1. **Introducere**
   1. **Contextul lucrării**

Trăim într-o lume dinamică, care este într-o continuă dezvoltare și unde, pe măsura ce populația planetei se află într-o creștere accelerată, la fel este și consumul de energie electrică la nivel mondial**[1]**. Energia electrică este vitală pentru a încuraja dezvoltarea economică și pentru a ușura viețile de zi cu zi ale oamenilor.

O problemă majoră care apare din cauza acestei creșteri a consumului de energie electrică este poluarea. Multe din sursele folosite pentru producerea energiei electrice sunt mari poluatori, cum ar fi petrolul și gazele naturale. Spre exemplu, în Statele Unite ale Americii aproximativ 62.7% din energia electrică produsă este obținută prin arderea de combustibili fosili**[2]**.

Cum multe din activitățile sociale și economice ale omenirii au devenit din ce în ce mai digitalizate, în ultimii ani a apărut nevoia pentru a avea acces la o putere de calcul din ce în ce mai mare. Așa au apărut centrele de date, care nu sunt altceva decât conglomerate de calculatoare care, de obicei, se aflată în aceeași încăpere și care lucrează împreună pentru realizarea de activități diverse. Unele dintre cele mai cunoscute centre de date sunt cele ale marilor companii din domeniul IT, cum ar fi cele de la Google**[3]**, Amazon**[4]** sau Microsoft**[5]** care sunt folosite și pentru a furniza servicii de cloud computing.

Pe măsură ce aceste centre de date cresc ca număr și ca putere de calcul, la fel crește și consumul lor de energie electrică. La momentul actual, centrele de date și rețele de transmisie a datelor consumă aproximativ 1% din energia electrică generată global**[6]**. Acest procent ar fi mult mai mare, deoarece, în ultimii ani consumul de energie al centrelor de date a crescut exponențial, dacă nu s-ar fi fost făcut atât de multe îmbunătățiri în eficientizarea consumului de energie electrică al acestor centre de date.

În prezent, deși s-au făcut foarte multe descoperiri în domeniul eficientizării consumului de energie electrică din centrele de date, acest domeniu este încă în plin avânt, deoarece, din cauza poluării și efectelor ei dăunătoare asupra mediului înconjurător, cercetătorii caută diferite metode prin care pot să eficientizeze consumul de energie și să reutilizeze tot ce se poate reutiliza care urmare a consumului de energie electrică. În cazul centrelor de date, acest lucru poate fi căldura generată în încăpere datorată încălzirii unităților care intră în componența acestor centre de date.

Datorită creșterii exponențiale a numărului de centre de date și a consumului lor de electricitate, a crescut la fel de exponențial și căldura care este generată de ele. Această căldură ar putea fi integrată în instalațiile locale de termoficare pentru a fi folosită de clădirile din vecinătate pentru încălzire. În acest caz, centrele de date ar putea fi considerate și ca producători de energie electrică.

Dacă s-ar putea prezice temperatură generată de un astfel de centru de date, s-ar putea crea un plan pentru includerea acestei călduri în instalațiile locale de termoficare. În acest fel, energia electrică folosită pentru încălzirea apei ar putea fi folosită cu alt scop, maximizând eficiența pe care un centru de date l-ar putea avea ca producător de căldură și consumator de energie electrică.

Din fericire, datorită dezvoltării accelerate ale inteligenței artificiale și în special a deep learning-ului, precum și a numeroaselor date pe care le avem la momentul actual despre consumul de energie electrică a centrelor de date și a căldurii generate de ele, se pot dezvolta modele matematice care să prezică cu exactitate căldura pe care urmează să fie generată de un centru date, acest lucru permițându-ne să creăm acel plan de distribuire a căldurii pe care l-am menționat în paragraful anterior. Această realizare ar duce la o mai bună gestionarea a producției și folosirii de energie electrică, într-un mod mult mai sustenabil.

* 1. **Motivația lucrării**

După cum am specificat în secțiunea anterioară, un centru de date poate fi considerat și ca un producător de căldură. Această căldură ar putea fi folosită pentru încălzirea altor clădiri din apropierea centrului de date sau pentru includerea ei în circuitele locale de termoficare. Acest lucru ar fi posibil doar dacă s-ar putea prezice temperatura generată de către centrul de date, pentru a putea întocmi un plan solid de distribuire a căldurii.

Așadar, motivația principală a acestei lucrări este ilustrarea unor metode de predicție a temperaturii generate de către un centru de date pornind de la un set date existent care conține mai multe informații despre temperatura din centrul de date la anumite intervale de timp, precum și compararea metodelor ilustrate în această lucrare cu alte metode deja existente. Predicția ar trebui să fie cu cât mai aproape de realitate cu putință.

O motivație secundară ar putea fi considerată găsirea unei metode care poate fi generalizată și pentru alte probleme de genul acesta, spre exemplu, predicția consumului de energie pentru un centru de date.

* 1. **Conținutul lucrării**

În aceasta secțiune se va prezenta conținutul acestei lucrări pentru fiecare capitol în parte. Așadar, conținutul lucrării este următorul:

* Capitolul 1 – conține prezentarea contextului și motivației care stau la baza întocmirii acestei lucrări
* Capitolul 2 – conține descrierea obiectivelor principale și secundare ale proiectului
* Capitolul 3 – conține studiul bibliografic a materialelor luate în considerare pentru implementarea soluțiilor discutate în această lucrare
* Capitolul 4 – conține noțiunile teoretice care au fost luate în considerare pentru implementare
* Capitolul 5 – conține exemplificarea detaliilor de proiectare și implementarea soluțiilor alese
* Capitolul 6 – conține ilustrarea rezultatelor matematice și grafice obținute pentru soluțiile alese
* Capitolul 7 – conține setul de instrucțiuni pentru rularea soluțiilor exemplificate
* Capitolul 8 – conține concluziile la care s-a ajuns, precum și eventuale dezvoltări ulterioare

1. **Obiectivele proiectului**

**2.1 Obiectivul principal**

În capitolul anterior am scos în evidență nivelul ridicat de energie electrică pe care centrele de data îl consumă, precum și impactul pe care producția de energie electrică îl are asupra mediului înconjurător. În acest caz, centrele de date pot fi considerate ca mari consumatori de energie electrică. Dar, tot în capitolul anterior, am ilustrat și faptul că centrele de date pot fi considerate și ca producători de căldură, datorită căldurii generate de către componentele care intră în alcătuirea lor. Am văzut cum, teoretic, această căldură ar putea fi folosită pentru încălzirea clădirilor vecine, reducând astfel consumul de energie folosit pentru încălzirea apei din instalațiile de încălzire și, implicit, reducând poluarea și maximizând reutilizabilitatea.

Totuși, pentru a putea folosi această căldură generată de centrele de date este nevoie să știm dinainte câtă căldură poate genera centrul de date respectiv, precum și pe ce perioadă de timp. Este nevoie de aceste cunoștințe pentru a putea întocmi un plan, de care furnizorii de căldură trebuie să fie conștienți. Ei trebuie să știe cât din necesarul de căldură poate să fie acoperit de centrul de date.

Așadar, putem spune că obiectivul principal al acestei lucrări este ilustrarea unor metode de predicție temperaturii generate de către un centru de date. Aceste predicții se vor baza pe un set de date deja existent care va fi descris în capitolele următoare. Toate abordările ilustrate în această lucrare sunt bazate pe algoritmi de deep learning, mai precis, pe rețelele neuronale recurente de tip LSTM (Long short-term memory). Reformulat, obiectivul principal al acestui proiect sună cam așa: **predicția temperaturii dintr-un centru de date folosind rețele neuronale recurente de tip LSTM, pornind de la un set de date, care conține istoricul temperaturilor generate, deja existent**.

**2.2 Obiectivele secundare**

Din obiectivul principal mai rezultă și un set de obiective secundare care vor fi prezentate în această secțiune.

Deoarece o să existe mai multe abordări pentru predicția temperaturii apare și nevoia de a le compara. De asemenea, în această lucrare se încearcă și ilustrarea unei metode generice pentru predicția temperaturii dintr-un centru de date.

***2.2.1 Compararea diferitelor abordări de predicție***

Tot procesul de identificare a unui model potrivit pentru orice tip de predicție se bazează pe un procedeu de încercare a diferitelor metode, a diferiților parametrii pentru algoritmii respectivi, precum și prin preprocesarea a setului de date inițial pentru obținerea de rezultate mai bune.

Pentru acest caz procedura de identificare a unui model potrivit este similară, prin compararea a mai multor abordări. Aceste abordări variază prin configurația rețelei neuronale recurente pe care o folosesc sau prin setul de date folosit (cel inițial sau o variantă preprocesată a sa). Toate rezultatele vor fi comparate prin folosirea unor metrici.

În concluzie, primul obiectiv secundar prezentat poate fi rezumat în următorul fel: **elaborarea unui studiu de comparație între metodele de predicție a temperaturii dintr-un centru de date.**

***2.2.2 Elaborarea unei metode generice de predicție***

Prin încercarea diferitor metode de predicție se va elabora un studiu de comparație între aceste metode. Pe baza acestor rezultate se dorește elaborarea, la nivel teoretic, a unei metode generice de predicție a temperaturii dintr-un centru de date și nu numai.

Prin metodă generică de predicție, se va înțelege o metodă de predicție a temperaturii bazată pe rețele neuronale recurente de tip LSTM care poate fi folosită și pentru alte tipuri de probleme, nu doar pentru centrele de date. Rețele neuronale de tip LSTM sunt foarte bune pentru predicția care se bazează pe date care sunt influențate de datele din trecut (ex. temperatura de afară de azi e influențată de temperatura de afară de ieri). De aceea, abordările prezentate aici se pot preta și pentru alte probleme similare.

Așadar, al doilea obiectiv secundar poate fi formulat după cum urmează: **elaborarea unei metode generice de predicție a temperaturii.**

1. **Studiul bibliografic**

În acest capitol se va prezenta un studiu asupra direcției de cercetare identificată în literatura de specialitate pentru predicția consumului de energie electrică al centrelor de date și, în special, pentru predicția temperaturii generate de centrele de date. Acest studiu pune accentul pe metodele folosite, algoritmii folosiți, caracteristicile luate în considerare, parametrii folosiți, metricile folosite și nivelul la care se face predicția sau predicțiile. S-a încercat identificarea doar a acelor metode care merg pe o abordare care folosește rețele neuronale pentru efectuarea predicțiilor, deoarece comparația cu abordarea ilustrată în această lucrare să fie cât mai relevantă cu putință.

Capitolul curent este împărțit în trei subcapitole care reprezintă cele trei categorii de articole studiate:

1. În prima categorie se studiază metodele de predicție a consumului de energie la un nivel general
2. În cea de-a doua categorie se studiază metodele de predicție a temperaturii generate de un centru de date
3. În ultima categorie se studiază diverse predicții folosind rețele neuronale recurente de tip LSTM

La prima categorie se vor studia trei articole relevante pentru predicția consumului de energie (nu doar energie electrică și nu doar pentru centrele de date). Toate metodele de predicție se bazează pe abordări care folosesc rețele neuronale. Cele trei studii pot fi rezumate în următorul fel:

* În primul articol studiat**[7]** se ilustrează o abordare pentru predicția consumului de energie obținută din petrol în Iordania pe o perioadă de câțiva ani
* În cel de-al doilea articol**[8]** se ilustrează o abordare pentru predicția consumului de energie electrică dintr-un supermarket pe o perioadă de câteva zile
* În cel de-al treilea articol**[9]** studiat aici se prezintă o abordare generică pentru predicția consumului de energie pe o perioadă de câțiva ani

Pentru cea de-a doua categorie se vor studia două articole care ilustrează metode de predicție ale temperaturii dintr-un centru de date bazate tot pe abordări care folosesc rețele neuronale clasice. Aceste articole sunt cele mai relevante din punctul de vedere al subiectului propus de către acest proiect, deoarece setul de date folosit aici este folosit și în metodele ilustrate în cel de-al doilea articol prezentat. Așadar, comparațiile care se vor efectua în această lucrare vor fi făcute în raport cu aceste articole. Ele pot fi descrise pe scurt după cum urmează:

* În primul articol studiat**[10]** se prezintă o abordare de predicție a temperaturii dintr-un centru de date care se folosește o simulare bazată pe dinamica fluidelor
* În cel de-al doilea articol**[11]** studiat se prezintă o abordare similară cu abordarea prezentată în primul articol, diferind prin structura rețelei neuronale propuse

Pentru ce-a de-a treia și ultima categorie se vor studia un număr de patru articole care ilustrează diferite tipuri de predicții bazate pe rețele neuronale recurente de tip LSTM. Predicțiile ilustrate în aceste articole nu sunt predicții pentru temperatură, dar abordările ilustrate aici se pot folosi cu brio și pentru predicțiile de temperatură, cum va fi demonstrat în această lucrare. Pe scurt, articolele pot fi rezumate astfel:

* În primul articol studiat**[12]** ...
* În cel de-al doilea articol**[13]** studiat ...
* În ce de-al treilea articol**[14]** studiat ..
* În ultimul articol**[15]** studiat pentru această categorie ..
  1. **Predicția consumului de energie**

În primul articol studiat se ilustrează a abordare bazată pe rețetele neuronale pentru predicția consumului de energie din Iordania pe perioada 2010-2025. Modelul folosește un se de date colectat între anii 1976 și 2008, setul de date având următoarele caracteristici folosite în predicție: produsul intern brut, populație, exporturile și importurile, consumul de petrol, iar ca variabilă de ieșire consumul de energie. Rețeaua neuronală care a avut cel mai bun scor a fost folosită pentru predicție. S-a observat o creștere majoră a consumului pe perioada de predicție (2010-2025)**[7]**.

În cel de-al doilea articol s-au studiat un număr de patru modele folosite pentru predicția consumului de energie electrică.

* Primul model ilustrat folosește trei variabile de intrare ale consumului de energie electrică și este de tipul MLP
* Al doilea model este un model parțial conectat și are ca variabile de intrare atât consumul de energie, cât și temperatura
* Cel de-a treilea model are ca variabile de intrare atât consumul de energie electrică, cât și numărul de importuri
* Ultimul model folosit are ca variabile de intrare consumul de energie, temperatura și numărul de importuri

Pentru construirea modelelor s-a trecut prin trei stadii: antrenare, testare și evaluare. Algoritmul de antrenare folosit este algoritmul de backpropagation. Acest algoritm permite ca variabilele de intrare să fie transmise prin straturile ascunse până la ieșire, iar eroarea care este calculată la ieșire este propagată înapoi spre intrări. În timpul propagării se ajustează ponderile a fiecărui neuron din rețea. Setul de date folosit pentru antrenare a fost colectat pe o perioadă de câțiva ani și reprezintă consumul energie electrică din Liban între ianuarie 1995 și decembrie 1999. S-a observat că modele care au variabilele de intrare mai diverse au furnizat predicții cu un scor mai bun**[8]**.

În ultimul articol studiat pentru această categorie este prezentată o comparație între predicția consumului de energie folosind rețele neuronale și alte metode de predicție, în special regresie liniară. Aceste predicții s-au realizat pentru sistemele de răcire și de încălzire sau de aer condiționat. Scopul final este dezvoltarea unui sistem care să controleze aceste sisteme pe baza predicțiilor realizate. De asemenea, se mai urmărește identificarea a celor mai importante variabile de intrare pentru facilitarea predicțiilor online, care la rândul lor vor permite implementarea unui sistem de diagnostic, de control al proceselor, de optimizare și de management eficient al energie din supermarket-uri. La finalul articolului ne sunt prezentate rezultatele obținute cu diferite variabile de intrare, precum și o comparație între metodele de predicție alese**[9]**.

Tabel 3.1 Rezumatul studiului pentru predicția consumului de energie

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nr. abordare | Variabile de intrare (caracteristici) | Variabile de ieșire | Observații modele | Metrici folosite |
| 1 | PIB, populație, exporturi, importuri, consumul de petrol | Consumul energie | Diferite tipuri de arhitecturi MLP:  5-5-1, 5-10-1, 5-5-5-1, 5-5-10-1 | Eroarea medie pătratică (MSE) |
| 2 | Consumul de energie din trecut, temperatura, numărul de importuri | Consumul de energie | 4 arhitecturi de rețele neuronale cu diferite variabile de intrare | Eroarea medie pătratică (MSE), deviație medie absolută (MAD), eroarea medie pătratică procentuală (MPSE),  Eroarea medie absolută procentuală (MAPE) |
| 3 | Materialul de construcție, temperatura, umiditatea, mediul intern | Consumul de energie | 7 arhitecturi de rețele neuronale și mai multe tipuri de modele bazate pe regresie liniară | Coeficientul de corelație |

* 1. **Predicția temperaturii generate de către centrele de date**

Pentru prima abordare studiată, centrele de date sunt considerate ca producători de căldură și se explorează posibila lor integrare cu structuri de energie inteligente care să poată refolosi această căldură generată. Tot aici se ilustrează un model al proceselor termoelectrice care apar în interiorul centrelor de date, permițând în felul acesta centrelor de date să își adapteze căldura generată la diferitele cerințe de energie termală. Pe lângă acest model, în cadrul acestei lucrări s-a dezvoltat o serie de simulări bazate pe dinamica fluidelor pentru determinarea parametrilor sistemului de răcire, în așa fel încât să permită creșterea temperaturii în centrul de date fără ca să se pună în pericol operațiunile serverelor sau distribuirea sarcinilor între servere. Acest lucru permite setarea unei mai mari valori ale temperaturii pentru centrul de date pentru perioade scurte de timp folosind mecanisme de pre-încălzire și post-răcire ca mecanisme de adaptare al profilului termic al centrului de date. Pentru reducerea complexității computaționale s-au folosit rețele neuronale antrenate cu rezultatele simulării bazate pe dinamica fluidelor. Experimentele au fost realizate pentru o încăpere 24 mp care conține 24 de servere organizate în 4 rackuri. Rezultatele scot în evidență potențialul centrelor de date de a genera destulă energie termală pentru încălzirea clădirilor din imediata vecinătate**[10]**.

Al doilea articol studiat adresează, la fel ca primul, problema eficientizării costurilor generate de un centru de date prin reutilizarea căldurii produse de către instalațiile de energie locale. Aici s-au studiat cerințele de energie termală și potențialul de furnizare a acestei energii de către un centru de date. Pentru acest lucru s-a definit un model de reutilizare care simulează procesele termodinamice din camera în care sunt situate serverele. Modelul definit permite estimarea, folosindu-se de simulări bazate pe dinamica fluidelor, a temperaturii aerului cald recuperat de pompele de căldură din încăpere, permițându-le acestora să opereze într-un mod mai eficient. Tot ca la prima abordare studiată în acest subcapitol, pentru reducerea complexității computaționale s-a folosit un o rețea neuronală MLP cu scopul de a prezice distribuirea temperaturii aerului cald în camera serverelor. Antrenarea acestei rețele s-a făcut pe baza rezultatelor obținute prin intermediul simulărilor. Pentru obținerea datelor de antrenare s-a modelat o camera virtuală cu un volum de 48 și două rackuri tipice 42U. Rezultatele obținute arată că distribuirea căldurii din camera serverelor poate fi prezisă cu o eroarea mai mică de 1 . În felul acesta, centrele de date pot să estimeze în avans cantitatea de căldură reziduală care poate fi folosită pentru eficientizarea operațiilor pompei de căldură**[11]**.

Tabel 3.2 Rezumatul studiului pentru predicția temperaturii

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nr. abordare | Variabile de intrare (caracteristici) | Variabile de ieșire | Observații modele | Metrici folosite |
| 1 | Munca executată de server, programarea acțiunilor de pe server, fluxul de aer al sistemului de răcire, temperatura inițială, timpul simulării | Temperatura de ieșirii a camerei,  Temperatura din jurul serverului | Rețea neuronală MLP cu N + 3 variabile de intrare, 2 straturi ascunse și N + 1 variabile de ieșire | Eroarea medie pătratică (MSE), eroarea medie absolută procentuală (MAPE) |
| 2 | Temperatura inițială a camerei, temperatura aerului rece, fluxul de aer, temperatura de la diferite rackuri, temperatura aerului cald pentru diverse probe | Temperatura pentru un număr de secunde | Rețea neuronală MLP cu 4 straturi | Eroarea medie pătratică (MSE) |

* 1. **Predicția folosind LSTM**

Pentru prima abordare studiată ..

În cel de-al doilea articol studiat ..

Pentru cea de-a treia abordare studiată la..

În cazul ultimei abordări ...

Tabel 3.1 Rezumatul studiului pentru predicția folosind LSTM

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Nr. abordare | Variabile de intrare (caracteristici) | Variabile de ieșire | Observații modele | Metrici folosite |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

1. **Analiză și fundametare teoretică**
2. **Proiectare de detaliu și implementare**
3. **Experimente și validare**
4. **Manual de instalare și utilizare**
5. **Conluzii și dezvoltări ulterioare**
6. **Glosar**

|  |  |
| --- | --- |
| Termen | Definiție |
| Cloud computing | **-** |
| Deep learning | **-** |
| LSTM | Prescurtarea de la Long short-term memory – un tip de rețele neuronale recurente |
| MLP | Multi layer perceptron |
| Backpropagation |  |

1. **Bibliografie**

[1] IEA, Data and statistics – Energy Consumption, <https://www.iea.org/data-and-statistics?country=WORLD&fuel=Energy%20consumption&indicator=Electricity%20consumption>

[2] EIA, What is U.S. electricity generation by energy source?, <https://www.eia.gov/tools/faqs/faq.php?id=427&t=3>

[3] Google global infrastructure, <https://www.google.com/about/datacenters/locations/>

[4] Amazon global infrastructure, <https://aws.amazon.com/about-aws/global-infrastructure/>

[5] Microsoft Azure global infrastructure, <https://azure.microsoft.com/en-us/global-infrastructure/>

[6] IEA, Data centres and data transmission networks, <https://www.iea.org/reports/tracking-buildings/data-centres-and-data-transmission-networks>

[7] Bassam M. AbuAl-Foul, *Forecasting Energy Demand in Jordan Using Artificial*

*Neural Networks*, 2012

[8] D. Datta, S. A. Tassou, D. Marriott, *Application of Neural Networks for the*

*Prediction of the Energy Consumption in a Supermarket*, 1997

[9] G. E. Nasr, E. A. Badr and M. R. Younes, *Neural Networks in Forecasting Electrical*

*Energy Consumption*, 2001

[10] Marcel Antal, Tudor Cioara, Ionut Anghel, Claudia Pop and Ioan Salomie, *Transforming Data Centers in Active Thermal Energy Players in Nearby Neighborhoods,* MDPI, 2019

[11] Marcel Antal, Tudor Cioara, Ionut Anghel, Radoslaw Gorzenski, Radoslaw Januszewski, Ariel Oleksiak, Wojciech Piatek, Claudia Pop, Ioan Salomie and Wojciech Szeliga, *Reuse of Data Center Waste Heat in Nearby Neighborhoods: A Neural Networks-Based Prediction Model,* MDPI,2019

[12] Hașim Sak, Andrew Senior, Francoise Beaufays, *Long Short-Term Memory Recurrent Neural Network Architectures for Large Scale Acoustic Modeling, Google*, USA

[13] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, *Predicting Stock Prices Using LSTM,* International Journal of Science and Research, 2015

[14] Xuan Zhang, Xun Liang, Aakas Zhiyuli, Shusen Zhang, Rui Xu, and Bo Wu, *AT-LSTM: An Attention-based LSTM Model for FinancialTime Series Prediction*, 2019

[15] Felix A. Gers, Jurgen Schmidhuber, Fred Cummins, *Learning to Forget: Continual Prediction with LSTM*, 1999