

---

# Automatic emotion identification in Russian text messages

Marina Kazyulina, Aleksandr Babii and  
Alexey Malafeev





# Plan

- Introduction
- Related work
  
- Data overview
- Corpus annotation
- Model development
  
- Human performance
- Feature importance analysis
- Error analysis
- Search for more features
  
- Conclusion

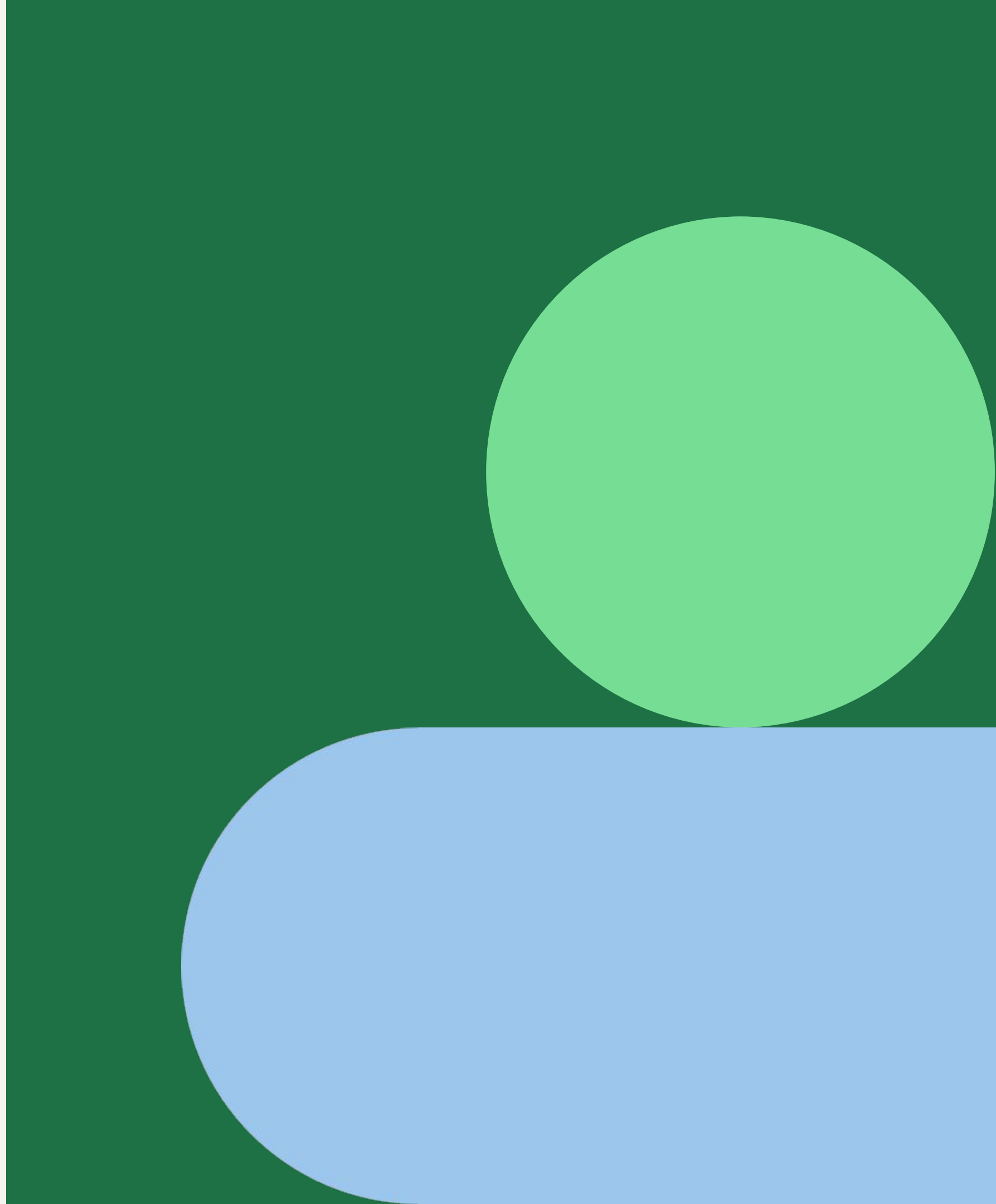
OCTOBER 2020

# Introduction

— PART 1

AIST CONFERENCE

---



# Problem

WHAT WE WANT TO SOLVE

In this paper, we address the issue of identifying emotions in Russian informal text messages. We also report on the importance of different linguistic features of the text messages for the task of automatic emotive analysis.

# Applications

WHY DO WE WANT TO SOLVE IT

- Partial automation of linguistic text analysis and psychological testing in medicine, forensics, and HR management
- New capabilities in marketing (targeting)
- Chat-bots, dialogue systems



# Emotion classification

EKMAN VS HANDCRAFTED

- Happiness
- Surprise
- Sadness
- Anger
- Fear
- Disgust
- Contempt

- Happiness
- Sadness
- Aggression
- Uncertainty
- Neutrality

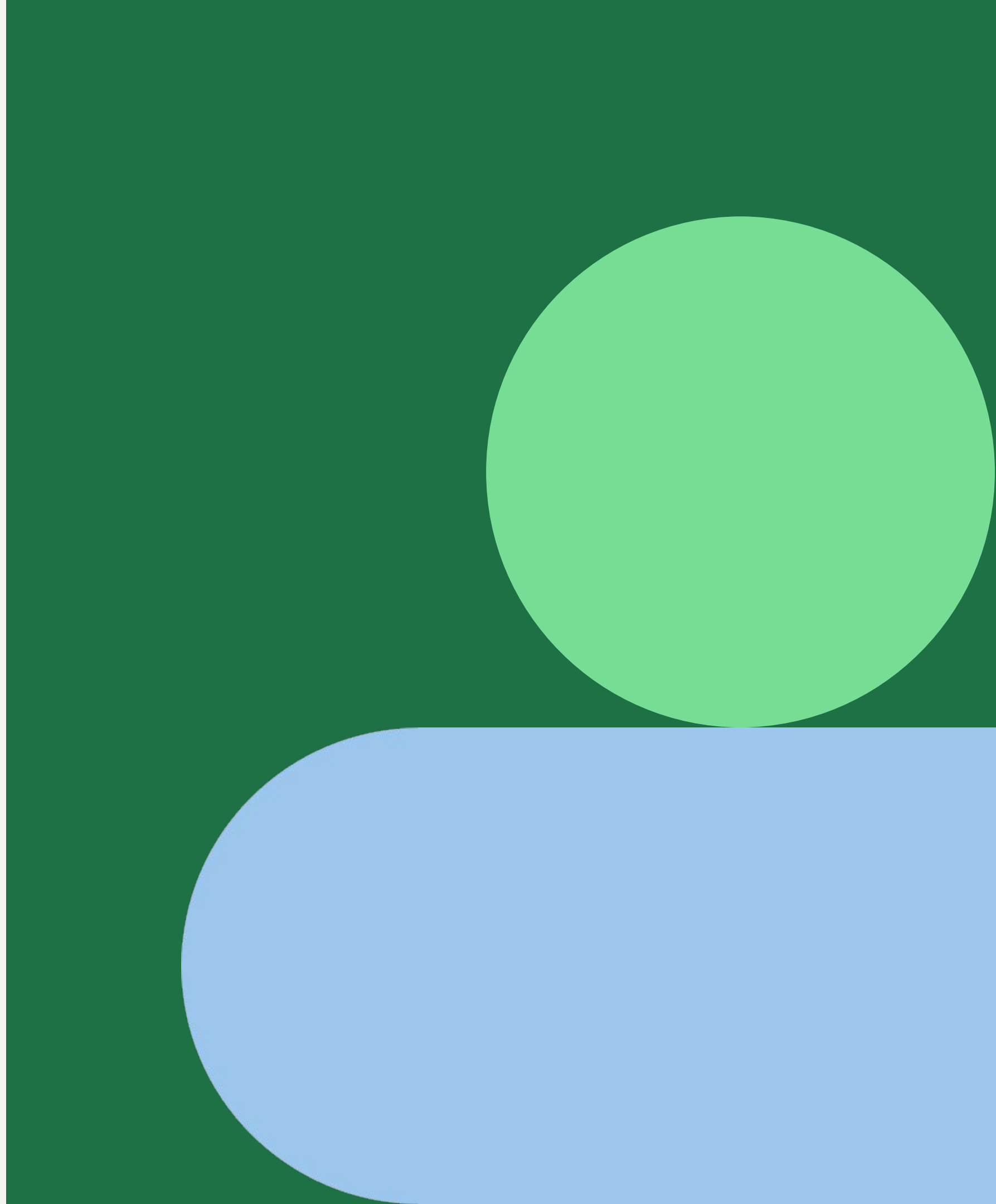
OCTOBER 2020

# Data

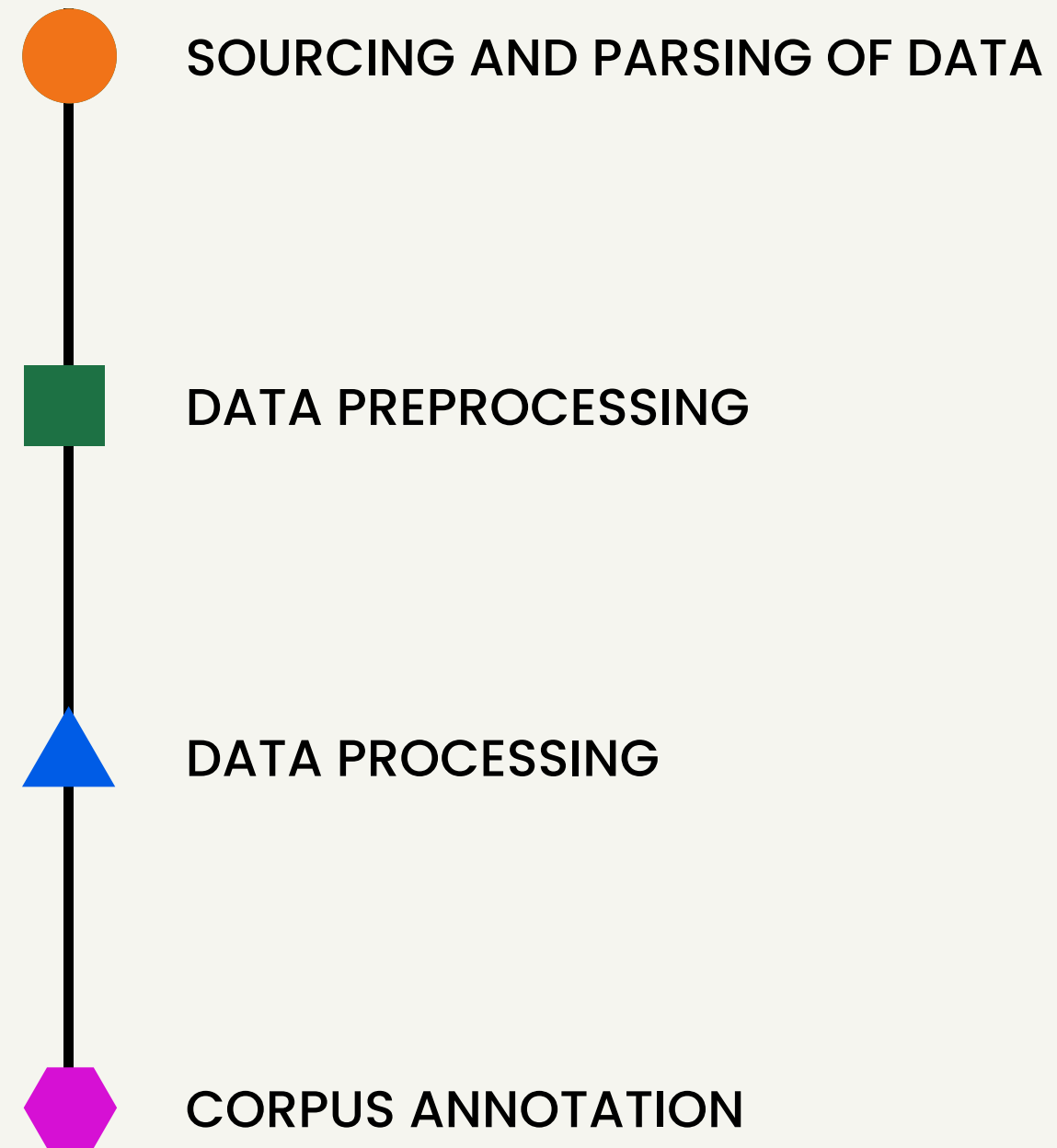
— PART 2

AIST CONFERENCE

---



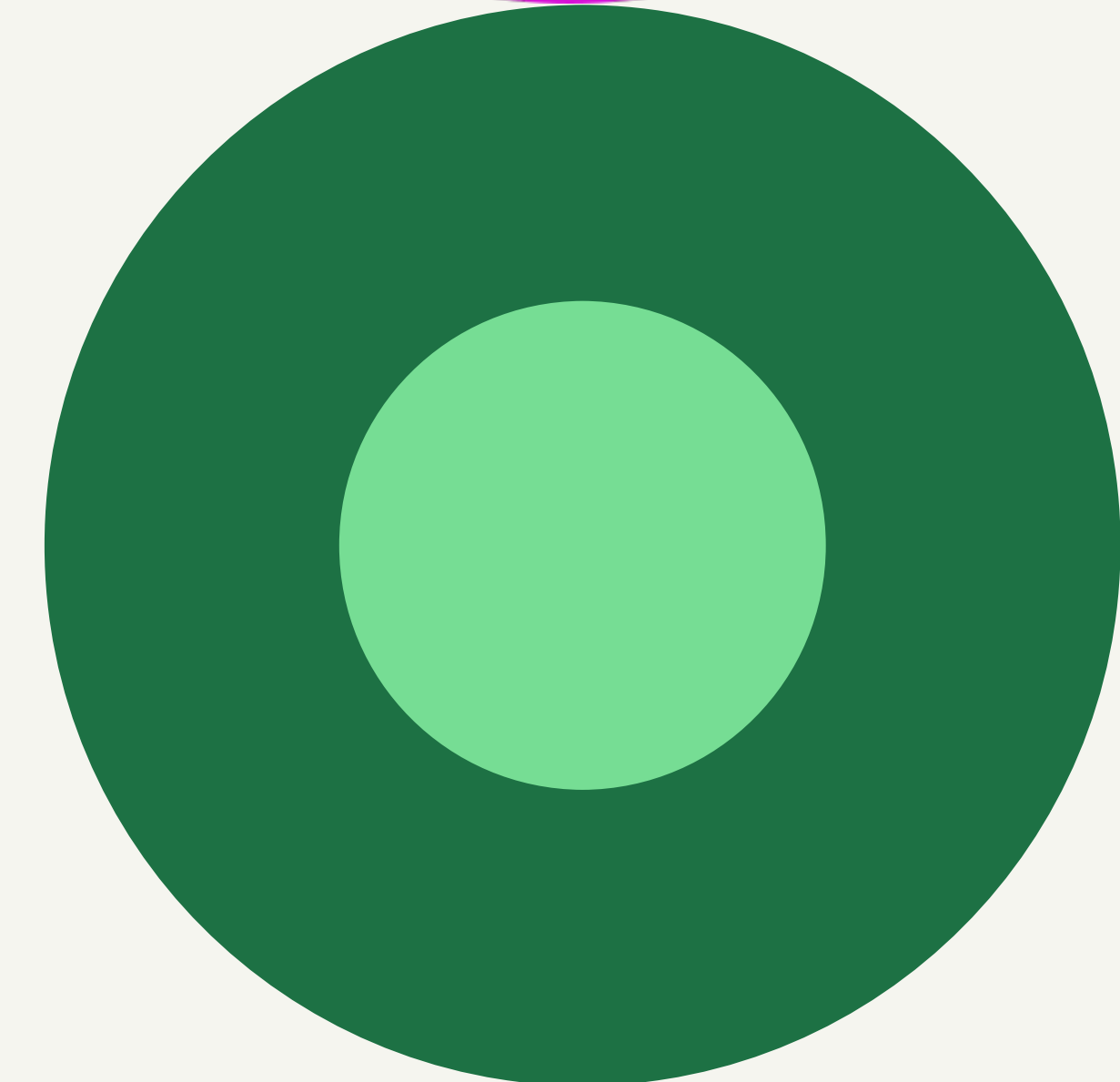
# Stages



# Data overview

## STATISTICAL POPULATION AND THE DATASET SIZE

The statistical population is represented by the sample of 4584 people. Statistically, most of the parsed messages are 1-9 words long and the average length is 4-5 words. The final dataset of lemmatized (but not labeled) messages consists of 1,800,000 messages.





---

# Corpus annotation

## STAGES

- Choice of emoji ideograms as an objective emotion indicator
- Construction of the binary classifiers
- Creation of the final dataset

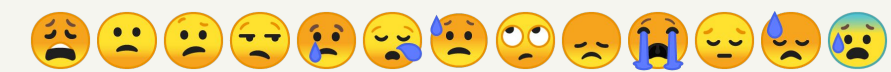


# Emoji sets

## 01 Happiness



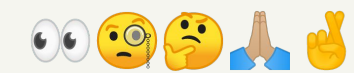
## 02 Sadness



## 03 Anger



## 04 Uncertainty



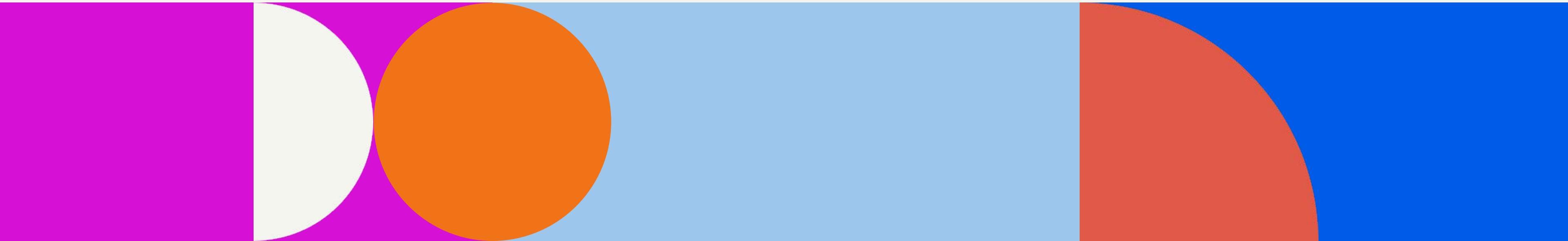
05 Neutral messages were annotated manually because they do not explicitly express emotions and include emoji quite rarely

# Classifiers and the final dataset

- 01 Exclusion of messages containing non-standard (irony, sarcasm) use of emoji (4500 out of 11287 messages containing emoji left)
- 02 Train/Validation sets split (2300 and 2200 messages, respectively)
- 03 Binary classifiers creation and evaluation (F1-score of 90-93%)
- 04 Resulting dataset (19.000 - 24,000 examples of each emotion and 110.000 items in total)

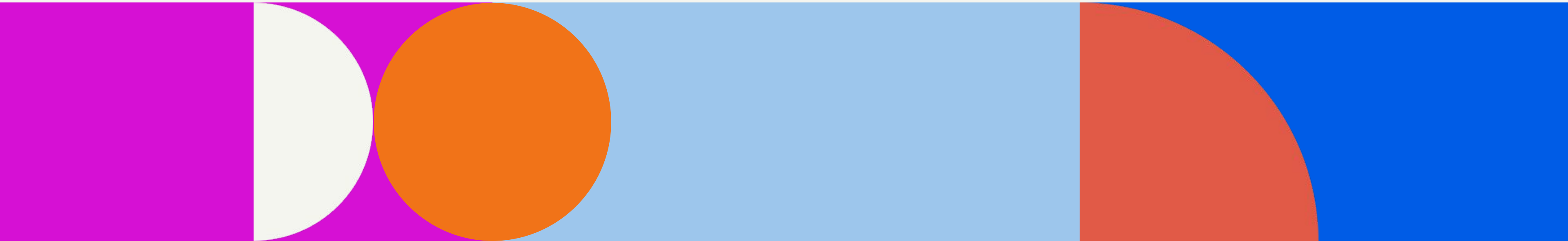
# Model development

	Precision	Recall	F1-Score
Logistic Regression	0.72	0.69	0.70
Naïve Bayes	0.73	0.68	0.69
Random Forest	0.67	0.65	0.65



# Model development

	Precision	Recall	F1-Score
Joy	0.90	0.80	0.85
Sadness	0.72	0.59	0.65
Anger	0.50	0.72	0.59
Interest	0.59	0.71	0.64
Neutrality	0.48	0.59	0.53
Weighted average	0.74	0.71	0.718



OCTOBER 2020

# Discussion of results

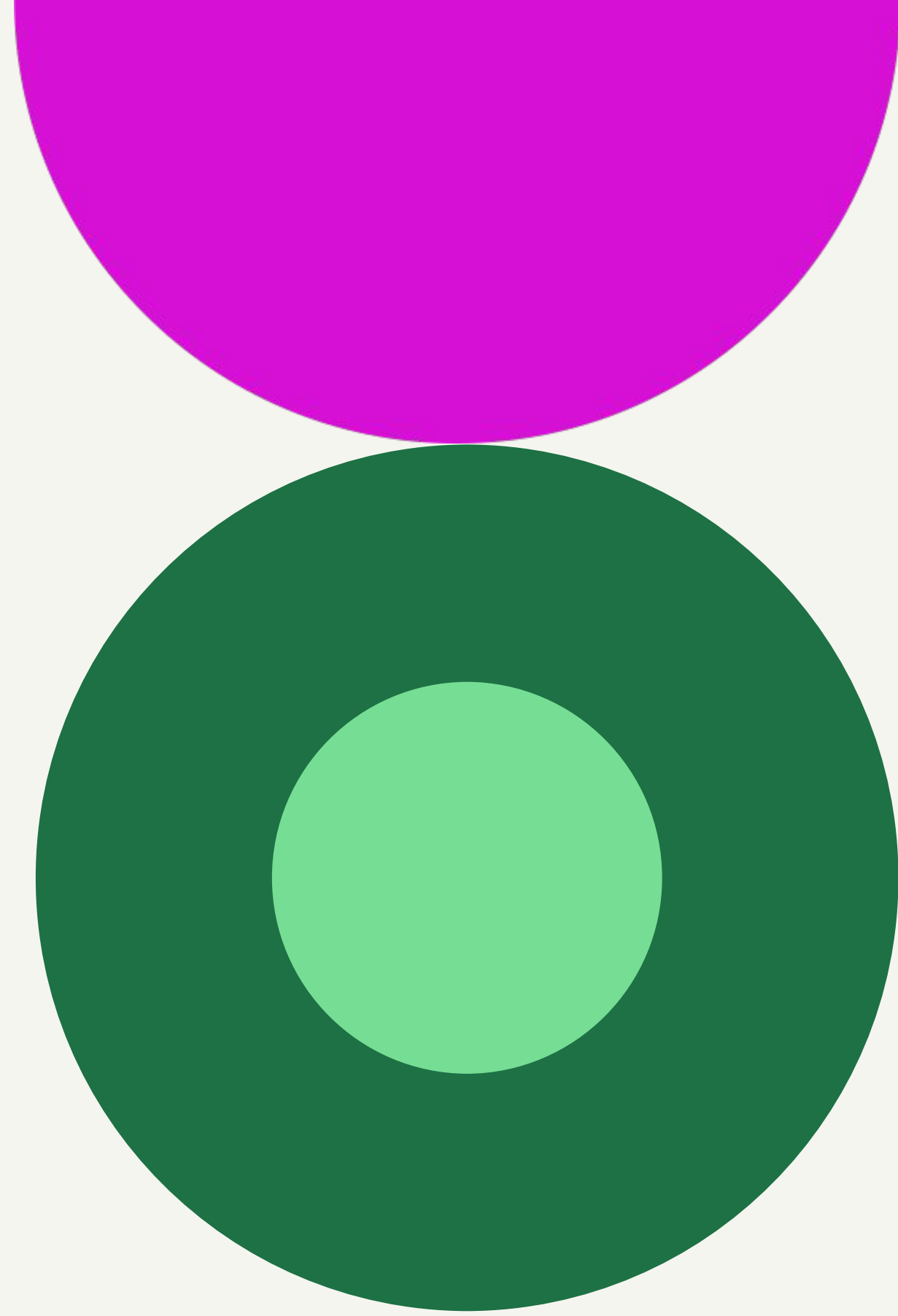
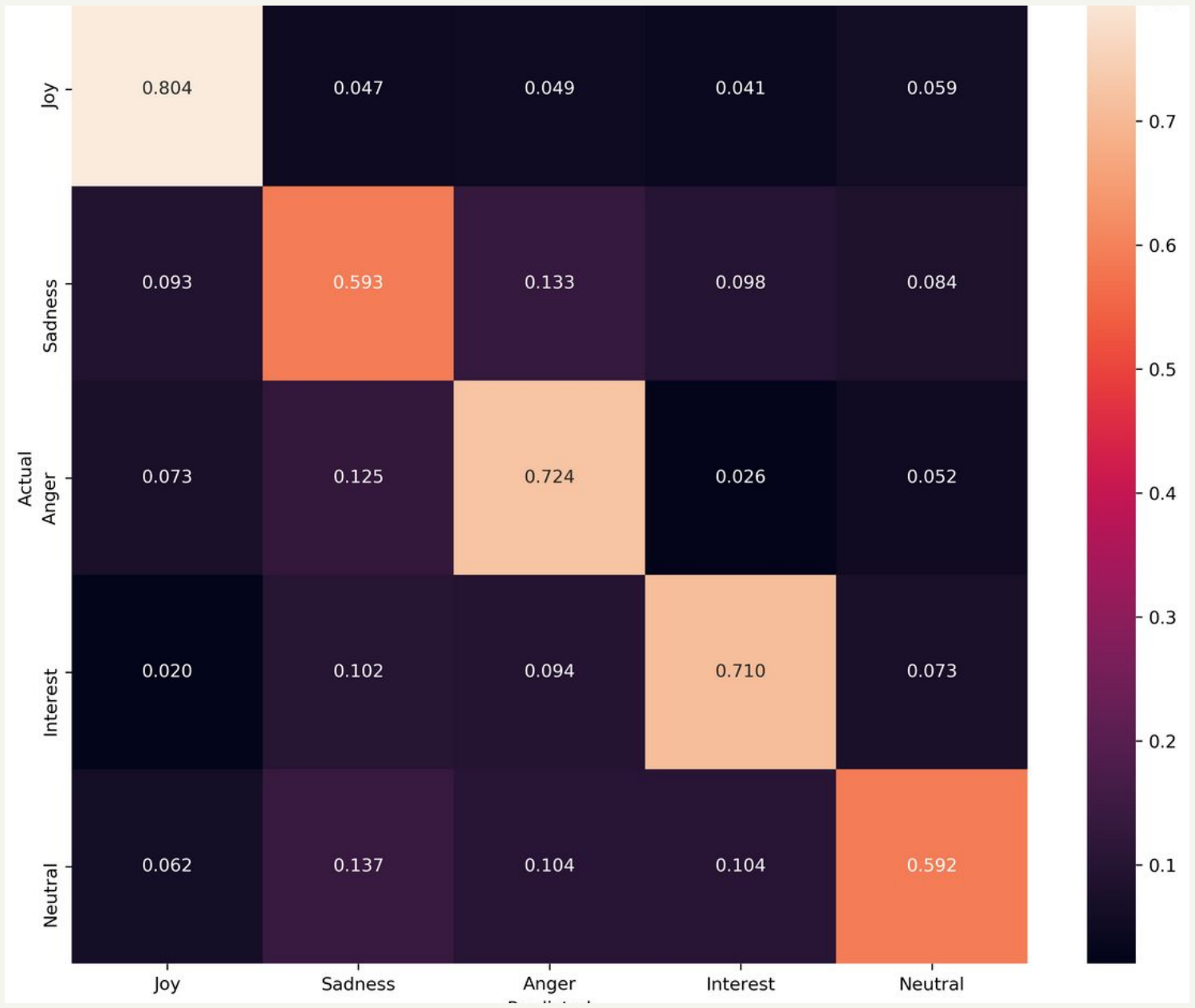
— PART 3



# Human performance

- <sup>01</sup> Human performance — 74%  
(F1-Score = 0.74)
- <sup>02</sup> Agreement between different  
annotators — 70.1%

# Confusion matrix







# Lexical features

Happiness: молодец (well done), ура (yay), солнышко (sweetheart), благодарить, (to thank), обожать (to adore), любить (to love), добрый (kind), думать (to think), удача (fortune), рад (glad).

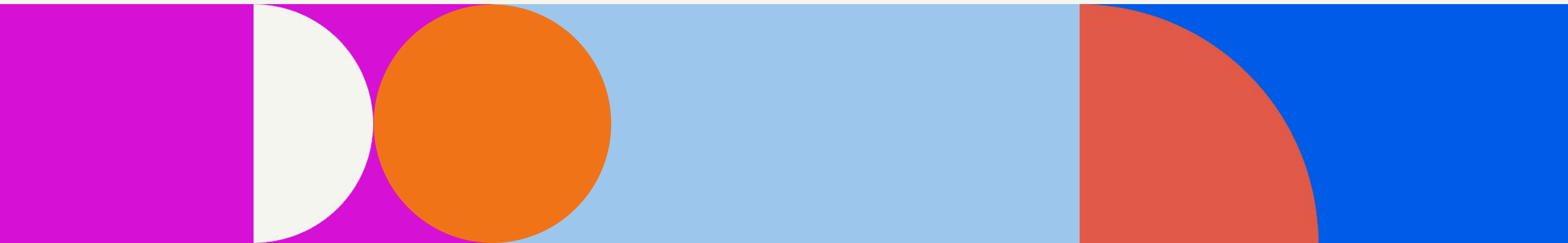
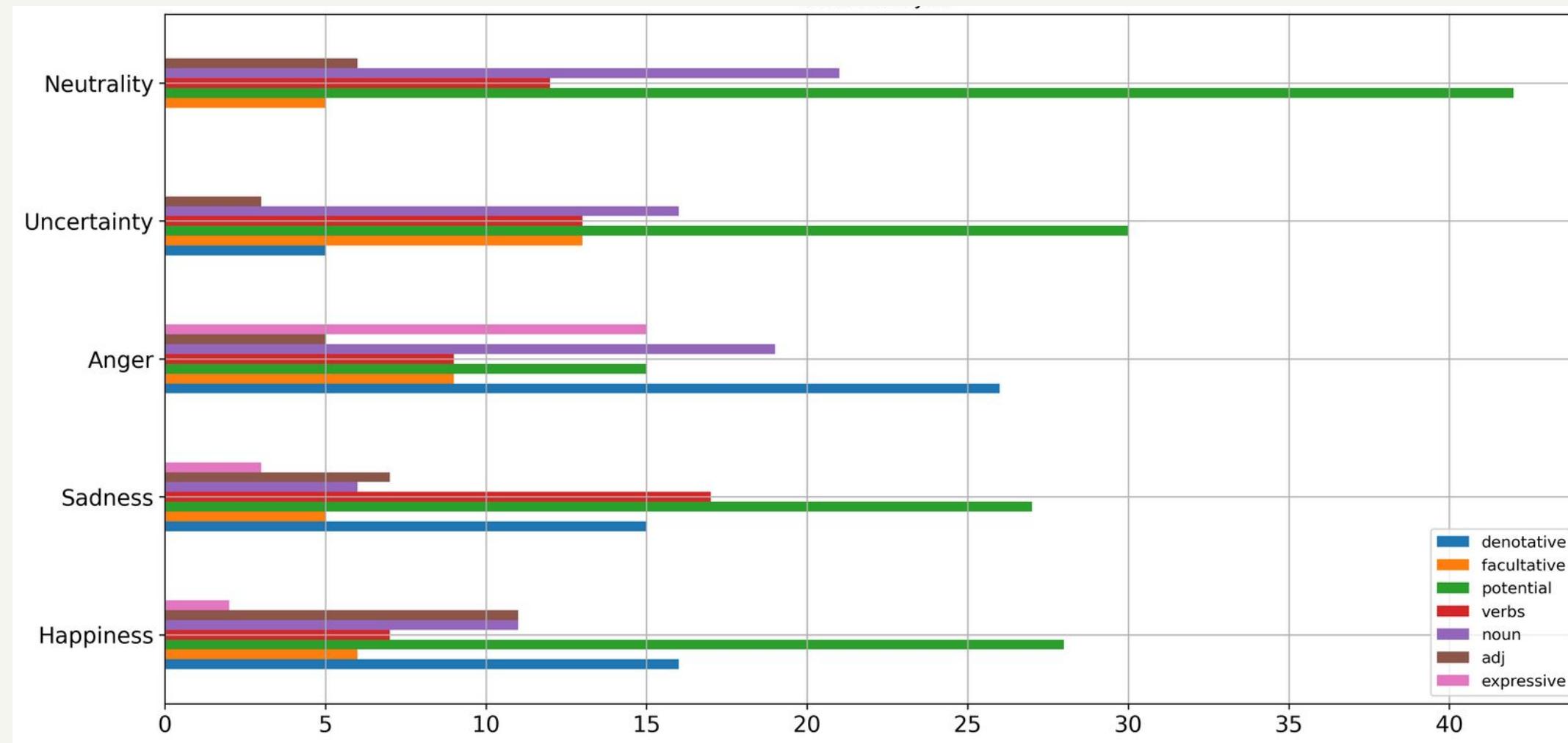
Sadness: скучать (to long for, to feel bored), испортить (to ruin), (never), болеть (to be sick, to hurt), грустный (sad), ничего (nothing), обидеться (to get offended), бедный (poor), жаль (too bad), прощать (to forgive), плакать (to cry), завидовать (to envy), никто (nobody).

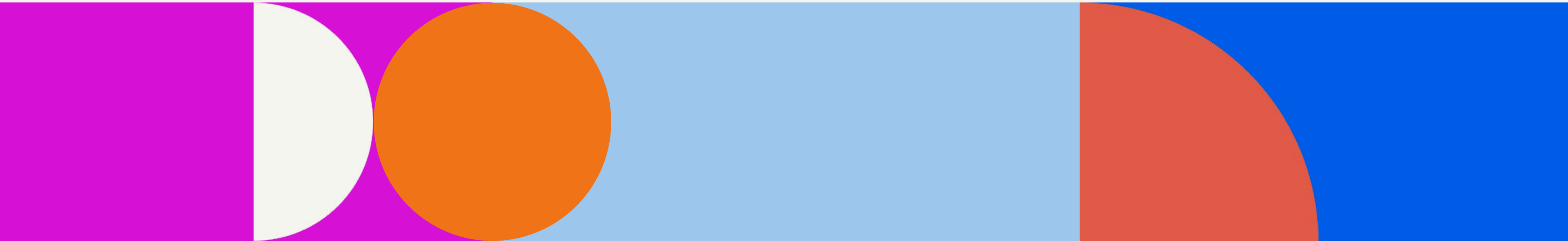
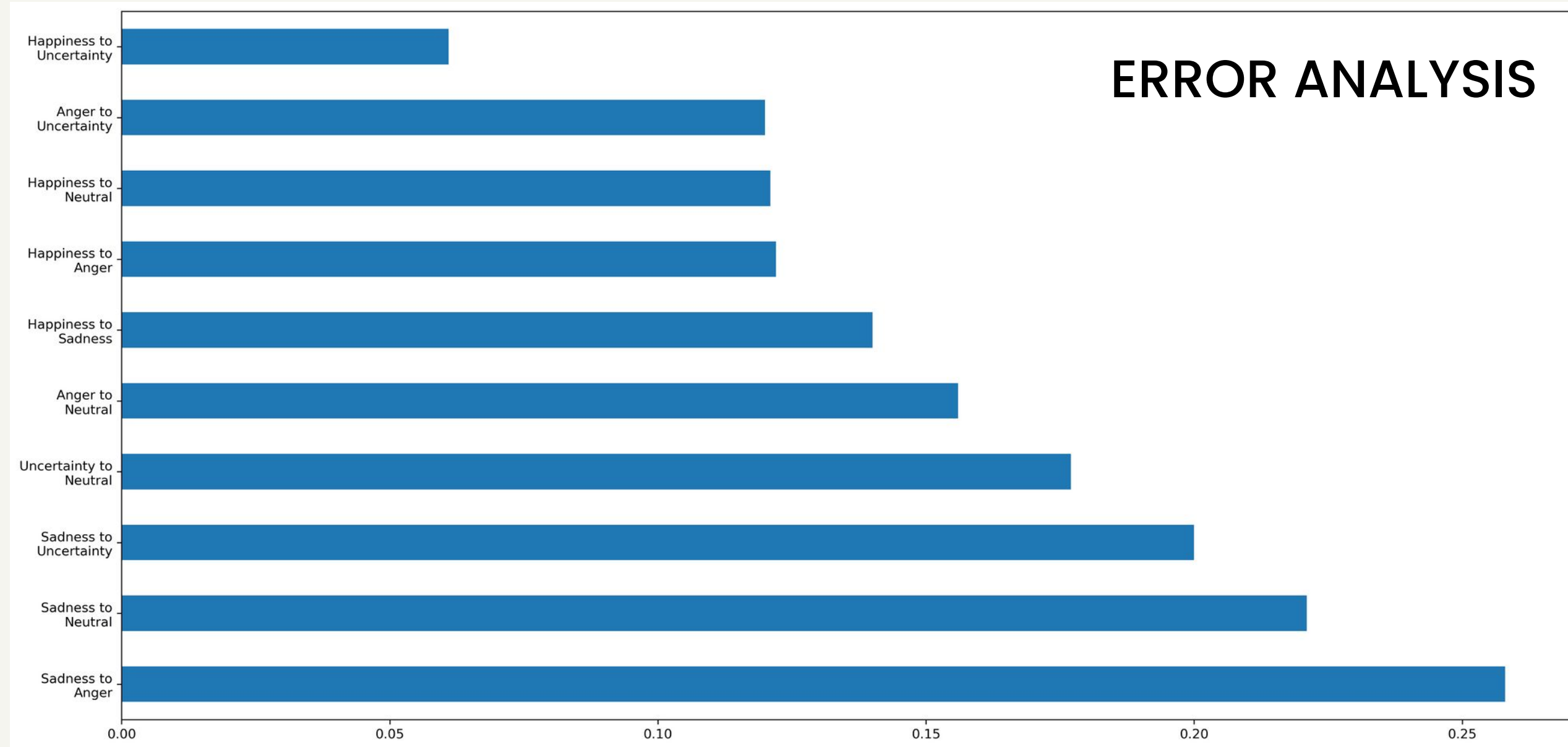
Anger: дебилка (d\*mbass), презирать (to despise), вылететь (to get kicked out, to fly out), ненавидеть (to hate), жесть (brutal, rough, creepy), жирно (greasy, brutal), странный (weird), тупой (stupid), фу (ew), ух (ooh, ugh).

Uncertainty: мб (short of may be), хм (erm), почему (why), казаться (to seem), думать (to think), сколько (how many, how much), интересно (interesting), разве (interrogative particle), ли (interrogative particle), вроде (particle expressing uncertainty, close to I think, probably).

Neutrality: странный (strange, weird), искренний (sincere), уйти (to go away), чисто (purely, simply), география (geography), который (which), знакомый (familiar), дз (short of homework), подняться (to go up), отличаться (to be different).

# FEATURE ANALYSIS





# Steps to eliminate errors



Correct the model's behavior concerning explicit language.



Exclude random word features from Neutrality and manually find and add more distinctive markers (not only words, but probably other linguistic levels as well).



Manually add a greater number of denotative emotives to the list of features.



Use a different emotion classification and/or a way to account for multiple emotionsexpressed in one message.

---

# Search for more features

WHAT MAY HELP THE CLASSIFIER TO MAKE RIGHT CHOICES?

- Number of nouns,
- number of спасибо (thanks),
- ratio of function words,
- message length in words,
- verb ratio,
- maximal length of a word in the message,
- presence of function words,
- the number of interrogative pronouns,
- message length in symbols.



# Recap

— PART 4



- 
- Semi-automatically generated dataset (1.800.000 unlabeled messages)
  - Classifier of emotions with F1-score of 0.718 which is found to be close to human level performance
  - Denotative emotives, level of expressiveness, morphological composition and stylistic markers play a significant role
- 