

Knowledge discovery in distance relay event report: a comparative data-mining strategy of rough set theory with decision tree

ABSTRACT

A protective relay performance analysis is only feasible when the hypothesis of expected relay operation characteristics as decision rules is established as the knowledge base. This has been meticulously accomplished by soliciting the relay knowledge domain from protection experts who are usually constrained by their experience and expertise. Manually analyzing an event report is also cumbersome due to the tremendous amount of data to be perused. This paper addresses these issues by intelligently divulging the knowledge hidden in the relay recorded event report using a data-mining strategy based on rough set theory and a rule-quality measure under supervised learning to discover the relay decision algorithm and association rule. The high prediction accuracy rate and the close-to-unity areas under ROC curve value of the relay operating characteristic curve of the discovered relay decision algorithm verifies its generalized ability to predict trip status in an expert system of relay performance analysis. The relay association rule that was subsequently discovered by using the rule-quality analysis had also been verified as being a reliable hypothesis of the relay operation characteristics. This hypothesis helps the protection engineers understand the behavior of the distance relay. These rules would then be compared with and validated by benchmarking decision-tree-based data-mining analysis.

Keyword: Association rule; Data mining; Digital protective relay; Distance protection; Knowledge discovery in database; Rough set theory