

# L3S at the NTCIR-12 Temporalia

## TDR and TID

### Query Subtopic Classification and Initial Retrieval

- ▶ Joint classification of subtopics to appropriate temporal intent
- ▶ Retrieval of pseudo relevant documents using a unigram language model
- ▶ List-wise L2R setup with NDCG@20 as the target measure
- ▶ Features: Expected Temporal Distance, Temporal Density, Verb Tense Features

### Temporal Diversification Approach

- ▶ Use earth mover's distance to measure distance between temporal distributions of two documents
- ▶ We use the top 100 documents retrieved for each temporal intent as the candidates.
- ▶ The diversified list is created by greedily maximizing the earth-movers distance

Run	NDCG@20					P@20				
	Atemporal	Future	Past	Recency	All	Atemporal	Future	Past	Recency	All
Manual L2R	0.7264	0.6511	0.7005	<b>0.7151</b>	0.6983	<b>0.7960</b>	0.7360	0.7710	<b>0.7970</b>	<b>0.7750</b>
Auto Param Sum	0.6109	0.6932	0.7127	0.6758	0.6731	0.7330	<b>0.7790</b>	<b>0.8000</b>	0.7760	0.7720
Auto L2R	<b>0.7299</b>	0.6508	0.6998	0.7116	0.6980	0.7960	0.7360	0.7700	0.7930	0.7737
LM	0.7052	<b>0.7151</b>	<b>0.7297</b>	0.6865	<b>0.7076</b>	0.7690	0.7850	0.7940	0.7580	0.7416

**Table 1: Per-Class results for all TDR Runs. For every temporal class, the highest value is indicated in bold.**

Run	Average Absolute Loss	Cosine Similarity
L3S-TID-E-1	<b>0.2031</b>	<b>0.7307</b>
L3S-TID-E-2	0.2452	0.6673

**Table 3 : Evaluation Results of TID Formal Runs**

Run	D#-NDCG@20	I-rec@20
Manual L2R	0.8262	0.9850
Auto Param Sum	0.6852	<b>0.9900</b>
Auto L2R	<b>0.8423</b>	0.9850

**Table 2: Diversified Results of TDR Formal Runs**

### Temporal Intent Disambiguation

- ▶ Rule based voting method
- ▶ Each rule is made based on a feature. If the rule is obeyed by a temporal class then it is awarded one vote
- ▶ Votes are normalized across the classes to get a probability distribution
- ▶ Features used:
  - ▶ Temporal Distance
  - ▶ Linguistic Features like verb tense and modality
  - ▶ N-grams

### Rules for Voting

- ▶ Decision tree trained using only n-grams to predict the temporal class of a query. Class with the highest confidence gets a vote
- ▶ Verb tense of the query counts as a single vote. If no verb tense is detected then the atemporal class is awarded a vote.
- ▶ Class of the time mention in the query gets a vote.
- ▶ If the standard deviation is low then temporal distance is used as the deciding vote.