

A Dataset of Program Educational Objectives Mapped to ABET Outcomes: Data Cleansing, Exploratory Data Analysis and Modeling

Authors : Addin Osman, Anwar Ali Yahya, Mohammed Basit Kamal

Abstract : Datasets or collections are becoming important assets by themselves and now they can be accepted as a primary intellectual output of a research. The quality and usage of the datasets depend mainly on the context under which they have been collected, processed, analyzed, validated, and interpreted. This paper aims to present a collection of program educational objectives mapped to student's outcomes collected from self-study reports prepared by 32 engineering programs accredited by ABET. The manual mapping (classification) of this data is a notoriously tedious, time consuming process. In addition, it requires experts in the area, which are mostly not available. It has been shown the operational settings under which the collection has been produced. The collection has been cleansed, preprocessed, some features have been selected and preliminary exploratory data analysis has been performed so as to illustrate the properties and usefulness of the collection. At the end, the collection has been benchmarked using nine of the most widely used supervised multiclass classification techniques (Binary Relevance, Label Powerset, Classifier Chains, Pruned Sets, Random k-label sets, Ensemble of Classifier Chains, Ensemble of Pruned Sets, Multi-Label k-Nearest Neighbors and Back-Propagation Multi-Label Learning). The techniques have been compared to each other using five well-known measurements (Accuracy, Hamming Loss, Micro-F, Macro-F, and Macro-F). The Ensemble of Classifier Chains and Ensemble of Pruned Sets have achieved encouraging performance compared to other experimented multi-label classification methods. The Classifier Chains method has shown the worst performance. To recap, the benchmark has achieved promising results by utilizing preliminary exploratory data analysis performed on the collection, proposing new trends for research and providing a baseline for future studies.

Keywords : ABET, accreditation, benchmark collection, machine learning, program educational objectives, student outcomes, supervised multi-class classification, text mining

Conference Title : ICDMBDDDS 2018 : International Conference on Data Mining, Big Data, Database and Data System

Conference Location : Toronto, Canada

Conference Dates : June 21-22, 2018