EventMiner

Michal Starý
EventMiner

- extract dates from text
- interpret (= normalize) dates to absolute values
- connect dates to relevant events
- deliver results

- Both EN, CS supported
Motivation
Temporal entity extraction
- quick recap
Temporal entity extraction process

Entity recognition

Entity normalization

- absolute
- relative (needs correct reference date)
Kate,

please, can you confirm the meeting at 6 PM tomorrow? It’s about time to prepare our talk at the NLP conference, it’s on the 27th of November, remember? We need to submit the presentation 2 weeks before the conference itself.

PS. I’ve talked to John on Wed. and the vacation won’t be a problem.

See you,

Michal
Existing tools

SUTime

Heideltime

Dateparser

Duckling

(Natty)
SUTime

**Rule based** temporal entity extractor

Great performance on more formal and shorter texts (legal records), the less formal and the longer the worse performance was observed.

Uses rules for matching and interpreting, POS tags for past/future distinction and filtering.

Problems/limitations:

- Not handling shortcuts, typos
- Only one reference date for the whole document
- Only EN!
Kate, please, can you confirm the meeting at 6 PM tomorrow? It’s about time to prepare our talk at the NLP conference, it’s on the 27th of November, remember? We need to submit the presentation 2 weeks before the conference itself. PS. I’ve talked to John on Wed. and the vacation won’t be a problem. See you, Michal
HeidelTime

Similar to SUTime, but keeps evolving

Looks for the reference date inside the texts.

For EN, DE, ES, IT, NE, ...(~10 languages) have high F1 (manual rules)

Extended for all languages using rules automatic translation (not so useful)

- Good baseline
- “Note that recall is much worse than precision for all languages and that the normalization of extracted expressions works quite well (value acc.).”

Kate, please, can you confirm the meeting at 6 PM tomorrow? It’s about time to prepare our talk at the NLP conference, it’s on the 27th of November, remember? We need to submit the presentation 2 weeks before the conference itself. PS. I’ve talked to John on Wed. and the vacation won’t be a problem. See you, Michal
<table>
<thead>
<tr>
<th>Model</th>
<th>F1</th>
<th>P</th>
<th>R</th>
<th>strict F1</th>
<th>value F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heidel\Time-t</td>
<td>90.30</td>
<td>93.08</td>
<td>87.68</td>
<td>81.34</td>
<td>77.61</td>
</tr>
<tr>
<td>Heidel\Time-bf</td>
<td>87.31</td>
<td>90.00</td>
<td>84.78</td>
<td>78.36</td>
<td>72.39</td>
</tr>
<tr>
<td>Heidel\Time-1.2</td>
<td>86.99</td>
<td>89.31</td>
<td>84.78</td>
<td>78.07</td>
<td>72.12</td>
</tr>
<tr>
<td>NavyTime-1.2</td>
<td>90.32</td>
<td>89.36</td>
<td>91.30</td>
<td>79.57</td>
<td>70.97</td>
</tr>
<tr>
<td>ManTIME-4</td>
<td>89.66</td>
<td>95.12</td>
<td>84.78</td>
<td>74.33</td>
<td>68.97</td>
</tr>
<tr>
<td>ManTIME-6</td>
<td>87.55</td>
<td>98.20</td>
<td>78.99</td>
<td>73.09</td>
<td>68.27</td>
</tr>
<tr>
<td>ManTIME-3</td>
<td>87.06</td>
<td>94.87</td>
<td>80.43</td>
<td>69.80</td>
<td>67.45</td>
</tr>
<tr>
<td>SUTime</td>
<td>90.32</td>
<td>89.36</td>
<td>91.30</td>
<td>79.57</td>
<td>67.38</td>
</tr>
<tr>
<td>ManTIME-1</td>
<td>87.20</td>
<td>97.32</td>
<td>78.99</td>
<td>70.40</td>
<td>67.20</td>
</tr>
<tr>
<td>ManTIME-5</td>
<td>87.20</td>
<td>97.32</td>
<td>78.99</td>
<td>69.60</td>
<td>67.20</td>
</tr>
<tr>
<td>ManTIME-2</td>
<td>88.10</td>
<td>97.37</td>
<td>80.43</td>
<td>72.22</td>
<td>66.67</td>
</tr>
<tr>
<td>ATT-2</td>
<td>85.25</td>
<td>98.11</td>
<td>75.36</td>
<td>78.69</td>
<td>65.57</td>
</tr>
<tr>
<td>ATT-1</td>
<td>85.60</td>
<td>99.05</td>
<td>75.36</td>
<td>79.01</td>
<td>65.02</td>
</tr>
<tr>
<td>ClearTK-1,2</td>
<td>90.23</td>
<td>93.75</td>
<td>86.96</td>
<td>82.71</td>
<td>64.66</td>
</tr>
<tr>
<td>JU-CSE</td>
<td>86.38</td>
<td>93.28</td>
<td>80.43</td>
<td>75.49</td>
<td>63.81</td>
</tr>
<tr>
<td>KUL</td>
<td>83.67</td>
<td>92.92</td>
<td>76.09</td>
<td>69.32</td>
<td>62.95</td>
</tr>
<tr>
<td>KUL-TE3RunABC</td>
<td>82.87</td>
<td>92.04</td>
<td>75.36</td>
<td>73.31</td>
<td>62.15</td>
</tr>
<tr>
<td>ClearTK-3,4</td>
<td>87.94</td>
<td>94.96</td>
<td>81.88</td>
<td>77.04</td>
<td>61.48</td>
</tr>
<tr>
<td>ATT-3</td>
<td>80.85</td>
<td>97.94</td>
<td>68.84</td>
<td>72.34</td>
<td>60.43</td>
</tr>
<tr>
<td>FSS-TimEx</td>
<td>85.06</td>
<td>90.24</td>
<td>80.43</td>
<td>49.04</td>
<td>58.24</td>
</tr>
<tr>
<td>TIPSem (TE2)</td>
<td>84.90</td>
<td>97.20</td>
<td>75.36</td>
<td>81.63</td>
<td>65.31</td>
</tr>
</tbody>
</table>

Table 3: Task A - Temporal Expression Performance.

Duckling

Haskell library for generic entities interpretation

Does not support entity recognition

Temporal entity interpretation for 20+ languages
Common problem - Reference date acquisition

Relation extraction task

The **hardest task** in whole temporal entity extraction process

**Solutions:**

- Create own model to get a correct reference date for each relative entity
  - Either heuristics, machine learning or combination of both could be used
- Be satisfied with HeidelTime reference date acquisition
I’m looking forward to Saturday's tournament. It all starts at 8 AM and ends before noon. Make sure to be there!

Saturday 2019-11-23  8 AM 2019-11-26T08:00  noon 2019-11-26T12:00

I’m looking forward to Saturday's tournament. It all starts at 8 AM and ends before noon. Make sure to be there!

Saturday’s 23-11-2019   8 AM 23-11-2019T08:00
Temporal entity extraction
- progress
Work done

1) Explored EN, CS datasets
2) Evaluated different NER tools/models on related datasets
3) Tried several approaches
4) Built Albert-biLSTM-CRF NER model for English
5) Prepared Iterative normalization scheme
Findings
1) Datasets - temporal entities

a) CS
   i) None

b) EN
   i) TBAQ
   ii) TimeBank and Aquaint corpora
   iii) well annotated
   iv) very good dataset
   v) ~1800 datetime entities
1) Datasets - general NE including date/time

a) CS
   i) CNEC
   ii) not so good
   iii) date/time annotation is focused only on more obvious absolute values
   iv) ~3k datetime entities
   v) not suitable for training

b) EN
   i) Ontonotes
   ii) surprisingly good
   iii) very well annotated
   iv) ~15k datetime entities
   v) suitable for training
2) Tools / models

Evaluation of recognition on TBAQ

Manual observation
2) Tools - evaluation on TBAQ

Metrics - soft:

a) preprocess
b) BOW approach
c) remove entity after match
d) add prefix “the”
e) check if reference ent is subword in predicted entity
f) erase errors caused by “now” keyword - easy to match manually
2) Tools - evaluation score

<table>
<thead>
<tr>
<th>Tool</th>
<th>soft_precision</th>
<th>soft_recall</th>
<th>soft_F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sutime</td>
<td>0.805</td>
<td>0.754</td>
<td>0.778</td>
</tr>
<tr>
<td>HeidelTime</td>
<td>0.887</td>
<td>0.741</td>
<td>0.808</td>
</tr>
<tr>
<td>Multilingual BERT NER</td>
<td>0.860</td>
<td>0.752</td>
<td>0.803</td>
</tr>
<tr>
<td>BERT NER</td>
<td>0.725</td>
<td>0.695</td>
<td>0.710</td>
</tr>
</tbody>
</table>
2) Tools - observation

multilingual BERT based model recover also surrounding with temporal function

multilingual BERT based model is surprisingly good for CS

HeidelTime and Sutime are much faster and lighter

Bonus - general NER for CS (simplified boolean matching):

<table>
<thead>
<tr>
<th>Model</th>
<th>soft F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>multiNER (NER currently used at FI)</td>
<td>0.580</td>
</tr>
<tr>
<td>multilingual BERT NER</td>
<td>0.675</td>
</tr>
</tbody>
</table>
Albert based NER
Why Albert based NER?

BERT ~ 5GB RAM

Albert > BERT (performance)

Albert < BERT (memory need)

(Hopefully) soon - multilingual Albert
Properties

Trained on Ontonotes dataset (finalized on TBAQ)

~ 20 entities (including date, time)

Useful for following tasks in the pipeline as well
Outcome
1) HeidelTime + Sutime [+Albert based] (recognition and basic normalization)
2) Iterative normalization
1) Deeppavlov multilingual BERT NER (recognition)
2) Feature engineering on top of duckling (basic normalization)
3) Iterative normalization
Iterative normalization
Issues to solve

Need of domain knowledge - catch and interpret the specific entities with high accuracy; past vs. future focus

- “due XXX”, “within XXX”, ...

Way to properly distinguish the absolute values of several orders

- totally absolute (specific date): DD/MM/YYYY, DD/MM
- partially absolute (specific interval): MM/YYYY, YYY

Model choice
Draft

Finetuned by hand/ ML model

Features:

- distance
- anaphora
- syntactic tree relationship
- POS
- modifying keyword in surrounding
- paragraph?
- topics?
- ... any suggestions?