

Data Mining - Blatt 03

Thomas, Manuel, Marius

November 14, 2017

1 Nr 1

Die Ausgabe .csv für jedes Datenset und jeden Threshold liegt in output_nr1. In den Spalten ClosedSet und FreeSet stehen die jeweiligen Sets pro Dataset. Mit Hilfe des ClosedSets ist es möglich nicht nur alle Solution Patterns wie bei den Boardern zu finden. Sondern es ist auch möglich die Frequencies der Pattern zu rekonstruieren. Das FreeSet ist ein Spezialfall von δ -freeSets und die ClosedSets sind die closures der FreeSets.

2 Nr 2

Die Ausgabe von dem FPGrowth-Algorithmus liegt im Ordner output_nr2 der Java Code im Ordner Nr2Weka. Die Zeiten von dem FPGrowth-Algorithmus sind wesentlich schneller als die von dem uns programmierten Apriori-Algo.