Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования

«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

(ФГБОУ ВО «КубГУ»)

Факультет компьютерных технологий и прикладной математики

Кафедра информационных технологий

ОТЧЕТ

О выполнении лабораторной работы № 7

По дисциплине «Администрирование в Linux»

Выполнил: ст. гр. 4ИТ/2

Хижний Евгений

Проверил: преподаватель КИТ Прутский А.С.

Краснодар

2021

Лабораторная работа №7.

Цель: реализовать скрипт для автоматической генерации больших массивов данных для машинного обучения. На вход подается ссылка на материал в сети интернет, они скачиваются, верифицируются, при необходимости конвертируются и сохраняются на виртуальной машине.

Выполнение работы:

Для нашего датасета необходимы данные только в формате “csv”, формат наименования: “dataset\_[i].csv”, где “i” номер датасета. Само значение будет храниться в файле var.txt. Скачивание файла будет происходить при помощи команды wget, все проверки, связанные с загрузкой файлов выполняются этой командой.

Код:

#!/bin/bash

# read $url

echo "Введите url:"

read url # получение url

if [ -e var.txt ] # проверяем существование файла с переменной индекса датасетов

then # если существует, то записываем в переменную

read -r var < var.txt

else # если не существует, то задаем переменную со значением 1

var=$(( 1 ))

fi

echo ${url: -4}

if [ ${url: -4} == '.csv' ]

then

wget -O dataset\_"$var".csv "$url"# скачиваем файл и называем dataset\_[i].csv

var2=$(( $var + 1 )) # увеличиваем переменную индекса и записываем в файл

echo $var2 > var.txt

else

echo "Файл не в формате csv!"

fi