

Temporal Tagging

Ein Vergleich von standalone Taggern mit ChatGPT.

Elias Wassertheurer & Wolfgang Prett

Temporal Tagging

Was ist das eigentlich?

- Extraktion von zeitbezogenen Informationen aus Text:
 - Identifizieren von Zeitausdrücken (“now”, “two hours”, “Friday”..)
 - Extrahieren von temporalen Informationen (Datum, Uhrzeit, Zeitzone, zeitl. Relationen)
 - Anwendung der Erkenntnisse

TimeML

Standard zur elektronischen Verarbeitung von Dokumenten

- **Timex3 Events** (“every week”, “last Monday”, “on the first of the Month”)
- **Signale** (“when”, “in”, “after”)
- **Temporal Link (TLink, “vor-“, “nach-“, “während-“** Beziehung von Events zueinander)
- **Aspectual Link (ALink, bspw. “Mary completed the marathon”)**
- **Subordinate Link (SLink, markieren kausale/konditionale/subordinative Beziehungen zwischen Events und Zuständen) -> Der Unfall **verursachte** eine Verzögerung**

Timex3 - Events

News - Beispiel

Jan 3 (Reuters) - Xiamen Intretech Inc : * SAYS IT AIMS TO RAISE UP TO 1.7 BILLION YUAN (\$261.90 million) AT 22.5 YUAN PER SHARE FOR ITS INITIAL SHENZHEN SHARE OFFERING

Lösung:

```
<TIMEX3 tid="t1" type="DATE" value="2018-01-03">Jan 3</TIMEX3>
```

(DCT: 2018-01-03)

Temporal Tagging

Anwendungen

- Information Retrieval (“Neuigkeiten von letzter Woche”)
- Natural Language Understanding (Verstehen von Bedeutungen in Text, Temporale Beziehungen von Events zueinander)
- Frage- / Antwort-Systeme
- ChatBots
- BI

Temporal Tagging

Vergleich verschiedener Ansätze

- Regelbasiert (Grammatik Regeln, Tree Parsing, Regelbasierte Normalisierung)
- Machine Learning (Hohe Genauigkeit, Kontextadaptierung, Disambiguierung)
 - Deep - Learning Modelle (State of the Art Performance, Erkennen von long-range dependencies)
- Hybride Lösungen (Kombination von Regelbasiert und ML)

Ziel des Projekts

- Vergleich von Regelbasierten und ML Systemen
 - UWTime: State of the Art regelbasierter Temporal Tagger (Almasian, S., et al. (2021). BERT got a Date: Introducing Transformers to Temporal Tagging)
 - GPT 4.o: State of the Art LLM

Versuchsaufbau

Datensatz - Professor Heideltime

- (Huggingface: [hugosousa/ProfessorHeidelTime](#))
- Entwickelt an der Universität Heidelberg
- Trainingsdatensatz für den Temporal Tagger Heideltime
(Sousa, H., Campos, R., & Jorge, A. (2023). TEI2GO: A Multilingual Approach for Fast Temporal Expression Identification)
- Mehr als 100 000 Datensätze (33.000 deutsch, ca 25.000 englisch)

Versuchsaufbau

UWTime - University of Washington Time

- Lokale Installation von UWTime (standalone):
 - Bereitstellung in einer Java Anwendung
 - HTTP Server auf Port 10001
 - Python Client der HTTP Anfragen an diesen Port schickt
 - Antwort als TimeML getaggtter Text (XML Tags in String)
 - Verarbeitung der Antwort mittels XML Element Tree (XML Parser, extrahiert XML Elemente, macht XML - Properties zugänglich und bietet Filter-Funktion)
 - Gegenüberstellung der Ergebnisse in eigener ComparisonInstance - Klasse (vergleicht zwei Collections von Timex3 - Datensätzen)

Versuchsaufbau

ChatGPT 4o mini - OpenAI

- Ansteuerung der ChatGPT API:
 - Schnittstelle der OpenAI/ChatGPT mittels Python Library und API Key
 - Python Modul koordiniert Anfragen an die API und responses
 - Vordefinierter Prompt mit Platzhaltern als Anfrage
 - Antwort als TimeML getaggtter Text (XML Tags in String)
 - Verarbeitung der Antwort mittels XML Element Tree (XML Parser, extrahiert XML Elemente, macht XML - Properties zugänglich und bietet Filter-Funktion)
 - Gegenüberstellung der Ergebnisse in eigener ComparisonInstance - Klasse (vergleicht zwei Collections von Timex3 - Datensätzen)

Versuchsaufbau

ChatGPT 4o mini - OpenAI

- Chat-GPT Prompt:

You are a temporal tagger. Given the following text , find temporal expressions. As answer provide only a list of timex3 xml tags:

{{message}}

Versuchsaufbau


ChatGPT 4o mini - OpenAI

You are a temporal tagger. Given the following text , find temporal expressions. As answer provide only a list of timex3 xml tags:

Sony's "super mid-range" Xperia XA1 phone is now available to preorder through Amazon and Best Buy. It'll go on sale at other retailers, like B&H and Fry's, on May 1st. The phone will cost \$299.99. The company announced the device during Mobile World Congress this year. Here are the specs again: The XA1 and the larger XA1 Ultra also feature something called "Xperia Actions," which Sony says can learn users' habits to automatically change settings and manage apps. For instance, the phone's software could learn your bedtime and automatically adjust screen brightness and volume.



xml

 Copy

```
<TIMEX3 tid="t1" type="DATE" value="PRESENT_REF">now</TIMEX3>  
<TIMEX3 tid="t2" type="DATE" value="2017-05-01">May 1st</TIMEX3>  
<TIMEX3 tid="t3" type="DATE" value="2017">this year</TIMEX3>
```


Versuchsaufbau

Definition Resultate

- Positive (P): Mengen von Timex-Ausdrücken sind gleich (= Gold-Label)
- False Positive (FP): Predictions sind Superset von Goldlabel
- False Negative (FN): Goldlabel ist Superset von Predictions
- N: sonst

Vergleich von Temporal Taggern

UWTime vs GPT 4.o

```
=====
```

```
UWTime Stats
```

```
=====
```

```
Calculating statistics for instance set
```

```
Amount of instances tested: #275
```

```
Precision: 0.6973684210526315
```

```
F-Score: 0.6198830409356725
```

```
Recall: 0.5578947368421052
```

```
Accuracy: 0.7636363636363637
```

```
=====
```

```
ChatGPT stats:
```

```
=====
```

```
Calculating statistics for instance set
```

```
Amount of instances tested: #1000
```

```
Precision: 0.9958634953464323
```

```
F-Score: 0.9963786859803415
```

```
Recall: 0.9968944099378882
```

```
Accuracy: 0.993
```