



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА СИСТЕМЫ ОБРАБОТКИ ИНФОРМАЦИИ И УПРАВЛЕНИЯ

Лабораторная работа

по дисциплине: «Методы машинного обучения в АСОИУ»

Студент	Саргсян Ованнес Гагикович
Группа	ИУ5-21М
Название	Парсинг данных
Вариант	Номер варианта

Студент	_____	Саргсян О.Г.
	<i>подпись, дата</i>	<i>фамилия, и.о.</i>

Преподаватель	_____	Григорьев Ю.А.
	<i>подпись, дата</i>	<i>фамилия, и.о.</i>

Оценка _____

Москва, 2024 г.

СОДЕРЖАНИЕ

1. Описание задания	3
2. Текст программы.....	3
3. Экранные формы с примерами выполнения программы	7

1. Описание задания

- Выбрать набор данных (XML датасет).
- Применить методы парсинга к данным для получения HTML формата

2. Текст программы

Для формирования метаграфа в требуемом виде произведем импорт данных (см листинг 1).

Листинг 1. Импорт библиотек и данных

```
from bs4 import BeautifulSoup
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

with open('mg.xml', 'r', encoding='utf-8') as f:
    file = f.read()

soup = BeautifulSoup(file, 'xml')
```

Далее производим генерацию HTML кода, прогоняя каждый тег-элемент через специальные циклы.

Листинг 2. Генератор HTML

```
x = 0
y = 0
z = 0

# <a-box position="0 0 0" scale="1 1 1" material="opacity: 0.5; color:
#0000FF;"></a-box>

# <a-text position="0 1 0" value="Hello, World!" color="black" scale="2
2 2"></a-text>

# <a-text value="Ссешь" font="custom-msdf.json" color="black" ne-
gate="false"></a-text>
```

```

cc = []
s=0
for i, name in enumerate(soup.find_all('Metavertex')):
    x = df.loc[i, 'x']
    y = df.loc[i, 'y']
    z = df.loc[i, 'z']

    # print(name.attrs)
    s = df.loc[i, 2]
    # s = len(list(name.children))
    cc.append(s)
    # s = s*(1/3)

    print(f""""<a-entity mixin="sphere changeScale" position="{x} {y} {z}"
geometry="radius: {s/2};"></a-entity>""")
    # print(f""""<a-text position="{x} {y+s} {z}" value="{name.at-
trs['name']}" color="black" scale="1 1 1"></a-text>""")
    # print(f""""<a-text position="{x} {y+s} {z}" value="{name.text}"
color="black" scale="1 1 1" ></a-text>""")
    # print(f""""<a-text position="{x} {y+s} {z}" value="{name.text}"
font="custom-msdf.json" color="black" negate="false"></a-text>""")

    # x+=1
    # y+=1

```

```

old_vertex = 'semesters'

c=0

for name in soup.find_all('Attribute'):

    if name.parent.name == 'Metavertex':

        if old_vertex == name.parent["name"]:

            c+=0.2

        else:

            c=0

        old_vertex = name.parent["name"]

        mask = df[1] == name.parent["name"]

        x, y, z = df[mask][['x', 'y', 'z']].values[0]

        s = df[mask][2].values[0]

        if s>14:

            print(f""""<a-text  position="{x-s*0.5}  {y+s*0.4-c}  {z+s*0.5}"
value="{name["name"]} +: '+' name.text}" scale="5 5 5" color="black" font="cus-
tom-msdf.json" negate="false"></a-text>""")

        else:

            print(f""""<a-text  position="{x-s*0.5}  {y+s*0.4-c}  {z+s*0.5}"
value="{name["name"]} +: '+' name.text}"  color="black" font="custom-
msdf.json" negate="false"></a-text>""")

            # print( name.parent["name"], c, x, y, z)

cc = np.array(cc)

for name in soup.find_all(['StartVertexRef', 'EndVertexRef']):

    x = df[df[1] == name["ref"]].x.values[0]

    y = df[df[1] == name["ref"]].y.values[0]

```

```
z = df[df[1] == name["ref"]].z.values[0]
```

```
if name.name == "StartVertexRef":
```

```
    print(f""""<a-entity line="start: {x} {y} {z}; """, end="")
```

```
else:
```

```
    print(f""""end: {x} {y} {z}; color: green; opacity: 0.3;"></a-entity>""")
```

```
for name in soup.find_all(['StartVertexRef', 'EndVertexRef']):
```

```
    x = df[df[1] == name["ref"]].x.values[0]
```

```
    y = df[df[1] == name["ref"]].y.values[0]
```

```
    z = df[df[1] == name["ref"]].z.values[0]
```

```
if name.name == "StartVertexRef":
```

```
    # print(f""""<a-entity line="start: {x} {y} {z}; """, end="")
```

```
    sx = x
```

```
    sy = y
```

```
    sz = z
```

```
else:
```

```
    ex = x
```

```
    ey = y
```

```
    ez = z
```

```
    x = (sx+ex)/2
```

```
    y = (sy+ey)/2
```

```
    z = (sz+ez)/2
```

```
print(f""<a-text position="{x} {y} {z}" value="{name.parent.Attribute.text}" color="black" font="custom-msdf.json" negate="false"></a-text>""")

# print(f""end: {x} {y} {z}; color: green; opacity: 0.3;"></a-entity>""")
```

3. Экранные формы с примерами выполнения программы

Листинг 3 представляет вывод программы.

Листинг 3. Вывод программы

```
<!-- Семестры -->

<a-entity mixin="sphere changeScale" position="-91.0 -7.0 -122.0" geometry="radius: 15.0;"></a-entity>

<a-entity mixin="sphere changeScale" position="-92.66805466155743 -14.048175986861073 -115.10501567048253" geometry="radius: 1.0;"></a-entity>

<a-entity mixin="sphere changeScale" position="-92.43405759291883 -14.84528258692184 -115.96717148001596" geometry="radius: 1.0;"></a-entity>

<a-entity mixin="sphere changeScale" position="-94.66342312094461 -15.852048909348547 -119.13284790368044" geometry="radius: 1.0;"></a-entity>

<a-entity mixin="sphere changeScale" position="-90.98589674758738 -7.007399500160495 -131.99998731727538" geometry="radius: 1.0;"></a-entity>

<a-entity mixin="sphere changeScale" position="-98.71102559615095 -6.123478986861557 -128.30648833892775" geometry="radius: 1.0;"></a-entity>
```

```

    <a-entity mixin="sphere changeScale" position="-97.21400043430101 -
13.6384018894375 -126.16146836547142" geometry="radius: 1.0;"></a-entity>
    <a-entity mixin="sphere changeScale" position="-84.5446900859149 -
14.522400246880379 -123.3200259235517" geometry="radius: 1.0;"></a-en-
tity>
    <a-entity mixin="sphere changeScale" position="-89.13814405027541 -
16.80731330474253 -122.59169093714148" geometry="radius: 1.0;"></a-en-
tity>
    <a-entity mixin="sphere changeScale" position="-42.0 39.0 -249.0" geom-
etry="radius: 15.0;"></a-entity>
    <a-entity mixin="sphere changeScale" position="-44.33475620401434
33.78299438604239 -257.2055935733956" geometry="radius: 1.0;"></a-entity>
    <a-entity mixin="sphere changeScale" position="-32.076405842555005
39.381326438174156 -247.82659932424045" geometry="radius: 1.0;"></a-en-
tity>
    <a-entity mixin="sphere changeScale" position="-47.998682022985804
46.13340730395855 -245.37642245523327" geometry="radius: 1.0;"></a-en-
tity>
    <a-entity mixin="sphere changeScale" position="-43.00033156812354
35.251676611203536 -258.216800341052" geometry="radius: 1.0;"></a-entity>
    <a-entity mixin="sphere changeScale" position="-46.0 73.0 -282.0" geome-
try="radius: 1.0;"></a-entity>
    ...
    <a-text position="-0.12924592889244302 28.685206478968603 -
264.93061971943484" value="ведет кафедра" color="black" font="custom-
msdf.json" negate="false"></a-text>

```



```

    <a-text    position="5.5089732666646025    33.67653856070696    -
269.50354287178135" value="ведет кафедра" color="black" font="custom-
msdf.json" negate="false"></a-text>

    <a-text    position="3.409911361395242    29.29449271388946    -
272.07456518803593" value="ведет кафедра" color="black" font="custom-
msdf.json" negate="false"></a-text>

    <a-text    position="-22.988548498363066    66.3345405643869    -
255.17761288410827" value="порядок" color="black" font="custom-msdf.json"
negate="false"></a-text>

    <a-text    position="-22.19205286078726    72.49002965922602    -
262.12630798492944" value="порядок" color="black" font="custom-msdf.json"
negate="false"></a-text>

    <a-text    position="-18.652895570499574    73.09931589414688    -
269.2702534535306" value="порядок" color="black" font="custom-msdf.json"
negate="false"></a-text>

    <a-text    position="3.0357612394970097    29.076377728859363    -
262.32609632108574" value="ведет кафедра" color="black" font="custom-
msdf.json" negate="false"></a-text>

    <a-text    position="5.4113494166164795    30.523707458737924    -
271.90260335086083" value="ведет кафедра" color="black" font="custom-
msdf.json" negate="false"></a-text>

    </a-entity>

    <a-entity directed="false" line="start: 239 911 476; end: -258 668 436;
color: red" name="e_graduation_sem8_dep">

    </a-entity>
</a-entity>

```