



RECUNOAȘTEREA formelor

Radu-Daniel Vatavu

Serialul dedicat recunoșterii formelor se încheie cu prezentarea unor procedee de învățare nesupravegheată. Vor fi descrise diferite modalități de abordare cum ar fi metoda celor mai mici pătrate, metoda nucleelor dinamice, metode de clasificare ierarhice etc. De asemenea, vom realiza o prezentare comparativă a clasificatorilor descriși pe parcursul celor trei episoade ale serialului.

Învățarea nesupravegheată

Învățarea nesupravegheată are drept scop împărțirea unei mulțimi de forme într-un anumit număr de clase, în funcție de gradul de "asemănare" sau de similaritate dintre acestea. Spre deosebire de contextul supravegheat al recunoașterii formelor, caracterizat de prezența unui "profesor" sau "supraveghetor" astfel încât apartenența formelor din setul de învățare la clase este cunoscută, în cadrul contextului nesupravegheat singurele informații de care se dispune sunt legate doar de măsurătorile efectuate asupra formelor. Ca urmare, fie F mulțimea formelor de care se dispune și care trebuie clasificate într-un număr de m clase:

$$F = \{x_0, x_1, \dots, x_{p-1}\}.$$

Pentru început, vom considera numărul m de clase în care vor fi împărțite formele cunoscute.

Deci, problema constă în partiționarea mulțimii F în m clase cât mai omogene posibil.

$$F = \{F_0, F_1, \dots, F_{m-1}\}$$

$$F_i \cap F_j = \emptyset \quad \forall i, j \in \{0, 1, \dots, m-1\}$$

$$i \neq j, F = \bigcup_{j=0}^{m-1} F_j.$$

Metoda celor mai mici pătrate

O primă abordare de rezolvare a problemei are drept punct de plecare metoda celor mai mici pătrate. În acest context, gruparea formelor în cele m clase se va face urmărindu-se minimizarea unui criteriu E de tipul:

$$E = \sum_{k=0}^{m-1} E_k, E_k = \sum_{x \in F_k} d^2(x, m_k),$$

unde d este distanța euclidiană dintre forma x și forma medie m_k (sau centrul de greutate) al clasei k , de cardinal n_k :

$$d(x, m_k) = \|x - m_k\|, m_k = \frac{1}{n_k} \sum_{x \in F_k} x.$$

Fie o partiționare oarecare a mulțimii F în m clase și fie E valoarea criteriului aferent acestei partiționări. În continuare, să considerăm modificarea care se produce asupra acestei valori E în cazul în care se modifică clasificarea unei forme x de la clasa i la clasa j . Noua valoare E' aferentă noii partiționări va fi:

$$\begin{aligned} E' &= E - \left\| x - \frac{n_i \cdot m_i - x}{n_i - 1} \right\|^2 + \left\| x - \frac{n_j \cdot m_j + x}{n_j + 1} \right\|^2 = \\ &= E - \left(\left\| \frac{n_i}{n_i - 1} \cdot x - \frac{n_i}{n_i - 1} \cdot m_i \right\|^2 - \left\| \frac{n_j}{n_j + 1} \cdot x - \frac{n_j}{n_j + 1} \cdot m_j \right\|^2 \right) = \\ &= E - \left(\left(\frac{n_i}{n_i - 1} \right)^2 \cdot \|x - m_i\|^2 - \left(\frac{n_j}{n_j + 1} \right)^2 \cdot \|x - m_j\|^2 \right) = \\ &= E - \Delta E(x) \end{aligned}$$

Pentru a urmări minimizarea criteriului E , va trebui aleasă acea clasă j care să maximizeze $\Delta E(x)$ sau echivalent, să minimizeze cel de-al doilea termen al expresiei $\Delta E(x)$:

$$e_j = \left(\frac{n_j}{n_j + 1} \right)^2 \cdot \|x - m_j\|^2,$$

primul termen fiind constant pentru o aceeași valoare a lui x aparținând inițial clasei i .

Algoritmul bazat pe metoda celor mai mici pătrate este prezentat în continuare:

Se alege o partiționare inițială F_0, F_1, \dots, F_{m-1} a mulțimii F . Se calculează valoarea E și centrele de greutate m_k ale claselor, $k=0, \dots, m-1$.

repetă

clasificat $\leftarrow 1$

pentru $s \leftarrow 0, p-1$ **execută**

 fie F_i clasa de care aparține forma x_s

 fie $j \in \{0, 1, \dots, m-1\}$ pentru care e_j este minim



dacă $i \neq j$ **atunci**

$F_i \leftarrow F_i - \{x_s\}$

$F_j \leftarrow F_j + \{x_s\}$

se actualizează centrele de greutate m_k

clasificat $\leftarrow 0$

sfârșit dacă

sfârșit pentru

până când clasificat = 1

Rezultatul furnizat de algoritm depinde de ordinea de parcurs a mulțimii formelor F precum și de partiționarea inițială, având tendința de a se bloca într-o situație de minim local. Rezultate mai bune se pot obține cu o metodă des utilizată și foarte simplu de implementat, prezentată în continuare.

Metoda nucleelor dinamice

Metoda nucleelor dinamice (cunoscută și sub denumirile de metoda celor k medii sau metoda **ISODATA**) constă în clasificarea unei forme x la cea clasă k pentru care distanța dintre forma x și centrul de greutate m_k al clasei este minimă. Procedul este repetat până când structura claselor rămâne neschimbată.

Algoritmul este prezentat în continuare:

Se aleg m forme din mulțimea F drept centrele de greutate m_k ale claselor, $k=0, \dots, m-1$.

repetă

clasificat $\leftarrow 1$

pentru $s \leftarrow 0, p-1$ **execută**

fie F_i clasa de care aparține x_s

fie $j \in \{0, 1, \dots, m-1\}$ pentru care $d(x_s, m_j)$ este minimă

dacă $i \neq j$ **atunci**

$F_i \leftarrow F_i - \{x_s\}$

$F_j \leftarrow F_j + \{x_s\}$

clasificat $\leftarrow 0$

sfârșit dacă

sfârșit pentru

se actualizează centrele de greutate m_k

până când clasificat = 1

Metoda prezintă la rândul ei anumite dezavantaje, rezultatele depinzând de:

- alegerea inițială a celor m centre de greutate. Pentru a înălțura această problemă, algoritmul poate fi reluat cu seturi diferite de puncte de plecare;
- distanța d folosită (a se vedea discuțiile de la distanțe privind normarea caracteristicilor cu abaterea medie pătratică);
- numărul m al claselor (pentru m egal cu numărul p al formelor se obține regula NN).

O variantă a acestei metode este reprezentată de actualizarea secvențială a centrelor de greutate ale claselor. Modificarea este preferabilă atunci când nu se dispune inițial

de toate formele mulțimii F , acestea fiind obținute pe parcurs.

Algoritmul este prezentat în continuare:

Se aleg primele m forme drept centrele de greutate m_k ale claselor, $k=0, \dots, m-1$.

$n_k \leftarrow 1, k=0, \dots, m-1$

repetă

fie x noua formă ce trebuie clasificată

fie $j \in \{0, 1, \dots, m-1\}$ pentru care $d(x, m_j)$ este minimă

$F_j \leftarrow F_j + \{x\}$

$m_j \leftarrow (n_j \cdot m_j + x) / (n_j + 1)$

$n_j \leftarrow n_j + 1$

până când <condiție_oprire>

Clasele produse de metodele prezentate mai sus alcătuiesc o așa-numită partiție "grea" a mulțimii F a formelor, în sensul că fiecare formă x_i din F aparține sau nu unei anumite clase j .

În contrast cu partițiile "grele" se află partițiile fuzzy caracterizate prin faptul că o formă x_i poate avea grade de apartenență diferite la clase diferite. În continuare, se va nota prin $u_{i,j}$ gradul de apartenență al unei forme x_i la clasa j considerat a fi de forma:

$$u_{i,j} = \frac{\exp(-\|x_i - m_j\|^2)}{\sum_{k=0}^{m-1} \exp(-\|x_i - m_k\|^2)}$$

Se observă că pentru orice formă x_i este îndeplinită relația:

$$\sum_{j=0}^{m-1} u_{i,j} = 1.$$

Acest model este mai apropiat de modul de funcționare al clasificatorului uman, existând situații când și acesta ia decizii de tipul: o formă aparține unei clase într-o proporție de 70% și altei clase în proporție de 30%.

Cu considerațiile de mai sus, centrele de greutate ale claselor se vor calcula cu ajutorul mediei ponderate, astfel:

$$m_k = \frac{\sum_{x_i \in F_k} u_{i,k} \cdot x_i}{n_k}.$$

Algoritmul care implementează abordarea fuzzy a nucleelor dinamice este identic cu cele prezentate mai sus.

Metode ierarhice de clasificare

Metodele ierarhice de clasificare au avut drept punct de plecare taxonomia biologică, propunând o serie de grupări (divizări) între clase pe baza unor similarități (disimilarități) inter-clasă. Aceste metode pot fi clasificate în două categorii:

- metode descendente (divizive) ce au drept punct de plecare o singură clasă formată din toate formele mulțimii F care este divizată recursiv;
- metode ascendente (de aglomerare) care procedează la gruparea succesivă a claselor, având inițial un număr de clase egal cu cardinalul mulțimii F .



Un algoritm aparținând celei de-a doua categorii de metode este prezentat în continuare. Numărul de clase este reprezentat de variabila m . Prin d_{ij} s-a notat distanța inter-clasă dintre clasele i și j .

Se formează clasele $F_k = \{x_k\}$, $k=0, \dots, p-1$

Se calculează distanțele d_{ij}

$m \leftarrow p$

repetă

 fie $i, j \in \{0, 1, \dots, m-1\}$ pentru care d_{ij} este minimă

$F_i \leftarrow F_i \cup F_j$

$m \leftarrow m-1$

 Se recalculează distanțele d_{ij}

până când <condiție_oprire>

Condiția de oprire poate fi determinată de: atingerea unei anumite valori m pentru numărul de clase, situarea tuturor distanțelor d_{ij} sub o anumită valoare prag etc.

Distanțele inter-clasă d_{ij} pot fi calculate folosind una dintre următoarele formule:

$$d_{i,j} = \max \{d(x, y) / x \in F_i, y \in F_j\}$$

$$d_{i,j} = \min \{d(x, y) / x \in F_i, y \in F_j\}$$

$$d_{i,j} = d(m_i, m_j)$$

$$d_{i,j} = \frac{n_i \cdot n_j}{n_i + n_j} \cdot d(m_i, m_j)$$

Semnificația variabilelor care apar în aceste formule este următoarea:

- d reprezintă o distanță oarecare (de exemplu, distanța euclidiană);
- m_i și m_j reprezintă centrele de greutate ale celor două clase i și j ;
- n_i și n_j reprezintă numărul de elemente ale celor două clase i și j .

Se poate observa similitudinea algoritmului prezentat anterior cu algoritmul de determinare a arborelui de cost minim al unui graf având drept noduri cele p clase inițiale și drept muchii distanțele inter-clasă. La fiecare pas, se adaugă un nou nod în arborele de cost minim.

Plecând de la această similitudine, să considerăm în continuare un graf neorientat (numit **graful formelor**) având un număr p de noduri corespunzătoare celor p forme, iar matricea de adiacență $A_{p \times p} = (a_{ij})$ definită astfel:

$$a_{i,j} = \begin{cases} 1, & \text{pentru } d(x_i, x_j) \leq D \\ 0, & \text{altfel} \end{cases},$$

unde d este o distanță oarecare, iar D este o valoare de prag.

Cu aceste considerente, rezultă un nou algoritm de clasificare a formelor mulțimii F , bazat pe determinarea componentelor conexe ale grafului construit cu specificațiile de mai sus.

Procesul de împărțire al formelor în clase este controlat prin valoarea de prag D .

Analiza comparativă a clasificatorilor

Evident, modelul de clasificator va fi ales în funcție de tipul aplicației, numărul și tipul caracteristicilor, mărimea setului de învățare, existența unor informații prealabile privind distribuția formelor etc.

În continuare se vor prezenta performanțele obținute cu o serie de clasificatori pentru o aplicație de recunoaștere a formelor. Trebuie precizat faptul că discuțiile ce vor urma au drept scop doar punerea în evidență a metodelor practice de construire a unor clasificatori și nu au pretenția de a rezolva problema specifică avută în vedere.

Având considerațiile făcute, fie următoarea aplicație de recunoaștere a formelor având drept obiectiv: recunoașterea corpusculilor Danne din imagini de microscopie electronică a serului bolnavilor de *hepatită virală B*. O astfel de imagine este prezentată în figura 1.

Imaginea a fost preluată din [7] (pag. 298). *Virusul hepatic B* este considerat a fi reprezentat de particula virală numită *corpuscul Danne*. După cum se observă și din imagine, în serul bolnavilor de *hepatită virală B* s-au pus în evidență în microscopie electronică trei forme ale *antigenelor virale B*: particule sferice mici cu un diametru de 22 nm, forme tubulare de grosime 22 nm și lungime 20-250 nm și *corpusculi Danne* de formă circulară cu un diametru de 42 nm.



Figura 1

Prin utilizarea unor tehnici specifice procesării imaginilor (filtrare, segmentare etc.) se obține o imagine prelucrată, binară, prezentată în figura 2, care permite identificarea celor trei tipuri de particule cu ajutorul unui algoritm de determinare a componentelor conexe.

Pentru caracterizarea celor trei tipuri de forme se vor folosi două caracteristici: circularitatea și aria.

Circularitatea este definită ca fiind proporțională cu raportul dintre aria și pătratul perimetrului unei figuri geometrice: $C = \frac{4 \cdot \pi \cdot A}{P^2}$, unde A reprezintă aria figurii, iar P

perimetrul. **Circularitatea** are valoarea 1 pentru un disc și o valoare subunitară pentru orice altă figură geometrică. Ea va distinge dintre cele două forme circulare și formele tubulare, iar aria va realiza discriminarea dintre cele două forme circulare (de diametre 22 și respectiv 42 nm).

În figura 3 sunt reprezentate formele în spațiul celor două caracteristici (circularitatea și aria).

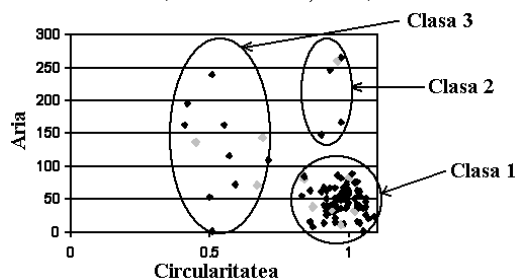


Figura 3



Din această reprezentare se observă gruparea relativă a formelor din cele trei ($m = 3$) clase și anume: clasa 1 - corpusculii circulari mici 22 nm (arie mică și circularitatea ≈ 1), clasa 2 - corpusculii Danne 42 nm (arie mare și circularitate ≈ 1), clasa 3 - corpusculii de formă tubulară (aria variază datorită lungimii lor între 20-250 nm, iar circularitatea ≈ 0.5).

Pentru contextul învățării supravegheate, cele p ($p = 110$) forme au fost împărțite în două seturi: setul de învățare (100 de puncte - punctele negre din planul caracteris-

ticilor) și setul de test (10 puncte - punctele gri). În tabelul 1 sunt prezentate avantajele și dezavantajele folosirii câtorva clasificatori aparținând contextului supravegheat al recunoașterii formelor.

În figura 4 sunt prezentate rezultatele utilizării clasificatorilor bazați pe NN , KNN ($k = 3$) și funcții discriminant pentru câteva elemente aparținând mulțimii de test.

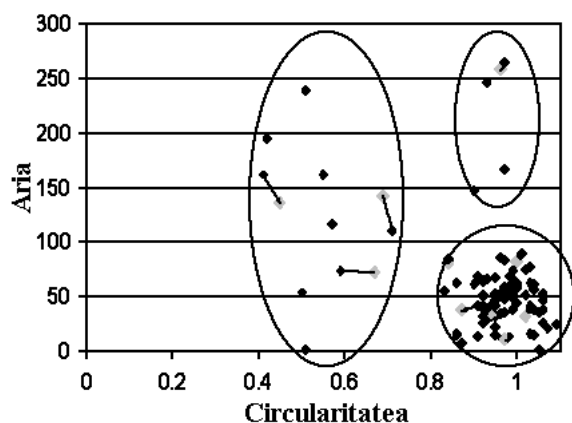
În cele ce urmează sunt prezentate comparativ două metode din cadrul contextului nesupravegheat al recunoașterii formelor:

Clasificator	Complexitate învățare	clasificare	Avantaje	Dezavantaje
<i>Nearest Neighbour</i> (NN)	n/a^*	$O(np)$	<ul style="list-style-type: none"> Implementare ușoară, rapidă. Nu necesită învățare și nici alte analize efectuate asupra setului de învățare. 	<ul style="list-style-type: none"> Este necesară memorarea setului de învățare. Timp mare de calcul (este parcurs întreg setul de învățare pentru clasificarea unei singure forme). Sensibilitatea la zgomot. Sensibilitatea la distanța folosită.
<i>K Nearest Neighbours</i> (KNN)	n/a^*	$O(k \times n \times p)$	<ul style="list-style-type: none"> Implementare rapidă. Performanțe mai bune decât NN fiind mai puțin sensibil la zgomot. 	<ul style="list-style-type: none"> Este necesară memorarea întregului set de învățare. Timp de clasificare mult prea mare.
<i>Funcții discriminant liniare</i>	$O(n \times p)$	$O(m \times n)$	<ul style="list-style-type: none"> Nu este necesară memorarea setului de învățare. Timp scăzut de clasificare comparativ cu celelalte două metode. 	<ul style="list-style-type: none"> Pot fi aplicate doar în cazul în care clasele sunt liniar separabile. Necesită memorarea a $n \times (m + 1)$ coeficienți.
<i>Clasificatori Bayes</i>	$O(p)$	$O(m \times n)$	<ul style="list-style-type: none"> Minimizează rata medie a erorii. 	<ul style="list-style-type: none"> Sunt necesare informații privind $Pr\{x j\}$ și $Pr\{j\}$.
<i>Perceptronul multistrat</i> (3 straturi)	variabilă**	$O(h \times (n + m))$	<ul style="list-style-type: none"> Performanțe superioare comparativ cu celelalte metode. Posibilități foarte bune de generalizare. Suprafețe de separare complexe. 	<ul style="list-style-type: none"> Determinarea numărului h de neuroni din stratul ascuns se face experimental. Performanțele depind fundamental de setul de învățare. Există posibilitatea ca rețeaua să se oprească într-un optim local. Există posibilitatea ca rețeaua să învețe foarte bine ("pe de rost") formele din mulțimea de învățare ceea ce conduce la o slabă generalizare.

* Nu se aplică în acest caz.

** Depinde de: mărimea p a setului de învățare, de rata de eroare ce se dorește a fi obținută, etc.

Tabelul 1: Analiza comparativă a clasificatorilor în învățarea supravegheată



a) regula NN

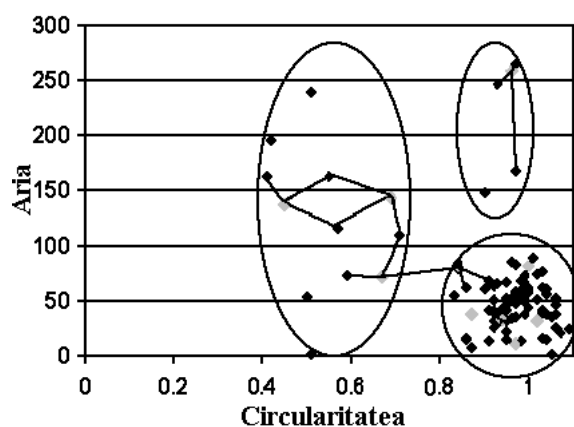
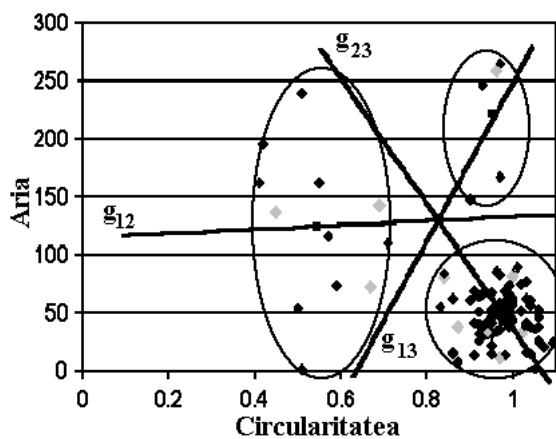
b) regula KNN ($k = 3$)c) funcții discriminant. Clase separate de suprafețele de separare g_{12} , g_{13} și g_{23} .

Figura 4

• metoda nucleelor dinamice

Complexitatea algoritmului de clasificare depinde de concentrarea formelor din setul de învățare și numărul m de clase. Are ca avantaj minimizarea inerției intra-clasă, iar ca dezavantaj faptul că rezultatele depind de valorile inițiale, distanța utilizată și numărul m al claselor.

• metoda grafului formelor

Complexitatea algoritmului de clasificare este $O(p^2 \cdot n)$. Printre avantajele acestei metode se află faptul că permite alegerea valorii de prag D și clasele obținute sunt omogene.

Timpul de calcul mare pentru un număr mare p de forme aparținând setului de învățare constituie unul dintre principalele dezavantaje ale utilizării metodei grafului formelor.

În figura 5 este ilustrată aplicarea metodei nucleelor dinamice prin conturarea celor trei clase. Formele sunt grupate în jurul centrelor de greutate.

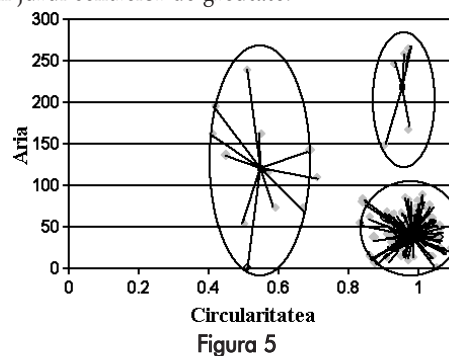


Figura 5

Concluzii

Acest serial s-a dorit a fi o introducere în recunoașterea statistică a formelor, prezentând câteva dintre tehnicile și metodele utilizate în cadrul celor două tipuri de abordări: învățarea supravegheată și nesupravegheată.

Cu toate că exemplele abordate au avut în vedere aspecte specifice procesării imaginilor, recunoașterea formelor include un domeniu mult mai vast al aplicațiilor permițând, după cum s-a prezentat, clasificarea și distingerea dintre obiecte și fenomene supuse unui același model matematic abstractizat.

Bibliografie

1. Duda R., Hart P., *Pattern Classification and Scene Analysis*, John Wiley & Sons, 1973
2. Belaid A., Belaid Y., *Reconnaissance des formes: Methodes et applications*, InterEditions, 1992
3. Milgram M., *Reconnaissance des formes: Methodes numeriques et connexionnistes*, Armand Colin, 1993
4. Pentiu Șt. Gh., *Aplicații ale recunoașterii formelor în diagnosticul automat*, Ed. Tehnică, 1997
5. Sonka M., Hlavac V., Boyle R., *Image Processing, Analysis and Machine Vision, 2nd Edition*, PWS Publishing, 1998
6. Dumitrescu D., Hariton Costin, *Rețele neuronale: Teorie și aplicații*, Teora, 1996
7. Georgescu L., Tudose N., Potencz E., *Morfopatologie*, Editura Didactică și Pedagogică, București, 1982

Radu-Daniel Vatavu este student în anul IV la Universitatea Ștefan cel Mare din Suceava. Poate fi contactat prin e-mail la raduvro@yahoo.com.