**Федеральное агентство связи**

**Государственное бюджетное образовательное учреждение высшего образование**

**Ордена Трудового Красного Знамени**

**«Московский технический университет связи и информатики»**

**Кафедра “Математической кибернетики и информационных технологий”**

Отчет по учебной практике

Natural language processing

Выполнила студентка группы БФИ1801

Калатанова Е.С.

Москва 2020

**1. Задачи:**

1) Разработать функцию на Python, которая принимает строку и число N и выдает “True”, если длина строки больше числа N, “False”, если длина строки меньше числа N.

2) Разработать функцию на Python, которая принимает строку и выдает кортеж, который содержит “True”, если данный вопрос связан с количеством баллов на направление, и номер направления, если он указан в вопросе, или “False”, в противном случае.

3) Запустить модель Named Entity Recognition (NER) ner\_rus DeepPavlov.

**2. Алгоритм решения задачи**

1) Сравниваем значения длины строки и числа N, если больше, то выдаем “True”, иначе “False”.

Для данной задачи нам необходимо изучить как строятся функции в Python, как работать со строками и условиями.

2) Идет поиск в строке подстроки «Балл» или «балл», при нахождении данной подстроки ищем шаблон направления \d{2}.\d{2}.\d{2}, если нашел только подстроку, то кортеж будет состоять из “True”, если еще и направление, то будет состоять из “True” и номера направления, иначе “False”.

Для данной задачи нам необходимо изучить что такое кортеж в Python, регулярные выражения и воспользоваться модулем re, который ищет номер направления с помощью шаблона.

3) Для начала необходимо запустить виртуальное окружение venv и активировать его. Дальше в изолированной среде установить DeepPavlov, после чего устанавливаем модель ner\_rus. При успешной установке используем модель с помощью кода, который указан в листинге 3 программы.

Для данной задачи нам необходимо изучить как работать с виртуальным окружением в Python и как запустить модель DeepPavlov.

**3. Листинг программы**

1)a = input("Введите строку: ")

b = input("Введите число: ")  
 def strt(a,b):  
 if (len(a)>int(b)):  
 d = True  
 else:  
 d = False  
 return d  
 print(strt(a,b))

2)import re

vopros = input("Введите вопрос: ")

def proverka(vopros):

if (vopros.find("балл") != -1) or (vopros.find("Балл") != -1):

d = (True,);

result = re.findall('\d{2}.\d{2}.\d{2}', vopros)

if str(\*result) != "":

d = (True, \*result);

else:

d = (False,);

return d

print(proverka(vopros))

3)python -m venv env  
 .\env\Scripts\activate.bat  
 pip install DeepPavlov   
 python -m deeppavlov install ner\_rus  
 from deeppavlov import configs,build\_model  
 ner\_model = build\_model(configs.ner.ner\_rus, download=True)

**4. Результат работы**

Результат работы первой программы представлен на рисунке 1.

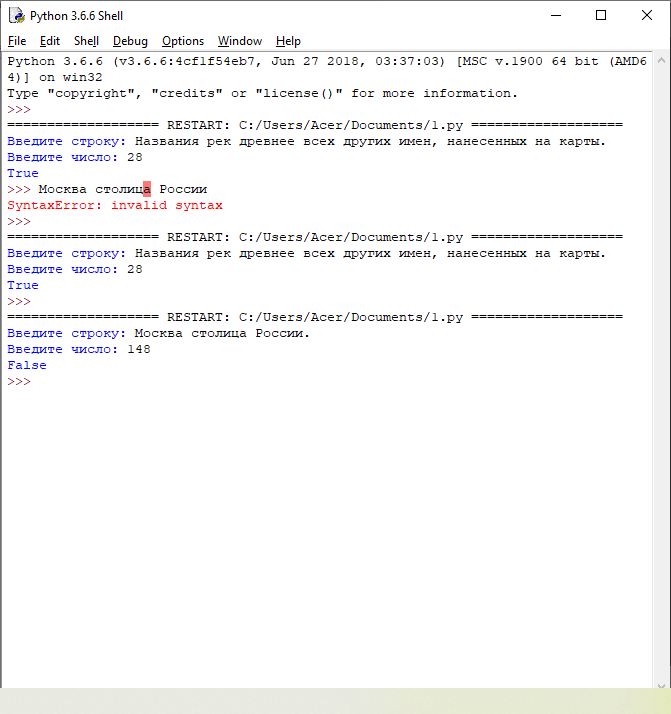


Рисунок 1 – Работа первой программы

Результат работы второй программы представлен на рисунке 2.

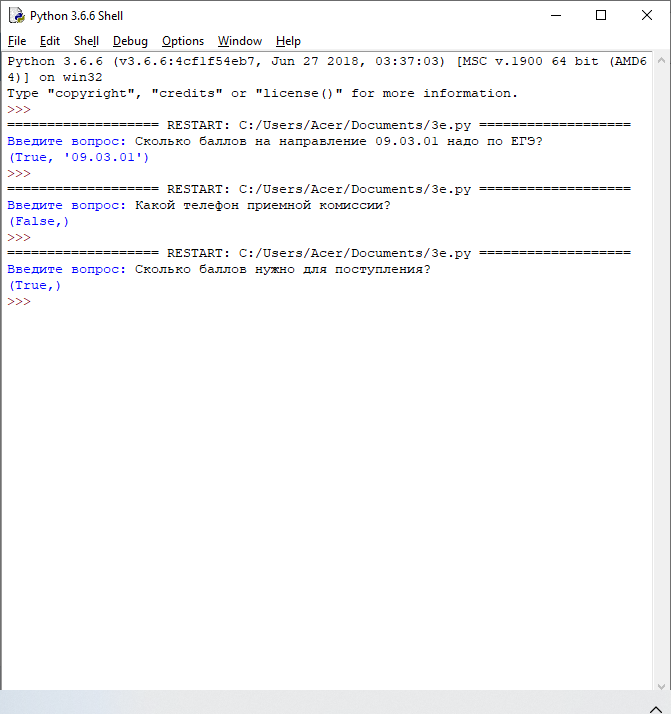


Рисунок 2 – Работа второй программы

Результат работы третьей программы представлен на рисунке 3.

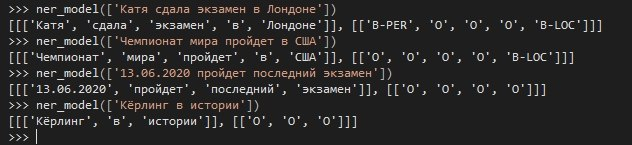


Рисунок 3 – Работа третьей программы

**Вывод**

В ходе практики были выполнены все задачи, изучены основы работы с Python (как строятся функции, как работать со строками, как подключать модули и т.д.), изучены основы работы с регулярными выражениями, получены практические навыки запуска модели DeepPavlov.