Integrated Dimensionality Reduction and Sequence Prediction using LSTM

Poster · March 2018
DOI: 10.13140/RG.2.2.28577.30563

2 authors, including:

Lambert Schomaker
University of Groningen
220 PUBLICATIONS 4,282 CITATIONS

Some of the authors of this publication are also working on these related projects:

Making Sense of Illustrated Handwriten Archives View project

MPS - Medieval Paleographic Scale View project
Integrated Dimensionality Reduction and Sequence Prediction using LSTM

Emmanuel Okafor and Lambert Schomaker
Institute of Artificial Intelligence and Cognitive Engineering, University of Groningen, The Netherlands

**Problem**
- Most industrial or complex processes present temporal dependencies which stretch over a long time.
- The underlying patterns in these processes can be extremely non-linear.
- Use of linear predictive model (ARMA/ARIMA) is not suitable.
- Hidden Markov Model has prediction limitation when dealing with temporal dependencies that stretch over long durations.

**Objectives**
- Use of external and a proposed integrated dimensionality reduction LSTM predictive systems for predicting message logs from industrial machines.
- Conversion of nominal codes (raw codes) to other vectorial paradigms to obtain better correlated patterns.

**Methods**
- External Methods: Recurrent Neural Networks (RNN) [3-7]
  - ID-LSTM Prediction on OHE codes during training and testing phases (left plot) and index predictions (right plot) over a duration of 10K time-counts.
  - Use of linear predictive model (ARMA/ARIMA) is extremely non-linear.
  - Hidden Markov Model has prediction limitation when dealing with temporal dependencies that stretch over long durations.

**Proposed Method:** Integrated Dimensionality-reduction LSTM

**Results**
- ID-LSTM Prediction on OHE codes during training and testing phases (left plot) and index predictions (right plot) over a duration of 10K time-counts.
- The left and right plots show the confusion matrix, that is, the plot of the output predictions against their target values for both training and testing phases respectively for subset 9.

<table>
<thead>
<tr>
<th>Subset</th>
<th>Time counts</th>
<th>No. of Index</th>
<th>No. of Machine</th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subset 1</td>
<td>0-1.54M</td>
<td>948</td>
<td>20</td>
<td>0.9826</td>
<td>0.9751</td>
</tr>
<tr>
<td>Subset 2</td>
<td>1.54-3.09M</td>
<td>606</td>
<td>30</td>
<td>0.9979</td>
<td>0.9695</td>
</tr>
<tr>
<td>Subset 3</td>
<td>3.09-4.63M</td>
<td>535</td>
<td>36</td>
<td>0.9886</td>
<td>0.9624</td>
</tr>
<tr>
<td>Subset 4</td>
<td>4.63-6.18M</td>
<td>619</td>
<td>48</td>
<td>0.9981</td>
<td>0.9921</td>
</tr>
<tr>
<td>Subset 5</td>
<td>6.18-7.73M</td>
<td>620</td>
<td>62</td>
<td>0.9837</td>
<td>0.9806</td>
</tr>
<tr>
<td>Subset 6</td>
<td>7.73-9.27M</td>
<td>675</td>
<td>109</td>
<td>0.9962</td>
<td>0.9347</td>
</tr>
<tr>
<td>Subset 7</td>
<td>9.27-10.8M</td>
<td>648</td>
<td>64</td>
<td>0.9205</td>
<td>0.9293</td>
</tr>
<tr>
<td>Subset 8</td>
<td>10.8-12.3M</td>
<td>679</td>
<td>95</td>
<td>0.9973</td>
<td>0.9976</td>
</tr>
<tr>
<td>Subset 9</td>
<td>12.3-13.9M</td>
<td>717</td>
<td>116</td>
<td>0.9943</td>
<td>0.9681</td>
</tr>
<tr>
<td>Subset 10</td>
<td>13.9-15.4M</td>
<td>624</td>
<td>263</td>
<td>0.9871</td>
<td>0.9268</td>
</tr>
</tbody>
</table>

**Data Representations**
- One-Hot-Encoding Codes
- 2-DIM Principal Component Analysis (PCA) Codes

**Future Directions**
- We suggest that it may be possible to combine the proposed model with an early anomaly detection algorithm.
- To allow continuous prediction of physical problems in the machines generating the message logs.
- Optimization of LSTM-based feature dimensionality reduction in a realistically large dataset.

**References**

This research was supported by the MANTIS