⋘ TRANSFORMERS ⋙
THE DETECTION OF MALICIOUS DOMAINS

Bc. Filip Bučko

supervisor
Ing. Radek Hranický, Ph.D.

CHALLENGE

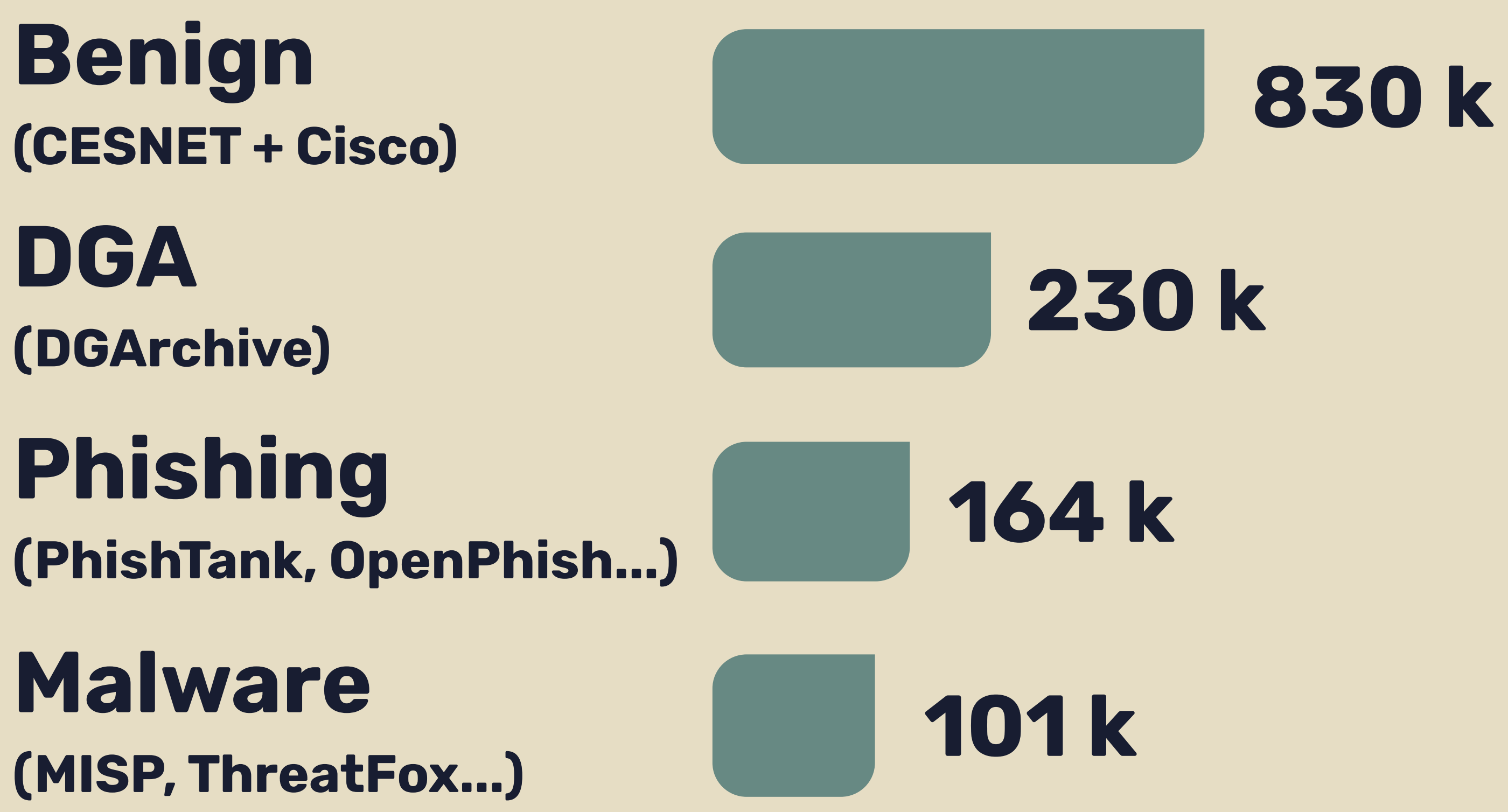
Effective malicious domain detection using current machine learning techniques demands significant expert knowledge for feature engineering, a time-consuming process that attackers continuously exploit.

SOLUTION

The Transformer model trains directly on raw domain text, removing the need for manual, time-consuming feature engineering. It adapts quickly to new threat patterns and delivers high-accuracy, real-time detection.

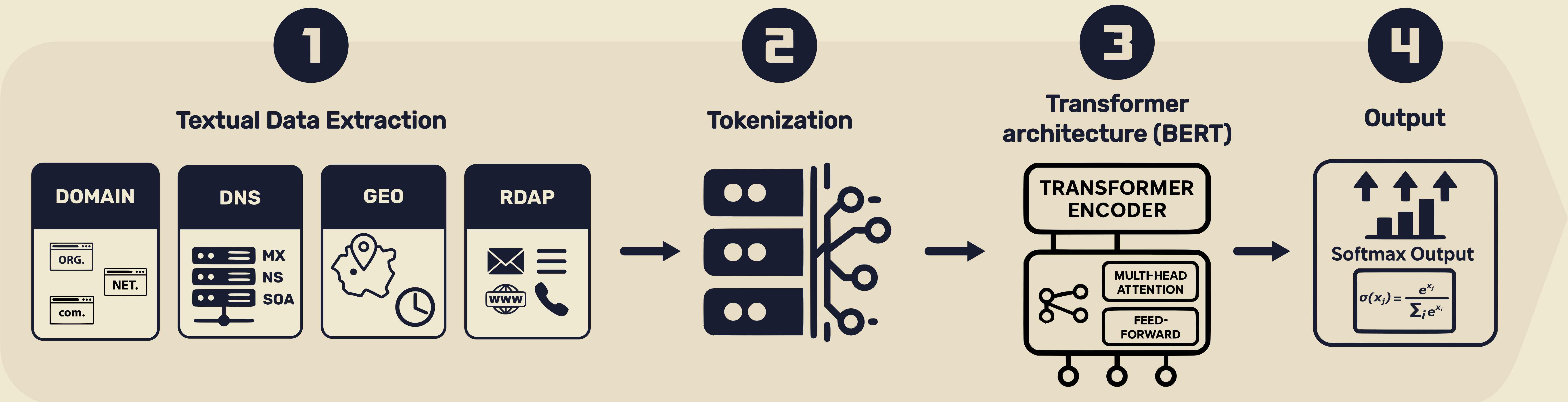
DATASETS

1.33 M labeled domains



RESULTS (F1-scores)

	Malware	Phishing	DGA
Domain name	89%	93%	98.6%
RDAP	95%	98%	-
DNS	95.6%	97.7%	-
Geo Data	95.3%	97.8%	-



1 Textual Data Extraction

DOMAIN · DNS · GEO · RDAP

2 Tokenization

3 Transformer architecture (BERT)

TRANSFORMER ENCODER

MULTI-HEAD ATTENTION

FEED-FORWARD

4 Output

Softmax Output

$$\sigma(x_j) = \frac{e^{x_j}}{\Sigma_j e^{x_i}}$$

CONTRIBUTIONS

- ☑ Experimentally Selected the Best Lightweight Transformer Architecture.
- ☑ Experimentally Determined Optimal Tokenizer (Pre-trained, N-grams, Character-level Tokenization).
- ☑ Feature-less Design Enables Easy Automation for Learning New Threats.
- ☑ Achieved State-of-the-Art Accuracy in Malicious Domain Detection.

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2025

MINISTRY OF THE INTERIOR OF THE CZECH REPUBLIC

NES@FIT