

UTILIZAREA ALGORITMILOR FUZZY LA PRELUCRAREA SERIILOR NUMERICE

Sergiu CHILAT

lector universitar, magistru

Universitatea Tehnică a Moldovei

***Abstract:** This paper presents a new method for making forecasting process, it is used a neural network and fuzzy controller. At the first stage, the neural network based forecasting performs internal parameters, then fuzzy controller adapts to changes in external parameters results. The main problem of neural networks in forecasting is unable to provide external parameters change, this problem was solved by fuzzy controller reacts to any change in external parameters and adjusts the data generated by the neural network forecasting.*

***Index terms:** Neural network, fuzzy logic, fuzzy controller, rules base, time series, diagnostication.*

1. Introducere

A prognoza înseamnă a prevedea modul în care se va desfășura în timp un fenomen, un proces etc., bazându-ne pe studiul împrejurărilor care îi determină apariția și evoluția. A prevedea valorile viitoare ale unor mărimi din prezent este o provocare răspândită și foarte utilă pentru diferite domenii ale activității umane, cum sunt medicina, tehnica, economia etc.

Complexitatea procesului de prognozare a secvențelor discrete se datorează faptului că, spre deosebire de procedurile de interpolare care sunt foarte bine algoritmizate, prognozarea necesită o extrapolare a datelor din trecut în viitor. În plus, trebuie să se țină cont de legitățile necunoscute care influențează secvența discretă. Există foarte multe studii dedicate elaborării modelului matematic al prognozei. Cele mai multe sunt bazate pe instrumente probabilistice și statistice. Însă, utilizarea lor, necesită o cantitate

considerabilă de date experimentale, acumularea cărora nu este posibilă în unele cazuri.

În ultimii ani, se observă o creștere a interesului față de rețelele neuronale la rezolvarea problemelor de prognozare. Acestea sunt privite ca un model maxim apropiat de creierul uman, fiind învățate să identifice legități necunoscute anterior.

În literatura de specialitate metodele de prognozare sunt separate în 4 grupe:

- bazate pe date concrete, reprezentate ca șiruri temporare;
- bazate pe informație euristică, obținută de la specialiștii cu calificare înaltă din domeniu;
- bazate pe analogii matematice, biologice sau istorice (Hassoun 1995);
- metode complexe, bazate pe combinarea diferitor metode și medii de realizare (Srinivasan 1998).

În această lucrare va fi descris un model de prognozare a vânzărilor, creat în baza unei rețele neuronale și a unui controler fuzzy, conectate în serie. La prima etapă, rețeaua neuronală va genera prognoza în baza datelor despre vânzările anterioare (parametri interni). La o a doua etapă, datele generate de rețeaua neuronală vor fi adaptate automat la influența factorilor externi a prognozei generate de rețeaua neuronală.

Astfel, procesul de prognozare devine unul continuu și adaptiv, iar rezultatele sunt modelate astfel ca să fie obținute rezultatele optime, fiind luată în considerație orice modificare a factorilor externi.

2. Descrierea modelului de prognozare

Stabilirea obiectivelor este una din sarcinile de bază ale prognozării. Selectarea obiectivelor se face în rezultatul analizei problemei, pentru care vor fi utilizate rezultatele prognozării. După selectare are loc crearea

arborelui obiectivelor și clasificarea pe baza coeficienților de prioritate (Fig. 1).

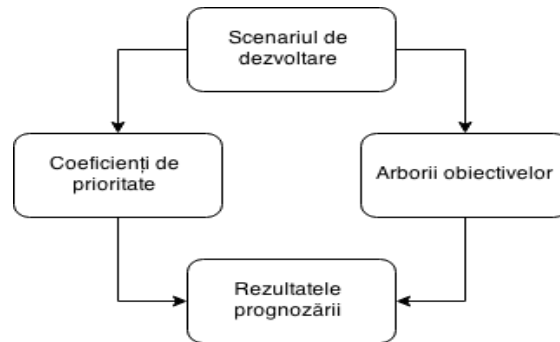


Fig. 1. Obținerea rezultatelor prognozărilor

Va fi cercetată metoda de creare a sistemelor de prognozare bazate pe tehnologiile intelectuale actuale: teoria mulțimilor fuzzy, rețele neuronale, metode ale logicii fuzzy și algoritmi genetici.

Cele mai importante probleme în crearea modelelor de prognozare sunt:

- între factorii externi și interni există legături neclare (fuzzy);
- lipsa stabilității statistice atât a parametrilor de intrare cât și a celor de ieșire.

Astfel, utilizarea metodelor clasice, cum ar fi, extrapolarea și analiza regresivă nu este practică, fiindcă dezavantajul lor principal este necesitatea prezenței unui volum mare de date statistice, pentru un număr mare de parametri și pentru mai multe perioade de timp (Morariu 2009:216).

Mulțimile fuzzy oferă posibilitatea formalizării valorilor, determinării relațiilor *cauză-efect* între parametri și factorii care îi influențează și formularea unei prognoze în condiții de incertitudine.

În fig. 2 este prezentată diagrama modelului prognozărilor, care este bazată pe metodele logicii fuzzy.

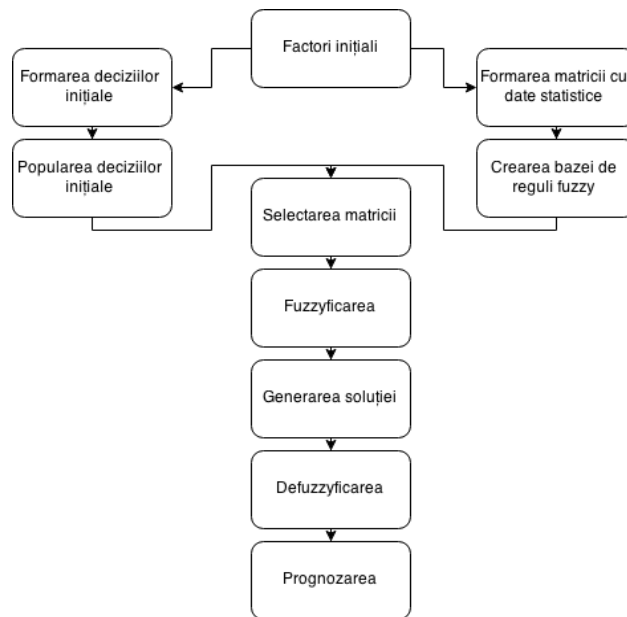


Fig. 2. Diagrama modelului prognozării

Algoritmul de creare a modelului matematic al prognozării constă în:

- determinarea parametrilor de intrare și ieșire;
- în baza sistemului expert se formează mulțimile fuzzy și se selectează unitatea de măsură;
- se definesc funcțiile de apartenență la mulțime;
- se formează baza de reguli;
- se selectează algoritmul pentru luarea deciziei.

La crearea modelului matematic, obiectul este cercetat în formă de funcție:

$$P = FuzzyPrognoza(t_1, t_2, \dots, t_n) \quad \square \square \square$$

unde n – numărul de termeni (parametri de intrare), P – soluția găsită.

Astfel, algoritmul de lucru al funcției *FuzzyPrognoza* este următorul:

1. colectarea setului de parametri de intrare care pot influența asupra soluției obținute;
2. generarea variabilelor lingvistice pentru fiecare din parametrii de intrare și a corespondențelor respective;

3. crearea unui graf în care este reflectată clasificarea parametrilor. În acest graf sunt incluse punctele de vedere ale experților din domeniu și căile prin care se poate ajunge la rezolvarea problemei. În dependență de situație, graful poate conține pînă la cîteva mii de noduri și poate avea mai multe variante de rezolvare. Metoda propusă va cerceta toate căile posibile și o va alege pe cea optimă;
4. colectarea datelor statistice;
5. fuzzyficarea, adică transformarea valorilor exacte ale parametrilor de intrare în variabile lingvistice fuzzy;
6. funcțiile de apartenență a variabilelor lingvistice se dau în forma (2)

$$fA(X) = \frac{1}{1 + \sum_{i=1}^n \left[\frac{x - P_i}{P_i} \right]^2} \quad (2)$$

unde P_i sunt parametrii de configurare;

7. obținerea rezultatelor modelării; se realizează prin transformarea mulțimii fuzzy în valori exacte, la etapa de defuzzyficare.

Configurarea modelului fuzzy se realizează prin completarea bazei de reguli și configurarea funcției de apartenență(2) și are forma prezentată în fig. 3.

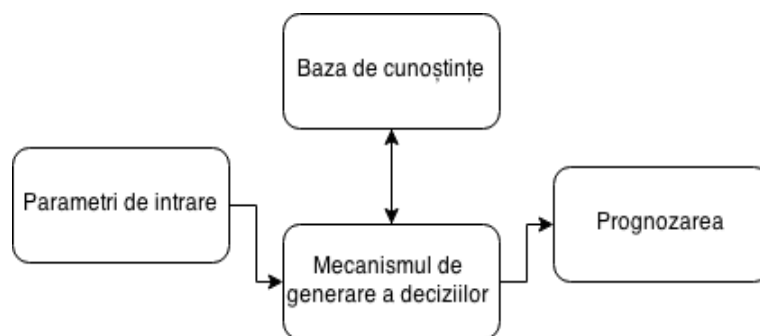


Fig. 3. Modelul fuzzy de luare a deciziilor

Baza de reguli este formată din date primite de la experți, în formatul *cauză-efect*(3).

$$S : \left\{ \begin{array}{l} S_1, \text{ dac\u0103 } A_1 \text{ atunci } B_1 \\ S_2, \text{ dac\u0103 } A_2 \text{ atunci } B_2 \\ \dots \\ S_n, \text{ dac\u0103 } A_n \text{ atunci } B_n \end{array} \right\} \quad (3)$$

Generarea matricei (3) are loc folosind abord\u0103ri din categoria algoritmilor genetici: \u00een rezultatul \u00endeplinirii unui num\u0103r finit de itera\u021bii ale algoritmului genetic se alege solu\u021bia (cromozomul) cu nivelul maxim de calitate, \u00een baza criteriilor enun\u021bate la \u00enceput.

Fiecare itera\u021bie a algoritmului genetic, const\u0103 din c\u00eteva opera\u021bii consecutive:

- formarea setului de cromozomi;
- crearea cromozomilor urma\u021bi, ca urmare a combin\u0103rii cromozomilor de baz\u0103, cu anumite devieri (muta\u021bii);
- calcularea nivelului de calitate pentru cromozomul nou creat (Mahfoud 1996:548).

Pentru determinarea subiectului prognoz\u0103rii este necesar s\u0103 se precizeze variabilele care sunt analizate \u0219i prognozate. Este foarte important s\u0103 fie cunoscut nivelul de detaliere, care este influen\u021bat de mai mul\u021bi factori: disponibilitatea \u0219i exactitatea datelor, costurile analizei preferin\u021belor utilizatorilor sistemului. \u00een situa\u021biile \u00een care cea mai bun\u0103 combina\u021bie de variabile nu este clar\u0103, pot fi \u00incercate diferite alternative \u0219i se va selecta cea care ofer\u0103 cele mai bune rezultate.

Al doilea pas important \u00een construirea unui sistem predictiv bazat pe re\u021ele neuronale este definirea urm\u0103torilor trei parametri: perioada prognoz\u0103rii, intervalul prognoz\u0103rii \u0219i orizontul prognoz\u0103rii.

Perioada prognoz\u0103rii este unitatea de baz\u0103 de timp pentru care se face prognoza.

Orizontul prognoz\u0103rii reprezint\u0103 num\u0103rul de perioade din viitor pentru care se face prognozarea. De exemplu, poate fi necesar\u0103 o prognoz\u0103 pentru o perioad\u0103 de o s\u0103pt\u0103m\u00e2n\u0103, cu prezentarea datelor pe fiecare zi \u00een parte. \u00een acest caz, perioada este de 1 zi, iar la orizontul - 7 zile.

Intervalul prognozării - frecvența cu care se face o nouă prognoză.

Selectarea corectă a perioadei și orizontului prognozării este una din cele mai dificile etape ale prognozării bazate pe rețele neuronale (Basaran 2007:67).

Exactitatea prognozei necesară pentru o problemă concretă are un impact substanțial asupra sistemului de predicție. De asemenea, un impact enorm asupra exactității prognozei îl are setul de date, care este ales pentru învățarea rețelei.

3. Formularea problemei

Va fi cercetat un exemplu concret: prognozarea volumului vânzărilor unei întreprinderi de distribuție a produselor alimentare. Mediul este non-deterministic, deoarece metodele convenționale nu permit de a determina cu certitudine ce se va întâmpla în următorul moment de timp și nu se pot determina toți factorii care ar putea influența acest lucru.

Sunt cunoscuți următorii indicatori financiari:

Activitatea întreprinderii

- istoria vânzărilor (cantitatea, sumele);
- istoria operațiilor de la depozit;
- indicatorii activității publicitare.

Factorii externi

- prețurile concurenților;
- starea pieței;
- inflația;
- cursul valutar;
- indicii bursieri.

Există și alți factori secundari (tabela 1) care, de asemenea, pot influența rezultatele vânzărilor. Acești parametri au o semnificație diferită, valorile lor sunt diferite ca natură și provin din diferite surse.

TABELA I. FACTORII SECUNDARI

Parametrul	Importanța
P1 = Produse la depozit	I(P1) = 30
P2 = Intensitatea publicității	I(P2) = 20
P3 = Prețul materiei prime	I(P3) = 60
P4 = Amplasarea geografică	I(P4) = 10
P5 = Inflația	I(P5) = 40
P6 = Prețurile concurenților	I(P6) = 30
P7 = Importul	I(P7) = 10
P8 = Cota de piață	I(P8) = 30
P9 = Concentrația cumpărătorilor	I(P9) = 10
P10 = Cursul valutar	I(P10) = 20

Ca urmare a analizei, se poate observa că, unii parametri nu pot fi incluși în model din cauza imposibilității de a obține date veridice, iar unii nu au o influență puternică asupra dinamicii modelului și, prin urmare, pot fi excluși din model fără o pierdere semnificativă a exactității.

Trebuie de menționat faptul că, doar istoria vânzărilor poate oferi 50-60% din volumul necesar de informație pentru învățarea rețelei neuronale (Charytoniuk 1995).

Problema prognozării vânzărilor companiei are caracteristici care fac utilizarea rețelelor neuronale eficientă:

- volumul de date poate fi relativ mic (datele despre vânzările produselor noi);
- în baza de date pot fi ”goluri”;
- datele pot fi distorsionate;

- este necesară o adaptare a modelului la apariția datelor noi;
- este complicat de obținut un model algebric liniar;
- numărul de produse este mare.

4. Descrierea algoritmului de prognozare

În tabela 2 sunt prezentate date despre istoria vânzărilor (mii unități), pentru perioada 01.01.2013 – 31.12.2014, grupate pe luni. Aceste date vor fi utilizate ca parametri de intrare a rețelei neuronale.

TABELA II. ISTORIA VÎNZĂRILOR

Perioada	Vânzări	Perioada	Vânzări
01.2013	2,0	01.2014	2,3
02.2013	2,5	02.2014	3,1
03.2013	2,5	03.2014	3,2
04.2013	2,7	04.2014	3,5
05.2013	3,2	05.2014	4
06.2013	3,7	06.2014	4,3
07.2013	3,8	07.2014	3,5
08.2013	3,9	08.2014	4,3
09.2013	2,8	09.2014	3
10.2013	2,4	10.2014	2,5
11.2013	3,8	11.2014	4
12.2013	4,5	12.2014	4,8

După învățarea rețelei neuronale, se obțin următoarele date după pentru eșantionul de verificare (fig. 4) selectat pentru perioada 01.2014 – 12.2014.

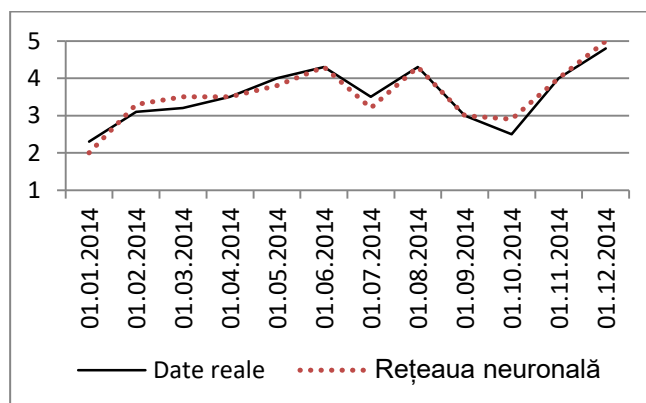


Fig. 4. Eșantionul de verificare

Eroarea maximă a rețelei este de 0,3 unități sau 6%, iar eroarea medie este de 0,04 unități sau 0,8%, deci poate fi utilizată în prognozare.

Următorul pas este generarea prognozei pentru primele 6 luni ale anului 2015. Utilizând rețeaua neuronală, se obțin următoarele rezultate(tabela 3):

TABELA III. REZULTATELE PROGNOZĂRII

Perioada	Vânzări
01.2015	2,35
02.2015	3,0
03.2015	3,1
04.2015	3,7
05.2015	4,2
06.2015	4,7

5. Adaptarea rezultatelor la factorii externi

Datele obținute sunt generate doar în baza parametrilor interni. Însă, aceste rezultate pot fi influențate și de factori externi(tabela 1), care pot să se modifice aleatoriu, fiind imposibil de prevăzut.

Astfel, apare o problemă, datele generate de rețeaua neuronală pot fi utilizate doar în cazul în care factorii externi au un efect nul sau sunt ignorați. Însă, în realitate acești parametri nu pot fi ignorați, iar în unele cazuri pot

modifică esențial dinamica vânzărilor, deci, datele prezentate în tabelul 3 își pot pierde actualitatea.

Pentru a rezolva această problemă, se propune utilizarea unui controler fuzzy, care la intrare va primi datele generate de rețeaua neuronală (tabela 3) și valorile pentru parametrii externi (tabela 1) și va realiza o reglare automată a datelor de intrare în fiecare zi (Garch 2005).

La prima etapă, va fi compusă baza de reguli a controlerului (utilizând tabela 1):

Pentru parametrul P1 - Produse la depozit

DACĂ produse la depozit = puține ATUNCI vânzările scad cu $I(P1)$ procente

DACĂ produse la depozit = destule SAU produse la depozit = multe ATUNCI vânzările se mențin la nivelul prognozat

Pentru parametrul P2 - Intensitatea publicității

DACĂ publicitate = nu există ATUNCI vânzările scad cu $I(P2)$ procente

DACĂ publicitate = puțină ATUNCI vânzările scad cu $I(P2)/2$ procente

DACĂ publicitate = multă ATUNCI vânzările cresc cu $I(P2)/2$ procente

Pentru parametrul P3 - Prețul materiei prime

DACĂ prețul materie prime = mare ATUNCI vânzările scad cu $I(P3)$ procente

DACĂ prețul materiei prime = mic ATUNCI vânzările cresc cu $I(P3)/2$ procente

Pentru parametrul P5 - Inflația

DACĂ inflația = mare ATUNCI vânzările scad cu $I(P5)$ procente

DACĂ inflația = mică ATUNCI vânzările cresc cu $I(P5)/2$ procente

Pentru parametrul P6 - Prețurile concurenților

DACĂ prețurile concurenților = mari ATUNCI vânzările cresc cu I(P6)/3 procente

DACĂ prețurile concurenților = mici ATUNCI vânzările scad cu I(P3) procente

Pentru parametrul P8 - Cota de piață

DACĂ cota de piață = ridicată ATUNCI vânzările cresc cu I(P8)/4 procente

DACĂ cota de piață = scăzută ATUNCI vânzările scad cu I(P8) procente

Pentru parametrul P10 - Cursul valutar

DACĂ cursul valutar = ridicat ATUNCI vânzările scad cu I(P10) procente

DACĂ cursul valutar = scăzut ATUNCI vânzările cresc cu I(P10) procente

Baza de reguli prezentată poate fi extinsă sau modificată, în dependență de cerințele și strategia companiei. Poate fi modificat atât numărul de parametri cât și baza de reguli, în scopul de a mări eficacitatea controlerului.

Datorită faptului că procesul de adaptare este unul continuu, aplicarea controlerului fuzzy va crește exactitatea prognozei, deoarece se iau în considerație toți factorii dinamici și imprevizibili care pot afecta vânzările.

6. Concluzii

Cu ajutorul unui sistem inteligent bazat pe rețele neuronale, pot fi generate prognoze referitor la vânzări, în baza datelor statistice despre vânzările anterioare. Unul din cele mai importante avantaje al acestei metode este lipsa necesității unei specificații stricte a modelului matematic.

La momentul prognozei pentru o perioadă de timp de câteva luni, este destul de complicat (chiar și în cazul rețelelor neuronale) de prevăzut modificările factorilor externi.

Pentru soluționarea problemei a fost propusă metoda de creare a sistemelor inteligente de prognozare cu utilizarea rețelei neuronale și a controlerului fuzzy conectate în serie, care presupune divizarea procesului de prognozare în 2 etape:

1. Prognozarea – rețeaua neuronală generează șirurile temporale în baza parametrilor interni (istoria vânzărilor pentru ultimii 2 ani);
2. Adaptarea la factorii externi – controlerul fuzzy adaptează datele oferite de rețeaua neuronală la acțiunea factorilor externi.

Etapa a doua este un proces continuu, rezultatele prognozei fiind adaptate la modificarea parametrilor externi.

În consecință, principala problemă a rețelelor neuronale în prognozare - imposibilitatea de a prevedea schimbările parametrilor externi, este rezolvată de controlerul fuzzy, care reacționează la orice modificare a parametrilor externi și ajustează datele prognozei generate de rețeaua neuronală.

Bibliografie:

1. Srinivasan D., Tan S.S., Chang C.S., Chan E.K. Practical implementation of a hybrid fuzzy neural network for one-day-ahead load forecasting // IEE Proc. Gener. Transm. Distrib. 1998. Vol. 145. № 6
2. Box G.E.P. and Jenkins G.M. "Time series analysis: Forecasting and control", San Francisco: Holden-Day, 1970..
3. Hassoun M.H. "Fundamentals of Artificial Neural Networks", A Bradford book, The MIT Press Cambridge Massachusetts, 1995.
4. Morariu N., Iancu E., Vlad S. A neural network model for time series forecasting // Romanian Journal of Economic Forecasting. 2009, No. 4. P. 213 – 223.
5. Mahfoud S., Mani G. Financial Forecasting Using Genetic Algorithms // Applied Artificial Intelligence. 1996, Vol. 10, No.6. P. 543 – 560

6. Basaran Filik U.,Kurban M. A New Approach for the Short-Term Load Forecasting with Autoregressive and Artificial Neural Network Models // International Journal of Computational Intelligence Research. 2007, No.3. P. 66 – 71.
7. Charytoniuk W., Chen M.S. Short-term Forecasting in Power Systems Using a General Regression Neural Network // IEEE Trans. on Power Systems. 1995. Vol. 7. № 1.
8. GARCH A. Forecasting Model to Predict Day-Ahead Electricity Prices / R.C. Garcia [at al.] // IEEE Transactions on Power Systems. 2005, Vol. 20, No. 2. P. 867 – 880