

Rough Set Theory and Association Rule Mining for Detecting Interactions and Improving Machine Learning Predictions for Weather Prediction in Kenya.

Isaac Kega¹, Lawrence Nderu², Ronald Mwangi², and Dennis Njagi²

¹Zetech University

²Jomo Kenyatta University of Agriculture and Technology

August 11, 2023

Abstract

Recently, the ever-increasing complexity of datasets has necessitated the development of sophisticated techniques to uncover meaningful patterns and interactions within the data. This paper investigates the synergy between Rough Set Theory and Association Rule Mining, which is a potent approach to detecting interactions and enhancing the prediction capabilities of machine learning models. The proposed framework leverages the Greedy Heuristic Method for reduct generation, an established technique in Rough Set Theory, to efficiently identify relevant features and reduce the dimensionality of the dataset. Furthermore, Association Rule Mining extracts association rules from the data, revealing interesting relationships and dependencies among the features. These association rules are transformed into binary values, representing the detected interactions, to create a concise yet informative representation of the data's intrinsic relationships. This binary representation is ideal for integration into machine learning models, enabling them to exploit the discovered interactions and gain a more comprehensive understanding of the underlying patterns. To assess the effectiveness of our proposed framework, we propose a comprehensive experiment involving a weather dataset scraped from www.wunderground.com for Kariki farm in the Juja sub-county, Kiambu County, Kenya. Using detected interactions, we modelled them to base machine learning models, including Naive Bayes, Decision Trees, Support Vector Machines (SVM), and Logistic Regression models. We compared the performance of these models while using the detected interactions versus not using the detected interactions. Through extensive experimentation, we demonstrate that our proposed approach is more effective than traditional machine learning models without interaction detection. Our results indicate that our interaction detection method framework significantly improves the prediction accuracy of the tested models on the benchmark datasets. This enhancement in accuracy highlights the practical relevance and potential benefits of adopting our approach to uncover valuable insights from datasets.

Hosted file

RoughSet feature selection.docx available at <https://authorea.com/users/505858/articles/659498-rough-set-theory-and-association-rule-mining-for-detecting-interactions-and-improving-machine-learning-predictions-for-weather-prediction-in-kenya>