

UNIVERSITATEA BABEȘ-BOLYAI  
FACULTATEA DE MATEMATICĂ ȘI INFORMATICĂ

Clase speciale de probleme de optimizare discretă

**Rezumatul tezei de doctorat**

Coordonator științific:  
prof. univ. dr. Militon Frențiu

Doctorand:  
Alexandra Băicoianu

Cluj-Napoca  
2016



# Cuprins

<b>1</b>	<b>Introducere</b>	<b>19</b>
<b>2</b>	<b>Probleme de croire și de acoperire de tip dreptunghi-dreptunghi</b>	<b>27</b>
2.1	Noțiuni de bază privind problemele de croire/acoperire de tip dreptunghi-dreptunghi	27
2.2	Rezolvarea problemei de acoperire de tip dreptunghi-dreptunghi folosind algoritmul lui Euclid	30
2.2.1	Formularea problemei	30
2.2.2	Pașii algoritmului de rezolvare	30
2.3	Determinarea restricțiilor de tip ghilotină pentru o problemă de acoperire de tip dreptunghi-dreptunghi	37
2.3.1	Formularea problemei	37
2.3.2	Determinarea tăieturilor	41
2.3.3	Algoritm pentru verificarea restricțiilor de tip ghilotină	43
2.4	Determinarea restricțiilor de tip ghilotină pentru o problemă de croire de tip dreptunghi-dreptunghi	56
2.4.1	Formularea problemei	56
2.4.2	Determinarea tăieturilor	59
2.4.3	Algoritm pentru verificarea restricțiilor de tip ghilotină	61
2.5	Contribuții și rezultate	70
<b>3</b>	<b>Probleme de împachetare de tip container-container</b>	<b>72</b>
3.1	Noțiuni de bază privind problemele de împachetare	73
3.2	Ordinea topologică într-o problemă de împachetare de tip container-container	74
3.2.1	Formularea problemei	75
3.2.2	Graful atașat unei probleme de împachetare de tip container-container	83
3.2.3	Algoritm de sortare topologică	86
3.3	Determinarea unor condiții de tip ghilotină pentru o problema de împachetare de tip container-container	89

---

3.3.1	Formularea problemei	89
3.3.2	Determinarea tăieturilor	92
3.3.3	Test pentru verificarea restricțiilor de tip ghilotină	95
3.4	Contribuții și rezultate	107
<b>4</b>	<b>Optimizare cu metodologia "Logical Analysis of Data"</b>	<b>108</b>
4.1	Aspecte teoretice	111
4.1.1	Metodologia "Logical Analysis of Data"	111
4.1.2	Algoritmi de clasificare	122
4.2	Experimente computaționale	126
4.2.1	Seturi de date - Exemple și comparații	126
4.2.2	Experimente	129
4.2.3	Analiza rezultatelor	135
4.3	Concluzii și perspective de dezvoltare	138
<b>5</b>	<b>Concluzii</b>	<b>140</b>
	<b>Anexe</b>	<b>145</b>

# Listă publicații

- **A. Băicoianu:** A Comparative Study of Some Classification Algorithms using WEKA and LAD Algorithm, Annals of the Tiberiu Popoviciu Seminar of Functional Equations, Approximation and Convexity, ISSN 1584-4536, Volume 12, 2014, pp. 7381.
- **A. Băicoianu, A. Vasilescu, R. Pândaru:** Upon the performance of a Haskell parallel implementation, Bulletin of the Transilvania University of Brașov, **2013**, Series III: Mathematics, Informatics, Physics, pp. 61-72.
- (textbook) **A. Băicoianu, I. Plajer:** Formal Languages and Automata - practice, Transilvania University Press, Brașov, Romania, 2012.
- D. Marinescu, **A. Băicoianu, S. Dimitriu:** Software for Plagiarism Detection in Computer Source Code, Proc. of the 7th International Conference on Virtual Learning ICVL, 2012, pp. 373-379.
- D. Marinescu, **A. Băicoianu, D. Simian:** The determination of the guillotine restrictions for a Rectangular Three Dimensional Bin Packing Pattern, Recent Researches in Computer Science, Proc. of the 15-th WSEAS International Conference on Computers CSCC, Corfu, 2011, Volume 1, pp. 491-496.
- D. Marinescu, **A. Băicoianu:** A test of the guillotine restrictions determination for a Rectangular Three Dimensional Bin Packing Problem, WSEAS Transactions on Information Science and Applications, Volume 8, Issue 6, 2011, pp. 253-263.
- A. Vasilescu, **A. Băicoianu:** Algebraic Model for the CPU Logic Unit Behaviour, Recent researches in computer science, Proc. of the 15-th WSEAS International Conference on Computers, Corfu, 2011, pp. 521-526.
- D. Marinescu, **A. Băicoianu:** The determination of the guillotine restrictions for a rectangular cutting-stock pattern, Latest Trends on Computers, Proc. of the 14-th WSEAS International Conference on Computers CSCC, ISSN: 1792-4251, Corfu, 2010, Volume 1, pp. 121-126.

- D. Marinescu, **A. Băicoianu**: An algorithm for the guillotine restrictions verification in a rectangular cutting-stock pattern, WSEAS Transactions on Computers, Volume 9, Issue 10, 2010, pp. 1160-1169.
- D. Marinescu, **A. Băicoianu**: The determination of the guillotine restrictions for a rectangular cutting-stock pattern, Latest Trends on Computers, Volume 1, 14th WSEAS International Conference on Computers, 2010, pp. 121-126.
- D. Marinescu, **A. Băicoianu**: An algorithm for determination of the guillotine restrictions for a rectangular cutting-stock pattern, Journal WSEAS Transactions on Computers archive, Volume 9 Issue 10, October 2010, pp. 1160-1169.
- **A. Băicoianu**, S. Dumitrescu: Data mining meets Economic Analysis: Opportunities and Challenges, Bulletin of the Transilvania University of Braşov, **2010**.
- D. Marinescu, **A. Băicoianu**: An Algorithm for the guillotine restrictions verification in a rectangular covering Model, WSEAS Transactions on Computers, ISSN: 1109-2750, Issue 8, Volume 8, 2009, pp. 1306-1316.
- D. Marinescu, **A. Băicoianu**: The determination of the guillotine restrictions for a rectangular covering model, Proceedings of the 13th WSEAS International Conference on Computers, Rodos, 2009, pp. 307-312.
- D. Marinescu, P. Iacob, **A. Băicoianu**: A Topological Order for a Rectangular Three Dimensional Bin Packing Problem, Proceedings of the 12th WSEAS International Conference on COMPUTERS, Heraklion, Greece, July 23-25, 2008, WSEAS Press, ISSN: 1790-5109, pp. 285-290. (ISI Web of Knowledge)
- D. Marinescu, P. Iacob, **A. Băicoianu**: A plan for loading the Boxes for a Three Dimensional Bin Packing Model, WSEAS Transactions on System, ISSN: 1109 – 2777, Issue 10, Volume 7, 2008, pp. 830-839.
- P. Iacob, D. Marinescu, **A. Băicoianu**: The generating of the cutting-covering receipts using Euclid's algorithm, Proceedings of the 12th WSEAS International Conference on COMPUTERS, Heraklion, Greece, 2008, WSEAS Press, ISBN 978-960-6766-85-5, pp. 280-284. (ISI Web of Knowledge)

# Cuvinte cheie

- optimizare discretă
- probleme de croire-acoperire
- probleme tridimensionale de încărcare/împachetare
- NP-hard
- model
- set de date
- data mining
- clasificare
- LAD
- WEKA
- OCTAVE

# Rezumat

În contextul problemelor actuale de cercetare în domeniul optimizării discrete, obiectivele cercetării acestei teze au fost orientate pe probleme de optimizare discretă care au aplicații practice. Rezolvarea unor astfel de probleme implică căutarea unei soluții optime într-o mulțime finită de potențiale soluții.

Teza reflectă propria cercetare în domeniul optimizării discrete, mai precis în studiul următoarelor tipuri speciale de probleme particulare de optimizare: probleme de croire-acoperire dreptunghiulară, probleme de încărcare a containerelor cu containere (tip special de problemă de croire tridimensională). De asemenea, s-a studiat poziția metodologiei LAD față de alți algoritmi de clasificare. Pe baza rezultatelor teoretice obținute, s-au dezvoltat algoritmi pentru rezolvarea problemelor amintite. S-a efectuat un studiu comparativ între metodologia LAD și cinci algoritmi de clasificare cunoscuți în domeniul data mining.

Teza este structurată în cinci capitole, urmate de o bibliografie, ce include 85 de titluri, autoarea tezei fiind autoare și coautoare la 17 dintre acestea. Teza conține și trei anexe, ce au rolul de a asigura o înțelegere deplină a noțiunilor din capitolul patru și dovedesc acuratețea diferitelor implementări făcute.

**Capitolul 1.** Este o introducere în domeniul optimizării discrete, mai precis în problemele de croire-acoperire considerate. Scopul capitolului este de a furniza informațiile de bază necesare parcurgerii acestei teze. De asemenea, introducerea precizează și principalele contribuții ale autorului în domeniul tratat de lucrare, acela al optimizării discrete.

**Capitolul 2.** Conține rezultatele teoretice noi, care au condus la generarea următorilor algoritmi: algoritm pentru rezolvarea problemelor de croire dreptunghiulară folosind algoritmul lui Euclid de calcul al celui mai mare divizor comun a două numere, algoritm pentru verificarea restricțiilor de tip ghilotină utilizând descompunerea unui graf în componente conexe, algoritm pentru verificarea restricțiilor de tip ghilotină când este permisă și existența unor spații.

În primul paragraf se prezintă problema generală de croire/acoperire.

În paragraful doi se formulează problema de croire-acoperire de tip dreptunghi-dreptunghi și se introduce un algoritm de soluționare a problemei tratate, algoritm polinomial care se bazează pe algoritmul lui Euclid de calcul al celui mai mare divizor comun. Am optimizat complexitatea algoritmului clasic folosind un algoritm de tip



Greedy. Demonstarea complexității metodei introduse se bazează pe algoritmul lui Euclid și pe teorema lui Lamé. Am evidențiat proprietățile algoritmului introdus de către noi în Teoremele 1, 2 și Lema 3, și am construit o mulțime de soluții. Toate rezultate din acest paragraf sunt incluse în lucrarea [34].

Al treilea paragraf tratează problemele de croire-acoperire de tip dreptunghi-dreptunghi cu restricții de tip ghilotină. Utilizând reprezentarea sub formă de graf a modelului de croire-acoperire, în această secțiune prezentăm o metoda analitică de a testa satisfacerea restricțiilor de tip ghilotină, bazându-ne pe descompunerea unui graf în componente conexe.

Reprezentarea sub forma de graf este necesară pentru a demonstra conexiunea dintre restricțiile de tip ghilotină și componentele conexe ale unui graf, în conformitate cu Teoremele 13 și 14. Plecând de la această conexiune am dezvoltat un algoritm care poate fi utilizat pentru a verifica restricțiile de tip ghilotină ale unui model de problemă de acoperire de tip dreptunghi-dreptunghi. Complexitatea algoritmului este  $O(km)$ , unde  $m$  este numărul de arce ale grafului atașat modelului iar  $k$  este numărul de bucăți dreptunghiulare. Teza cuprinde și un exemplu extins, cu toate iterațiile, în vederea unei mai bune ilustrări a algoritmului propus.

În ultimul paragraf al acestui capitol se studiază o problemă de croire de tip dreptunghi-dreptunghi, în care se permite existența unor spații, dar nu se permit suprapuneri. Folosind rezultatele din secțiunea precedentă, modelului  $i$  se atașează un graf cu costuri pentru relațiile de adiacență considerate. În acest scop s-au introdus definițiile 15, 16, 17 și 18. Am propus o metodă pentru a verifica satisfacerea restricțiilor de tip ghilotină core-spunzătoare unui model care permite spații și care nu permite suprapuneri.

Algoritmul generat de aceasta metoda utilizează ca date de intrare cele două grafuri cu costuri, care caracterizează modelul dreptunghiular considerat. Rezultatul algoritmului este o reprezentare a modelului ca o formulă în formă poloneză prefixată. Ea indică ordinea în care trebuie efectuate tăieturile în vederea obținerii modelului dorit. Algoritmul folosește parcurgerea în preordine a unui graf, de aici obținându-se și complexitatea algoritmului  $O(k^2)$ , unde  $k$  este numărul de itemi considerați.

**Capitolul 3.** Acesta este dedicat studierii problemelor de croire tridimensionale de tip încărcare/împachetare container-container. Tipul de problemă studiat corespunde situației concrete ce constă în încărcarea unui container finit tridimensional cu containere tridimensionale de dimensiuni mai mici date, fără spațiere sau suprapunere. Obiectivul urmărit a fost acela de a determina o ordine pentru încărcarea containerelor, considerând următoarea restricție: un container se poate încărca numai dacă nu sunt spații libere sub container. Poziția de așezare a containerului va fi determinată, după o regulă fixată, ținând cont de containerele situate la Vest și Nord de el.

Prima secțiune a acestui capitol prezintă noțiuni de bază privind problemele tridimensionale de încărcare/împachetare.

În a doua secțiune prezentăm o generalizare a unui algoritm dezvoltat pentru problemele bidimensionale, algoritm care stabilește ordinea corectă de încărcarea a unor containere astfel încât să nu existe spațiere sau suprapunere. Algoritmul prezentat se bazează

pe construirea anterioară a unui graf cu costuri atașat problemei, respectiv pe forma unei matrice atașate grafului. Complexitatea algoritmului este liniară. Originalitatea metodei este garantată de faptul că ea se bazează pe modele matematice nou introduse. Corectitudinea ei este susținută și de faptul că grafurile utilizate extind la cazul tridimensional tipul de grafuri introdus pentru cazul bidimensional.

A treia secțiune tratează problema tridimensională de încărcare în care un container este umplut cu o mulțime de containere, fără suprapunere, dar acceptând existența unor locuri goale. Una dintre cele mai frecvente restricții impuse asupra unei soluții pentru problemele de încărcare de tip container-container este restricția de tip ghilotină. De aceea, obiectivul acestei secțiuni este de a obține o metoda pentru a verifica dacă o soluție a unei probleme de încărcare satisface restricțiile de acest tip sau nu.

Pentru aceasta, s-a generat reprezentarea sub forma unui graf cu costuri a modelului și s-a generalizat metoda obținută în capitolul precedent pentru problema bidimensională. În vederea ilustrării noțiunilor nou introduse și a algoritmului generat, s-a construit un exemplu care permite parcurgerea tuturor iterațiile algoritmului propus.

**Capitolul 4.** Este dedicat studierii unei metodologii care implică optimizare discretă și data mining, cunoscută sub denumirea de LAD (Logical Analysis of Data). Obiectivul este de a prezenta această metodologie, care face clasificare bazată pe principii data mining, combinatorică, funcții booleene și cercetări operaționale. De asemenea, am comparat metodologia LAD cu câțiva dintre cei mai cunoscuți algoritmi de clasificare: C4.5, regresia logistică, Random Forest, Perceptron Multistrat și SVM (Support Vector Machine). Algoritmii considerați au fost aleși ținându-se cont de 2 criterii: să aibe caracteristici asemănătoare cu LAD și să fie algoritmi de clasificare validați și cunoscuți în domeniul data mining. Metodologia LAD este destul de recentă, de aceea o comparare cu acești algoritmi este absolut justificată. Principalele motive pentru care autorul aceste teze a ales WEKA au fost flexibilitatea și dorința de a testa într-un mediu comun toți algoritmii considerați. Implementările din OCTAVE vin ca și adăugare, spre a consolida rezultatele obținute cu WEKA.

În prima secțiune a capitolului s-au prezentat principalele noțiuni care stau la baza metodologiei LAD, precum și diverse rezultate obținute în domeniul data mining cu aceasta. Etapele obținerii modelelor LAD implică cunoașterea unor noțiuni de optimizare discretă, cercetări operaționale și data mining. Tot în această secțiune, vom prezenta succint noțiuni de bază despre utilizarea WEKA și OCTAVE, precum și elemente despre funcționalitatea algoritmilor de clasificare aleși în comparațiile ulterioare.

În a doua secțiune, în prima parte s-au descris seturile de date utilizate în comparații, dar și cele două direcții de comparare propuse de către autor. Tot aici, amintim aportul original al autorului în dezvoltarea algoritmilor propuși, precum și modificările realizate în implementarea originală a LAD-ului. În continuare, s-a realizat efectiv comparația cu algoritmii de clasificare amintiți și am expus toate rezultatele obținute sub forma unor tabele comparative. Implementarea clasică folosită pentru LAD a fost modificată de către autor pentru obținerea mai rapidă a rezultatelor dorite. În Anexa 3 sunt prinse toate aceste modificări. Pentru algoritmii de clasificare am folosit WEKA, care oferă o

implementare rapidă și destul de simplă de utilizat pentru o familie generoasă de algoritmi de data mining/machine learning. În urma comparației realizate pe trei seturi de date diferite, s-a observat faptul că rezultatele obținute cu LAD sunt absolut remarcabile, în sensul că sunt foarte apropiate de cele obținute cu algoritmi de clasificare menționați, uneori chiar mai bune. Pentru anumite cazuri particulare, modelele obținute cu LAD sunt și foarte intuitive și ușor de justificat.

Testele realizate cu WEKA, OCTAVE și LAD dovedesc faptul că alegerea LAD-ului ca și o unealtă data mining utilă este justificată prin:

- noțiunile (relative simple) pe care se construiesc modelele LAD → vezi probleme de acoperire, funcții booleene, simplificarea funcțiilor booleene, funcții parțial definite, metoda simplex, etc;
- eficiența patternurilor oferite de LAD → LAD oferă (ca toți ceilalți algoritmi) niște modele care acoperă observațiile pozitive și observațiile negative, însă spre deosebire de ceilalți algoritmi considerați, aceste modele sunt justificabile, adică cele pozitive acoperă numai observații validate ca și pozitive, respectiv nimic din cele negative, și invers. În aplicarea unui algoritm de clasificare, sunt foarte importante calitatea și transparența modelelor generate, LAD conferă aceste două proprietăți, oferind modele ușor de înțeles pentru specialiștii din diferitele domenii în care se aplică;
- acuratețea oferită de LAD → în urma experimentelor realizate de către autor, s-a observat faptul că rezultatele oferite de LAD sunt clare și cu un procentaj ridicat al acurateței, apropiat sau care depășește procentul obținut cu alți algoritmi de clasificare cunoscuți;
- timpul de execuție → având la bază procesarea de funcții booleene, generarea unor modele valide este un proces cu o complexitate redusă;

Rezultatele extraordinare obținute în domeniul medicinei, și nu numai, ne-au determinat să alegem această metodologie pentru clasificare de instanțe. Pentru a justifica concluziile prezentate relative la poziția metodologiei LAD față de algoritmi de clasificare considerați, în Anexa 3 au fost aleși doi dintre algoritmi de clasificare și s-a realizat și o implementare în OCTAVE. Și în cazul acestor implementări s-a observat faptul că timpul de execuție al LAD-ului și a algoritmilor implementați în OCTAVE este comparabil. Recomandăm LAD pentru seturile de date cu conținut medical, fără înregistrări nevalide/incomplete. Răspândirea LAD-ului în domeniul medical este explicat prin faptul că LAD generează modele justificabile/explicabile, adică modele ce pot fi înțelese de specialiștii în medicină. În acest domeniu, o astfel de metodologie este importantă, pentru că aduce după sine automatizarea procesului de repartizare a unor bolnavi, respectiv de clasificare în bolnavi/sănătoși. Clar, acest lucru înseamnă timp redus de diagnosticare. Bineînțeles că în alte domenii, obținerea unor modele "cu sens" nu e neapărat o prioritate, important fiind numai procentajul efectiv de clasificare corectă.

În ultima secțiune a acestui capitol, descriem perspectivele de dezvoltare. Există o implementare clasică pentru LAD, însă aceasta are câteva neajunsuri importante, amintim: codul nerefactorizat scris în C++, metode neoptime pentru diferitele etape ale metodologiei (exemplu: generarea modelelor de clasificare), implementarea procedurală. Vom încerca dezvoltarea unui plugin pentru LAD, în vederea unei mai largi utilizări ale acestei metodologii. Tot ca o perspectivă de viitor, vom încerca dezvoltarea unor algoritmi de machine learning sub HASKELL. Există deja câteva rezultate obținute cu biblioteca HLearn [12], intenționăm o corelare cu metodologia LAD.

**Capitolul 5.** Conține concluziile autoarei tezei desprinse din cercetările efectuate și prezintă opțiunile acestora relative la cercetările ulterioare pe care dorește să le întreprindă.

**Anexa 1** conține codul în OCTAVE scris de autoarea tezei pentru regresia logistică precum și pentru perceptronul multistrat. Algoritmii au fost implementați într-un limbaj care oferă ușurința în lucrul cu matrici și vectori, precum și operațiile dintre acestea. Obiectivul implementării a fost de a sublinia că timpul de execuție și acuratețea LAD-ului sunt foarte apropiate de cele ale acestor algoritmi implementați, indiferent de limbajul folosit. Pentru fiecare algoritm s-a precizat și acuratețea rezultatului obținut.

**Anexa 2** conține o parte din modificările implementării clasice a LAD-ului, prelucrări făcute de autoarea tezei. Amintim faptul că s-au realizat mai multe modificări, dintre care enumerăm:

- adăugarea unei funcționalități noi, mai precis utilizarea aplicației pe Windows;
- convertirea aplicației dintr-o aplicație din linia de comanda, într-o aplicație de tip consola;
- rescrierea unor metode folosite pentru pașii metodologiei LAD;
- adăugarea unor metode noi, în vederea unei comparații complete cu cei cinci algoritmi de clasificare (metodă pentru "k-fold cross validation", metodă pentru deviația standard, etc).

**Anexa 3** conține rezultate computaționale obținute de autoarea tezei, utilizate în capitolul patru pentru a trage concluziile privitoare la deosebiri și asemănări între metoda LAD și următorii cinci algoritmi de clasificare: C4.5, Random Forest, Perceptron Multistrat, SVM, regresie logistică.

**Contribuțiile originale** introduse în teză se află în Capitolele 2, 3 și 4 și sunt următoarele:

În capitolul 2:

- algoritmul pentru rezolvarea unor probleme de croire de tip dreptunghi-dreptunghi, folosind algoritmul lui Euclid. Am construit o mulțime de soluții și am stabilit că soluția construită este o construcție geometrică a algoritmului lui Euclid, teorema

1. De aceea, am concluzionat că metoda bazată pe numărul de diviziuni din algoritmul lui Euclid se bazează pe teorema lui Lamé, iar complexitatea algoritmului propus este:  $O(5 * MAX(numberOfDigits(r_a), numberOfDigits(r_b)))$ , unde  $r_a$  și  $r_b$  sunt resturile din ecuațiile propuse, 1.

- teoremele 13 și 14 care conduc la un algoritm pentru verificarea restricțiilor de tip ghilotină, utilizând descompunerea unui graf în componente conexe. Corectitudinea algoritmului este dovedită prin teoremele amintite iar complexitatea metodelor introduse, V-CUT și H-CUT, s-a demonstrat a fi  $O(m)$ , unde  $m$  este numărul de arce.
- rezultatele din teorema 22 pe baza cărora s-a construit în teză un algoritm pentru verificare restricțiilor de tip ghilotină, în cazul mai general când sunt permise spații.  
Algoritmul construiește un arbore sintactic pentru o reprezentare a modelului de croire considerat, iar la fiecare pas verifică dacă se poate face o tăietură verticală sau orizontală (procedurile VCUT, HCUT). Corectitudinea este asigurată de 22, iar complexitatea obținută este  $O(k^2)$ .

În capitolul 3:

- s-au generalizat tipurile de adiacență folosite pentru modelele bidimensionale și s-au obținut trei tipuri de grafuri, în direcțiile  $Ox$ ,  $Oy$  și  $Oz$ .
- s-au introdus modelele matematice noi și prin teoremele 32, 33 s-au stabilit proprietăți importante pentru tipurile de grafuri atașate. S-a demonstrat, teorema 35, că grafurile sunt aciclice.
- s-a prezentat o posibilă ordine a containerelor la încărcarea lor - algoritmul de sortare OVERDIAG-3D.
- s-a introdus un nou tip de graf, cu costuri pentru modelarea soluției problemei tridimensionale, 43. Prin Teorema 45 s-au demonstrat noi proprietăți ale grafurilor atașate modelului. Rezultatele din teorema sugerează un algoritm pentru a verifica restricțiile de tip ghilotină, în cazul unui model tridimensional cu spații, dar fără suprapunere. S-a remarcat faptul că acest algoritm poate fi aplicat și pentru cazul fără spații, iar pentru aceasta s-a considerat un exemplu extins.

În capitolul 4, am realizat o comparație între metodologia LAD și cinci algoritmi de clasificare. În această comparație s-a folosit o variantă modificată a implementării clasice a LAD-ului, variantă construită de autoarea acestei teze și implementarea în WEKA și OCTAVE a celor cinci algoritmi. Comparația s-a făcut utilizând cea mai nouă variantă de WEKA și urmărind două direcții principale: realizarea testelor cu parametrii default furnizați de WEKA - Explorer, repetare teste cu parametrii optimizați - Experimenter. În urma comparației, s-a dovedit faptul că, cel puțin din cele două perspective

de comparație: acuratețe (cu deviație standard luată în considerare), matrici de confuzie, LAD este o metodologie care se apropie foarte bine de algorimii considerați, uneori chiar cu rezultate mai bune. Direcțiile de comparație, precum și testele, implementările și concluziile, aparțin autorului.

# Bibliografie

- [1] G. Alexe, S. Alexe, P.L. Hammer: Pattern-Based Clustering and Attribute Analysis, *Soft Computing Springer Journal*, 2006, pp. 442-452.
- [2] G. Alexe, S. Alexe, T.O. Bonates, A. Kogan: Logical analysis of data the vision of Peter L. Hammer, *Annals of Mathematics and Artificial Intelligence*, 49(1-4):265-312, 2007.
- [3] G. Alexe, P.L. Hammer: Spanned patterns for the logical analysis of data, *Discrete Appl. Math.*, 154 (2006), pp. 1039 - 1049.
- [4] G. Alexe, P.L. Hammer: Accelerated algorithm for pattern detection in logical analysis of data, *Discrete Appl. Math.*, 154 (2006), pp. 1050 - 1063.
- [5] G. Alexe, S. Alexe, P.L. Hammer, A. Kogan: Comprehensive vs. comprehensible classifiers in logical analysis of data, *Discrete Applied Mathematics*, in press, 2007.
- [6] S. Alexe, et al.: Coronary Risk Prediction by Logical Analysis of Data. *Annals of Operations Research*, 119:15 - 42, 2003.
- [7] R.R. Amossen, D. Pisinger: Multi-dimensional Bin Packing Problems with Guillotine Constraints, *Computers & OR* 37, No. 11, pp. 1999-2006, 2010.
- [8] M.Z. Arslanov, D.U. Ashigaliev, E.E. Ismail: Polynomial algorithms for guillotine cutting of a rectangle into small rectangles of two kinds *European, Journal of Operational Research*, Vol. 185, Issue 1, 2008, pp. 105-121.
- [9] R.E. Banfield: A comparison of decision tree ensemble creation techniques, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29 (2007), pp. 173-180.
- [10] A. Băicoianu, I. Plajer: *Formal Languages and Automata - practice*, Transilvania University Press, Braşov, Romania, 2012.
- [11] A. Băicoianu: A Comparative Study of Some Classification Algorithms using WEKA and LAD Algorithm, *Annals of the Tiberiu Popoviciu Seminar of Functional Equations, Approximation and Convexity*, ISSN 1584-4536, vol 12, 2014, pp. 7381.

- [12] A. Băicoianu, A. Vasilescu, R. Pândaru: Upon the performance of a Haskell parallel implementation, Bulletin of the Transilvania University of Braşov, **2013**, Series III: Mathematics, Informatics, Physics, 61-72.
- [13] A. Băicoianu, S. Dumitrescu: Data mining meets Economic Analysis: Opportunities and Challenges, Bulletin of the Transilvania University of Braşov, 2010.
- [14] T.O. Bonates, P.L. Hammer: Pseudo-Boolean Regression, RRR 3-2007, January, 2007.
- [15] E. Boros, P.L. Hammer, T.I Baraki, A. Kogan: Logical analysis of numerical data, Mathematical Programming, 79 (1997), pp. 163 - 190.
- [16] E. Boros, P.L. Hammer, et al.: An implementation of logical analysis of data. IEEE Transactions on Knowledge and Data Engineering 12 (2), 2000,292 – 306.
- [17] E. Ciurea, L. Ciupală: Algoritmi - Introducere în algoritmica fluxurilor în reţele, Ed. Matrix ROM Bucureşti, 2006.
- [18] P. Chen, Y. Chen, M. Goel, F. Mang: Approximation of Two-Dimensional Rectangle Packing, CS270 Project Report, 1999, [www.cad.eecs.berkeley.edu/~fmang/cs270/report.pdf](http://www.cad.eecs.berkeley.edu/~fmang/cs270/report.pdf).
- [19] T.H. Cormen, C.E. Leiserson, R.R Rivest: Introduction to Algorithms, MIT Press, 1990.
- [20] I.G. Czibula, G. Şerban: A Hierarchical Clustering Algorithm for Software Design Improvement, KEPT 2007, Knowledge Engineering: Principles and Techniques, Babeş-Bolyai University, June 6-8, 2007, Cluj-Napoca, pp. 316-323.
- [21] I.G. Czibula, G.S. Cojocar: A Hierarchical Clustering Based Approach in Aspect Mining, Computing and Informatics, Vol. 29, 2010, 881900.
- [22] D. Dumitrescu, H.F. Pop, C. Sârbu: Fuzzy hierarchical cross-classification of Greek muds, Journal of Chemical Information and Computer Science, 35, pp. 851-857, 1995.
- [23] D. Dumitrescu: Principiile matematice ale teoriei clasificării, Academiei Române, Bucureşti, 1999.
- [24] H. Dyckhoff: A typology of cutting and packing problems, European Journal of Operational Research 44 (1990) 145-159.
- [25] C. Filippi, A. Agnetis: An asymptotically exact algorithm for the high-multiplicity bin packing problem, Mathematical Programming, 2005.



- [26] M. R. Garey, D. S. Johnson: *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, 1979.
- [27] S. Godoy-Calderon, J.D. Batiz, M. Lazo-Cortez: *A non-Classical View of Coverings and its Implications in the Formalization of Pattern Recognition Problems*, Proceedings of the WSEAS Conference, 2003, paper 459-151.
- [28] P.L. Hammer: *The Logic of Cause-effect Relationships*, Lecture at the International Conference on Multi-Attribute Decision Making via Operations Research-based Expert Systems, Passau, Germany, 1986.
- [29] P.L. Hammer, T. Bonates: *Logical Analysis of Data: From Combinatorial Optimization to Medical Applications*, Rutgers Center for OR.
- [30] J. Han, M. Kamber: *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, August 2000.
- [31] D.J. Hand, H. Mannila, P. Smyth: *Principles of Data Mining*. MIT Press, Cambridge, MA, 2001.
- [32] P. Iacob, D. Marinescu, C. Luca: *Covering a rectangle with rectangular pieces*, Proceedings - Strategies and Applications, Thessaloniki, Greece, October 18-20, 2001, pp. 506-513.
- [33] P. Iacob, D. Marinescu, K. Kiss-Jakab: *Some algorithms for generating receipts in the rectangular cutting-covering problem*, NAUN, International Journal of Mathematical Models and Methods in Applied Science, Issue 3, Vol. 1., 2007, pp. 182-187.
- [34] P. Iacob, D. Marinescu, A. Băicoianu : *The generating of the cutting-covering receipts using Euclid's algorithm*, International Conference on Computers, Heraklion, Greece, July 23-25, 2008.
- [35] P. Iacob, D. Marinescu, C. Luca: *L-Shape room*, Proceeding of WMSCI , Orlando, Florida, USA, 2005, Vol III, pp. 175-179.
- [36] H.W. Ian, E. Frank: *Data Mining: Practical machine learning tools and techniques*. 2nd edition. Morgan Kaufmann, San Francisco, 2005.
- [37] Z. Jin, T. Ito: *The three Dimensional Bin Packing Problem and its Practical Algorithm*, *JSME*, Series , Vol. 46, No.1, 2003.
- [38] S.B. Kotsiantis: *Supervised Machine Learning: A Review of Classification Techniques*, *Informatica* 31(2007) 249-268, 2007.
- [39] M. Lejeune: *Pattern definition of the p-efficiency concept*, *Annals of Operations Research* 200(1):1-14, November 2012.

- [40] M. Lauer, S. Alexe, et al.: Use of the Logical Analysis of Data Method for Assessing Long-Term Mortality Risk After Exercise Electrocardiography. *Circulation*, 106:685 - 690, 2002.
- [41] P. Lémaire: The ladoscope gang: Tools for the Logical Analysis of Data. <http://www.kamick.org/lemaire/LAD/>
- [42] J.E. Lewis, R.K. Ragade, A. Kumar, W.E. Biles: A distributed chromosome genetic algorithm for bin-packing, 14th Intern. Conf. on Flexible Automation and Intelligent Manufacturing Vol. 21, Issues 4-5, Aug-Oct 2005, pp. 486-495.
- [43] D. Marinescu: An integer programming model for two-dimensional cutting stock-problems, Babeş-Bolyai University, Fac. Math. Phys., Res. Semin. 8, pp. 65-74 (1987).
- [44] D. Marinescu: A representation problem for a rectangular cutting - stock model, *Foundations of Computing and Decision Sciences*, Vol. 32, 2007, No. 3, pp. 239-250.
- [45] D. Marinescu: A bidimensional Turing machine for the cutting stock problem with guillotine restrictions, *Proceeding of the scientific symposium with the contribution of teachers and researchers from the Republic of Moldova, Braşov 1991*, pp. 73-83.
- [46] D. Marinescu: A s-picture language for a cutting-stock model with guillotine restrictions, *Bulletin of the Transilvania University of Braşov - seria C*, Vol XXXIII 1991 pp. 39-45.
- [47] D. Marinescu: Graphs attached to a rectangular cutting-stock model (French), *Itinerant Seminar of Functional Equation, Approximation and Convexity, Cluj-Napoca 1988*, Preprint No. 6, pp. 209 -212.
- [48] D. Marinescu, A. Băicoianu: The determination of the guillotine restrictions for a rectangular cutting-stock pattern, *Latest Trends on Computers, Proc. of the 14-th WSEAS International Conference on Computers CSCC*, ISSN: 1792-4251, Corfu, 2010, Vol. 1, pp. 121-126.
- [49] D. Marinescu, A. Băicoianu: An algorithm for the guillotine restrictions verification in a rectangular cutting-stock pattern, *WSEAS Transactions on Computers*, Volume 9 Issue 10, Oct. 2010, pp. 1160-1169.
- [50] D. Marinescu, A. Băicoianu: An Algorithm for the guillotine restrictions verification in a rectangular covering Model, *WSEAS Transactions on Computers*, ISSN: 1109-2750, Issue 8, Vol. 8, Aug. 2009, pp. 1306-1316.
- [51] D. Marinescu, A. Băicoianu: The determination of the guillotine restrictions for a rectangular covering model, *Proc. of the 13-th WSEAS International Conference on Computers, Crete, 2009*, pp. 307-312.
- [52] D. Marinescu, A. Băicoianu: The determination of the guillotine restrictions for a rectangular cutting-stock pattern, *Proceeding of the 14-th WSEAS CSCC, Corfu, 2010*, pp. 121 - 126.

- [53] D. Marinescu, A. Băicoianu: An algorithm for determination of the guillotine restrictions for a rectangular cutting-stock pattern, *Journal WSEAS Transactions on Computers archive*, Volume 9 Issue 10, October 2010, pp. 1160-1169.
- [54] D. Marinescu, A. Băicoianu, S. Dimitriu: Software for Plagiarism Detection in Computer Source Code, *Proc. of the 7th International Conference on Virtual Learning ICVL 2012*, pp. 373-379.
- [55] D. Marinescu, P. Iacob, K. Kiss-Jakab: A topological order for a rectangular covering model, *Proc. of the 11-th WSEAS CSCC, Crete, Vol. 4, 2007*, pp. 82-85.
- [56] D. Marinescu, P. Iacob, A. Băicoianu: A Topological Order for a Rectangular Three Dimensional Bin Packing Problem, *Proceedings of the 12th WSEAS International Conference on COMPUTERS, Heraklion, Greece, July 23-25, 2008*, WSEAS Press, ISSN: 1790-5109, pp. 285-290.
- [57] D. Marinescu, A. Băicoianu, D. Simian: The determination of the guillotine restrictions for a Rectangular Three Dimensional Bin Packing Pattern, *Recent Researches in Computer Science, Proc. of the 15-th WSEAS International Conference on Computers CSCC, Corfu, 2011, Vol. 1*, pp. 491-496.
- [58] D. Marinescu, A. Băicoianu: A test of the guillotine restrictions determination for a Rectangular Three Dimensional Bin Packing Problem, *WSEAS Transactions on Information Science and Applications, Volume 8 Issue 6, June 2011*, pp. 253-263.
- [59] D. Marinescu, P. Iacob, A. Băicoianu: A plan for loading the Boxes for a Three Dimensional Bin Packing Model, *WSEAS Transactions on System, ISSN: 1109 – 2777, Issue 10, Vol. 7, Oct. 2008*.
- [60] E. Mayoraz: C++ Tools for Logical Analysis of Data, *Technical Report, RTR 1 – 95, Rutgers University, July 1995*.
- [61] T. Papke, A. Bortfeldt, H. Gehring: Software Programs for Solving a Cutting Problem in the Wood-Working Industry. A Case Study, *WSEAS Transactions on Information Science & Applications, Issue 5, Volume 4, May 2007*, pp. 932-938.
- [62] S.A. Paschos, V.Th. Paschos, V. Zissimopoulos: A model to approximately cover an area by antennas, *Proc. of WSEAS Conference, 2003*, pp. 458-187.
- [63] Z. Pawlak: *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers Norwell, USA, 1992.
- [64] H.F. Pop, C. Sârbu, O. Horowitz, D. Dumitrescu: A fuzzy classification of the chemical elements, *Journal of Chemical Information and Computer Sciences* 36, 3 (1996), pp. 465-482.
- [65] H.F. Pop: *Sisteme inteligente în probleme de clasificare*, PhD Thesis, Babeş-Bolyai University, 1995.

- [66] H.F. Pop, M. Frențiu: Applications of Principal Components Methods, (ISI) Proceedings of the International Conference "Complexity and Intelligence of the artificial and Natural Complex Systems. Medical Applications of the Complex Systems, Biomedical Computing", IEEE Computer Society Press, 2009, pp. 103-109, ISBN 78-0-7695-3621-7.
- [67] J.R. Quinlan: C4.5: Programs for Machine Learning, Morgan Kaufmann, 1993.
- [68] K.H. Rosen: Handbook of discrete and combinatorial mathematics, CRC Press, 2000.
- [69] X. Song, C. Chu, Y. Nie: A Heuristic Dynamic Programming Algorithm for 2D Unconstrained Guillotine Cutting, Proceedings of WSEAS Conference , 2004, pp. 484-108.
- [70] P. Sweeney, E. Paternoster: Cutting and packing problems: a categorized application-oriented research bibliography, Journal of the Operational Research Society , 43/7 (1992) 691-706.
- [71] A. Vasilescu, A. Băicoianu: Algebraic Model for the CPU Logic Unit Behaviour, Recent researches in computer science, Proc. of the 15-th WSEAS International Conference on Computers, Corfu, 2011.
- [72] I.H. Witten, E. Frank: Data Mining: Practical Machine Learning Tools with Java Implementations, Morgan Kaufmann, San Francisco, 1999.
- [73] Three dimensional Packing Problems (2003) The Institute for Algorithms and Scientific Computing (SCAI, Germany).
- [74] \*\*\*, <https://hackage.haskell.org/packages/>
- [75] \*\*\*, <http://mathworld.wolfram.com/OptimizationTheory.html>
- [76] \*\*\*, <http://dictionary.sensagent.com>
- [77] \*\*\*, <http://rutcor.rutgers.edu/pub/LAD/>
- [78] \*\*\*, <http://www.cs.waikato.ac.nz/~ml/weka>
- [79] \*\*\*, <https://weka.wikispaces.com/Optimizing+parameters>
- [80] \*\*\*, <http://www.cs.waikato.ac.nz/ml/weka/documentation.html>
- [81] \*\*\*, [https://en.wikipedia.org/wiki/Confusion\\_matrix](https://en.wikipedia.org/wiki/Confusion_matrix)
- [82] \*\*\*, <https://izbicki.me/public/papers>
- [83] \*\*\*, <http://archive.ics.uci.edu/ml/datasets.html>
- [84] \*\*\*, <http://www.gnu.org/software/octave/download.html>
- [85] \*\*\*, <http://www-mdp.eng.cam.ac.uk/web/CD/engapps/octave>