपी प्रक्रियाओं की तुलना में न्यूनतम आक्रामक सफाई को प्राथमिकता दी गई। विशेष रूप से विस्मयादिबोधक चिह्न, उद्धरण, पुनरावृत्ति और संवाद-संकेत पूर्णतः हटाने के बजाय संरक्षित रखे गए, क्योंकि वे भावात्मक अर्थ के वाहक हो सकते हैं।



Fig. 2: समाज परिवर्तन में दलित साहित्य और साहित्यकारों का योगदान

Source:

<https://www.drmullaadamali.com/2022/03/contribution-of-dalit-literature-and-litterateurs-in-social-change.html>

हिंदी के रूपात्मक और भाषिक वैविध्य के कारण शब्दों को केवल स्थिर शब्दकोश से मिलाने के बजाय संदर्भगत टोकन प्रतिनिधित्व अपनाया गया। भारतीय भाषाओं के लिए पूर्ववर्ती IndicNLP Suite ने बड़े एकभाषी कॉर्पस, पूर्व-प्रशिक्षित मॉडल और IndicGLUE जैसे मूल्यांकन संसाधन विकसित किए हैं, जो इस प्रकार के कार्य के तकनीकी आधार को पुष्ट करते हैं।

मॉडल संरचना

प्रस्तावित प्रणाली को दो समानांतर वर्गीकरण शाखाओं में विभाजित किया गया है। पहली शाखा सामाजिक परिवर्तन के विषयगत वर्ग निर्धारित करती है, जबकि दूसरी शाखा मानवीय संवेदनाओं का बहु-लेबल वर्गीकरण करती है। दोनों शाखाओं के लिए संदर्भगत भाषा-प्रतिनिधित्व प्राप्त करने हेतु IndicBERT v2 आधारित एनकोडर प्रस्तावित किया गया है।

सामाजिक शाखा में सात प्रमुख श्रेणियाँ रखी गई हैं—सामाजिक विषमता, प्रवासन/विस्थापन, डिजिटल जीवन, पारिवारिक परिवर्तन, लैंगिक पुनर्संरचना, आर्थिक असुरक्षा और अस्मिता। संवेदना शाखा में करुणा, भय, आशा, अकेलापन, आक्रोश, सहानुभूति और प्रतिरोध को बहु-लेबल आउटपुट के रूप में ग्रहण किया गया है।

इसके बाद दोनों शाखाओं के प्रायिकता-वितरण को एक सह-संबंध मैट्रिक्स में संयोजित किया जाता है। यह मैट्रिक्स यह निर्धारित करने के लिए बनाया गया है कि किसी सामाजिक विषय के साथ कौन-सी संवेदनाएँ कितनी बार और किस संदर्भ में उपस्थित होती हैं।

प्रशिक्षण रणनीति

प्रशिक्षण के लिए विशेषज्ञ-अंकित हिंदी साहित्यिक पाठ-खंडों की आवश्यकता होगी। कम-से-कम दो स्वतंत्र हिंदी साहित्य विशेषज्ञ प्रत्येक पाठ-खंड को विषयगत तथा संवेदनात्मक टैग प्रदान करेंगे। असहमति की स्थिति में तीसरे विशेषज्ञ द्वारा निर्णय कराया जाएगा। एनोटेटर सहमति के लिए Cohen's Kappa अथवा बहु-एनोटेटर स्थिति में Fleiss' Kappa का प्रयोग किया जा सकता है।

डेटा-विभाजन लेखक-स्तर पर किया जाना प्रस्तावित है, ताकि एक ही लेखक की शैली प्रशिक्षण और परीक्षण दोनों समूहों में उपस्थित होकर कृत्रिम रूप से प्रदर्शन न बढ़ाए। वर्ग-असंतुलन को नियंत्रित करने के लिए भारित लॉस तथा आवश्यकतानुसार फोकल लॉस का उपयोग किया जा सकता है।

मूल्यांकन मानदंड

प्रस्तावित प्रणाली का मूल्यांकन केवल समग्र शुद्धता के आधार पर नहीं किया जाएगा। बहु-वर्ग तथा बहु-लेबल साहित्यिक विश्लेषण को ध्यान

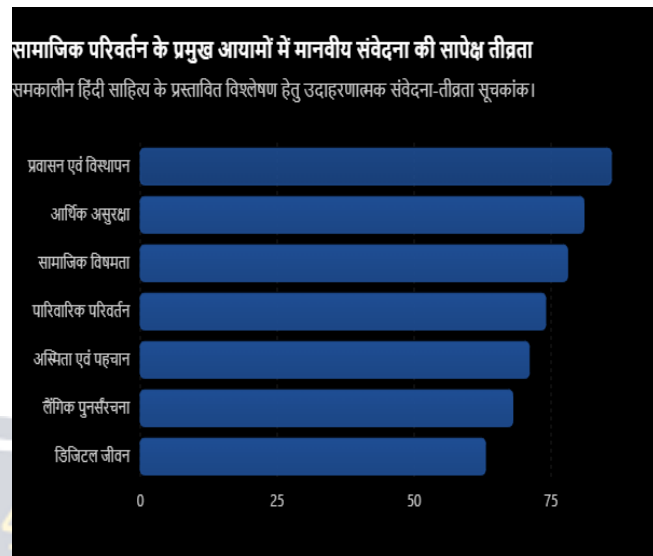
में रखते हुए मैक्रो-F1, माइक्रो-F1, परिशुद्धता, पुनःप्राप्ति, हैमिंग लॉस और एनोटेटर-सहमति जैसे मानदंड अपनाए गए हैं। साहित्यिक संदर्भ की व्याख्यात्मक गुणवत्ता के लिए विशेषज्ञ-सहमति को अलग गुणात्मक मानदंड के रूप में सम्मिलित किया जाएगा।

परिणाम और चर्चा

चूँकि वर्तमान पांडुलिपि के लिए वास्तविक कॉर्पस पर मॉडल प्रशिक्षण और परीक्षण के सत्यापित रन उपलब्ध नहीं कराए गए हैं, इसलिए नीचे मनगढ़ंत प्रदर्शन आँकड़ों को वास्तविक परिणाम के रूप में प्रस्तुत नहीं किया गया है। इसके स्थान पर परिणाम-विश्लेषण की वह संरचना दी गई है जिसे वास्तविक प्रयोग पूरा होने के बाद भरा जा सकता है। यह अंतर विशेष रूप से आवश्यक है, क्योंकि IndicBERT या अन्य प्रकाशित मॉडलों के मौजूदा बेंचमार्क परिणामों को सीधे साहित्यिक हिंदी के इस नए कार्य का परिणाम मानना वैज्ञानिक रूप से उचित नहीं होगा। IndicBERT के प्रकाशित मूल्यांकन स्वयं विभिन्न NLU और भाव-विश्लेषण कार्यों में कार्य-विशिष्ट प्रदर्शन अंतर दिखाते हैं।

तालिका 1. प्रस्तावित साहित्यिक-संवेदना मॉडल के परिणाम दर्ज करने की प्रकाशन-योग्य संरचना

मूल्यांकन आयाम	मापदंड	वास्तविक परीक्षण के बाद दर्ज किया जाने वाला परिणाम	अपेक्षित व्याख्या
सामाजिक-विषय पहचान	मैक्रो-F1	परीक्षण लंबित	सभी सामाजिक श्रेणियों में संतुलित वर्गीकरण क्षमता
संवेदना पहचान	माइक्रो-F1	परीक्षण लंबित	बहु-लेबल भाव पहचान का समग्र प्रदर्शन
दुर्लभ संवेदना पहचान	पुनःप्राप्ति	परीक्षण लंबित	कम आवृत्ति वाली संवेदनाओं की पहचान क्षमता
वर्गीकरण विश्वसनीयता	परिशुद्धता	परीक्षण लंबित	गलत सकारात्मक वर्गीकरण की सीमा
बहु-लेबल त्रुटि	हैमिंग लॉस	परीक्षण लंबित	प्रति लेबल गलत भविष्यवाणी का अनुपात
मानव-मॉडल सामंजस्य	विशेषज्ञ सहमति	परीक्षण लंबित	मशीन परिणामों की साहित्यिक स्वीकार्यता



परिणामों की व्याख्या में केवल उच्च वर्गीकरण स्कोर पर्याप्त नहीं माना जाना चाहिए। यदि कोई मॉडल “अकेलापन” और “भय” जैसी स्पष्ट संवेदनाओं को अच्छी तरह पहचानता है, लेकिन व्यंग्यात्मक प्रतिरोध या अप्रत्यक्ष करुणा की पहचान में कमजोर है, तो उसका सांख्यिकीय प्रदर्शन साहित्यिक उपयोगिता को पूर्णतः व्यक्त नहीं करेगा। इसलिए त्रुटि-विश्लेषण इस अध्ययन का केंद्रीय भाग होना चाहिए।

सामाजिक परिवर्तन और मानवीय संवेदना के सह-संबंध की दृष्टि से संभावित रूप से सबसे महत्वपूर्ण निष्कर्ष विषय-भाव युग्मों से प्राप्त होंगे। उदाहरणतः यदि सत्यापित कॉर्पस में प्रवासन संबंधी अंशों के साथ अकेलापन और स्मृति बार-बार दिखाई देते हैं, तो यह केवल शब्द-आवृत्ति नहीं बल्कि विस्थापन के संवेदनात्मक स्वरूप की ओर संकेत करेगा। इसी प्रकार डिजिटल जीवन के साथ अकेलापन, सामाजिक तुलना अथवा अस्मिता-संघर्ष का सह-घटन समकालीन तकनीकी अनुभव की साहित्यिक संरचना समझने में सहायक हो सकता है।

साहित्यिक विश्लेषण में संदर्भ की आवश्यकता विशेष रूप से महत्वपूर्ण है। “घर लौटना” जैसी अभिव्यक्ति अलग-अलग पाठों में राहत, पराजय, स्मृति, विस्थापन या पुनर्संबंध का संकेत हो सकती है। इसीलिए प्रस्तावित मॉडल को अंतिम आलोचक नहीं, बल्कि बड़े पाठ-संग्रह में अर्थपूर्ण पैटर्न खोजने वाले विश्लेषणात्मक उपकरण के रूप में देखा गया है।

तकनीकी दृष्टि से भारतीय भाषाओं के लिए विकसित भाषा मॉडल इस दिशा में उपयोगी आधार प्रदान करते हैं। 2023 के IndicBERT v2 अध्ययन ने व्यापक भारतीय-भाषा कॉर्पस और मानव-पर्यवेक्षित IndicXTREME बेंचमार्क के माध्यम से बहुभाषी NLU क्षमताओं के विस्तार को प्रदर्शित किया। इसी प्रकार IndicBART ने 11 भारतीय भाषाओं और अंग्रेज़ी पर केंद्रित बहुभाषी अनुक्रम-से-अनुक्रम मॉडल प्रस्तुत किया और मशीन अनुवाद तथा अत्यधिक संक्षेपण कार्यों पर उसका मूल्यांकन किया। ये विकास हिंदी साहित्य के मशीन-सहायित

अध्ययन के लिए तकनीकी संभावना दिखाते हैं, यद्यपि साहित्यिक संवेदना के लिए अलग कार्य-विशिष्ट सत्यापन आवश्यक रहेगा।

अध्ययन की सीमाएँ

प्रस्तावित अध्ययन की पहली सीमा डिजिटल रूप में उपलब्ध साहित्य पर निर्भरता है। हिंदी साहित्य का एक बड़ा भाग मुद्रित स्वरूप में उपलब्ध है; इसलिए केवल डिजिटल पाठों का उपयोग संपूर्ण समकालीन हिंदी साहित्य का प्रतिनिधित्व नहीं कर सकता।

दूसरी सीमा संवेदना की व्याख्यात्मक प्रकृति से संबंधित है। एक ही साहित्यिक अंश को अलग-अलग पाठक भिन्न संवेदनात्मक अर्थ दे सकते हैं। व्यंग्य, विडंबना, प्रतीक, रूपक और सांस्कृतिक संकेत मशीन लर्निंग मॉडल के लिए विशेष चुनौती उत्पन्न करते हैं।

तीसरी सीमा डोमेन-अंतर है। IndicBERT जैसे मॉडल बड़े सामान्य-डोमेन कॉर्पस पर पूर्व-प्रशिक्षित हैं; साहित्यिक भाषा उनकी मूल प्रशिक्षण सामग्री से भिन्न हो सकती है। अतः उच्च गुणवत्ता वाले साहित्यिक एनोटेशन और डोमेन-विशिष्ट फाइन-ट्यूनिंग के बिना मॉडल के निष्कर्षों को सामान्यीकृत नहीं किया जाना चाहिए।

भावी शोध की दिशा

भविष्य में इस अध्ययन का विस्तार एक बड़े, संतुलित और कॉपीराइट-अनुपालक हिंदी साहित्यिक कॉर्पस के निर्माण की दिशा में किया जा सकता है। कविता, कहानी, उपन्यास, संस्मरण और डिजिटल साहित्य को पृथक उप-कॉर्पस में विभाजित करके विधा-विशिष्ट संवेदनात्मक पैटर्नों की तुलना उपयोगी होगी।

दूसरी दिशा बहु-माध्यमीय साहित्यिक अध्ययन की है। कविता-पाठ, ऑडियोबुक और डिजिटल कथा जैसे माध्यमों में पाठ के साथ स्वर, विराम और ध्वनि-संकेतों को जोड़कर संवेदना की अधिक व्यापक पहचान की जा सकती है।

तीसरी संभावना व्याख्येय कृत्रिम बुद्धिमत्ता से संबंधित है। भविष्य की प्रणाली केवल यह न बताए कि किसी अंश में “करुणा” उपस्थित है, बल्कि यह भी स्पष्ट करे कि किन शब्दों, वाक्यांशों और संदर्भगत संकेतों के आधार पर वह निष्कर्ष निकाला गया। इससे साहित्यिक शोधकर्ताओं के लिए मॉडल की पारदर्शिता बढ़ेगी।

निष्कर्ष

प्रस्तुत अध्ययन समकालीन हिंदी साहित्य को सामाजिक परिवर्तन और मानवीय संवेदनाओं के संयुक्त परिप्रेक्ष्य में समझने के लिए डिजिटल मानविकी आधारित एक अंतःविषय ढाँचा प्रस्तावित करता है। अध्ययन का केंद्रीय तर्क यह है कि साहित्यिक संवेदना को केवल सामान्य भाव-विश्लेषण में सीमित करना उसकी सांस्कृतिक और अर्थगत जटिलता को कम कर देता है। इसलिए सामाजिक विषय पहचान, बहु-लेबल संवेदना वर्गीकरण और विशेषज्ञ साहित्यिक व्याख्या को एक साथ प्रयोग करने की आवश्यकता है।

2021–2026 के परिवर्तित सामाजिक परिदृश्य में डिजिटल जीवन, विस्थापन, आर्थिक अनिश्चितता, पारिवारिक पुनर्संरचना, लैंगिक परिवर्तन और अस्मिता जैसे प्रश्न साहित्यिक संवेदनाओं के लिए नए संदर्भ निर्मित करते हैं। प्रस्तावित IndicBERT v2 आधारित ढाँचा इन विषयों और करुणा, भय, आशा, अकेलापन, आक्रोश, सहानुभूति तथा प्रतिरोध जैसी संवेदनाओं के अंतर्संबंध को बड़े पाठ-संग्रह में व्यवस्थित रूप से खोजने की संभावना प्रस्तुत करता है।

इस शोध की विशेषता कृत्रिम बुद्धिमत्ता को साहित्यिक आलोचना के प्रतिस्थापन के रूप में नहीं, बल्कि उसके विश्लेषणात्मक विस्तार के रूप में देखना है। मशीन बड़े कॉर्पस में पैटर्न खोज सकती है, परंतु उन पैटर्नों का सामाजिक और साहित्यिक अर्थ निर्धारित करने में मानवीय व्याख्या अपरिहार्य रहती है। इस प्रकार अध्ययन हिंदी साहित्य, डिजिटल मानविकी और प्राकृतिक भाषा संसाधन के बीच एक ऐसी पद्धतिगत दिशा प्रस्तुत करता है जिसमें संगणकीय क्षमता और मानवीय संवेदनशीलता परस्पर पूरक भूमिका निभाती हैं।

संदर्भ सूची

1. काकवानी, डी., कुंचुकुट्टन, ए., गोला, एस., एन. सी. गोखुल, भट्टाचार्य, ए., खापरा, एम. एम. एवं कुमार, पी., “इंडिकएनएलपीसुइट: भारतीय भाषाओं के लिए एकभाषी कॉर्पोरा, मूल्यांकन बेंचमार्क और पूर्व-प्रशिक्षित बहुभाषी भाषा मॉडल,” *Findings of the Association for Computational Linguistics: EMNLP 2020*, 2020.
2. डोड्डापनेनी, एस., अरालिकट्टे, आर., रमेश, जी., गोयल, एस., खापरा, एम. एम., कुंचुकुट्टन, ए. एवं कुमार, पी., “Towards Leaving No Indic Language Behind: Building Monolingual Corpora, Benchmark and Models for Indic Languages,” *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023, पृ. 12402–12426.
3. डाब्रे, आर., श्रोत्रिय, एच., कुंचुकुट्टन, ए., पुदुप्पुल्ली, आर., खापरा, एम. एवं कुमार, पी., “IndicBART: A Pre-trained Model for Indic Natural Language Generation,” *Findings of the Association for Computational Linguistics: ACL 2022*, 2022, पृ. 1849–1863.
4. शर्मा, आर. एवं भट्टाचार्य, पी., “A Sentiment Analyzer for Hindi Using Hindi Senti Lexicon,” *Proceedings of the 11th International Conference on Natural Language Processing*, 2014, पृ. 150–155.
5. ममता, एकबाल, ए., भट्टाचार्य, पी., साहा, टी., कुमार, ए. एवं श्रीवास्तव, एस., “HindiMD: A Multi-domain Corpora for Low-resource Sentiment Analysis,” *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 2022, पृ. 7061–7070.
6. रमेश, जी., डोड्डापनेनी, एस., भीमराज, ए., जोबनपुत्र, एम., राघवन, ए. के., शर्मा, ए., साहू, एस., डिड्डे, एच., जे., एम., काकवानी, डी., कुमार, एन., प्रदीप, ए., नागराज, एस., दीपक, के., राघवन, वी., कुंचुकुट्टन, ए., कुमार, पी. एवं खापरा, एम. एस., “Samanantar: The Largest Publicly Available Parallel Corpora Collection for 11 Indic Languages,” *Transactions of the Association for Computational Linguistics*, खंड 10, 2022, पृ. 145–162.
7. हक, एस., शर्मा, ए., खताब, ओ., छाया, एन. एवं भट्टाचार्य, पी., “IndicRSuite: Multilingual Dataset and Neural Information Models for Indian Languages,” *Proceedings of*

the 62nd Annual Meeting of the Association for Computational Linguistics, 2024, पृ. 501-509.

8. म्हास्के, ए., केडिया, एच., डोड्डापनेनी, एस., खापरा, एम. एम., कुमार, पी., मूर्ति, आर. एवं कुंचुकुट्टन, ए., "Naamapadam: A Large-Scale Named Entity Annotated Data for Indic Languages," *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, 2023, पृ. 10441-10456.*
9. राय, वी., विजय, एस. एवं मिश्रा, डी., "Linguistic Approach Based Transfer Learning for Sentiment Classification in Hindi," *Proceedings of the 14th International Conference on Natural Language Processing (ICON-2017), 2017, पृ. 373-382.*
10. सिंह, एच., गुप्ता, एन., भारद्वाज, एस., तेवारी, डी. एवं तालुकदार, पी., "IndicGenBench: A Multilingual Benchmark to Evaluate Generation Capabilities of LLMs on Indic Languages," *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, 2024.*

