A Project Description

On

**HITS Algorithm**

**RAHUL JAUHARI 2020A7PS0106H**

**SANCHIT GUPTA 2020A7PS2069H**

**ABHIRAJ KHARE 2020A7PS0161H**

Under the supervision of

**FACULTY: Dr. N L Bhanumurthy**

**SUBMITTED AS AN EVALUATION COMPONENT OF**

**Course: Information Retrieval CS F469**

A picture containing text, compass, clipart, cup

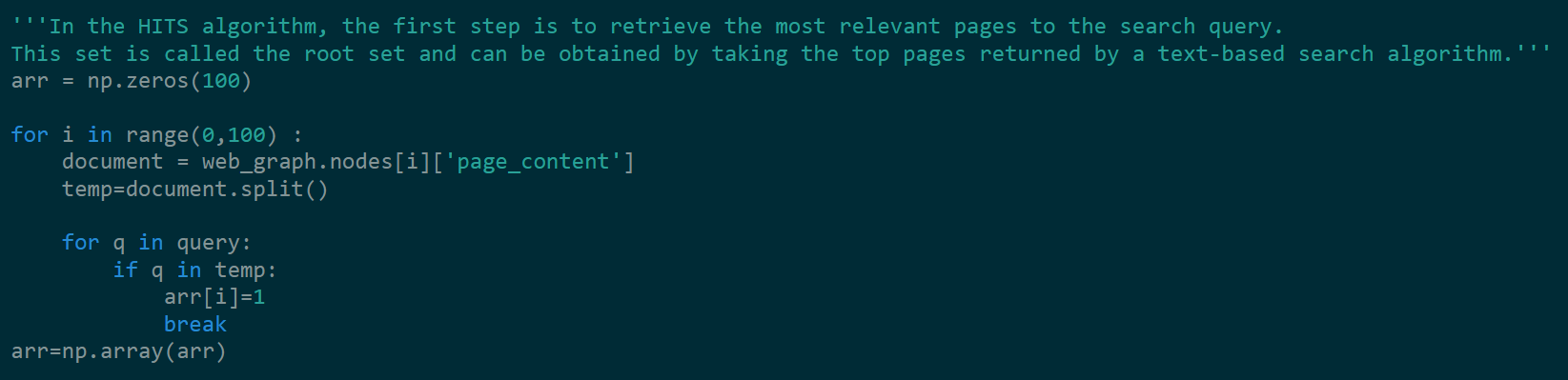
Description automatically generated

**BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE PILANI**

**HYDERABAD CAMPUS**

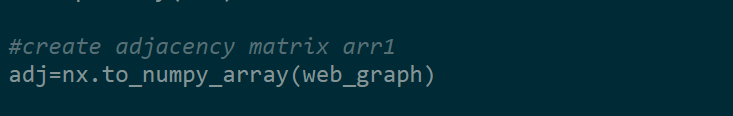
**Obtaining the root set for the query:**

A root set is obtained by taking the top pages returned by a text-based search algorithm.

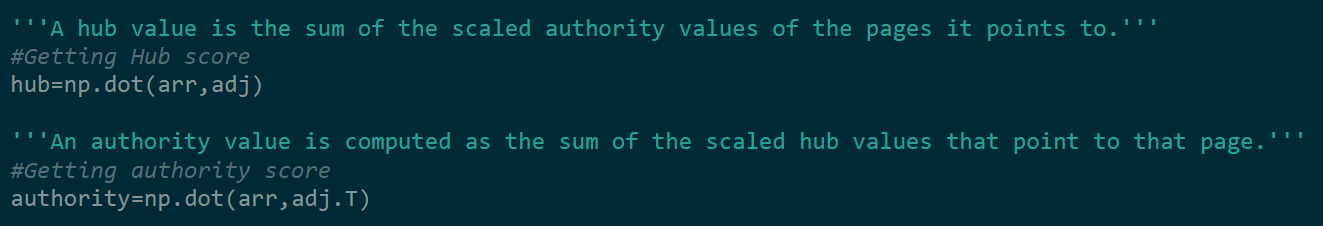


**Creating the Adjacency matrix:**

The elements of the matrix indicate whether pairs of vertices are adjacent or not in the graph.

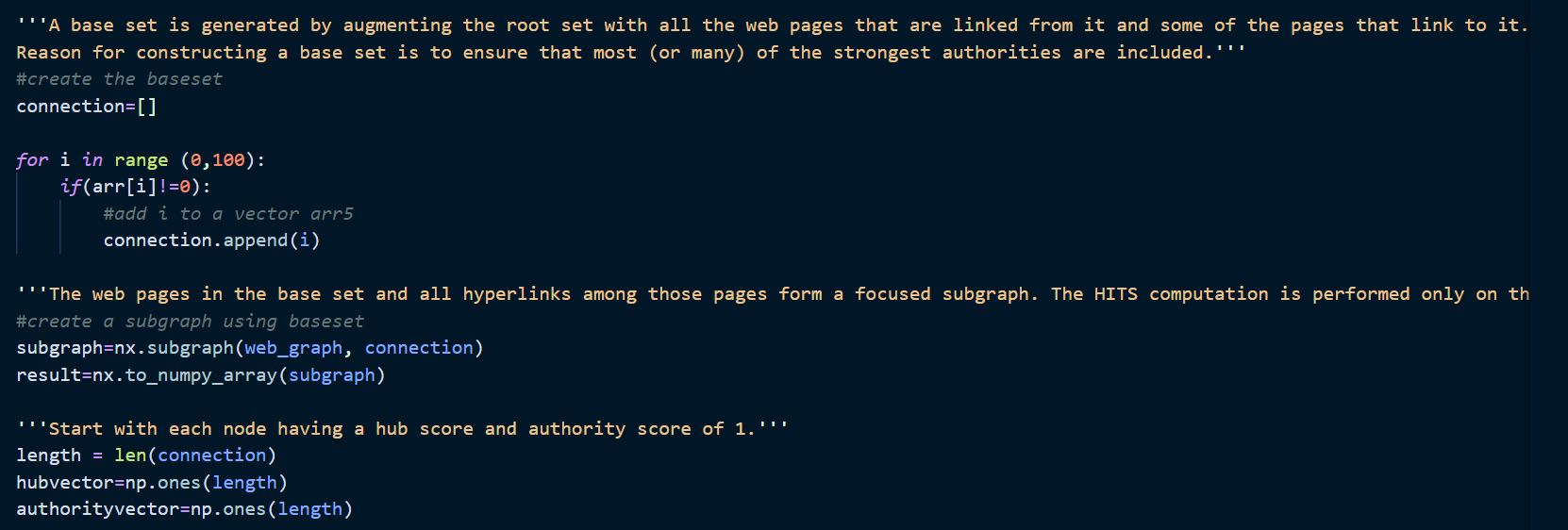


**Calculating Hub and Authority score:** A hub value is the sum of scaled authority values of the pages it points to and the authority value is computed as the sum of the scaled hub values that point to that page.

****

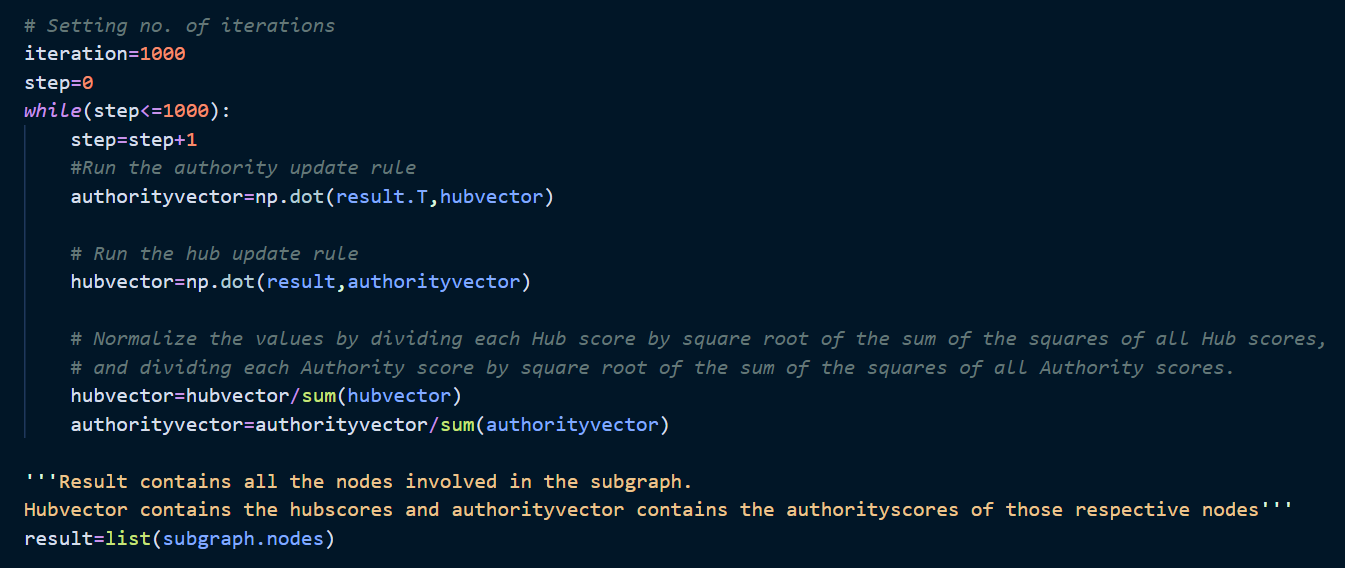
**Creating the base set:**

A base set is generated by augmenting the root set with all the web pages that are linked from it and some of the pages that are linked to it.

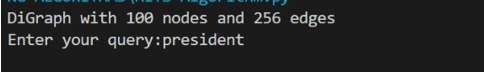


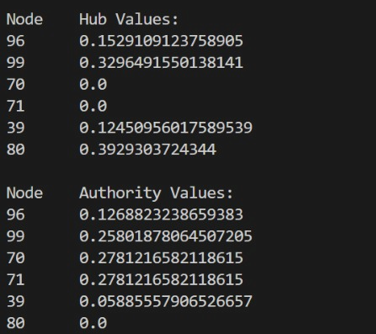
**Number of iterations for calculating hubs and authority scores:**

The process of calculating hub score from authority values and authority values from hub, is repeated multiple times till we obtain some constant value of hub and authority values.

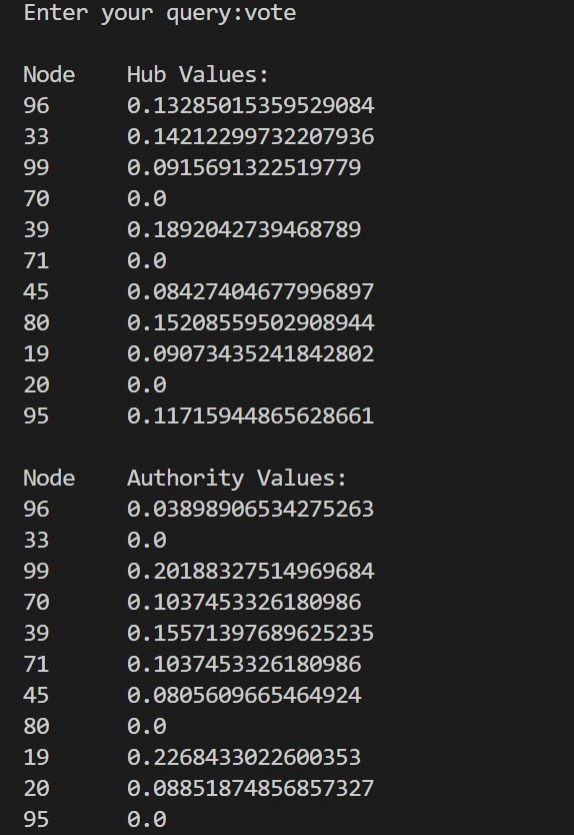


**Test Case 1:**

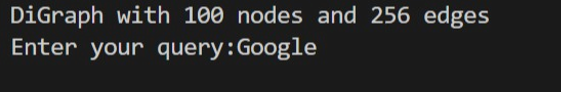
****

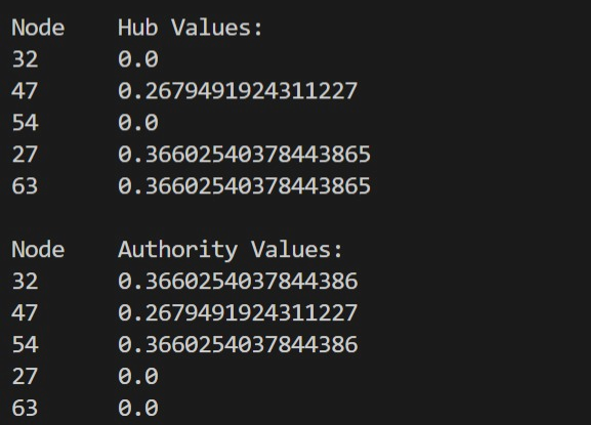


**Test Case 2:**



**Test Case 3:**

****

****