

PHASE 1 : DESCRIPTION OF EXISTING SYSTEM

Project Link : <https://github.com/anubhavs11/Fake-Product-Review-Monitoring>

TASK 1 : IDENTIFYING ENTITIES AND CONTEXT DIAGRAM

In recent times there is a rapid increase in the online purchases in various e-commerce sites. User is estimating the quality of the product based on the product ratings and reviews. Simultaneously, there might be a high chance of coming across fake reviews. This impacts negatively on the user trusting the product and the website. Similarly, the negative fake reviews also affects the merchant business. Hence, author is aiming to create a system which detects and monitor these fake reviews.

EXTERNAL ENTITIES :

Users / Bots : The persons / systems who are providing reviews on the e-commerce website.

Admin : To monitor the fake reviews detected by the system.

E-Commerce Database : Used to store information about the existing reviews and submitted reviews by the user and also user information.

INPUTS:

Product Reviews: Information or feedback from various users on e-commerce websites either existing or new reviews submitted by the end-users.

User Data: Information about the persons / systems (e.g., IP addresses).

OUTPUTS:

Identification of Review Patterns : Based on the processes performed by the system, we get the fake reviews patterns in the data.

Reports: Summary of the Fake Reviews detected.

RESPONSIBILITIES OF THE SYSTEM:

The main objective of the author's system is to detect the fraudulent reviews in the e-commerce websites using Natural Language processing techniques and machine learning models.

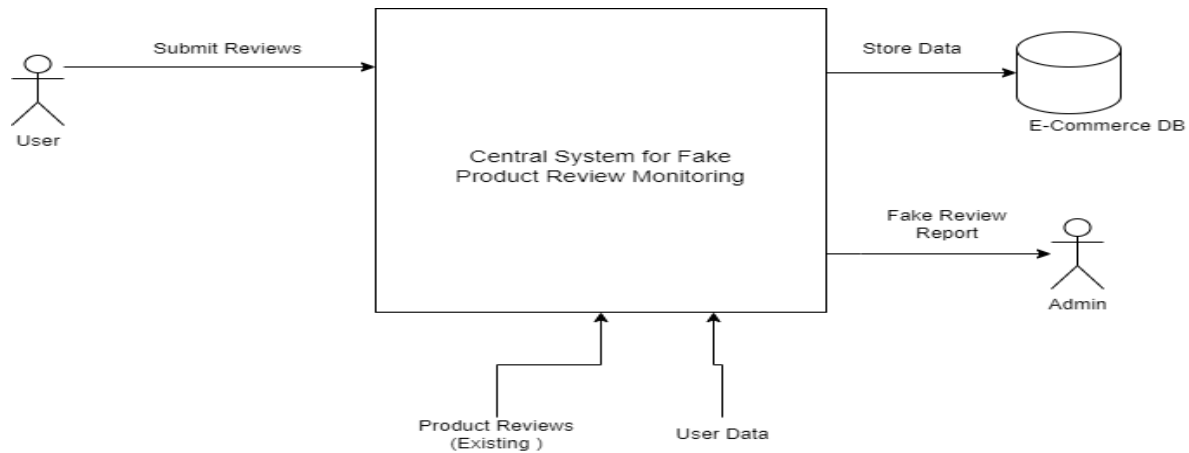
Processing Of Reviews : Using Natural Language processing (NLP) techniques, preparing data collected from websites for data analysis.

Perform Analysis with the Reviews submitted by the end-user : 3 set of analysis are made by the system which includes Sentimental analysis, which is made to assess the user's perspective on review of the product and capture the review patterns. Performing Content Similarity with the help of cosine similarity, detecting the duplicate checks and fake responses. Performing Latent Semantic Analysis (LSA) to eliminate the similar content reviews.

Identifying the Patterns : Based on factors like similar IP addresses, Flooding by same user, User's tone in multiple reviews. In our project, 9 factors are determined to identify the patterns of Fake Reviews

Providing Outputs : After making the required analysis by the system, the system provides the review patterns and fake reviews which helps in maintaining the trust of the review platform.

CONTEXT DIAGRAM OF SYSTEM: When the end-user provides reviews in the e-commerce website, the system processes the required analysis and provides information about the fakeness as the output. The admin, based on the output, monitors the review platform and takes required action.

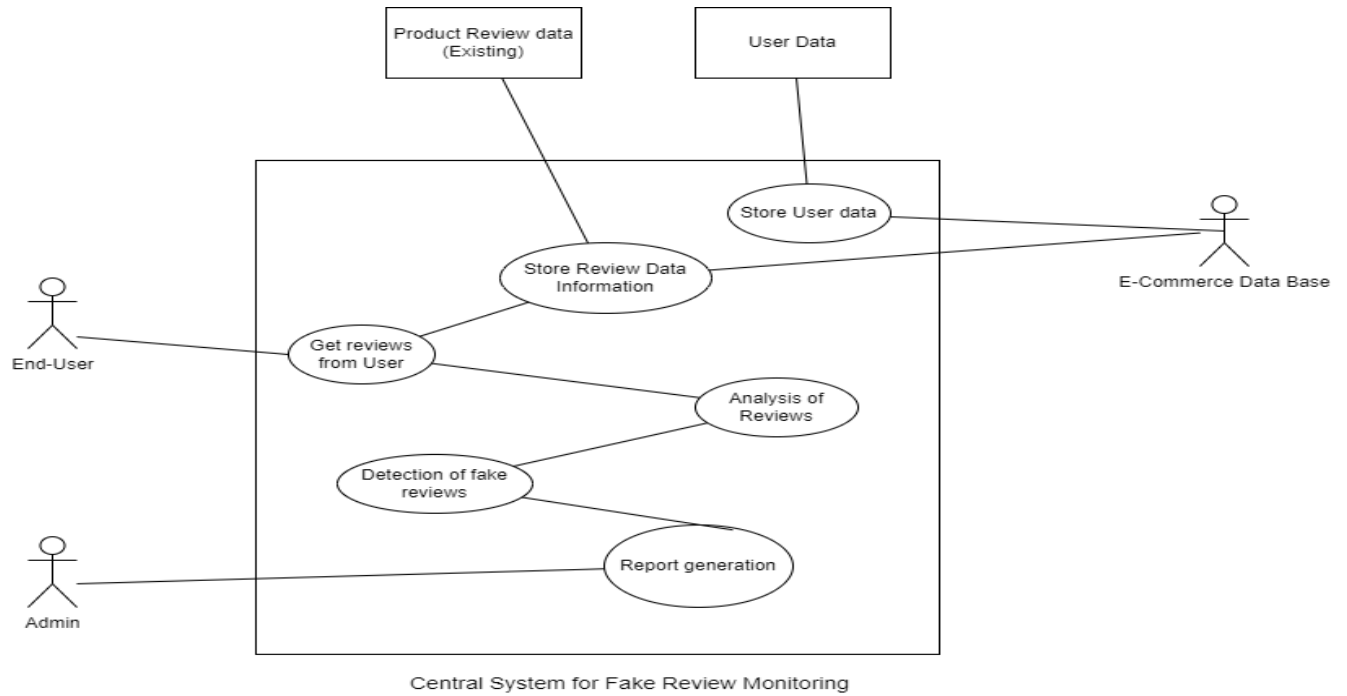


TABLES TO DESCRIBE ENTITIES:

Entities	Responsibilities	Input Data	Output Data
User	Submission of reviews for the product	Reviews	None
Admin	Monitors the Fake reviews submitted by the user.	Information about the Fake Review Patterns	Validation of Reviews
E-commerce Database	Stores information about the User, Review data	Product data and End-User data	Data
Central System	Analyzing the data and Detection of hoax reviews and a report about the review.	User data and Product Review Data	Information about Review , Notification to Admin for Monitoring

TASK 2:BUILDING OF UML USE-CASE AND COMPONENT DIAGRAM BASED ON ARCHITECTURE RECOVERY

UML- USE-CASE DIAGRAM



UML COMPONENT DIAGRAM

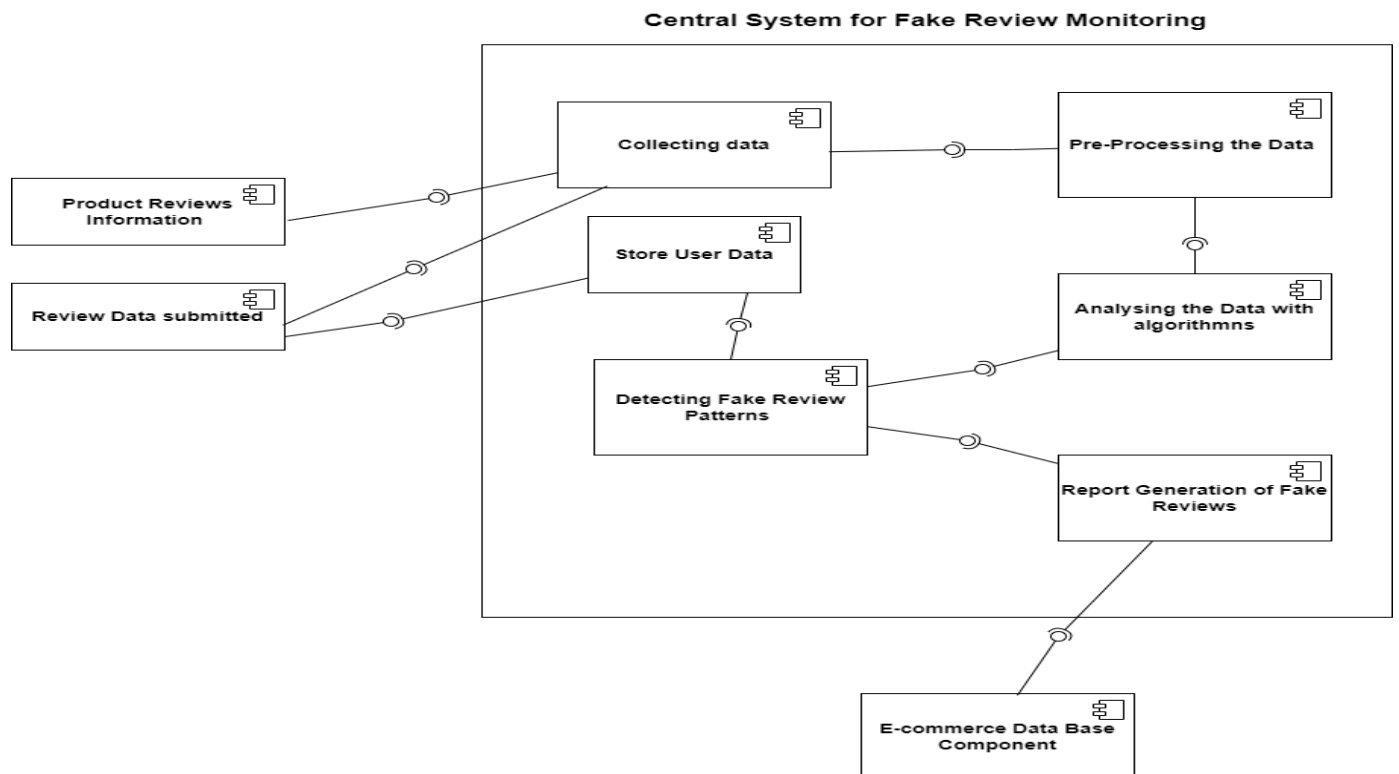


TABLE ABOUT THE COMPONENTS.

Component	Description
Review Data submitted	Product feedback information provided by the user
Product Reviews Information	Existing information about the product in the e-commerce website
Collecting data	Data receiving from the end-user
Store User Data	Safekeeping the end user's IP-address,tone of review,context of the feedback of the product.
Pre-Processing the Data	Perform pre-compiling operations like Feature_Extraction,Tokenization on the input data
Analysing the Data with algorithmns	Performing of Sentimental,Content and Latent Semantic Analysis on pre-processed data
Detecting Fake Review Patterns	With help of user information and analysed data, identify the patterns to detect the false feedbacks received.
Report Generation of Fake Reviews	Validation of Review received from the user.
E-commerce Data Base Component	Stores details about the Reviews and also the User information

TASK 3: CHALLENGES FACED IN TASK-1 AND TASK-2

In Task 1, Identifying the responsibilities of the system and how they interact with the input and providing the output was the major challenge we faced.

- Another issue with identifying the key components and the systems processes like Sentimental analysis,LSA and their interactions with the external entities like users.
- As the context diagram focuses on overview representation of the system,the challenge was how to eliminate the technical aspects which dominate the context diagram.
- Maintaining Understandability of systems is another major task. Need to include all the required entities,interactions with the system to decrease the complexity in understanding the system and it will be easy to propose similar systems.

In Task 2: While building Use-case diagrams, The main challenge was abstracting the technical information and concentrating only on the required actions performed by the actors and interactions with and within the system.

- While building the component diagram,extracting the components from the code and understanding the system functionalities was a major task.
- In both diagrams, making the clear understandability of UML diagrams was little complicated.

LESSONS LEARNED :

- Documentation of UML diagrams needs to be precise,clear and should be flexible to use the existing diagrams to propose similar systems.
- Extracting the details for use-case and component diagram from recovery of code,requires good understanding of the flow of the process.
- The actors in the use-case diagram will NOT be represented in the Component diagram.
- Representing the code recovered in the UML-Component diagram makes it easier to understand the process involved in the system,interactions and relationships between the components.
- By profiling the UML diagram from code,it makes it easier for the developers in getting the missing requirements and implementing the functionalities.