**Capitolul 1: Introducere**

În general, un proces generează un rezultat, majoritatea rezultatelor formând un comportament normal al acelui proces și ajutând la stabilirea unui model care definește standardul, adică norma. O **anomalie** este o observație din rezultatele procesului care se diferențiază atât de mult de celelalte rezultate, trezind suspiciunii că acea observație a fost generată de un alt mecanism. Termenul **outlier** este de asemenea utilizat în literatura de specialitate pentru descrierea unei anormalitați, unii cercetători utilizând termenii de **weak outlier** și **strong outlier**(Figura 1.1) pentru a distinge observațiile care sunt cu adevărat anormale de cele care doar par a fi anormale, această distincție realizându-se prin analiza nivelului de deviere a observației față de normă. În acest sens, termenul **noise** este folosit pentru a descrie observațiile care se abat puțin de media grupului, iar termenul **anomalie** este utilizat pentru a descrie observațiile care se abat semnificativ de la media grupului și reprezintă cu adevărat o anormaliate.[1]

Graphical user interface, application

Description automatically generated

Figura 1.1: Spectrul de la date normale la anomalii (Imagine preluată din [1])

Recunoașterea comportamentelor neobișnuite sau anormale este utilizată într-o varietate de domenii(Figura 1.2), câteva domenii de aplicabilitate sunt :

**Securitatea cibernetică** reprezintă o preocupare majoră în era digitală în care trăim, amenințarile cibernetice devenind din ce în ce mai sofisticate si tot mai frecvente, astfel de atacuri putând genera consecințe grave atât pentru corporații cât și pentru viața personală a oamenilor. În această sferă, detecția anomaliilor este folosită în:

Detectarea atacurilor de tip ***Denial of Service***, care sunt o metodă prin care se incearcă supraîncarcarea unui server prin trimiterea unui număr mare de pachete, acesta neputând procesa această cantitate de trafic și ducând la afectarea disponibilitații și a performanței serviciului, putând să aducă pierderi financiare pentru entitatea asupra căreia a avut loc atacul.

Monitorizarea ***activităților de autentificare******neobișnuite***, cum ar fi încercările de acces cu parole greșite sau încercarile de acces de la locații geografice neobișnuite, pentru a identifica probleme de securitate și de a bloca accesul neautorizat.

Identificarea activităților neobișnuite ale unui ***program malware*** care încearcă sa modifice fișiere critice de sistem.

**Finanțele** reprezintă un alt domeniu în care detectarea anomaliilor a căpătat un rol extrem de important, multe atacuri cibernetice focusându-se pe obținerea și abuzarea de date și resurse financiare ale unor persoane naive. Unul dintre aceste abuzuri este ***frauda cu cardul de credit***, aceasta constând in utilizarea neautorizată a cardului de credit sau debit în urma pierderii, furtului acestuia sau prin obținerea de informații de pe card, cum ar fi numărul cardului, data expirării și codul CVV, aceastea putând fii obținute prin intermediului **phishing**-ului sau al altor metode frauduloase.

**Medicina** este un domeniu care beneficiază de evoluția continuă a tehnologiei, evoluție care a adus la schimbări notabile in ceea ce privește modalitățile de diagnosticare, tratament și gestiune a bolilor. Printre aceste evoluții apar și tehnici de detecatre a anomaliilor în diferite arii din acest domeniu cum ar fi:

**Diagonsticarea** pacienților, care se bazează pe analiza datelor, semnalând comportamente anormale în semnele vitale. *Contracția ventriculară prematură* este o *aritmie* cardiacă, adică o neregularitate in datele electrogramei ***EKG***-ului, în acest caz aceasta constând în apariția unor contracții suplimentare ale ventriculelor inimii, care se produc prematur. Detectarea acesti anormalitați se realizează bazându-se pe serii de timp și prin compararea cu datele unui EKG normal, astefel reducându-se semnificativ volumul de muncă al cardiologilor și crescând rata de detectare a contracției ventriculare premature [1]. **Diagnosticarea cancerului**. Probelma constă în difernțierea tumorilor *benigne* de cele *maligne* prin date de la imagini radiografice , ținându-se cont de numărul cazurilor reltativ mic de tumori maligne descoperite. Asupra acestei probleme a fost aplicat **machine learning** și **algoritmi de clasificare**, dar problema poate fi rezolvată și prin aplicarea tehnicilor de detecție a anomaliilor, presupunând că benignele reprezintă ceva “normal” se poate construii un model, astfel malignele putând fi diferențiate [1].



Figura 1.2: Domenii aplicabilitate(Imagine preluata din [6])

Luând in considerare aceste domenii enumerate mai sus și multe altele, consider că studiiul tehnicilor pentru detectia anomaliilor reprezinta un taină importanta. Astfel aceasta lucrare se va ocupa de aceasta problema, dupa cum urmeaza: In capitolul 1…..capitolul 2….capitolul2

Prin această lucrare, voi efectua un studiu comparativ a diferitelor tehnici de detecție a anomaliilor...........................................................................................................................

**Capitolul 2:**

**Bibliografie**

[1] Charu C. Aggarwal - Outlier Analysis-Springer (2013)

[2] Anomaly Detection Principles and Algorithms (Kishan G. Mehrotra,Chilukuri K. Mohan etc.)

[3] LOCI: Fast Outlier Detection Using the Local Correlation Integral ( Spiros Papadimitriou Hiroyuki Kitagawa Phillip B. Gibbons Christos Faloutsos, November 2002)

[4] Side channel monitoring: Attack detection in industrial systems( Bolboacă Roland)

[5] Master (Roland Bolboacă)

[6] Machine Learning for Time-Series with Python: Forecast, Predict, and Detect Anomalies with State-of-the-art Machine Learning Methods(Ben Auffarth, 2021)

[7] Programming with MATLAB for Scientists A Beginner’s Introduction(Eugeniy E. Mikhailov)

[8] <https://scikit-learn.org/stable/auto_examples/neighbors/plot_lof_outlier_detection.html>

[9] https://scikit-learn.org/stable/modules/outlier\_detection.html#outlier-detection