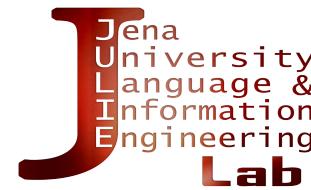




FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA



JULIE Lab & Med Uni Graz @ TREC 2019

Precision Medicine Track

Erik Faessler¹ and Michel Oleynik²

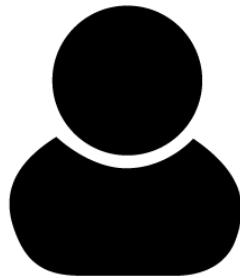
¹Jena University Language &
Information Engineering (JULIE) Lab
Friedrich Schiller University Jena,
Jena, Germany

<http://www.julielab.de>

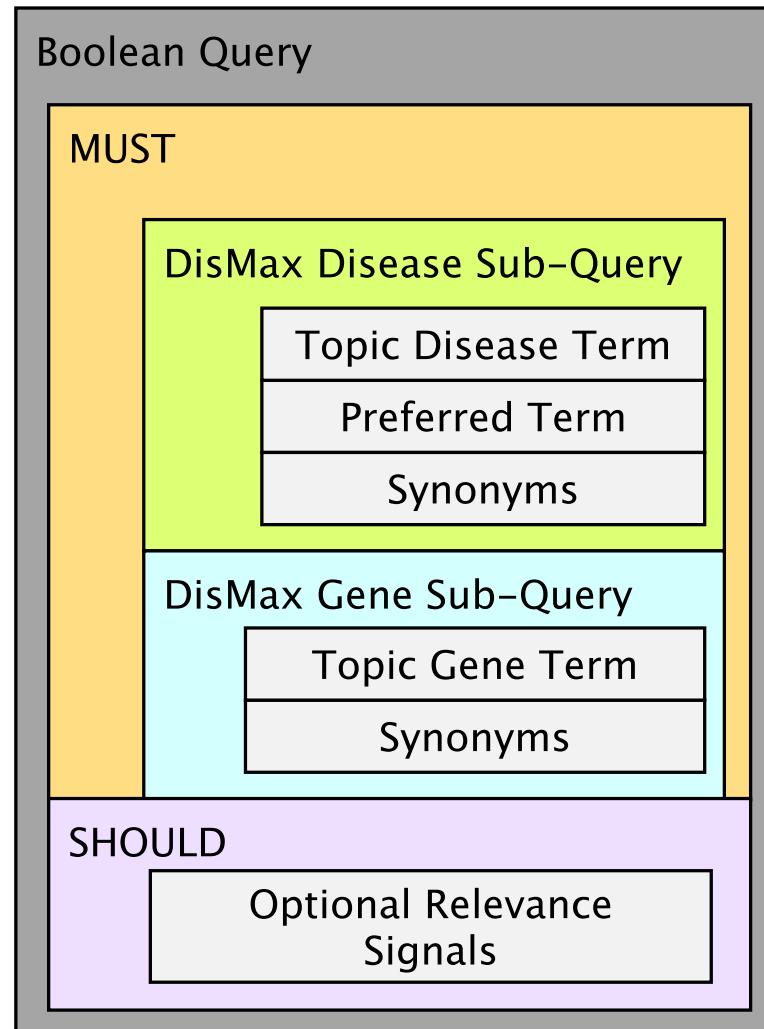
²Institute for Medical Informatics,
Statistics and Documentation,
Medical University of Graz
Graz, Austria

<https://www.medunigraz.at/imis>

Topics



Query Structure



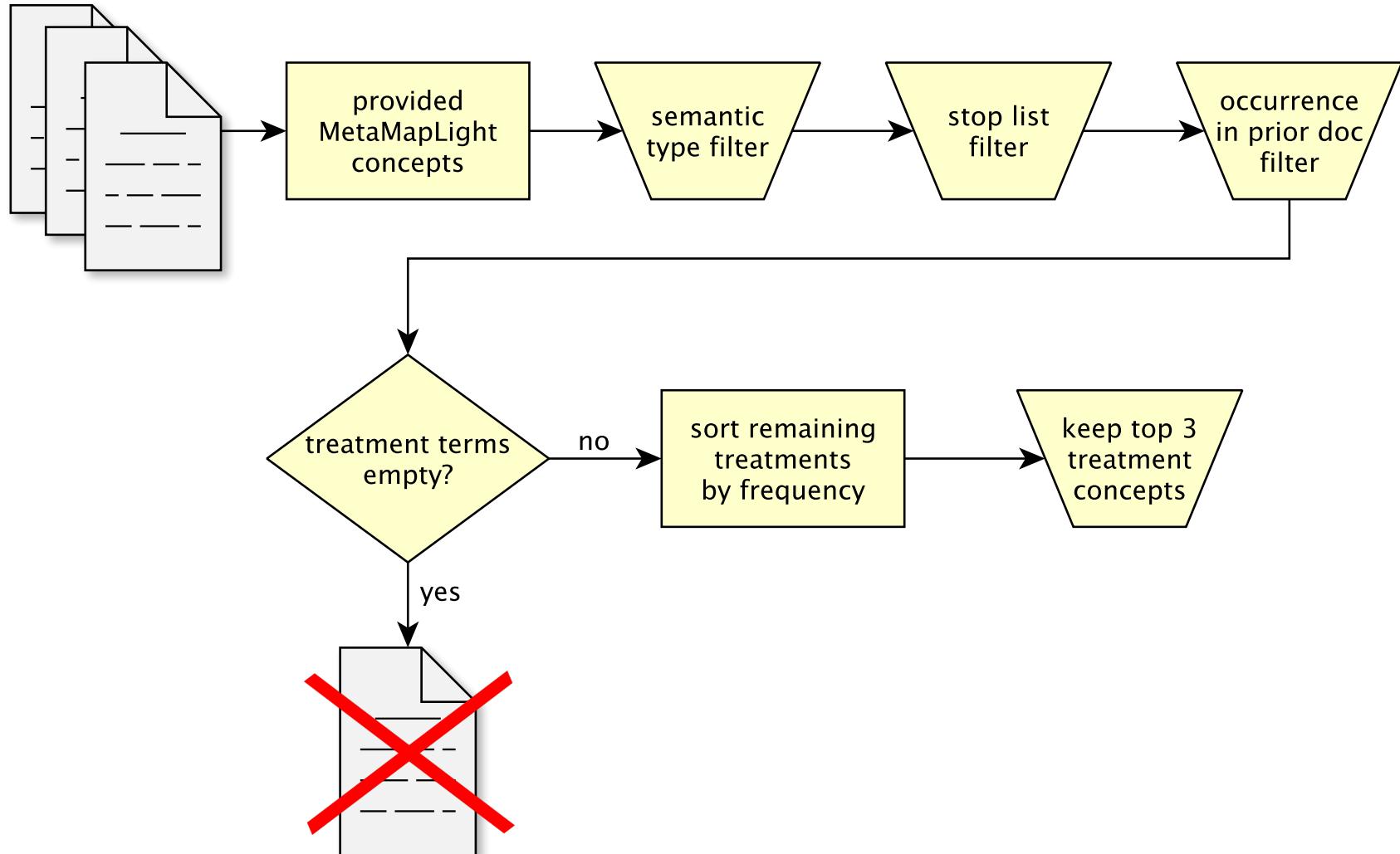
New this Year (group name last year: hpi-hc)

- Biomedical Abstracts: New index field for treatments mentioned in the document, boosting documents with treatments (see next slide)
- Clinical Trials: Added a field containing NLP-extracted gene mentions (already done for BA last and this year)

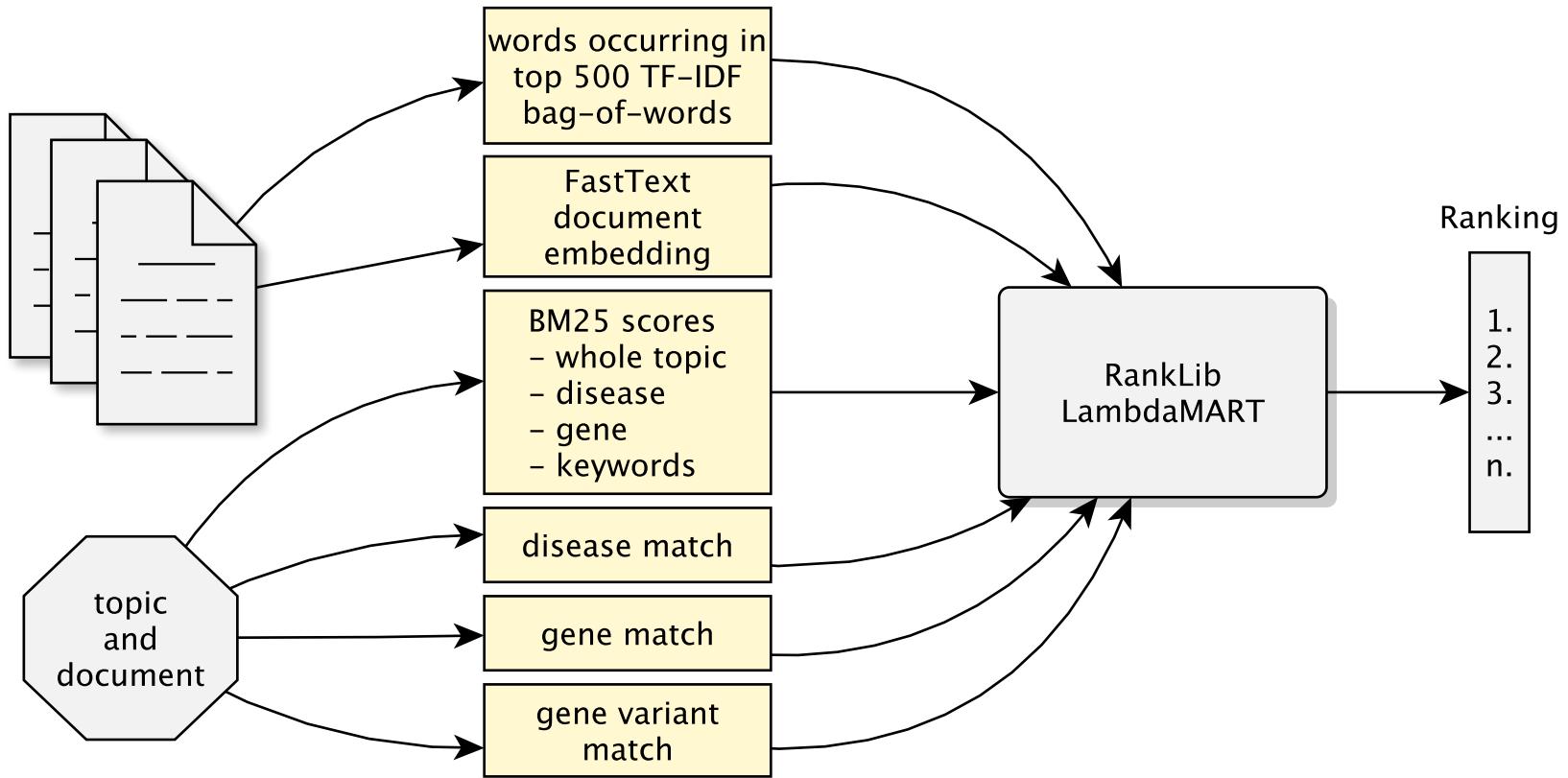
Treatment Task

- Restriction of the MetaMapLite results by semantic type
 - Procedures: T061, Therapeutic or Preventive Procedure
 - Chemicals & Drugs: T121, Pharmacologic Substance
 - Chemicals & Drugs: T200, Clinical Drug
- Further concept filtering by 230 item stop list for results
 - duration, tandem, others, balance, color...

Treatment Concepts Assignment



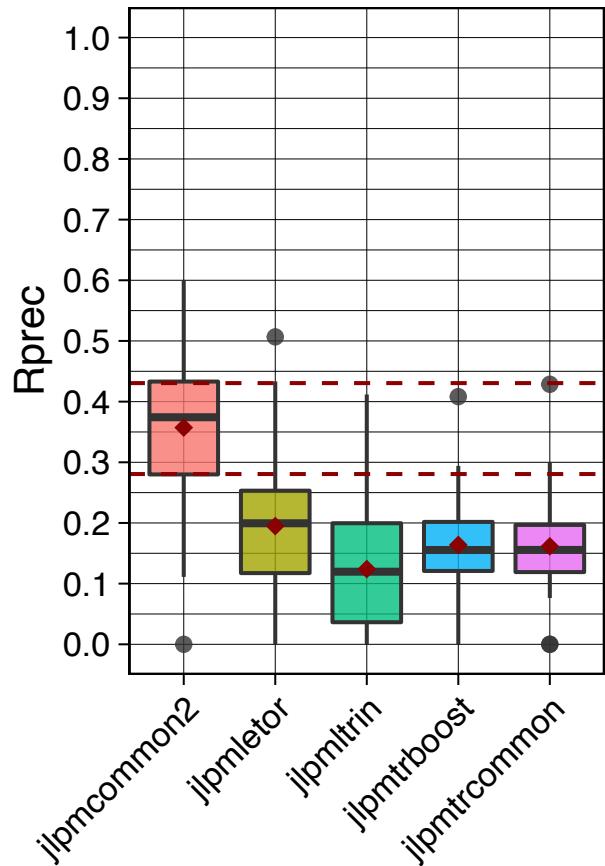
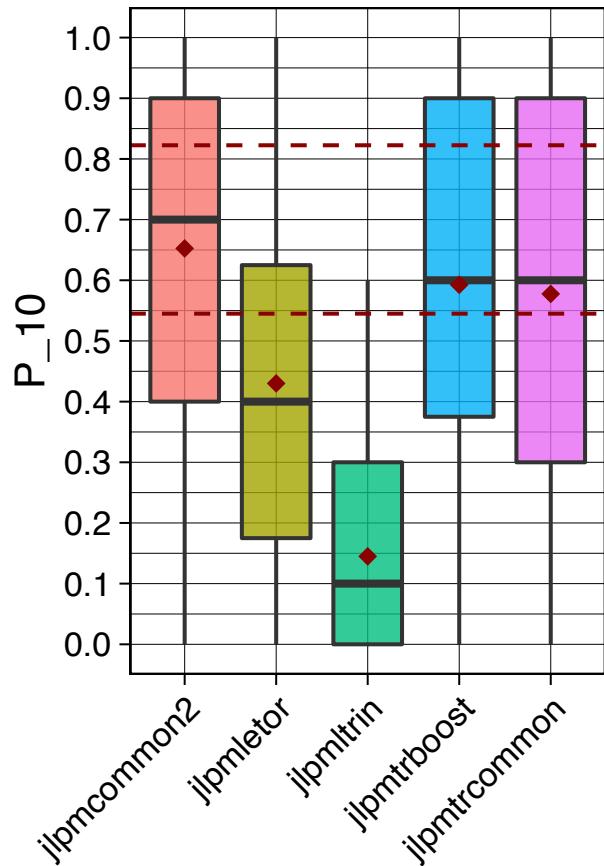
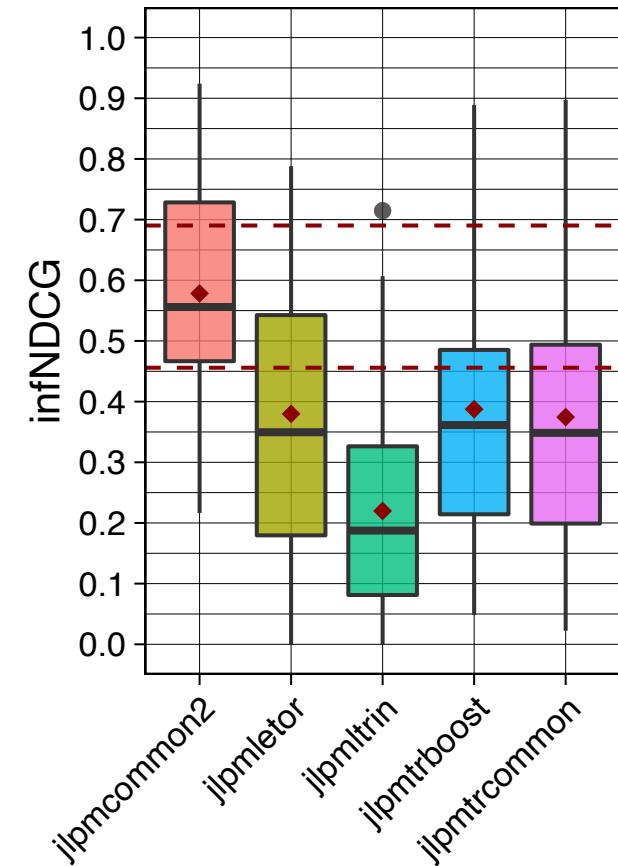
Learning to Rank



Training data

- 2017/2018 gold standard
- 2019 internal annotations (405 abstracts, 403 trials)
- For abstracts and clinical trials

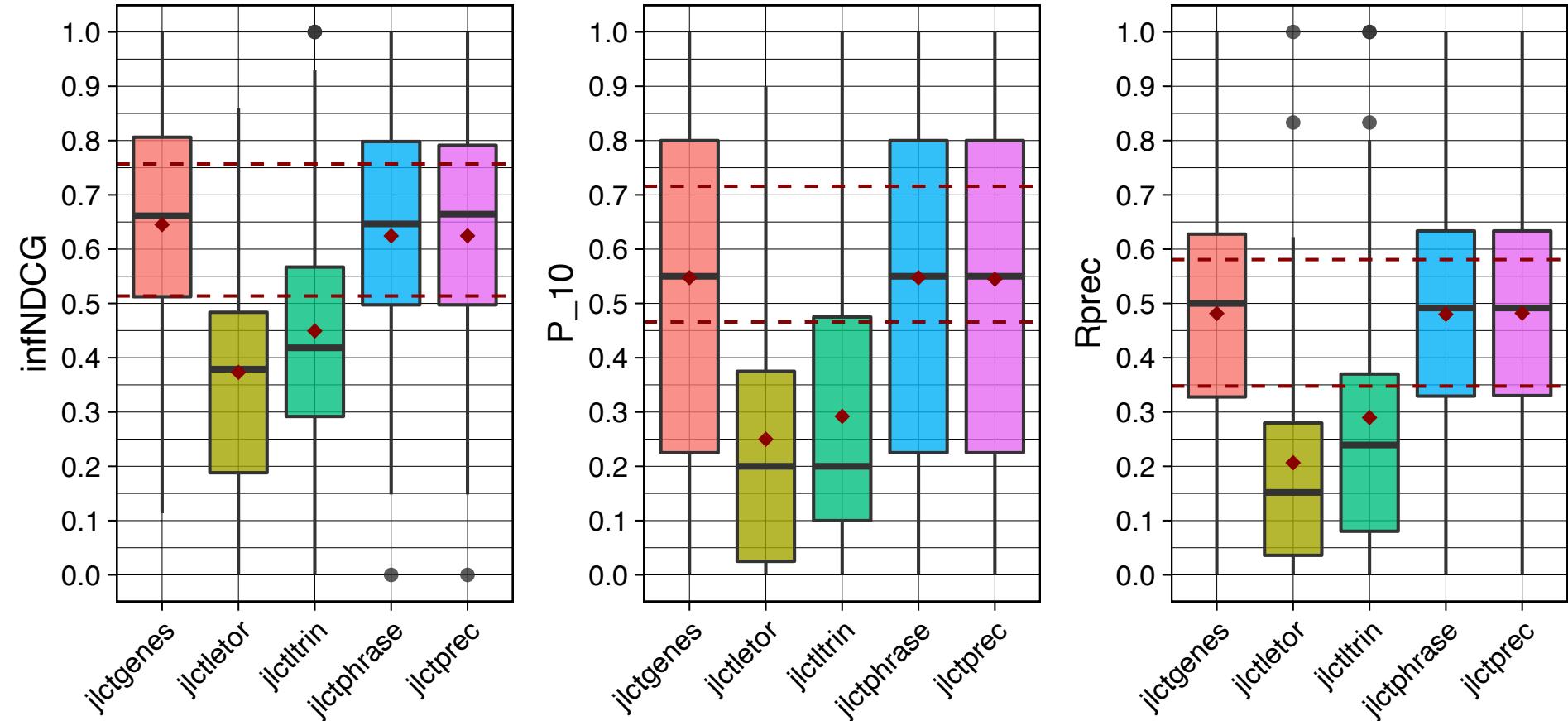
Results: Biomedical Abstracts



Results: Treatments

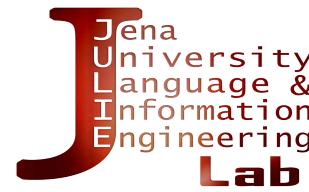
Strategy	trboost	trcommon
Baseline strategies	Y	Y
Valid treatment exists	Y	N
<hr/>		
# Treatments	11,128	9,443
Recall@10	0.2857	0.2698
F ₁ @10	0.3118	0.3019
Recall@25	0.4603	0.4469
F ₁ @25	0.3793	0.3716

Results: Clinical Trials





FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA



Thank you!



¹Jena University Language &
Information Engineering (JULIE) Lab
Friedrich Schiller University Jena,
Jena, Germany

<http://www.julielab.de>

²Institute for Medical Informatics,
Statistics and Documentation,
Medical University of Graz
Graz, Austria

<https://www.medunigraz.at/imis>

Results: Biomedical Abstracts

Strategy	common2	letor	jlpm ltrin	trboost	trcommon
Baseline strategies	Y	Y	Y	Y	Y
Valid treatment exists	Y	N	N	Y	N
LETOR training data	-	2017/18	2019	-	-
# Treatments				11,128	9,443
Recall@10				0.2857	0.2698
F ₁ @10				0.3118	0.3019
Recall@25				0.4603	0.4469
F ₁ @25				0.3793	0.3716
infNDCG	0.5783	0.3798	0.2198	0.3876	0.3745
P@10	0.6525	0.4300	0.1450	0.5925	0.5775
R-Prec	0.3572	0.1948	0.1241	0.1639	0.1615

Results: Clinical Trials

Strategy	genes	letor	jlct ltrin	phrase	prec
Baseline strategies	Y	Y	Y	Y	Y
Match all	Y	Y	Y	Y	N
Match extracted genes	Y	N	N	N	N
LETOR training data	-	2017/18	2019	-	-
infNDCG	0.6451	0.3731	0.4492	0.6244	0.6245
P@10	0.5474	0.2500	0.2921	0.5474	0.5447
R-Prec	0.4814	0.2066	0.2898	0.4799	0.4820