

MODELING DUTCH MEDICAL TEXTS FOR DETECTING FUNCTIONAL CATEGORIES AND LEVELS OF COVID-19 PATIENTS

Jenia Kim, Stella Verkijk, Edwin Geleijn, Marike van der Leeden, Carel Meskers, Caroline Meskers, Sabina van der Veen, Piek Vossen, Guy Widdershoven

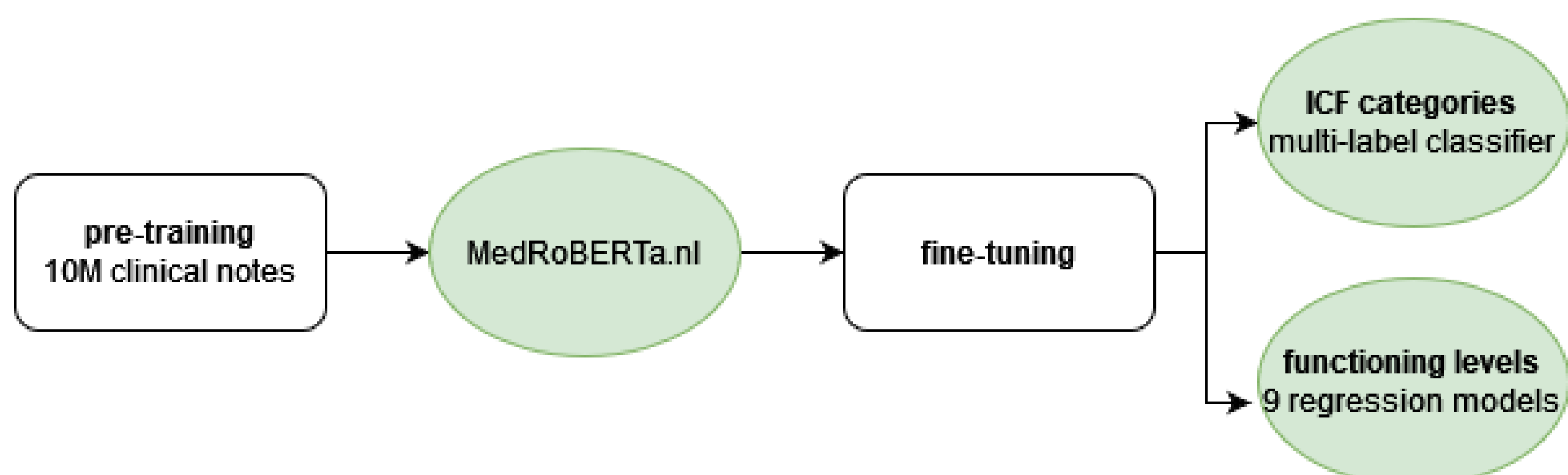
Abstract

Electronic Health Records (EHRs) contain a lot of information in natural language that is not expressed in structured clinical data. Especially in the case of new diseases such as COVID-19, this information is crucial to get a better understanding of patient recovery patterns and factors that may play a role in it. However, the language in these records is very different from standard language and generic natural language processing tools cannot easily be applied out-of-the-box. In this paper, we present a fine-tuned Dutch language model specifically developed for the language in these health records that can determine the functional level of patients according to a standard coding framework from the World Health Organization: the International Classification of Functioning, Disability and Health (ICF). We provide evidence that our classification performs at a sufficient level (F1-score above 80% for the main categories and error rates of less than 1 level on a 5-point Likert scale for levels) to generate patient recovery patterns. These patterns can be used to analyse factors that contribute to the rehabilitation of COVID-19 patients and to predict individual patient recovery of functioning.

Method

The language in EHRs deviates in many aspects from standard Dutch. Therefore, we relied on the transformer Language Model (LM) 'MedRoberta.nl' that was built from scratch on millions of Dutch hospital notes in EHRs (Verkijk and Vossen, 2022).

The MedRoberta.nl model was then fine-tuned with ICF labels to capture 9 different functional categories and levels of functioning in each category.



ICF categories and levels

ICF code	Category	Abbrev.	Functioning levels scale
b1300	Energy level	ENR	0-4
b140	Attention functions	ATT	0-4
b152	Emotional functions	STM	0-4
b440	Respiration functions	ADM	0-4
b455	Exercise tolerance functions	INS	0-5
b530	Weight maintenance functions	MBW	0-4
d450	Walking	FAC	0-5
d550	Eating	ETN	0-4
d840-d859	Work and employment	BER	0-4

Annotation

The annotation consists of assigning a label to a phrase that describes one of the 9 ICF categories and another label to a phrase that describes the level of functioning. For example, in the sentence *Concentration is still slightly diminished*, the word *concentration* is marked with the category label ATT (Attention) and the phrase *slightly diminished* is marked with the level label att-3, which indicates a mild functioning problem. **Inter-annotator agreement:**

category	F1-score	ADM	ATT	BER	ENR	ETN	FAC	INS	MBW	STM
level	MAE	.25	.32	.38	.39	.28	.17	.30	.32	.31

Datasets

	Total number sentences	Sentences with ICF categories
train	239,153	17,731
dev	21,742	1,281
test	22,082	2,256
total	282,977	21,268

Example recovery trajectory

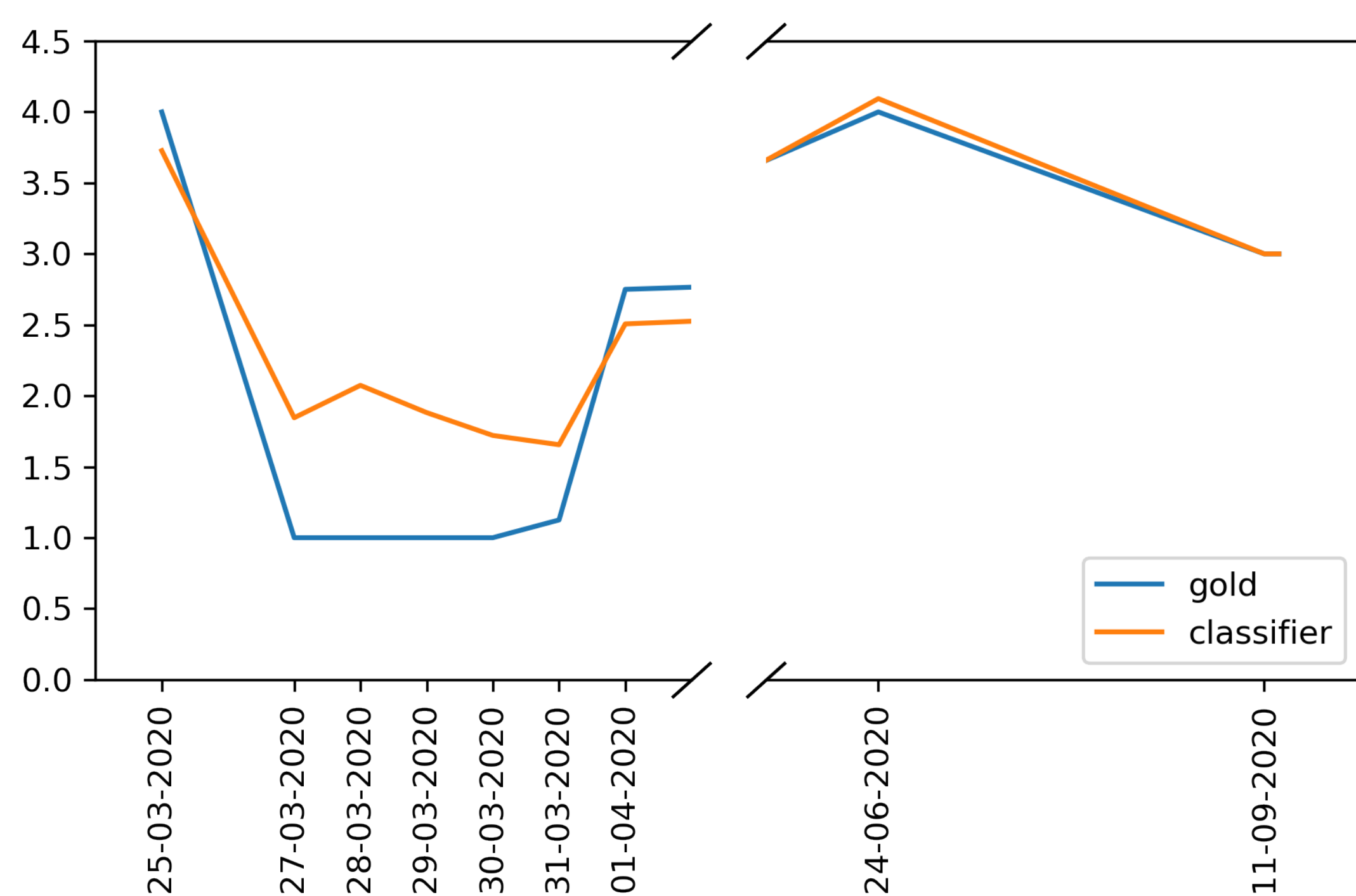


Fig. 5: ADM levels of a COVID-19 patient over time

Results: categories

		P	R	F1	support
sents	ADM	.98	.49	.66	775
	ATT	.98	.41	.58	39
	BER	.56	.29	.35	54
	ENR	.96	.57	.72	160
	ETN	.92	.49	.63	382
	FAC	.84	.71	.76	253
	INS	.89	.26	.41	287
	MBW	.79	.62	.70	125
	STM	.70	.75	.72	181
notes	ADM	1.0	.89	.94	231
	ATT	1.0	.56	.71	27
	BER	.66	.44	.50	34
	ENR	.96	.70	.81	92
	ETN	.95	.72	.82	165
	FAC	.84	.89	.86	95
	INS	.95	.46	.61	116
	MBW	.87	.87	.87	64
	STM	.80	.87	.84	94

Results: levels

		MAE	MSE	RMSE	support
sents	ADM	.48	.55	.74	421
	ATT	.99	1.35	1.16	32
	BER	1.56	3.06	1.75	26
	ENR	.48	.49	.70	100
	ETN	.59	.65	.81	183
	FAC*	.70	.91	.95	139
	INS*	.69	.80	.89	136
	MBW	.81	.83	.91	60
	STM	.76	1.03	1.01	155
notes	ADM	.37	.34	.58	200
	ATT	1.03	1.47	1.21	21
	BER	1.49	2.85	1.69	22
	ENR	.43	.42	.65	70
	ETN	.50	.47	.68	123
	FAC*	.66	.93	.96	79
	INS*	.61	.64	.80	74
	MBW	.60	.56	.75	41
	STM	.68	.87	.93	84

Conclusion

- We described a fine-tuned Dutch language model for the medical domain that assigns functional level classifications to Electronic Health Records of COVID-19 patients.
- We showed that our classifier has sufficient performance to generate potentially reliable patient recovery patterns that can be used to search for factors that impact recovery.
- Our approach can be applied to other languages, which will help the understanding of patient functioning in an international context.
- Our models are freely available on <https://huggingface.co/clt1> and the code is available on <https://github.com/clt1/a-proof-zonmw>.

References

Verkijk, S. and Vossen, P. (2022). MedRoBERTa.nl: A language model for Dutch electronic health records. *Computational Linguistics in the Netherlands Journal*, 11.