



**UNIVERSITATEA "POLITEHNICA"  
TIMIȘOARA**

**TEZE DE DOCTORAT**

**CONTRIBUȚII  
LA ÎMBUNĂȚĂȚIREA TEHNOLOGIEI  
DE CLASIFICARE A DATELOR**

**Ing. Claudiu-Raul Robu**

**2012**

**Seria 14: Calculatoare și Tehnologia Informației**

**Nr. 6**



**EDITURA POLITEHNICA**





**EDITURA POLITEHNICA**

ISSN-L: 2069-8216

ISSN: 2069-8216

ISBN: 978-606-554-504-5



Robu, Claudiu-Raul

## **Contribuții la îmbunătățirea tehnologiei de clasificare a datelor**

Teze de doctorat ale UPT, Seria 14, Nr. 6, Editura Politehnica, 2012, 182 pagini, 51 figuri, 67 tabele.

ISSN: 2069-8216

ISSN-L:2069-8216

ISBN:978-606-554-504-5

Cuvinte cheie:

clasificare, data mining, Ant Colony Optimization, algoritmi genetici, predicții, Weka

Rezumat:

Capitolul 1 prezintă definiții și terminologie, etapele procesului de data mining, domeniile de aplicare, instrumente utilizate, noțiunea de clasificare a datelor, algoritmi de clasificare, seturi de date utilizate la analiza algoritmilor, etc. Capitolul 2 tratează preprocesarea datelor și prezintă o sinteză a tehnicilor de preprocesare, aspecte legate de instrumentul conceput și implementat Arff Convertor și realizează o analiză statistică a datelor unor studenți. Capitolul 3 prezintă algoritmul ACO; se analizează comportamentul lui rulând un număr mare de configurații de parametri, iar în urma analizei se emit recomandări privind alegerea parametrilor. Se prezintă diverse variante ale lui ACO și se propun 4 îmbunătățiri care se implementează în algoritmul original Ant-r-Miner. În capitolul 4 se prezintă aspecte legate de algoritmii genetici utilizați la clasificarea datelor și se propune un nou algoritm genetic care se implementează în Weka. Se analizează comportamentul său la diferite valori ale parametrilor de intrare și se emite o listă de recomandări privind alegerea acestora. În capitolul 5, este concepută o nouă metodă de a realiza predicții în Weka, cu ajutorul unei interfețe dinamice, funcție de setul de date analizat. În capitolul 6 sunt analizate date legate de nașterile care au avut loc la Clinica de Obstetrică și Ginecologie Bega din Timișoara, în anul 2010. Se construiesc cu algoritmii propuși clasificatori în funcție de indicii apgar.