

Perfecționarea sistemelor de management al relațiilor cu clienții folosind tehnologii Data Mining

The improvement of Customer Relationship Management systems by using Data Mining techniques

Prof.dr. Ioan RADU
Asist.drd. Mihai CIOC

Rezumat

Tehnica mineritului datelor permite analiza unor volume mari de informații stocate în depozite de date, în vederea rezolvării unor probleme de natură decizională. Din punctul de vedere al aplicațiilor software utilizate, tehnica își are originile în trei categorii de aplicații: cele statistice, cele specifice inteligenței artificiale, care au la bază algoritmi neuro-fuzzy și cele de învățare automată de tip “machine learning”. În domeniul managementului relațiilor cu clienții, tehnica DM este utilizată pentru analiza informațiilor istorice specifice activității de vânzări, în vederea identificării unor modele comportamentale ale consumatorilor care să permită efectuarea de previziuni cu privire la modul în care aceștia vor răspunde la ofertele lansate pe piață de către companie.

Abstract

Data mining is a technique that consists in analysing large volumes of informations stored in data warehouses, in order to resolve decision problems. The tehnique is derived from three categories of software applications: the statistical ones, artifficial intelligence applications based on neuro-fuzzy algorithms and the ones based on automated machine learning. In the field of customer relationship management, the DM technique is used to analyse hystorical informations related to sales activity, in order to identify models for consumer choice that allows predictions about the way they will respond to new offers.

1. Depozitele de date și tehnica mineritului datelor

Până recent, specialiștii considerau că sistemele clasice de prelucrare a informațiilor de tipul ERP reprezintau o soluție eficientă atât pentru gestiunea datelor cât și pentru efectuarea de analize economico-financiare. În fapt, sistemele de planificare a resurselor ofereau o viziune integrată asupra activității organizației și permiteau vizualizarea unor situații informatice sintetice sub forma tablourilor de bord. Principalul dezavantaj era acela că ERP-urile nu permiteau realizarea unor corelații complexe a datelor din sistem, ceea ce limita mulțimea rapoartelor ce puteau fi obținute de către utilizator la cele cu un grad mediu-ridicat de parametrizare. Cu alte cuvinte, sistemele nu erau capabile să răspundă la niște cerințe punctuale și specifice, impuse de complexitatea proceselor decizionale desfășurate în cadrul organizației.

De asemenea, cvasitotalitatea aplicațiilor informatice pentru management dezvoltate în cadrul sistemelor ERP funcționează, după cum se cunoaște, pe principiul băncilor de date și, implicit, a sistemelor de gestiune care asigură crearea și gestionarea acestora. Caracteristica fundamentală a acestor aplicații este dată de faptul că **elementul**

primar al prelucrărilor este informația, generată întotdeauna în sistemul informațional, prelucrată în cel informatic și interpretată, în forma sa finală, fie în subsistemul decizional, fie tot la nivelul celui informațional. În același timp, aceste aplicații nu au capacitatea de a genera la rândul lor informații suplimentare pornind de la cele stocate la un moment dat în baza de date.

Aceste restricții au impus un nou trend în dezvoltarea de produse software și anume acela al aplicațiilor din sfera inteligenței afacerilor (Business Intelligence). Datorită faptului că răspund unor cerințe decizionale de natură complexă, aplicațiile de BI se adresează în primul rând managerilor de nivel mediu și superior din cadrul companiilor sau instituțiilor publice și mai puțin personalului operațional. În sfera acestora tehnologii se detașează ca importanță sistemele OLAP, tehnicile depozitelor de date și mineritului datelor sau sistemele de management al cunoștințelor de tip KMS.

Depozitul de date sau *“data warehouse”* reprezintă o bază de date dezvoltată în scopul satisfacerii cerințelor unui sistem de inteligență a afacerilor. Din punctul de vedere al structurii, depozitul de date integrează un număr mare de categorii de date, provenite din diverse surse externe. De regulă, depozitele de date au capacități de stocare foarte mari, de ordinul terabytes și înmagazinează date care au un caracter istoric. În general, într-o organizație există trei tipuri de depozite de date:

- *Depozitul de date central*, care gestionează principalele informații legate de activitatea organizației și asigură suportul pentru desfășurarea proceselor de fundamentare a deciziilor și este utilizată de către managerii de nivel superior.
- *Depozitul de date operațional*, care gestionează date de natură operațională actualizate în timp real.
- *Depozite de date de tip Datamart*, care reprezintă subcomponente ale magaziei de date operaționale, dezvoltate la nivel de entitate organizatorică structurală sau funcțională.

Prelucrarea datelor stocate în depozite de date se realizează cu ajutorul tehnicii mineritului datelor, cunoscută în literatura de specialitate sub denumirea de *Data Mining* (DM) sau *Knowledge Discovery in Databases* (KDD). În opinia lui Shapiro, mineritul datelor este *tehnica de extragere din date a unui conținut informațional necunoscut la momentul respectiv*. Alți specialiști consideră că Data Mining definește *explorarea și analiza, folosind mijloace automatizate și semiautomatizate, a unor mari cantități de date în scopul identificării unor relații semnificative între acestea*.

Pentru a putea însă face o analiză pertinentă a conceptului, este necesară o scurtă analiză a evoluției acestei tehnologii, mai ales dacă ținem cont de faptul că, ținând cont de gradul de nouitate al instrumentelor utilizate în Data Mining, majoritatea teoreticienilor și practicienilor au opinii divergente. Tehnica mineritului datelor își are rădăcinile în trei categorii de aplicații din sfera tehnologiei informației și comunicației:

- *Aplicații statistice*, respectiv cele bazate pe concepte precum regresie, distribuție standard, deviație standard, abatere medie pătratică sau intervale de confidență, utilizate pentru analiza datelor și a relațiilor dintre acestea.
- *Aplicații specifice inteligenței artificiale*, care au la bază principiile euristicii și iau de regulă forma sistemelor expert, folosind ingineria neuro-fuzzy precum și alte tehnici derivate din teoria mulțimilor vagi.
- *Aplicații de învățarea automată de tip “machine learning”*, aflate la intersecția primelor două categorii și care combină modelele euristice cu cele statistice. De

fapt, în esență Data Mining-ul reprezintă o adapare a tehnicilor de învățare automată la nivelul aplicațiilor de afaceri sau din mediul administrației publice.

Din cele de mai sus, rezultă că principala contribuție a mineritului datelor în procesul de fundamentare a deciziilor este capacitatea acestuia de a ușura procesul de analiză a datelor. Evident că succesul procesului este dependent de intuiția utilizatorilor și de capacitatea acestora de a rafina informația primită de la sistem și de a formula în mod corespunzător regulile de filtrare a datelor. Misiunea principală a Data Mining-ului este, în această situație, de a descoperi asocieri sau modele ale datelor, de corelații și tendințe folosind tehnici statistice și matematice, în timp de decidenții vor selecta din informațiile primite alternativa decizională optimală. Pentru a putea realiza acest obiectiv, aplicațiile din sfera Data Mining-ului îndeplinesc trei categorii de funcții:

- **Funcții de grupare a datelor**, care sunt utilizate, de regulă, în prima fază a mineritului datelor și care vizează identificarea grupurilor de date care descriu caracteristici sau comportamente similare.
- **Funcții de ierarhizare a datelor**, folosite în vederea clasificării datelor în funcție de semnificația și importanța pe care o au relativ la o anumită situație decizională.
- **Funcții de interogare**, care permit utilizatorului să solicite furnizarea anumitor rezultate ale procesării datelor, în funcție de un anumit set de cerințe.

Datorită acestor funcționalități, aplicațiile de Data Mining se utilizează pe scară largă sub forma unor componente integrate în cadrul unor sisteme precum CRM (Customer Relationship Management) unde are rolul de a analiza caracteristicile unice ale clienților unei organizații în baza istoricului acestora, în vederea identificării unor trenduri în evoluția segmentelor de consumatori. De asemenea, mineritul datelor este folosit și pentru analiza datelor în arhitecturile multidimensionale de tip OLAP sau pentru transformarea datelor în cunoștințe în cadrul aplicațiilor din sfera KMS (Knowledge Management Systems).

2. Sisteme de management al relațiilor cu clienții (CRM)

Planificarea resurselor organizației (ERP) trebuie să pornească de la principala verigă a lanțului de furnizare și anume consumatorul final. În vederea optimizării modului în care o companie interacționează cu clienții săi, implementarea de sisteme ERP este însoțită de cele mai multe ori de dezvoltarea unei componente CRM. Un sistem de management al relațiilor cu clienții (CRM Customer Relationship Management) reprezintă o un ansamblu de aplicații informatice și proceduri care vizează identificarea principalelor așteptări și preferințe ale clienților organizației și utilizarea eficientă a informațiilor acumulate în scopul perfecționării relațiilor cu aceștia.

Implementarea unui astfel de sistem are drept obiectiv fundamental creșterea gradului de satisfacere a cerințelor explicite și implicite ale clienților și presupune dezvoltarea a două **componente**:

- **componenta managerială** reprezentată de ansamblul metodelor și tehnicilor utilizate pentru integrarea și utilizarea datelor referitoare la comportamentul consumatorilor;
- **componenta informatică** care include totalitatea echipamentelor hardware și a componentelor software utilizate pentru colectarea, stocarea și gestionarea

informațiilor. În această categorie sunt cuprinse aplicațiile **CIS (Customer Interaction Software)** care prezintă, drept principale funcționalități : preluarea datelor referitoare la clienți utilizând machete standardizate pentru comenzi, cereri, solicitări, reclamații etc.; evidența clienților existenți și a celor potențiali, prin prisma unui set de caracteristici unice, păstrându-se și un istoric al evoluției acestora; consolidarea datelor colectate în scopul analizei și gruparea informațiilor obținute în rapoarte de sinteză și reprezentări grafice.

Din punct de vedere arhitectural, un sistem de management al relațiilor cu clienții asistat de calculator integrează o gamă largă de tehnologii informaționale precum ICT(Integrated Computer Telephony) sau IVR(Intelligent Voice Responce). **Principalele componente ale sistemelor CRM** sunt :

- a. **Ghișeu unic** reprezintă punctul de intrare în sistem a informațiilor legate de solicitările și sesizările clienților, de regulă sub forma unor documente, care urmează să fie apoi preluate și operate în cadrul un sistem de management al documentelor și fluxurilor de lucru.
- b. **Centrul de contact** (Call Center/ Contact Center/ Help Desk) reprezintă o entitate din cadrul unei organizații, care acordă asistență de specialitate clienților care solicită informații referitoare la produsele și serviciile furnizate. Dezvoltarea unei astfel de componente generează o serie de beneficii : micșorarea numărului de apeluri pierdute prin distribuirea inteligentă a apelurilor, creșterea productivității personalului implicat în activități de vânzări și marketing, amplificarea satisfacției clientului prin creșterea valorii percepute de acesta, monitorizarea gradului de satisfacere a cerințelor clienților.
- c. **Componenta eCRM** vizează utilizarea tehnologiilor Internet în vederea perfecționării relațiilor cu clienții organizației, folosind instrumente specifice precum :
 - Listele de adrese poștale și serviciile de e-mail personalizate. Prin intermediul acestor aplicații sunt transmise mesaje standardizate (buletine informative, formulare electronice etc.) clienților existenți într-o bază de date.
 - Serviciile de chat sau dialoguri interactive. Reprezintă servicii WEB prin care un client poate solicita asistență on-line de la personal de specialitate din organizație.
 - Forumurile. Sunt discuții pe marginea unor teme de interes general pentru consumatori precum: noile produse sau servicii oferite de organizație, procedura care trebuie parcursă pentru obținerea unor documentații etc.

Pe plan mondial, cele mai utilizate pachete de aplicații eCRM sunt : Marketing Automation Systems (MAS), Sales Force Automation (SFA) și Customer Service Automation Systems (CSAS).

Ca principale avantaje ale implementării sistemelor CRM enumerăm: eficientizarea activităților de deservire a comenzilor primite de la consumatori, îmbunătățirea calității serviciilor prestate de organizație pentru clienții săi, comunicarea cu clientul la un nivel calitativ superior folosind canale de comunicare multiple (telefon, ghișeu unic, web, e-mail), diminuarea costurilor de comunicare cu clientul, reducerea

timpului de răspuns la problemele sesizate, realizarea unei imagini mai bune a organizației în relația cu clienții.

3.Utilizarea tehnicii DM in sistemele CRM

După cum s-a mai precizat, principalul obiectiv strategic al implementării CRM îl reprezintă maximizarea valorii percepute de client, în contextul înțelegerii comportamentului acestuia și promovării unor acțiuni pentru fidelizarea sa. În acest context, **fluxurile informaționale** modelate în cadrul sistemelor CRM includ, în general, trei faze:

- a. Clientul contactează compania, în scopul de a lansa o comandă pentru un produs, de a obține suport tehnic, de a face o reclamație sau o sugestie etc.
- b. Compania contactează clientul prin intermediul unei campanii de marketing, unui studiu de piață sau folosind comerțul electronic.
- c. Compania analizează informațiile obținute de la clienți în vederea unei mai bune înțelegeri a cerințelor reale ale acestora.

În practică, însă, mai ales în cazul companiilor mari, aplicarea tehnicilor specifice CRM presupune parcurgerea unei succesiuni de etape:

ETAPA 1. Identificarea clienților organizației și gruparea acestora pe categorii în funcție de preferințe sau comportamente. Principalele criterii de clasificare sunt :

- Specificul cerințelor : consumatori, agenți economici, instituții publice etc.
- Apartenența clienților la un anumit grup: clienți cu cerințe generale (nevoile comune ale unui grup de indivizi sau organizații), sau cu cerințe individuale (specifice fiecărui individ sau organizație).
- Gradul de stabilitate în timp a cerințelor: clienți cu cerințe uniforme sau neuniforme în timp.

Luând în considerare aceste criterii de clasificare, putem împărți clienții unei organizații în patru grupe:

- GRUPA I: Clienți cu cerințe generale și care au un caracter uniform.
- GRUPA II: Clienți cu cerințe specifice și care au un caracter uniform.
- GRUPA III: Clienți cu cerințe generale și care au un caracter neuniform.
- GRUPA IV: Clienți cu cerințe specifice și care au un caracter neuniform.

ETAPA 2. Stabilirea informațiilor necesare și proiectarea arhitecturii sistemului. În cadrul acestei etape are loc proiectarea bazei de date de gestiune a clienților care conține, în general, informații legate de: datele de identificare ale persoanei; pregătirea profesională; situația socială; apartenența la o anumită categorie de clienți; atitudini și percepții; comportamente în diferite situații; cereri, reclamații, sesizări depuse de client etc. Desigur că, de cele mai multe ori, legislația în vigoare nu permite stocarea și utilizarea tuturor categoriilor de date. Din acest motiv, în paralel cu alegerea informațiilor, este necesar să fie elaborată o **politică de confidențialitate** care să conțină reglementări precum: descrierea amănunțită a modului în care datele colectate vor fi

exploatate, definirea informațiilor care nu vor putea fi utilizate de către organizație sau definirea unor proceduri prin care clientul să poată preveni utilizarea anumitor informații.

ETAPA 3. Identificarea modalităților de culegere a informațiilor presupune dezvoltarea unui instrumentar de metode și tehnici prin care informațiile care descriu comportamentele clienților să fie colectate și introduse în baza de date. Preluarea datelor se poate realiza fie prin mijloace convenționale, fie automatizat, cu ajutorul tehnicii de calcul. Principalele metode utilizate pentru culegerea datelor includ : fax, telefon, e-mail, Internet etc.

ETAPA 4. Culegerea informațiilor și actualizarea bazei de date constă în utilizarea tehnicilor definite în cadrul celei de-a doua etape în vederea consolidării bazei de date de clienți. O metodă clasică de colectare a acestor informații o reprezintă interviul structurat, care poate fi derulat fie de operatori umani, fie de operatori electronici (sondajul electronic).

ETAPA 5. Operaționalizarea de modificări în plan organizațional în vederea amplificării satisfacției clienților prin îmbunătățirea și diversificarea serviciilor furnizate, acționând concomitent atât în planul acoperirii cerințelor generale cât și în planul celor individuale. Studiile efectuate relevă faptul că amplificarea gradului de satisfacere a cerințelor clienților generează îmbunătățirea imaginii prin care organizația este percepută pe piață, însă numai până la un anumit punct de maxim, dincolo de care imaginea sa începe să se deterioreze.

Aplicarea acestei metodologii presupune utilizarea unui instrumentar care să permită parcurgerea datelor care descriu comportamentul clienților în vederea descoperirii unor trenduri care să permită efectuarea de predicții privind evoluția fiecărui segment de piață. Tehnicile derivate din Data Mining pot fi utilizate cu succes, mai ales deoarece se pune problema unei abordări multidimensionale care poate, spre exemplu, să includă **trei dimensiuni**:

- ierarhia produselor (brand, clasă, categorie, produs);
- ierarhia perioadelor (ani, trimestre, luni, date calendaristice);
- ierarhia clienților (regiuni, zone clase clienți).

În practică, această abordare este operaționalizată cu succes prin intermediul unui concept modern care stă astăzi la baza majorității sistemelor de asistare a procesului decizional și anume cel de **OLAP (On-Line Analytic Processing)**, care au la bază **tehnica analizei multidimensionale a datelor**. Termenul de OLAP a fost introdus în 1993 de către Codd, împreună cu cele 18 reguli de funcționare specific acestei tehnologii. Codd făcea o diferențiere clară între tehnologia sistemelor multidimensionale și cea a sistemelor tranzacționale. De altfel, Codd a fost cel care a dat și definiția consacrată a termenului de OLAP: “OLAP reprezintă “o categorie de instrumente software, care permit analiștilor, managerilor și directorilor să înțeleagă esența datelor printr-un acces rapid, consistent și interactiv la o mare varietate de viziuni posibile ale informațiilor, care au fost obținute prin transformarea datelor primare astfel încât să reflecte dimensiunile reale ale întreprinderii așa cum o percepe și o înțelege utilizatorul”.

Dacă ne raportăm la cerințele utilizatorilor sistemelor CRM, sistemele OLAP asigură suportul pentru satisfacere în timp real a unor cereri specifice ale acestora,

deoarece anticipează momentul și conținutul interogării și oferă combinația optimă între rezultatele antecalulate și cele calculate la momentul solicitării informației. În acest scop, la nivelul sistemelor CRM, OLAP utilizează un **instrumentar specific**, motiv pentru care majoritatea specialiștilor consideră că reprezintă cel mai potrivit mediu pentru implementarea unor modele informaționale funcționale în baza principiilor dinamicii sistemelor:

- 1. Instrumente pentru importul datelor privind comportamentul clienților.** Importul presupune extragerea datelor din surse externe, de regulă din baze de date relaționale sau orientate obiect și se realizează prin intermediul unor interfețe standard. Procesul este consumator de timp pentru că presupune trecerea la modul de scriere a datelor în sistem multidimensional și este cu atât mai dificil cu cât sistemul de inteligență a afacerilor trebuie integrat cu sistemul informatic existent, dezvoltat pe o arhitectură clasică de tip SGBD.
- 2. Instrumente pentru stocarea datelor.** Stocarea presupune încărcarea datelor în depozitul de date, folosind o serie de aplicații utilitatea care presează o serie de servicii suplimentare, cum ar fi: verificarea integrității datelor, agregarea, sortarea și indexarea acestora etc.
- 3. Instrumente pentru integrarea datelor.** Desigur că trecerea la modelul multidimensional generează o serie de erori în procesul de extragere a datelor din bazele de date tradiționale, ceea ce implică utilizarea unor instrumente de tip *data mining* care să permită identificarea și eliminarea deficiențelor în integritatea datelor apărute odată cu importul acestora.
- 4. Instrumente pentru actualizarea datelor.** De regulă, în majoritatea aplicațiilor OLAP, depozitul de date este modificat zilnic, ceea ce generează o serie de probleme legate de momentul și modul în care se realizează actualizarea. Există o serie de instrumente care facilitează acest proces, care au la bază *tehnica încărcării incrementale* care presupune ca actualizarea să fie compusă dintr-o secvență de modificări cu durată redusă. Alte instrumente, cum ar fi *data shipping* sau *transaction shipping*, utilizează principiul replicării sau “fotografierii” datelor.
- 5. Instrumente pentru prelucrarea datelor, interogare și generarea rapoartelor.** Aceste instrumente utilizează conceptul de *metadata* utilizate în cadrul tehnicilor de data mining, definite ca fiind “*datele ce descriu semnificația datelor*”. Metadatale reprezintă elemente esențiale în procesul de interogare a bazei de date, atât prin prisma faptului că pot conține informații de natură administrativă (profiluri de utilizatori, drepturi de acces, locația datelor) cât și de natură funcțională (relații între diferitele tipuri de date, informații statistice etc.).
- 6. Instrumente pentru administrarea sistemului.** În această categorie intră instrumentele utilizate pentru gestiunea depozitelor de date, a capacității de stocare, a sistemului de comunicație etc.

Concluzionând, putem aprecia că, la momentul actual, aproape toate organizațiile mari dețin o platformă Intranet care, împreună cu anumite extensii și instrumente specifice, asigură funcționalitățile de bază ale aplicațiilor de Business Intelligence, cum ar fi organizarea informațiilor în depozite de date și prelucrarea acestora folosind tehnici de Data Mining. Mai mult decât atât, unele organizații au

dezvoltate platforme de tip Groupware care permit lucrul in echipă și organizarea eficientă a activităților colaborative. Numeroase funcții specifice Data Mining sunt deja implementate sub formă de componente ale arhitecturii Intranet sau a unor soluții specifice cum ar fi cele de tip CRM (Customer Relationship Management). În aceste condiții, o intensificare a utilizării funcțiilor DM existente și dezvoltarea de funcții noi s-ar putea face prin proiectarea unui sistem de management al cunoștințelor de tip KMS.

În ceea ce privește piața Business Intelligence, aceasta este una destul de confuză și dinamică ca urmare a faptului că, la momentul actual, nu există o soluție standard cu un nivel ridicat de interoperabilitate. Mai mult decât atât, achiziționarea unui pachet standard ar putea conduce la pierderea avantajului competitiv ca urmare a nivelului redus de satisfacere a cerințelor specifice. Din acest motiv, de regulă, organizațiile recurg la capacitatea de dezvoltare proprie. O altă problemă este cea a integrării soluțiilor DM implementate cu alte aplicații cum ar fi cele de management al documentelor și fluxurilor de lucru. În acest sens, o soluție tehnică la care recurg tot mai multe companii este cea a dezvoltării de portaluri de cunoștințe care pot facilita integrarea prin utilizarea funcțiilor inteligente în procesul de manipulare a cunoștințelor.

Bibliografie:

1. **Adelman S., Moss L.** - *Data Strategy*, London: Prentice Hall, 2005.
2. **Alavi M., Leidner D.** - *Knowledge Management and Knowledge Management Systems*, New York: MIS Quarterly, 2001.
3. **Avison D., Fitzgerald G.** - *Information Systems Development: Methodologies, Techniques and Tools*, London: McGraw Hill, 2002.
4. **Bazerman H.** - *Judgment in Managerial Decision Making*, New York: John Wiley and Sons, 2005.
5. **Date C.** - *Database Relational Model: A Retrospective Review and Analysis: A Historical Account and Assessment of E. F. Codd's Contribution to the Field of Database Technology*, New York: Addison Wesley, 2000.
6. **Doukidis G., Mylonopoulos N.A.** - *Social and Economic Transformation in the Digital Era*, London: Idea Group Publishing, 2004.
7. **Kamber M., Jiawei H.** - *Data Mining: Concepts and Techniques*, London: Morgan Kaufmann, 2005.
8. **Linoff G., Berry M.** - *Mastering Data Mining*, New York: John Wiley and Sons, 2000.
9. **Mora M., Forggione G.** - *Decision Making Support Systems*, New York: Idea Group Publishing, 2003.
10. **Shapiro P., Fayyad M.** - *Advantage in Knowledge Discovery and Data Mining*, Washington: AAI Press, 1996.