

PROGNOZAREA DE DATE PENTRU APLICAȚII SOFTWARE

Maria CERNEI

Universitatea Tehnică a Moldovei, Facultatea Electronică și Telecomunicații, Departamentul Telecomunicații și Sisteme electronice, grupa RST-171, Chișinău, Republica Moldova

Autorul corespondent: Cernei, Maria, maria.cernei@tlc.utm.md

Rezumat. În lucrare este prezentat rezultatul unui studiu de testare și implementare a interfețelor de programare a aplicațiilor software destinate prognozării de date pentru eficientizarea și automatizarea proceselor de funcționare a rețelelor inteligente fotovoltaice și a activităților din sectorul serviciilor publice. A fost demonstrată posibilitatea utilizării modelelor de prognozare „Naive Forecast” și „Seasonal Naive Forecast” pentru date statistice cu și fără periodicitate temporală. Aplicațiile testate și implementate oferă date prognozate cu probabilități care corespund satisfăcător datelor statistice reale.

Cuvinte cheie: prognozare de date, analiză predictivă, aplicații software.

Introducere

Avansările tehnologice care se produc treptat în mai toate domeniile de activitate aduc beneficii companiilor din întreaga lume, care prin investiții au posibilitatea de a eficientiza multe dintre activitățile sale. Aplicațiile software create special pentru soluționarea și optimizarea unor funcționalități au revoluționat complet numeroase sectoare industriale și nu numai, motiv pentru care au devenit atât de populare în lume. Fiecare dintre aceste aplicații sunt dezvoltate cu scopul de a îmbunătăți calitatea serviciilor oferite utilizatorului final, dar și cu scopul de a eficientiza timpul de lucru al angajaților.

Inteligența artificială (AI) reprezintă abilitatea computerelor și mașinilor de a îndeplini sarcini în mod asemănător cu inteligența umană. Un sistem AI trebuie să posedă capacitatea de a raționa, abilitatea de a rezolva unele probleme și de a învăța din experiențe anterioare. Analiza predictivă este un termen care descrie utilizarea datelor din trecut pentru a dezvolta presupuneri în cunoștință de cauză cu privire la rezultatele viitoare.

În lucrare este prezentat un studiu de cercetare a metodelor și modelelor de prognozare a datelor cu scopul eficientizării proceselor și activităților din diverse domenii ale serviciilor publice și industriale. Ca obiectiv a servit identificarea, utilizarea și implementarea unor interfețe de programare a aplicației (API) deja existente pentru prognozarea de date dintr-un set de date actuale statistice cu scopul eficientizării activităților din domeniul serviciilor publice și energetică.

Modele de prognozare a seriilor temporale și non-temporale de date

În cadrul prognozării seriilor temporale de date, datele recente și actuale pot fi folosite pentru a genera predicții despre valorile obiectivelor viitoare. Datele din serii non-temporale pot proveni de la multe surse și conțin valori care nu variază în timp și sunt analizate conform valorii originale [1].

Există mai multe modele de învățare automată în analiza seriilor temporale de date. Cele mai utilizate sunt [2]:

- ARIMA: Reprezintă o combinație de trei modele diferite, „AR”, „MA” și „I”, unde:
 1. „AR” reflectă variabila de interes în evoluție, care este regresată pe propriile sale valori anterioare;
 2. „MA” deduce eroarea de regresie, care este combinația liniară a valorilor termenilor de eroare care au avut loc în diferite etape ale timpului anterior;
 3. „I” - înlocuiește valorile datelor cu diferența dintre valorile lor și valorile anterioare.

- ARCH/GARCH: oferă prelucrarea variațiilor dinamice ale volatilității din seriile temporale.
- Modelul „Vector AutoRegresiv” („VAR”): oferă independențe între diferite date din seriile temporale.
- LSTM („Long-Short Term Memory”): este o rețea neuronală recurentă (RNN) pentru citirea dependențelor de secvență, care creează predicții în funcție de datele anterioare.

Rezultate de testare și implementare a aplicațiilor de prognozare

Prognozarea puterii produse de rețelele inteligente de generare a energiei fotovoltaice oferă posibilitatea de a înțelege modul în care noile regimuri de funcționare ar afecta producția rețelelor inteligente și, de asemenea, de a estima costurile viitoare ale sistemului, veniturile și profitul.

Pentru prognozarea de date a rețelelor fotovoltaice a fost selectat un API existent („Unplugg”), în care sunt implementate modele de prognozare a seriilor temporale. Studiul și testarea a fost realizată pe baza datelor colectate din rețele inteligente reale montate în Republica Moldova, scopul fiind determinarea celor mai bune condiții de utilizare a rețelelor. Mai jos sunt prezentate mai multe cazuri de date statistice și prognozate pentru diferiți agenți, perioadă de timp și ore.

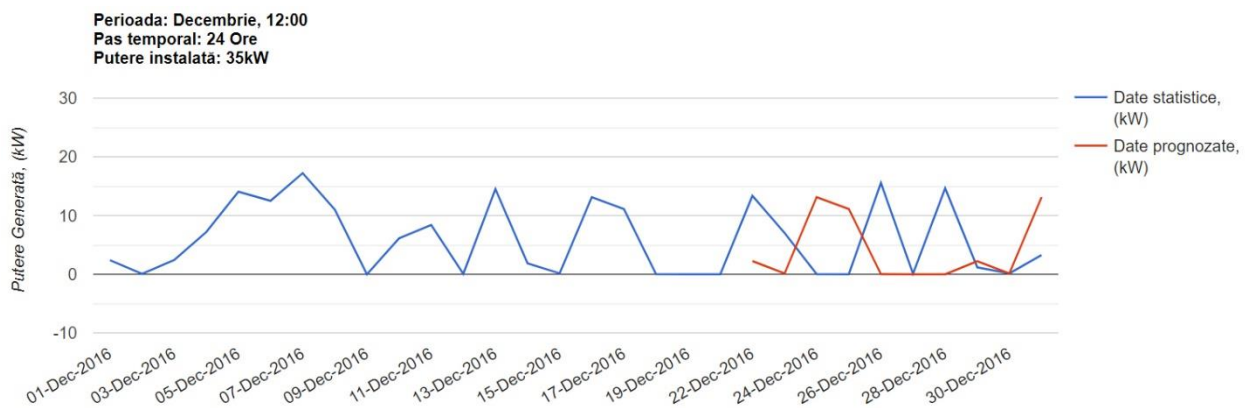


Figura 1. Putere produsă și prognozată pentru o perioadă de 10 zile. (Perioada statistică procesată - 20 de zile anterioare. Perioada prognozată - 10 zile).

În Fig. 1 este reprezentată puterea prognozată pentru o perioadă de 10 zile. 20 de zile anterioare au fost selectate ca perioada de date statistice procesate (payload). Este evident din figură că datele prognozate au o mare discrepanță cu datele statistice reale.

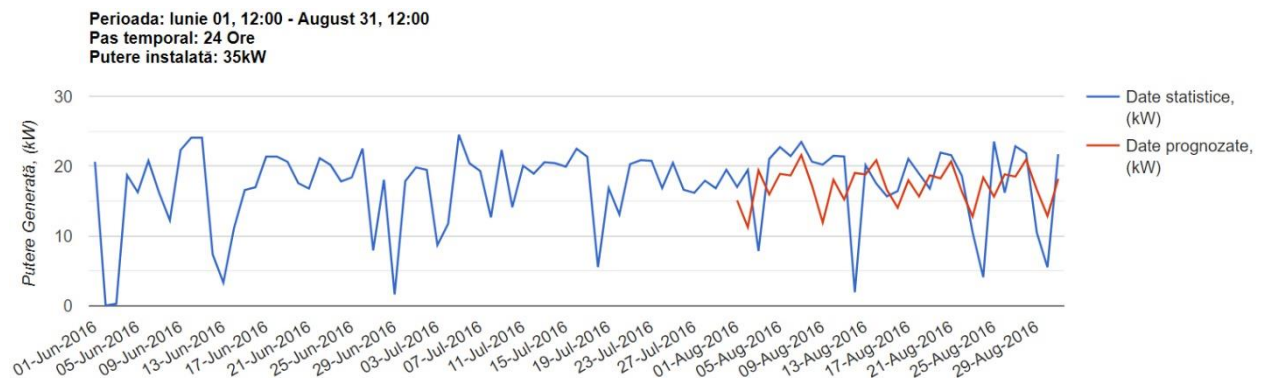


Figura 2. Puterea produsă și prognozată pentru o perioadă de 1 lună (perioada de vară). Perioada statistică procesată - 2 luni anterioare. Perioada prognozată - 1 lună.

În Fig. 2 este reprezentată puterea prognozată pentru o perioadă de 1 lună. 2 luni anterioare au fost selectate ca perioada de date statistice procesate (payload). Valorile datelor puterii prognozate repetă datele statistice reale. Mai mult decât atât, forma dependenței prognozate repetă forma dependenței statistice. Totuși, există valori care diferă considerabil de datele statistice, care pot fi cauzate de condițiile meteorologice la momentul în care au fost efectuate măsurători.

Rezultatele de prognoză obținute sunt satisfăcătoare pentru cazurile, în care perioada de date statistice procesate (payload) nu depășește un interval de 2-3 luni. Utilizarea datelor statistice sezoniere (vară, iarnă) oferă rezultate corespunzătoare celor statistice. Un aspect bun al acestui API reprezintă repetabilitatea și periodicitatea datelor pentru puterea produsă.

În cazul serviciilor sau proceselor în care datele statistice sunt achiziționate în momente de timp diferite, modelele de prognozare a seriilor temporale de date nu oferă rezultate satisfăcătoare, iar cele mai potrivite sunt modelele de prognozare non-temporale cu regresie liniară, așa ca „Naive Forecast” sau „Seasonal Naive Forecast”. Pentru modelul „Naive Forecast” valoarea prognozată reprezintă valoarea celei mai recente observații. Această metodă este deseori utilizată pentru evaluarea performanței prognozelor mai sofisticate. Modelul „Seasonal Naive Forecast” este similar cu modelul „Naive Forecast”, însă, valoarea prognozată reprezintă ultima valoare observată din același sezon al perioadei de timp.

Pentru prognozarea acestui tip de date a fost selectat un API existent („Algorithmia”) și implementat un alt API în 2 versiuni („NA Predict v.3” și „NA Predict v3.1”), cu aplicarea unor modificări și ajustări pentru sporirea veridicității prognozei. Studiul și testarea a fost realizată pe baza datelor colectate în cadrul unor procese și activități din domeniul serviciilor publice, scopul fiind automatizarea proceselor și activității prin înlăturarea componentei umane din proces.

În Fig. 3 sunt prezentate rezultatele de testare a API-ului „Algorithmia”. Setul de date procesat pentru prognozare (payload) a fost constituit din date statistice selectate independent de ora și ziua săptămânii în care au fost înregistrate. În Fig. 3, A este prezentat numărul de cazuri statistice pentru modificarea Valorii 3 în Valorile 1, 2 sau 3, iar în Fig. 3, B este prezentat rezultatul prognozării de modificare a Valorii 3 în Valorile posibile 1 sau 2 (probabilitatea 0.2 și 0.8, respectiv), pentru o anumită zi și oră selectată aleatoriu. Rezultatul obținut după testarea API-ului „Algorithmia” nu este destul de satisfăcător, deoarece nu reflectă în mod dorit rezultatele statistice.

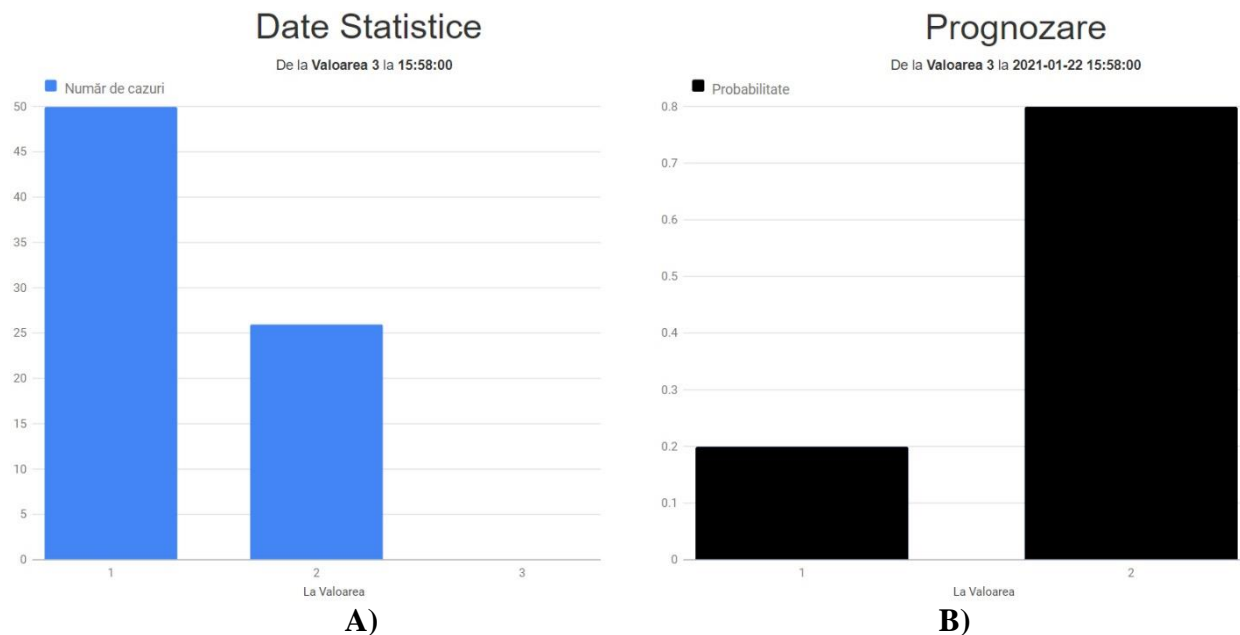


Figura 3. Numărul de cazuri statistice procesate (A) și rezultatul prognozării (B) pentru API-ul „Algorithmia”.

Rezultatele obținute au demonstrat necesitatea utilizării în API a unui model de tip „Seasonal Naive Forecast”. Datele statistice procesate în acest API („NA Predict v3.1”) corespund orei și zilei din săptămâna selectată aleatoriu și orelor adiacente într-un interval de ± 2 ore (Fig. 4, A).

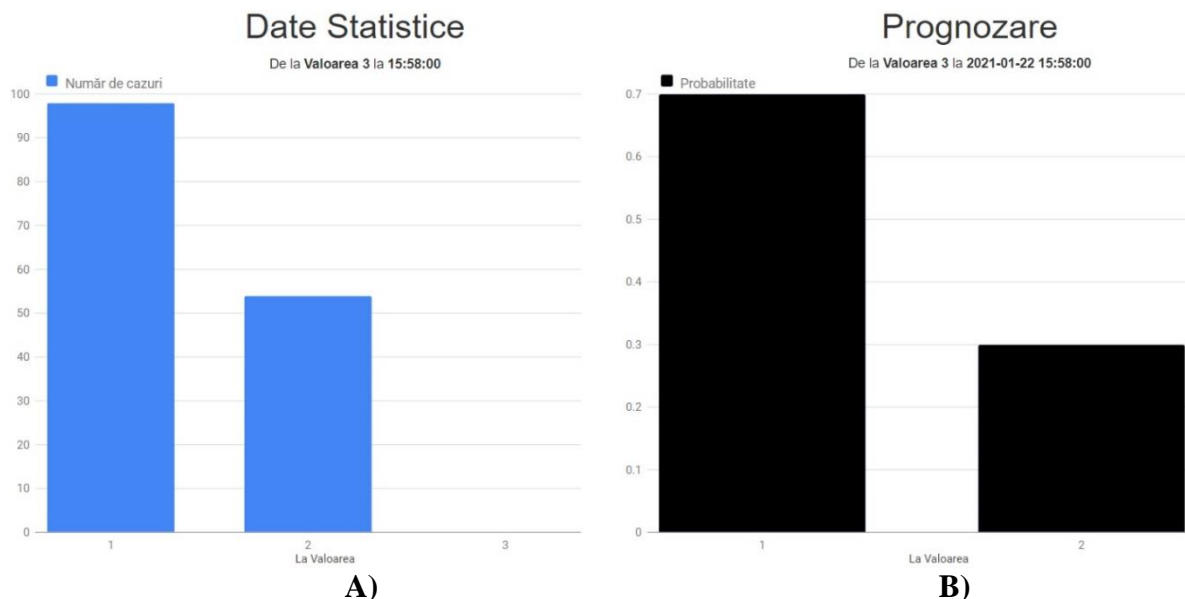


Figura 4. Numărul de cazuri statistice procesate (A) și rezultatul prognozei (B) pentru API-ul „NA Predict v.3.1”.

Rezultatele obținute în acest caz corespund destul de satisfăcător datelor statistice procesate cu o probabilitate de 0.7 și 0.3 pentru Valorile posibile 1 și 2, respectiv (Fig.4, B). În cadrul acestui API s-a utilizat un set de date procesate (payload) mult mai variat și extins, dar și o diferență între probabilități mult mai esențială. Acest API („NA Predict v3.1”) s-a dovedit a fi promițător pentru utilizare în cadrul automatizării activității și proceselor propuse.

Concluzii

Prognozarea de date a rețelelor de generare a energiei fotovoltaice a demonstrat un rezultat satisfăcător pentru perioade de date statistice procesate (payload) care nu depășește intervalul de 2-3 luni. Utilizarea datelor statistice procesate sezoniere (vară, iarnă) oferă rezultate corespunzătoare celor istorice, datorită dependenței lor de condițiile meteorologice.

Interfețele de programare a aplicației (API) în care se utilizează modelele de prognozare „Naive Forecast” și „Seasonal Naive Forecast” s-au dovedit a fi mai potrivite pentru procesarea datelor statistice fără periodicitate temporală (non-temporale). Utilizarea API-ului implementat „NA Predict v3.1” poate asigura automatizarea și optimizarea proceselor și activităților din domeniul serviciilor publice. Implementarea aplicațiilor software de prognozare poate contribui la îmbunătățirea calității produselor și serviciilor din diverse domenii ale economiei și industriei prin excluderea componentei umane din proces, automatizarea și eficientizarea multiplelor procese tehnice și tehnologice.

Referințe

1. *ThingWorx Analytics, ThingWorx Foundation Help Center*. [online]. [accesat 16.09.2020]. Disponibil: https://support.ptc.com/help/thingworx_hc/thingworx_analytics_8.
2. *Forecasting: Principles and Practice*, Monash University, Australia. [online]. [accesat 23.08.2020]. Disponibil: <https://otexts.com/fpp2/>.