

Use of neural networks for pattern recognition in modeling events on financial markets

Valentin-Dragoș MILITARU
Consilier LogiCons SA

The efficient use of artificial neural networks (ANNs) in predicting the evolution of the financial markets is a highly controversial subject. This paper presents the results of a study performed on the FOREX market, namely on a time series of the EUR/USD rate that covered the entire year 2002. The objective was to build a classification model, able to recognize behavioral patterns in market data. The model was built using a Multilayer Full Feed Forward neural network architecture and it made possible highly accurate predictions (more than 90% accuracy) about the short-term significant market movements. Started from an idea exposed by Prof. Dr. Doyne Farmer in the July 1994 issue of Wired Magazine, this research proved that regions of "local predictability" can be found even in highly complex systems, like the financial markets. It also represents a supplementary proof that the Random Walk Theory and the Efficient Market Theory are obsolete and that consistent predictions can be made about the future evolution of the stock market.

Keywords: Artificial neural networks, FOREX, financial market prediction, pattern recognition, classification models, technical analysis, decision systems.

Preambul

Într-un interviu realizat pentru ediția din Iulie 1994 a revistei *Wired Magazine*, Doyne Farmer, profesor de fizică la Santa Fe Institute, afirma că și sistemele cu grad ridicat de complexitate, precum piețele financiare, sunt caracterizate de regiuni de "predictibilitate locală". Deși astfel de sisteme complexe manifestă un comportament haotic în cea mai mare parte a timpului, fiind astfel impredictibile, anumite părți ale sistemelor pot fi previzionate pentru scurte perioade de timp. Analiza *pattern*-urilor prezente în evoluția piețelor financiare ar trebui în consecință să se concentreze exclusiv asupra acestor regiuni. Prezentul material descrie un studiu care confirmă încă o dată teoria Prof. Farmer. El reprezintă o dovadă în plus că *Random Walk Theory* și *Efficient Market Theory* sunt lipsite de susținere din partea realității.

1. Obiectivul studiului

Studiul descris în prezentul material a fost realizat în vederea explorării posibilităților de identificare a unor corelații între trecut și viitor în evoluția pieței FOREX. Având în vedere că cel mai reprezentativ indicator al acestei piețe este cursul EUR/USD, a fost considera-

tă o serie cronologică conținând valorile sale la intervale de 15 minute pe parcursul întregului an 2002. Obiectivul proiectului a fost stabilit la a construi un model de clasificare care să identifice situațiile în care cursul EUR/USD ar urma să înregistreze într-un orizont de timp de 75 de minute o creștere sau scădere cu cel puțin 5 PIPS față de momentul prezent.

2. Câteva considerații metodologice

2.1. Considerații asupra cercetărilor anterioare privind modelarea piețelor financiare

De mulți ani, modelarea liniară și regresia au fost utilizate în nenumărate încercări de a descoperi modele de previziune aplicabile piețelor financiare. Eșecul acestor încercări a făcut să se încetățenească ideea că piețele financiare sunt sisteme haotice, care evoluează aleator și pentru care evoluția viitoare nu poate fi previzionată în baza datelor despre evoluția trecută.

Avantajul major pe care rețelele neuronale îl prezintă în comparație cu modelele de regresie constă în faptul că aparatul matematic care stă la baza instruirii și utilizării rețelelor permite modelarea facilă și mult mai exactă a

relațiilor neliniare, de unde și capacitatea rețelelor neuronale de a reproduce funcții matematice extrem de complexe.

Cu toate acestea, două greșeli majore sunt foarte comune în aproape toate studiile legate de construirea sistemelor decizionale pentru piețe financiare:

a) Alegerea unei serii cronologice relativ scurte (de multe ori sub 1.000 de înregistrări), pe un trend ascendent, pentru construirea și testarea modelelor. Statistic, majoritatea sistemelor decizionale tind să ofere rezultate bune în perioadele de creștere a pieței pentru ca în perioade de recesiune să genereze pierderi importante.

b) Aprecierea pozitivă a unor modele care determină decizii corecte de investiții numai în 55-65% din cazuri. În realitate, un sistem decizional ar trebui să funcționeze în egală măsură pe piețe aflate în ascensiune și pe piețe aflate în declin și ar trebui să determine investiții profitabile în cel puțin 75-85% din cazuri. În definitiv, chiar și un model care garantează o rată de 80% previziuni corecte trebuie utilizat cu precauție, dacă avem în vedere că o singură pierdere substanțială ar putea anula profitul acumulat printr-o serie lungă de tranzacții încheiate favorabil.

2.2. Considerații asupra datelor aferente piețelor financiare

Previzionarea comportamentului piețelor financiare prin instrumente specifice analizei tehnice este recunoscută a fi o provocare atât pentru cercetători, cât și pentru practicieni, datorită îndeosebi componentei de zgomot care se regăsește în datele de piață, dar și datorită caracterului non-staționar al seriilor cronologice. Din acest motiv, primul pas în crearea unui model previzional pentru piața financiară este eliminarea acestui dezavantaj prin organizarea și pre-procesarea adecvată a datelor primare.

În utilizarea rețelelor neuronale, prezența zgomotului în date limitează capacitatea de generalizare, indiferent de volumul de date avut la dispoziție pentru instruire.

Transformarea variabilelor continue în variabile discrete reprezintă de multe ori o soluție eficientă de reducere a zgomotului și a caracterului nestaționar al seriilor cronologice.

Transformările aplicate asupra datelor în cadrul prezentului studiu sunt în concordanță cu această premiză, iar transformarea variabilelor continue în variabile discrete transformă problema de previziune într-una de clasificare.

2.3. Considerații asupra selectării arhitecturii rețelei

Instruirea unei rețele neuronale artificiale este o procedură de multe ori empirică, bazată pe un număr cât mai mare de experimentări. Și în cazul de față, efortul de optimizare a arhitecturii rețelei utilizate a avut o pondere importantă în volumul total de muncă.

Au fost luate în considerare cinci tipuri de arhitecturi, însă o atenție specială a fost acordată rețelelor Feed Forward. Dat fiind faptul că setul de date rezultat după pre-procesare conținea atât variabilele de intrare, cât și de ieșire, au fost testate metode de instruire supervizată. Back Propagation a reprezentat punctul de început al acestor teste, însă este larg acceptată ideea ca această metodă se bazează pe un algoritm care prezintă marele dezavantaj al blocării în zone de optim local. Din acest motiv, experimentarea a fost extinsă la algoritmi de instruire care prezintă variații față de algoritmul standard, pentru ca în final alegerea să se oprească la *Jacob's Back Propagation*, care s-a dovedit a avea cea mai bună convergență.

3. Construirea modelului de clasificare

3.1. Pre-procesarea datelor

Seturile de date utilizate în acest studiu au fost construite având în vedere următorul obiectiv: fiecare moment de timp să poată fi descris atât în relație cu evoluția din trecut a cursului EUR/USD, cât și în relație cu evoluția viitoare.

Datele inițiale erau conținute într-o singură serie cronologică (valoarea la fiecare 15 minute a cursului EUR/USD pe durata anului 2002). Datelor inițiale li s-au aplicat transformări semnificative, astfel încât, prin încercări succesive, au fost construite trei atribute care să descrie fiecare din cele peste 24.000 de înregistrări. Pentru construcția acestor atribute s-au utilizat formule de calcul apropiate de cele utilizate frecvent de traderi în

calculul indicatorilor de analiză tehnică. Aceste trei atribute for fi numite în continuare astfel: PEI1 și PEI2 (PEI = acronim de la *Past Evolution Indicator*) sunt cei doi indicatori care vor furniza inputurile rețelei și care

descriu evoluția trecută a cursului EUR/USD, iar al treilea indicator, FEI (FEI = *Future Evolution Indicator*) reprezintă outputul rețelei și oferă informații cu privire la evoluția viitoare.

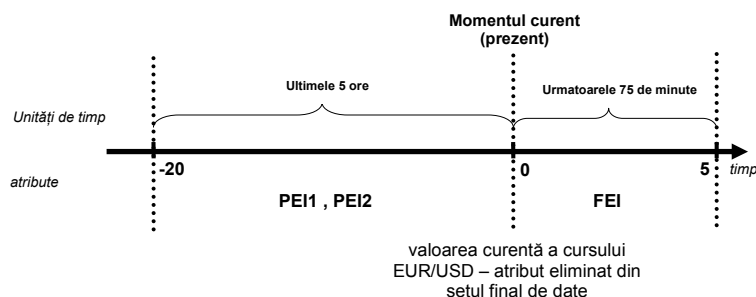


Fig. 1. Reprezentarea variabilelor independente (PEI1 și PEI2) și a variabilei dependente (FEI) în timp. În setul de date final, valoarea zilnică de închidere a indicelui a fost eliminată, pentru reducerea zgomotului din date.

Variabilele independente, PEI1 și PEI2 au fost construite în așa fel încât să descrie fiecare înregistrare din setul de date prin prisma evoluției cursului în ultimele 20 de unități de timp (o unitate = 15 minute). Semnificația fiecărei variabile este următoarea:

a) PEI1 are rolul de a elimina caracterul non-staționar al seriei cronologice inițiale; formula de calcul se bazează pe valoarea de închidere a indicelui Dow Jones în ziua curentă și pe media valori de închidere pentru ultimele 20 de unități de timp. În funcție de evoluția indicelui, PEI1 ia valori pozitive sau negative.

b) PEI2 are rolul de a poziționa valoarea curentă a indicelui în raport cu valorile minime și maxime înregistrate în ultimele 20 de unități de timp. Intervalul de variație al lui PEI2 este $[0;100]$.

Variabila dependentă, FEI, descrie evoluția cursului EUR/USD pentru următoarele 5 momente de timp (totalizând 75 de minute). FEI este o variabilă discretă, care ia valori în mulțimea $\{-1,0,+1\}$. Valorile sale codifică următoarele situații:

- i) evoluția indicelui în următoarele 75 de minute este pozitivă (cursul mediu depășește cu cel puțin 5 PIPS valoarea curentă) \rightarrow FEI = +1;
- ii) evoluția indicelui în următoarele 75 de minute este negativă (cursul mediu scade cu

cel puțin 5 PIPS sub valoarea curentă) \rightarrow FEI = -1;

iii) evoluția indicelui în următoarele 75 de minute este neutră (cursul mediu pentru viitoarele 75 de minute oscilează cu mai puțin de 5 PIPS în jurul valorii curente) \rightarrow FEI = 0.

În conformitate cu aceste transformări aplicate datelor, problema de predicție s-a transformat într-una de clasificare (*pattern recognition*), având drept obiectiv plasarea fiecărei înregistrări într-una din cele trei categorii de evoluție descrise de FEI.

Din setul inițial de date de peste 24.000 de înregistrări a fost selectat în mod aleator un sub-set reprezentând numai 10% din volumul de date aferent întregului an 2002. Structura acestui sub-set a inclus numai cele trei variabile a cărei calculație a fost descrisă anterior. Cele 2.400 de înregistrări selectate au fost împărțite în 4 seturi de date a câte 600 de înregistrări. Instanțele din aceste patru seturi de date sunt distincte, iar din punct de vedere cronologic, seturile de date și implicit înregistrările se succed în ordinea firească (înregistrările din primul set de date le preced cronologic pe cele din setul al doilea ș.a.m.d.). Această organizarea a datelor a asigurat ulterior o evaluare corectă a performanțelor modelului.

3.2. Arhitectura rețelei neuronale

Identificarea arhitecturii optime de rețea, capabilă să genereze cel mai eficient model de clasificare, s-a făcut prin testări repetate. Au fost avute în vedere cinci tipuri de arhitecturi: *Multilayer Normal Feed Forward* (MNFF), *Multilayer Full Feed Forward* (MFFF), *Total Recurrent*, *Prior Recurrent* și *Cascade*. Experimentele au evidențiat că arhitectura MFFF oferă cel mai precis model de clasificare în cazul de față.

În comparație cu arhitectura de tip *Multilayer Normal Feed Forward*, o rețea *Multilayer Full Feed Forward* prezintă două diferențe substanțiale:

- i) în primul rând, fiecare nivel de neuroni primește ca intrări și semnalele de input externe
- ii) în al doilea rând, fiecare nivel primește ca intrări semnalele provenind de la fiecare nivel situat anterior lui în rețea.

Prin urmare, alegerea finală s-a oprit asupra unei rețele MFFF cu 4 niveluri de neuroni: nivelul de input, cu 2 neuroni (PEI1 și PEI2), două niveluri ascunse, cu 4 respectiv 3 neuroni fiecare și nivelul de output, cu un singur neuron (FEI).

De o manieră similară, experimente repetate au condus la identificarea celui mai eficient algoritm de instruire pentru rețeaua neuronală. Cea mai rapidă convergență și în același timp a fost și cea mai mică eroare a demonstrat să aibă o versiune optimizată a algoritmului *Back Propagation*.

3.3. Instruirea rețelei neuronale

Procesul de instruire supervizată s-a derulat folosind PEI1 și PEI2 ca vector de intrare și FEI ca output. Pentru instruirea rețelei s-a utilizat primul set de 600 de înregistrări, iar pentru testarea preciziei modelului s-a apelat la celelalte trei seturi. Pentru a simplifica referirea, aceste seturi denumite în continuare TRAIN, respectiv TEST1, TEST2 și TEST3. Este important de subliniat că obiectivul urmărit în cadrul experimentului reprezintă un caz tipic de clasificare, outputul dorit încadrându-se în mulțimea de valori $\{-1;0;+1\}$. Pe de altă parte, funcția de transfer utilizată pentru nivelul de ieșire al rețelei (*Threshold Linear Bipolar*) produce rezultate

continue, în intervalul $[-1;+1]$. Din acest motiv, cea mai rezonabilă opțiune pentru calculul erorii rețelei pe parcursul instruirii a fost utilizarea abaterii medii pătratică.

Pentru etapa de instruire, ajustarea ponderilor s-a făcut folosind setul de date TRAIN. În paralel, aplicarea rețelei asupra setului de date TEST1 a furnizat o primă confirmare a eficienței modelului.

În ce privește evoluția procesului de instruire, este important de subliniat că arhitectura aleasă asigură o convergență rapidă. După 100 de iterații, descreșterea erorii a devenit mai lentă, ceea ce indică faptul că rețeaua se apropie de soluția optimă. Instruirea a fost oprită după 2000 de iterații, când rata de descreștere a erorii a devenit aproape nulă.

Diferența foarte mică între eroarea înregistrată pe cele două seturi de date (TRAIN și TEST1) demonstrează că atât arhitectura aleasă, cât și structurarea setului de date care a stat la baza instruirii au asigurat o bună generalizare a problemei.

După 2000 de iterații, rețeaua a fost testată folosind toate cele trei seturi de date special selectate (TEST1, TEST2 și TEST3). Rezultatele sunt discutate în secțiunea următoare.

4. Verificarea performanței rețelei. Semnificația modelului pentru adoptarea de decizii pe piețele financiare

4.1. Generalități cu privire la măsurarea performanței

Analiza performanței modelului de clasificare ar trebui să ia în considerare două aspecte. În primul rând, așa s-a mai precizat, variabila dependentă (FEI) are valori discrete, în mulțimea $\{-1;0;+1\}$, în timp ce outputul rețelei este de natură continuă, în intervalul $[-1;+1]$. Datorită zgomotului conținut în datele de input și datorită arhitecturii rețelei neuronale, aceasta va avea tendința de a clasifica corect cazurile de evoluție pozitivă (outputul = +1) sau negativă (outputul = -1), spre deosebire de cazurile de evoluție neutră. Pe de altă parte, caracterul continuu al outputului rețelei și caracterul discret al valorilor variabilei dependente (tipic de altfel pentru o problemă de clasificare), fac posibilă o interpretare flexibilă a rezultatelor. Repartizarea fiecărei in-

stanțe pe una din cele trei clase ar putea fi făcută luând în calcul o marjă de toleranță, aleasă de analist. Această marjă ar putea fi definită ca fiind diferența absolută dintre outputul dorit și outputul generat de rețea. Totuși, pentru simplificarea interpretării rezultatelor, s-a renunțat la această posibilitate. În al doilea rând, performanța modelului ar trebui evaluată strict în corelație cu necesitățile și așteptările unei persoane care adoptă decizii de investiții pe piața financiară. Din punctul de vedere a unui participant la piață, prezintă interes numai instanțele asociate cu o evoluție viitoare pozitivă sau negativă (creșteri, respectiv scăderi ale valorii cursului cu cel puțin 5 PIPS față de momentul prezent). În plus, un participant la piață este preocupat exclusiv de oportunitățile de tranzacționare pe care le-a fructificat, deschizând poziții în piață. Cu alte cuvinte, cazurile de evoluție pozitivă sau negativă care nu au fost semnalate (recunoscute) de rețeaua neuronală, sunt lipsite de importanță, din moment ce ele nu generează profit sau pierdere reală. *In extremis*, aceste situații ar putea fi asociate unor costuri de oportunitate, ele reprezentând ocazii de investiție ratate.

Tabelul 1. Structura celor trei seturi de date în funcție de valorile variabilei dependente (FEI)

Numărul de instanțe	TEST1	TEST2	TEST3
Evoluție negativă (FEI=-1)	94	167	125
Evoluție pozitivă (FEI = +1)	119	188	122
Evoluție neutră (FEI = 0)	387	245	351
Nr. total de instanțe	600	600	600

Rețeaua neuronală funcționează ca un generator de semnale de vânzare/cumpărare. Un semnal de cumpărare este asociat cu un output egal cu 1. Un semnal de vânzare este asociat cu un output egal cu -1. Performanța modelului a fost evaluată prin compararea outputurilor rețelei aparținând mulțimii de valori $\{-1; +1\}$ cu valorile efective ale atributului FEI din cele trei seturi de date. Egalitatea dintre outputul rețelei și valoarea atributului FEI este asociată cu un semnal corect de vânzare sau cumpărare, în timp ce orice diferență dintre output

Următoarele două tabele oferă o imagine statistică a performanței modelului, pentru cele

Pe de altă parte, semnalarea eronată a unui caz de evoluție pozitivă sau negativă determină deschiderea unei poziții necorespunzătoare în piață și, ulterior, pierderi financiare.

În concluzie, capacitatea modelului de a recunoaște pattern-uri de evoluție pozitivă sau negativă în piață trebuie analizată numai prin prisma raportului de instanțe clasificate corect ca semnale de evoluție viitoare pozitivă sau negativă, *versus* instanțe clasificate incorect.

4.2. Informații statistice cu privire la performanța modelului și interpretarea economică a rezultatelor

După cum s-a menționat în secțiunea precedentă, modelul de clasificare construit pe baza setului de date TRAIN a fost testat prin aplicarea rețelei asupra altor trei seturi de date (TEST1, TEST2 și TEST3). Din acest motiv, discuția privind performanța modelului va face referire în egală măsură la fiecare din aceste trei seturi de date.

Distribuția înregistrărilor aparținând fiecăreia din cele trei clase (FEI=-1; FEI=0 și FEI=1) este următoarea:

trei seturi de date.

Prin convenție, să considerăm eroarea de clasificare ca fiind dată de procentajul cazurilor greșit clasificate în numărul total de semnale de vânzare/ cumpărare generate de rețea.

După cum a fost deja precizat în secțiunea 4.1, cea mai importantă caracteristică a unui sistem de decizie cu aplicație pe piețele financiare este să nu genereze semnale false. Prin prisma acestei condiții, modelul prezentat aici se bucură de o precizie mai mult decât rezonabilă (vezi Tabelul 2).

Spre exemplu, eroarea modelului aplicat asupra setului de date TEST2 a atins valoarea de 5.2%, ceea ce înseamnă că 94.5% din cele

116 semnale de vânzare/ cumpărare generate din semnalele generate la aplicarea modelului au fost corecte. Similar, numai 3, respectiv 4 pe seturile TEST1 și TEST3 sunt eronate.

Tabelul 2: Informații statistice privind eficiența aplicării modelului de clasificare asupra seturilor de date folosite la testare

A	Cazuri de evoluție pozitivă (semnale de cumpărare)		Cazuri de evoluție negativă (semnale de vânzare)		Statistică generală		
	B	C	D	E	F=B+C+D+E	G=C+E	H=G / F
Setul de date	Nr. cazuri clasif. Corect	Nr. cazuri clasif. incorect	Nr. cazuri clasif. corect	Nr. cazuri clasif. incorect	Semnale vânz./cump. (total)	Nr. semnale incorecte	Eroare de clasificare (%)
TEST1	20	3	1	0	24	3	12.5%
TEST2	81	5	41	1	116	6	5.2%
TEST3	34	3	17	1	55	4	7.3%

5. Concluzii

În primul rând, acest studiu pune în evidență posibilitățile de utilizare cu succes a rețelelor neuronale artificiale de tip Multilayer Full Feed Forward în probleme de clasificare. În al doilea rând, el dovedește că rețelele neuronale pot trata aspecte complexe de recunoaștere a *pattern*-urilor în cadrul modelării evoluției piețelor financiare. Totuși, cel mai important, acest studiu reprezintă o dovadă în plus că evoluția trecută a piețelor financiare este corelată cu cea viitoare. Existența unor regiuni de "predictibilitate locală" în datele aferente pieței elimină ipoteza valabilității unor teorii precum *Random Walk* sau *Efficient Market Hypothesis*.

Metodologia discutată aici a fost aplicată nu numai asupra unei serii cronologice aferente evoluției cursului EUR/USD, ci și pentru alte instrumente tranzacționate pe piețe reglementate (indici, acțiuni, mărfuri, valute...). O performanță chiar mai bună a fost obținută în cazul unui experiment realizat asupra indicelui Dow Jones Industrial Average, în care *pattern*-uri de evoluție au fost izolate cu precizie de 99%. Se poate afirma deci că aceste studii demonstrează o dată în plus că includerea rețelelor neuronale în sisteme decizionale complexe aplicate piețelor financiare poate îmbunătăți semnificativ performanța activității de tranzacționare.

Bibliografie

1. Chester, D.L. - *Why Two Hidden Layers are Better than One*, IJCNN-90-WASH-DC, Lawrence Erlbaum, 1990, volumul 1, 265-268.
2. Cohen, P.R. - *A Survey of the Eighth National Conference on Artificial Intelligence: Pulling Together or Pulling Apart?*, AI Magazine, Vol. 12, No. 1, 1992, pag17-41
3. Devroye, L., Györfi, L., and Lugosi, G. - *A Probabilistic Theory of Pattern Recognition*, Springer, New York, 1996.
4. Kelly, Kevin - *Cracking Wall Street*; Wired Magazine, Issue 2.07 / July 1994
5. Prechelt, L - *A Quantitative Study of Experimental Evaluations of Neural Network Learning Algorithms: Current Research Practice*, Neural Networks, 9(3), 1996, pag457-462.
6. Tichy, W.F.; Lukowicz, P.; Prechelt, L.; Heinz, A. - *Experimental Evaluation in Computer Science: A Quantitative Study*, Journal of Systems and Software, 28(1):9-18. 1995.
7. Tichy, W. F. - *Should Computer Scientists Experiment More? 16 Reasons to Avoid Experimentation*, IEEE Computer, 31(5), 1998, pag32- 44.
8. Tino, Peter; Schittenkopf, Christian; Dorffner Georg – *Volatility Trading via Temporal Pattern Recognition in Quantized Financial Time Series*; NEC ResearchIndex, 2001(<http://citeseer.nj.nec.com/506945.html>)
9. White, H. - *Connectionist Nonparametric Regression: Multilayer Feed Forward Networks Can Learn Arbitrary Mappings*, Neural Networks, 1990, Issue 3, pag. 535-550 (retipărit în *Artificial Neural Networks: Approximation and Learning Theory*, Blackwell, 1992.