Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение

высшего образования «Кубанский государственный университет»

Кафедра вычислительных технологий

**ОТЧЕТ**

о выполнении лабораторной работы №2

по дисциплине “Основы компьютерной лингвистики”

Выполнил: ст. гр. 36/2

Прозоров М.С.

Проверил: Вишняков Ю.М

Краснодар

2022

**Краткое описание пакетов и функций**

pymorphy2 написан на языке Python. Он умеет:

1) приводить слово к нормальной форме;

2) ставить слово в нужную форму;

3) возвращать грамматическую информацию о слове.

При работе используется словарь OpenCorpora; для незнакомых слов строятся гипотезы. Библиотека достаточно быстрая: в настоящий момент скорость работы – от нескольких тысяч слов в секунду до более, чем 100 тысяч слов в секунду; потребление памяти – от 10 до 20 Мб; полностью поддерживается буква ё.

Морфологический анализ – это определение характеристик слова на основе того, как это слово пишется. При морфологическом анализе не используется информация о соседних словах.

В pymorphy2 для морфологического анализа русских слов есть класс MorphAnalyzer.

С помощью метода MorphAnalyzer.parse() можно разобрать слово.

**Листинг к заданию 2**

from nltk import word\_tokenize

from nltk import sent\_tokenize

from pymorphy2 import MorphAnalyzer

path = "dataset.txt"

file = open(path, "r", encoding="utf8")

text = file.readlines()

words = []

file.close()

def wordsFromSentences():

for line in text:

sentences = sent\_tokenize(line)

for i in range(len(sentences)):

words.extend(word\_tokenize(sentences[i]))

def morphAnalyze():

morph = MorphAnalyzer()

for i in range(50):

print("Слово " + str(i) + ": " + words[i])

print(str(morph.parse(words[i])))

print()

def main():

wordsFromSentences()

morphAnalyze()

if \_\_name\_\_ == "\_\_main\_\_":

main()

В качестве примера работы программы представлен морфологический разбор первых 50-ти слов из датасета.

**Результат работы программы**

Слово 0: Говорим

[Parse(word='говорим', tag=OpencorporaTag('VERB,impf,tran plur,1per,pres,indc'), normal\_form='говорить', score=0.971428, methods\_stack=((DictionaryAnalyzer(), 'говорим', 415, 2),)), Parse(word='говорим', tag=OpencorporaTag('PRTS,impf,pres,pssv masc,sing'), normal\_form='говорить', score=0.028571, methods\_stack=((DictionaryAnalyzer(), 'говорим', 415, 124),))]

Слово 1: Flutter

[Parse(word='flutter', tag=OpencorporaTag('LATN'), normal\_form='flutter', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'Flutter'),))]

Слово 2: —

[Parse(word='—', tag=OpencorporaTag('PNCT'), normal\_form='—', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), '—'),))]

Слово 3: подразумеваем

[Parse(word='подразумеваем', tag=OpencorporaTag('VERB,impf,tran plur,1per,pres,indc'), normal\_form='подразумевать', score=0.888888, methods\_stack=((DictionaryAnalyzer(), 'подразумеваем', 215, 2),)), Parse(word='подразумеваем', tag=OpencorporaTag('PRTS,impf,pres,pssv masc,sing'), normal\_form='подразумевать', score=0.111111, methods\_stack=((DictionaryAnalyzer(), 'подразумеваем', 215, 97),))]

Слово 4: Dart

[Parse(word='dart', tag=OpencorporaTag('LATN'), normal\_form='dart', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'Dart'),))]

Слово 5: Многие

[Parse(word='многие', tag=OpencorporaTag('NPRO plur,nomn'), normal\_form='многие', score=0.679245, methods\_stack=((DictionaryAnalyzer(), 'многие', 2044, 0),)), Parse(word='многие', tag=OpencorporaTag('ADJF,Apro plur,nomn'), normal\_form='многий', score=0.276729, methods\_stack=((DictionaryAnalyzer(), 'многие', 1009, 19),)), Parse(word='многие', tag=OpencorporaTag('ADJF,Apro inan,plur,accs'), normal\_form='многий', score=0.044025, methods\_stack=((DictionaryAnalyzer(), 'многие', 1009, 22),))]

Слово 6: люди

[Parse(word='люди', tag=OpencorporaTag('NOUN,anim,masc plur,nomn'), normal\_form='человек', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'люди', 3166, 6),))]

Слово 7: уже

[Parse(word='уже', tag=OpencorporaTag('ADVB'), normal\_form='уже', score=0.692307, methods\_stack=((DictionaryAnalyzer(), 'уже', 3, 0),)), Parse(word='уже', tag=OpencorporaTag('PRCL'), normal\_form='уже', score=0.153846, methods\_stack=((DictionaryAnalyzer(), 'уже', 22, 0),)), Parse(word='уже', tag=OpencorporaTag('NOUN,anim,masc sing,loct'), normal\_form='уж', score=0.076923, methods\_stack=((DictionaryAnalyzer(), 'уже', 242, 5),)), Parse(word='уже', tag=OpencorporaTag('COMP,Qual'), normal\_form='узкий', score=0.076923, methods\_stack=((DictionaryAnalyzer(), 'уже', 3079, 31),))]

Слово 8: пользуются

[Parse(word='пользуются', tag=OpencorporaTag('VERB,impf,intr plur,3per,pres,indc'), normal\_form='пользоваться', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'пользуются', 140, 6),))]

Слово 9: приложениями

[Parse(word='приложениями', tag=OpencorporaTag('NOUN,inan,neut plur,ablt'), normal\_form='приложение', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'приложениями', 77, 20),))]

Слово 10: ,

[Parse(word=',', tag=OpencorporaTag('PNCT'), normal\_form=',', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), ','),))]

Слово 11: написанными

[Parse(word='написанными', tag=OpencorporaTag('PRTF,perf,tran,past,pssv plur,ablt'), normal\_form='написать', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'написанными', 959, 67),))]

Слово 12: на

[Parse(word='на', tag=OpencorporaTag('PREP'), normal\_form='на', score=0.998961, methods\_stack=((DictionaryAnalyzer(), 'на', 24, 0),)), Parse(word='на', tag=OpencorporaTag('PRCL'), normal\_form='на', score=0.000849, methods\_stack=((DictionaryAnalyzer(), 'на', 22, 0),)), Parse(word='на', tag=OpencorporaTag('INTJ'), normal\_form='на', score=0.000188, methods\_stack=((DictionaryAnalyzer(), 'на', 21, 0),))]

Слово 13: Flutter

[Parse(word='flutter', tag=OpencorporaTag('LATN'), normal\_form='flutter', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'Flutter'),))]

Слово 14: ,

[Parse(word=',', tag=OpencorporaTag('PNCT'), normal\_form=',', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), ','),))]

Слово 15: но

[Parse(word='но', tag=OpencorporaTag('CONJ'), normal\_form='но', score=0.999185, methods\_stack=((DictionaryAnalyzer(), 'но', 20, 0),)), Parse(word='но', tag=OpencorporaTag('INTJ'), normal\_form='но', score=0.000814, methods\_stack=((DictionaryAnalyzer(), 'но', 21, 0),))]

Слово 16: не

[Parse(word='не', tag=OpencorporaTag('PRCL'), normal\_form='не', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'не', 22, 0),))]

Слово 17: замечают

[Parse(word='замечают', tag=OpencorporaTag('VERB,impf,tran plur,3per,pres,indc'), normal\_form='замечать', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'замечают', 215, 6),))]

Слово 18: этого

[Parse(word='этого', tag=OpencorporaTag('NPRO,neut sing,gent'), normal\_form='это', score=0.729, methods\_stack=((DictionaryAnalyzer(), 'этого', 3240, 1),)), Parse(word='этого', tag=OpencorporaTag('ADJF,Subx,Apro,Anph neut,sing,gent'), normal\_form='этот', score=0.261, methods\_stack=((DictionaryAnalyzer(), 'этого', 3241, 14),)), Parse(word='этого', tag=OpencorporaTag('ADJF,Subx,Apro,Anph masc,sing,gent'), normal\_form='этот', score=0.008, methods\_stack=((DictionaryAnalyzer(), 'этого', 3241, 1),)), Parse(word='этого', tag=OpencorporaTag('ADJF,Subx,Apro,Anph anim,masc,sing,accs'), normal\_form='этот', score=0.002, methods\_stack=((DictionaryAnalyzer(), 'этого', 3241, 4),))]

Слово 19: .

[Parse(word='.', tag=OpencorporaTag('PNCT'), normal\_form='.', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), '.'),))]

Слово 20: Например

[Parse(word='например', tag=OpencorporaTag('CONJ,Prnt'), normal\_form='например', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'например', 2109, 0),))]

Слово 21: ,

[Parse(word=',', tag=OpencorporaTag('PNCT'), normal\_form=',', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), ','),))]

Слово 22: мини-приложение

[Parse(word='мини-приложение', tag=OpencorporaTag('NOUN,inan,neut sing,nomn'), normal\_form='мини-приложение', score=0.5, methods\_stack=((DictionaryAnalyzer(), 'мини-приложение', 121, 0),)), Parse(word='мини-приложение', tag=OpencorporaTag('NOUN,inan,neut sing,accs'), normal\_form='мини-приложение', score=0.5, methods\_stack=((DictionaryAnalyzer(), 'мини-приложение', 121, 3),))]

Слово 23: Gmail

[Parse(word='gmail', tag=OpencorporaTag('LATN'), normal\_form='gmail', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'Gmail'),))]

Слово 24: ,

[Parse(word=',', tag=OpencorporaTag('PNCT'), normal\_form=',', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), ','),))]

Слово 25: приложения

[Parse(word='приложения', tag=OpencorporaTag('NOUN,inan,neut sing,gent'), normal\_form='приложение', score=0.485714, methods\_stack=((DictionaryAnalyzer(), 'приложения', 77, 2),)), Parse(word='приложения', tag=OpencorporaTag('NOUN,inan,neut plur,nomn'), normal\_form='приложение', score=0.285714, methods\_stack=((DictionaryAnalyzer(), 'приложения', 77, 13),)), Parse(word='приложения', tag=OpencorporaTag('NOUN,inan,neut plur,accs'), normal\_form='приложение', score=0.228571, methods\_stack=((DictionaryAnalyzer(), 'приложения', 77, 18),))]

Слово 26: BMW

[Parse(word='bmw', tag=OpencorporaTag('LATN'), normal\_form='bmw', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'BMW'),))]

Слово 27: ,

[Parse(word=',', tag=OpencorporaTag('PNCT'), normal\_form=',', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), ','),))]

Слово 28: eBay

[Parse(word='ebay', tag=OpencorporaTag('LATN'), normal\_form='ebay', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'eBay'),))]

Слово 29: ,

[Parse(word=',', tag=OpencorporaTag('PNCT'), normal\_form=',', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), ','),))]

Слово 30: Tencent

[Parse(word='tencent', tag=OpencorporaTag('LATN'), normal\_form='tencent', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'Tencent'),))]

Слово 31: .

[Parse(word='.', tag=OpencorporaTag('PNCT'), normal\_form='.', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), '.'),))]

Слово 32: В

[Parse(word='в', tag=OpencorporaTag('PREP'), normal\_form='в', score=0.999327, methods\_stack=((DictionaryAnalyzer(), 'в', 393, 0),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,gent'), normal\_form='в', score=0.000249, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 1),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,loct'), normal\_form='в', score=5.7e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 5),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,nomn'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 0),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,datv'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 2),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,accs'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 3),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,ablt'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 4),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,nomn'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 6),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,gent'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 7),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,datv'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 8),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,accs'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 9),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,ablt'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 10),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,loct'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 11),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Name,Fixd,Abbr,Init sing,nomn'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Name,Fixd,Abbr,Init sing,gent'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Name,Fixd,Abbr,Init sing,datv'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Name,Fixd,Abbr,Init sing,accs'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Name,Fixd,Abbr,Init sing,ablt'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Name,Fixd,Abbr,Init sing,loct'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Name,Fixd,Abbr,Init sing,nomn'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Name,Fixd,Abbr,Init sing,gent'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Name,Fixd,Abbr,Init sing,datv'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Name,Fixd,Abbr,Init sing,accs'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Name,Fixd,Abbr,Init sing,ablt'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Name,Fixd,Abbr,Init sing,loct'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedFirstNameAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Name,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Patr,Fixd,Abbr,Init sing,nomn'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Patr,Fixd,Abbr,Init sing,gent'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Patr,Fixd,Abbr,Init sing,datv'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Patr,Fixd,Abbr,Init sing,accs'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Patr,Fixd,Abbr,Init sing,ablt'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,masc,Sgtm,Patr,Fixd,Abbr,Init sing,loct'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Patr,Fixd,Abbr,Init sing,nomn'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Patr,Fixd,Abbr,Init sing,gent'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Patr,Fixd,Abbr,Init sing,datv'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Patr,Fixd,Abbr,Init sing,accs'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Patr,Fixd,Abbr,Init sing,ablt'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),)), Parse(word='в', tag=OpencorporaTag('NOUN,anim,femn,Sgtm,Patr,Fixd,Abbr,Init sing,loct'), normal\_form='в', score=0.0, methods\_stack=((AbbreviatedPatronymicAnalyzer(letters='АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЭЮЯ', score=0.1, tag\_pattern='NOUN,anim,%(gender)s,Sgtm,Patr,Fixd,Abbr,Init sing,%(case)s'), 'В'),))]

Слово 33: Яндексе

[Parse(word='яндексе', tag=OpencorporaTag('NOUN,inan,masc,Orgn sing,loct'), normal\_form='яндекс', score=1.0, methods\_stack=((DictionaryAnalyzer(), 'яндексе', 48, 5),))]

Слово 34: ,

[Parse(word=',', tag=OpencorporaTag('PNCT'), normal\_form=',', score=1.0, methods\_stack=((PunctuationAnalyzer(score=0.9), ','),))]

Слово 35: в

[Parse(word='в', tag=OpencorporaTag('PREP'), normal\_form='в', score=0.999327, methods\_stack=((DictionaryAnalyzer(), 'в', 393, 0),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,gent'), normal\_form='в', score=0.000249, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 1),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,loct'), normal\_form='в', score=5.7e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 5),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,nomn'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 0),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,datv'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 2),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,accs'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 3),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,ablt'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 4),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,nomn'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 6),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,gent'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 7),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,datv'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 8),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,accs'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 9),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,ablt'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 10),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,loct'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 11),))]

Слово 36: частности

[Parse(word='частности', tag=OpencorporaTag('NOUN,inan,femn sing,loct'), normal\_form='частность', score=0.980988, methods\_stack=((DictionaryAnalyzer(), 'частности', 13, 5),)), Parse(word='частности', tag=OpencorporaTag('NOUN,inan,femn plur,nomn'), normal\_form='частность', score=0.007604, methods\_stack=((DictionaryAnalyzer(), 'частности', 13, 6),)), Parse(word='частности', tag=OpencorporaTag('NOUN,inan,femn sing,gent'), normal\_form='частность', score=0.003802, methods\_stack=((DictionaryAnalyzer(), 'частности', 13, 1),)), Parse(word='частности', tag=OpencorporaTag('NOUN,inan,femn sing,datv'), normal\_form='частность', score=0.003802, methods\_stack=((DictionaryAnalyzer(), 'частности', 13, 2),)), Parse(word='частности', tag=OpencorporaTag('NOUN,inan,femn plur,accs'), normal\_form='частность', score=0.003802, methods\_stack=((DictionaryAnalyzer(), 'частности', 13, 9),))]

Слово 37: в

[Parse(word='в', tag=OpencorporaTag('PREP'), normal\_form='в', score=0.999327, methods\_stack=((DictionaryAnalyzer(), 'в', 393, 0),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,gent'), normal\_form='в', score=0.000249, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 1),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,loct'), normal\_form='в', score=5.7e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 5),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,nomn'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 0),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,datv'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 2),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,accs'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 3),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr sing,ablt'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 4),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,nomn'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 6),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,gent'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 7),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,datv'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 8),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,accs'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 9),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,ablt'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 10),)), Parse(word='в', tag=OpencorporaTag('NOUN,inan,masc,Fixd,Abbr plur,loct'), normal\_form='в', score=1.9e-05, methods\_stack=((DictionaryAnalyzer(), 'в', 349, 11),))]

Слово 38: Яндекс

[Parse(word='яндекс', tag=OpencorporaTag('NOUN,inan,masc,Orgn sing,nomn'), normal\_form='яндекс', score=0.764705, methods\_stack=((DictionaryAnalyzer(), 'яндекс', 48, 0),)), Parse(word='яндекс', tag=OpencorporaTag('NOUN,inan,masc,Orgn sing,accs'), normal\_form='яндекс', score=0.235294, methods\_stack=((DictionaryAnalyzer(), 'яндекс', 48, 3),))]

Слово 39: Про

[Parse(word='про', tag=OpencorporaTag('PREP'), normal\_form='про', score=0.988212, methods\_stack=((DictionaryAnalyzer(), 'про', 24, 0),)), Parse(word='про', tag=OpencorporaTag('NOUN,inan,femn,Sgtm,Fixd,Abbr sing,nomn'), normal\_form='про', score=0.001964, methods\_stack=((DictionaryAnalyzer(), 'про', 586, 0),)), Parse(word='про', tag=OpencorporaTag('NOUN,inan,femn,Sgtm,Fixd,Abbr sing,gent'), normal\_form='про', score=0.001964, methods\_stack=((DictionaryAnalyzer(), 'про', 586, 1),)), Parse(word='про', tag=OpencorporaTag('NOUN,inan,femn,Sgtm,Fixd,Abbr sing,datv'), normal\_form='про', score=0.001964, methods\_stack=((DictionaryAnalyzer()al\_form='на', score=0.000188, methods\_stack=((DictionaryAnalyzer(), 'на', 21, 0),))]

Слово 49: Dart[Parse(word='dart', tag=OpencorporaTag('LATN'), normal\_form='dart', score=1.0, methods\_stack=((LatinAnalyzer(score=0.9), 'Dart'),))]

**Приложение**

**Датасет, используемый в ЛР №2**

Говорим Flutter — подразумеваем Dart

Многие люди уже пользуются приложениями, написанными на Flutter, но не замечают этого. Например, мини-приложение Gmail, приложения BMW, eBay, Tencent.

В Яндексе, в частности в Яндекс Про, фреймворк Flutter используется для написания мобильных приложений на Dart. Поэтому внутри мы их не разделяем — работа с Flutter предполагает знание Dart.

Фреймворк — это набор готовых абстракций, участков кода, функций и методов, которые можно подключить в Dart-проект всего одной строчкой. А ещё Flutter — это целая идеология и набор правил: как писать код и какие приёмы использовать. Это целый мир, существующий по своим законам и традициям.

Преимущества Flutter для разработчиков

— Кроссплатформенность. На Flutter легко переходить тем, кто уже пробовал создавать мобильные приложения на iOS или Android. При этом код получается качественным, потому что у команды появляется больше времени на доработку и тестирование: не нужно распылять внимание на несколько приложений, чтобы проверить, как в них реализованы одни и те же функции.

— Работать на Flutter может одна команда вместо двух или даже трёх. У нас в команде есть люди, которые пришли с хорошим знанием Flutter, есть бывшие Android- и iOS-разработчики. Все они ценны тем, что знают специфику разных приложений и при этом решают совместные задачи.

— Для того чтобы сделать хорошее коммерческое приложение, достаточно выучить всего один язык — Dart. Для сравнения: в нативной разработке нужно разобраться в специфике Android и выучить минимум два языка — Java и XML. К тому же придётся повозиться с Android-специфичными штуками. В случае с Flutter достаточно задать в консоли проект и писать его на Dart, чтобы создать приложение, которое работает «из коробки». Да, в дальнейшем для некоторых фич, связанных с платформенной спецификой, понадобится и нативность, но базово будет достаточно Dart.

— У Dart низкий порог вхождения, поэтому освоить Flutter можно довольно быстро. Особенно если вы уже знаете несколько языков программирования. Так было со мной: когда я начинал учить Flutter, я просто открыл документацию Dart, прочёл ее часа за три, и начал на нём писать. Какие-то сверхнавыки, вроде функционального программирования, здесь не нужны. Освоите базу — и это уже даст вам возможность начать программировать на Flutter. Легко переходить на Flutter и тем, кто пробовал писать приложения на React. JS и Dart похожи реактивным UI, подходами к управлению состоянием и вёрсткой в коде.

— Flutter позволяет больше времени уделять тому, что важно для пользователя: приятному UI, стабильной работе, так как реализация многих фич на Flutter занимает ощутимо меньше времени, чем их нативные аналоги.

— Этот фреймворк идеален для прототипирования, особенно в том случае, когда хочется просто проверить какую-то гипотезу: можно за пару дней сделать приложение и протестировать его. Flutter обеспечивает очень высокую скорость разработки, если вы умеете с ним обращаться и хорошо знакомы с API Flutter.

Кроссплатформенность является не только главным преимуществом Flutter, но и причиной, по которой он появился. Часто стоит задача разработать одинаковые приложения для iOS, Android и браузера. Большая компания может позволить себе несколько команд, а игрокам поменьше важно оптимизировать процесс работы. Так получился Flutter. Он похож на бензин, который подойдёт любому автомобилю: с ним заведётся и Android, и iPhone, и десктоп, и Fuchsia, и даже в вебе всё будет работать. Правда, понадобятся дополнительные усилия для того, чтобы приложение было нативным и удобным для пользователей.

Заменяет ли Flutter нативную разработку

Приложение, написанное на Flutter, везде выглядит и работает одинаково — и на Xiaomi, и на iPhone. Это удобно, но в то же время теряется платформенно-специфичный опыт и нативность.

В мобильной разработке нативным считается приложение, написанное под конкретную платформу и до мелочей учитывающее её специфику. Например, для iOS есть набор дизайн-гайдлайнов, описывающих, как должно выглядеть и работать хорошее приложение. В такие гайды входит анимация, способы прокручивания списков, индикаторы загрузки и другие стандарты.

Существенная разница между операционными системами состоит в концептуальных различиях: по-разному называются и работают одни и те же сущности, отличаются ограничения. С Flutter легко создавать универсальный базовый интерфейс, но для разработки более серьёзных функций — отправки уведомлений, получения геолокации и работы виджетов — нужно знать платформенную специфику.

Для пользователя нативность приложения проявляется в том, что оно выглядит узнаваемым, в нём интуитивно понятный интерфейс, к которому не нужно долго привыкать. Конечно, во Flutter можно написать код так, чтобы приложение выглядело максимально органично, но всё равно останется ощущение, что «что-то не то». Особенно это характерно для Flutter-приложений в iOS. Так что Flutter не является полной заменой нативной разработке.