

**Московский государственный технический
университет им. Н.Э. Баумана.**

Факультет «Информатика и управление»

Кафедра «Системы обработки информации и управления»

Курс «Базовые компоненты интернет-технологий»

Отчет по индивидуальному домашнему заданию

Выполнил:

студент группы ИУ5-32

Морозенков Олег

Подпись и дата:

Проверил:

преподаватель каф. ИУ5

Гапанюк Ю.Е.

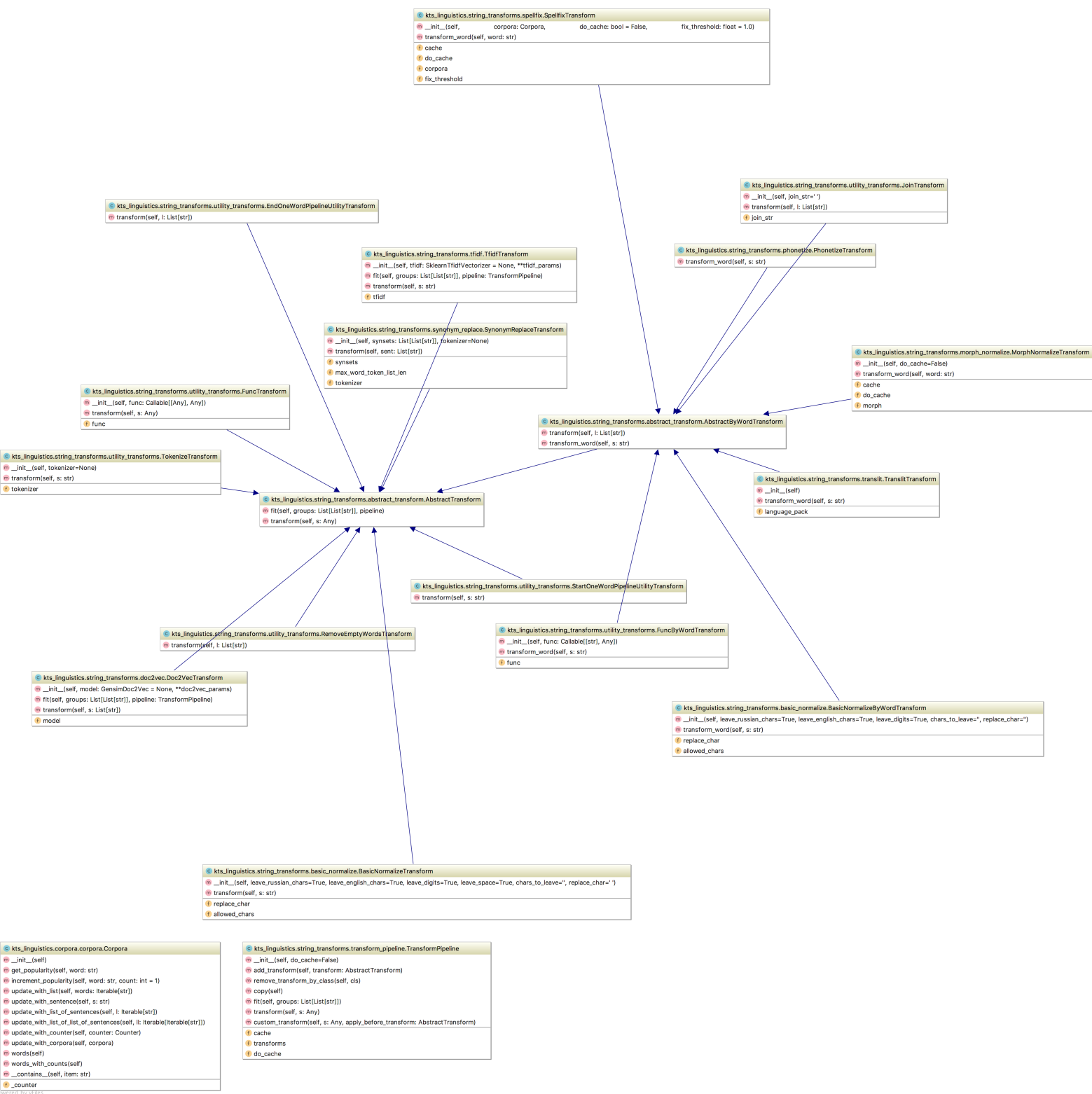
Подпись и дата:

г. Москва, 2018 г.

Описание задания:

Разработать программу, умеющая исправлять опечатки в тексте.

Диаграмма классов:



Текст программы (только функция main):

```
from kts_linguistics.corpora.loaders import load_precomputed_corpora
from kts_linguistics.string_transforms.basic_normalize import
BasicNormalizeByWordTransform
from kts_linguistics.string_transforms.spellfix import SpellfixTransform
from kts_linguistics.string_transforms.transform_pipeline import
TransformPipeline
from kts_linguistics.string_transforms.translit import TranslitTransform
from kts_linguistics.string_transforms.utility_transforms import
TokenizeTransform, JoinTransform

if __name__ == '__main__':
    # setup
    pipeline = TransformPipeline()

    corpora = load_precomputed_corpora()

    pipeline.add_transform(TokenizeTransform())
    pipeline.add_transform(TranslitTransform())
    pipeline.add_transform(BasicNormalizeByWordTransform())
    pipeline.add_transform(SpellfixTransform(corpora=corpora,
fix_threshold=0.3))
    pipeline.add_transform(JoinTransform())

    # test
    while True:
        input_str = input()
        if input_str == '' or input_str == 'END':
            break
        fixed_str = pipeline.transform(input_str)
        print(input_str, '->', fixed_str)
```

Примеры:

```
Run: example x
/usr/local/bin/python3.7 /Users/reo7sp/Documents/code/_kts/kts_linguistics/examples/example.py
Приложение с ашибками
Приложение с ашибками -> предложение с ошибками
Привет
Привет -> привет
Превед
Превед -> превед
МГТУ им Н Э Боумана
МГТУ им Н Э Боумана -> мгту им н э баумана
Раз джва три
Раз джва три -> раз два три

Process finished with exit code 0
|
```