Photo-Z redshift reconstruction using a constructive multilayer perceptron

Motivations: Photo-Z redshift reconstruction

On the realistic validation of photometric redshifts, or why Teddy will never be Happy

R. Beck1*, C.-A. Lin2,3, E. E. O. Ishida4, F. Gieseke5, R. S. de Souza6,7, M. V. Costa-Duarte7,8, M. W. Hattab9, A. Krone-Martins10, for the COIN Collaboration

1Department of Physics of Complex Systems, Eotvos Loránd University, Budapest 1117, Hungary
2Service d’Astrophysique, CEA Saclay, Orme des Merisiers, Bât 509, 91191 Gif-sur-Yvette, France
3Fenglin Veteran Hospital, 2 Zhongzheng Rd. Sec. 1, Fenglin Township, Hualien 97544, Taiwan
4Laboratoire de Physique Corpusculaire, Université Clermont-Auvergne, 4 Avenue Blaise Pascal, 63178, Aubière Cedex, France
5University of Copenhagen, Segersåsvej 11, 2200 Copenhagen, Denmark
6MTA Eötvös University, EIRSA “Lendület” Astrophysics Research Group, Budapest 1117, Hungary
7Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, R. do Matão 1226, 05508-090, SP, Brazil
8Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands.
9Center for Biomarker Research and Personalized Medicine, Virginia Commonwealth University, Richmond, VA, USA

- **Challenge?**
  - Lack of spectroscopic coverage in feature space (e.g. colours and magnitudes)
  - Mismatch between photometric error distributions associated with the spectroscopic and photometric samples.
Motivations: Photo-Z redshift reconstruction

- **Regression task**
  - Techniques:
    - Decision tree, IBL, Bayesian networks, Lattice-based,
    - SVM, Artificial Neural Networks (ANN), …

- Different models of ANN:
  - …
  - Multilayer perceptron
  - …

- **Choice of the ANN structure?**
  - Application to photo-Z redshift reconstruction
Motivations

This story might be apocryphal, but it doesn't really matter. It is a perfect illustration of the biggest problem behind neural networks. Any automatically trained net with more than a few dozen neurons is virtually impossible to analyze and understand. One can't tell if a net has memorized inputs, or is 'cheating' in some other way. A promising use for neural nets these days is to predict the stock market. *Even though initial results are extremely good, investors are leery of trusting their money to a system that nobody understands*.

Neil Fraser [2003]
Problem: Settings

Multilayer perceptron process

- Input layer (n neurons)
- Output layer (m neurons)

Process:

1. Choose a topology
2. Choose an activation function
   - (sigmoid, …)
3. Initialize the connection weights
4. Train the neural network
   - (BackPropagation, …)
Problem: Settings

Multilayer perceptron *topology*

- Input layer (n neurons)
- Output layer (m neurons)

How many hidden layers?
- How many neurons per layer?
- What connection policy?

Input layer (n neurons) → Hidden layer → Output layer (m neurons)
Assumptions:

- MLP Feedforward networks are universal approximators
Problem: Background

- **Existing approaches**
  - Adhoc approach
    - One hidden layer: number of units equal to the average between the number of output units and the number of input units
    - ....

- **Automatic approaches**
  - Dynamic construction of ANN from the training set
  - Use of an apriori domain knowledge (set of rules)
  - Concept-lattice based ANN
Propose an automatic approach of defining interpretable ANN architecture when the domain knowledge is not available

- E. Mephu Nguifo et al., M-CLANN: Multiclass Concept Lattice-Based Artificial Neural Network. *Constructive Neural Networks* 2009: 103-121.

- Lauraine Tiogning Kueti et al., Boolean factors based Artificial Neural Network. *IJCNN 2016* : 819-825

- Norbert Tsopzé et al., Towards a generalization of decompositional approach of rule extraction from multilayer artificial neural network. *IJCNN 2011* : 1562-1569
Our proposal

Architecture of ANN:

- Input layer = input data
  - One (neuron) unit for each attribute [+ bias]

- Hidden layer: one
  - One neuron = one Boolean factor
  - Input layer are fully connected to hidden layer

- Output = one neuron (Regression)
## Experimentations

<table>
<thead>
<tr>
<th>Jeu de données</th>
<th>Nombre d'objets</th>
<th>Taille (après prétraitement)</th>
<th>Nombre de bandes</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHAT2</td>
<td>316</td>
<td>52 Ko</td>
<td>18</td>
</tr>
<tr>
<td>PHAT1</td>
<td>515</td>
<td>84 Ko</td>
<td>18</td>
</tr>
<tr>
<td>PHAT0</td>
<td>1984</td>
<td>328 Ko</td>
<td>18</td>
</tr>
<tr>
<td>SDSS DR9 Data 1</td>
<td>5000</td>
<td>621 Ko</td>
<td>10</td>
</tr>
<tr>
<td>SDSS DR9 Data 2</td>
<td>12000</td>
<td>1477 Ko</td>
<td>10</td>
</tr>
<tr>
<td>Deep2 DR4_sans_1</td>
<td>10838</td>
<td>735 Ko</td>
<td>6</td>
</tr>
<tr>
<td>Deep2 DR4_sans_0_1</td>
<td>11784</td>
<td>677 Ko</td>
<td>6</td>
</tr>
<tr>
<td>SDSS DR10 Data 1</td>
<td>300 000</td>
<td>32 119 Ko</td>
<td>10</td>
</tr>
<tr>
<td>SDSS DR10 Data 2</td>
<td>500 000</td>
<td>53 536 Ko</td>
<td>10</td>
</tr>
</tbody>
</table>
Experimentations
Experimentations
Experimentations (Results)
## Experimentations (Results in the Literature)

| class      | $\text{std}(\Delta z_{\text{norm}})$ | $\text{bias}(\Delta z_{\text{norm}})$ | $|\Delta z_{\text{norm}}| > 0.15$ |
|------------|--------------------------------------|---------------------------------------|----------------------------------|
| [2] galaxies | 0.041                                | -0.003                                | 0.99%                            |
| [3] galaxies | $\sigma_{68} = 0.03$                | -0.001                                | 1.56%                            |
| [9] galaxies | $\sigma_{68} = 0.0248$              | 0.0008                                | 0.73%                            |
| [13] quasars | 0.15                                 | 0.032                                 | $> 0.3 : 6.53\%$                 |
| [7] galaxies | 0.0490                               | -0.0081                               | 7.6%                             |
| [10] galaxies | 0.024                                | 0.0                                    | 1.51%                            |
| [6] galaxies | 0.0205                               | 0.00005                               | 4.11%                            |
Conclusion

- BF-ANN, new approach to find **interpretable** ANN architecture when domain knowledge is not available
  - Semantic of neuron
  - Two variants
  - Preliminary validation seems promising

- Next:
  - Rules extraction (Tsopze et al. IJCNN 2011)
Thanks !