कृषि में मशीन लर्निंग की भूमिका

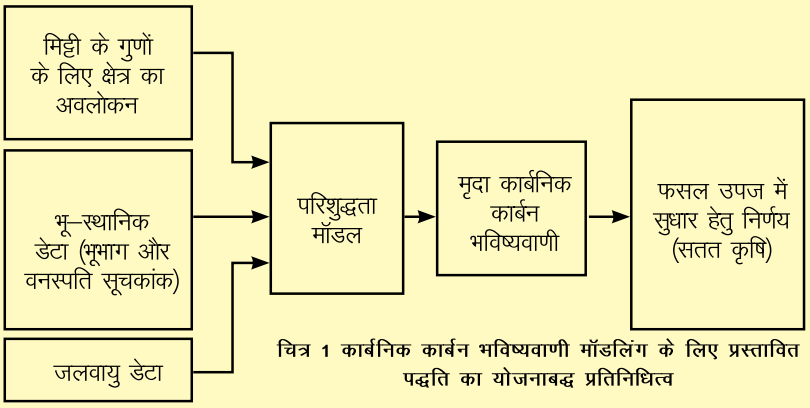
डॉ. पूजा जैन
एसिस्टेंट प्रोफेसर, आई आई टी नागपुर

कृषि को वैश्विक अर्थव्यवस्था का प्रमुख स्तंभ माना जाता है। स्थायी संतुलन स्थापित करने के लिए कृषि उत्पादन को बढ़ाना होगा। फसल की स्थायी रूप से उत्पादन करने की क्षमता ज्यादातर मिट्टी, पानी, जलवायु और परिवेश सहित कई कारकों पर निर्भर होती है। कृत्रिम बुद्धिमत्ता फसल की निगरानी करने और स्वचालित परिणाम में उपज की भविष्यवाणी करने में सहायता करती है। आज की यह अत्याधुनिक तकनीक कृषि प्रणाली को अनुकूलित करके सटीक कृषि में अपनी महत्वपूर्ण भूमिका निभा रही है।

कृषि में एमएल के कई अनुप्रयोग हैं। हाल के अध्ययन के अनुसार, तीन मानक श्रेणियों की पहचान की गई जो फसल, जल और मिट्टी प्रबंधन को संदर्भित करती हैं। जहाँ तक फसल प्रबंधन का सवाल है, इसमें उपज का पूर्वानुमान, रोग और खरपतवार का पता, फसल की पहचान और फसल की गुणवत्ता का पता लगाना सम्मिलित है।

बीमारियाँ पैदावार को नष्ट कर देती हैं, इसलिए, एमएल तकनीकों का उपयोग करके किसान को समस्याओं का पता लगाना, सचेत करना और निदान करना एक बड़ी सहायता है। इमेजिंग प्रसंस्करण प्रौद्योगिकियों के साथ मिलकर मशीन लर्निंग एल्गोरिदम लक्ष्य खरपतवारों की वास्तविक समय की पहचान और मानचित्रण की अनुमति देते हैं। इससे खेतों में अत्यधिक छिड़काव के बजाय विशिष्ट क्षेत्रों में शाकनाशी लगाने में सहायता मिलती है। दोषों की पहचान की जा सकती है और उपज को वजन, रंग, आकार और फसल की परिपक्वता के आधार पर क्रमबद्ध किया जा सकता है।

सतत कृषि कार्बनिक कार्बन (ओसी) के आस–पास घूमती है, जो कई मिट्टी के कार्यों और पारिस्थितिक–तार्किक विशेषताओं के लिए

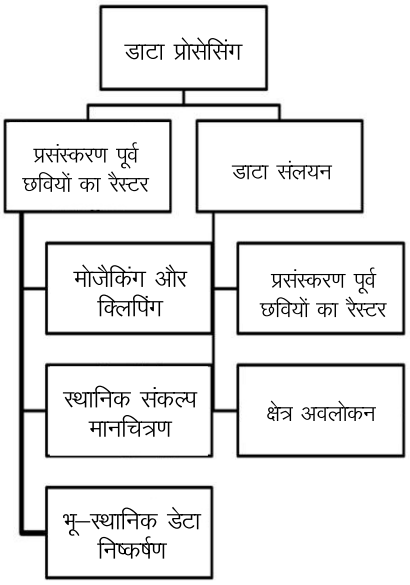


चित्र 1 कार्बनिक कार्बन भविष्यवाणी मॉडलिंग के लिए प्रस्तावित पद्धति का योजनाबद्ध प्रतिनिधित्व

https://doi.org/10.52151/aet2024481.1714

आवश्यक है। उपज और मिट्टी के स्वास्थ्य को बढ़ाने के लिए, विशिष्ट क्षेत्रों में किसान मिट्टी की कार्बनिक कार्बन को बनाए रखने और बढ़ाने में रुचि रखते हैं। मेरा शोध सतत कृषि के लिए मशीन लर्निंग तकनीकों और भू–स्थानिक डेटा का उपयोग करके मिट्टी में कार्बनिक कार्बन की भविष्यवाणी करता है। 30 मीटर रिजॉल्यूशन पर मिट्टी के आर्गेनिक कार्बन का पूर्वानुमान लगाने के लिए, चार मशीन लर्निंग मॉडल– रैंडम फॉरेस्ट, सपोर्ट वेक्टर मशीन, एडाप्टिव बूस्टिंग और के–नियरेस्ट नेबर का उपयोग किया गया, जिसने काफी अच्छे परिणाम दिए।

शोध का उद्देश्य मिट्टी में मौजूद कार्बनिक कार्बन (ओसी) के पूर्वानुमान के लिए मिट्टी की स्थलाकृति और गुणों के आधार पर एक मॉडल बनाना था, जो किसानों को फसल की उपज में उचित सुधार करने में सहायता करेगा। कई पर्यावरणीय और जैविक कारक, साथ ही उनकी परस्पर क्रियाएं, मिट्टी में पोषक तत्वों की मात्रा को नियंत्रित करती हैं। स्थलाकृति, जलवायु और रिमोट सेंसिंग का प्रतिनिधित्व करने वाले सहसंयोजकों के एक समूह को मिट्टी के कार्बनिक कार्बन का पूर्वानुमान लगाने के लिए संभावित भविष्यवक्ता कारकों के रूप में चुना जाएगा। समर्थन आकार में परिवर्तन करके, इस डेटा को विभिन्न स्थानिक, वर्णक्रमीय और रेडियोमेट्रिक रिजॉल्यूशन के साथ एकीकृत किया जाना चाहिए। जैसा कि चित्र 1 में दिखाया गया है, प्रस्तावित पद्धति में मुख्य रूप से ओसी का अनुमान निर्धारित करने के लिए डेटा प्रोसेसिंग और मॉडल निर्माण चरण सम्मिलित हैं जिनका उपयोग टिकाऊ कृषि के लिए उपयुक्त निर्णय लेने के लिए किया जा सकता है।



चित्र 2 डेटा प्रोसेसिंग के चरण

मशीन लर्निंग मॉडलिंग के लिए मौलिक चरणों में इनपुट डेटा एकत्र करने के लिये की आवश्यकता होती है। जैसा कि चित्र 2 में दर्शाया गया है। प्री–प्रोसेसिंग में वे क्रियाएं सम्मिलित होती हैं जो प्राथमिक डेटा विश्लेषण और मिट्टी की जानकारी निष्कर्षण से पहले प्रायः आवश्यक होती हैं। डेटा एक्सट्रैक्शन में मुख्य रूप से दो चरण होते हैं, एक है रैस्टर स्टैकिंग जहाँ रैस्टर को भविष्यवक्ता मानों को पुनः प्राप्त करने के लिए संयोजित किया जाता है और दूसरा है सैंपलिंग रैस्टर मान जिसमें रैस्टर मानों को स्टैक्ड रैस्टर से निकाला जाता है, और मॉडलिंग के लिए उपयुक्त प्रारूप में अस्थायी फाइल में संग्रहीत किया जाता है।

विभिन्न स्रोतों से डेटा इकट्ठा करने के बाद, उस डेटा को मॉडल इनपुट के रूप में उपयोग करने के लिए दोबारा सैंपल लिया गया। डेटा सैंपलिंग और मॉडल विश्लेषण के लिए, हमने विभिन्न पायथन पैकेजों का उपयोग किया है।

भारत के राज्य छत्तीसगढ़ के धमतरी जिले में ओसी के पूर्वानुमान के लिए चार मॉडलों के प्रदर्शन का मूल्यांकन किया गया और रैंडम फॉरेस्ट एल्गोरिदम ने सर्वोत्तम परिणाम दिखाए, जैसा कि क्रॉस वैलिडेशन (सीवी) के साथ और उसके बिना तालिका 1 में दर्शाया गया है।

ट्रेन–टेस्ट स्प्लिट पद्धति का उपयोग करके ऑर्गेनिक कार्बन के पूर्वानुमान में मॉडलों के प्रदर्शन को मान्य किया गया था। डेटासेट को दो हिस्सों में विभाजित किया गया, परीक्षण के लिए 20 प्रतिशत और प्रशिक्षण के लिए 80 प्रतिशत। यह सुनिश्चित करने के लिए कि मॉडल स्थिर था, डेटासेट को प्रशिक्षित करने के लिए 10 गुना के साथ क्रॉस सत्यापन संचालित किया गया था। मॉडल का मूल्यांकन करने के लिए, दो त्रुटि मेट्रिक्स का उपयोग किया गया है, अर्थात् R^2 और RMSE एक सांख्यिकीय माप है जिसका उपयोग प्रतिगमन मॉडल की फिट की गुणवत्ता का आँकलन करने के लिए किया जाता है। निर्धारण के गुणांक, R^2 की व्याख्या आश्रित चर में विचरण के अनुपात के रूप में की जाती है जो स्वतंत्र चर से अनुमानित है। आरएमएसई एक सांख्यिकीय माप है जो त्रुटि के विचलन का अनुमान लगाता है।

अध्ययन के निष्कर्ष दर्शाते हैं कि मिट्टी से संबंधित सटीक डेटा तैयार करने में डिजिटल मृदा–मानचित्रण विधियाँ कितनी प्रभावी हैं। किसान अन्य सूचना और संचार प्रौद्योगिकी (आईसीटी) प्रौद्योगिकियों के अलावा डिजिटल प्लेटफॉर्म और मोबाइल एप्लिकेशन के माध्यम से ज्ञान प्राप्त कर सकते हैं।

एमएल मॉडल	सी.वी. के साथ		सी.वी. के बिना	
	आर^2	आरएमएसई	आर^2	आरएमएसई
आरएफ	0.1526	0.2175	0.1281	0.2207
एसवीएम	0.0031	0.2359	0.0030	0.2360
केएनएन	0.2141	0.2604	0.2145	0.2604
अडाबूस्ट	0.0633	0.2287	0.0724	0.2447

तालिका 1 : परीक्षण डेटासेट के लिए मॉडल प्रदर्शन