

Workshop lansare proiect KBS-Weld: Sistem bazat pe cunoștințe pentru structuri sudate și tehnologii

PROCESAREA COGNITIVĂ A DATELOR EXPERIMENTALE PE BAZA INFERENȚEI STATISTICE

Autor: Ing.fiz. Sandu CRĂȘTEȚI

Timisoara, 31 Mai 2018

Introducere

Inferența - componentă a unui proces logic care derivă o concluzie dintr-o premisă, adică extrage o consecință necesară, o informație specifică, dintr-o descriere de stare dată.

Inferența (deducția)

- *reprezintă componenta esențială a unui raționament*
- *analizează conținutul informațional inițial disponibil, formulat într-o propoziție sau lanț de propoziții*
- *cercetează conținutul fiecărei propoziții, eventualei relații interne sau cu alte propoziții*
- *caută dependențe între predicatele caracterizante de stare ale fiecărei propoziții*
- *face legături corelante (inferențiale) maxim probabile sau necesare, între propoziții sau predicatele unei propoziții*
- *obține sau derivă o informație suplimentară*

Introducere

Extragerea de cunoștințe din date (*data mining*) este un proces de analiză a unor cantități mari de date și de extragere informațiilor relevante din acestea folosind metode matematice și statistice.

Termenul este utilizat în cadrul organizațiilor ce se ocupă cu prelucrarea informațiilor despre companii și de către analiștii financiari, dar este folosit din ce în ce mai mult și în domeniul științific pentru extragerea informațiilor din volumuri mari de date, generate de exemplu de experimente moderne.

Data mining a fost descrisă ca "*extragerea netrivială a informațiilor implicite, anterior necunoscute și potențial utile din date*", precum și ca "*știința extragerii informațiilor utile din volume de date mari sau din baze de date*".

Data mining reprezintă, de asemenea, o modalitate de analiza statistică și logică a unor mari volume de date, în căutarea unor șabloane care pot ajuta procesul de luare a deciziilor.

Introducere

Procesul de identificare a cunoștințelor (elemente cognitive) din bazele de date (cognitive) poate fi definit și caracterizat de următoarele patru etape:

- *Selecția datelor*
- *Pre-procesarea datelor*
- *Transformarea datelor*
- *Extragerea informațiilor (Data mining)*
- *Interpretarea / evaluarea datelor*

Introducere

Într-o variantă simplificată a procesului:

- *pre-procesarea*
- *data mining*
- *interpretarea rezultatelor*
- *validarea rezultatelor*

Pre-procesarea

- Anterior utilizării algoritmilor de extragere a informațiilor este necesară identificarea mulțimii datelor semnificative.
- Mulțimea datelor semnificative trebuie să fie suficient de mare pentru a permite identificarea consistentă a elementelor (formelor, patterns) de interes.
- Pre-procesarea este o etapă esențială pentru analiza caracteristicile datelor anterior procesului de extragere a informațiilor.
- Pe parcursul pre-procesării are loc și procesul de “curățare” a datelor – eliminarea elementelor considerate “zgomote de fond” și a datelor incomplete.

Procesarea cognitivă a datelor

Elementele caracteristice ale procesării cognitive a datelor (procesul de extragere a informațiilor):

- *detectia anomaliilor (deviațiilor)*
- *identificarea datelor ce necesită investigații suplimentare*
- *identificarea relațiilor dintre date*
- *gruparea datelor în funcție de criteriile de interes*
- *clasificarea datelor*
- *identificarea unei funcții potențiale de modelare (pe baza minimizării erorilor) și estimare a relațiilor dintre elementele setului de date analizat*
- *generarea concluziilor*

Validarea rezultatelor

Pot apare situații în care rezultatele procesării cognitive a datelor pot fi, aparent, semnificative fără a fi utilizabile în procesul de predicție fenomenologică în cadrul analizei procesului de interes.

Aceasta se poate datora unei investigații prea extensive, ce are ca obiectiv testarea prea multor ipoteze simultan fără a se analiza totalitatea elementelor specifice testării statistice a ipotezelor.

Validarea rezultatelor

Etapa finală ce caracterizează procesul de identificare a elementelor cognitive din bazele de date este reprezentat de verificarea modului în care rezultatele generate de algoritmi *data mining* pot fi regăsite la o analiză a mulțimii extinse de date.

Nu toate formele detectate pe baza algoritmilor sunt în mod necesar valide. Pentru a se evita aceste situații este necesară testarea pe o mulțime de date diferite de cele pe baza cărora au fost elaborați (antrenați, învățați) algoritmi.

Validarea rezultatelor

Dacă eficiența procesului nu este satisfăcătoare, se impune o re-evaluare și modificare a modalităților de pre-procesare și identificare.

Dacă eficiența procesului este considerată satisfăcătoare, etapa finală este reprezentată de interpretarea rezultatelor și transferul acestora în baza de cunoștințe validate.

Sisteme inferențiale

Sistemele bazate pe paradigmele specifice inteligenței artificiale au fost primele elemente care au funcționat pe baza inferenței logice, generând aplicații industriale sub forma sistemelor expert - ulterior și sisteme funcționale pe baza unor reguli prestabilite și sisteme de demonstrare automată a teoremelor.

Sistemele inferențiale sunt destinate și extinderii bazelor de date cognitive, considerate ca o mulțime de propoziții destinate a descrie percepția sistemului privind domeniul de interes în care acesta funcționează.

Abordarea propusă în cadrul KBS-Weld

Obiectivul general propus: dezvoltarea, aplicarea și combinarea unor tehnici specifice paradigmelor inteligenței artificiale pentru implementarea și validarea, la nivel de laborator experimental, a unui sistem de procesare cognitivă destinat planificării și optimizării proceselor de sudare, avându-se în vedere complexitatea proceselor în contextul cerințelor utilizatorilor.

Formulare alternativă: generarea unui model pentru corelarea datelor de intrare specifice proceselor de sudare (materiale și parametri de sudare) cu elementele reprezentând datele de ieșire (proprietățile îmbinărilor sudate, caracterizate prin sudabilitate).

Abordarea propusă în cadrul KBS-Weld

Cazul 1: cunoscând-se calitatea, și caracteristicile mecanice și structurale ale materialelor de sudat, precum și alte condiții și cerințe impuse inițial, sistemul urmează a genera (sintetiza), ca date de ieșire, configurația optimă a elementelor specifice procesului de sudare, reprezentat printr-o specificație de sudare (WPS) optimă.

Cazul 2: luând în considerare caracteristicile materialelor de sudat, precum și procedurile aplicabile, sistemul va avea abilitatea de a estima (predictiv) proprietățile mecanice ale îmbinării sudate.

Cazul 3: sistemul va putea contribui la optimizarea procesului de sudare pe baza reducerii pierderilor datorate unor variații impredictibile ale parametrilor în cursul desfășurării procesului.

Abordarea propusă în cadrul KBS-Weld

Elemente inovative:

i. crearea unei baze de date extinse (Big Data) pentru accesabilă de către specialiștii în domeniul sudurii

ii. implementarea unor elemente inovative specifice procesării statistice și analizei de date – inclusiv în ceea ce privește reprezentarea și formalizarea elementelor cognitive - pe bazele paradigmelor inteligenței artificiale

iii. module hibride destinate modelării datelor de intrare și a celor de ieșire specifice procesului de sudare, pe baza unor elemente cognitive generate în cadrul programului experimental

iv. identificarea parametrilor optimi de sudare

Abordarea propusă în cadrul KBS-Weld

Beneficii estimate:

- *reducerea timpului de realizare a unor produse inovative*
- *optimizarea procesului de sudare și a proprietăților elementelor sudate*
- *generarea de elemente cognitive noi, pe baza utilizării elementelor deja existente*
- *automatizarea activităților de rutină, oferind mai mult timp pentru activitățile de implementare a soluțiilor inovative*

Vă mulțumesc pentru atenție!