Problem St:

Write a microservice in python that takes a URL of a static website as an input and identifies all the unique words and how many times they occur on a web page. The output is a list of words in JSON format along with the frequency of occurrence (number of times the word is repeated).

Solution:

from flask import Flask, request, jsonify

from bs4 import BeautifulSoup

import requests

import json

app = Flask(\_\_name\_\_)

@app.route('/word\_frequency', methods=['GET'])

def word\_frequency():

# Get the URL from the request

url = request.args.get('url')

# Fetch the webpage content

page = requests.get(url)

# Parse the HTML content using BeautifulSoup

soup = BeautifulSoup(page.content, 'html.parser')

# Extract the text content from the page

text = soup.get\_text()

# Split the text into individual words

words = text.split()

# Create a dictionary to store the word frequency counts

word\_counts = {}

# Iterate over each word and update the dictionary with the frequency count

for word in words:

if word in word\_counts:

word\_counts[word] += 1

else:

word\_counts[word] = 1

# Convert the dictionary to a JSON object and return it

return jsonify(word\_counts)

if \_\_name\_\_ == '\_\_main\_\_':

app.run(debug=True)