



FACULTY  
OF INFORMATICS

Masaryk University

# E-mail classification

---

Part 4

Maroš Dubíny

# Process 1

Last semester

- Self made LSTM network
- Shallow
- Dead end

## Process 2

### BERT

- Bidirectional Encoder Representations from Transformers
- Best success rate
- Configurable
- Optimized for TPU

## Process 3

### ALBERT

- A Lite BERT
- Too weak
- Rigid
- Still slow

## Process 4

### RoBERTa

- A Robustly Optimized BERT Pretraining Approach
- Best
- High GPU requirements
- Not tested yet

## simpletransformers

Built on top of PyTorch library

```
model = ClassificationModel('albert', 'albert-base-v1', num_labels=4)
```

```
model.train_model(train_set)
```

```
result, model_outputs, wrong_predictions =
```

```
model.eval_model(test_set)
```

## Results

Manual LSTM - too shallow

BERT - No hardware, huge memory requirements

ALBERT - weak , 30% accuracy

RoBERTa - Still running :(

Results will be served on REST API