



PROJECT REPORT

TOPIC OF PROJECT

“MANTRA & GITA MINER”

Submitted by: Jayesh Sukdeo Patil

(MSc Computer Science Part-1 (SEM-2))

Academic Year: 2021-2022

Project Guide: Dr. Trupti Wani

University of Mumbai



SIES (Nerul) College of Arts, Science and Commerce

NAAC Re-Accredited 'A' Grade

Sri Chandrasekarendra Saraswathy Vidyapuram,

Plot 1-C, Sector V,

Nerul, Navi Mumbai-400 706.

Certificate

This is to certify that the Project entitled **“MANTRA & GITA MINER”** successfully completed by **JAYESH SUKDEO PATIL** of Part-1(Sem-2) Masters in Science (Computer Science) as per the requirement of University of Mumbai in part fulfillment for the completion of PG Degree of Master's of Science (Computer Science). It is also to certify that this is the original work of the candidate done during the academic year 2021-2022.

Roll No: 24

Date of Submission: 27-04-2022

Subject: Web Mining

Prof.Trupti Wani

Date: 27-04-2022

(Project Guide)

INDEX

Sr.No.	Contents
1	Abstract
2	Introduction
3	Objectives
4	Literature Review
5	Methodology
6	Results
7	Limitations
8	Future scope
9	Conclusion
10	References

MANTRA & GITA MINER

JAYESH SUKDEO PATIL

SIES (Nerul) Navi Mumbai, Maharashtra

Email: jpatil045@gmail.com

ABSTRACT:

This Research paper focuses on the use of web mining techniques to create a Mining Model for extracting Hindu Gods mantras and Srimad Bhagavat Gita Slokas for user in an Interactive way. Using Web mining libraries this mining model can extract Mantras from website and stored it as text file and display according to user wish threw tkinter interface.

KEYWORDS:

Mantras, Srimad Bhagavat Gita, Beautiful Soup, Web mining, Tkinter

INTRODUCTION:

Web mining can widely be viewed as the application of adapted data mining methods to the web, whereas data mining is represented as the application of the algorithm to find patterns on mostly structured data fixed into a knowledge discovery process.

Web mining has a distinctive property to support a collection of multiple data types. The web has several aspects that yield multiple approaches for the mining process, such as web pages including text, web pages are connected via hyperlinks, and user activity can be monitored via web server logs.

- **The Web seems to be too large for efficient data warehousing and data mining** – The size of the Web is in the order of hundreds of terabytes and is still growing rapidly. Some organizations and societies place several public-

accessible data on the Web. It is applicable to set up a data warehouse to replicate, save, or integrate some data on the Web.

- **The complexity of Web pages is far greater than that of any traditional text document collection** – Web pages lack a unifying structure. They contain far more authoring style and content variations than any set of books or other traditional text-based documents.
- The Web is treated as a huge digital library; but, the tremendous number of records in this library are not arranged according to any specific sorted order. There is no index by the element, nor by title, author, cover page, table of contents, etc. It can be very challenging to search for the information you desire in such a library.
- **The Web is a highly dynamic information source** – It does not only do the Web grow rapidly, but its information is also constantly updated. News, stock markets, weather, sports, shopping, company advertisements, and numerous other Web pages are updated regularly on the Web. Linkage information and access records are also updated frequently.
- **The Web serves a broad diversity of user communities** – The Internet currently connects more than 100 million workstations, and its user community is still rapidly expanding. Users can have multiple backgrounds, interests, and usage goals.

OBJECTIVES:

1. To build a Mining model to mine Slokas and Mantras according to users need.
2. To provide spiritual solutions for material problems through Srimad Gita slokas through an interactive application.

LITERATURE REVIEW:

Data mining is database knowledge discovery (KDD) and is primarily defined as a discovery method. The underlying design from the vast amount of data stored in a huge database. Huge data development sums up World WideWeb (WWW) where huge web pages created as often as

possible are useful for research and analysis Information is a fair test. The WWW consists of billions of interconnected web pages distributed by so many people. A web page is a report that can be reasonably displayed on the WWW using a web program [1]. WCM Extract important information and knowledge from the content of your website. Web pages include text, images, Recording, sound, or other structured content such as a table whose purpose is to pass this content to a designated customer.

WSM is a technique for retrieving and extracting basic information from weblogs, and mining techniques work at two levels. It depends on the type of baseline data used in the analysis process: hyperlinks and web page level [1]. Table 1 shows that different types of web mining classes are stored. The nature of what was mined and distant knowledge. WCM focuses on website content. B. Website text, Images and other connected media, data mining methods applied to this class include web page preprocessing. Text content, and other content such as images, depends on the needs of your application. WCM methods can be used in many web applications that aim to find web objects with basic attributes. Examples such as comparison points (topics) and collections of web pages that explain comparison web images Includes specific objects such as logos, watermarks, faces, etc.

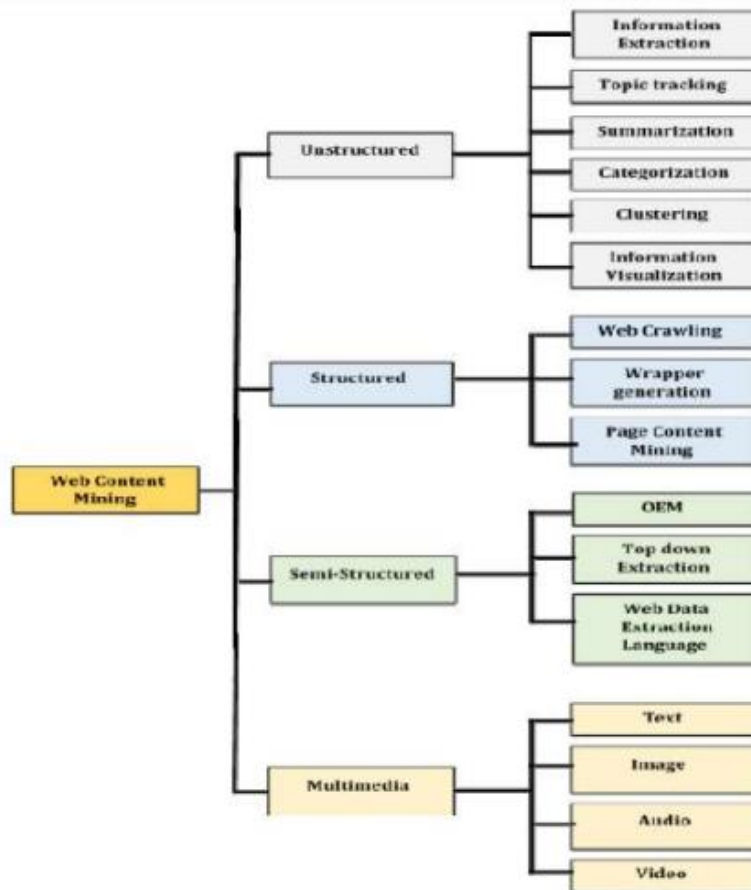


Figure 1: Web Content Mining Techniques and Applications

METHODOLOGY:

Sources used for extraction:

For Extracting Spiritual Mantras:

<https://www.drikpanchang.com/vedic-mantra/mantra-collection.html>

For Extracting Srimad Bhagavat Gita Slokas:

<https://vedabase.io/en/library/bg/>

The following steps were followed for building the Miner Model:

1.Web Scrapping:

First the Miner Model is built to gather the information from website using BeautifulSoup library of python

2. Data Storing

After scrapping data from website, it is stored as text file for further operations.

3. Data Display using Tkinter

Using Tkinter the categories are displayed using button which upon clicking will display the extracted data in text area.

Libraries used:

Beautiful Soup:

Beautiful Soup is a library that makes it easy to scrape information from web pages. It sits atop an HTML or XML parser, providing Pythonic idioms for iterating, searching, and modifying the parse tree. BeautifulSoup helps you pull particular content from a webpage, remove the HTML markup, and save the information. It is a tool for web scraping that helps you clean up and parse the documents you have pulled down from the web.

Steps to use:

Installation of Python BeautifulSoup

Accessing of the HTML Through a Webpage

Parsing of the Content HTML

Code:

```
from bs4 import BeautifulSoup
```

Requests:

The [requests](#) library is the de facto standard for making HTTP requests in Python. It abstracts the complexities of making requests behind a beautiful, simple API so that you can focus on interacting with services and consuming data in your application. Requests is ready for the demands of building robust and reliable HTTP-speaking applications, for the needs of today.

- Keep-Alive & Connection Pooling
- International Domains and URLs
- Sessions with Cookie Persistence
- Browser-style TLS/SSL Verification
- Basic & Digest Authentication
- Familiar dict-like Cookies
- Automatic Content Decompression and Decoding
- Multi-part File Uploads
- SOCKS Proxy Support
- Connection Timeouts
- Streaming Downloads
- Automatic honoring of .netrc
- Chunked HTTP Requests

Code:

```
import requests
```

Sys:

The **sys module** in Python provides various functions and variables that are used to manipulate different parts of the Python runtime environment. It allows operating on the interpreter as it provides access to the variables and functions that interact strongly with the interpreter.

Code:

```
import sys
```

Python OS:

Python OS module allows us to use the operating system dependent functionalities and to interact with the underlying operating system in several different ways. For example, we can work with files, change the environment variables, and we can move files around, etc. This is as same as overriding all the os built-in functionalities in a module and using them in a file I/O and system handling.

Code:

```
import os
```

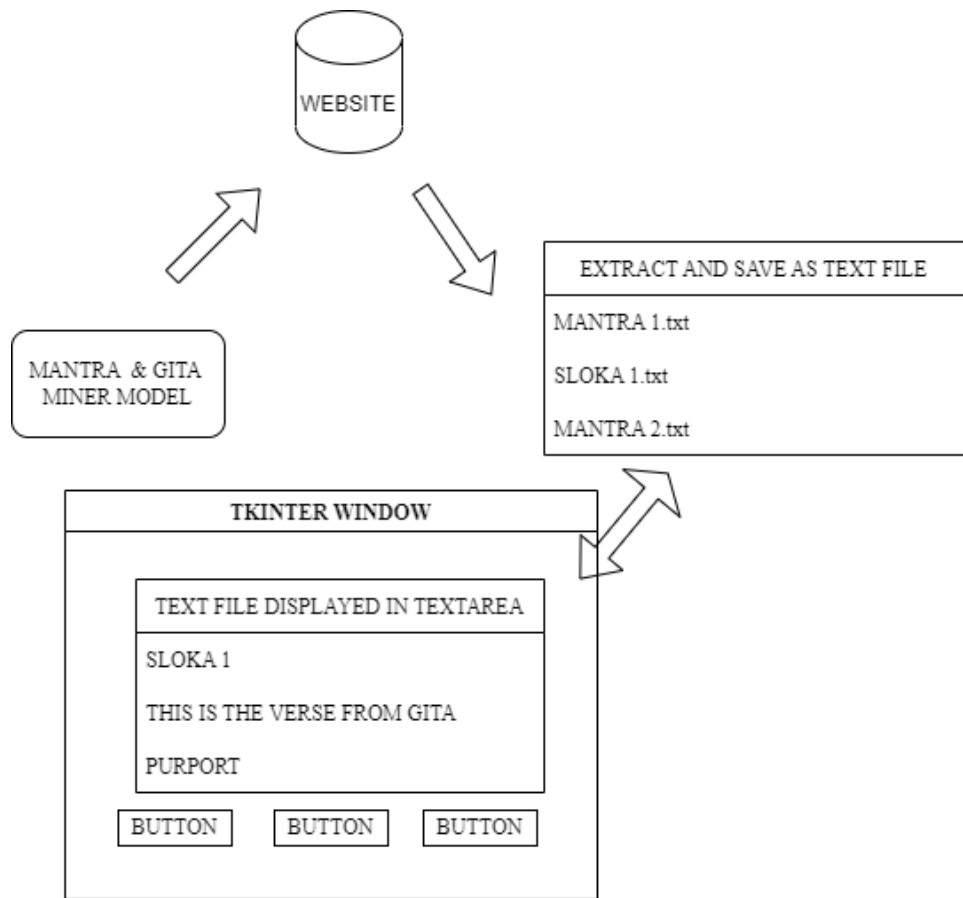
Tkinter:

Tkinter is the most commonly used library for developing GUI (Graphical User Interface) in Python. It is a standard Python interface to the Tk GUI toolkit shipped with Python. As Tk and Tkinter are available on most of the Unix platforms as well as on the Windows system, developing GUI applications with Tkinter becomes the fastest and easiest.

Code:

```
from tkinter import *
```

Flow of Model:



RESULTS:

MANTRA MINER

[illegible]

Goddess Saraswati

U

ॐ नमो भगवते वासुदेवाय ॥

Goddess Saraswati

GITA MINER

 GITA MINER

Spiritual Solutions for your Material Problems

Anger

Fear

Lust

Confusion

Envy

Greed

Peace

Demotivated

Laziness

Depression

Loneliness

Uncontrolled Mind

 GITA MINER

Spiritual Solutions for your Material Problems

BG Chapter.verse:
Bg. 16.1-3

Sanskrit: श्रीभगवानुवाच अभयं सत्त्वसंशुद्धिर्ज्ञानयोगव्यवस्थितिः । दानं दमश्च यज्ञश्च स्वाध्यायस्तप आर्जवम् ॥ १ ॥ अहिंसा सत्यमक्रोधस्त्यागः शान्तिरपेक्षन्म् । दया भूतेष्वलोलुपत्वं मार्दवं ह्रीरचापलम् ॥ २ ॥ तजः क्षमा धृतिः शौचमद्रोहो नातिमानिता । भवन्ति सम्पदं देवीमभेजातस्य भारत ॥ ३ ॥

English: śrī-bhagavān uvāca abhayaṁ sattva-saṁśuddhir jñāna-yoga-vyavasthitiḥ dānaṁ damaś ca yajñaś casvādhyāyas tapa ārjavam

Translation: The Supreme Personality of Godhead said: Fearlessness; purification of one's existence; cultivation of spiritual knowledge; charity; self-control; performance of sacrifice; study of the Vedas; austerity; simplicity; nonviolence; truthfulness; freedom from anger; renunciation; tranquillity; aversion to faultfinding; compassion for all living entities; freedom from covetousness; gentleness; modesty; steady determination; vigor; forgiveness; fortitude; cleanliness; and freedom from envy and from the passion for honor – these transcendental qualities, O son of Bharata, belong to godly men endowed with divine nature.

Purport: In the beginning of the Fifteenth Chapter, the banyan tree of this material world was explained. The extra roots coming out of it were compared to the activities of the living entities, some auspicious, some inauspicious. In the Ninth Chapter, also, the devas, or godly, and the asuras, the ungodly, or demons, were explained. Now, according to Vedic

Anger

Fear

Lust

Confusion

Envy

Greed

Peace

Demotivated

Laziness

Depression

Loneliness

Uncontrolled Mind

LIMITATIONS:

- The respective miner model is not in a user-friendly mode as the project misses the API through which user can interact and get needed data.
- The Miner model does not have a database linked through which user can stored and retrieve data and also keep as backup of extracted data.

FUTURE SCOPE:

This model can be used to make a full-fledged android application where user can get needful information in a more interactive way. Further this model can be used to make a repository of spiritual texts which are not digitalized from where user can get needed text easily without searching among huge data.

CONCLUSION:

As there was not a Miner model to extract Mantras and slokas this model can be used to make a user-friendly application. This model helps user to get needed content with an Tkinter interface. With some limitations this model can provide a wide variety of applications in domain specific knowledge.

Acknowledgement

I would like to express my deep gratitude to Dr. Trupti Wani, my research supervisor, for her patient guidance, enthusiastic encouragement and useful critiques of this research work. A special thanks to my friends as well who constantly supported me and acted as peer reviewers for this research. I would also like to extend my thanks to the technicians of the laboratory department for their help in offering me the resources in running the program. Finally, I wish to thank my parents for their support and encouragement throughout my study.

REFERENCES:

Mrs. S. Kulandai Terasa - "*AN EMPIRICAL STUDY ON WEB MINING TECHNIQUES AND APPLICATIONS*", ISSN NO: 2347-3150, Mukta Shabd Journal, Volume IX, Issue VII, JULY/2020

Bing Liu - "Web Data Mining -Exploring Hyperlinks, Contents, and Usage Data" Second Edition ISBN 978-3-642-19459-7 Department of Computer Science

Rao, T. V., & Hiranmayi, D. (2017). Methodologies and techniques of web usage mining. *Advances in Data Mining and Database Management*, 275–296. <https://doi.org/10.4018/978-1-5225-0613-3.ch011>

