



Universitatea Tehnica din Cluj-Napoca

Facultatea de Automatică și Calculatoare

Departamentul de Automatică

2023

Propunere de proiect pentru admiterea la studii de master

1. Date personale ale candidatului:

1.1. Nume	Stoica
1.2. Prenume	Marius-Adrian

2. Date referitoare la forma de învățământ absolvită de candidat:

2.1. Instituția de învățământ	Universitatea Tehnică Cluj-Napoca
2.2. Facultatea	Automatică și Calculatoare
2.3. Specializarea	Calculatoare
2.4. Anul absolvirii universității	2020

3. Titlul propunerii de cercetare (în limba română):

Descoperirea infecțiilor bacteriene din spitale folosind senzori electrochimici și o soluție bazată pe machine learning

4. Titlul propunerii de cercetare (în limba engleză):

Finding bacterial infections in hospitals using electrochemical sensors and a solution based on machine learning

5. Termeni cheie:

1	Senzor electrochimic
2	Microcontroller (NodeMCU)
3	Rețea neuronală artificială
4	Aplicație mobile
5	Bază de date non-relatională (Firebase)

6. Durata proiectului 2 ani.

7. Prezentarea propunerii de cercetare:

[ANEXA 1]

8. Date referitoare la lucrarea de licență:

8.1. Titlul lucrării de licență:

Detectia codului numeric personal folosind o soluție hibridă bazată pe rețele neuronale

8.2. Rezumatul lucrării de licență:

Scopul lucrării de licență a fost implementarea unui sistem, care preia o imagine a cărții de identitate, localizează cifrele codului numeric personal, și clasifică fiecare cifră, folosind o rețea neuroanală convolutională. Cifrele sunt detectate, iar rezultatele convoluțiilor aplicate pe acestea, sunt trimise la placa de dezvoltare Basys3, folosind protocolul de transmisie UART. Rețeaua neuronală artificială implementată pe placa FPGA primește atât rezultatele intermedii, cât și modelul rețelei, și clasifică fiecare cifră în parte.

Soluția propusă implică implementarea unei rețele neuronale artificiale în VHDL. Input-ul și modelul rețelei sunt trimise de la calculator la placa Basys3, folosind protocolul UART. De asemenea, a fost creată o aplicație pentru automatizarea procesului de transmitere a input-ului și modelului rețelei, la placa de dezvoltare FPGA. Valorile, reprezentate în fixed point, au fost salvate în două memorii BRAM distincte, una pentru input-uri, alta pentru weights-uri. Rețeaua neuronală artificială a clasificat, rând pe rând, fiecare cifră din CNP. Această rețea neuronală artificială a fost parte a rețelei neuronale convoluționale, antrenată pe calculator, menită să clasifice cifrele din CNP. Rețeaua neuroanală convolutionala are 3 layere de convoluție, intercalate de 2 layere de max-pooling, și o rețea neuronală artificială, complet conectată, alcătuia dintr-un input layer și un output layer. De asemenea, a fost creată o aplicație desktop, menită să îmbunătățească experiența utilizatorului sistemului. Astfel, utilizatorul are posibilitatea de a se înregistra, respectiv autentifica în aplicație. Utilizatorul poate alege o imagine cu actul de identitate, pe care să o trimită spre procesare, pentru a detecta cifrele care formează CNP-ul. Imaginea editată, cu cifrele CNP-ului încadrate, poate fi salvată. Informații privind utilizatorul, imaginile salvate de acesta sau istoricul evenimentelor din aplicația desktop sunt stocate folosind o baza de date relațională.

În ceea ce privește rezultatele obținute, cifrele care formează codul numeric personal au fost detectate și încadrate, folosind un boundary box. Rețeaua neuronală artificială clasifică cele 13 cifre ale CNP-ului. CNP-ul este afișat pe SSD-ul (Seven Segment Display) plăcii de dezvoltare. Cifrele sunt afișate în zecimal, câte 4 cifre deodată, fiind shiftate cu o poziție la stânga, la fiecare 2 secunde. De asemenea, există posibilitatea afișării celor 10 output-uri pentru fiecare din cele 13 cifre ale CNP-ului, folosind ledurile și un numărător. Aceste valori pot fi comparate cu cele obținute de rețeaua neuronală convoluțională.

Testarea rezultatelor rețelei neuronale artificiale s-a făcut în două moduri. În primul rând, au fost testate rezultatele fiecărui output din cele 10 output-uri posibile, pentru fiecare cifră din CNP, din cele 13 cifre posibile. Așadar, un total de 130 de output-uri au fost salvate într-un array. Valorile acestor output-uri au fost afișate pe led-uri, folosind un counter, și comparate cu valorile prezise de rețeaua neuronală în Python. Rezultatul clasificării cifrelor din CNP a fost afișat pe SSD, putând fi comparat cu CNP-ul din imaginea din actul de identitate. După ce a avut loc antrenarea rețelei neuronale convoluționale, s-a încercat o predicție pe baza imaginilor de test. S-a dovedit că rețeaua neuronală convolutională are o acuratețe de 98%. De asemenea, a fost creat un confusion matrix. Aceasta relevă de câte ori au fost prezise corect valorile, fiind o posibilitate bună de testare. În ceea ce privește detecția cifrelor din CNP, algoritmul OCR (Optical Character Recognition) a preluat, pe lângă caractere, o serie de informații precum coordonatele x,y ale caracterului, width-ul și height-ul acestuia. Folosind aceste informații, a fost posibilă localizarea cifrelor și încadrarea lor într-un bounding box pentru testare.

Principalele contribuții aduse proiectului de licență au fost în ceea ce privește rețeaua neuronală artificială în VHDL. Modalitatea de transmitere a datelor de la calculator la placă, reprezentarea numerelor pe 16 biți în fixed point, salvarea input-urilor și weights-urilor în memorii BRAM, clasificarea cifrelor care formează CNP-ul folosind un automat cu stări finite, implicit o abordare mai degrabă sevențială decât concurrentă, face ca aplicația VHDL să aibă o amprentă proprie.

9. Activitatea științifică a candidatului:

[ANEXA 2]

DATA: 18.07.2023

TITULAR DE PROIECT,
Nume, prenume: **Stoica Marius-Adrian**
Semnatura:



7. Prezentarea programului de cercetare:

7.1. STADIUL ACTUAL AL CUNOAȘTERII ÎN DOMENIU PE PLAN NAȚIONAL ȘI INTERNAȚIONAL, RAPORTAT LA CELE MAI RECENTE REFERINȚE DIN LITERATURA DE SPECIALITATE.

Resursele bibliografice găsite, sub forma lucrărilor de disertație [2], articolelor științifice [1] și a articolelor de pe site-uri [3], au ajutat în crearea pipeline-ului (flow-ului) pe care doresc să îl implementez în încercarea de a detecta bacteriile din spitale. În cele ce urmează, voi prezenta, succint, fiecare din cele trei resurse bibliografice alese, obiectivul lucrării, și modul în care a folosit în crearea pipeline-ului propus de mine.

În [1] se prezintă ideea că artificial intelligence și machine learning pot fi folosite pentru a detecta bacteriile din spitale. Algoritmii de machine learning sunt mai performanți decât abordarea statistică, convențională. Pentru antrenare, poate fi folosită învățarea supervizată (supervised learning), învățarea nesupervizată (unsupervised learning) și reinforcement learning. În proiectul meu, voi folosi învățarea supervizată, folosindu-mă de datele de la UMF pentru a antrena rețeaua neuronală. În [2] este prezentată o soluție pentru detecția bacteriilor din spitale, bazată pe învățarea supervizată. Prezenta lucrare folosește algoritmi de machine learning (support vector machine, gradient tree boosting) pentru a detecta bacteriile. S-au folosit rapoarte ale pacienților, în format text, din care au fost preluate informații utile algoritmilor, sub forma de feature-uri. Dataset-ul a fost alcătuit pe baza a 300 de rapoarte. Ideea de bază pe care am preluat-o din [2] este că, pe lângă rețelele neuronale artificiale, pot folosi algoritmi de machine learning (support vector machine, gradient tree boosting) pentru a detecta bacteriile.

Resursa [3] folosește senzorul electrochimic pentru a testa bacteriile din probele prelevate de la pacienți. Se aplică proba pe senzorul electrochimic, iar output-ul este sub forma unui waveform. Acest waveform diferă în funcție de bacteria din proba aleasă. Acest aspect va fi folosit și în implementarea pe care am propus-o, dorind ca din acel waveform, după data cleaning, să obțin feature-uri care pot fi fit-uite în rețeaua neuronală artificială.

7.2. OBIECTIVELE PROIECTULUI

Obiectivul principal al proiectului este detecția rapidă a bacteriilor din spitale, folosind senzori electrochimici. Se va folosi un senzor electrochimic, conectat la un microcontroller. Propunerea mea de microcontroller este un NodeMCU V3 Lolin, în primul rând datorită conectivității WiFi. Pe acel senzor, se va aplica o probă, de urină sau sânge. Probele și datele vor fi prelevate/obținute de la Universitatea de Medicină și Farmacie din Cluj-Napoca. Datele generate pe baza citirii de la senzor, vor fi salvate într-o bază de date non-relațională (JSON) în Firebase. Astfel, este necesară o conectivitate NodeMCU(WiFi)-Firebase(Internet). Datele din Firebase vor fi introduse ca și input într-o rețea neuronală artificială, antrenată în prealabil, folosind date de la UMF. Output-ul acelei rețele neuronale va fi una din urmatoarele bacterii: *P. aeruginosa*, *S. aureus*, *E. coli*, and *K. pneumoniae*. Folosind input-ul primit de la baza de date non-relațională (Firebase), și rețeaua neuronală artificială, se va face un forwarding, rezultând una din cele 4 bacterii. Rezultatul ar urma să fie salvat înapoi în Firebase. De asemenea, va fi creată o aplicație de mobil (framework-ul Flutter), care va citi informațiile din Firebase (conectivitate Flutter-Firebase) și va afișa pe telefonul mobil bacteria corespunzătoare probei aplicate.

Obiectivele intermediare (granularizate), reiterate pentru a putea fi mapate cu obiectivele de la punctul 7.5 "Obiectivele și activitățile de cercetare din cadrul proiectului"

1. Obținerea datelor (și a probelor) de la UMF, care stau la baza antrenării rețelei neuronale artificiale.
2. Crearea și antrenarea rețelei neuronale artificiale, pe baza datelor obținute de la UMF.
3. Crearea unei aplicații C/C++ pentru a citi datele de la senzorii electrochimici.
4. Crearea unei baze de date non-relaționale în Firebase.
5. Conectarea aplicației care citește datele de la senzor, folosind un microcontroller cu capabilități WiFi, la baza de date Firebase.
6. Crearea unei aplicații Python care citește informațiile din Firebase, le prelucrează, și le introduce în rețeaua neuronală artificială deja antrenată.
7. Returnarea rezultatului dat de rețeaua neuronală artificială, în Firebase.
8. Crearea unei aplicații de mobil, și conectarea ei la baza de date Firebase.
9. Citirea rezultatului din Firebase (bacteria corespunzătoare probei aplicate) în aplicația de mobil și afișarea acesteia.
10. Crearea unui GUI (Graphic User Interface) user-friendly, pentru a îmbunătăți experiența utilizatorului aplicației: folosind tool-uri precum Adobe XD sau Figma, GIMP.
11. Înglobarea senzorului electrochimic și a microcontrollerului într-o carcăsă.

7.3. DESCRIEREA PROIECTULUI

În cele ce urmează, voi încerca crearea unei paralele între informațiile obținute, deduse din research-ul pe care l-am efectuat în scopul completării secțiunii 7.1, pe de o parte, și a obiectivelor pe care le-am propus spre realizare, pe de altă parte. Referințele [1] și [2] au stat la baza primelor două obiective propuse: „Obținerea datelor de la UMF care stau la baza antrenării rețelei neuronale artificiale” și „Crearea și antrenarea rețelei neuronale artificiale”. Proiectul meu va folosi învățarea supervizată, propunere regăsită și în [1]. Spre deosebire de [2], acolo unde se propune folosirea unui algoritm de machine learning, abordarea mea se va îndrepta spre folosirea unei rețele neuronale artificiale. Referința [3] a stat la baza următoarelor 6 obiective propuse. Folosirea senzorilor electrochimici în detecția bacteriilor este o idee regăsită și în implementarea propusă de mine. În acest scop, voi crea o aplicație C/C++ care va citi datele de la senzorii electrochimici. Senzorul va fi conectat la un microcontroller, iar datele vor fi salvate într-o bază de date Firebase. Această bază de date va comunica, pe de o parte, cu un script în Python, care va introduce datele, după data cleaning, ca și input în rețea neuronală artificială. Rezultatul clasificării va fi returnat în baza de date Firebase. Această bază de date va comunica și cu aplicația de mobil pe care o voi implementa, idee regăsită în [3]. Rezultatul, sub forma bacteriei detectate, va fi afișat în aplicația de mobil. De asemenea, voi încerca crearea unui sistem încorporat, însemnând că senzorul electrochimic și microcontrollerul vor fi înglobate într-o carcăsă.

7.4. REFERINȚE BIBLIOGRAFICE

- [1] Omar Sayyouh, Machine Learning Application to Combat Superbugs in Hospitals: A Primer to Infection Prevention Practitioners, *Biomedical Journal of Scientific & Technical Research*, Volume 45-Issue 1, 2022.
- [2] Markus Nasman, Detecting Hospital Acquired Infections using Machine Learning, Degree project in Computer Science, KTH Royal Institute of Technology, Stockholm, 2013
- [3] Andy Extance, Electrochemical DNA sensor on a chip detects what caused UTIs faster, *Chemistry World*, 2021

7.5. OBIECTIVELE ȘI ACTIVITĂȚILE DE CERCETARE DIN CADRUL PROIECTULUI

An	Obiective științifice (Denumirea obiectivului)	Activități asociate
An1	1. Crearea și antrenarea rețelei neuronale artificiale care ar urma să clasifice tipul de bacterie.	1. Obținerea datelor de la UMF, care stau la baza antrenării rețelei neuronale 2. Crearea rețelei neuronale (a modelului). 3. Antrenarea rețelei neuronale, folosind datele de la UMF.
	2. Crearea unei aplicații C/C++ pentru a citi datele de la senzorii electrochimici și a-i salva în baza de date non-relatională.	1. Crearea unei aplicații C/C++ pentru a citi datele de la senzori. 2. Crearea unei baze de date non-relatională (JSON), în Firebase 3. Salvarea datelor citite de la senzorii electrochimici, în baza de date Firebase.
An2	1. Crearea unei aplicații Python care citește informațiile din Firebase și le introduce ca și input în rețeaua neuronală artificială.	1. Citirea informațiilor din Firebase. 2. Prelucrarea informațiilor (data cleaning), pentru a putea fi introduse ca și input al rețelei neuronale artificiale. 3. Returnarea rezultatului dat de forwarding (bacterie corespunzătoare) în baza de date non-relațională (Firebase).
	2. Crearea unei aplicații de mobil pentru a afișa rezultatul detecției de bacterii.	1. Crearea aplicației de mobil în Flutter. 2. Conectarea aplicației de mobil la Firebase, pentru a putea prelua rezultatul detecției. 3. Crearea unui GUI (Graphic User Interface) user-friendly, pentru a îmbunătăți experiența utilizatorului aplicatiei.
	3. Înglobarea microcontrollerului și a senzorului electrochimic într-o carcăsă.	1. Proiectarea carcasei. 2. Crearea/Construirea carcasei. 3. Adăugarea (înglobarea) componentelor în carcăsă.
	4. Publicarea unui articol științific la o conferință indexată IEEE/Scopus	

7.6. CONSULTANȚI

Prof. Dr. Ing. Adrian Groza, Universitatea Tehnică Cluj-Napoca
Dr. Cristea Cecilia Victoria, Universitatea Medicină și Farmacie "Iuliu Hațieganu" Cluj-Napoca

9. Activitatea științifică a candidatului:**9.1. PREMII OBȚINUTE LA MANIFESTĂRI ȘTIINȚIFICE.****9.2. PARTICIPAREA CU LUCRĂRI LA SESIUNI DE COMUNICĂRI ȘTIINȚIFICE.**

„Educație, știință și societate în secolul XXI oportunități și provocări”, 24 noiembrie 2020

9.3. PUBLICAȚII.

Valentin-Florin Ciotea, Margareta Ciotea, Marius Stoica, Alexandra Ciotea, Aplicație informatică pentru managementul documentelor dintr-o organizație, *Educație, știință și societate în secolul XXI oportunități și provocări*, Editura RISOPRINT, Cluj-Napoca, 24 noiembrie 2020

Valentin-Florin Ciotea, Margareta Ciotea, Marius Stoica, Alexandra Ciotea, Aplicație informatică pentru managementul proceselor organizaționale, *Educație, știință și societate în secolul XXI oportunități și provocări*, Editura RISOPRINT, Cluj-Napoca, 24 noiembrie 2020

Copia după cele două articole va fi adusă la interviul din 21.07.2023

9.4. PARTICIPAREA ÎN PROGRAME DE CERCETARE-DEZVOLTARE NAȚIONALE ȘI INTERNAȚIONALE

(nume proiect/director proiect/cadru didactic care a supervizat cercetarea – pentru proiecte din UTCN)
(nume proiect/director proiect/instituția în care s-a derulat cercetarea – pentru proiecte din afara UTCN)

9.5. BURSE OBȚINUTE

- FINANȚATORUL;
- PERIOADA ȘI LOCUL;
- PRINCIPALELE REZULTATE ȘI VALORIZAREA LOR.