

Universitatea POLITEHNICA din București

Facultatea de Automatică și Calculatoare,
Departamentul de Calculatoare



LUCRARE DE DIPLOMĂ

Instrument informatic pentru procesarea
și vizualizarea electroencefalogramelor

Autor:

Dochian Alexandru Adrian

Conducători Științifici:

sl. dr. ing. Alexandru Sorici
sl. dr. ing. Mihai Trăscău
sl. dr. ing. Irina Andra Tache

București, Iulie 2022

Această lucrare nu ar fi fost posibilă fără ajutorul catedrei de fiziologie din cadrul Facultății de Medicină, UMF "Carol Davila", București care a achiziționat electroencefalograme intracraniene de calitate superioară pe regretați soareci și le-a împărtășit cu mine și coordonatorii mei.

De asemenea adresez mulțumiri coordonatorilor mei, Domnului sl. dr. ing. Mihai Trăscău, Domnului sl. dr. ing. Alexandru Șorici și Doamnei sl. dr. ing. Irina Andra Tache pentru timpul, energia și cunoștințele dedicate acestei lucrări.

Compendiu

Electroencefalogramele sunt semnale care poartă informații despre evoluția potențialului electric într-un punct ales din cortexul subiectului și sunt reprezentate ca simple serii de timp. Ele sunt utilizate în domeniul medical pentru diagnosticarea sau predicția de patologii. În lucrare de față este prezentat un sistem informatic de analiză a electroencefalogramelor. Analiza presupune procesarea informațiilor utilizând mai multe fluxuri configurabile formate la rândul lor dintr-un număr variabil de stagii de procesare. În vederea analizei, informațiile inițiale trec prin fluxuri configurabile de preprocesare și procesare. Fiecare flux este la rândul său format dintr-un număr variabil de stagii procesatoare. Setul de date constă în electrocorticograme din lobii frontali și parietali și electrocardiograme achiziționate pe șoareci. Pe parcursul colectării datelor, subiecții au fost stimulați cu impulsuri electrice, de asemenea înregistrate. Fluxul de preprocesare modifică semnalele în domeniul timp, iar fluxul de procesare extrag caracteristici din domeniile timp și frecvență. Caracteristicile extrase sunt analizate și vizualizate cu ajutorul algoritmului t-SNE(t-Distributed Stochastic Neighbor Embedding). Sistemul prezentat este un instrument informatic ce poate fi folosit pentru a analiza astfel de experimente. Arhitectura sistemului este flexibilă și permite extinderea facilă a fluxurilor existente.

Cuprins

Mulțumiri	i
Compendiu	ii
1 Introducere	1
1.1 Electroencefalografia	1
1.1.1 Fundamente fizice	1
1.1.2 Istorie	1
1.1.3 Aplicații	2
1.2 Prezentarea sistemului	2
1.2.1 Datele folosite	2
1.2.2 Analiza clasică a electroencefalogramelor	3
1.2.3 Analiza electroencefalogramelor cu ajutorul sistemului	3
2 Arhitectură	4
2.1 Arhitectura codului	4
2.1.1 Pachetul engine	4
2.1.2 Pachetul pipelines	5
2.1.3 Pachetul exceptions	7
2.1.4 Pachetul configuration	7
3 Execuția sistemului	8
3.1 Inițializarea sistemului	8
3.2 Fluxul de preprocesare	9
3.2.1 SplitInWindows	10
3.3 Fluxul de procesare	10
3.3.1 FeatureExtractor	10
3.4 Execuția algoritmului T-SNE	10
3.4.1 Despre T-SNE	10
3.4.2 T-SNE pe un alt set de date	11
3.4.3 T-SNE în cadrul sistemului curent	11
4 Rezultate și Concluzii	12
4.1 Lobii frontali și parietali ai primului șoarece(mouse1.acq)	12
4.2 Lobii frontali și parietali ai celui de-al doilea șoarece(mouse2.acq)	12
4.3 Lobii frontali și parietali ai celui de-al treilea șoarece(mouse3.acq)	13
4.4 Lobii frontali și parietali ai celui de-al patrulea șoarece(mouse4.acq)	13
4.5 Lobul frontal pentru toți șoarecii	14
4.6 Lobul parietal pentru toți șoarecii	14
4.7 Lobii frontali și parietali pentru toți șoarecii	15
5 Viitor	16

5.1	Comunicare peste HTTP	16
5.2	Interfață grafică	16
A	Configurații de execuție	17
A.1	Configurația standard de execuție	17

Capitolul 1

Introducere

1.1 Electroencefalografia

Electroencefalografia este o modalitate prin care se relevă informații despre activitatea electrică a stratului de la suprafața creierului. Informația rezultată în urma procesului de achiziționare este evoluția în timp a potențialului electric în punctul de interes.

Electroencefalografia înregistrează date de pe scalpul subiecților și de aceea este predispusă la zgomot. Electroencefalografia sau electroencefalografia invazivă a fost inventată și rafinată. Ea presupune achiziționarea datelor direct de pe cortexul subiectului.

1.1.1 Fundamente fizice

Electroencefalogramul este un dispozitiv care folosește electrozi plasați pe scalpul subiectului pentru a înregistra fluctuațiile de voltaj din punctul de interes. Existența acestui voltaj ce poate fi măsurat se datorează existenței unui curent ionic între neuroni.

Electrozii sunt conductori electrici folosiți pentru a face contact cu părți nemetalice ale unui circuit electric. Pot fi construiți din mai multe materiale dintre care menționăm: Litiu (Li), Mangan (Mn), Cupru (Cu) și Zinc (Zn)

Un ion este un atom care prezintă un dezechilibru de sarcină electrică. Anionul este un ion format din mai mulți electroni decât protoni, astfel având o sarcină electrică negativă. Cationul este opusul anionului, fiind format din mai mulți protoni decât electroni și implicit are o sarcină electrică pozitivă. Încheind această prezentare a fundamentelor fizice, curentul ionic este un flux de de electroni produs de grupări atomice de ioni și întâlnit la diverse nivele ale materiei.

1.1.2 Istorie

În anul 1875 fizicianul englez Richard Canton a prezentat în jurnalul British Medical Journal descoperile sale despre fenomenele electrice relevate pe cortexuri de iepuri și maimuțe.

De-a lungul vremii oameni de știință ca fiziologul polonez Adolf Beck sau fiziologul ucrainean Vladimir Vladimirovich Pravdich-Neminsky au studiat aceste fenomene electrice ale creierului. În experimentele lor subiecții erau animale supuse la diverși stimuli exteriori.

Un eveniment de o importanță majoră în istoria electroencefalografiei este produs de fiziologul și psihiatrul german Hans Berger care reușește să achiziționeze primele date pe un subiect uman. Numește dispozitivul propus de el "Electroencefalogram" și astfel netezește și potențează această calea științifică ce și-a dovedit până în zilele noastre potențialul.

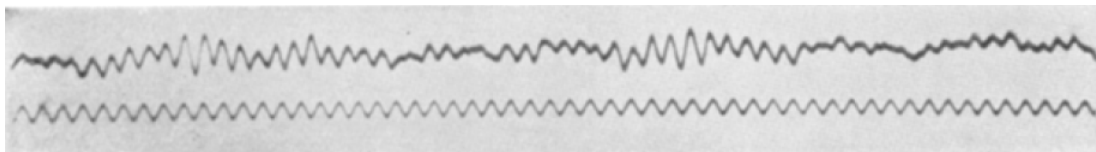


Figura 1.1: Primul EEG uman achiziționat de către Hans Berger în 1924. Semnalele din imagine sunt eșantionate la o frecvență de 10 Hz.

1.1.3 Aplicații

Activitatea creierului observată cu ajutorul acestei metode s-a dovedit prolifică în domeniul medical. Electroencefalogramele sunt utile în diagnosticarea și tratarea unor boli ca epilepsia, tumorile pe creier sau accidentele vasculare cerebrale. Unele boli ca apoplexia pot fi chiar și prezise utilizând activitatea cerebrală relevantă.

1.2 Prezentarea sistemului

1.2.1 Datele folosite

Catedra de fiziologie de la Facultatea de Medicină, UMF "Carol Davila", București ne-a onorat cu împărtășirea unor electroencefalograme invazive (electrocorticograme) obținute din experimente efectuate pe șoareci.

Pentru achiziționarea electrocorticogramelor, părțile anatomice care proieau cortexul subiecților au fost eliminate. Electrozi au fost plasați direct pe cortex în zonele corespunzătoare lobilor frontali și parietali și anestezice au fost administrate direct în creier prin injecție. Patru experimente au fost efectuate pe șobolani diferiți astfel obținându-se câte o electrocorticogramă pentru fiecare experiment.

De-a lungul acestei lucrări experimentele vor fi recunoscute prin fișierele rezultate corespunzătoare: "mouse1.acq", "mouse2.acq", "mouse3.acq" și "mouse4.acq".

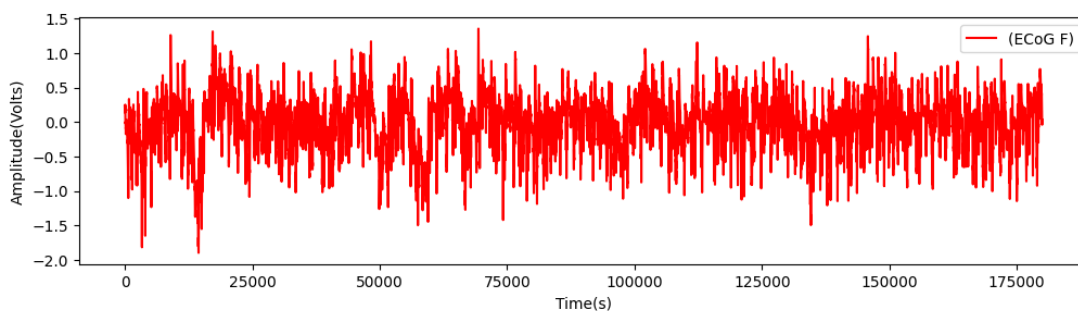


Figura 1.2: Electrocorticograma lobului frontal al șobolanului 2 de-a lungul a aproximativ 180 secunde

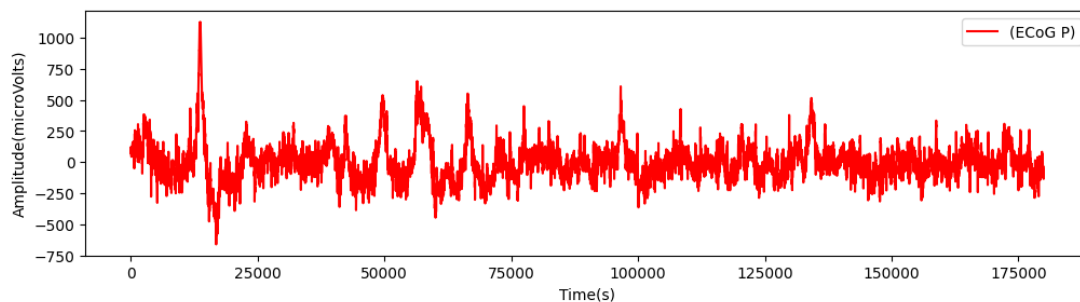


Figura 1.3: Electrocuticograma lobului parietal al șobolanului 2 de-a lungul a aproximativ 180 secunde

1.2.2 Analiza clasică a electroencefalogramelor

În mod normal astfel de informații se analizează ochiometric de către oameni specializați în acest sens. Un neurolog se antrenează timp îndelungat să dezvolte abilitățile necesare interpretării acestor semnale. Fără un sistem de calcul, analiza umană se poate face doar prin observarea unor lungi secvențe de semnale.

1.2.3 Analiza electroencefalogramelor cu ajutorul sistemului

Cu această lucrare, încercăm să oferim o nouă perspectivă pentru analiza semnalelor electroencefalografice ca Lobul Frontal sau Lobul Parietal ale șobolanului 2.

Sistemul informatic prezentat în această lucrare procesează electroencefalograme cu ajutorul învățării automate nesupervizate și le proiectează într-un spațiu bidimensional.

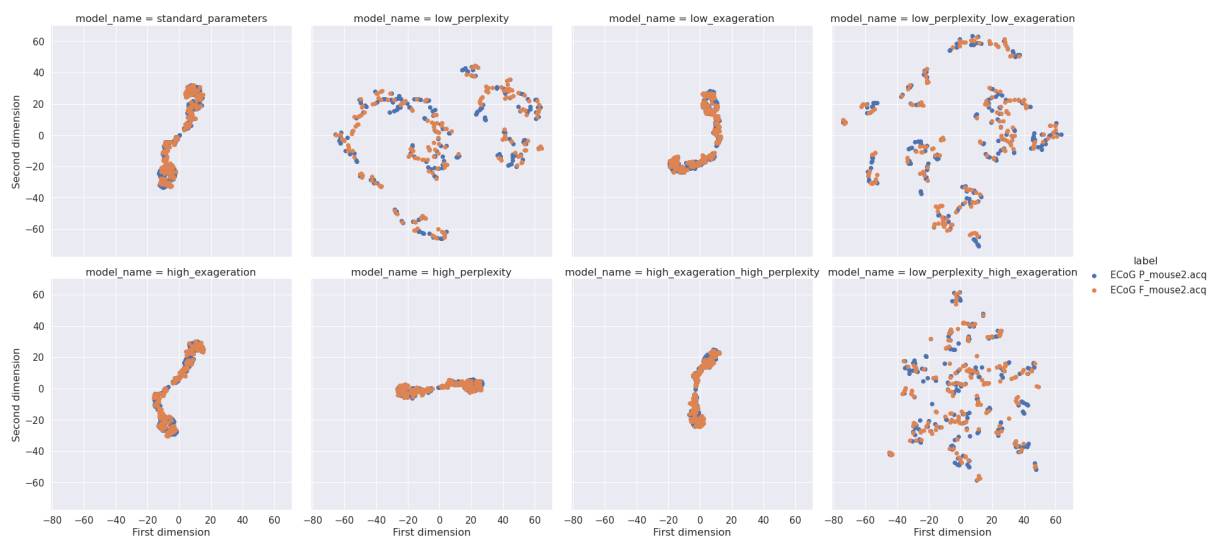


Figura 1.4: Execuția sistemului pe electroencefalogramele prezentate în Secțiunea 1.2.1

Modul în care sistemul produce aceste rezultate va fi explicat în următoarele secțiuni ale lucrării.

Capitolul 2

Arhitectură

În acest capitol se va prezenta și explicita aplicația din diverse unghiuri și nivele de adâncime.

2.1 Arhitectura codului

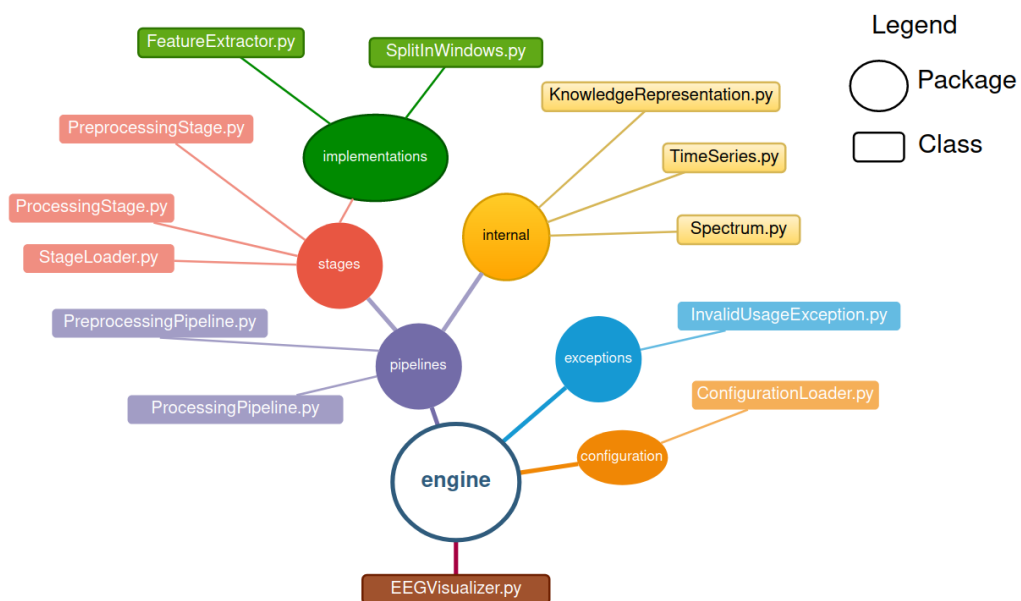


Figura 2.1: Structura codului

Sistemul este implementat folosind limbajul de programare Python.

Fiecare fișier din ierarhia prezentată în Structura codului conține doar o clasă de python cu același nume.

2.1.1 Pachetul engine

Pentru inițierea sistemului trebuie creat un obiect din clasa **EEGVisualizer**. Obiectul expune metodele necesare producerii de rezultate și creează la nevoie obiectele necesare din restul sistemului.

În mod implicit electroencefalogramele prezentate în Secțiunea 1.2.1 sunt încărcate în memorie la instanțierea clasei **EEGVisualizer**.

În continuare o să exemplific modul în care s-a folosit sistemul pentru generarea rezultatelor prezentate în Figura 1.2.3.

```

1
2 from engine.EEGVisualizer import EEGVisualizer
3
4 eeg_visualizer = EEGVisualizer()
5 eeg_visualizer.build(files_name=["mouse2.acq"], eeg_channels=["ECoG_
    F", "ECoG_P"])
6 eeg_visualizer.execute()
7 eeg_visualizer.plot()

```

Listarea 2.1: Execution

2.1.2 Pachetul pipelines

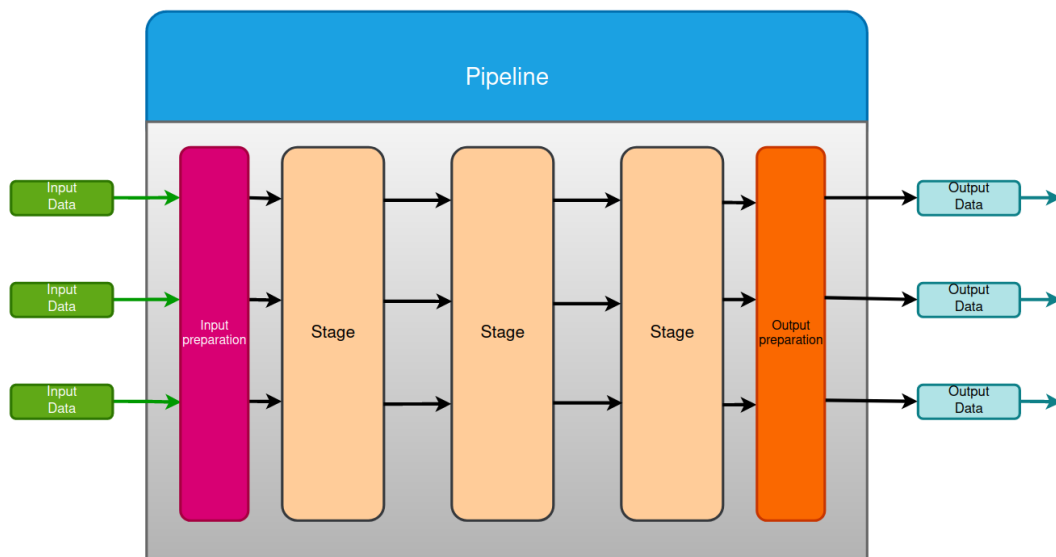


Figura 2.2: Perspectivă globală asupra funcționalității unui flux

În Figura de mai sus observăm comportarea abstractă a unui flux de procesare.

În acest sistem au fost implementate două astfel de fluxuri:

- Flux de preprocesare
- Flux de procesare

Fiecare flux își are propriul protocol de comunicare cu restul sistemului și conține un număr variabil de stagii interne de comportament corespunzător. Merită menționat că un stagiu din fluxul de preprocesare nu poate fi folosit în fluxul de procesare sau invers deoarece implementează interfețe diferite.

Pachetul stages

Aici sunt plasate clasele abstracte care definesc comportamentul ce trebuie respectat de un stagiile efective:

- Stagiu de preprocesare
- Stagiu de procesare

Stagiile instanțiabile ce intră, din perspectiva execuției, în componența fluxurilor sunt definite în pachetul implementations.

Pachetul implementations

Clasele definite în acest pachet trebuie să respecte unul dintre protocoalele impuse de cele 2 stagii abstracte ale sistemului. Sistemul a fost construit într-o manieră deschisă la extensie. În eventualitatea îmbogățirii sistemului cu noi stagii și integrarea lor în execuție ar trebui urmați următorii pași simpli:

1. Crearea clasei care implementează protocolului impus de superclasa aleasă
2. Actualizarea configurației de execuție servită de pachetul configuration

Pachetul internal

Electroencefalogramele pe care sistemul le primește din exterior sunt în formatul bibliotecii bioread. Un obiect de tipul **bioread.biopac.Datafile** conține informații achiziționate pe mai multe canale de tipul **bioread.biopac.Channel**.

Sistemul traduce aceste intrări într-o logică proprie, generică, **KnowledgeAcquisition**.

Problema procesării electroencefalogramelor se reduce în mod implicit la a procesa semnale în domeniile timp și frecvență. Considerând acest lucru am construit clasele **TimeSeries** și **Spectrum**.

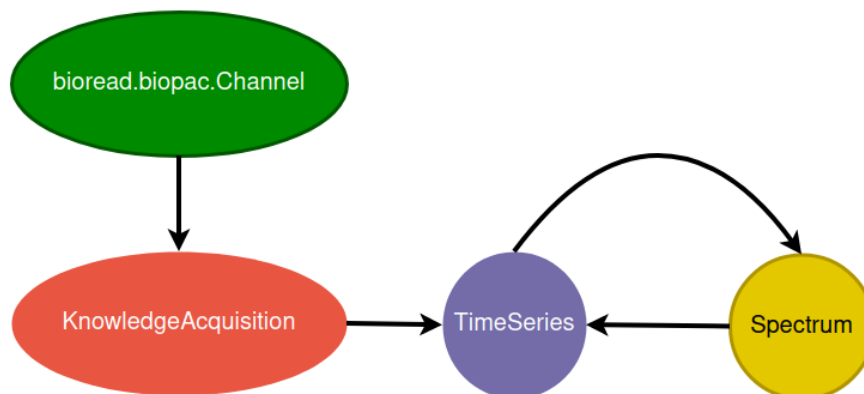


Figura 2.3: Relaționarea obiectelor din pachetul **internal**

TimeSeries și **Spectrum** conțin din punct de vedere matematic aceeași informație esențială.

Sistemul permite, cu ajutorul **transformatei Fourier** obținerea unei instanțe **Spectrum** dintr-o instanță **TimeSeries** și viceversa.

2.1.3 Pachetul exceptions

Excepțiile definite la nivel de sistem au fost plasate în acest pachet pentru modularitate și pentru înlesnirea dezvoltării următoarelor versiuni.

2.1.4 Pachetul configuration

Pachetul respectiv oferă restului aplicației clasa `ConfigurationLoader`. La nevoie, alte părți ale sistemului cer din acest pachet configurația de execuție a sistemului. Un exemplu de configurație de execuție servită de acest pachet este prezentată în [Listarea A.1](#)

Capitolul 3

Execuția sistemului

În acest capitol se va explica pas cu pas modul în care sistemul procesează electroencefalogramele.

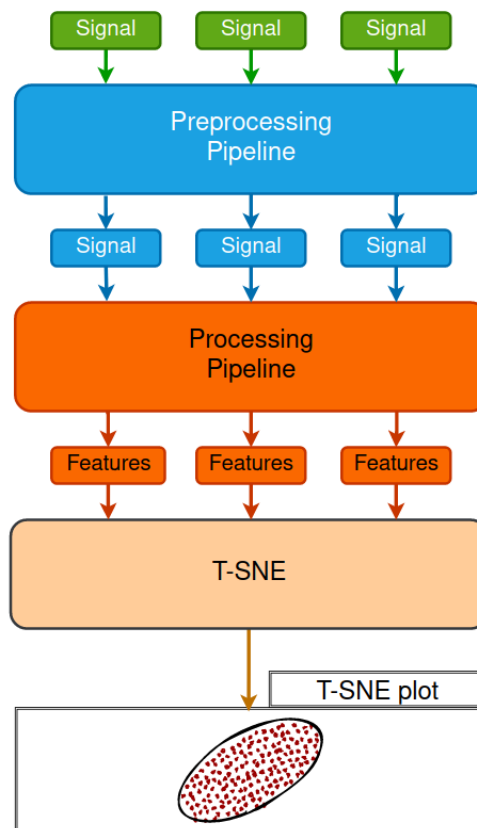


Figura 3.1: Structura execuției

3.1 Inițializarea sistemului

Pentru inițializarea sistemului este nevoie de:

- O configurație de execuție ca cea din [Listarea A.1](#)

- Fișiere cu extensia **.acq** create de programul AcqKnowledge ce vor fi încărcate în memorie
- Alegerea unor fișiere și a unor canale de achiziție comune tuturor fișierelor alese

În mod implicit electroencefalogramele prezentate în [Secțiunea 1.2.1](#) sunt încărcate în memorie la instanțierea clasei **EEGVisualizer** și dacă nu primește o nouă configurație de execuție o va folosi pe cea din [Listarea A.1](#).

Configurația de execuție este reprezentată în formatul json(JavaScript Object Notation) și conține 3 chei.

preprocessing_pipeline_stages și **processing_pipeline_stages**

Primele două chei, **preprocessing_pipeline_stages** și **processing_pipeline_stages**, corespund celor două fluxuri de execuție din [Figura 3](#) și conțin colecții cu 0 sau mai multe obiecte de tipul:

```

1 {
2   "stage_name": "SplitInWindows",
3   "constructor_kwargs": {
4     "window_size": 1000
5   }
6 }
```

Listarea 3.1: Exemplu unui stadiu definit la nivelul configurației de execuție

Existența unui astfel de obiect în configurația unui flux implică, în timpul procesării, instanțierea clasei corespunzătoare cu parametrii specificați.

t_sne_models

Ultima cheie din configurația de execuție, **t_sne_models**, conține o colecție cu cel puțin un obiect de următoarea structură:

```

1 {
2   "model_name": "standard_parameters",
3   "parameters": {
4     "n_jobs": 6,
5     "learning_rate": "auto",
6     "init": "random",
7     "n_components": 2,
8     "early_exaggeration": 12,
9     "perplexity": 30
10  }
11 }
```

Listarea 3.2: Exemplu unui model T-SNE definit la nivelul configurației de execuție

Numărul de modele regăsite în configurație implică numărul de tablouri rezultate. Am gândit o astfel de funcționalitate deoarece hiperparametrii unui model T-SNE influențează mult rezultatul.

3.2 Fluxul de preprocesare

Fluxul de procesare și stagiile sale interne respectă același protocol de comunicare, primesc și produc o colecție de obiecte de tipul **TimeSeries**. În acest punct al execuției se fac procesările

dorite în domeniul timp. Sistemul poate fi ușor extins în viitor cu alte stagi, cu condiția să respecte protocolul impus.

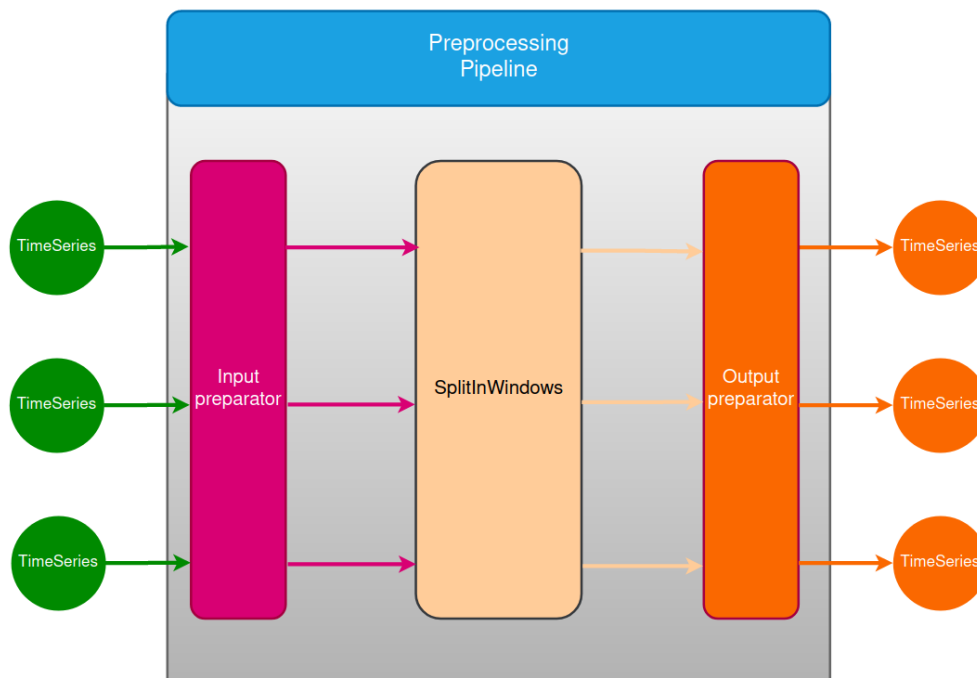


Figura 3.2: Modul de funcționare al fluxului de preprocesare

3.2.1 SplitInWindows

Dacă sistemul ar executa cu un flux vid de preprocesare atunci pentru fiecare semnal de intrare sistemul va produce un punct în tabloul rezultat.

Stagiul de preprocesare SplitInWindows multiplică un semnal în mai multe semnale, ferestrea ale celui inițial, păstrând etichetarea. Motivul pentru care a fost creat este rezolvarea numărului mic de exemple în setul de date folosit în această lucrare.

3.3 Fluxul de procesare

TODO

3.3.1 FeatureExtractor

TODO

3.4 Execuția algoritmului T-SNE

TODO

3.4.1 Despre T-SNE

TODO

3.4.2 T-SNE pe un alt set de date

TODO

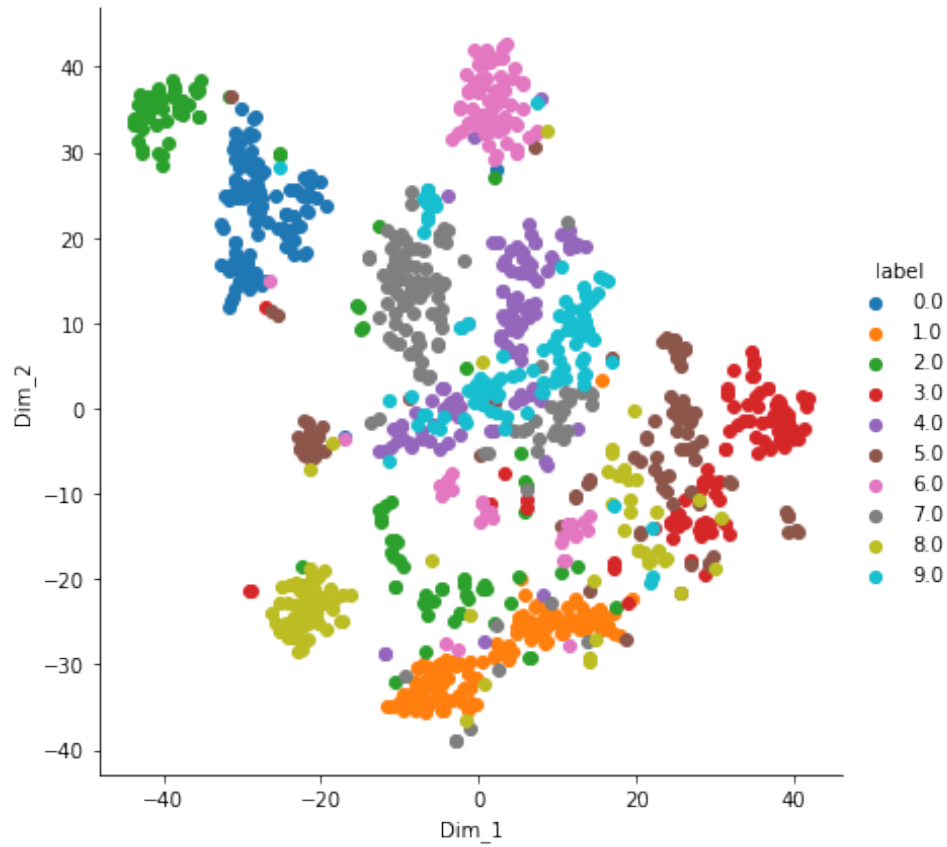


Figura 3.3: Execuția algoritmului T-SNE pe setul de date MNIST

3.4.3 T-SNE în cadrul sistemului curent

TODO

Capitolul 4

Rezultate și Concluzii

TODO

4.1 Lobii frontali și parietali ai primului șoarece(mouse1.acq)

TODO

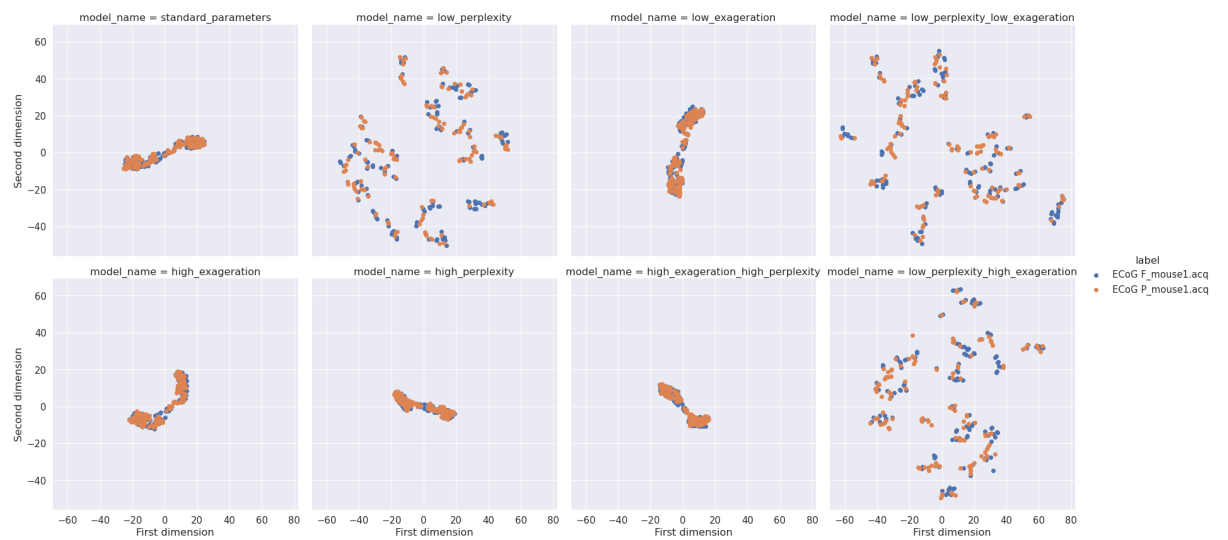


Figura 4.1: Tabloul lobiilor frontali și parietali ai primului șoarece

4.2 Lobii frontali și parietali ai celui de-al doilea șoarece(mouse2.acq)

TODO

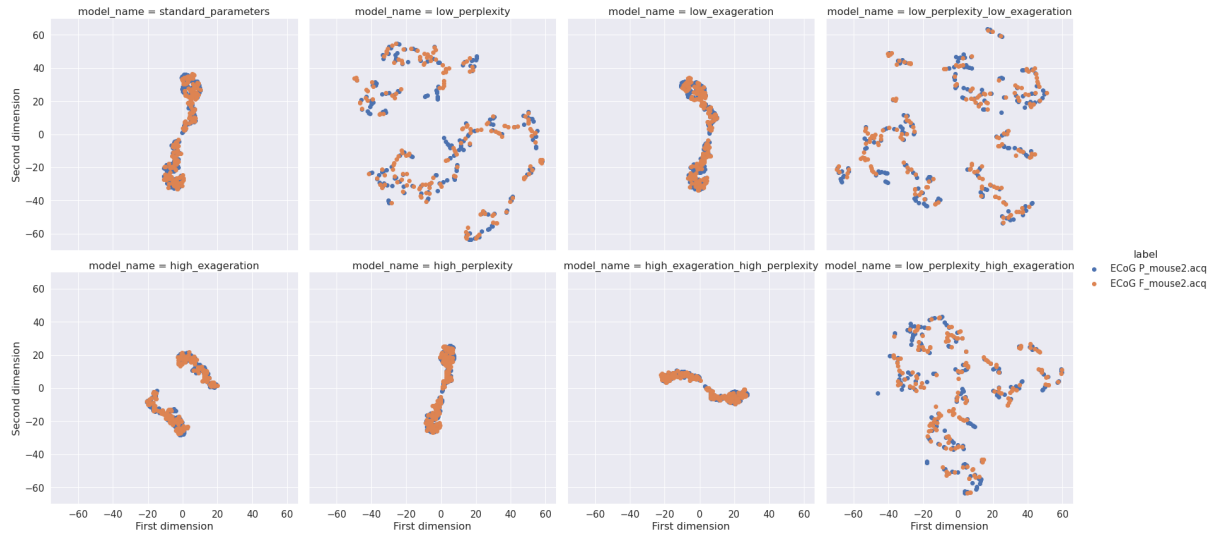


Figura 4.2: Tabloul lobiilor frontali și parietali ai celui de-al doilea șoarece

4.3 Lobii frontali și parietali ai celui de-al treilea șoarece(mouse3.acq)

TODO

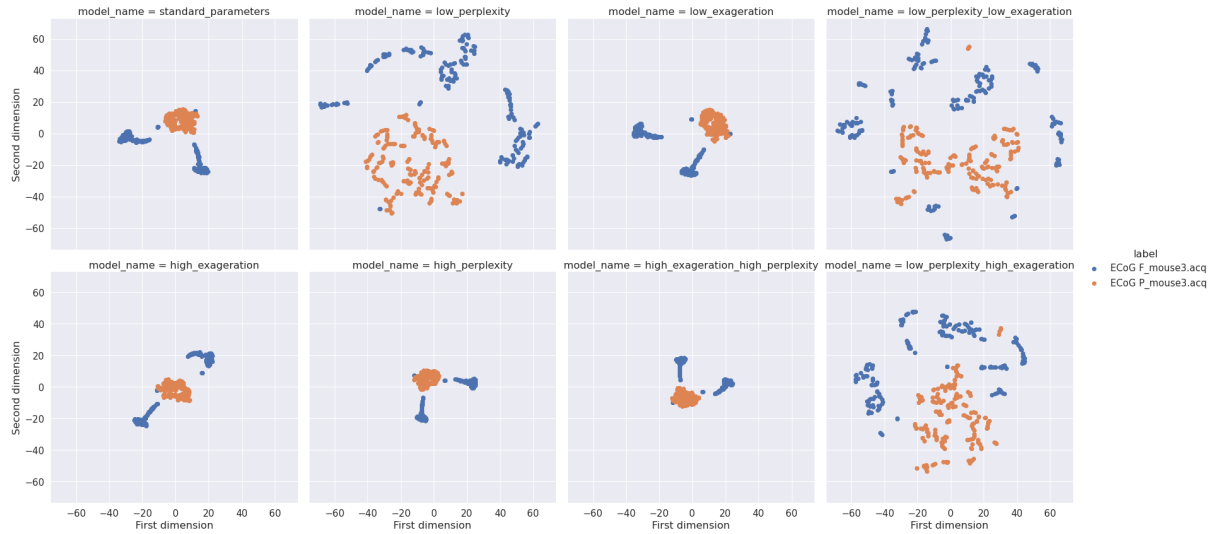


Figura 4.3: Tabloul lobiilor frontali și parietali ai celui de-al treilea șoarece

4.4 Lobii frontali și parietali ai celui de-al patrulea șoarece(mouse4.acq)

TODO

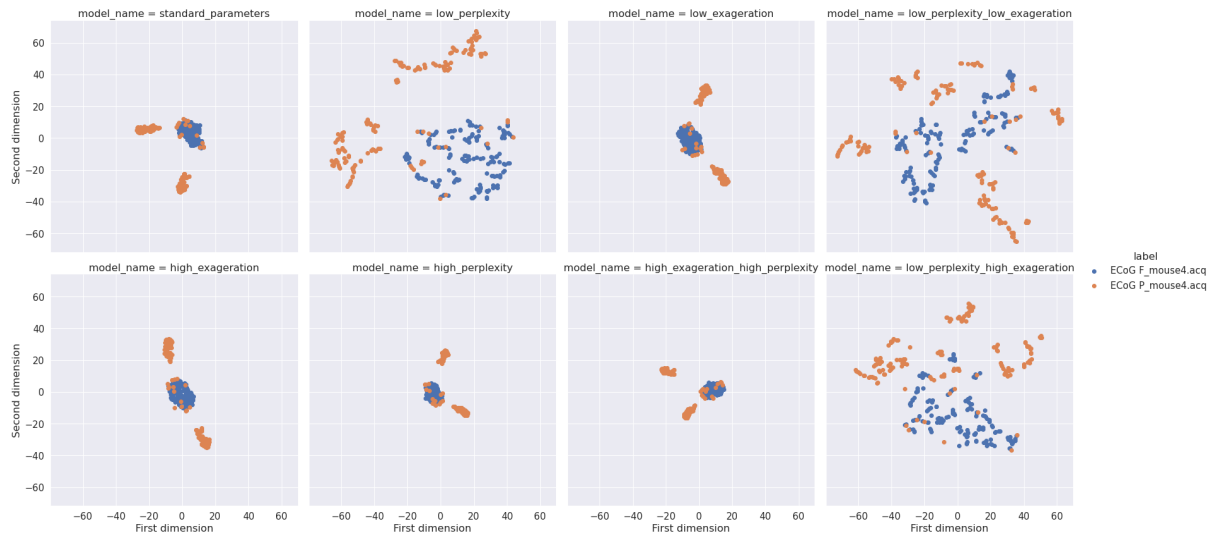


Figura 4.4: Tabloul lobiilor frontali și parietali ai celui de-al patrulea șoarece

4.5 Lobul frontal pentru toți șoarecii

TODO

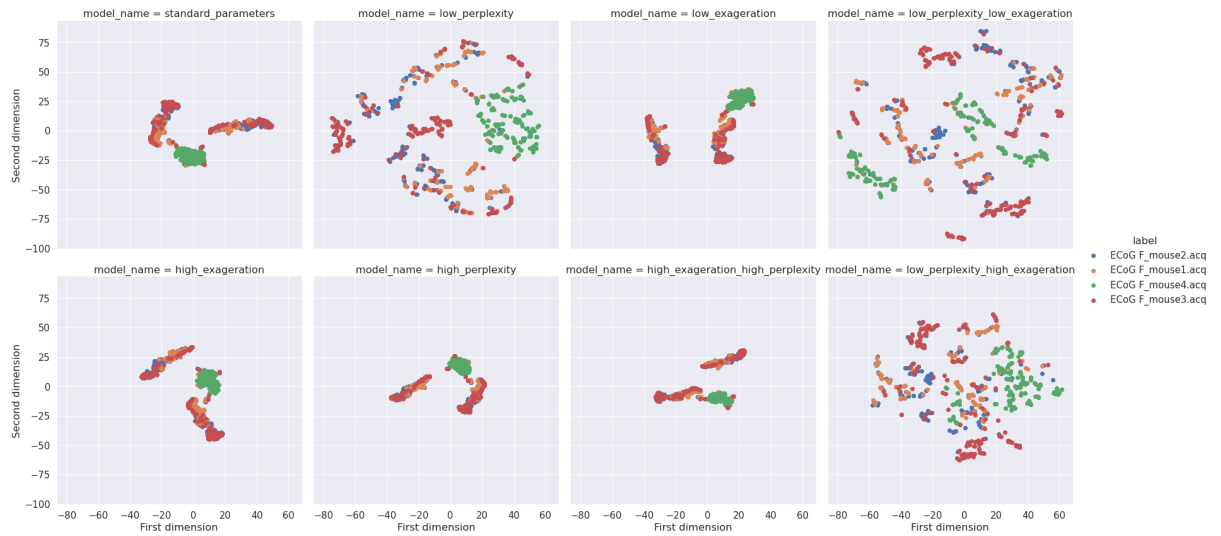


Figura 4.5: Tabloul lobiilor frontali ai tuturor șoarecilor

4.6 Lobul parietal pentru toți șoarecii

TODO

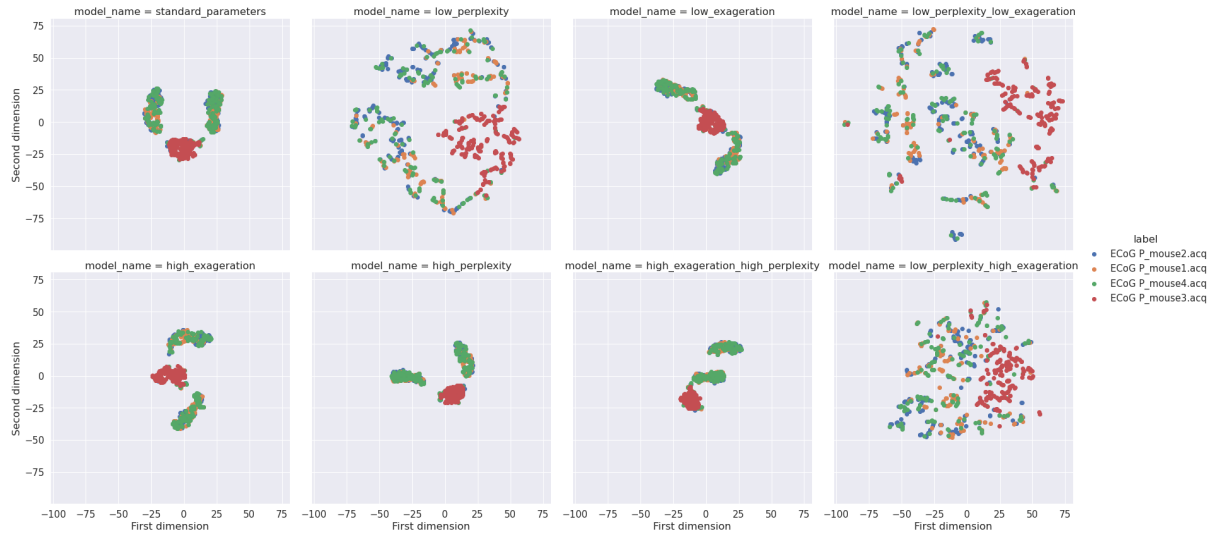


Figura 4.6: Tabloul lobiilor parietali ai tuturor șoarecilor

4.7 Lobii frontali și parietali pentru toți șoarecii

TODO

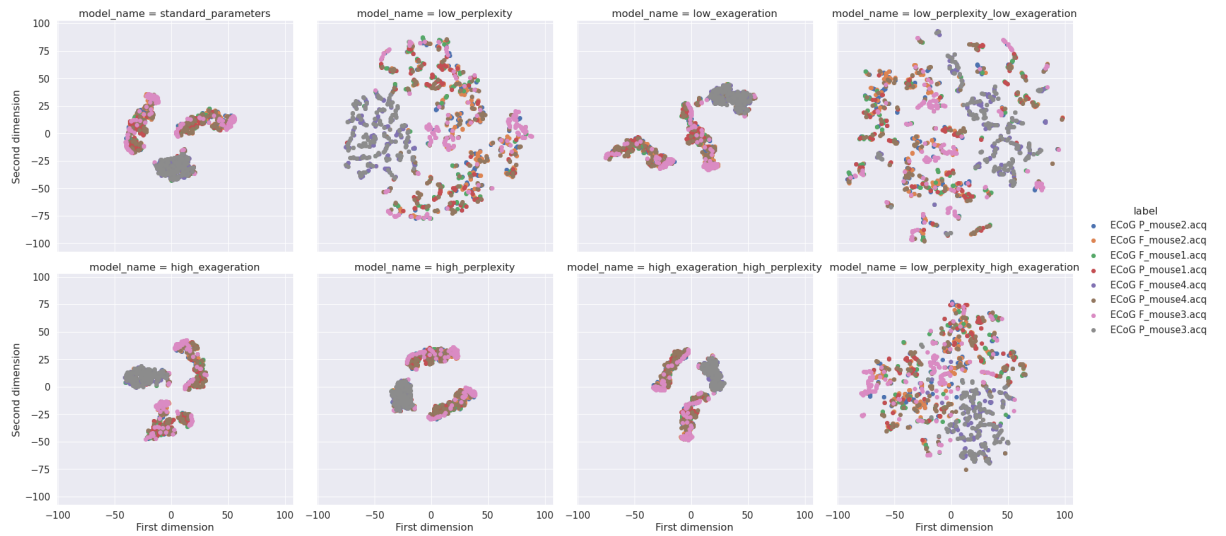


Figura 4.7: Tabloul lobiilor frontali și parietali ai tuturor șoarecilor

Capitolul 5

Viitor

TODO:

5.1 Comunicare peste HTTP

TODO:

5.2 Interfață grafică

TODO:

Appendix A

Configurații de execuție

A.1 Configurația standard de execuție

```
1 {
2   "preprocessing_pipeline_stages": [
3     {
4       "stage_name": "SplitInWindows",
5       "constructor_kwargs": {
6         "window_size": 1000
7       }
8     }
9   ],
10  "processing_pipeline_stages": [
11    {
12      "stage_name": "FeatureExtractor",
13      "constructor_kwargs": {
14        "features": [
15          "alpha_spectrum_coefficients_sum",
16          "beta_spectrum_coefficients_sum",
17          "low_gamma_spectrum_coefficients_sum",
18          "high_gamma_spectrum_coefficients_sum",
19          "delta_spectrum_coefficients_sum",
20          "theta_spectrum_coefficients_sum",
21          "mean",
22          "standard_deviation"
23        ]
24      }
25    }
26  ],
27  "t_sne_models": [
28    {
29      "model_name": "standard_parameters",
30      "parameters": {
31        "n_jobs": 6,
32        "learning_rate": "auto",
33        "init": "random",
34        "n_components": 2,
35        "early_exaggeration": 12,
```

```
36         "perplexity": 30
37     }
38 },
39 {
40     "model_name": "low_perplexity",
41     "parameters": {
42         "n_jobs": 6,
43         "learning_rate": "auto",
44         "init": "random",
45         "n_components": 2,
46         "early_exaggeration": 12,
47         "perplexity": 5
48     }
49 },
50 {
51     "model_name": "low_exaggeration",
52     "parameters": {
53         "n_jobs": 6,
54         "learning_rate": "auto",
55         "init": "random",
56         "n_components": 2,
57         "early_exaggeration": 1,
58         "perplexity": 30
59     }
60 },
61 {
62     "model_name": "low_perplexity_low_exaggeration",
63     "parameters": {
64         "n_jobs": 6,
65         "learning_rate": "auto",
66         "init": "random",
67         "n_components": 2,
68         "early_exaggeration": 1,
69         "perplexity": 5
70     }
71 },
72 {
73     "model_name": "high_exaggeration",
74     "parameters": {
75         "n_jobs": 6,
76         "learning_rate": "auto",
77         "init": "random",
78         "n_components": 2,
79         "early_exaggeration": 300,
80         "perplexity": 30
81     }
82 },
83 {
84     "model_name": "high_perplexity",
85     "parameters": {
86         "n_jobs": 6,
87         "learning_rate": "auto",
88         "init": "random",
```

```
89         "n_components": 2,  
90         "early_exaggeration": 12,  
91         "perplexity": 40  
92     }  
93 },  
94 {  
95     "model_name": "high_exaggeration_high_perplexity",  
96     "parameters": {  
97         "n_jobs": 6,  
98         "learning_rate": "auto",  
99         "init": "random",  
100        "n_components": 2,  
101        "early_exaggeration": 300,  
102        "perplexity": 40  
103    }  
104 },  
105 {  
106     "model_name": "low_perplexity_high_exaggeration",  
107     "parameters": {  
108         "n_jobs": 6,  
109         "learning_rate": "auto",  
110         "init": "random",  
111         "n_components": 2,  
112         "early_exaggeration": 300,  
113         "perplexity": 5  
114     }  
115 }  
116 ]  
117 }
```

Listarea A.1: Configurația standard de execuție

Bibliografie

- [1] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.