



UNIUNEA EUROPEANĂ



Instrumente Structurale
2014-2020

Achiziții și prelucrări de date IoT în CloudUT



UNIVERSITATEA TEHNICĂ
DIN CLUJ-NAPOCA



Conf. dr. ing. Anca Hangan
Departament Calculatoare
Universitatea Tehnica din Cluj-Napoca
Anca.Hangan@cs.utcluj.ro

Proiect CLOUDUT



- Titlu: Cloud Cercetare UTCN – CLOUDUT
(<http://cloudut.utcluj.ro>)
- MySMIS ID: 124493
- Contract no.: 235/ 21.04.2020
- Tip Proiect: Program Operational Competitivitate 2014-2020 (POC)
- Axa prioritara 1: Cercetare, dezvoltare tehnologică și inovare (CDI) în sprijinul competitivității economice și dezvoltării afacerilor
- Acțiunea 1.1.2: Dezvoltarea unor rețele de centre CD, coordonate la nivel național și racordate la rețele europene și internaționale de profil și asigurarea accesului cercetătorilor la publicații științifice și baze de date europene și internaționale
- Finantare: Fonduri Europene pentru Dezvoltare Regională, Valoarea totală: 4.955.000 RON, din care 4.950.000 RON din fonduri Europene.



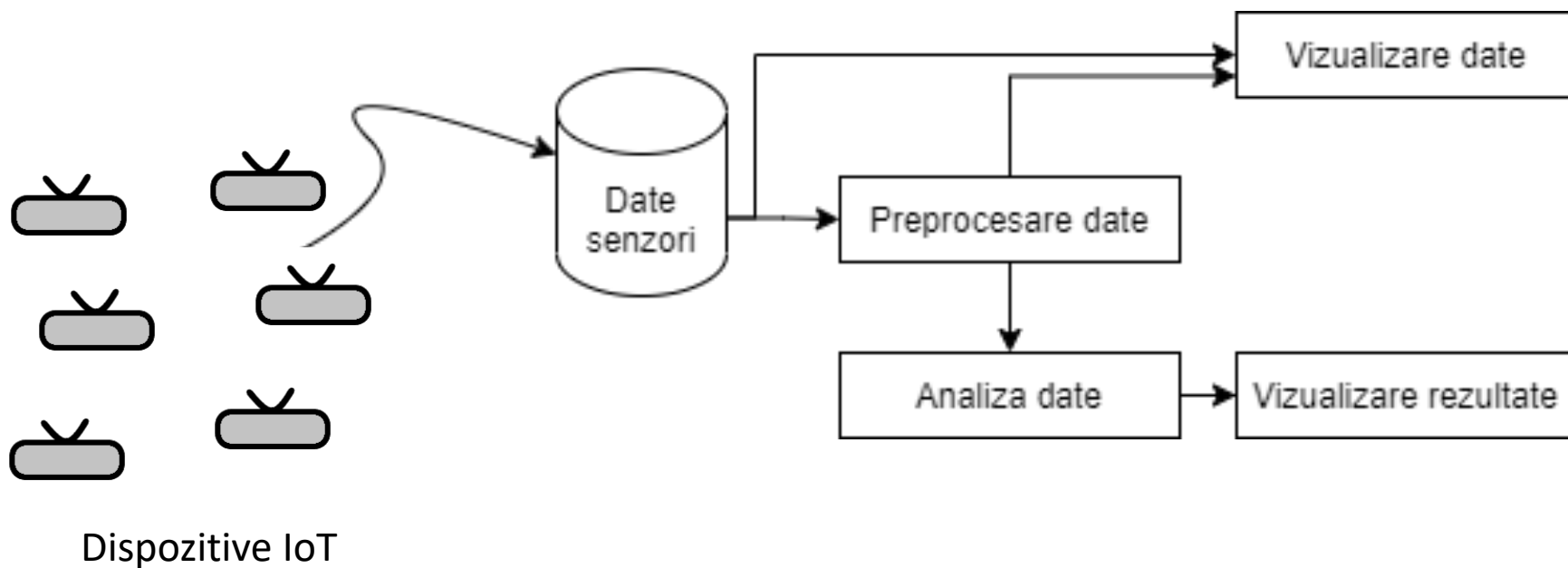
Conținut

1. Proiecte de monitorizare cu senzori: colectarea, vizualizarea, procesarea și analiza datelor.
2. Motive pentru mutarea proiectului de monitorizare în CloudUT.
3. Solicitarea și modul de utilizare al resurselor în CloudUT.
4. Servicii IoT în CloudUT: prezentarea specificațiilor setului de servicii.
5. Experiment de utilizare a serviciilor IoT:
 - a. Colectarea datelor de la un dispozitiv IoT.
 - b. Detecția anomaliilor într-un set de date de calitate a apei din rețeaua de distribuție.
6. Sesiune de întrebări și discuții.

Context

- Monitorizarea cu dispozitive IoT - aplicații în direcții de cercetare de interes în contextul actual din domeniul CTI și în direcții de cercetare interdisciplinare
- Direcții de cercetare care folosesc monitorizare cu dispozitive IoT:
 - Asistență pentru autonomie la viață (Ambient Assisted Living)
 - Automatizarea casei/clădirilor
 - Transport inteligent
 - Orașul inteligent
 - Distanțare socială în contextul pandemiei covid-19

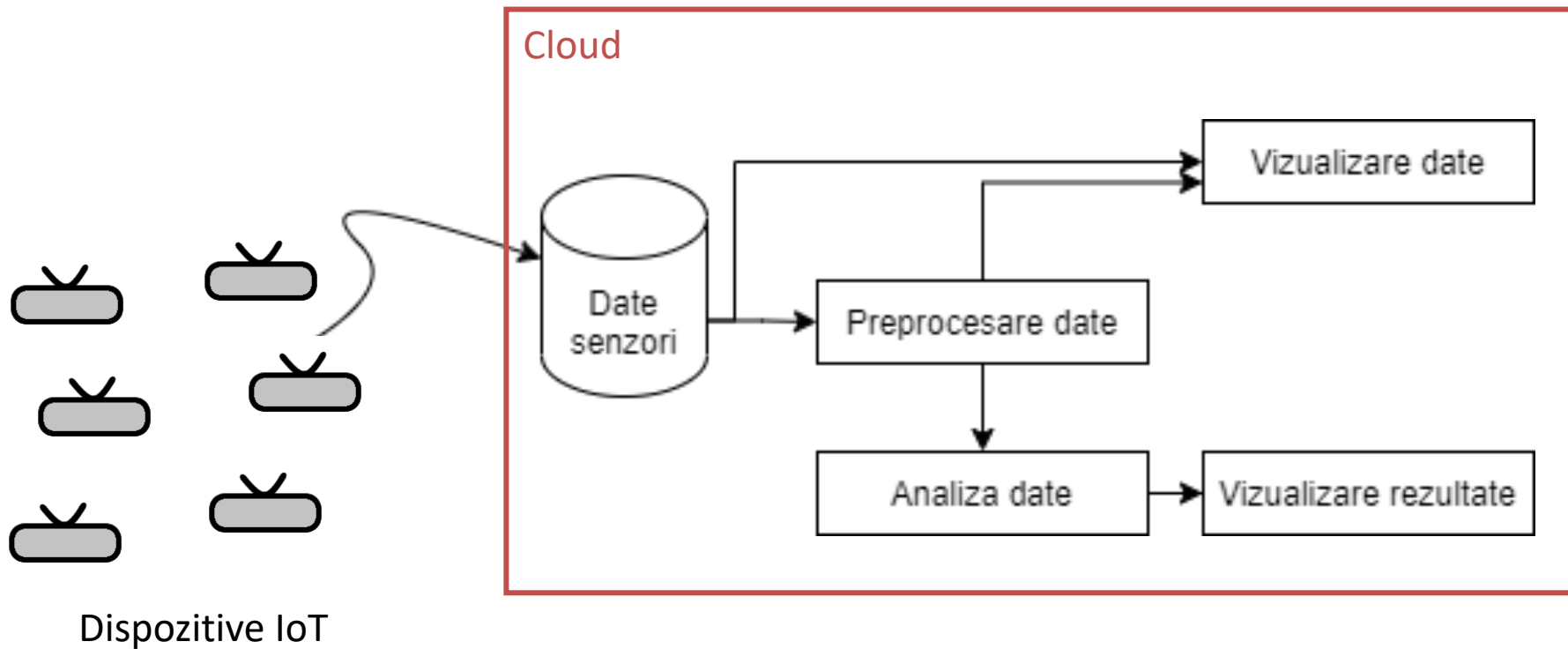
Proiecte de monitorizare cu dispozitive IoT



Componente importante

- *Colectarea datelor de la dispozitive aflate la distanță*
 - Dispozitive pentru achiziția datelor de la senzori
 - Colectarea și stocarea datelor
 - *Câte date* se generează/stochează?
- *Vizualizarea, procesarea și analiza seturilor de date colectate*
 - Unele *distincte* pentru vizualizarea/preprocesarea/analiza datelor
 - Preprocesarea datelor: date lipsă, normalizare, filtrare, etc.
 - *Analiza datelor*: detecția de modele/șabloane, detecția de anomalii, etc.
 - Metode de inteligență artificială, învățare automată
 - Alți algoritmi

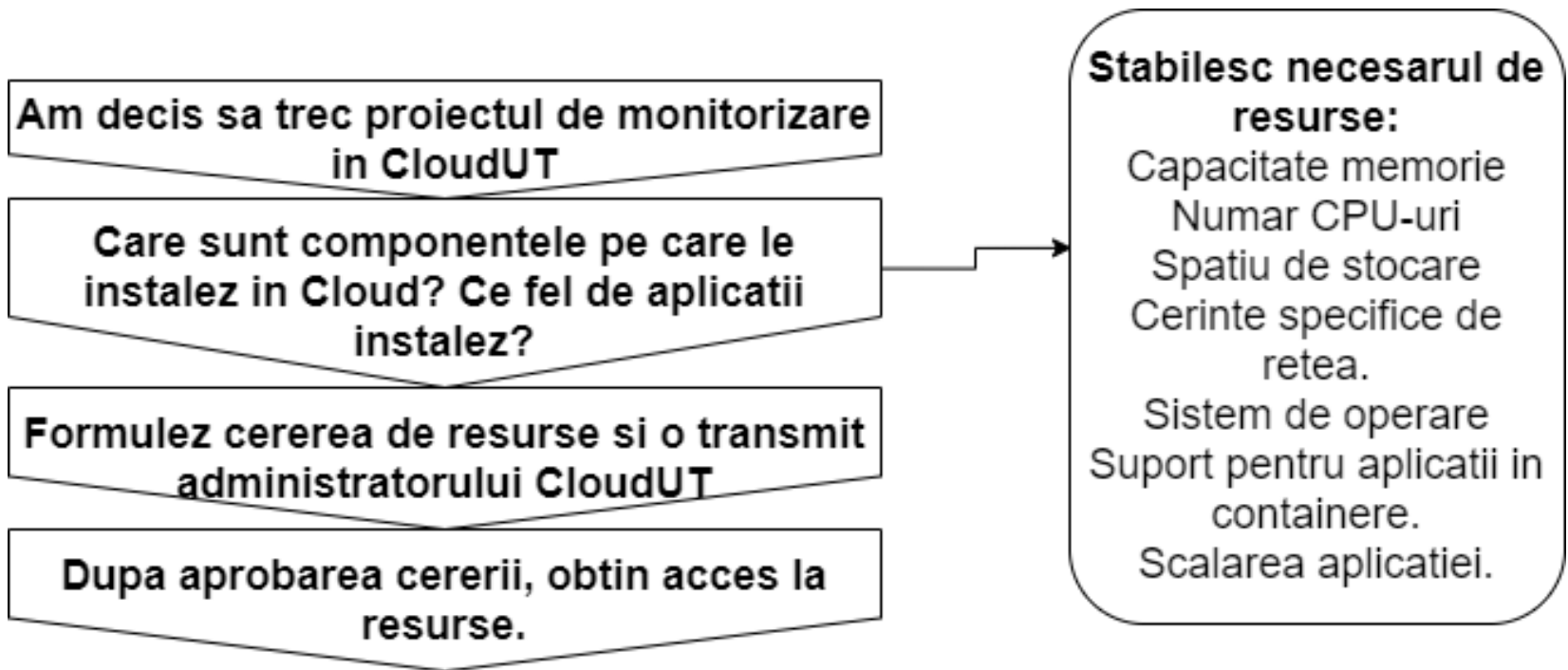
Proiect de monitorizare in Cloud



Utilizarea infrastructurii CloudUT

- Este recomandat să folosesc CloudUT dacă:
 - Proiectul de monitorizare este de **dimensiuni mari**
 - Am multe dispozitive de monitorizare (=> trafic mare de date)
 - Se generează seturi de date de dimensiuni foarte mari (=> necesită spațiu de stocare de capacitate mare)
 - Procesarea și analiza datelor necesită **multe resurse de calcul** (ex: trebuie să antrenez un model de învățare automată)
 - Cerințe de **scalabilitate** și **fiabilitate**
- CloudUT - suport pentru:
 - Realizarea proiectelor de cercetare în care se desfășoară activități de monitorizare
 - Realizarea de cercetări experimentale care pot fi folosite ulterior ca bază pentru propuneri de proiecte

Solicitarea resurselor



Modul de utilizare a resurselor

Dezvolt si testez aplicatia pe resursele locale

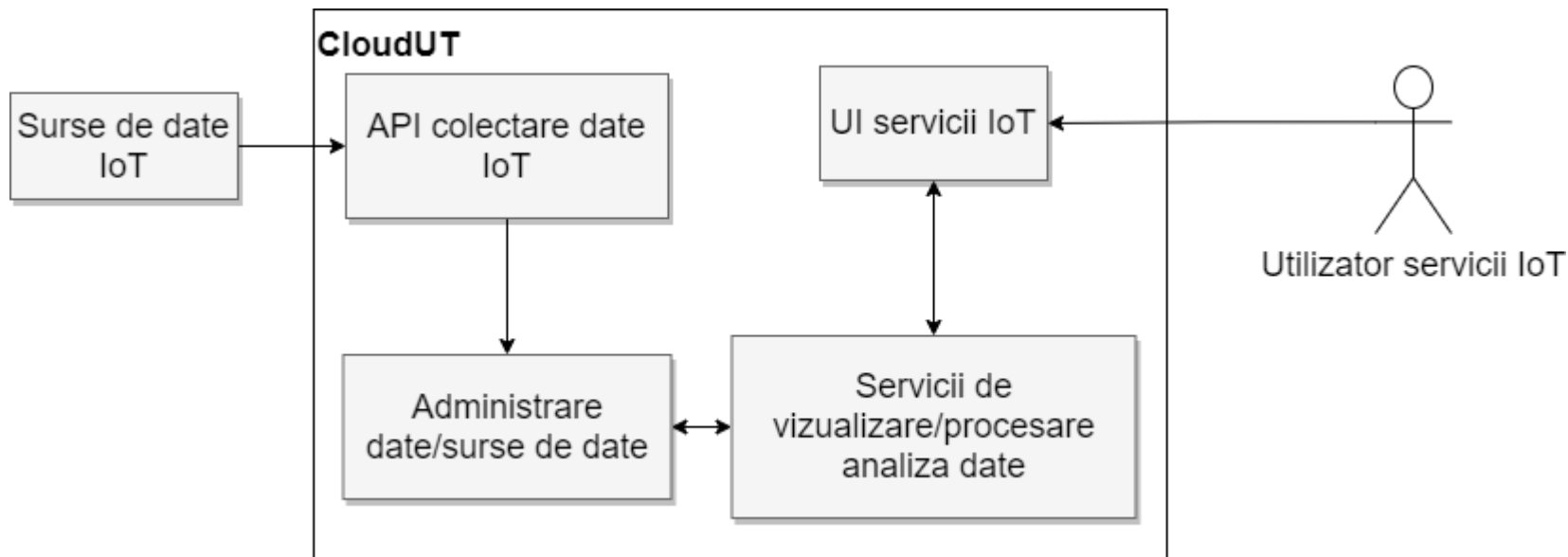
Instalez aplicatia in CloudUT

Utilizez resursele pe perioada desfasurarii experimentelor

La finalizarea experimentelor, salvez datele local si eliberez resursele din CloudUT

- *Obiectiv: set integrat de servicii pentru colectarea, stocarea si analiza datelor provenite de la dispozitive IoT, publicabil prin CloudUT*
- *Utilizatori:* colective de cercetare care desfășoară activități de monitorizare cu rețele de senzori/dispozitive IoT, în cadrul proiectelor de cercetare sau în studiile premergătoare unor propuneri de proiecte.
- *Scopul final:* infrastructura suport pentru realizarea proiectelor de cercetare care implică task-uri de monitorizare

Servicii IoT in CloudUT



Arhitectură bazată pe servicii instalate în containere

Surse de date

- Surse eterogene:
 - dispozitive individuale care realizează măsurători (senzori) și care se pot conecta în mod direct la Internet
 - rețele de senzori care folosesc diverse tehnologii pentru transmitia de date, conectate la Internet prin dispozitive de tip gateway (multi-protocol)
 - surse externe de date generate de alte sisteme
- Scenarii de monitorizare:
 - Dispozitive IoT fixe sau mobile, conectate într-o rețea locală, care au acces la Internet printr-un dispozitiv gateway.
 - Dispozitive IoT fixe sau mobile, conectate individual la Internet.
 - Dispozitive IoT mobile, care se conectează la Internet numai în momentul în care există o conexiune disponibilă.

API de colectare a datelor

- Interfața HTTP pentru preluarea datelor
- Dispozitive înregistrate transmit date prin cereri HTTP
- Probleme de luat în considerare:
 - Interfață pentru preluarea datelor trebuie să permită conectarea unei game largi de dispozitive
 - *API Standard* pentru preluarea datelor IoT (OGC SensorThings si implementările sale open source)

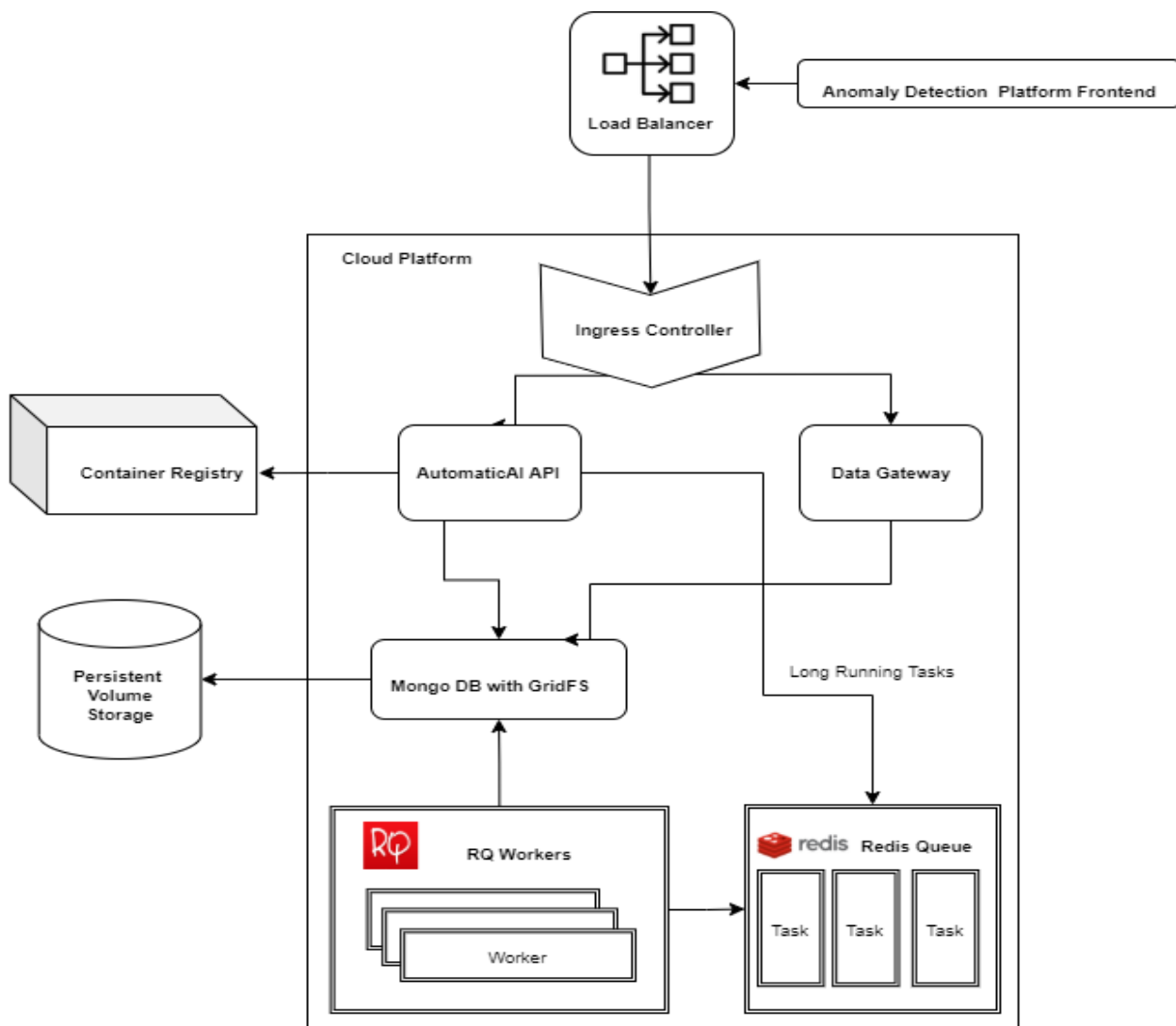
Serviciu administrare date/surse de date

- *Stocarea datelor* colectate de la dispozitive IoT
- *Administrarea surselor de date* (dispozitivelor): păstrarea de informații despre starea surselor de date, și eventual, de control acestora de la distanță
- *Administrarea datelor*: primirea, decriptarea, decodificarea și stocarea datelor trimise de sursele de date
- Probleme de analizat:
 - Formatul datelor primit/acceptat de dispozitive
 - Modelul folosit pentru stocarea datelor (modele standard)

Serviciul de vizualizare, procesare și analiza a datelor

- *Reproiectarea platformei pentru detecția anomaliilor* dezvoltată în proiectul intern UTCN: „Detecția anomaliilor în rețele de senzori de dimensiuni mari”.
- Include:
 - Algoritmi de preprocesare a datelor
 - Algoritmi de clasificare
 - Selecție automată a algoritmului de clasificare, reglare automată a parametrilor
 - Execuție paralela a task-urilor de procesare
 - Vizualizarea datelor și a rezultatelor

Arhitectura platformei de vizualizare, procesare și analiză a datelor



Interfața utilizator a serviciului

- Configurare
- Managementul datelor (download, upload date)
- Vizualizarea datelor sub diferite forme (ex: serii de timp, histograme, etc.)
- Definirea secvenței de pași pentru procedurile de procesare și analiză a datelor
- Vizualizarea rezultatelor procesărilor

Experimentul de azi

- Prototip funcțional al serviciilor IoT CloudUT
- Colectarea datelor de la un dispozitiv de tip NodeMCU (conectat la Internet)
- Detectia anomaliilor in date de calitate a apei din rețeaua de distribuție
 - Setul de date utilizat -> Competiția GECCO 2017*
 - Pipeline preprocesare + antrenare model cu PSO-SA pentru detectia anomaliilor
 - Evaluarea rezultatelor: aplicarea modelului antrenat pe datele de validare - > vizualizarea rezultatelor

Setul de date

- Coloane:
 - Tp - temperatura apei in grade Celsius
 - pH - valoare pH
 - Leit - conductivitate
 - Trueb - turbiditate
 - Redox - potențial de oxidare
 - Cl, Cl2 - concentrația de clor
 - Fm, Fm2 - debit
 - Event: true (anomalie)/false (valoare normală)
- Setul de date împărțit in:
 - Set de antrenare
 - Set de validare

Pipeline de procesare a datelor

- Obiectiv: etichetare automată - true (anomalie) / false (valoare normală)
- *3 pași de preprocesare:*
 - Completare date lipsa
 - Normalizare
 - Over/Under sampling (număr mic de anomalii in setul de antrenare)
- ***Antrenare modele*** de învățare automată cu algoritmul PSO-SA
 - Alege modelul cu rezultatele cele mai bune pentru datele de antrenare furnizate
 - Modelul este salvat, apoi aplicat pe datele de validare
- ***Vizualizarea rezultatelor***
 - Matrice (valori reale/prezise)
 - Tabel

Întrebări?

Mulțumesc pentru atenție!