सूचना सिद्धांत में विचलन माप का विश्लेषण

पारुल खंडेलवाल, संध्या पारीक
गणित विभाग,
एपेक्स विश्वविद्यालय

सारांश (Abstract)

सूचना सिद्धांत (Information Theory) सूचना के परिमाणीकरण, संचरण और संपीड़न को गणितीय रूप देता है। इस क्षेत्र का एक प्रमुख विचार विचलन माप (Divergence Measure) है, जो दो प्रायिकता वितरणों के बीच असमानता को मापता है। यह शोध-पत्र प्रमुख विचलन मापों—**कुलबैक–लाइबलर विचलन (Kullback–Leibler Divergence), जेन्सेन–शैनन विचलन (Jensen–Shannon Divergence), हेलिंजर दूरी (Hellinger Distance) तथा टोटल वेरिएशन दूरी (Total Variation Distance)**—का गणितीय विश्लेषण प्रस्तुत करता है। इनके अनुप्रयोग मशीन लर्निंग, प्राकृतिक भाषा प्रसंस्करण, डेटा संपीड़न और जैव-सूचना विज्ञान में रेखांकित किए गए हैं। साथ ही इनके लाभ, सीमाएँ और भविष्य की संभावनाएँ भी दर्शाई गई हैं।

कीवर्ड्स (Keywords) - सूचना सिद्धांत, KL Divergence, JS Divergence, हेलिंजर दूरी, टोटल वेरिएशन, मशीन लर्निंग, NLP, बायोइन्फॉर्मेटिक्स

1. परिचय (Introduction)

सूचना सिद्धांत का प्रस्ताव सबसे पहले क्लॉड शैनन ने 1948 में किया था, जहाँ उन्होंने सूचना को एंट्रॉपी (Entropy) द्वारा मापने की संकल्पना दी (Shannon, 1948)। सूचना सिद्धांत के विस्तार के साथ, यह स्पष्ट हुआ कि केवल एंट्रॉपी ही पर्याप्त नहीं है; विभिन्न वितरणों की तुलना करना भी आवश्यक है। इसी उद्देश्य से विचलन माप विकसित किए गए।

Kullback और Leibler (1951) ने KL Divergence प्रस्तुत किया, जिसने दो वितरणों के बीच असमानता को मापने का गणितीय आधार प्रदान किया। बाद में Lin (1991) ने Jensen–Shannon

Divergence विकसित किया, जो KL का सममित और सीमित (bounded) संस्करण है। इसके अलावा Hellinger Distance और Total Variation Distance जैसे माप भी व्यापक रूप से उपयोग किए जाते हैं।

2. अनुप्रयोग (Applications)

2.1 मशीन लर्निंग

KL Divergence का उपयोग Variational Autoencoders में होता है (Goodfellow et al., 2014)। JS Divergence Generative Adversarial Networks में प्रयुक्त होता है।

2.2 संचार एवं डेटा संपीड़न

KL Divergence सूचना हानि मापने के लिए उपयोगी है (Cover & Thomas, 2006)।

2.3 प्राकृतिक भाषा प्रसंस्करण (NLP)

JS Divergence डॉक्यूमेंट क्लस्टरिंग और भाषा मॉडलिंग में प्रभावी है।

2.4 जैव-सूचना विज्ञान (Bioinformatics)

Hellinger Distance DNA अनुक्रमों और जीनोमिक डेटा तुलना में प्रयुक्त होती है।

3. केस स्टडी (Case Studies)

· भाषा मॉडल तुलना: KL Divergence से यह पता लगाया जा सकता है कि एक अनुमानित भाषा मॉडल वास्तविक डेटा से कितना मेल खाता है।

- DNA अनुक्रम तुलना: Hellinger Distance का प्रयोग दो जैविक अनुक्रमों के बीच समानता मापने के लिए किया जाता है।
- डेटा संपीड़न: KL Divergence यह बताता है कि किसी संपीड़न एल्गोरिद्म से कितनी सूचना हानि हुई।

4. चुनौतियाँ (Challenges)

· KL Divergence Q(x)=0 होने पर परिभाषित नहीं है।

· उच्च-आयामी डेटा (High-dimensional data) में गणना जटिल हो जाती है।

· वास्तविक दुनिया के डेटा शोरयुक्त (noisy) होते हैं, जिससे Divergence मान अस्थिर हो सकता है।

5. भविष्य की संभावनाएँ (Future Directions)

- Quantum Information Theory में नए Divergence Measures का विकास।
- Explainable AI में Divergence के आधार पर मॉडल पारदर्शिता बढ़ाना।
- Big Data Analytics में स्केलेबल approximation तकनीकों का विकास।
- चिकित्सा विज्ञान में Genomic Divergence Measures का प्रयोग।

6. निष्कर्ष (Conclusion)

विचलन माप सूचना सिद्धांत का एक महत्वपूर्ण हिस्सा हैं। KL Divergence सूचना हानि का माप है, JS Divergence आधुनिक मशीन लर्निंग में अत्यंत उपयोगी है, Hellinger Distance जैव-सूचना विज्ञान में प्रभावी है और Total Variation Distance सांख्यिकी में सरल और लोकप्रिय है। जैसे-जैसे डेटा जटिल और विशाल होता जाएगा, Divergence Measures का महत्व और भी बढ़ेगा।

References (APA 7th Edition)

1. Cover, T. M., & Thomas, J. A. (2006). Elements of information theory. Wiley.

2. Csiszár, I., & Shields, P. C. (2004). Information theory and statistics: A tutorial. Now Publishers.

3. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. Advances in Neural Information Processing Systems, 27.

4. Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. Annals of Mathematical Statistics, 22(1), 79–86.

5. Lin, J. (1991). Divergence measures based on the Shannon entropy. IEEE Transactions on Information Theory, 37(1), 145–151.

6. Nielsen, F. (2020). An elementary introduction to information geometry. Springer.

7. Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27(3), 379–423.