

Übungen zur Vorlesung
Wissensentdeckung in Datenbanken
Sommersemester 2007
Blatt 4

Aufgabe 4.1

Die folgende Aufgabe behandelt den in der Vorlesung vorgestellten Algorithmus FP-growth. Bei Verständnisproblemen lesen Sie am besten das Papier von Han et. al.: "Mining Frequent Patterns without Candidate Generation".

Aufgrund des Feiertages (am 01.05.) ist der Abgabetermin dieses Übungszettels erst am 09.05.2007

Verein	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Schalke	X	X	X	X	X	X	X	X	X		X	X
Gladbach		X		X	X	X			X	X	X	
Bochum	X					X	X	X				
Hamburg		X			X	X	X					
Berlin				X				X				
Hannover	X				X		X		X			
Nuernberg	X		X	X	X	X	X	X	X		X	X
Stuttgart			X	X				X	X			X
Bayern										X		

- Bestimmen Sie die Header-Tabelle sowie den FP-tree aus der angegebenen Transaktionstabelle. Arbeiten Sie (auch in den folgenden Aufgaben) mit einem minimum-support von $1/3$.
- Bestimmen Sie alle conditional pattern bases zum FP-tree.
- Bestimmen Sie nun zu den conditional pattern bases die conditional FP-trees. Achten Sie auch hier auf den minimum-support.
- Bestimmen Sie anhand der conditional FP-trees rekursiv die frequent patterns. Zeigen Sie die Erfassung der frequent patterns jeweils an der Entwicklung der conditional pattern bases sowie den conditional FP-trees.