Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Team 4 Bravo°Д°

Table of Contents

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Abstract

In recent years, with the upsurge of healthy life, more and more people begin to try to keep fit. Then there are all kinds of fitness blogs springing up. They want to promote their fitness products, such as protein and vitamin supplements, by introducing their sports experience and shaping plans. But many new entrants don't know much about this field, so they are easy to be misled by one family's opinions and buy products of poor quality or unsuitable for themselves.

This project is going to come out an automatic processing method to filter related products for a certain fitness demand, which provides purchase suggestions by integrating the evaluation and user experience of each platform. It aims to give the total impression of specific fitness products by extracting and summarizing the key words in the evaluation blog. The feedback information of user experience is gained by analyzing the emotion of product comments and the repurchase rate of products. Combining this information, it will finally give appropriate purchase suggestions to customers.

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Project Objective,Scope and Deliverables

Project Objective

This project is going to provide purchase suggestions about fitness products by integrating the evaluation and user experience of each platform.

Project Scope

The project scope will be focused on database design. Implement database structure that can show the real relationship between blogs' attitude, customers' feedback, fitness product price. As the rest of the project, there will be the suggestion keywords extraction, sentiment analysis based on text and price trend prediction.

Key Deliverables

1. A database that stores blog URL, some product feedback and price history from different planforms.
2. An emotion analysis model based on text and a price prediction ML model will be implemented
3. A suggestion system that gives proper purchase advice based on input will be achieved.

Comparative analysis of existing solutions

1. Precision: As for now, if a person wants to know the situation about one fitness product, he has to check the sale planform feedback and some evaluation blogs. And sometimes one family words will mislead people. This project combines different planforms information to give a purchase suggestion, which will be more preciser
2. Analysis angle: Most of suggestion systems are implemented by analyzing the similarities between customers' behavior. This may gather the same type customers and give them advice. But this system may not be able to consider the evaluation about the product. This system is different in combining evaluation blogs with customers' feedback, which is a new analysis angle to solve suggestion problem.
3. Convenience: If a customer wants to check status about a product as much as possible, he need to spend lots of time to read the comments and evaluation articles. But this system will save lots of time by giving a glimpse about a fitness product by combining many planforms' information.

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Information helping validate the possibility of the application

1. Natural Language Processing:
   Having defined number of tags of emotion: Positive, Negative or Neutral. We can use NLP, text classifiers which can analyze text and can allocate set of pre - defined tags or division based on its content automatically.

2. Decision tree:
   We can use a decision tree to correlate blogs to certain fitness goal. Through this all possible outcomes would be taken under consideration to help decision maker come up with final decision.

Team Information

SAURIN SHAH

Saurin is an Engineering Management Graduate Student with an interest in Project management and Analysis. He recently completed his project management coop where he mainly focussed on BAAN ERP system to extract, analyse and process data according to engineering change required in the product. His skills include team management, marketing and leadership in the field of project management. After completing his undergrad in mechanical engineering, he started working as an Operations engineer for almost 2 years. He was assigned to coordinate sales and service team, optimize goods transportation, check on production schedule, managing work force and assist in production.

Zixiao, Wang

Tentative role: Member

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Responsibilities: I will do research about emotion analysis based on text and ML prediction about number sales price. After that, I will implement there two parts in jupyter notebook.

Key background: A first year graduate student in Information System. Have experience in statical analysis and ML. Build a forest regression model to predict Beijing PM2.5. And a main contributor to the NEU Skunkworks workshop with EM Lyon during Nov, 2019.

Expertise: Python, R, Java, HTML, CSS, JS, Jupyter notebook.

Nityasheee Vijay Kumar

Tentative Role: Member

Responsibility: I will work on creating and designing the database.Work on collecting the structured and unstructured data, integrating the data and creating a structured dataset by creating the variables and assigning values to those variables based on the research. Also, help in analysing the outcome of the data.

Key background: First year graduate student in Information Systems. Have experience in data analysis. Worked for Insurance and Investment bank as data analyst.

Expertise : SQL, SAS, Tableau, MS Office

Jalal Jamaluddain A Hassan

Tentative Role: Member

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

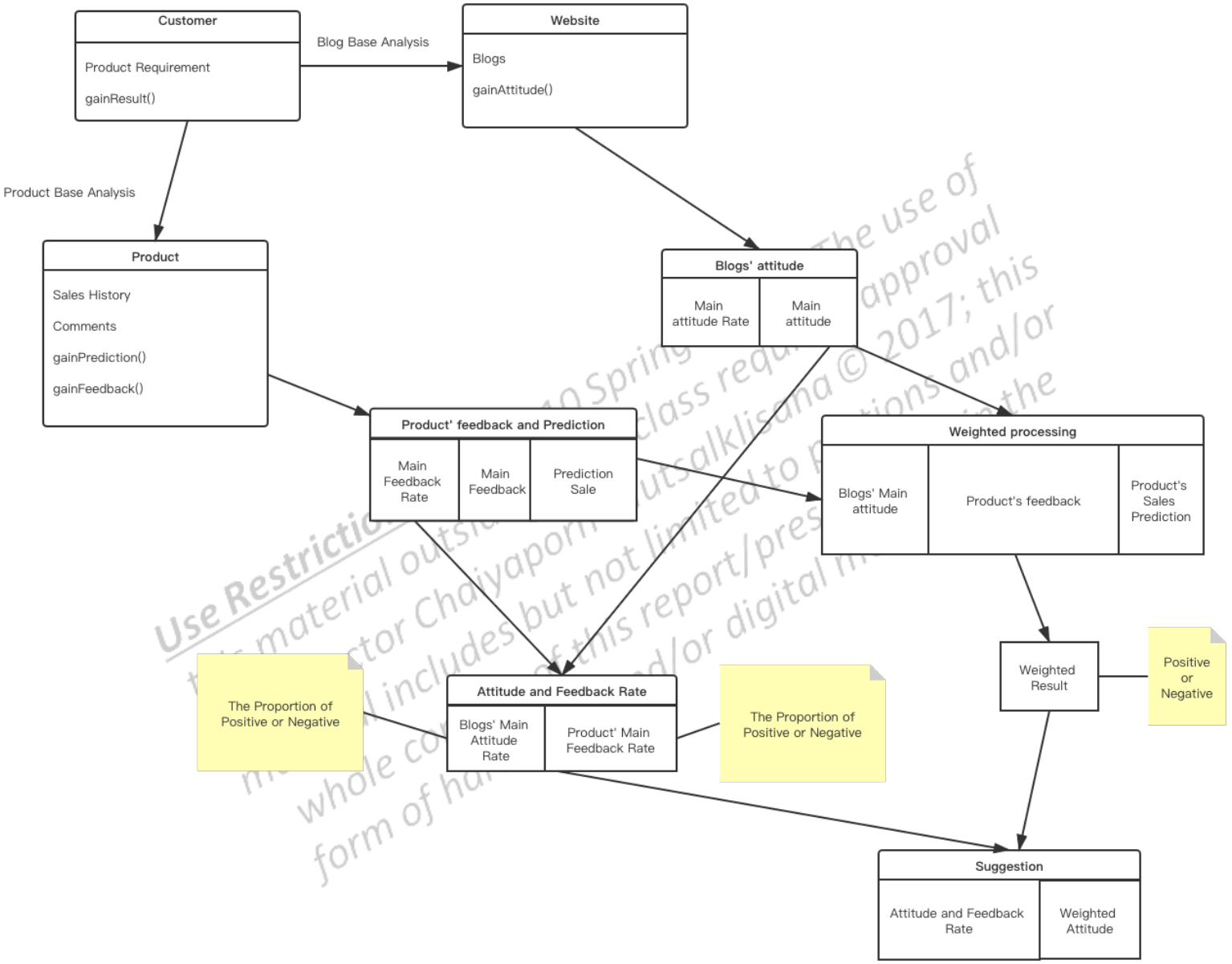
Responsibility: I will manage the whole project. Lists schedule and control time nodes. List the plans for the next week and summarize the results of the previous week. I will also help Zixiao to analysis the data and provide some technical directions.

Key background: Second year graduate student in Engineering Management.

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Other items to include

Top level activity diagram (hand sketch/visio/other tools) and/or sequential statements



Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Assumptions and Constraints

Since there are lots of supplements available in the market for fitness, collecting reliable data and analysing them would be a challenge as the analysis can be done only for limited products. Below are a few scenarios where the project might not give 100% reliability.

Scenario 1: Loyalty towards a brand makes a huge difference in the analysis here. Some brands will have customers who are loyal to their brands and some bloggers are paid by the brand to market them. In such situations we can expect biasing about the product of particular brand.The assumption of the project is to consider all the blogs are genuine and honest.

Scenario 2: Since the project is focusing on globally known products there is a higher chance of leaving out not so known products. For eg: Patanjali products are very popular in India and most Indians use those products. Even though it is doing well in Indian market it will not be the most recommended product as it is not doing very well in the global market.

So, the constraint here is that the project does not focus on a particular geographical area but focus on the global customers.Hence the project will not work merely for people in one region. It serves all the people across the globe and for the globally known products.

Scenario 3: The data extract is constraint only to certain blogging websites,feedbacks and comments from different E-Commerce websites, product sales history. So, the suggestion is based on only these resources.

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

Integrated Master Schedule/Milestones

Jan 24 Proposal

Jan 31 Do research about data resource

Feb 7 Collect data

Feb 14 Design database

Feb 21 Pre-process the data

Feb 28 Mathematical about price prediction

Mar 6 Build the automatically data collecting program

Mar 13 Build the price prediction model

Mar 20 Mathematical about emotion analysis

Mar 27 Build the emotion analysis model

Apr 4 Tuning the models to get better result

Apr 11 Finish final presentation

Bravo with
Nityashree Vijay Kumar, Jalal Jamaluddain A Hassan, Saurin Nitinbhai Shah

References

Jacob, C. (2020). 24 Fitness Experts, Authors, and Blogs to Follow in 2019. Retrieved 22 January 2020, from https://upjourney.com/fitness-experts-authors-and-blogs-to-follow-in-2019

MATSUO, Y., & ISHIZUKA, M. (2004). KEYWORD EXTRACTION FROM A SINGLE DOCUMENT USING WORD CO-OCCURRENCE STATISTICAL INFORMATION. *International Journal On Artificial Intelligence Tools*, *13*(01), 157-169. doi: 10.1142/s0218213004001466

https://www.worldscientific.com/doi/pdf/10.1142/S0218213004001466

Chaffar, S., & Inkpen, D. (2011). Using a Heterogeneous Dataset for Emotion Analysis in Text. *Advances In Artificial Intelligence*, 62-67. doi: 10.1007/978-3-642-21043-3_8

https://link.springer.com/content/pdf/10.1007%2F978-3-642-21043-3_8.pdf

Automated Keyword Extraction from Articles using NLP. (2020). Retrieved 22 January 2020, from https://medium.com/analytics-vidhya/automated-keyword-extraction-from-articles-using-nlp-bfd864f41b34