Data science tools in neuroscience

1. Circular statistics
2. Spatial statistics
3. Clustering
4. Classification
5. Bayesian networks
6. Optimization
1. Circular statistics. Branching angles of dendritic trees

- Dendritic arbors of layer III pyramidal neurons
- Seven cortical areas: M1, M2, S1, S2, V1, V2, PrL/II

Bielza, Benavides-Piccione, López-Cruz, Larrañaga, DeFelipe (2014). Branching angles of pyramidal cell dendrites follow common geometrical design principles in different cortical areas. Scientific Reports, 4, article 5909
2. Spatial statistics. Distribution of synapses

- 1695 synaptic junctions from ten 3D samples
- Neuropil of layer III of the young rat somatosensory cortex

2. Spatial statistics. Distribution of synapses

3. Clustering of human pyramidal cell dendritic spines

- 3D reconstructed 7,000 spines from layer III pyramidal neuron human cingulate cortex (aged 40 and 85)
- Each spine is characterized with 54 morphological variables (some of them directional variables)

3. Clustering. Probabilistic

Finite mixture modeling

\[
f(x \mid \theta) = \sum_{k=1}^{K} \pi_k f_k(x \mid \theta_k)
\]

\[
f_k(x \mid \theta_k) \sim \mathcal{N}(x \mid \mu_k, \Sigma_k)
\]
3. Probabilistic clustering of human pyramidal cell dendritic spines

Cluster 1

Cluster 2

Cluster 3

Cluster 4

Cluster 5

Cluster 6

Cluster 7

Cluster 8

4. Classification. Interneuron classification

4. Classification. Interneuron classification

4. Classification. Interneuron classification

<table>
<thead>
<tr>
<th></th>
<th>$X_1=$Length</th>
<th>$X_2=$Terminal degree</th>
<th>...</th>
<th>$X_{2,886}=$Partition asymmetry</th>
<th>$C=F_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2947.20</td>
<td>1648</td>
<td></td>
<td>9.67</td>
<td>Martinotti</td>
</tr>
<tr>
<td>2</td>
<td>3545.45</td>
<td>1942</td>
<td></td>
<td>5.33</td>
<td>Large basket</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>1684.50</td>
<td>480</td>
<td></td>
<td>6.00</td>
<td>Chandelier</td>
</tr>
</tbody>
</table>

**Algorithm**

**Model**

**Question**

**Answer**

<table>
<thead>
<tr>
<th></th>
<th>$X_1=$Length</th>
<th>$X_2=$Terminal degree</th>
<th>...</th>
<th>$X_{2,886}=$Partition asymmetry</th>
<th>$C=F_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>n+1</td>
<td>1619.28</td>
<td>1160</td>
<td></td>
<td>4.00</td>
<td>Chandelier</td>
</tr>
</tbody>
</table>

---

4. Classification. Algorithms

K nearest neighbors

Classication trees

Naïve Bayes

Neural networks

Support vector machines

Random forest

4. Classification. Feature subset selection

Filter approach

Wrapper approach

5. Bayesian networks. Factorization of the JPD

\[
p(A, N, S, D, P) = p(A)p(N|A)p(S|A)p(D|N, S)p(P|S)
\]

5. Bayesian networks. Reasoning via inference

5. Bayesian networks. Reasoning via inference

Evidence: “Stroke = yes”

5. Bayesian networks. Reasoning via inference

Evidence: “Stroke = yes, Neuronal Atrophy=yes”

5. Bayesian networks. Reasoning via inference

Evidence: “Stroke = yes, Neuronal Atrophy = yes, Age = young”

5. Bayesian networks. Modeling and simulation of pyramidal cells

- 3D reconstructions of 90 pyramidal cells
- Mouse neocortex layer III


PDQ-39 and EQ-5D: quality of life instruments to measure the degree of disability in PD

39-item Parkinson’s Disease Questionnaire: a specific instrument

PDQ-39 captures patient’s perception of his illness covering 8 dimensions:

1. Mobility
2. Activities of daily living
3. Emotional well-being
4. Stigma
5. Social support
6. Cognitions
7. Communication
8. Bodily discomfort


**European Quality of Life - 5 Dimensions: a generic instrument**

EQ-5D is a generic measure of health for clinical and economic appraisal

- **Mobility**
  - I have no problems in walking about
  - I have some problems in walking about
  - I am confined to bed

- **Self-care**
  - I have no problems with self-care
  - I have some problems washing and dressing myself
  - I am unable to wash and dress myself

- **Usual activities** (eg. work, study, housework, family or leisure activities)
  - I have no problems with performing my usual activities
  - I have some problems with performing my usual activities
  - I am unable to perform my usual activities

- **Pain/discomfort**
  - I have no pain or discomfort
  - I have moderate pain or discomfort
  - I have extreme pain or discomfort

- **Anxiety/depression**
  - I am not anxious or depressed
  - I am moderately anxious or depressed
  - I am extremely anxious or depressed

---


**Mapping PDQ-39 to EQ-5D**

<table>
<thead>
<tr>
<th>PDQ(_1)</th>
<th>PDQ(_2)</th>
<th>...</th>
<th>...</th>
<th>PDQ(_{39})</th>
<th>EQ(_1)</th>
<th>EQ(_2)</th>
<th>EQ(_3)</th>
<th>EQ(_4)</th>
<th>EQ(_5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1</td>
<td>...</td>
<td>...</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>...</td>
<td>...</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>...</td>
<td>...</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>...</td>
<td>...</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>...</td>
<td>...</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>...</td>
<td>...</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

\[ h : (PDQ_1, ..., PDQ_{39}) \rightarrow (EQ_1, ..., EQ_5) \]


6. Optimization. Optimal wiring of pyramidal cells

- Analyze the neuronal wiring of 288 3D reconstructed complete basal arborizations of pyramidal cells
- Layers II, III, IV, Va, Vb and VI of the hindlimb somatosensory cortical región of P14 rats

6. Optimization. Optimal wiring of pyramidal cells

![Mean wiring length (µm)](chart)

**Fig 2. Mean wiring length (real vs. optimized).** Mean wiring length (µm) of the 48 analyzed cells in each cortical layer (red) versus mean wiring length of the shortest arborizations found by our optimization algorithm for each layer (green). The optimization algorithm found an equal or slightly better (shorter) wiring for all the neurons in all the layers. We found the biggest difference with respect to the real wiring in layer Va, where the synthetic wiring was, on average, 2.06% shorter than the real wiring. The smallest difference occurred in layer IV, where the optimized wiring was, on average, 1.01% shorter than the real wiring.

I INTRODUCTION
1. Neuroscience

II STATISTICS
2. Exploratory data analysis
3. Probability theory and random variables
4. Probabilistic inference

III SUPERVISED PATTERN RECOGNITION
5. Performance evaluation
6. Feature subset selection
7. Non-probabilistic classifiers
8. Probabilistic classifiers
9. Meta-classifiers
10. Multi-dimensional classifiers

IV UNSUPERVISED PATTERN RECOGNITION
11. Non-probabilistic clustering
12. Probabilistic clustering

V PROBABILISTIC GRAPHICAL MODELS
13. Bayesian networks
14. Markov networks

VI SPATIAL STATISTICS
15. Spatial statistics
Computational Intelligence Group

http://cig.fi.upm.es/