RAPORT DE CERCETARE INTERMEDIAR NR. 2

PROIECT „SMART BUILDING. SMART CITY.”

ACADEMIA OAMENILOR DE ȘTIINȚĂ DIN ROMÂNIA (AOSR)

Modelarea și predicția consumului de energie electrică într-o clădire inteligentă utilizând rețele neuronale adânci

Coordonator proiect: Dr.-Ing. Vladimir Tanasiev

Autor: Dr.-Ing. Grigore Stamatescu

Septembrie 2018

Table of Contents

REZUMAT **3**

1. INTRODUCERE **4**

2. STADIUL ACTUAL AL CERCETĂRILOR **6**

3. METODE UTILIZATE **7**

4. STUDIU DE CAZ ȘI REZULTATE **11**

5. CONCLUZII **16**

6. BIBLIOGRAFIE **17**

Rezumat

Pe măsură ce strategiile de management la scară largă al energiei au început să pună accentul pe consumator în detrimentul producătorilor, mediul construit și clădirile inteligente în mod special joacă un rol determinant în managementul eficient al rețelelor electrice. Suplimentar, unele clădiri au devenit prosumatori prin integrarea capacităților de generare locale din surse regenerabile de energie introducând astfel elemente suplimentare de complexitate în operarea sistemelor energetice. Ca o alternativă la tehnicile convenționale de modelare ale consumului de energie o abordare intrare-ieșire de tipul black-box poate identifica tiparele de consum principale și trendurile asociate prin exploatarea cantităților mari de date generate și înregistrate de echipamentele inteligente prin care se realizează instrumentarea clădirilor moderne. În acest raport este prezentată o abordare de aplicare a rețelelor neuronale artificiale clasice si adânci („deep learning”), în special rețele neuronale recurente (RNN – Recurrent Neural Networks) implementate prin straturi de tipul Long Short-Term Memory (LTSM) pentru modelarea și predicția consumului de energie. Se pune accentul pe clădirile comerciale de dimensiuni medii și mari ce pot fi administrate centralizat de o manieră eficientă și în cazul cărora modelele de predicție ale consumului mai performante pot rezulta în economii semnificative de costuri și un impact mai larg la nivel social. Studiul de caz este ilustrat pe două clădiri academice amplasate în climate temperate de-a lungul unui an de operare și folosind seturi de date deschise ce permit rezultate replicabile. Rezultatele obținute pot fi utilizate mai departe în scheme robuste și algortmi de managementul sarcinilor cu efort redus de ajustare al ponderilor prin antrenare/învățare incrementală.

1. Introducere

Tendințele globale de urbanizare recente au condus la provocări inginerești semnificative pentru dezvoltarea și administrarea eficientă a mediului construit. Orașele inteligente („Smart Cities”) reprezintă unul dintre exemplele proeminente ale unei noi paradigme care reunește la un loc elemente de senzori și măsurări, calcul, comunicație și control automat pentru îmbunătățirea operării diferitelor sisteme componente precum și a calității vieții locuitorilor acestora. Una dintre direcțiile critice de dezvoltare într-un oraș inteligent este sectorul energetic și al rețelelor electrice de utilități, cu rolul de a asigura o alimentare cu energie fiabilă, curată și de cost redus pentru nevoi urbane în continuă creștere. În plan specific raportul curent se axează pe clădirile comerciale de dimensiuni medii și mari care pot juca un rol esențial în calitate de consumatori, prosumatori sau entități de echilibrare pentru stabilitatea rețelelor electrice. Astfel este relevată nevoia unor modele precise care să capteze tiparele si trendurile principale ce determină consumul de energetic și care pot fi utilizate ulterior pentru prediciția unor curbe de sarcină cu scopul de a îmbunătăți performanța algoritmilor de control de nivel înalt. Principalul obiectiv este adresarea provocărilor actuale în domeniul stabilității rețelei și cele de mediu prin dezvoltarea, implementarea și evaluarea unor algoritmi avansați de modelare și predicție.

Deoarece consumul de energie al clădirilor este într-o creștere continuă și a atins circa 40% din necesarul de energie primară în multe țări dezvoltate [1], clădirile moderne utilizează rețele de instrumentație extinse pentru măsurarea și controlul mai multor parametri energetici relevanți în operarea zilnică. Fluxurile de date rezultante sunt gestionate de sisteme informatice cum sunt sistemele de management al clădirilor (BMS – Building Management Systems) și sunt stocate în baze de date locale sau în cloud pentru prelucrări ulterioare sau generarea de decizii on-line. În clădirile existente și mai vechi înlocuirea contoarelor de energie convenționale cu contoare inteligente („smart meters”), care comunică prin rețele cu sau fără fir, oferă o modalitate practică de a colecta datele necesare. Datorită disponibilității în creștere a datelor și a scăderii efortului și costului de colectare al acestora există mai multe abordări pentru construirea unor modele precise pentru predicția și controlul consumului energetic.

Ca exemplu, în perioada recentă, metodele de învățare statistică au consemnat o adopție în creștere atât în cercetare cât și în aplicații industriale, stimulate fiind de disponibilitatea unor seturi de date de calitate și de creșterea exponențială a resurselor de calcul disponibile, inclusiv prin intermediul sistemelor cloud. Rețelele neuronale artificiale reprezintă un exemplu proeminent de tehnică de învățare statistică care oferă rezultate bune în multe tipuri de aplicații. Acestea pot include atât sarcini de clasificare cât și pentru metode de regresie în cazul cărora obiectivul îl reprezintă predicția unei valori numerice de interes. Tehnicile de învățare adâncă dezvoltă arhitecturile existente, bine studiate, de rețele neuronale artificiale, la un nivel de complexitate superior. Aceasta se realizează prin includerea unui număr foarte mare de straturi ascuse in topologia rețelei precum și a multor tipuri de unități de prelucrare a datelor la fiecare nivel. Deși implementate inițial în cadrul unor inițiative comerciale din zona sistemelor de prelucrare multimedia audio/video și de traduceri, o serie de alte aplicații tehnice pot beneficia de disponibilitatea exista a unor algoritmi și utilitare open-source, validate în comunitățile tehnice de origine. Accentul în acest raport este pus pe clădirile comerciale de dimensiune medie și mare, în contrast cu o zonă similară adresată unităților locative individuale și a mediului rezidențial. Aceasta de datorează în special motivațiilor economice și randamentului investițiilor în tehnologii și proiecte de îmbunătățire a eficienței energetice atunci când valori relative mici pot fi translatate în câștiguri absolute semnificative pentru proprietarii și/sau operatori clădirilor vizate.

Principalele contribuții ale acestui raport pot fi sintetizate după cum urmează:

- ilustrarea unor abordări de modelare și predicție ce utilizează rețele neuronale clasice și adânci pentru consumul energetic al clădirilor comerciale de medii și mari dimensiuni, ca alternativă la tehnicile convenționale de modelare;
- prezentarea unui studiu de caz experimental pentru utilizarea tehnicilor de învățare adâncă în predicția robustă a consumului energetic al clădirilor.

Raportul este organizat în continuare după cum urmează. Capitolul 2 prezintă stadiul actual al cercetărilor prin referire la publicații actuale relevante din domeniul modelării consumului energetic. Capitolul 3 ilustrează baza teoretică a rețelelor neuronale clasice și a rețelelor neuronale recurente implementate prin straturi de unități LTSM precum și aplicabilitatea acestora. Un studiu de caz este prezentat în Capitolul 4 prin aplicarea tehnicilor ilustrate pe seturi de date de referință provenite de la clădiri reale și discutarea rezultatelor. Capitolul 5 sintetizează potențialul de dezvoltare al acestor metode.

2. Stadiul actual al cercetărilor

Aplicarea tehnicilor de învățare statistică pentru modelarea consumului energetic în contexte diferite. Sunt indentificați următorii factori pentru aceste tendințe:

- disponibilitatea îmbunătățită și în creștere a unor seturi de date de bună calitate și a resurselor de calcul suficiente și de cost redus care să permită testarea și validarea la scară largă a metodelor propuse;
- biblioteci de algoritmi comerciale și open-source precum și software cu documentație adecvată pentru o audiență tehnică mai larga în afara nucleelor de cercetători în inteligentță artificială din mediul academic și din industrie;
- un cadru de colaborare în creștere între experții în algoritmi, calculatoare și control și specialiștii din domeniile energiei și construcțiilor; aceasta a influențat proiectarea unor noi arhitecturi de învățare adâncă adaptate la direcții de aplicații specifice.

În [2] autorii abordează două tehnici de învățare adâncă pentru modelarea consumului de energie al clădirilor în speță mașini Boltzmann condiționale restricționate (CRBM – Conditional Restricted Boltzmann Machines) și mașini Boltzmann condiționale restricționate factorizate (FCRBM – Factored Conditional Restricted Boltzmann Machines). Rezultatele sunt prezentate de o manieră comparativă cu rețelele neuronale și mașinile de vectori suport (SVM – Support Vector Machines) precum și față de rețelele recurente. Este arătat cum pentru multe dintre scenariile de testare: date de consumuri energetice agregate și subcontorizate precum și la scări temporale și rezoluții diferite, metodele alese oferă performanțe superioare relative la eroarea medie pătratică (RMSE – Root Mean Squared Error) pe orizontul de predicție. Autorii [3] prezintă o abordare de învățare adâncă completă pentru predicția consumului energetic al clădirilor comerciale prin combinarea unor stive de autoencoderi (SAE – Stacked AutoEncoders) cu mașini de învățare extremp (ELM – Extreme Learning Machines). Componenta SAE este responsabilă de extragerea automată a trăsăturilor relevante din setul de date de intrare în timp ce ELM funcționează ca predictor. Avantajele acestei abordări specifice constau în obținerea directă a ponderilor de ieșire ale rețelei fără o propagare inversă iterativă. Două modele RNN bazate pe unități LTSM pentru consumul de energie al clădirilor

sunt evaluate în [4]. Un rezultat important al studiului este faptul că metodele RNN tind să ofere rezultate mai bune atunci când sunt aplicate pe date de energie agregate care fac dependențele pe termen lung să fie mai greu de identificat. De asemenea autorii folosesc modelul rezultat pentru a infera valorile lipsă din seria de timp originală.

În [5] a fost prezentat un cadru de testare al modelelor ARIMA și ANN clasice pe date de energie colectate local. Studiul diferitelor configurații ANN este discutat în [6] în timp ce modelele rezultante pot fi integrate ulterior într-un sistem de suport al deciziei [7].

3. Metode bazate pe rețele neuronale utilizate în modelarea și predicția consumului de energie electrică

Datorită aplicabilității largi precum și a performanțelor bune dovedite, rețelele neuronale recurente au devenit un instrument important în scenarii reprezentate de secvențe de evenimente, în cazul nostru un eveniment reprezentând un punct de date. RNN pot implementa același calcul pentru toate elementele dintr-o serie de intrări și, de exemplu, pot evalua probleme neliniare de serii de timp, cum este și cazul consumului de energie electrică și pot oferi predicții. RNN au o structură diferită de rețelele neuronale artificiale clasice (ANN – Artificial Neural Networks) și folosesc propagarea inversă în timp (BPTT – Back-propagation through Time) [8] sau algoritmul de învățare recurentă în timp real (RTRL – Real-time recurrent learning) [9] pentru a calcula evoluția gradientului după fiecare iterație. Astfel, RNN reprezintă un caz particular de ANN care adaugă ponderi suplimentare rețelei pentru a crea bucle închise în graful rețelei pentru a controla stările interne ale acesteia.

Rețelele neuronale LSTM reprezintă un caz particular de RNN. Similar cu cele menționate anterior sunt folosiți algoritmii de tipul BPTT sau RTRL pentru antrenare dar în unele cazuri cercetătorii au arătat că folosirea acestor metode poate conduce la problema gradientului excesiv. LSTM [10] oferă suport pentru utilizarea seriilor de timp și a altor date secvențiale ca date de intrare într-o rețea și structura acesteia ajută la propagarea gradientului peste secvențe lungi în timpul antrenării. LSTM adresează această problemă prin integrarea unor „porți” auto-conectate în unitățile ascunse. Într-o rețea LSTM fluxul de informație prin rețea este gestionat prin citirea, scrierea și ștergerea informațiilor din memorie [11,12].

Figura 1 ilustrează parcursul unei serii de timp x de lungime n, printr-un strat LSTM. În această diagramă, h reprezintă ieșirea celulei, sau starea ascunsă, c este starea celulei. Prima unitate LSTM preia starea inițială a rețelei: c0, h0, și prima valoare a secvenței x1 și ulterior calculează prima ieșire h1 și actualizarea stării celulei c1. Procesul este reluat la fiecare perioadă de eșantionare.

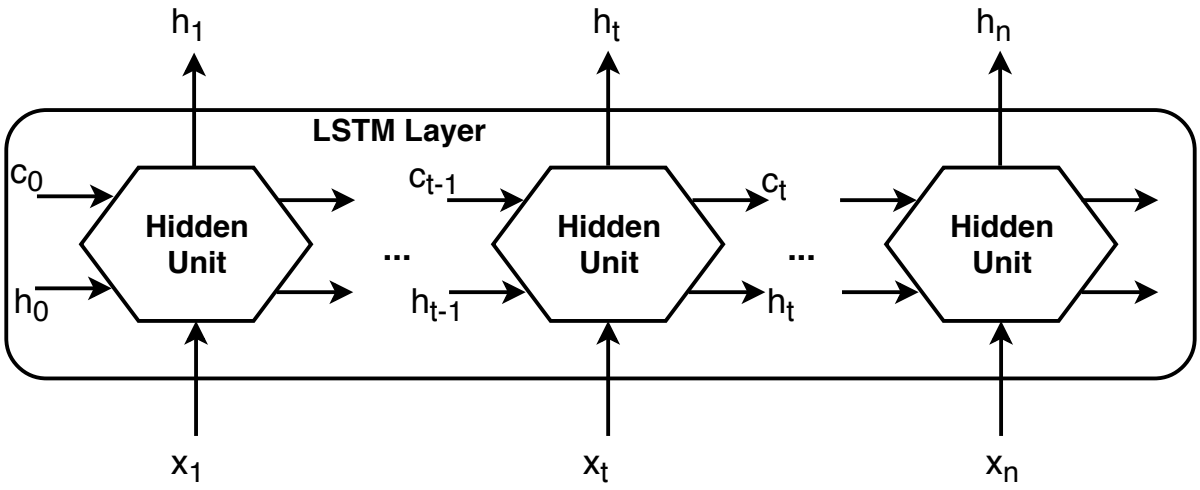


Figura 1 Diagrama unui strat LSTM

Starea unui strat LSTM constă în starea de ieșire h și starea celulei c. Pe de-o parte starea de ieșire la momentul de timp t conține ieșirea stratului LSTM pentru acest moment de timp, în timp ce starea celulei conține infomația învățată la momentele de timp anterioare. La fiecare iterație, stratul scrie informație în starea celulei sau șterge informație din aceasta, procesul fiind controla prin „porți". Acestea reprezintă componentele care controlează starea celulei și starea de ieșire a stratului. Există patru astfel de componente: poarta de intrare i care controlează nivelul de actualizare al stării celulei, intrarea stratului g care adaugă informații stării celulei, poarta de uitare f care controlează nivelul de resetare al stării celulei și poarta de ieșire o care controlează nivelul stării celulei adăugat la starea de ieșire [12].

Diagrama din Figura 2 ilustrează fluxul de date la un moment de timp t în cadrul unei unități ascunse LSTM.

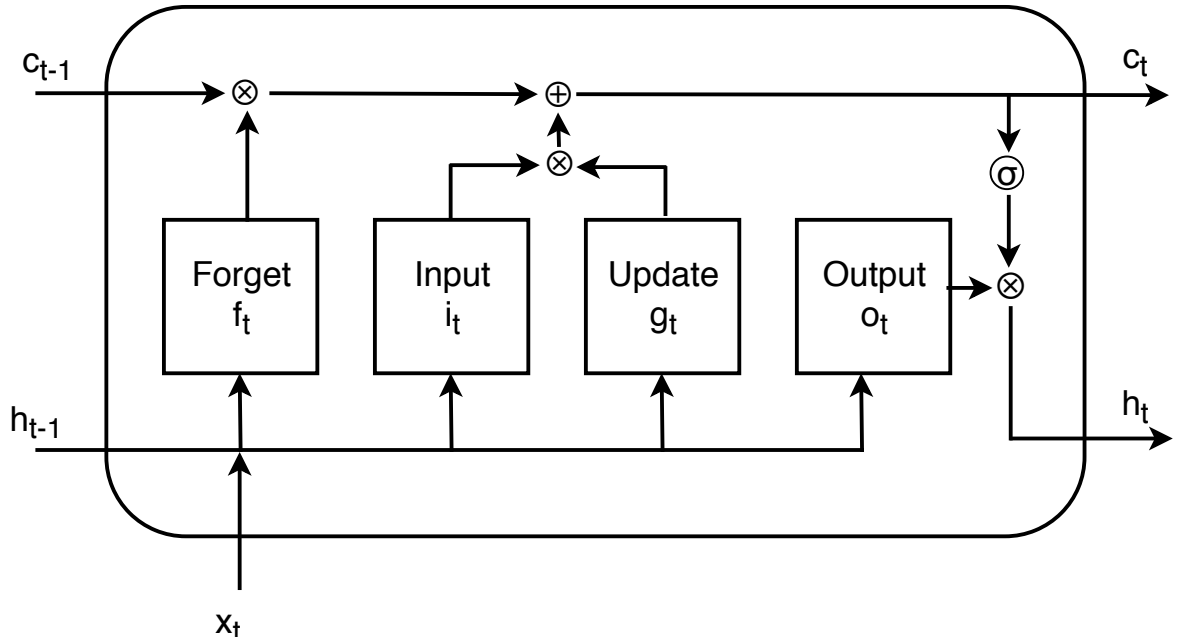


Figura 2 Diagrama unei unități ascunse LSTM

Un strat LSTM are următorii parametri de învățare: ponderile de intrare W, ponderile recurente R și factorul de bias b. W, R și b reprezintă matrice formate prin concatenarea ponderilor de intrare, a ponderilor reucrente și bias-ul fiecărei componente. Structura fiecărei matrice este următoarea:

$$W = \begin{pmatrix} W_i \\ W_f \\ W_g \\ W_o \end{pmatrix} R = \begin{pmatrix} R_i \\ R_f \\ R_g \\ R_o \end{pmatrix} b = \begin{pmatrix} b_i \\ b_f \\ b_g \\ b_o \end{pmatrix}$$

unde i, f, g și o reprezintă porțile de intrare, uitare, intrare strat și ieșire.

Starea memoriei celulei la momentul de timp t este actualizată recursiv folosind formula [10]:

$$c_t = f_t \otimes c_{t-1} \oplus i_t \otimes g_t$$

unde $\otimes$ este produsul Hadamard și ecuațiile pentru calculul fiecărei componente la momentul de timp t sunt următoarele:

$$f_t = \sigma\left(W_f x_t + R_f h_{t-1} + b_f\right)$$
$$i_t = \sigma(W_i x_t + R_i h_{t-1} + b_i)$$

$$g_t = \tanh\left(W_g x_t + R_g h_{t-1} + b_g\right)$$

σ reprezintă funcția sigmoidă în timp ce tanh reprezintă funcția tangentă hiperbolică.
Starea de ieșire la momentul de timp t este dată de poarta de ieșire o care implementează o funcție de citire combinată cu starea celulei c. Procesul este descris de următoarea ecuație:

$$h_t = o_t \otimes \tanh\left(c_t\right)$$

unde

$$o_t = \sigma(W_o x_t + R_o h_{t-1} + b_o)$$

Studiul de caz prezentat în acest raport prezintă o aplicație de modelare bazată pe LSTM folosind diferite configurații de rețea și evaluarea performanțelor pentru fiecare dintre modelele de predicție LSTM.
Seturile de date pentru consumul energetic sunt preluate din baza de date a grupului de Știința Datelor pentru Clădiri și mediul Urban (BUDS) de la Universitatea Națională din Singaporre – www.budslab.org Acestea sunt parte a unui efort la scară largă de colectare și publicare a datelor de consum energetic dintr-o serie de clădiri comerciale/academice cu scopul de a permite testarea unor abordări și algoritmi replicabili în cadrul comunității academice.
Datele de interes au fost colectare cu o perioadă de eșantionare de 60 de minute pe o durată de un an în două clădiri academice cu o suprafață construită de circa 9000 mp. Clădirile alese pentru acest studiu provind din campusuri universitare din Chicago, SUA și Zuerich, Elveția. După pre-procesarea acestora pentru reducerea zgomotului și completarea valorilor lipsă prin utilizarea unui filtru median, au fost obținute două serii de timp cu aproximativ 8670 puncte de date fiecare.
Alegerea clădirilor de interes din baza de date s-a bazat pe replicabilitatea la nivel local în viitor a studiului pe două clădiri academice moderne din campusul UPB. Factorii determinanți pentru selecție au fost, suprafața medie-mare construită, tipar de utilizare mixtă ce include birouri, laboratoare, săli de clasă și un climat temperat cu patru anotimpuri distincte [6].

4. Studiu de caz și rezultate

Seriile de timp de intrare pentru modelarea și testarea rețelelor LSTM proiectate sunt ilustrate în Figura 3. Graficul superior reprezintă energia activă consumată pe parcursul unui an de clădirea din Zuerich iar cel inferior consumul clădirii din Chicago.

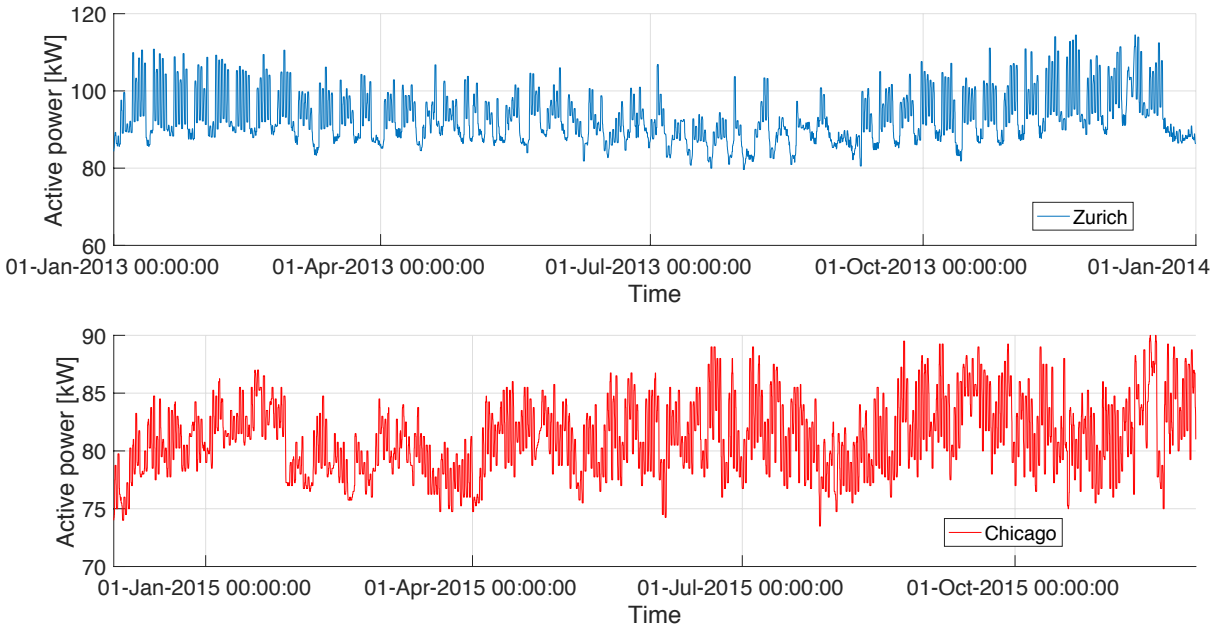


Figura 1 Seturile de date utilizate pentru antrenarea rețelei LSTM

În acest studiu de caz am propus o serie de configurații de rețele neuronale LSTM pentru modelarea și predicția consumului de energie electrică în acest tip de clădiri. Obiectivul este obținerea atât a unei evaluări a metricilor de performanță a algoritmilor cât și analiza efortului de calcul necesare pentru setul de date de referință.

Este utilizat algoritmul LSTM standard pentru metodologia de predicție. Toate arhitecturile de rețele neuronale folosesc unități LSTM tip cu o secvență de intrare, un strat LSTM, un strat conectat comple și un strat de regresie. Fiecare rețea are o configurație diferită reprezentată de numărul de unități ascunse din stratul LSTM. Următoarele structuri de rețea au fost implementate și validate:

- C0: 5 unități ascunse în stratul LSTM;
- C1: 25 de unități ascunse în stratul LSTM;

- C2: 50 de unități ascunse în stratul LSTM;
- C3: 100 de unități ascunse în stratul LSTM;
- C4: 125 de unități ascunse în stratul LSTM;

- Z0: 5 unități ascunse în stratul LSTM;
- Z1: 25 de unități ascunse în stratul LSTM;
- Z2: 50 de unități ascunse în stratul LSTM;
- Z3: 100 de unități ascunse în stratul LSTM;
- Z4: 125 de unități ascunse în stratul LSTM;

unde C reprezintă clădirea din Chicago iar Z reprezintă clădirea din Zuerich.

Referitor la procesul de antrenare, a fost utilizat algoritmul de estimare cu moment adaptiv (ADAM – Adaptime Moment Estimation). ADAM reprezintă un algoritm de optimizare cu gradient de ordinul întâi pentru funcții obiectiv stocastice cu moment. Algoritmul calculează rate de învățare care se pot adapta automat la funcția optimizată pentru fiecare parametru din estimările momentelor de ordinul întâi și doi ai gradienților. Acesta menține o medie alunecătoare la nivel de element a gradienților parametrilor și ale valorilor lor pătratice [14]. În literatura de specialitate a fost demonstrat că algoritmul este eficient în termeni de timp de calcul, necesar de memorie și este astfel adecvat problemelor de mari dimensiuni.

Legat de rata de învățare, după o serie de teste, s-a determinat următoarea abordare: rata inițială de învățare a fost setată la 0.1 și după aceasta a fost redusă cu un factor de 0.2 la fiecare 100 de epoci de antrenare. Numărul maxim de epoci de antrenare a fost ales 200.

Figurile 4 și 5 prezintă răspunsul predicțiilor rețelei LSTM cu 50 de unități ascunse pentru ambele clădiri studiate. Graficele ilustrează o bună modelare a seriilor de timp prin modele LSTM pentru aceste date de testare.

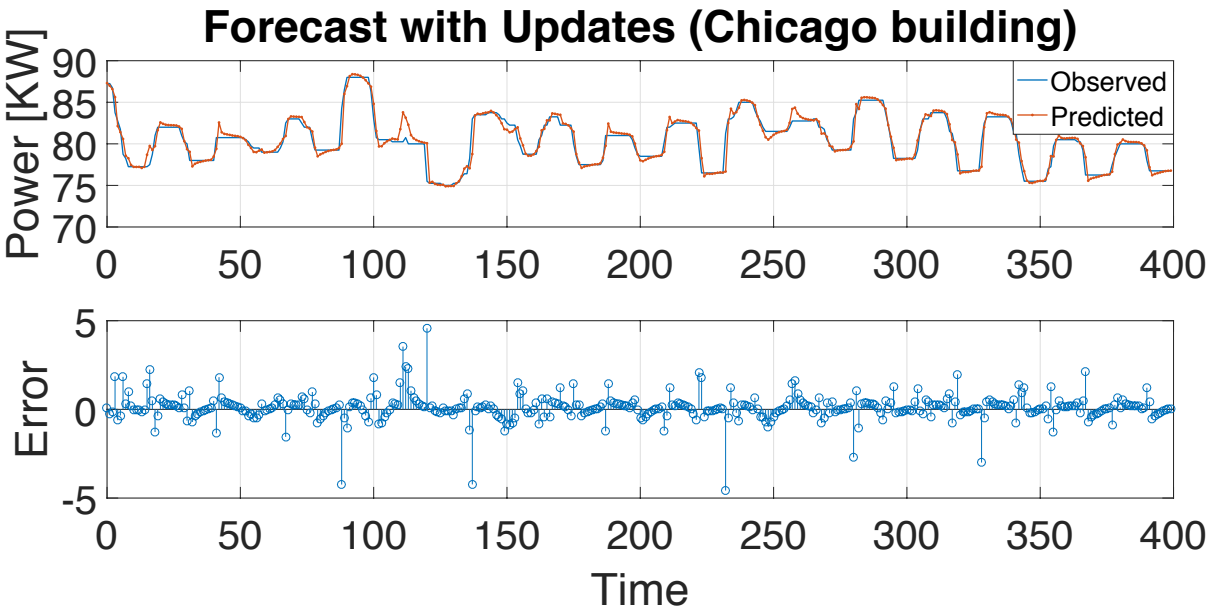


Figura 2 Rezultate de modelare cu rețea LSTM cu 50 de unități ascunse (Chicago)

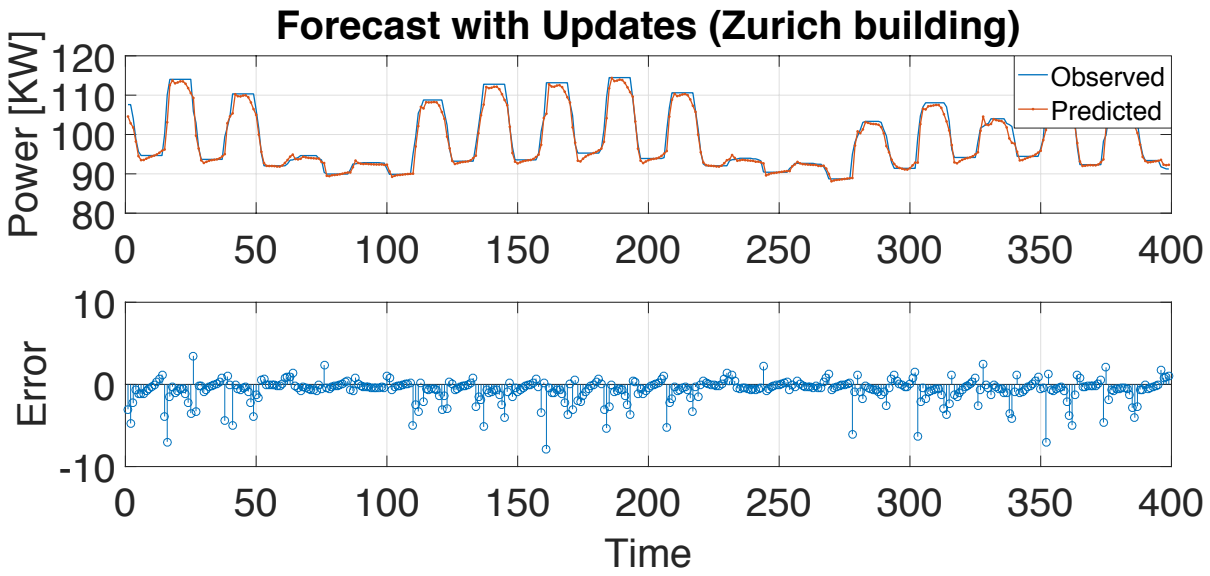


Figura 3 Rezultate de modelare cu rețea LSTM cu 50 de unități ascunse (Zuerich)

Pentru evaluarea modelelor de predicție au fost alese trei metrici de performanță: eroarea medie pătratică (MSE – Mean Squared Error), rădăcina erorii medii pătratice (RMSE – Root Mean Squared Error) și media erorii procentuale absolute (MAPE – Mean Absolute Percentage Error). Acestea sunt definite astfel:

$$MSE = \sum_{1}^{n} \frac{(Y_t - Y_{p_t})^2}{n}$$

$$RMSE = \sqrt{\sum_{1}^{n} \frac{(Y_t - Y_{p_t})^2}{n}}$$

$$MAPE = \frac{1}{n}\sum_{1}^{n}\left|\frac{Y_t - Y_{p_t}}{Y_t}\right| \cdot 100$$

unde n reprezintă numărul de eșantioane, Yt și Ypt reprezintă datele reale și datele de predicție respectiv.

Tabelele 1 și 2 prezintă timpul de calcul, erorile MSE, RMSE și MAPE pe seturile de testare pentru configurațiile LSTM analizate.

	C0	C1	C2	C3	C4
Timp (s)	75	93	143	247	330
MSE	0.6295	0.6132	0.5553	0.7486	0.9555
RMSE	0.7934	0.7831	0.7452	0.8652	0.9775
MAPE (%)	0.5623	0.5091	0.4945	0.5535	0.8177

Tabelul 1 Performanța de predicție LSTM (Chicago)

	Z0	Z1	Z2	Z3	Z4
Timp (s)	76	94	150	275	355
MSE	2.3846	2.1732	2.0506	2.2506	20.34
RMSE	1.5442	1.4742	1.432	1.5002	4.5107
MAPE (%)	0.9197	0.9008	0.828	0.9958	3.4684

Tabelul 2 Performanța de predicție LSTM (Zuerich)

Din punctul de al performanței de testare, după cum este ilustrat în Tabelele 1 și 2, cea mai mică eroare de predicție și cea mai bună precizie pentru ambele clădiri: Chicago și Zuerich se obține în configurația cu 50 de unități ascunse în stratul LSTM. Timpul de calcul este relvant în contextul sistemului de calcul pe care au fost rulați algoritmii cu un procesor quad-core i7 la 3.4GHz și 16GB. Mediul de implementare utilizat a fost MATLAB versiunea R2018a.

Figurile 6 și 7 prezintă evoluția erorii RMSE peste orizontul de antrenare de 200 de epoci pentru cazurile cele mai favorabile și defavorabile. Rezultatul este ilustrat pe setul de date Zuerich în condițiile în care cele mai bune valori se obțin în confiugrația Z2 cu 50 de unități și cele mai slabe rezultate în configurația Z4 cu 125 de unități ascunse. În acest caz se poate vedea cum RMSE oscilează pe orizontul de antrenare datorită sturcturii de rețea prea complexe ce conduce la overfitting. Cel mai bun algoritm converge către valoarea RMSE

minimă în mai puțin de 100 de iterații în timp ce pentru cel mai slab valoarea minimă este atinsă după 120 de iterații.

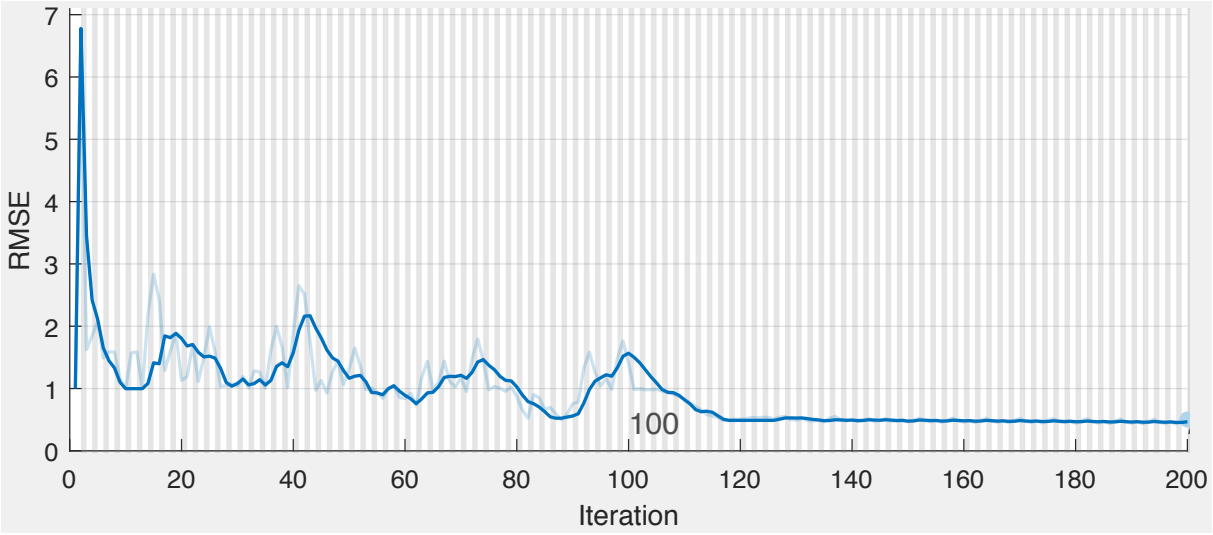


Figura 4 Performanța de antrenare Z4

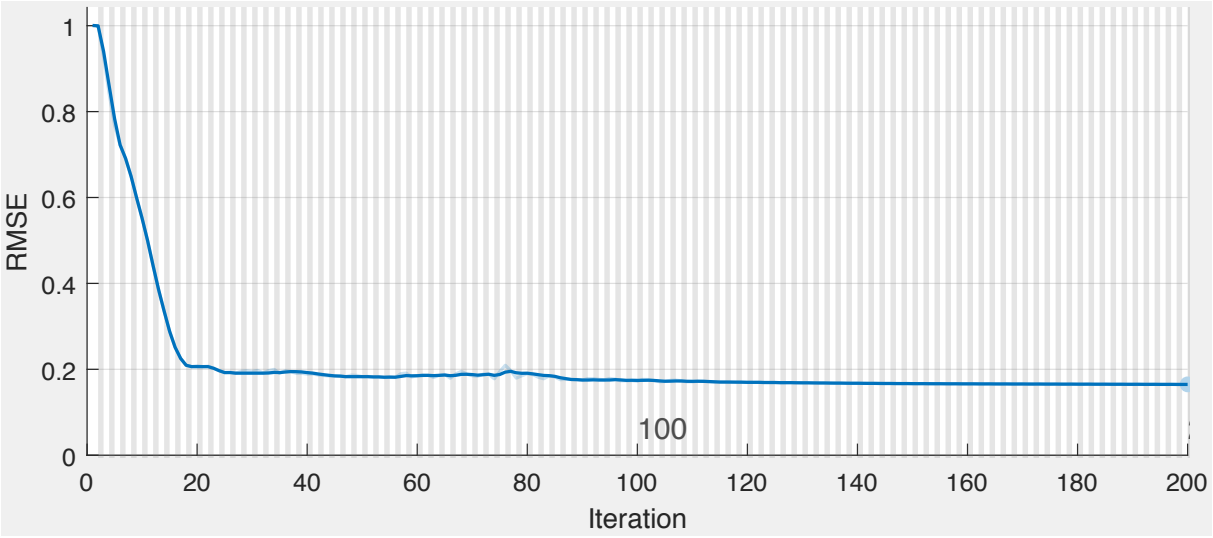


Figura 5 Performanța de antrenare Z2

5. Concluzii

Raportul a prezentat o abordare de aplicare a tehnicilor de învățare bazate pe rețele neuronale adânci la problema modelării și predicției consumului de energie electrică în clădirile comerciale de dimensiuni medii și mari. Tehnica aleasă se bazează pe rețele neuronale recurente cu stratul de unități LSTM de dimensiune variabilă. Studiul de caz prezentat este replicabil în contextul utilizării unor seturi de date de intrare deschise pe două clădiri academice reale. Rezultatele sunt investigate în funcție de complexitatea arhitecturilor de rețea și este identificată o arhitectură potrivită care oferă cea mai bună performanță pe setul de date de intrare, evitând overfitting-ul. Modelele implementate și evaluate se pretează la optimizarea online într-un sistem care asigură colectarea datelor în timp real precum și resursele de calcul adecvate.

Deoarece există o serie întreagă de modele documentate ce pot fi aplicate la problema studiată pentru a extrage indirect informații despre procesele determinante, caracteristicile clădirii și tiparele de consum, utilizarea expertizei specialiștilor de domeniu poate reprezenta de asemenea un ajutor semnificativ. Integrarea ulterioară a acestor modele în sistemele de control ale clădirilor inteligente poate conduce la îmbunătățirea semnificativă a strategiilor de management energetic, în special în cazul entităților de tipul prosumator cu producere locală de energie electrică din surse regenerabile și capacități de stocare.

6. Bibliografie

[1] U. Berardi, “Building energy consumption in us, eu, and bric countries,” *Procedia Engineering*, vol. 118, pp. 128 – 136, 2015, defining the future of sustainability and resilience in design, engineering and construction. [Online]. Available: http://www.sciencedirect.com/science/ article/pii/S1877705815020664

[2] E.Mocanu, P.H.Nguyen, M.Gibescu, and W.L.Kling,“Deeplearning for estimating building energy consumption,” *Sustainable Energy, Grids and Networks*, vol. 6, pp. 91 – 99, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2352467716000163

[3] C. Li, Z. Ding, D. Zhao, J. Yi, and G. Zhang, “Building energy consumption prediction: An extreme deep learning approach,” *Energies*, vol. 10, no. 10, 2017. [Online]. Available: http://www.mdpi.com/ 1996- 1073/10/10/1525

[4] A. Rahman, V. Srikumar, and A. D. Smith, “Predicting electricity consumption for commercial and residential buildings using deep recurrent neural networks,” *Applied Energy*, vol. 212, pp. 372 – 385, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/ pii/S0306261917317658

[5] C. Nichiforov, I. Stamatescu, I. Fagarasan, and G. Stamatescu, “Energy consumption forecasting using arima and neural network models,” in *2017 5th International Symposium on Electrical and Electronics Engineering (ISEEE)*, Oct 2017, pp. 1–4.

[6] C. Nichiforov, G. Stamatescu, I. Stamatescu, I. Fagarasan, and S. S. Iliescu, “Intelligent load forecasting for building energy management,” in *2018 14th IEEE International Conference on Control and Automation (ICCA)*, Jun 2018.

[7] I. Stamatescu, N. Arghira, I. Fagarasan, G. Stamatescu, S. S. Iliescu, and V. Calofir, “Decision support system for a low voltage renewable energy system,” *Energies*, vol. 10, no. 1, 2017. [Online]. Available: http://www.mdpi.com/1996- 1073/10/1/118

[8] P. J. Werbos, “Backpropagation through time: what it does and how to do it,” *Proceedings of the IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct 1990.

[9] R. J. Williams and D. Zipser, “Backpropagation,” Y. Chauvin and D. E. Rumelhart, Eds. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1995, ch. Gradient-based Learning Algorithms for Recurrent Networks and Their Computational Complexity, pp. 433–486. [Online]. Available: http://dl.acm.org/citation.cfm?id=201784.201801

[10] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997. [Online]. Available: http://dx.doi.org/10.1162/neco.1997.9.8.1735

[11] D. L. Marino, K. Amarasinghe, and M. Manic, “Building energy load forecasting using deep neural networks,” *CoRR*, vol. abs/1610.09460, 2016. [Online]. Available: http://arxiv.org/abs/1610.09460

[12] S. Srivastava and S. Lessmann, “A comparative study of lstm neural networks in forecasting day-ahead global horizontal irradiance with satellite data,” *Solar Energy*, vol. 162, pp. 232 – 247, 2018.

[13] C. Miller and F. Meggers, “The building data genome project: An open, public data set from non-residential building electrical meters,” *Energy Procedia*, vol. 122, pp. 439 – 444, 2017.

[14] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *CoRR*, vol. abs/1412.6980, 2014. [Online]. Available: http://arxiv.org/abs/1412.6980