

A SYNOPSIS ON

---

## **Redzon: A Subreddit Analysis and Community Exploration Tool**

---

Submitted in partial fulfilment of the requirement for  
the award of the degree of

**MASTER OF COMPUTER APPLICATIONS**

**(OMC306 Mini Project, Third Semester MCA)**

Submitted by:

**Name:** Deepankar Sharma      **Student ID:** 233512013

Under the Guidance of  
**Dr. Jayashree Kulkarni**  
**Professor, GEU**



**Department of Computer Applications**  
**Graphic Era (Deemed to be University)**  
**Dehradun, Uttarakhand**

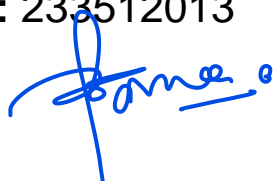
## CANDIDATE'S DECLARATION

I/we hereby certify that the work which is being presented in the Synopsis entitled "**Redzon: A Subreddit Analysis and Community Exploration Tool**" in partial fulfillment of the requirements for the award of the Degree of Master of Computer Applications (**OMC306 Mini Project, Third Semester MCA**) in the Department of Computer Applications of the Graphic Era (Deemed to be University), Dehradun shall be carried out by the undersigned under the supervision of **Guide Name**, Designation, Department of Computer Applications, Graphic Era (Deemed to be University), Dehradun.

**Name:** Deepankar Sharma

**Student ID:** 233512013

**Signature:**



The above-mentioned student shall be working under the supervision of the undersigned on the "**Redzon: A Subreddit Analysis and Community Exploration Tool**."

Signature

Signature

Supervisor

Head of the Department

### Internal Evaluation (By DPRC Committee)

Status of the Synopsis: Accepted / Rejected

Any Comments:

Name of the Committee Members:

1.

2.

Signature with Date

# Table of Contents

1. Introduction and Problem Statement
  2. Background/ Literature Survey
  3. Objectives
  4. Hardware and Software Requirements
  5. Possible Approach/ Algorithms
  6. Applications and Real-World Impact
  7. References
- 

## Chapter 1: Introduction and Problem Statement

### 1.1 Introduction

"Redzon" is an innovative, Streamlit-powered application designed to leverage the vast troves of data on Reddit for actionable insights. By tapping into Reddit's subreddit ecosystem via Python's PRAW (Python Reddit API Wrapper) library, this tool empowers users to analyze community trends, explore similar subreddits, and extract critical information to optimize strategies for online engagement, SEO, and content marketing.

Reddit, often dubbed "the front page of the internet" due to its dynamic aggregation of trending content and diverse discussions, boasts thousands of active subreddits covering virtually every topic imaginable. This nickname underscores its role as a central hub for real-time news, niche communities, and viral phenomena. Businesses, marketers, and researchers can benefit immensely from harnessing this data to identify trends, understand audience behavior, and target niche communities. However, existing tools often fail to provide user-friendly interfaces and holistic analyses of subreddit data.

Redzon bridges this gap by offering:

- **Subscriber Growth Analysis:** Understand how a subreddit's audience evolves over time.
- **Activity Metrics:** Monitor engagement trends, such as post frequencies and comment activity.
- **Topic Discovery and Community Exploration:** Discover new subreddits and analyze them based on similar topics.
- **Visualization and Reports:** Generate visual summaries of subreddit performance for better decision-making.

### 1.2 Problem Statement

As social media platforms like Reddit become critical for content discovery, brand engagement, and SEO, there is a growing demand for tools that can:

- Simplify subreddit data analysis.
- Present insights visually and intuitively for actionable use.
- Help businesses and creators identify communities to optimize their reach organically.
- Provide trend analyses to predict future activity and identify "rising" subreddits.

Existing tools are either too complex for non-technical users or lack comprehensive analytics, making it difficult for users to harness Reddit's potential for targeted content

and community engagement. "Redzon" addresses these gaps by providing a scalable, interactive, and easy-to-use solution.

---

## Chapter 2: Background/ Literature Survey

The exponential growth of Reddit as a hub for community-driven discussions has led researchers and marketers to explore its potential for audience insights. Tools like PRAW have simplified Reddit API integration, enabling developers to extract data programmatically. However, while raw data access is straightforward, interpreting and visualizing this data remains a challenge.

Marketing professionals increasingly rely on Reddit to drive organic traffic, identify influencers, and test audience responses to content. Reddit's unique structure—a combination of posts, comments, and community-specific rules—makes it an ideal platform for niche audience targeting. Yet, without a proper analysis tool, users miss opportunities to:

- Discover high-performing communities relevant to their goals.
- Identify trends and optimize content strategies for SEO.
- Leverage data for community building and campaign planning.

Redzon integrates these requirements into one cohesive platform, ensuring seamless access to insights that were previously difficult to attain.

---

## Chapter 3: Objectives

The objectives of the proposed work are as follows:

1. **Develop an Interactive Subreddit Analysis Dashboard:** Create a user-friendly interface using Streamlit for engaging with subreddit data.
  2. **Enable Real-Time Data Fetching:** Implement functionality using PRAW to provide up-to-the-minute subreddit insights. This dynamic feature ensures that metrics like activity trends and subscriber counts are always current, avoiding the pitfalls of outdated batch data processing.
  3. **Provide Growth Trend Analysis:** Analyze and visualize the growth trends of subscribers and activity metrics over time.
  4. **Implement Similarity-Based Recommendations:** Develop algorithms to suggest related subreddits based on user interests and community overlap.
  5. **Visualize Key Subreddit Data:** Use charts and tables to present data in an intuitive manner for easier understanding.
  6. **Optimize Subreddit Data for SEO Applications:** Identify trending topics and high-engagement posts to assist users in content creation and SEO strategies.
- 

## Chapter 4: Hardware and Software Requirements

### 4.1 Hardware Requirements

Sl. No	Name of the Hardware	Specification
1	Development Machine	Intel i5 Processor, 8 GB RAM, 256 GB SSD
2	Backup Storage	External HDD with 1 TB capacity

## 4.2 Software Requirements

Sl. No	Name of the Software	Specification
1	Operating System	Ubuntu 20.04 LTS
2	Programming Language	Python 3.8 or later
3	Streamlit	Streamlit 1.3.0 or later
4	Libraries	PRAW, Matplotlib, Pandas, Plotly
5	External APIs	Reddit API, PRAW

# Chapter 5: Possible Approach/ Algorithms

## 5.1 Data Collection and Preprocessing

- **API Integration:** Utilize PRAW to fetch subreddit data including subscriber counts, posts, comments, and activity trends.
- **Data Cleaning:** Handle missing or inconsistent data entries to ensure accurate analysis.
- **Feature Engineering:** Derive meaningful features such as engagement ratios, trending keywords, and growth rates.

## 5.2 Visualization Techniques

- **Time Series Analysis:** Use line charts to illustrate subscriber and activity growth trends.
- **Bar and Pie Charts:** Represent subreddit rankings and activity metrics visually.
- **Interactive Dashboards:** Leverage Streamlit's components to create a dynamic user experience for data exploration.

## 5.3 Topic and Similarity Analysis

- **Keyword Extraction:** Implement frequency-based techniques to identify popular topics and trends.
- **Community Clustering:** Utilize algorithms to identify similar subreddits based on content themes and user engagement overlap.

# Chapter 6: Applications and Real-World Impact

1. **SEO and Organic Traffic Optimization:** Businesses can utilize the tool to pinpoint trending subreddits and relevant topics, tailoring content for specific audience segments.

2. **Content Marketing:** Marketers can target high-engagement communities identified through the analysis, enhancing the return on investment (ROI) of their campaigns.
  3. **Community Management:** Subreddit moderators can monitor activity trends and engagement metrics to foster a more vibrant community.
  4. **Trend Analysis:** Researchers and analysts can leverage the data to predict future activities, identify emerging subreddits, and understand shifts in community interests.
- 

## Chapter 7: References

1. **Reddit API Documentation:** <https://www.reddit.com/dev/api/>
  2. **PRAW Documentation:** <https://praw.readthedocs.io/en/stable/>
  3. **Streamlit Documentation:** <https://docs.streamlit.io/>
  4. **Matplotlib Library:** <https://matplotlib.org/>
  5. **Pandas Library:** <https://pandas.pydata.org/>
-