Syllogisms

- Quantified statements in one of four moods (All, Some, Some ..., not, and None)
- Classical syllogisms consist of two premises and one conclusion
- The premises contain terms that can be arranged in one of four so-called figures
- Structurally, there are 64 possible syllogisms
- First-Order Logics is not able to account for human performance

Research Question

- Does the representation of data have an impact on general model performance?
- Which features of the data are important/necessary?
- How dense is the information encoded in the data?
- Can we find encodings for behavioural data enabling the use of general models from computer science and artificial intelligence?
- Is machine learning generally suited to analyze behavioral data?
- Can we use representation learning techniques to infer meta-information from the data (e.g., dependencies, redundancies, noise, etc.)?

The Autoencoder

- Artificial neural network model finding minimized latent state representation for given inputs
  - The Encoder component represents a function to compress the inputs into a latent state
  - The Decoder component recovers the original data from the latent state
  - Trained via general gradient descent optimization algorithms
  - After training, encoder and decoder can be applied independently
  - Applied to reasoning data, the autoencoder can be used to find dense latent representations of data

Encodings

- Automated modeling approaches benefit from rich data
- Categorical data is not suited for numerical methods
- Standard encodings such as onehot encoding make categories accessible
- A syllogistic answer can be encoded using 9 bits:

<table>
<thead>
<tr>
<th>All</th>
<th>Some</th>
<th>Some not</th>
<th>None</th>
<th>NVC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ac</td>
<td>ca</td>
<td>ac</td>
<td>ac</td>
<td>ca</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- The usual experimental record of 64 answers produce a sparse onehot vector of dimensionality 576

Conclusions

- Neural networks are capable of learning from reasoning data
- The autoencoder manages to substantially compress the data (576 to 42)
- The results suggest that behavioral data in its categorical form is highly sparse
- By finding optimized representations, it might be possible to increase general model performance
- In prediction tasks, the autoencoder is able to recover 50% missing data with a state-of-the-art precision of 48% (random baseline of 11%)
- Individual patterns can be exploited (better performance than just following the most frequent answer)
- Autoencoder can be used as baseline model for future evaluations

References