

DETECTAREA ANOMALIILOR ÎN EVOLUȚIA ZILNICĂ A INDICILOR BURSIERI FOLOSIND REȚELE NEURONALE

Lucian Liviu Albu ¹, Radu Lupu ^{2*}

¹ *Academia Română, București, România*

² *Academia de Studii Economice din București, București, România*

Rezumat

Rețelele neuronale au fost folosite mult timp pentru analiza seriilor de timp în diferite aplicații. Dezvoltarea recentă a puterii computerelor și a ingineriei datelor a dus la apariția unei multitudini de algoritmi, care au dobândit denumirea mai largă de Știința datelor. Proliferarea acestor algoritmi se datorează utilizării lor pentru mai multe aplicații, care variază de la simple probleme de clasificare, cum ar fi identificarea corectă a e-mail-urilor de tip „spam”, până la detectarea formelor în imagini și videoclipuri și realizarea de prognoze. Modul lor de funcționare este desemnat de conceptul de *machine learning*, care atestă faptul că acești algoritmi sunt instruiți (învață) pe cantități mari de date, cu scopul de a extrage structuri repetitive cu potențial ridicat de persistență. Pe cale de consecință, este explicabilă atenția recentă acordată posibilității de utilizare a acestor algoritmi pentru aplicații de natură economică. Această lucrare folosește un algoritm, care are la bază rețele neuronale recurente, aplicat la date zilnice pentru mai mulți indici bursieri, cu scopul de a detecta comportamente anormale. Abordarea este înrădăcinată în literatura de specialitate care privește detectarea salturilor, care are ca obiectiv identificarea randamentelor logaritmice anormale de pe piața bursieră. Această abordare permite detectarea unor situații speciale, care au semnificație economică și care pot fi considerate ca anomalii prin comparație cu evoluția lor istorică.

Cuvinte cheie: detectarea anomaliilor, rețele neuronale, LSTM, piața bursieră

Clasificare JEL: C58; C53; G14

Introducere

Aplicațiile științei datelor s-au dezvoltat puternic în ultimii ani. Bazându-se pe o literatură amplă, dezvoltată în secolul XX și pierzând amploare pentru o perioadă de

* Autor de contact, **Lupu Radu** – radu.lupu@rei.ase.ro

timp până când tehnologia s-a dezvoltat, acest domeniu a permis dezvoltarea mai multor instrumente care s-au dovedit esențiale în diferite aplicații din viața reală, motiv pentru care au proliferat în domeniul informaticii și al ingineriei. Folosind tehnici probabilistice similare celor aplicate în econometrie, instrumentele nou dezvoltate servesc mai multor sarcini care ar putea fi împărțite în: probleme de clasificare, probleme de regresie sau probleme de identificare supravegheate (*supervised learning*) și nesupravegheate (*unsupervised learning*). Setul de instrumente care a captat atenția în ultimii ani este cel care utilizează rețelele neuronale, care reușesc să aibă o performanță neașteptată în îndeplinirea acestor sarcini. Prin urmare, este inevitabil faptul că aceste tehnici devin instrumente de analiză pentru domeniul finanțelor, care beneficiază de seturi lungi de date cu comportament aleatoriu și care dețin proprietăți variabile în timp cu persistență de natură neliniară.

Obiectivul articolului nostru este determinat de existența acestei direcții de cercetare, motivat fiind de provocarea de a compara tehnologia existentă în domeniul detecției salturilor pe piața de capital cu un set de modele neliniare, care sunt alimentate de rețele neuronale special concepute pentru seriile de timp. Folosind un set de date privind indicii bursieri cu frecvență zilnică, pentru perioada cuprinsă între ianuarie 2000 și septembrie 2020, s-a construit un model de rețea neuronală care captează dinamica neliniară existentă în date, pentru a identifica comportamentul non-normal sau anomaliile. Această abordare are potențialul de a îmbogăți paradigma actuală de detectare a salturilor, permițând posibilitatea de a detecta nu numai simple valori extreme, ci mai ales o perioadă de dinamică neregulată a seriei temporale respective. Utilitatea acestei abordări se referă la domeniul general al detecției evenimentelor de risc sistemic, crahuri bursiere care ar putea constitui sursa unor fenomene de contagiune, cu scopul de a seta corect scenariile aferente testelor de stres, pe de o parte, și de a investiga în continuare posibilitatea dezvoltării sistemelor de avertizare timpurie, pe de altă parte.

Articolul continuă cu o prezentare a literaturii de specialitate care abordează acest domeniu de cercetare, o descriere a datelor folosite și a metodologiei aplicate asupra acestor date, urmată de expunerea rezultatelor cercetării și o secțiune de concluzii.

1. Revizuirea literaturii științifice

În primul rând, datele financiare au unele caracteristici care le fac speciale: o cantitate mare de date și lipsa regularităților. În acest sens, Chen și colab. (2019) au dezvoltat o tehnică de curățare a datelor, apelând la un algoritm de separare recurent adaptiv, mai exact la o transformare *wavelet* de suprapunere maximală discretă, care permite trei acțiuni referitoare la: salturile dependente de timp, modelele consecutive în timp și perturbații marginale. După luarea în considerare a datelor simulate și, de asemenea, a pieței bursiere din SUA, autorii au concluzionat că precizia este îmbunătățită pentru modelele predictive.

În cercetarea financiară s-au folosit modele de difuzie, cu scopul de a sublinia apariția salturilor, dar s-a dovedit că nu sunt suficient de robuste. În consecință, au fost

propuse și investigate alte metode. O combinație de procese compuse Poisson și modelul de mișcare browniană pare să fi câștigat teren în finanțe și asigurări, prin dezvoltarea proceselor de difuzie cu salturi.

Algoritmii de învățare automată au fost folosiți de Au Yeung și colab. (2020) pentru a identifica salturi în seriile cronologice financiare; autorii au combinat această metodă cu un tip de rețele neuronale recurente, și anume Long Short-Term Memory (LSTM) și au derivat o metodă mixtă pentru a detecta anomalii, care nu sunt definite în prealabil în model. Această tehnică a fost aplicată pe datele pieței bursiere din 11 țări și rezultatele au fost comparate cu cele obținute din alte metode, cum ar fi algoritmul k-nearest neighbors, metoda Hampel sau testul Lee Mykland, concluzionând că noua metodă propusă este mai bună în ceea ce privește acuratețea datelor. Ideea caracteristicilor specifice prețurilor bursiere, în ceea ce privește procesele *path-dependent*, este promovată de Makinen și colab. (2018) după aplicarea unei noi metode, care se bazează pe o rețea neuronală LSTM convoluțională pentru cinci active financiare tranzacționate pe piața SUA. Aceștia consideră că rezultatele lor sunt mai bune decât cele obținute prin utilizarea rețelelor multistrat cu perceptroni sau a modelului LSTM simplu.

Contribuția lui Hawkes în domeniul salturilor (procesele Hawkes care se autoexcită sau se excită reciproc) a avut loc în anii 70, dar recent a fost recunoscută aplicarea lor în domeniul financiar ca „salturi contagioase”, în conformitate cu Hawkes (2020).

Progresul tehnologic a contribuit la disponibilitatea datelor de mari dimensiuni (*Big Data*), extinderea globalizării și, de asemenea, a dus la apariții de noi industrii. Prin aplicațiile disponibile datorită dezvoltării inteligenței artificiale (în special tehnici de învățare automată), anomaliile pot fi detectate în diferite sisteme economice. O istorie a cercetărilor efectuate în ultimii ani este prezentată de Garcia (2019). Rolul detectării anomaliilor este esențial pentru orice activitate economică / industrială, sistemul bancar sau perturbarea oricărei rețele, dar acestea sunt tratate diferit în funcție de dimensiunea datelor. Pentru cazul datelor de dimensiuni mari, Ramchandran și Sangaiah (2018) au propus o specificație mixtă pentru algoritmul de detecție nesupravegheată a anomaliilor, denumit DBN-K care permite o identificare în timp real.

2. Date și metodologie

Analiza noastră se bazează pe date zilnice pentru zece indici bursieri colectați pentru perioada cuprinsă între ianuarie 2000 și septembrie 2020. Am dezvoltat o investigație care a vizat un eșantion de aproximativ 5300 de observații pentru următorii indici: AEX (AEX Index de la Bursa de Valori din Amsterdam), ASX (indicele S&P / ASX 200 de la bursa australiană), ATX (indicele ATX de la bursa austriacă), BEL (indicele bursier BEL 20 de la Euronext Bruxelles), CAC (indicatorul de referință pentru piața bursieră franceză), DAX (indicele bursier german), HSI (indicele Hang Seng din Hong Kong), IBEX (indicele bursier spaniol), NIKKEI (indicele bursier japonez) și SP500 (indicele bursier american). Proprietățile statistice ale rentabilităților

și informațiile despre dimensiunea eșantionului pentru fiecare dintre acești indici sunt furnizate în (Tabelul nr.1). Variația mărimilor eșantionului se datorează sărbătorilor legale, care sunt diferite pe fiecare piața de capital națională.

Tabelul nr. 1: Proprietățile statistice ale randamentelor logaritmice pentru indicii bursieri

	Nr. Observații	medie	std	min	25%	50%	75%	max
AEX	5352	-3.2E-05	0,013983	-0.11376	-0.00603	0,000437	0,006476	0,100283
ASX	5299	0,000132	0,010263	-0.10203	-0,00449	0,000424	0,005305	0,067665
ATX	5253	0,000114	0,014348	-0.14675	-0,0063	0,000392	0,007381	0,12021
BEL	5349	-3E-06	0,012692	-0.15328	-0,00551	0,000255	0,006237	0,09334
CAC	5347	-3.4E-05	0,014406	-0.13098	-0,00656	0,000253	0,007114	0,105946
DAX	5319	0,000122	0,014864	-0.13055	-0,0066	0,000657	0,007386	0,107975
HSI	5206	0,000065	0,014495	-0.13582	-0,00657	0,000222	0,007176	0,134068
IBEX	5345	-9.9E-05	0,014689	-0.15151	-0,00711	0,000415	0,00716	0,134836
NIKKEI	5219	0,000042	0,014784	-0.12111	-0,00686	0,000074	0,007759	0,132346
SP500	5232	0,000167	0,012573	-0.12765	-0,0048	0,00059	0,005795	0,109572

Sursa: Date Bloomberg, calculele autorilor

2.1. Autoencoder

Analiza noastră a constat în utilizarea rețelelor neuronale recurente pentru a detecta posibile dinamici anormale în evoluția acestor indici. Motorul principal al modelului nostru este autoencoderul, care este un anumit tip de rețea neuronală utilizată pentru a efectua învățarea nesupravegheată a codificărilor de date. Principala caracteristică a acestei construcții este de a realiza o reducere a dimensionalității prin convertirea datelor cu scopul de a elimina zgomotul. Acest proces de reprezentare a datelor este însoțit de un altul, care urmărește reconstrucția sa din versiunea codificată. Cea mai simplă versiune a unui autoencoder este o rețea care conține mai multe straturi de neuroni: un strat de intrare, un strat ascuns și un strat de ieșire. Această rețea neuronală încearcă să se asigure că datele sunt copiate din stratul de intrare în stratul de ieșire, cât mai exact posibil, cu constrângerea că stratul ascuns extrage doar cele mai relevante aspecte ale datelor. În acest fel putem spune că stratul ascuns oferă codificarea propriu-zisă (o descriere a codului) a modului de eliminare a părții irelevante a datelor. În Figura nr. 1 este prezentată o schemă a unui codificator automat simplu.

Obiectivul general al replicării datelor, de la partea de intrare la partea de ieșire, necesită implementarea unei aproximări. Începând să fie folosit ca mijloc de reducere a dimensionalității, conceptul este utilizat, pe scară largă, ca parte a algoritmilor de inteligență artificială care plasează acest codificator în rețelele neuronale pentru a obține o precizie mai mare.

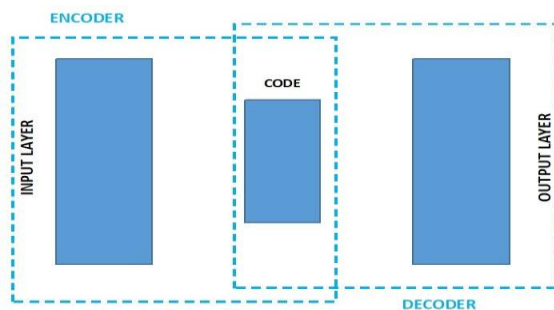


Figura nr. 1: Un Autoencoder simplu

Sursa: concepția autorilor

2.2. Rețea neuronală recurentă - RNN

Rețelele neuronale recurente sunt algoritmi care sunt dezvoltați pe baza mai multor straturi de neuroni (unități), cu scopul specific de a aborda problemele de învățare automată pentru datele secvențiale. Prin urmare, acestea sunt destinate să capteze complexitatea secvențierii temporale și dețin, de asemenea, o memorie internă care ajută la captarea persistenței pe termen lung. RNN-urile de bază sunt organizate în straturi ordonate de neuroni, astfel încât fiecare strat să fie conectat la fiecare nod din stratul următor. Fiecare astfel de neuron conține o funcție de activare care este dependentă de timp și toate conexiunile au ponderi modificabile care se schimbă pe măsură ce datele circulă prin rețea. RNN-urile au capacitatea de a-și actualiza starea curentă, în funcție de stările trecute și de noile date de intrare. Există multe încercări de configurare a arhitecturilor pentru RNN care sunt utilizate, în prezent, și dezvoltate în algoritmi mai complecși (Elman (1990), Jordan, Chen și colab. și Ster (2012)). O descoperire importantă a avut loc odată cu realizarea rețelelor neuronale specifice concepute pentru a gestiona „dependențele pe termen lung”. Hochreiter și Jürgen Schmidhuber (1997) au creat rețelele de memorie pe termen scurt, cunoscute și sub numele de LSTM. Datorită capacității puternice de învățare, rețelele LSTM sunt utilizate pe scară largă în activități, cum ar fi procesarea limbajului natural și modelarea seriilor de timp. O revizuire amănunțită a celulelor LSTM și a arhitecturilor de rețea este redată de Yu și colab. (2019).

2.3. Autoencoderul LSTM

Problemele de predicție au început cu cazul prognozei următoarei valori dintr-o secvență de date, care necesită dezvoltarea unui algoritm de prognoză ce poate intra în categoria „mulți-la-unul”. Cu toate acestea, cele mai importante probleme apar atunci când ne ocupăm de predicția mai multor momente din viitor, pe baza unei serii de date,

care este cunoscută sub numele de probleme de predicție secvență-secvență (seq2seq). Principala preocupare aici este realizarea unei rețele neuronale care ar trebui să aibă capacitatea de a utiliza o anumită dimensiune pentru eșantionul de intrare și de a produce o dimensiune diferită a eșantionului de ieșire, situație care este cunoscută ca „mulți-la-mulți”. Conform Cho (2006), arhitectura LSTM codificator-decodor este concepută pentru a surprinde probleme de predicție seq2seq. Ca autoencoder obișnuit, acest algoritm conține două modele: unul pentru analiza eșantionului de intrare și codificarea acestuia într-un vector și altul pentru decodificarea vectorului și generarea valorilor (*output*) prezise. Confrunțați cu o aplicație care se referă la reprezentarea video, Srivastava și colab. (2015) menționează utilizarea codificatorului-decodificatorului LSTM așa cum este prezentat în Figura nr. 2.

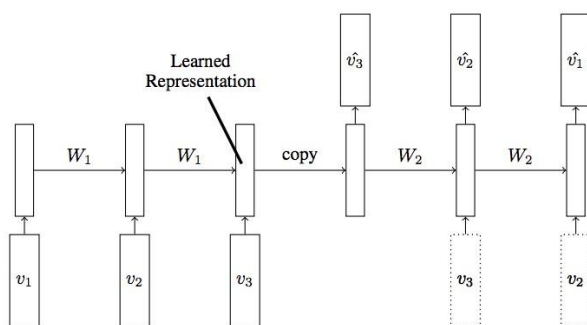


Figura nr. 2: Modelul LSTM Autoencoder

Sursa: Srivastava și colab. (2015)

Folosim acest principiu pentru a crea o arhitectură de auto-codificare LSTM care permite analiza rentabilităților pentru datele noastre de pe piața bursieră. Aceasta cuprinde o rețea LSTM cu 128 de neuroni pentru codificator, un dropout și un vector de repetare și o altă rețea LSTM cu 128 de neuroni pentru decodor. Datele au fost eșantionate în probe secvențiale de câte 30 de observații fiecare. Există 198.273 de parametri pe care acest autoencoder îi calibrează în timpul procesului de învățare, așa cum este prezentat în Figura nr. 3.

Layer (type)	Output Shape	Param #
lstm_38 (LSTM)	(None, 128)	66560
dropout_38 (Dropout)	(None, 128)	0
repeat_vector_19 (RepeatVect	(None, 30, 128)	0
lstm_39 (LSTM)	(None, 30, 128)	131584
dropout_39 (Dropout)	(None, 30, 128)	0
time_distributed_19 (TimeDis	(None, 30, 1)	129
Total params: 198,273		
Trainable params: 198,273		

**Figura nr. 3: Structura Autoencoderului LSTM folosită pentru fiecare serie
(extragere din algoritmul Python)**

Sursa: calculele autorilor

3. Rezultate și discuții

Calibrarea codificatorului LSTM a oferit rezultate structurate în următoarele diagrame. Modelul descris în secțiunea anterioară a fost antrenat până la 100 de epoci (deși, de obicei, antrenamentul a terminat în mai puțin de 20 de epoci în medie) în care monitorizăm pierderea din eșantionul de instruire și pierderea din eșantionul de validare. Am folosit 10% din date pentru setul de validare. Am folosit un parametru de *patience* de 3 pentru a verifica măsura în care am ajuns la un platou sau am găsit minimul global al valorilor de pierdere, în cazul eșantionului de validare. Rularea acestui algoritim pentru fiecare dintre seturile de date a redat un set de erori din eșantion pentru calibrările efectuate, prezentate în Figura nr. 4.

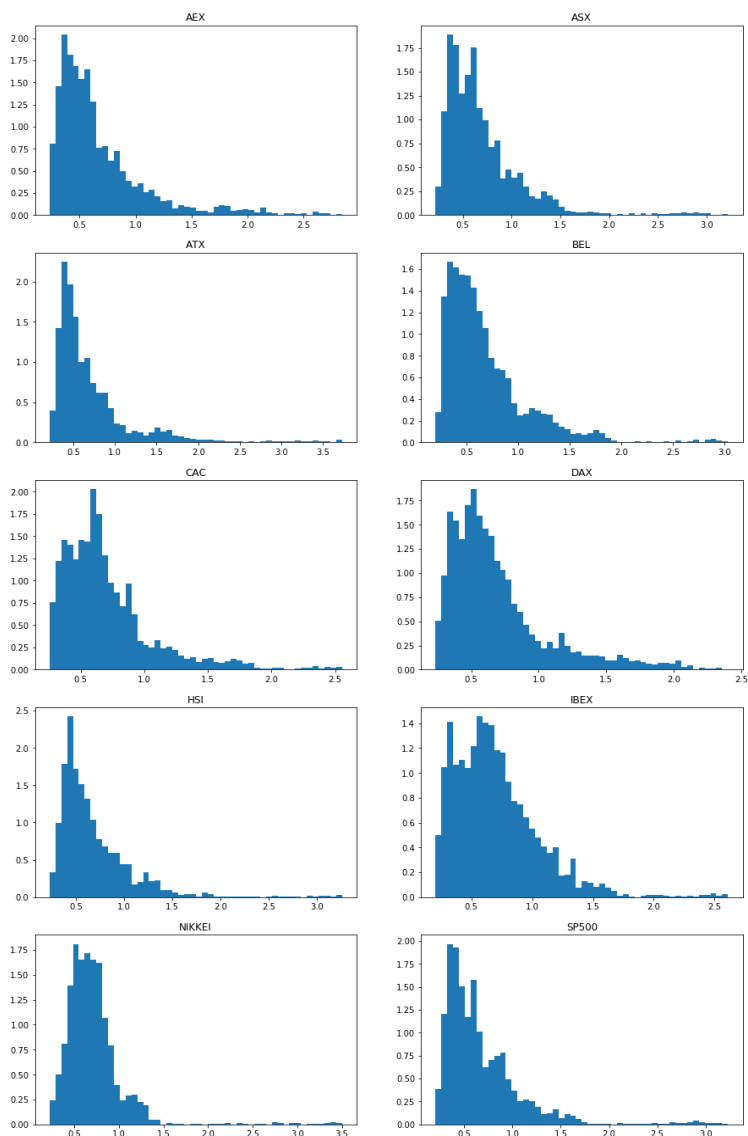


Figura nr. 4: Histograme ale erorilor de instruire (în eșantion) pentru fiecare serie de randamente logaritmice

Sursa: calculele autorilor

Principala observație interesantă este că distribuțiile prezentate în această diagramă (Figura nr. 4) sunt destul de similare, cu aproape aceeași scară atât pe axa

orizontală, cât și pe cea verticală. Toate sunt înclinate spre dreapta, ceea ce înseamnă că există situații în care randamentele noastre tind să fie mai puțin capturate de către autoencoder și pot duce la posibile realizări anormale.

Observăm vârfuri mai mari pentru cazul ASX, ATX, NIKKEI și SP500. Dacă considerăm analiza noastră ca o investigație a măsurii în care dinamica trecută tinde să se repete în viitor, putem spune că există anumite tipare în viitor, de obicei nu de dimensiuni mari, care păstrează un anumit nivel de incertitudine ce tinde să fie similar pentru fiecare dintre seriile noastre. Folosirea rețelelor neuronale este considerată importantă pentru analiza și prognozarea seriilor de timp datorită capacității sale neliniare de a capta tipare repetitive. Eșantionul de instruire folosește toate datele până la 1 ianuarie 2010 și setul de testare conține toate datele de la 1 ianuarie până la 15 septembrie 2020. Acest lucru înseamnă că algoritmul identifică un comportament care a acoperit principalele crize ale pieței bursiere de la începutul 2000 și marea criză financiară din 2007 și 2008. Putem spune că în acest fel rețeaua neuronală are șansa de a experimenta mai multe fenomene anormale cu modificări neliniare complexe. Pe baza setului de parametri deduși din aceste date de instruire, am folosit specificațiile rețelei neuronale pentru a genera prognoză pentru setul de date de testare. Setul de erori obținute pentru fiecare observație este grupat în histogramme, așa cum este prezentat în Figura 5. Putem remarca din nou faptul că, cu puține excepții, setul de erori pentru eșantionul de testare are forme de distribuție similare, cu aceeași asimetrie ca în cazul erorilor din eșantionul de instruire și scale similare pentru cele două axe. Observăm că axele orizontale mai mari aparțin ASX, ATX și SP500.

De asemenea, remarcăm că niciuna dintre aceste distribuții nu este multimodală, fapt care furnizează dovezi că algoritmul reușește să capteze tiparele existente cu o precizie rezonabilă. Având în vedere această capacitate de a capta modele repetitive în date, folosim acest algoritm pentru a extrage dinamica anormală în randamentelor logaritmice, care ar putea fi expresii ale manifestărilor de risc sistemic în date. Încercarea noastră de a identifica aceste anomalii se bazează pe această calibrare, pentru a stabili un prag care ne va ajuta să separăm erorile obișnuite de cele anormale. Am decis să estimăm acest prag pe baza distribuției erorilor în eșantionul de instruire (analiză în eșantion), alegând cuantila 95%. Figura nr. 5 prezintă histogrammele erorilor din setul de testare (în afara eșantionului) și pragurile care au fost obținute prin utilizarea acestei proceduri. Observăm că există o situație în care pragul este prea mare, respectiv cazul randamentelor logaritmice ale indicelui HSI. În acest caz putem spune că comportamentul setului de instruire a fost mai anormal decât cel din setul de testare, ceea ce înseamnă că a doua parte a seriei de date s-a comportat într-un mod mai normal.

Pe baza acestei metodologii am putut acum să identificăm momentele în care algoritmul nostru a observat dinamica neregulată. O reprezentare a acestor momente este ilustrată în Figura nr. 6. Cu doar câteva excepții, putem observa că anomaliile randamentelor logaritmice au fost detectate în principal în intervalul de timp aferent COVID-19, adică în martie 2020, când a început criza pandemiei. Numărul mare de anomalii detectate în această perioadă poate fi văzut ca o confirmare a acurateții algoritmului propus. În acest sens, putem susține că reacția pieței în acest interval de timp este cea mai importantă, asemănătoare cu criza amplă din 2007 și 2008. Un element important al acestor rezultate este faptul că algoritmul preia anomalii doar în

martie, care este clar compatibil cu sentimentul general al pieței. Aceasta înseamnă că, în perioada de după martie 2020, piețele bursiere au revenit la un tip destul de normal de dinamică, reflectând așteptările pentru o fază de „normalitate nouă”.

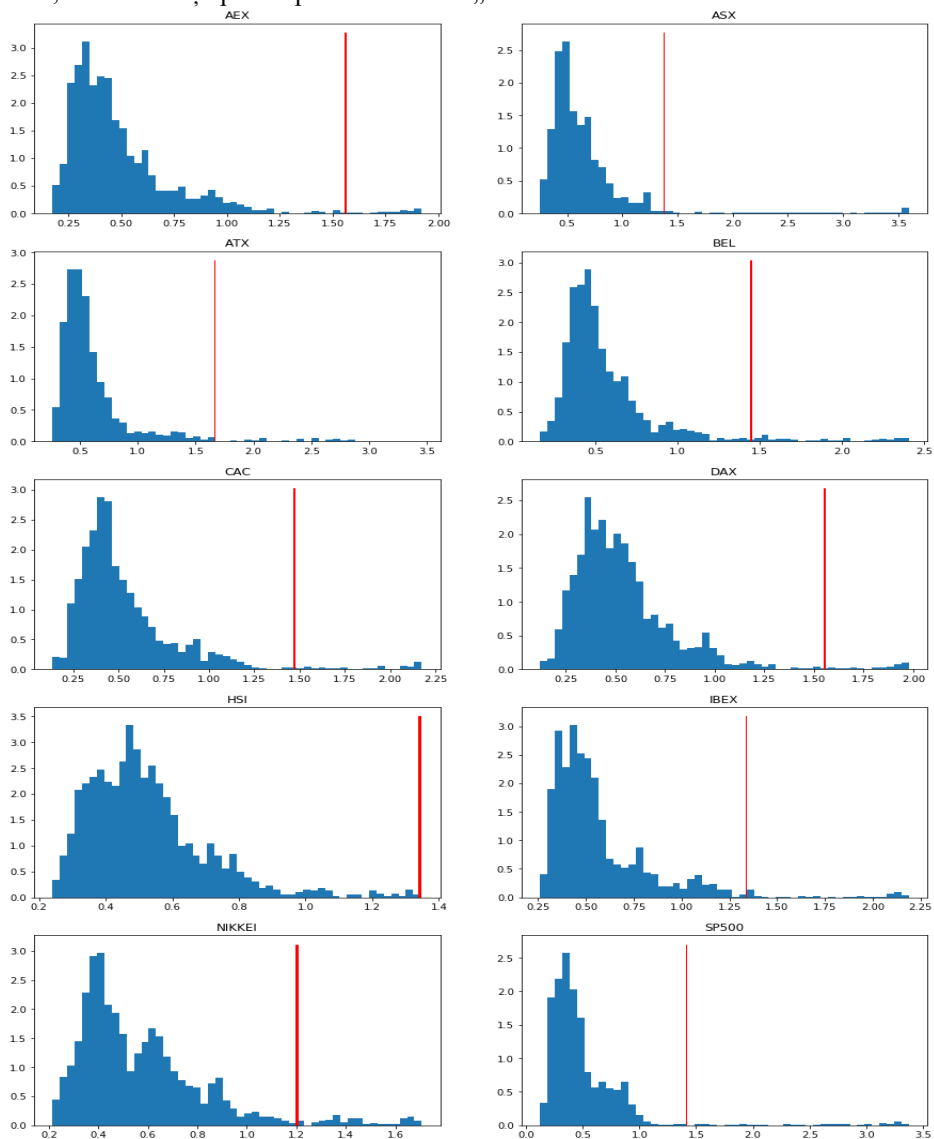
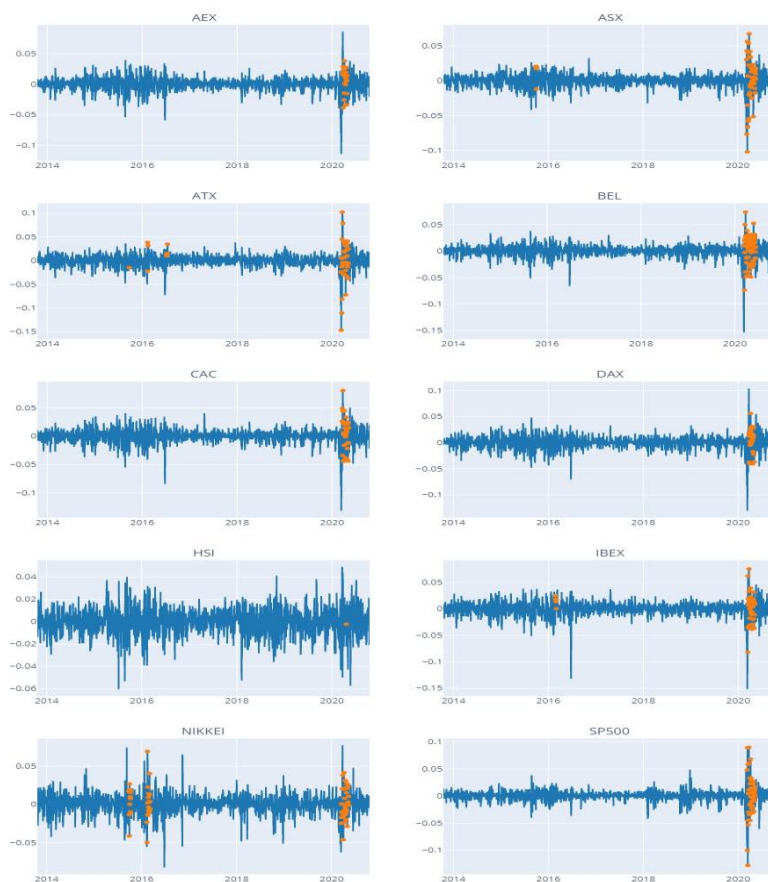


Figura nr. 5: Histograme ale erorilor de testare (în afara eșantionului) pentru fiecare serie de randamente logaritmice

Sursa: calculele autorilor

Cea mai mare acumulare de randamente anormale a fost capturată în cazul ASX, BEL, NIKKEI și, în special, SP500. Observăm că acestea corespund situațiilor în care randamentele logaritmice au fost cele mai mari (atât negative inițial, cât și pozitive ulterior), în comparație cu restul indicilor bursieri analizați. Șocurile au fost, de asemenea, capturate în alte perioade pentru ASX în 2016, ATX și în 2016, IBEX în aproximativ același timp cu cele două anterioare și, mai ales, pentru NIKKEI în același timp.



**Figura nr. 6: Randamentele logaritmice și anomaliile detectate cu
Autoencoderul LSTM**

Sursa: calculele autorilor

Observăm că acest comportament anormal preluat de algoritmul nostru este consecvent cu realizarea unor randamente mari în acel timp, dar nu neapărat în toate situațiile în care am avut aceste randamente mari negative sau pozitive, fapt care constituie un alt element interesant. Atunci când comparăm acest proces de detectare a anomaliilor cu un algoritm simplu de detecție a saltului, trebuie să afirmăm faptul că nu căutăm pur și simplu situații în care randamentele sunt mari, ci și dinamica generală anormală a unui anumit moment de timp.

Dacă literatura de detecție a salturilor încearcă să identifice situațiile în care randamentele logaritmice au atins niveluri mai ridicate în comparație cu volatilitatea locală (în acest fel ținând cont de condițiile locale de piață), abordarea studiului este capabilă să identifice o perioadă cu o dinamică perturbată de impactul prelungit al evenimentelor cu influență semnificativă asupra prețurilor acțiunilor.

Concluzii

Această lucrare folosește date zilnice pentru un set de zece indici bursieri, pe o perioadă de douăzeci de ani, pentru a realiza o analiză a posibilității de a detecta anomalii în randamentele logaritmice. Fiecare set de date este analizat separat cu un autoencoder LSTM, iar setul de date este divizat în eșantion de instruire și eșantion de testare prin luare în considerare a datelor de dinainte și de după data de 01 ianuarie 2010. Obiectivul studiului a fost înțelegerea utilității modelării non-liniare a datelor prin folosirea rețelelor neuronale pe serii temporale financiare, astfel încât să se poată obține mai mult decât randamente extreme, așa cum s-a realizat anterior în literatura de detectare a salturilor. Rezultatele studiului arată că această abordare oferă o perspectivă interesantă asupra aplicării unor astfel de tehnici de învățare automată. Nu numai că s-a reușit să se identifice un comportament anormal, dar s-a putut, de asemenea, să se releve faptul că un astfel de comportament este legat de evenimente influente clare, cum ar fi explozia crizei pandemiei din martie 2020. Datorită acestor rezultate, abordarea prezentată în această lucrare merită o atenție suplimentară pentru a fi utilizată ca instrument pentru detectarea sistemică a riscurilor și creează oportunitatea pentru un studiu ulterior al algoritmilor de avertizare timpurie.

Referințe bibliografice

- [1] Chen, T. B., and Soo, V. W. A comparative study of recurrent neural network architectures on learning temporal sequences. *In Proceedings of International Conference on Neural Networks (ICNN 1996)* (Vol. 4, pp. 1945-1950). IEEE. doi: 10.1109/ICNN.1996.549199
- [2] Chen, Y. T., Lai, W. N., & Sun, E. W. Jump Detection and Noise Separation by a Singular Wavelet Method for Predictive Analytics of High-Frequency Data. *Computational Economics*, 54(2):809–844, 8 2019. ISSN 15729974. doi: 10.1007/s10614-019-09881-3.

- [3] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation Kyunghyun. *Journal of Biological Chemistry*, 281(49):37275–37281, 2006. ISSN 00219258.
- [4] Elman, J. L. Finding Structure in Time. *Cognitive Science*, 14(2):179– 211, 3 1990. ISSN 03640213. doi: 10.1207/s15516709cog1402 1
- [5] Hawkes, A. G. Hawkes jump-diffusions and finance: a brief history and review. *European Journal of Finance*, 2020. ISSN 14664364. doi: 10.1080/1351847X.2020.1755712.
- [6] Hochreiter, S., and Schmidhuber, J. Long short-term memory. *Technical Report* 8, 1997.
- [7] Jordan, M. I. Attractor dynamics and parallelism in a connectionist sequential machine. *Artificial neural networks: concept learning*, pp. 112-127. 1990.
- [8] Mäkinen, Y., Kannianen, J., Gabbouj, M., and Iosifidis, A. Forecasting of Jump Arrivals in Stock Prices: New Attention-based Network Architecture using Limit Order Book Data. 10 2018.
- [9] García, N. M. Multi-agent system for anomaly detection in Industry 4.0 using Machine Learning techniques. *ADCAIJ: Advances In Distributed Computing And Artificial Intelligence Journal*, 8(4):33, 9 2019. ISSN 2255-2863. doi: 10.14201/adcaij2019843340.
- [10] Ramchandran, A., and Sangaiah, A. K. Unsupervised Anomaly Detection for High Dimensional Data—an Exploratory Analysis. In *Computational Intelligence for Multimedia Big Data on the Cloud with Engineering Applications*, pages 233–251. Elsevier, 1 2018. doi: 10.1016/B978-0-12-813314-9.00011-6.
- [11] Srivastava, N., Mansimov, E., and Salakhudinov, R. Unsupervised Learning of Video Representations using LSTMs. *32nd International Conference on Machine Learning, ICML 2015*, 1:843–852, 2 2015. URL <http://arxiv.org/abs/1502.04681>.
- [12] Šter, B. Selective Recurrent Neural Network. *Springer*, 2012. doi: 10.1007/s11063-012-9259-4. URL
- [13] Yeung, J. F. A., Wei, Z. K., Chan, K. Y., Lau, H. Y., and Yiu, K. F. C. Jump detection in financial time series using machine learning algorithms. *Soft Computing*, 24(3):1789–1801, 2 2020. ISSN14337479. doi: 10.1007/s00500-019-04006-2.
- [14] Yu, Y., Si, X., Hu, C., and Zhang, J.. A review of recurrent neural networks: Lstm cells and network architectures, 7 2019. ISSN: 1530888X. doi: https://doi.org/10.1162/neco_a_01199