

Ontológia bányászat

Önálló laboratórium feladat összefoglalója (II. Félév)

Székely Gergő Kristóf (XX2OAI)

Konzulens: Pataricza András, Kocsis Zsolt

**BME Méréstechnika és Információs Rendszerek Tanszék
Informatikai infrastruktúra tervezése szakirány, 2006/2007. II. félév**

A félév során új feladatot kezdtem Önálló laboratórium keretében, a témám címe Ontológia bányászat. Ez egy olyan új területe az adatbányászatnak, ahol megpróbáljuk kihasználni, hogy a vizsgált témáról előzetes információval rendelkezünk. Ennek a tudásnak a reprezentációja egy ontológia lehet, hiszen az ontológia egy tudásterület alapfogalmainak, valamint ezek összefüggéseinek géppel értelmezhető definícióit tartalmazza. A módszer arra próbál meg válaszolni, hogy az adatbányászati elemzés során kialakított modell (pl. Döntési fa) hol jelenik meg az ontológia struktúrájában.

A félév során irodalmazás keretében megismertem az adatbányászat legfontosabb technikáit, és algoritmusait. Elolvastam az OWL ontológia leíró nyelv W3C ajánlásait, valamint feldolgoztam a témához kapcsolódó számos tanulmányt. A kapcsolódó szakirodalom olvasása mellett a felhasznált szoftverek kezelését is elsajátítottam.

Elvégeztem egy, a tanszéken üzemelő Apache webservert logjainak adatbányászati elemzését SPSS Clementine eszközzel. A vizsgálat tárgya az volt, hogy a webservert működésében okoznak-e fennakadást a külső támadások, melyek nyomai a logokban megjelennek. Eredményül azt kaptam, hogy ezek a támadások nincsenek hatással a szerver működésére.

A félév során megismertem az ontológiákat. A Protegé eszközzel elkészítettem egy Pizza témakört leíró ontológiát. Az elkészült OWL dokumentum vizsgálatát Racer szoftverrel végeztem.

Hosszú távú cél egy olyan módszer kialakítása, melyben van egy hibamodellünk a meglévő rendszerhez, célunk pedig adatbányászati eszközökkel a rendszer logokból kinyerni azokat a területeket, amiket a hibamodell nem tartalmaz.