

# Application for detection of Fake News

Bc. Jan Zádrapa, 2024

(FakeDet)

## MANIPULATION TECHNIQUES



FIG 1: logo

- fear,
- labelling,
- propaganda,
- conspiracy,
- others.

## NLP PIPELINE AND MODELS



FIG 2: NLP pipeline

- implementing NLP pipeline,
- one model = one technique,
- models BERT and RoBERTa.

## DATASET

- serious and disinformation websites,
- 1000+ labelled paragraphs,
- novinky.cz, ceskenoviny.cz,
- protiproud.info, kresadlo.com, aeronet.news.

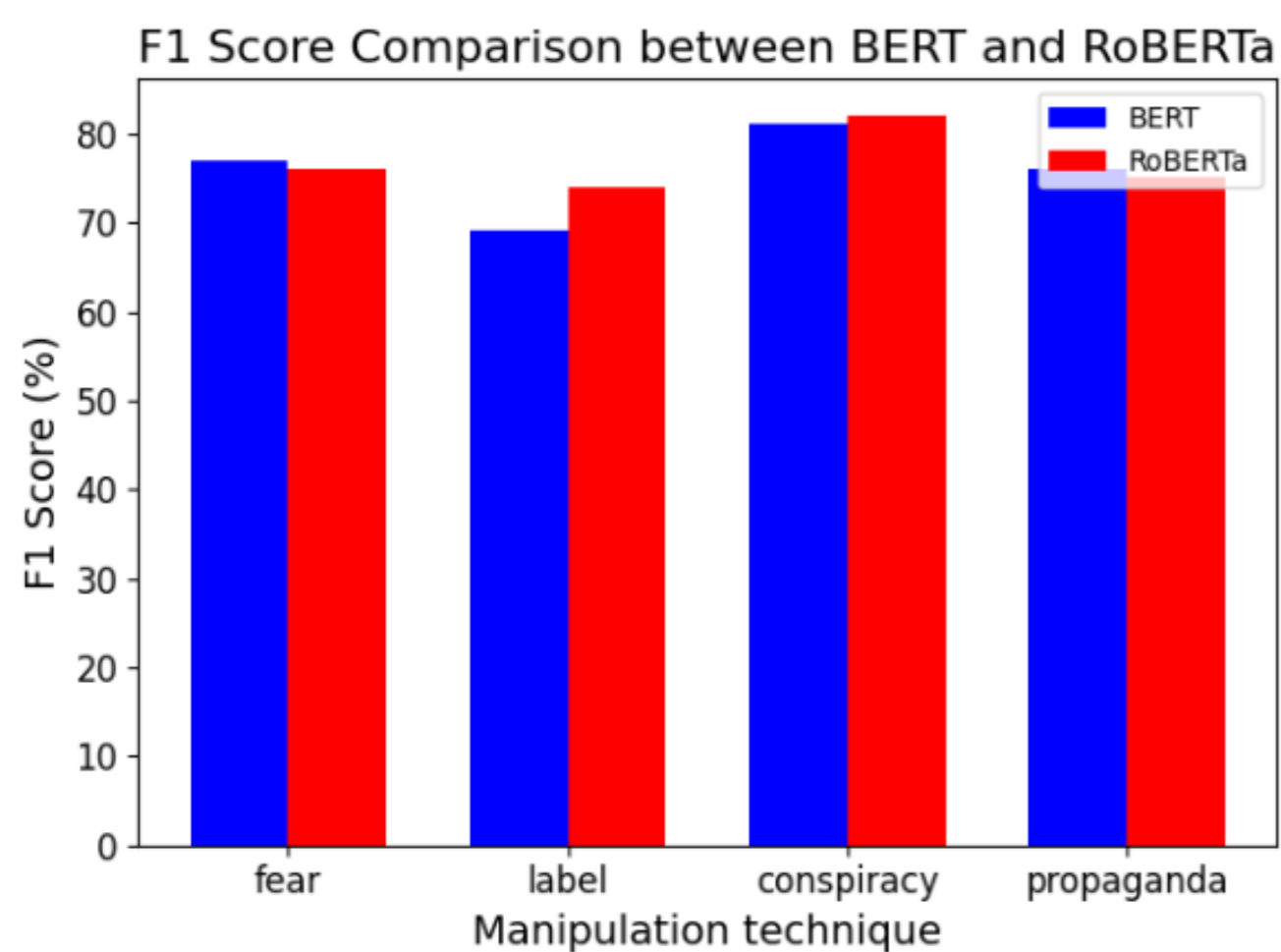


FIG 3: BERT vs. RoBERTa

## APPLICATION AND RESULTS

FIG 4: design

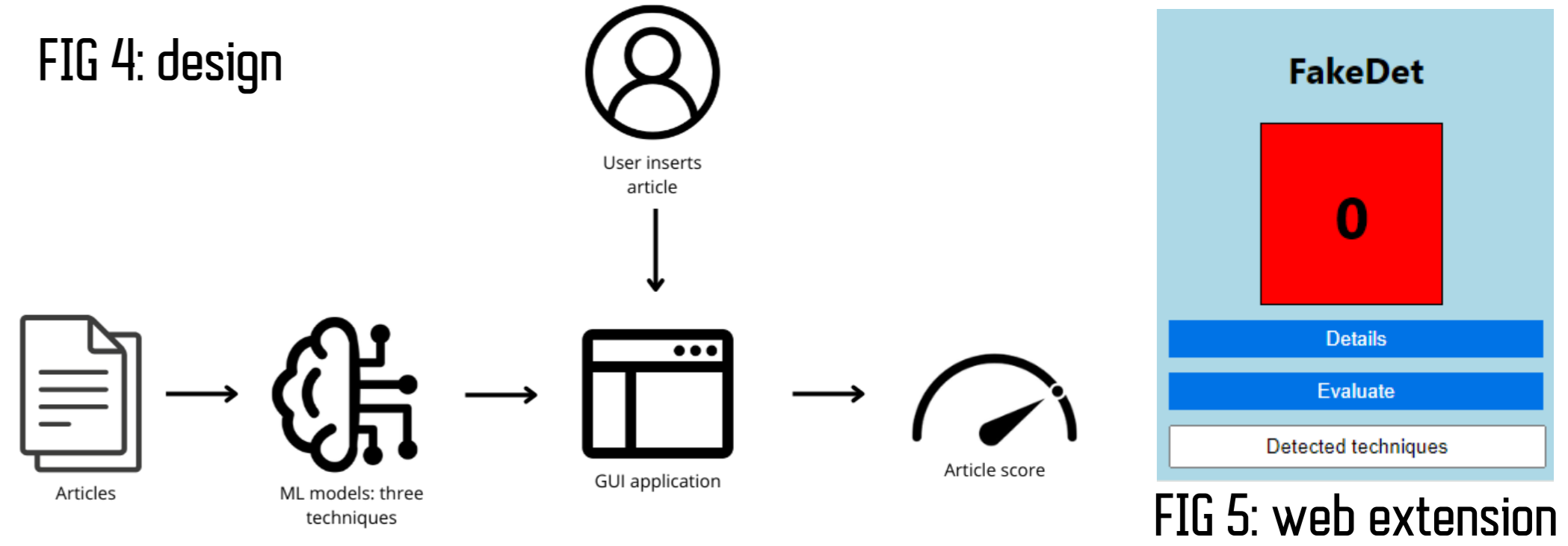
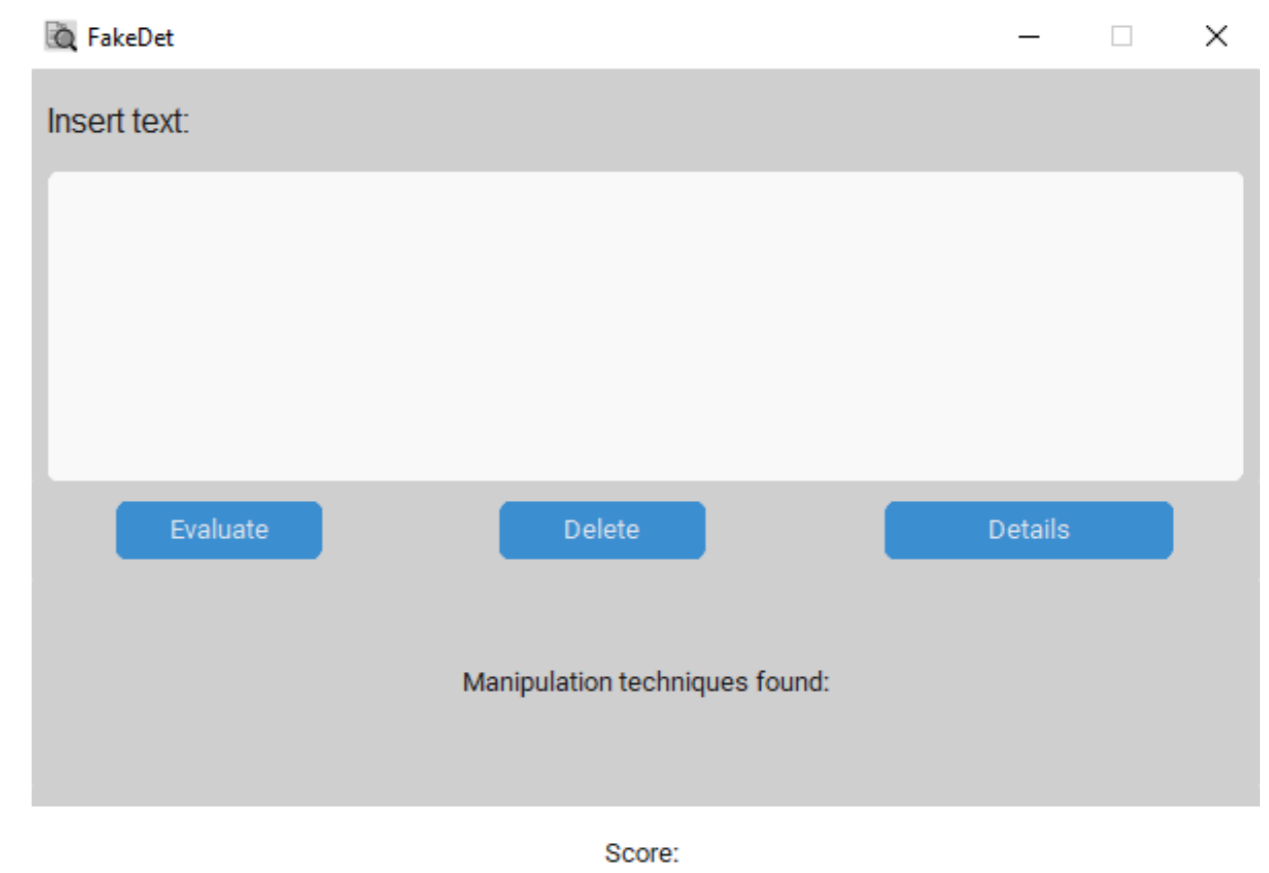


FIG 5: web extension

FIG 6: desktop app



- desktop app and Chrome extension
- all models over 80 % F1 score,
- scoring system -> trustworthiness,

	FakeDet	Verifée	FakeDet category
Novinky.cz	97.5	82.3	Highly trustworthy article
iDnes.cz	92.5	83.7	Highly trustworthy article
iRozhlas.cz	87.5	83.2	Highly trustworthy article
Aeronet.news	41.5	22.4	Unreliable
Novarepublika.cz	30.5	-	Not trustworthy
Protiproud.info	30	37.3	Not trustworthy

- application works correctly.

TAB 1: results

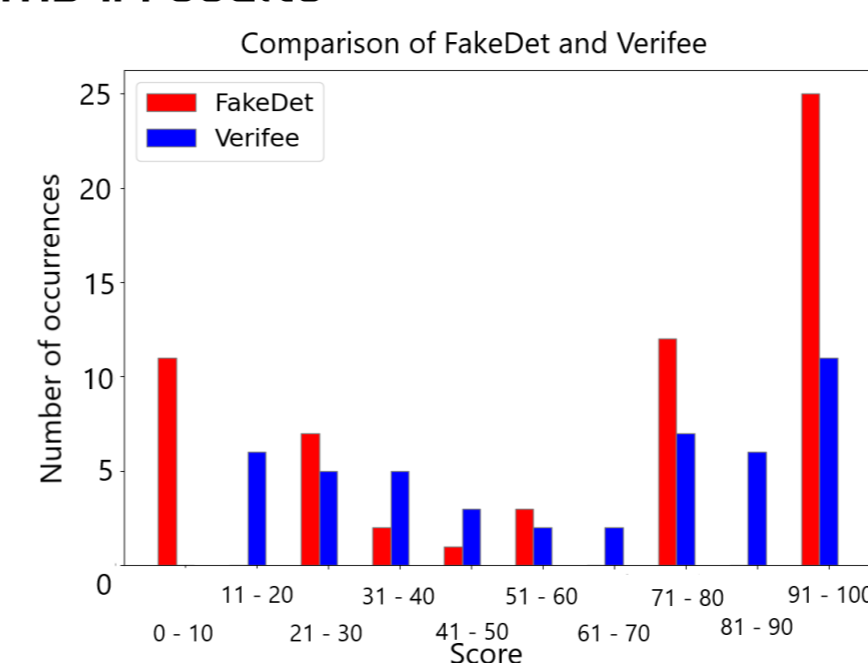


FIG 7: FakeDet vs. Verifée

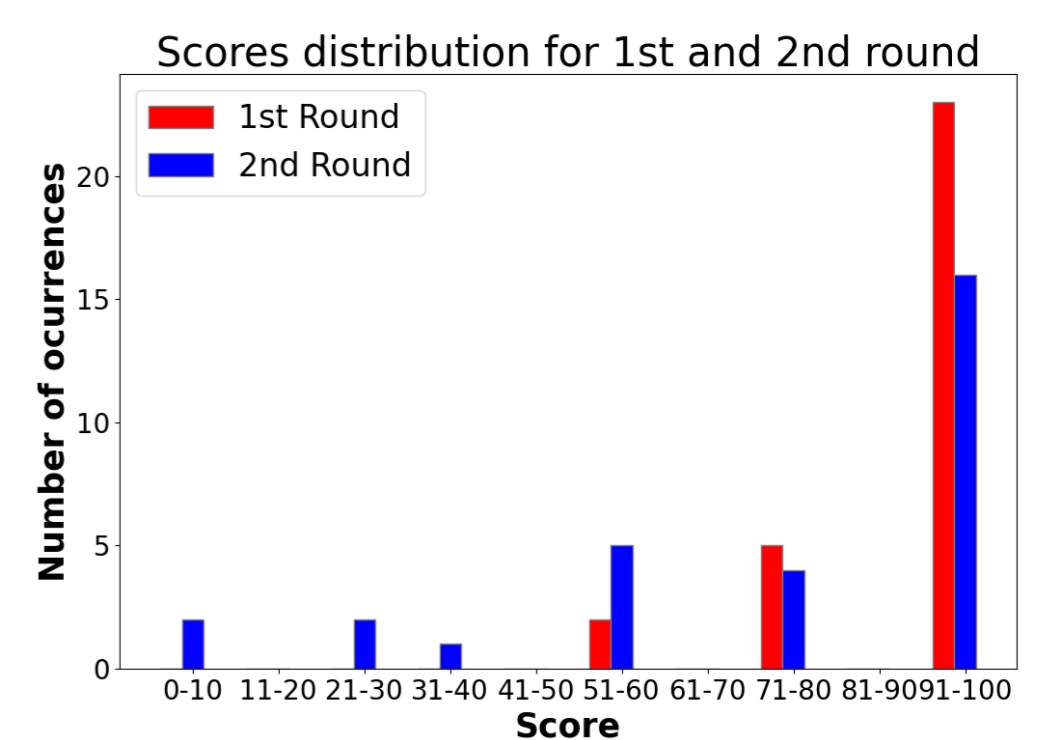


FIG 8: serious news score

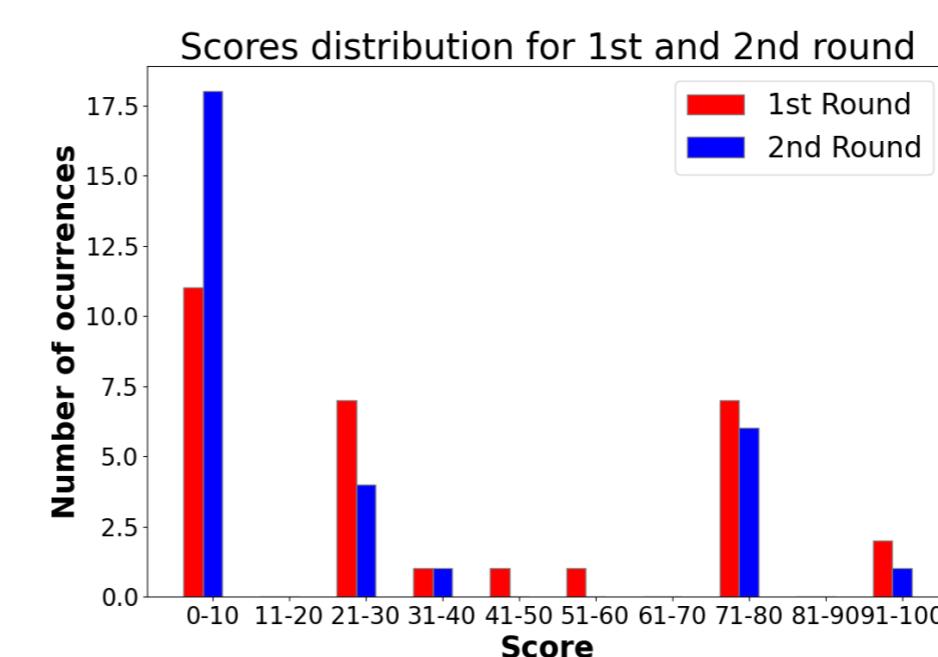


FIG 9: Fake News scores

- 2 testing rounds,
- comparison to Verifée
- 1500 characters vs. whole article (almost)

FIG 10: score examples

