Universidad Nacional del Altiplano
Escuela de Posgrado
Doctorado en Ciencias de la Computación

Data Mining

Unit 4. Data Mining with Complex Networks

Prof. Dr. Ivar Vargas Belizario

ivargasbelizario@gmail.com

2024 - I

Contenido

- Complex Networks
- Networks Modeling
- Network Measurements
- Community Detection
- Applications

Contenido

- Complex Networks
- Networks Modeling
- Network Measurements
- Community Detection
- Applications

Complex Networks

- They are the graph representation of a complex system

Complex System $\xrightarrow{\text{Graph Representation}}$ Complex Networks

Graph = Networks

Complex Networks

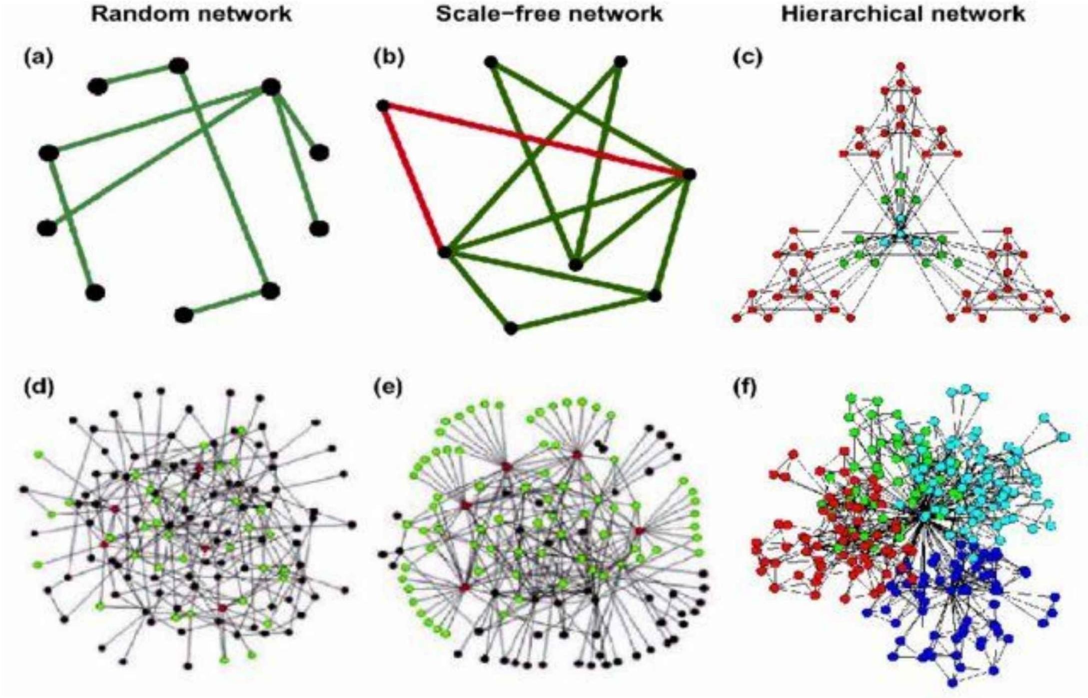
- Social Networks
- Communication
- Computer Science
  - Circuits
  - Image Processing and Analysis
- Internet
- Citations
- Electric power transmission systems
- Biomolecular Networks
- Epidemic spreading
- etc.

- Nodes ?
- Links ?

[3] https://doi.org/10.1080/00018732.2011.572452

Contenido

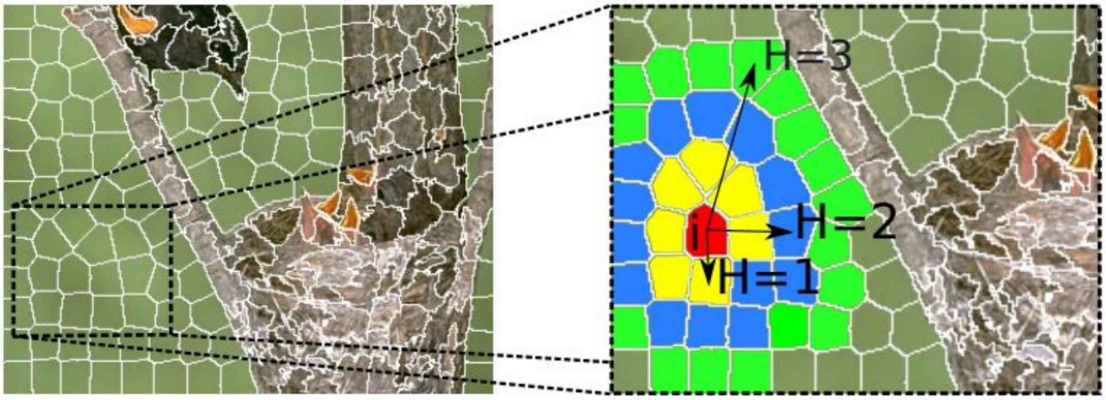
Networks Modeling



[32] https://doi.org/10.1142/9789812772367_0001

Networks Modeling

- Make an approximation to the real representation of the complex system.



(a) (b)

Figura 12 – Estratégia de criação de arestas em vários níveis. (a) Um subconjunto de superpixels; (b) Criação de arestas para o vértice i (superpixel vermelho) e seus j vizinhos para o primeiros níveis ($H = 3$) denotados com cores amarelo, azul, e verde.

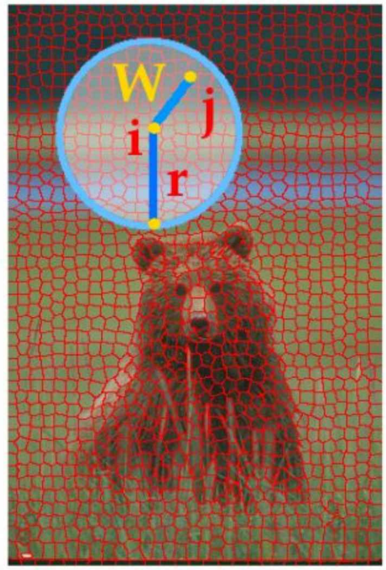
[33] https://doi.org/10.11606/T.55.2021.tde-09032021-123250

Networks Modeling

- Make an approximation to the real representation of the complex system.

$$w(i,j) = \begin{cases} e^{-\left(\frac{F^2}{\sigma^2}\right)}, & si\ j \in r, \\ 0, & \text{caso contrario} \end{cases}$$

$$w_{gEU}(i,j) = e^{-\left(\frac{Euclidean(i,j)^2}{\sigma^2}\right)}$$
$$w_{gMH}(i,j) = e^{-\left(\frac{Manhatatan(i,j)^2}{\sigma^2}\right)}$$
$$w_{gCH}(i,j) = e^{-\left(\frac{Chebyshev(i,j)^2}{\sigma^2}\right)}$$
$$w_{gCO}(i,j) = e^{-\left(\frac{1-Cosine(i,j)^2}{\sigma^2}\right)}$$
$$w_{gTA}(i,j) = e^{-\left(\frac{1-Tanimoto(i,j)^2}{\sigma^2}\right)}$$
$$w_{gFU}(i,j) = e^{-\left(\frac{1-Fu(i,j)^2}{\sigma^2}\right)}$$
$$w_{gMB}(i,j) = e^{-\left(\frac{Mahalanobis(i,j)^2}{\sigma^2}\right)}$$

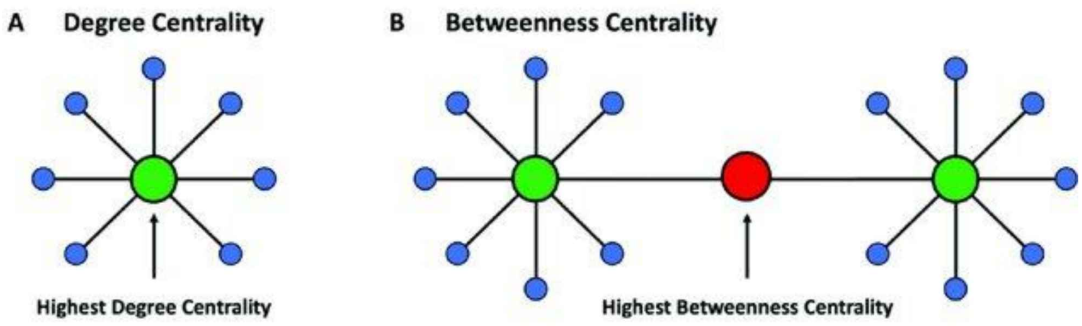


Shi and Malik (2000)

Contenido

- Complex Networks
- Networks Modeling
- Network Measurements
- Community Detection
- Applications

Network Measurements



[34] https://doi.org/10.3390/tomography8030116
[35] https://doi.org/10.1080/00018730601170527

Contenido

- Complex Networks
- Networks Modeling
- Network Measurements
- Community Detection
- Applications

Community Detection

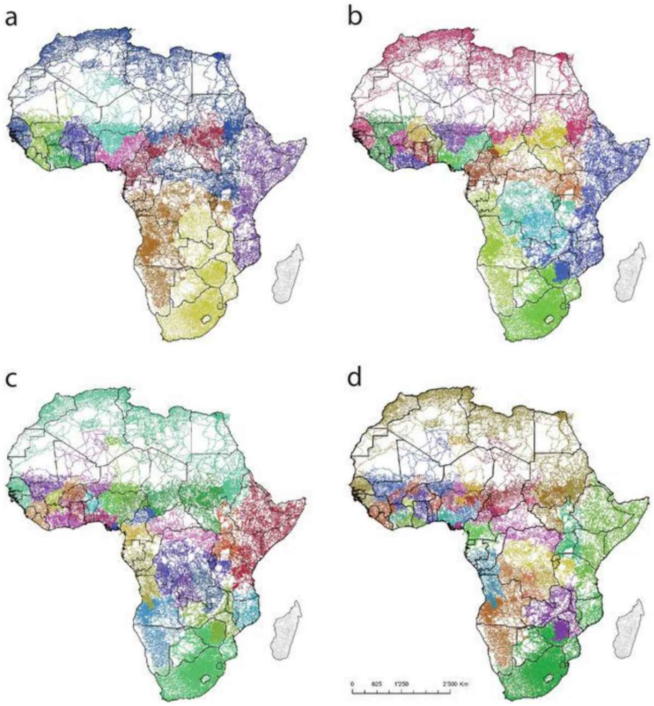
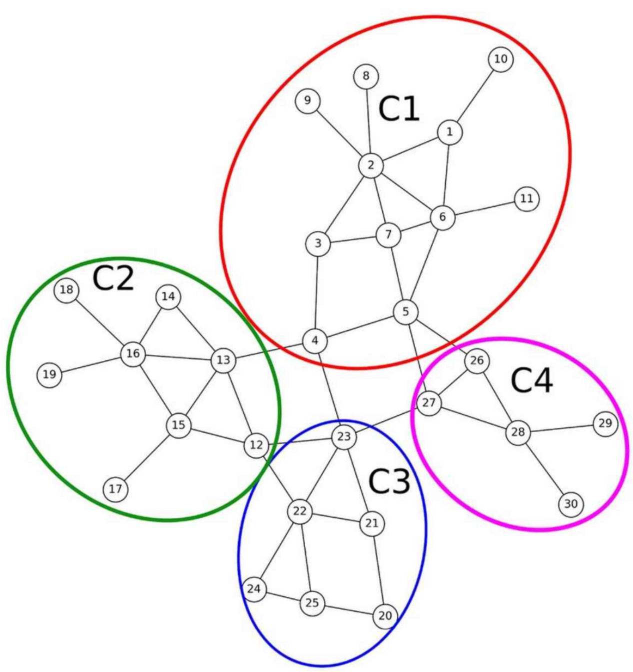


Fig. Example outputs of community detection on the unweighted Africa road network, constrained to (a) 10 communities; (b) 20 communities; (c) 30 communities and (d) 40 communities. Figure produced using ArcGIS v10.5 (www.arcgis.com).

[36] https://www.nature.com/articles/s41598-018-22969-4

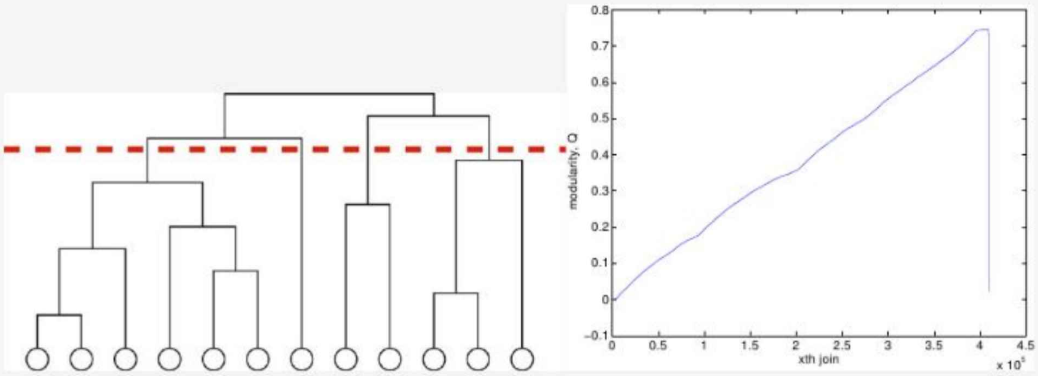
Community Detection



[37] http://dx.doi.org/10.1038/s41598-023-41460-3

Community Detection

Finding community structure in very large networks



(a) Dendrograma (b) $\triangle Q$

Figura: (a) Representação do dendrograma, para o algoritmo *fast greedy*, a linha vermelha representa o corte do dendrograma empregando o máximo valor de $\triangle Q$. (b) No eixo x é mostrado o número de uniões realizadas entre duas comunidades ao longo do algoritmo, nela é mostrada o acrescentamento de $\triangle Q$, onde $\triangle Q$ apresenta um único valor máximo.

[38] http://dx.doi.org/10.1103/PhysRevE.70.066111

Community Detection

Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks

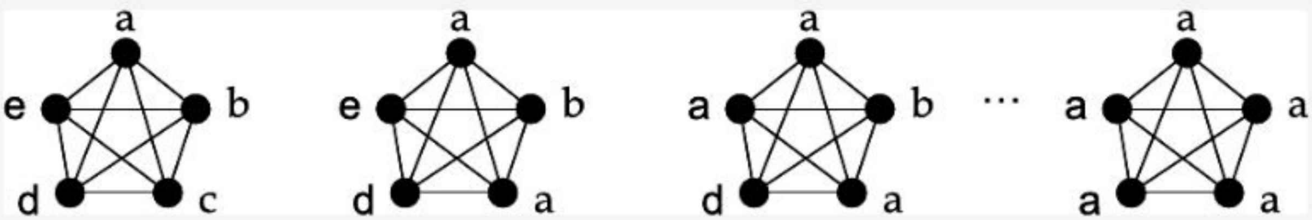


Figura: Atualização de rótulos no *label propagation*: na figura de esquerda a direita, os vértices são atualizados um por um. Neste caso existe uma grande densidade de arestas, isto faz possível que todos os vértices adquiram o mesmo rótulo.

[39] http://dx.doi.org/10.1103/PhysRevE.76.036106

Contenido

- Complex Networks
- Networks Modeling
- Network Measurements
- Community Detection
- Applications

Applications

Segmentation of Large Images with Complex Networks

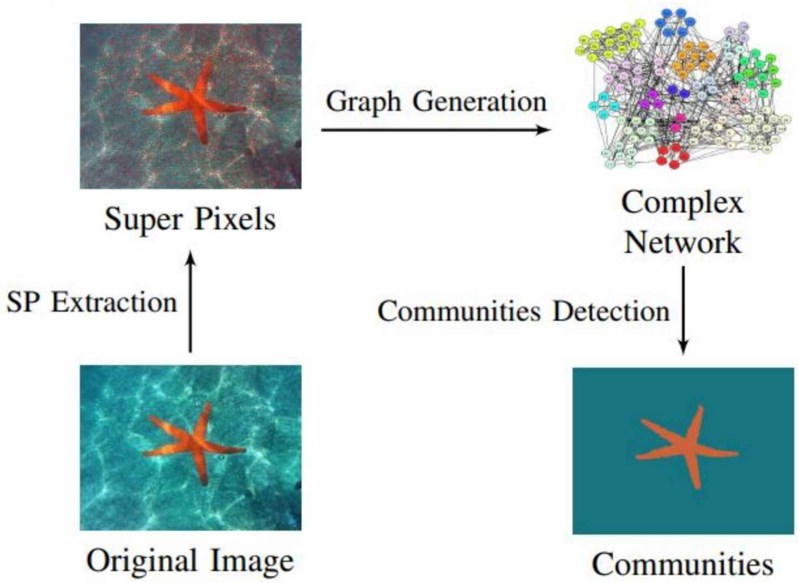


Fig. 4. Image segmentation approach based on complex network combined with super pixels. In the figure, SP = Super Pixel.

[40] https://doi.org/10.1109/SIBGRAPI.2012.13

Applications

Automatic image segmentation based on label propagation

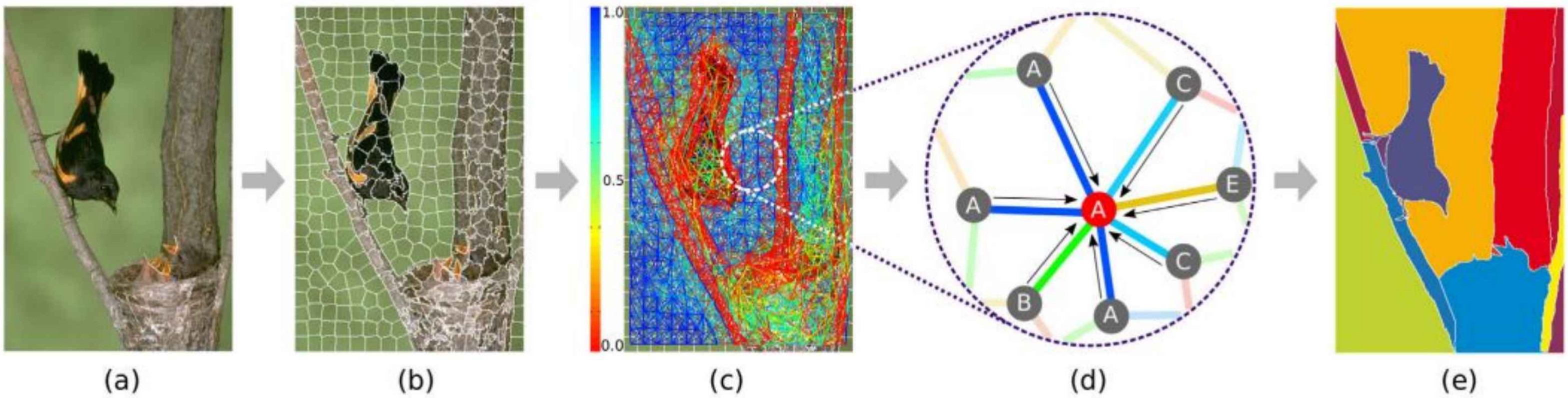


FIGURE 1 Illustration of our method. (a) Input image; (b) super-pixel extraction; (c) graph building. Blue edges indicate high similarity among vertices, whereas red edges represent low similarity; (d) label propagation. The red vertex is assigned the most frequent label in its neighbourhood, (e) segmentation result, that is, regions containing super-pixels with the same label

[41] https://doi.org/10.1049/ipr2.12242

Applications

Texture analysis and classification: A complex network-based approach

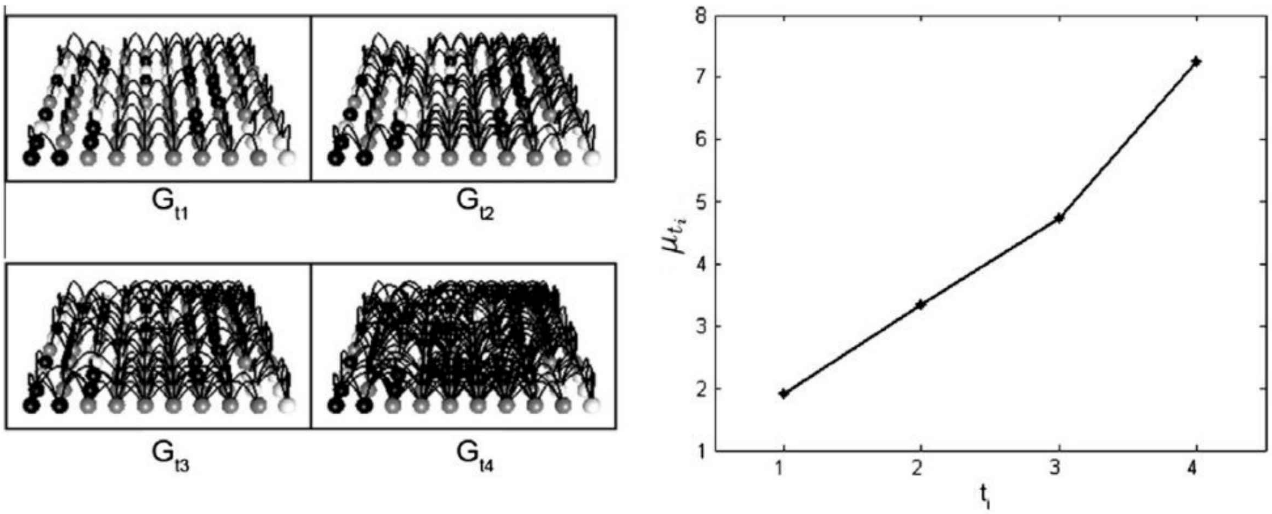


Fig. 2. Left: a texture example modeled as a Complex network, different threshold values ($t_1 \cdots t_4$) make different topological features. Right: complex network characterization by its evolution (mean degree at each threshold t_i.

[42] https://doi.org/10.1016/j.ins.2012.07.003

Applications

Texture analysis and classification: A complex network-based approach

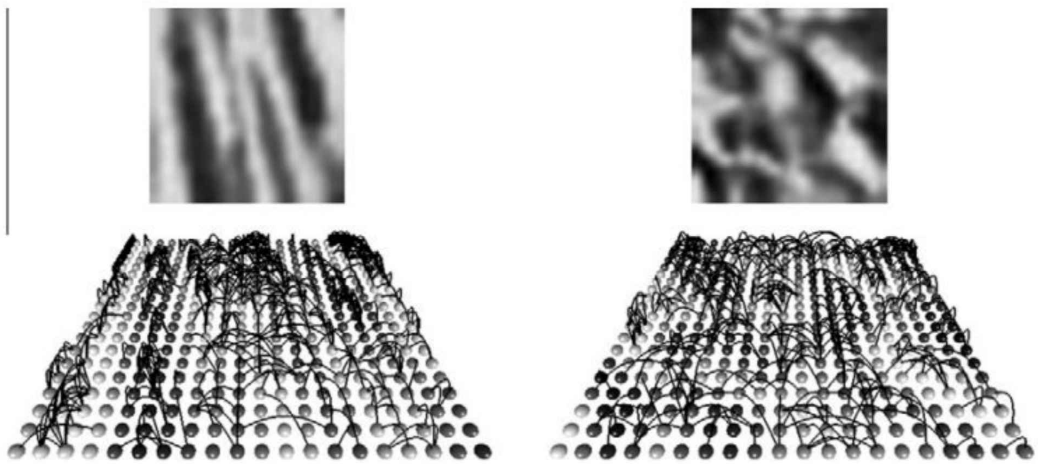


Fig. 3. Two complex networks, at same threshold value, for two different texture samples. The differences in their topological features results in measurements which can be used as texture descriptors.

[42] https://doi.org/10.1016/j.ins.2012.07.003

Applications

Multilayer complex network descriptors for color–texture characterization

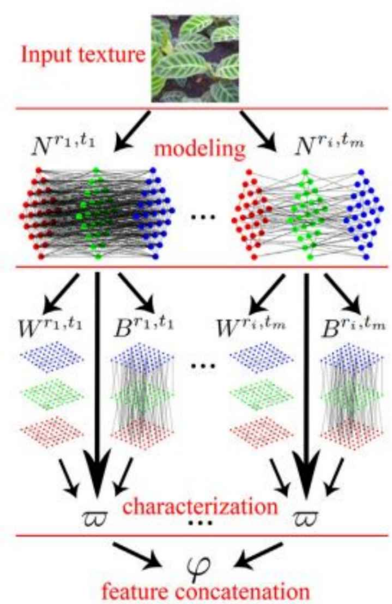


Fig. 6. Step-by-step of the proposed method. The modeling step considers the network dynamic evolution, thus combining each parameter r and t results in a set of networks N. The characterization consists on first obtaining the subnets W and B and then quantify the structure of all networks (N, W and B) with degree (k) and clustering (c) statistics ϖ. The final feature vector φ is the concatenation of statistics from all networks.

[43] https://doi.org/10.1016/j.ins.2019.02.060

Applications

Epilepsy Detection From EEG Using Complex Network Techniques: A Review

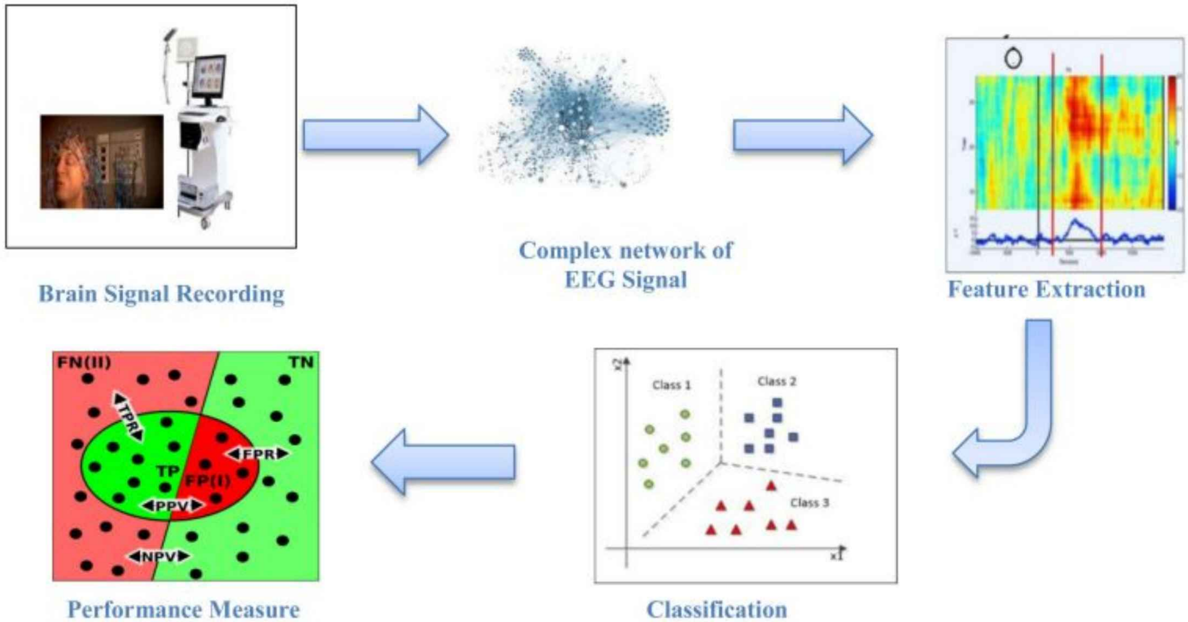


Fig. 2. Schematic representation of general sequence of steps followed by graph theory based approaches for epileptic seizure detection from brain EEG signals.

[44] https://doi.org/10.1109/RBME.2021.3055956

Universidad Nacional del Altiplano
Escuela de Posgrado
Doctorado en Ciencias de la Computación

Data Mining

Unit 4. Data Mining with Complex Networks

Gracias

Prof. Dr. Ivar Vargas Belizario

ivargasbelizario@gmail.com

2024 - I