Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

Danilo Dessì¹, Mauro Dragoni², Mirko Marras¹, Diego Reforgiato Recupero¹

¹Department of Mathematics and Computer Science, University of Cagliari
²Fondazione Bruno Kessler

{danilo_dessi, mirko.marras, diego.reforgiato}@unica.it  dragoni@fbk.eu
• Today more and more people express their opinions in online platforms
• The use of Artificial Intelligent approaches has showed remarkable improvements in many domains
• Recent methods to capture and represent knowledge are Word Embeddings
There is a technological transfer of teaching material, learning tools, lessons in e-learning platforms.

Students express their opinions about courses.

The analysis of students’ opinions is useful to evaluate courses quality and to make courses recommendation.
Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

Deep Learning

- The Deep Learning refers to Machine Learning methods inspired by biological nervous systems that are able to learn from data
- They have nodes that simulate neurons functions
- The output of a network depends on connections between network layers and functions that nodes implement
Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

Deep Learning VS Common Machine Learning

- In a Deep Learning approach each layer can see what others did in order to improve results
- Common approaches: Random Forests, Simple Neural Networks, Support Vector Machines
Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

Our Deep Learning Model

Data representation

Embeddings Layer

Bidirectional LSTM

Attention

Dense

Model

Double direction sequential classification of patterns that appear in data

Reference to the input network for contextualizing current results

Densely-connected layer that provides an unique result
Word Embeddings

- Words Embeddings are numerical vectors that represent words.
- They capture the syntactic and semantic of words functions.

\[ \text{King} - \text{Man} + \text{Woman} = \text{Queen} \]
Word Embeddings Generators

Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

04 October 2018
Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

The COCO Dataset

43K Courses

16K Instructors

2,5M Learners

4,5M Ratings

1,2M Reviews
Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

Our Approach

- **Review Splitter**
  - Embedding Set Splitter
  - Training/Test Set Splitter

- **Neural Word Embedding Generator**
  - Word2Vec Generator
  - GloVe Generator
  - FastText Generator

- **Review Pre-Processor**
  - Training Pre-Processor
  - Testing Pre-Processor

- **Trainer & Regressor**
  - Deep Neural Network Trainer
  - Deep Neural Network Regressor

Score

04 October 2018
Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

Deep Learning Improvements

Mean Square Error
Deep Learning and Word Embeddings Created from Online Course Reviews for Sentiment Analysis

Word Embeddings Comparison

MSE

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word2Vec Contextual</td>
<td>3200</td>
</tr>
<tr>
<td>Word2Vec Generic</td>
<td>4500</td>
</tr>
<tr>
<td>GloVe Contextual</td>
<td>3800</td>
</tr>
<tr>
<td>GloVe Generic</td>
<td>3600</td>
</tr>
<tr>
<td>FastText Contextual</td>
<td>4000</td>
</tr>
<tr>
<td>FastText Generic</td>
<td>4300</td>
</tr>
</tbody>
</table>

04 October 2018
Thank you for your attention