**Utilizarea retelelor de senzori si modelelor bazate pe Inteligenta Artificiala in automatizarea industriala**

[1 Introducere 2](#_Toc473097977)

[2 Cercetare – Faza 1 3](#_Toc473097978)

[2.1 Analiza stadiului actual al tehnologiei 3](#_Toc473097979)

[2.2 Modelul matematic utilizat pentru optimizarea sistemului 4](#_Toc473097980)

[2.2.1 Inferenta in reteaua neurala adanca propusa 8](#_Toc473097981)

[2.2.2 Antrenarea automatizata a retelei neurala adanci 9](#_Toc473097982)

[2.3 Algoritmul sistemului integrat de retea senzori si Inteligenta Artificiala 11](#_Toc473097983)

[2.4 Mediul de executie experimental 12](#_Toc473097984)

[3 Dezvoltare si experimente – Faza 2 14](#_Toc473097985)

[3.1 Modelul experimental aplicat 14](#_Toc473097986)

[3.2 Rezultatele testelor experimentale 15](#_Toc473097987)

[4 Referinte 16](#_Toc473097988)

# Introducere

Analiza informatiilor cu ajutorul retelelor de senzori este un important subiect atat de cercetare-dezvoltare in cadrul academic cat si cu aplicare directa in aplicatii de productie din mediul comercial. Daca retelele de senzori au o istorie de peste 50 de ani, in conditiile in care prima retea de senzori wireless (SOSUS) care poarta asemanari cu retelele moderne de senzori a aparut in anii ’50 dezvoltata de armata Statelor Unite cu scopul de a detecta si urmari submarinele sovietice [1], abia in ultimii 10 ani puterea de analiza si coordonare data de retelele de senzori a inceput sa devina cu adevarat impresionanta cu ajutorul evolutiei uriase a tehnicii de calcul in termeni de procesare, stocare, retragere, comunicare. Odata cu evolutia tehnologiei de calcul – in special tehnologia de calcul paralel si disponibilitatea acesteia la scara larga si la nivele de cost acceptabil – domeniul Inteligentei Artificiale a luat o amploare deosebita cu un accent deosebit in zona de predictie bazata. Fie ca este vorba de algoritmi genetici, machine learning clasic, algoritmi genetici de machine learning sau deep learning, toate aceste modele si tehnologii ce au ca scop determinarea unei predictii/inferente in urma unui proces de invatare stau actualmente la baza stadiului actual al tehnologiei in domeniul Inteligentei Artificiale. In final, cele doua domenii – cel al retelelor avansate de senzori si cel al Inteligentei Artificiale – se unesc in punctul in care apar atat in mediul academic cat si in mediul comercial din ce in ce mai multe proiecte atat aplicabile in domeniul Internetului Lucrurilor (Internet-of-Things) cat si in alte domenii. Astfel, propunerea noastra concreta consta in utilizarea retelelor de senzori si a modelelor de Inteligenta Artificiala bazate pe Retelele Neurale Artificiale Adanci pentru optimizarea si automatizarea avansata a depozitelor de distributie. Aceasta propunere inovativa combina capacitatea de colectare si centralizare automatizata in timp real a datelor telemetrice data de o retea de senzori de arhitectura proprie cu un model inovativ de retea neurala artificiala adanca propus de noi in baza celor mai noi cercetari din domeniul Deep Learning. Rezultatul va consta intr-un prototip de sistem avansat inovativ auto-adaptabil ce va fi capabil sa optimizeze activitatea in timp real in cadrul unui depozit industrial.

# Cercetare – Faza 1

## Analiza stadiului actual al tehnologiei

Scurta prezentare a problematicii depozitelor industriale (distributie, retail, etc)

Scurta prezentare a celui mai modern domeniu de Inteligenta Artificiala: Deep Learning

Analiza stadiul actual al tehnologiei plecand de la adaptarile moderne ale "travelling salesman" pana la modele deep learning

## Modelul matematic utilizat pentru optimizarea sistemului

In cele ce urmeaza vom face prezentarea modelului matematic principial de analiza a informatiilor si inferenta a deciziilor bazat pe invatarea cu propagare inversa in retele neurale adanci. Conform descrierii anterioare acest model va primi datele generate de reteaua de senzori, va combina aceste date cu datele introduse de operatorii umani (in cazul existentei acestora) si va realiza urmatoarele operatii: (1) auto-adaptare prin invatare in timp real (invatare on-line) a modificarii parametrilor de functionare a retelei de senzori sau a altor elemente din cadrul mendiului de productie (depozit) (2) generarea in timp real de predictii/inferente referitoare la succesiunea optima de operatii necesare pentru atingerea unui anumit scop – concret in cazul automatizarii unui depozit este necesara inferenta de: (a) rute optime la nivel de utilaj mobil, (b) sub-rute optime la nivel de operator uman, (c) optimizarea pick-list, (d) alegerea operatorului uman ideal pentru o anumita structura de pick-list, s.a.m.d.

Concret putem considera ca intregul sistem poate fi sintetizat sub forma unei singure functii non-liniare cu complexitate foarte mare care ia forma din Ecuatia (AI.1)

|  |  |
| --- | --- |
| ) | (AI.1) |
|  | (AI.2) |
|  | (AI.3) |

Astfel intreaga abordare se reduce la determinarea unui set de parametrii Ɵ ai functiei-ipoteza care sa determine un timp rezultat de proces minim (valoare reala continua) printr-o retea neurala adanca de regresie (figura AI.F1) precum si alti parametrii ca , , , s.a.m.d, de aceasta data utilizand o retea adanca de clasificare (figura AI.F2).

In vederea obtinerii acestor parametrii ai retelelor neurale vom determina o functie de calcul al erorii (ecuatia *(*AI.*2)* pentru cazul reteli neurala de regresie) care sa se poata minimiza conform ecuatiei *(*AI.*3)*. In vederea determinarii parametrilor optimi vom aplica un algoritm de cautare a acestora in spatiul vectorial de parametrii prin aplicarea metodei “coborarii” prin gradientele secventiale, metoda ideala pentru mediile de lucru online si in timp real conform ecuatiilor (AI.4) si (AI.5).

|  |  |
| --- | --- |
|  | (AI.4) |
|  | (AI.5) |

Figura AI.F – Structura grafica (naiva) a retelei neurale adanci

Revenind la functia non-liniara ipoteza – punctul central al modelului matematic propus – aceasta va fi definita de o retea neurala adanca a carei structura generica este descrisa in figura (AI.F1) unde primul nivel (intrare) de culoare albastra primeste datele de intrare de la reteaua de senzori, nivelele ulterioare desenate cu gri contin unitatile ascunse ale retelei neurale adanci iar ultimul nivel (verde) contine datele de iesire ale retelei neurale adanci. Din motive de prezentre grafica au fost omise din figura atat numarul corect de nivele ascunse si unitatile individuale din fiecare nivel precum si conexiunile intre diversele nivele ale retelei neurale adanci (singurele conexiuni prezentate sunt cele dintre nivelul de intrare si primul nivel ascuns). O imagine mai elocventa a retelei neurale adanci utilizate in cazul nostru este prezentata in figura (AI.F2) pentru unul din cazurile de arhitectura propuse in care avem de-a face cu urmatoarea structura a modelului:

* Un total de 8 intrari in retea
* 3 nivele total conectate ascunse ce utilizeaza activatori neurali ReLU
* Un nivel de clasificare a datelor de intrare cu 4 iesiri de tip Softmax conform ecuatiei (AI.7)

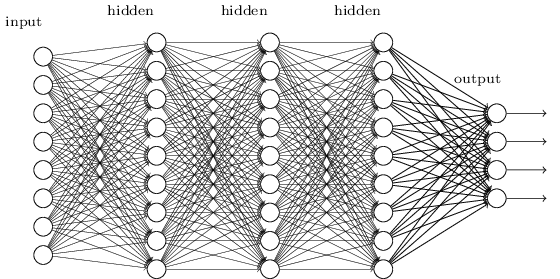


Figura AI.F 2 – Retea neurala adanca de clasificare

De remarcat ca spre deosebire de reteaua neurala adanca utilizata la evaluarea timpului (valoare reala continua) pentru analiza elementelor ca “operator optim”, “ambalaj optim”, s.a.m.d. este necesare utilizarea clasificarii care sa determina elementul cu predicitie maxima conform formulei (AI.7)

Modelul de propagare a informatiilor prin reteaua neurala se bazeaza in principal pe utilizarea unitatilor rectificate liniar (ReLU) [2] ce au proprietatea de a fi stabile in procesele de calcul numeric necesare pentru antrenarea retelei adanci. Astfel functia ReLU descrisa in detaliu in ecuatia (AI.6) va “activa” rezultatele obtinute la nivelul anterior al retelei neurale adanci dupa ce acestea au fost multiplicate cu parametrii nivelului actual.

|  |  |
| --- | --- |
|  | (AI.6) |
|  | (AI.7) |

### Inferenta in reteaua neurala adanca propusa

Concret seria de calcule matematice necesare pentru propagare inainte necesara predictiei valorii reale a timpului de procesare a comenzii in depozit este data de pasii prezentati mai jos. De mentionat faptul ca inferenta parametrilor aditionali ai unui proces complet in depozit (parametrii cum ar fi operatorul optim, structura de impachetare, etc) se determina prin reteaua neurala adanca ce se bazeaza pe nivelul de iesire softmax (clasificare). Ambele retele (exceptand nivelul de iesire) se bazeaza pe seria de ecuatii propuse mai jos is bazate pe ecuatia (AI.6).

De mentionat este faptul ca retelele clasice neurale utilizeaza functii de activare de tip sigmoid in timp ce majoritatea modelelor state-of-the-art in domeniul Deep Learning utilizeaza functia de activare ReLU datorita proprietatii acesteia de a fi stabila numeric pentru valori foarte mari (negative) ale parametrului z. O potentiala variatie a aplicarii modelului propus consta in utilizarea activarii neurala sigmoid pentru reteaua neurala adanca de clasificare si utilizarea activarii ReLU pentru reteaua neurala adanca de regresei. Pentru diversitate vom prezenta propagarea inainte in reteaua de regresie bazata pe activarea neurala ReLU si propagarea inapoi (invatarea retelei neurala adanci) in varianta utilizarii activarii neurala sigmoid la nivelul neuronilor din nivelele ascunse.

Calculul la nivelul de intrare 1 in reteaua neurala adanca de regresie:

(AI.8)

Calculul propagarii inainte in reteaua neurala adanca la nivelul ascus 2:

(AI.9)

(AI.10)

Calculul propagarii inainte in reteaua neurala adanca la nivelul ascus 3:

(AI.11)

(AI.12)

Calculul propagarii inainte in reteaua neurala adanca la nivelul ascus 4: (AI.13)

(AI.14)

Calculul propagarii inainte in reteaua neurala adanca la nivelul de iesire (5):

(AI.15)

(AI.16)

(AI.17)

### Antrenarea automatizata a retelei neurala adanci

Antrenarea modelului de retea neurala artificiala adanca prin propagarea inversa se face dupa principiile prezentate anterior bazate pe ecuatiile (AI.4) si (AI.5). In vederea simplificarii prezentarii vom prezenta cazul particular al calcului derivatelor partiale necesare determinarii directiei de modificare graduala a ponderilor axonilor pentru neuronii din nivelul 3 dupa cum urmeaza:

|  |  |
| --- | --- |
|  | (AI.18) |
|  | (AI.19) |
|  | (AI.20) |
|  | (AI.21) |
|  | (AI.22) |
|  | (AI.23) |
|  | (AI.24) |
|  | (AI.25) |
|  | (AI.26) |
|  | (AI.27) |

## Algoritmul sistemului integrat de retea senzori si Inteligenta Artificiala

In continuare este prezentat algoritmului general de functionare al intregului ansamblu inovativ bazat pe fluxului general de date.

**Algoritm 1** – ***Algoritmul general de functionare al sistemului***

1. **Intrarea datelor in sistem**
   1. **Reteaua de senzori genereaza informati *telemetrie***
   2. **Utilizatorii introduc *picklist*-ul si *optiunile***
2. **Predictie si inferenta**
   1. **Modelul neural adanc de regresie determina *timpul* optim bazat pe *picklist si optiuni***
      1. **Matricile de predictori sunt trimise (distribuite) prin off-load in unitatea de procesare masiva paralela de unde se primesc rezultatele (Mapare)**
      2. **Modelul neural adanc centralizeaza rezultatele primite din unitatea de procesare masiva paralela (Reducere)**
   2. **Modelul neural adanc de clasificare genereaza inferenta parametrilor de proces bazat pe *picklist*, *optiuni* si *timpul* optim prezis**
      1. **Matricile de predictori sunt trimise (distribuite) prin off-load in unitatea de procesare masiva paralela de unde se primesc rezultatele (Mapare)**
      2. **Modelul neural adanc centralizeaza rezultatele primite din unitatea de procesare masiva paralela (Reducere)**
   3. **Datele generate in etapa de predictie si inferenta sunt trimise catre operatorul uman**
3. **Executia procesului**
   1. **Reteaua de senzori impreuna cu terminalele mobile genereaza telemetrie**
   2. **Se masoara eroarea de regresie si clasificare generata in etapa de predictie si inferenta**
   3. **Se re-antreneaza cele doua retele neurale adanci cu rezultatele din mediul de productie**
      1. **Pentru fiecare retea se propaga inapoi plecand de la erorile de clasificare si regresie eroarea in vederea ajustarii ponderilor dendritelor**
         1. **Matricile de inferente, erori si ponderi sunt trimise catre unitatea de procesare masiva paralela de catre fiecare model (Mapare)**
         2. **Fiecare model neural adanc centralizeaza rezultatele primite din unitatea de procesare masiva paralela (Reducere)**
4. **Se reia pasul 1 pentru un nou proces**

## Mediul de executie experimental

Mediul de executie experimental pentru amsamblul inovativ propus este total adaptat pentru procesarea informatiilor si generarea de decizii in timp real.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | ***SI1***: Modul de achizitie a datelor de la reteaua de senzori si telemetria dispozitivelor mobile | | ***UI1***: Modul de introducere date utilizator | | ***AI1***: Modul de predictie a timpilor de procesare | ***AI2***: Modul de inferenta a parametrilor de proces optimi | | | ***MP1***: Modul de procesare paralela pentru executia in timp reala a task-urilor din ***AI1*** si ***AI2*** | | |   Figura AI.F 2 – Structura generala a sistemului |

Arhitectura generala a sistemului experimental prezentat in figura (AI.F3) se bazeaza pe 3 categorii de module principale dupa cum urmeaza:

1. Module de achizitie responsabile de coletarea datelor:
   1. Sistemul de colectare a datelor automate (retea de senzori) (SI1)
   2. Sistem de colectare date de la statii operate uman (UI1)
   3. Sistem de colectare telemetrie de pe dispozitive mobile (SI1)
2. Module bazate pe Inteligenta Artificiala ce determina inferente si predictii ale operatiunilor optime
   1. Modul de predictie a timpilor de procesare (AI1)
   2. Modul de inferenta a parametrilor de proces optimi (AI2)
3. Module de suport al executiei proceselor in mediu de procesare masiva paralela - sistem de tip High Performance Computing de procesare masiva paralela prin utilizarea de procesoare de calcul stiintific de tip GPU

# Dezvoltare si experimente – Faza 2

## Modelul experimental aplicat

Faza 2

## Rezultatele testelor experimentale

Faza 2

# Referinte

|  |  |
| --- | --- |
| [1] | I. Silicon Laboratories, "The Evolution of Wireless Sensor Networks," Silicon Labs, 2013. |
| [2] | R. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas and H. Seung, "Digital selection and analogue amplification coesist in a cortex-inspired silicon circuit," *Nature,* no. 405, p. 947–951, 2000. |