Learning Word Relatedness over Time

Guy Rosin*, Kira Radinsky* and Eytan Adar⁰

* CS, Technion, Israel
⁰ CSE, University of Michigan, USA

Introduction

Search systems are often focused on providing relevant results for the "now", assuming both corpora and user needs that focus on the present. However, this is not always true, and common search features cannot function well across all times when temporal variance exists.

We tackle the challenge of understanding the temporal intent of the user and retrieving the most relevant historical content.

Examples, courtesy of Google:

Learning Relatedness

Evaluation

We used YAGO2 as a source for temporal relations, and the New York Times archive as a temporal corpus – for creating word embeddings.

Temporal QE Evaluation

<table>
<thead>
<tr>
<th>Learning Relatedness Evaluation</th>
<th>Temporal QE Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>AUC</td>
</tr>
<tr>
<td>Baseline</td>
<td>0.31</td>
</tr>
<tr>
<td>Specific Classifier</td>
<td>0.56</td>
</tr>
<tr>
<td>Temporal Classifier</td>
<td>0.69</td>
</tr>
<tr>
<td>Temporal Model Classifier</td>
<td></td>
</tr>
</tbody>
</table>

Methods

The two charts above represent the Dynamics of their entities, which is defined as:

\[ \text{Dynamics}(e_1, e_2) = \cos(v^{e_1}_1, v^{e_2}_1), \ldots, \cos(v^{e_1}_n, v^{e_2}_n) \]

Where \( e_i \) are entities, and \( v_j^y \) are the vector representations of word \( j \) built upon news from year \( y \).

We introduce the task of learning temporal relatedness: given two entities, identify whether they relate to each other during a certain time period.

We present two classifiers to tackle this task.

In addition, we present the application of this task to query expansion. Our technique allows us to produce temporally appropriate expansions: we identify the temporal focus of the query and find semantically related terms from that time, to expand the query with.

Evaluation

<table>
<thead>
<tr>
<th>Learning Relatedness Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
</tr>
<tr>
<td>Baseline</td>
</tr>
<tr>
<td>Specific Classifier</td>
</tr>
<tr>
<td>Temporal Classifier</td>
</tr>
</tbody>
</table>

Contributions

- Described a novel problem of evaluating word relatedness over time and built datasets to evaluate this task.
- Presented novel representations and algorithms for evaluating this task.
- Presented the application of this task to query expansion along with several methods built on top of the temporal relatedness algorithms.

Future Work

Event-driven word embeddings:

Ariel Sharon

Lebanon War  Gaza Disengagement  Election

Unit 101  Kadima  Park

Gaza Disengagement  Likud  Deaths

Time