

# **Algoritmi de Reprezentare Rară, Învățarea Dicționarelor și Reconstrucție Rară**

## **Detectarea și Clasificarea Aritmiilor EKG pe 12 Derivații**

Tudor Pistol și Teofil Simiraș

January 16, 2025

# Cuprins

---

1. Introducere și Context
2. Problema Abordată
3. Preprocesarea Semnalelor EKG
4. Reprezentarea Rară (Sparse Coding)
5. Algoritmi Sparse (OMP și K-SVD)
6. Clasificarea Bătăilor EKG
7. Rezultate și Metodologie de Evaluare
8. Concluzii și Direcții

# Introducere și Context

---

- Semnalele EKG de 12 derivații sunt voluminoase și complexe.
- **Aritmiile cardiace** pot fi critice dacă nu sunt detectate rapid.
- Proiectul nostru:
  - Preprocesare (filtrare, segmentare)
  - **Sparse Coding** (ex. OMP + K-SVD)
  - Clasificare (SVM sau Logistic Regression)

# Problema Abordată

---

## Obiectiv principal

**Detectarea și clasificarea aritmiilor cardiace** pe EKG de 12 derivații prin *algoritmi de reprezentare rară* și *învățarea dicționarelor*.

## De ce e important?

- Aritmiile cardiace sunt frecvente și pot deveni periculoase.
- Sistem automat: reduce timpul de analiză și crește acuratețea (spitale, dispozitive portabile).

# Filtrare Butterworth / Chebyshev

---

- Folosim un filtru **band-pass** (0.5 – 40 Hz) pentru a reține componentele relevante din EKG.
- **Butterworth**: tranziție lină, fără ripple în banda de trecere.
- **Chebyshev**: roll-off mai abrupt, dar introduce ripple.

# Segmentare & Normalizare

---

- **Segmentare:** Detectăm complexul QRS (ex. algoritm Pan-Tompkins).
- Extragem fereastră (ex. 100 ms înainte, 300 ms după R-peak).
- **Normalizare:** Scalăm amplitudinile în intervalul  $[-1, 1]$ .

# Definiție Sparse Coding

---

**Reprezentare rară** a unui semnal:

$$y \in \mathbb{R}^m, \quad D \in \mathbb{R}^{m \times n}, \quad y = Dx,$$

unde  $\|x\|_0$  (# de elemente nenule) este mic.

- Dicționarul  $D$  conține *atomi* (coloane).
- Doar câțiva atomi (coeficienți) sunt nenuli în  $x$ .

# Probleme de optimizare

---

**1. Minimizare a erorii sub constrângere de raritate:**

$$\min_x \|y - Dx\|^2 \quad \text{s.t.} \quad \|x\|_0 \leq s.$$

**2. Minimizare a rarității sub constrângere de eroare:**

$$\min_x \|x\|_0 \quad \text{s.t.} \quad \|y - Dx\| \leq \varepsilon.$$



# OMP (Orthogonal Matching Pursuit)

---

- Algoritm *greedy* pentru a determina suportul vectorului rar.
- **Pași principali :**
  1. Reziduul inițial  $e = y$ , suportul  $S = \emptyset$ .
  2.  $k = \arg \max_{j \notin S} |e^T d_j|$ .
  3.  $S = S \cup \{k\}$ ; rezolvăm  $\min_{x_S} \|y - D_S x_S\|$ .
  4. Reziduu:  $e \leftarrow y - D_S x_S$ .
  5. Repetăm până  $\|e\|$  mic sau  $|S| = s$ .

# Antrenarea Dicționarului - K-SVD

---

- Obiectiv:

$$\min_{D, X} \|Y - DX\|_F^2 \quad \text{s.t.} \quad \|x_i\|_0 \leq s.$$

- Alternăm:

1. *Sparse coding* (OMP pentru fiecare coloană  $y_i$ ).
2. *Actualizare dicționar*:
  - Calculăm reziduu  $F = Y - DX$ .
  - Pentru fiecare atom  $d_j$ , se aplică SVD pe submatricea reziduală relevantă pentru a-l optimiza.

# Clasificare EKG (Normal vs. Aritmie)

---

- După ce fiecare bătaie EKG este reprezentată de un vector  $x$ , îl folosim drept *feature vector*.
- **Algoritm de clasificare:**
  1. **SVM** (kernel RBF): foarte folosit în probleme binare (normal vs. aritmie).
  2. **Regresie Logistică**: mai simplu, interpretabil.
  3. (Opțional) Random Forest, XGBoost, Rețele Neurale etc.
- Se antrenează pe train, se testează pe test; metrici: Acuratețe, Recall, F1, ROC/AUC.

## Rezultate Așteptate

---

- Acuratețe  $> 90\%$  în detectarea aritmiilor cardiace.
- Reducerea zgomotului prin *sparse reconstruction*.
- Coeficienții rari  $\implies$  separare mai bună între bătăile normale și anormale.

# Metodologie de Evaluare

---

- Împărțire **train** / **validation** / **test**.
- **Metrici:**
  - Acuratețe, Sensibilitate (Recall), Specificitate, F1, ROC/AUC.
- Comparație cu:
  - Fără sparse coding (features brute).
  - Dicționar random vs. K-SVD.

# Concluzii

---

- **Sparse Coding** (OMP + K-SVD) oferă un mod eficient de extragere a trăsăturilor EKG.
- **Clasificare** (SVM) a aritmiilor cu acuratețe ridicată.
- **Robustețe** la zgomot, potențial de compresie.

- **Detecție multi-clasă:** fibrilație atrială, flutter, blocuri AV etc.
- **Implementare embedded:** pe un dispozitiv Holter, monitorizare real-time.
- **Sparse Autoencoder:** antrenare neuronală cu constrângeri de raritate.

# Bibliografie

---

- [1] *A large scale 12-lead electrocardiogram database for arrhythmia study*, <https://physionet.org/content/ecg-arrhythmia/1.0.0/>
- [2] *Heart Arrhythmias*, [https://www.physio-pedia.com/Heart\\_Arrhythmias](https://www.physio-pedia.com/Heart_Arrhythmias)
- [3] *Cursul de Procesarea Semnalelor*, <https://cs.unibuc.ro/~crusu/ps/index.html>
- [4] *Cursul de Calcul Numeric*, <https://numeric.cs.unibuc.ro/cni.html>



**Vă mulțumim!**