

Closed and Maximal Tree Mining Using Natural Representations

José L. Balcázar

Albert Bifet

Antoni Lozano

balqui@lsi.upc.edu

abifet@lsi.upc.edu

antoni@lsi.upc.edu

Universitat Politècnica de Catalunya, Departament de Llenguatges i Sistemes Informàtics

1. Introduction

Mining frequent trees is becoming an important task, with broad applications including chemical informatics, computer vision, text retrieval, bioinformatics, and Web analysis. Many link-based structures may be studied formally by means of unordered trees [...]

the closed graph patterns discovered.

In the case of trees, there are two broad kinds of subtrees considered in the literature: subtrees which are just induced subgraphs, called induced subtrees, and subtrees where contraction of edges is allowed, called embedded subtrees[...]