

411 Project Proposal

Junyu Liu, Harry Yuan, Minjun Wen, Leon Heff, Meng-Hsin Cheng

Repository: <https://github.com/SorosWen/411project>

Idea 1:

The web page will allow users to input items that they want to purchase. Our site will then grab related product data from retailers (such as Walmart and Ebay) and e-commerce (such as Amazon) and list them side by side for users to compare. The user can also filter out undesired items by setting the price range or the estimated delivery time.

- a. **Database:** We will create a database for users to store their profile information and their recent searched keyword and items. We will use these data to analyze user behavior and improve our website.
- b. **Data Sets:**
 - i. Shop.com API / eBay Shopping API / Google Shopping Search API: These APIs allow us to view products that are for sale on the market. We will use one or more of these APIs to access the datasets of products that are for sale.
- c. **Third Party Authorization:**
 - i. OAuth: we will allow users to register to our web-application by signing in using Google through OAuth.
- d. **Decoupled Structure:** The front end will allow users to input the name of the product. The backend will then fetch all the information of the related products that are for sale using those APIs and then send it to the front end. The front end will then read those product information and display each product for users to compare. The filter will be done in the front end instead of the back end. However, we might change the way we implement the filter if we found it more efficient to do it the other way around.

Idea 2:

The webpage will retrieve data, tweets, and blogs from users' social media profile. Then it will conduct sentiment analysis using datasets such as Watson API or Google Cloud API. The analysis will give each post a positivity score. Once the analysis is done, the webpage will present a graph showing a timeline and the positivity score (positive, negative or neutral content) for each day.

- a. **Database:** Our database will store users' profile information and their previous sentiment analysis data for trending over time.
- b. **Data Sets**
 - i. Twitter API / Facebook Graph API: we will use one of these (or both) API to retrieve/access users' social media content.
 - ii. IBM Watson NLU API / Google Natural Language API: we will use datasets from one of these APIs to conduct sentiment analysis.
- c. **Third Party Authorization:** OAuth: for letting users log into Twitter, Facebook, or any of their social media accounts.

- d. **Decoupled Structure:** The front end will let users input data and will display the timeline graph containing the positivity score. The backend will retrieve user social media history, conduct sentiment analysis, and send the analyzed data to the front end.