Annif at SemEval-2025 Task 5: Traditional XMTC augmented by LLMs

Osma Suominen, Juho Inkinen and Mona Lehtinen / National Library of Finland, University of Helsinki

LLMs4Subjects task overview

Given:
- GND subject vocabulary (German monolingual terms)
- bibliographic records (titles and abstracts in German and/or English + GND subject labels) from the TIBKAT database

Predict GND subjects for *unseen* titles and abstracts in the test set - an Extreme Multilabel Text Classification (XMTC) task!

Variations:
- **all**: full GND (200k subjects), large train set (82k records)
- **tib-core**: GND subset (79k subjects), small train set (42k records)

Our approach

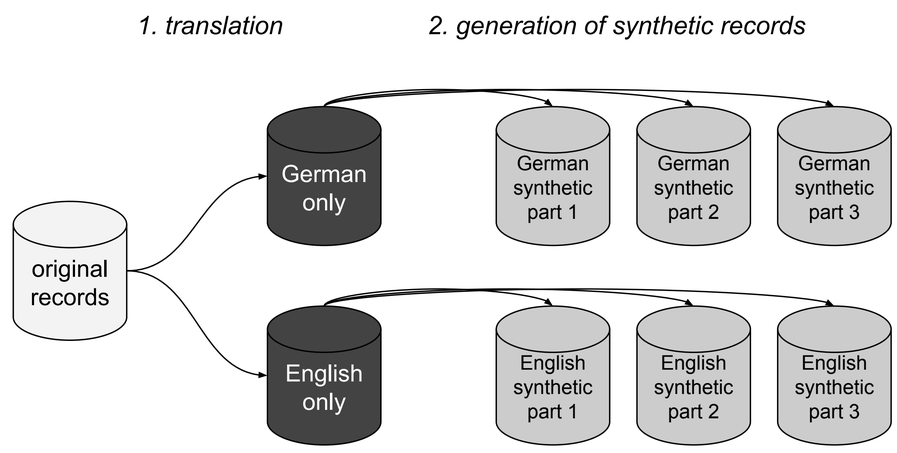
1. **Use the Annif automated subject indexing toolkit (by us, open source)**
2. **Pre-process and enhance data set using LLMs**
   - translate **GND terms** to English with **GPT-4o-mini**
   - translate **TIBKAT records** to monolingual German/English with **Llama-3.1-8B-Instruct**
   - generate additional **synthetic training data** with **Llama-3.1-8B-Instruct**
3. **Train ensemble of XMTC algorithms**
   - Omikuji Bonsai (partitioned label trees)
   - MLLM (lexical approach)
   - XTransformer (BERT based XMTC and ranking)
4. **Use monolingual (de/en) prediction pipelines in parallel**
   - merge their results for final system output

Quantitative results

Vocab	System	Run#	Ensemble	Lang	development set F1@5	development set nDCG@10	test set F1@5	test set Avg recall	Rank
all	Annif (ours)	1	BM simple	de	0.3174	0.5459	0.3108	0.5736	
		2		en	0.3312	0.5677	0.3184	0.5890	
		3		de+en	-	-	0.3376	0.6201	
		4	BM neural	de	0.3337*	0.5726*	0.3029	0.5447	
		5		en	**0.3504***	**0.6008***	0.3116	0.5599	
		6		de+en	-	-	0.3318	0.6005	
		7	BMX simple	de	0.3263	0.5614	0.3185	0.5859	
		8		en	0.3411	0.5842	0.3276	0.6038	
		9		de+en	-	-	**0.3432**	**0.6295**	1st
	DUTIR831	3					0.3346	0.6045	2nd
	RUC Team	1					0.3015	0.5856	3rd
	DNB-AI-Project	1					0.3231	0.5631	4th
	icip	1					0.2618	0.5302	5th
tib-core	Annif (ours)	1	BM simple	de	0.2821	0.5557	0.2796	0.5285	
		2		en	0.3009	0.5936	0.2984	0.5617	
		3		de+en	-	-	0.3113	0.5824	
		4	BM neural	de	0.3209*	0.6171*	0.2660	0.4864	
		5		en	**0.3467***	**0.6661***	0.2886	0.5217	
		6		de+en	-	-	0.3043	0.5559	
		7	BMX simple	de	0.2891	0.5684	0.2864	0.5385	
		8		en	0.3079	0.6051	0.3030	0.5719	
		9		de+en	-	-	0.3136	0.5899	2nd
	RUC Team	1					**0.3271**	**0.6568**	1st
	LA^{2}I^{2}F	2					0.2717	0.5794	3rd
	DUTIR831	2					0.3153	0.5599	4th
	icip	1					0.2370	0.4976	5th

Table 1: Quantitative evaluation results for the ensemble projects measured against the development and test sets. Top 5 systems included for comparison. Note that Lang refers to the project language, not to the indicated language of the records. *Unreliable score because the neural ensemble was trained on the development set it was evaluated on.

Pre-processing steps



Qualitative results

Case	System	Avg Recall	Rank
1	DNB-AI-Project	**0.5657**	1st
	DUTIR831	0.5330	2nd
	RUC Team	0.5199	3rd
	Annif (ours)	0.5024	4th
	jim	0.4928	5th
2	DNB-AI-Project	**0.5094**	1st
	DUTIR831	0.4851	2nd
	RUC Team	0.4645	3rd
	Annif (ours)	0.4484	4th
	jim	0.4258	5th

Table 2: Qualitative evaluation results for the top 5 teams in evaluation cases 1 and 2.

Pre-processing example

Original record

Record ID: 3A1002732751
all-subjects/train/Thesis/en

Wasserbauliche Systemanalyse zur Bewertung der Auswirkungen des Klimawandels für tidebeeinflusste Nebengewässer der Elbe

The present work focusses on three major hydraulic engineering issues at the tidal tributaries of the Lower Elbe (rivers Este and Krückau). These issues relate to i) flood protection, ii) morphodynamics and morphological development, as well as iii) the impact of climate change on runoff and morphodynamics in the tidal-influenced lower reaches.

Diese Arbeit behandelt drei wesentliche wasserbauliche Fragestellungen an den beiden Nebenflüssen der Unterelbe, Este und Krückau. Diese beziehen sich auf i) den Hochwasserschutz, ii) die Morphodynamik und die morphologische Entwicklung sowie iii) die Auswirkungen des Klimawandels auf das Abflussgeschehen sowie die Morphodynamik in den tidebeeinflussten Unterläufen.

Subjects: Auswirkung, Dynamische Geomorphologie, Flussbau, Klimaänderung, Nebenfluss, Hochwasserschutz

Translated/monolingual records

Analyse hydraulischer Systeme zur Bewertung der Auswirkungen des Klimawandels für tidebeeinflusste Nebengewässer der Elbe

Die vorliegende Arbeit konzentriert sich auf drei Hauptprobleme der Wasserbaukunst an den tidebeeinflussten Nebenflüssen der Niederelbe (Flüsse Este und Krückau). Diese Probleme beziehen sich auf i) den Hochwasserschutz, ii) die Morphodynamik und die morphologische Entwicklung sowie iii) die Auswirkungen des Klimawandels auf das Abflussgeschehen und die Morphodynamik in den tidebeeinflussten unteren Läufen.

Hydraulic Engineering System Analysis for Evaluating the Impacts of Climate Change on Tidal-Influenced Tributaries of the Elbe

The present work focuses on three major hydraulic engineering issues at the tidal tributaries of the Lower Elbe (rivers Este and Krückau). These issues relate to i) flood protection, ii) morphodynamics and morphological development, as well as iii) the impact of climate change on runoff and morphodynamics in the tidal-influenced lower reaches.

Synthetic records

Optimierung stromlinienförmiger Uferstrukturen zur Reduzierung der Auswirkungen von Sturmfluten im Kontext von Klimawandel und dynamischer Geomorphologie an tiefebenen Flussläufen mit Nebenflüssen

Diese Studie untersucht die Effektivität stromlinienförmiger Uferkonzepte bei der Minderung der Auswirkungen von Sturmfluten an tiefebenen Flussläufen mit Nebenflüssen unter Berücksichtigung der Auswirkungen des Klimawandels auf die Dynamische Geomorphologie und den Flussbau. Die Ergebnisse liefern wertvolle Erkenntnisse für die Planung und Implementierung effektiver Hochwasserschutzmässigkeiten in diesen sensiblen Gebieten.

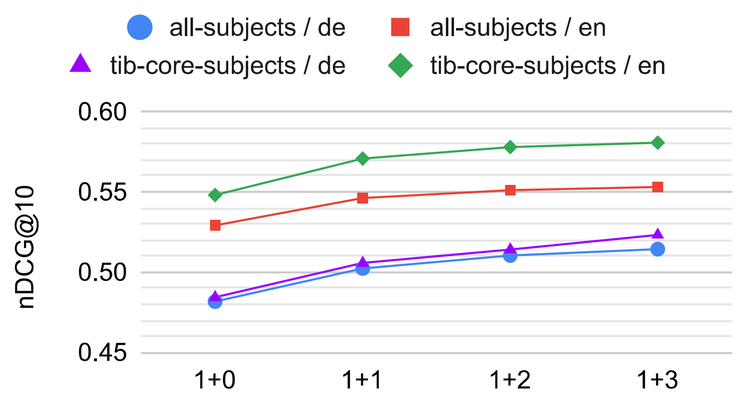
Subjects: + Stromlinie <Strömungsmechanik>

Assessing the Effects of Climate-Driven Phosphorus Complexes on Geomorphic Processes in Coastal Tributaries: A Study on Morphodynamic Adaptation and Flood Risk Mitigation Strategies

This study investigates the interplay between climate-driven phosphorus complex formation, dynamic geomorphic processes, and flow construction in coastal tributaries. The research aims to understand how these factors contribute to increased flood risk and evaluates the effectiveness of adaptive management strategies for mitigating such impacts under changing environmental conditions.

Subjects: + Phosphorkomplexe

Synthetic training data



Paper on arXiv

arxiv.org/abs/2504.19675