Mike Choi, fleetpro@gmail.com, Skype Name: live:3ac11b4a862664b5

Proposal – 08

Part. Appendix

1. Introduction

See the Main part of the proposal.

2. Background

See the Main part of the proposal.

Contents

3. Executive Summary

See the Main part of the proposal.

4. Terminology

See the Main part of the proposal.

5. Team, Vision, Mission

See the Main part of the proposal.

6. Use-Case View – Methodology

See the Main part of the proposal.

7. Use-Case View – Enumeration

See the Main part of the proposal.

8. Objectives

See the Main part of the proposal.

9. High-level Architecture

See the Main part of the proposal.

10. Management

See the Main part of the proposal.

11. Appendices

Appendix-1 Product Assessment & Comparison (PAC), Opportunities

This section aims to discuss business opportunities in respect of e-commerce Product Assessment & Comparison (PAC). It will contribute to the shaping and convincing of business vision for project members, sponsors, and other stakeholders.

E-commerce PAC helps shoppers make assessment and comparison of products. PAC analyzes product data and reviews to provide Consumer with concise yet comprehensive, personalized product insights enabling them to make informed decisions, have confidence to purchase products, and save time.

There are now massive amount of product data and reviews available on major e-commerce platform websites, like Amazon, Facebook, and Twitter, that PAC will explore and exploit. Natural Language Processing techniques are employed to summarize the product data and reviews to specific needs. See Appendices for technical discussion of Opinion Analysis techniques.

Figure. Typical e-commerce platforms provide shoppers with Product Insights and Recommendations based on Text Analysis of User Generated Content.
(Image source: aspectiva.com)

Before we discuss business opportunities arising from e-commerce PAC, we are going to have some case studies on major multi-vendor e-commerce platforms to have concrete ideas of the current NLP practices in e-commerce NLP.

Case Study 1. Amazon

Amazon.com is a typical multi-vendor e-commerce platform that employs Natural Language Processing techniques.

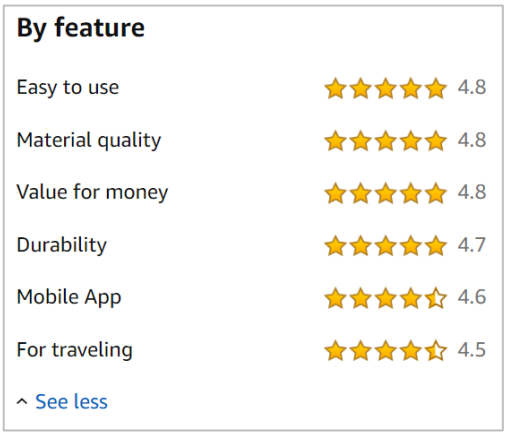
For a particular product picked by a shopper, Amazon assists the shopper with

- Customer reviews classified by feature (RC-by-Feature).
- Customer reviews classified by mention (RC-by-Mention).

Two examples of product assessment are illustrated below.

Amazon Example 1. ASIN: B0773JHC6N, AmazonBasics 60W 6-Port Multi USB Wall Charger.

Customer review ratings classified by feature.



Customer review ratings are classified by feature, as shown above. The features of this particular product are “Easy to use”, “Material quality”, “Value for money”, “Durability”, “Mobile App”, and “For travelling.”

Features are specific for a given product, and derived from reviews on the product. Features can inspire shoppers to what aspects they should take into consideration before deciding on purchase. It will be critical to make sure:

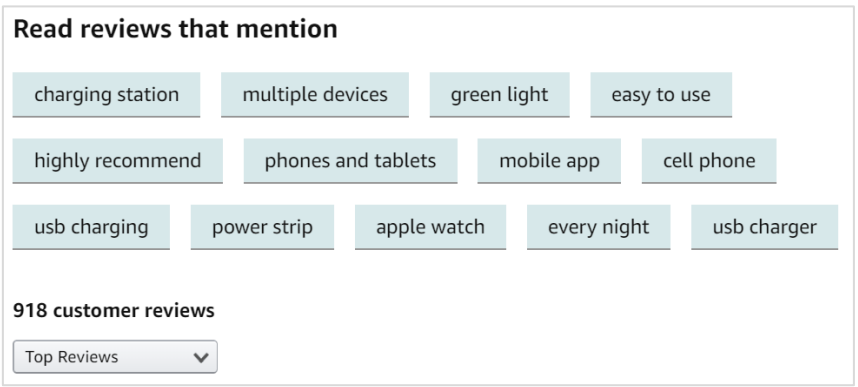
- The features are the principal components of customer opinion.
- The rating of a feature really represents the extent to which the feature is rated.

The words or phrases that represent a product feature are not necessarily explicitly mentioned in the product reviews. They are derived, rather than picked up, from the reviews, since many sentences imply, rather than physically contain, them.

For example, “Looking from the outside in, one would hesitate to enter this hole in the wall (really, that's what it is)” is criticizing a restaurant for its distracting appearance. *An intelligence that can derive “Appearance is distracting”, for example, from this sentence seems not yet fully developed, though there have been attempts at that.*

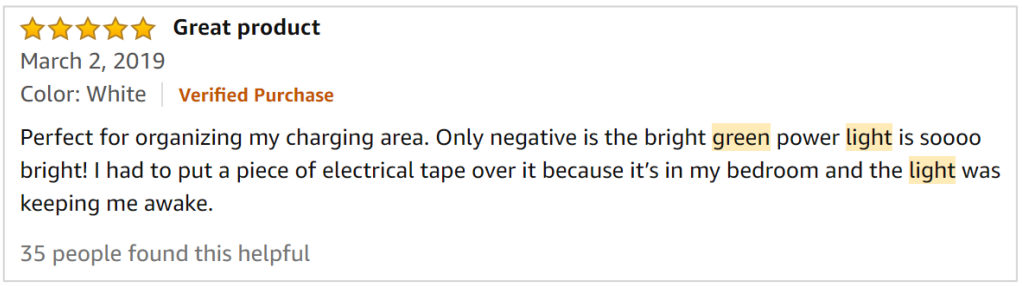
Choosing right words/phrase of right features for a given product is crucial in helping shoppers grasp the insights of the product. See Apendix-3 for more about Natural Language Processing.

Customer reviews classified by mention. (Mention is words or phrases that are mentioned in reviews.)



Amazon also classifies reviews by mention, as shown above, which is words/phrases nearly explicitly mentioned in reviews.

Clicking the mention button “green light”, for example, shows a list of reviews that seem to mention the “green light” of the product. The polarity of sentiment (positive or negative) is not identified here. The following is a review that is displayed by the mention button “green light”.



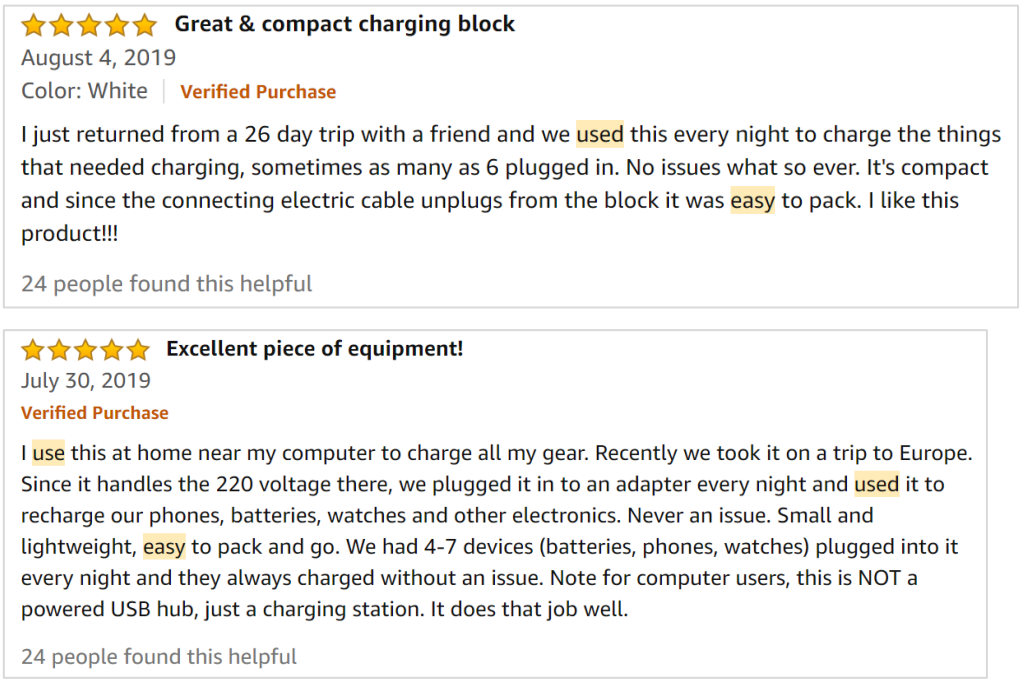
★★★★★ **Great product**
March 2, 2019
Color: White | Verified Purchase

Perfect for organizing my charging area. Only negative is the bright green power light is soooo bright! I had to put a piece of electrical tape over it because it’s in my bedroom and the light was keeping me awake.

35 people found this helpful

There is no alphabet string of “green light” in the review, though a green light is clearly mentioned. The NLP algorithm did a great job by excluding, from the mention list, the following nouns and noun phrases: light, power light, bright light, bright green light, bright power light, bright green power light, “electrical tape, a piece of electrical tape, and bedroom. *Features are, thus, derived, rather than picked, from reviews.*

The “easy to use” mention button serves the following reviews, the relevance of which is quite convincing.



★★★★★ **Great & compact charging block**
August 4, 2019
Color: White | Verified Purchase

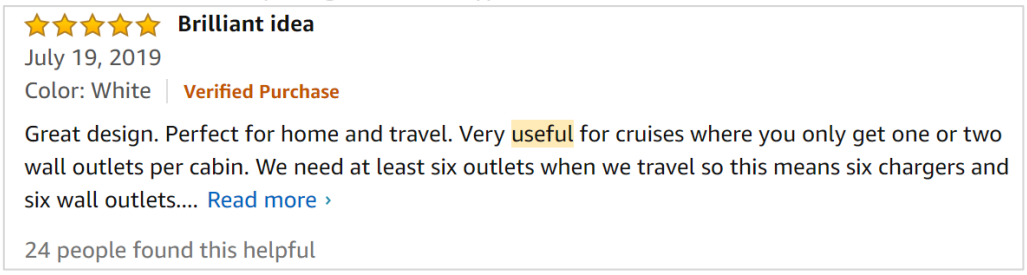
I just returned from a 26 day trip with a friend and we used this every night to charge the things that needed charging, sometimes as many as 6 plugged in. No issues what so ever. It's compact and since the connecting electric cable unplugs from the block it was easy to pack. I like this product!!!

24 people found this helpful

★★★★★ **Excellent piece of equipment!**
July 30, 2019
Verified Purchase

I use this at home near my computer to charge all my gear. Recently we took it on a trip to Europe. Since it handles the 220 voltage there, we plugged it in to an adapter every night and used it to recharge our phones, batteries, watches and other electronics. Never an issue. Small and lightweight, easy to pack and go. We had 4-7 devices (batteries, phones, watches) plugged into it every night and they always charged without an issue. Note for computer users, this is NOT a powered USB hub, just a charging station. It does that job well.

24 people found this helpful


★★★★★ **Brilliant idea**
July 19, 2019
Color: White | **Verified Purchase**
Great design. Perfect for home and travel. Very useful for cruises where you only get one or two wall outlets per cabin. We need at least six outlets when we travel so this means six chargers and six wall outlets.... Read more ›
24 people found this helpful

These reviews have many more candidate features and mentions, like

- (Features) travel use, charge electronics, multiple devices, charging station and home use; of which “travel use” and “charging station” are listed as a feature.
- (Mentions) night, compact, cable, block, computer, gear, trip, Europe, adapter, issue, phone, electronics, computer user, USB hub, charging station, job, home, travel, cruise, and outlet; of which “night”, “phone”, “charging station”, and “travel” are listed as a mention.

The NLP algorithm made the following right decisions:

- “easy to use” is derived from “we used …”, “easy to pack”.
- “charge electronics”, “adaptor”, and “electronics” are excluded, maybe because they are actually mentioned in the product category name and description.
- “every night” is correctly chosen, rather than simple “night”, to show a value name.
- “phone” is focused on among other (less important) accessories mentioned in parallel.
- “compact”, “cable”, “gear”, “Europe”, “USB hub”, and “outlet” are adequately excluded.
- General terms are listed as a feature, and specific terms as a mention.

The NLP algorithm made the following questionable decisions, which are wrong unless there are too few features or mentions to list:

- “mobile app” is not very relevant to an electric charger, as is “power strip”.
- “easy to use” is listed as both a feature and a mention.
- “highly recommended” is not a value name, but a value extent.
- “phones and tablets” and “cell phone” are redundant to each other, as are “usb charging” and “usb charger”.
- “multiple devices” and “usb charger” are redundant, as they are already mentioned in the category name.
- “apple watch” is too specific to be listed as an independent mention.

The mention button “power strip” is not of great help, as found in the following review which is served by the button.

★★★★★ **Great product at a very good price.**
December 15, 2018
Color: White | Verified Purchase

Makes life so much easier to charge all of your portable devices. Four one amp USB outlets for phones and four 2 amp USB outlets for tablets. All this with only one plug to put into your AC outlet. Now I can finally charge my two phones and three tablets at the same time from one power strip and still have room to plug in my lamp, fan etc. Fits the bill perfectly.

Amazon Example 2. ASIN: B01MSUOG4A, Amazon example of e-commerce NLP - BioShock Big Daddy Backpack.

Customer review ratings classified by customer group & interest, unlikely with the above example.

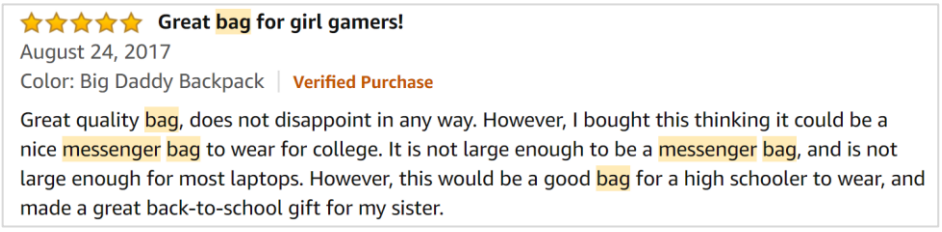
By customer groups & interests ⓘ

Gaming	★★★★½ 4.6
Backpacking	★★★★½ 4.5

Is this feature helpful? Yes No

The customer groups & interests can be derived by grouping product features, which, in turn, are derived from the reviews. Labeling a group of features, like “gaming”, is a challenging task.

The following two reviews are served by the mention button “messenger bag”. They seem to justify that the button is worth listing.



The relevance of the button “sister”, however, is not so obvious, because “the bag making a gift for one’s sister” may or may not reflect the features that interest shoppers. Note that “laptop” button could also be listed with greater relevance than “sister.” The fact that “The bag is not large enough for most laptops” would help shoppers more than does the fact that “the bag made a great gift for one’s sister.”

If the NLP engine took into consideration that the review’s sentiment for laptop is negative, and if negative sentiment and mentions were intentionally skipped for this specific product or shopper, then this would be an example of bias in product assessment and comparison that favors vendors. There could be a “laptop” mention button, instead of “sister”.

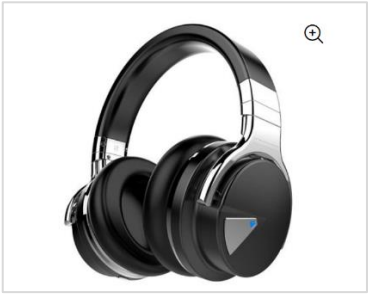
Case Study 2. Walmart

Walmart.com is also a multi-vendor e-commerce platform that employs Natural Language Processing techniques.

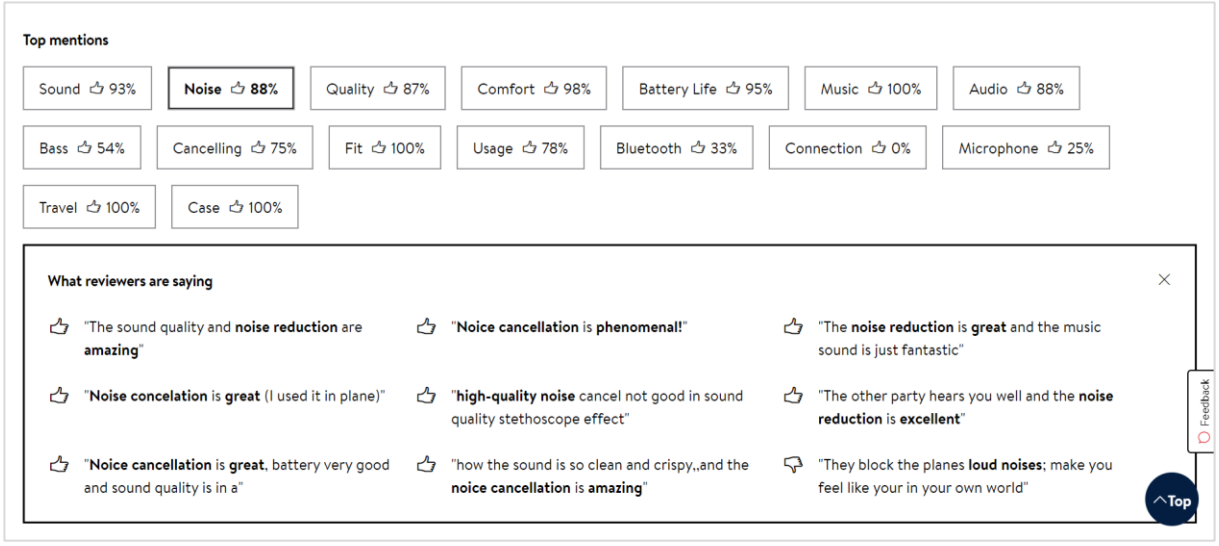
For a particular product picked by a shopper, Walmart assists the shopper with

- Customer reviews classified by mention.
- Most helpful positive/negative review.

Walmart Example 1. Active Noise Cancelling Headphones Bluetooth Headphones with Mic Deep Bass Wireless Headphones Over Ear

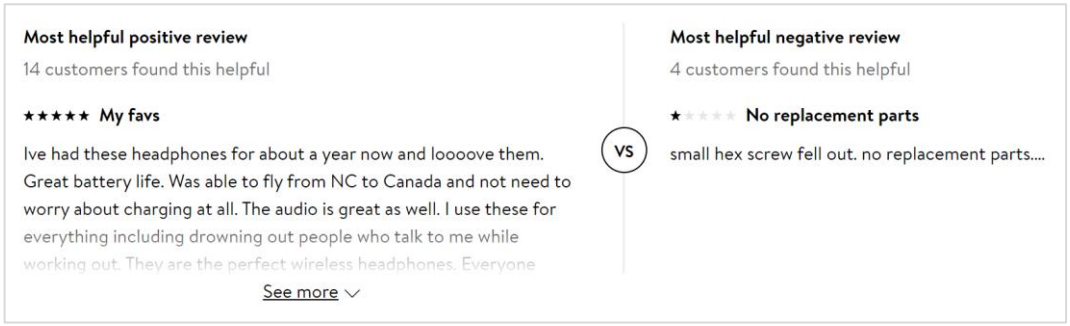


Customer review classification by mention is a bit different from that of Amazon, as shown in the below figure.

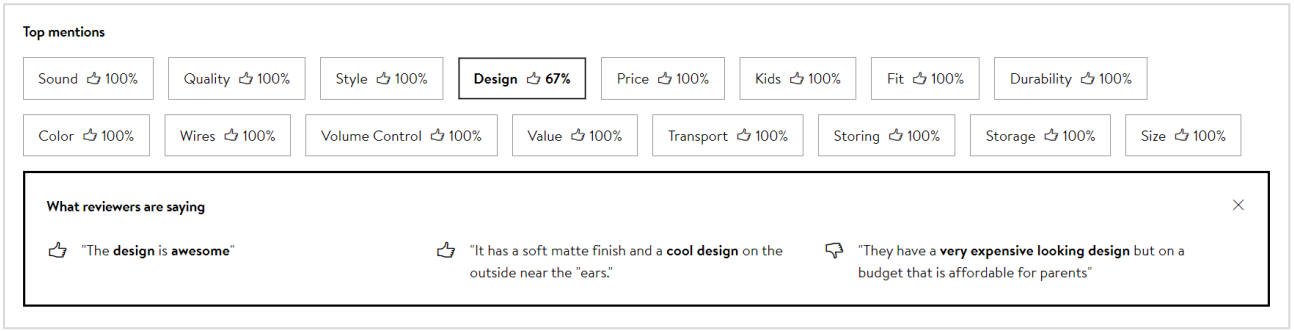


While Amazon has both Reviews Classified by Feature (RC-by-Feature) and Reviews Classified by Mention (RC-by- Mention), Walmart has RC-by-Mention only, with enhancement. Walmart also has Most Helpful Positive/Negative Review, as below.

The helpfulness score is calculated directly form the number of customer votes. The NLP engine simply classifies the sentiment of review.



Walmart Example2. Children's Wired Headphones with SharePort - Assorted Colors (Kids, Toddlers, Boys, Girls)



Summarizing case studies

E-commerce aspect/feature extraction can be defined as "finding a set of words/phrases that represent the important value name of a given product from the customer reviews on the product."

Looking into the Walmart RC-by-Mention, shown above, you find that its Amazon equivalence is Amazon RC-by-Feature, actually. Walmart RC-by-Mention is compared with Amazon RC-by-Feature as follows:

- Walmart RC-by-Mention lists neat, distinctive, and relevant mention words, like sound, noise, quality, battery life, music, audio, bass, etc., which sound like pre-defined and fixed per product category. But investigation shows the words are at least not fixed per category and are based on the reviews on particular product. If, further, the words are not pre-defined and are derived from reviews, then Walmart NLP engine excels that of Amazon.

- Walmart RC-by-Mention tags each mention with a percentage score, giving more intuitive comprehension to shoppers. The score represents the extent to which the mentioned feature is rated. You find that the score is the percentage of positive mention instances over total mention instances for the given mention. For example, if 10 reviews mentioned "quality", 7 of which did positively, and 3 of which negatively, then the score of the mention "quality" is 70%. It's a simple NLP task, yet seems to help shoppers much.

- Walmart RC-by-Mention displays the very sentence responsible for and relevant to a given mention, instead of the whole review text that has the mention. This helps shoppers focus on the chosen mention. It's a simple NLP task, yet helps shoppers much.

- Both platforms avoid spotting the opinion words that describe a given mentioned feature. For example, in the sentence "It is great in color," mentioned feature is "color", and opinion words is "great." Both platforms do not pick up the opinion words "great." Instead, Amazon gives the whole review, and Walmart the sentence. We know that extracting opinion words is more difficult than feature words.

- Walmart RC-by-Mention made a few mistakes, as shown in the above figure, in sentiment classification. This is almost unavoidable in NLP tasks, and suggests that we will have to improve NLP engines on an ongoing basis.

Opportunities discussed

Through the case-studies as well as other market research, we identify or confirm the following opportunities:

- The selection of reviews to search for features, together with the selection of features to show to customers, is up to the specific purpose. The purposes may include, for example:
  - ✓ encouraging shoppers to buy more: for a specific vendor, of a specific product
  - ✓ promoting vendors to sell more: of a specific product, to a specific customer group
  - ✓ indirectly boosting Value for Money: for a specific product category/vendor
  - ✓ personalizing to a specific shopper
  - ✓ adjusting the product competition
  - ✓ trading off between the interests of vendor, shopper, reviewer, and the platform itself

- Amazon and Walmart, and many more, make a good example/convention of multi-vendor e-commerce platform that exploit NLP techniques in product assessment. When it comes to comparison, however, Amazon does not make an explicit comparison of products either in its name or on behalf of customers. Walmart also serves a neat assessment of product by using NLP techniques, without comparison.

- It seems that the NLP adaptation curve is now at the matured Early Adopters Phase. Although leading e-commerce platforms have long been employing NLP techniques in PAC, the majority platforms are either waiting for clear signals of productivity gain or adopting NLP simply as a fashion. We need a breakthrough to motivate or accelerate the transition to Early Majority Phase.

- Once ignited with real productivity gain, no platform will be able to resist adopting NLP. Enhancing conversion rate is a kind of zero-sum game, as new e-commerce techniques can not significantly improve the society's total purchase. Early Adopters and Early Majority will adopt AI/NLP to get benefit until adoption process is over, whereas Later Majority and Laggards will have no way but adopt AI/NLP to stop the loss, which will otherwise be converted to pioneering adopters' gain. The growth rate of adopters increases constantly. Once fully adopted, there cannot be returning back. The society forces it and benefit from it, implicitly.

- There is a gap between e-commerce industry and academia. The quality of existing practice of AI/NLP in e-commerce has long way to go, whereas the academic community

does not pay much attention directly to e-commerce AI/NLP problems. The majority NLP research papers are racing for novel state-of-the-art performance in a few defined benchmarks, none of which closely relate to e-commerce problems.

- No platform seems to have PAC-by-Review yet, though it will unleash conversion rates for Early Adopters. PAC-by-Review makes product assessment and comparison based on product reviews, rather than vendor-controlled product specifications. Once satisfied by specifications, shoppers tend to refer to PAC-by-Review. For the majority consumers, comparison of word-of-mouth is the best decision factor.

- Apart from pure PAC-by-Review, there are behavioral parameters that govern. The parameters will control how much to favor particular vendors, shoppers, reviewers, and the platform itself, respectively. They will also affect ….

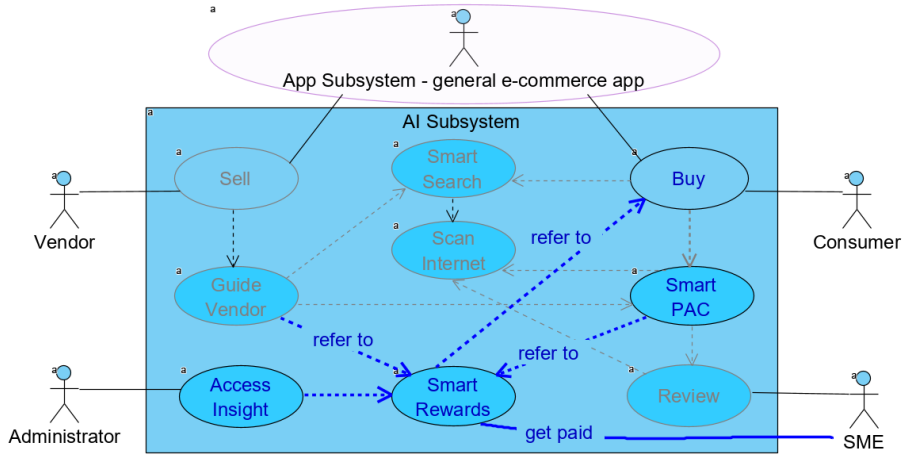
Taking opportunities

The most proactive action of taking the opportunities discussed in about subsection will be to

- develop and distribute PAC-by-Review functionality in any form,
- make it the decisive motivation for transition from Early Adopters to Early Majority Phase.

- STEP1. Fundamental PAC-by-Review will be developed by combining and improving the-stage-of-art NLP techniques. The following way of thinking is conventional in AI/NLP/Deep Learning:
  - Identifying numerical representation of the major shared/similar features of given set of products of the same category, (because numerical space is free from lack of words and is a playground of rich mathematical operations.)
  - Searching for best words/phrase that closely match the numerical representation of feature.

- STEP2. Smart Rewards will be developed to intelligently guide PAC-by-Review. From the Consumer behavior who does browse, compare, reject, and choose products; Smart Rewards, which passively guides PAC, will learn a deep insight of Consumer behavior and become smarter over time. Intelligently identifying a particular shopper deciding factors from his/her purchase behavior, for example, is crucial in all aspects. See Appendix-2 for more about Smart Rewards.

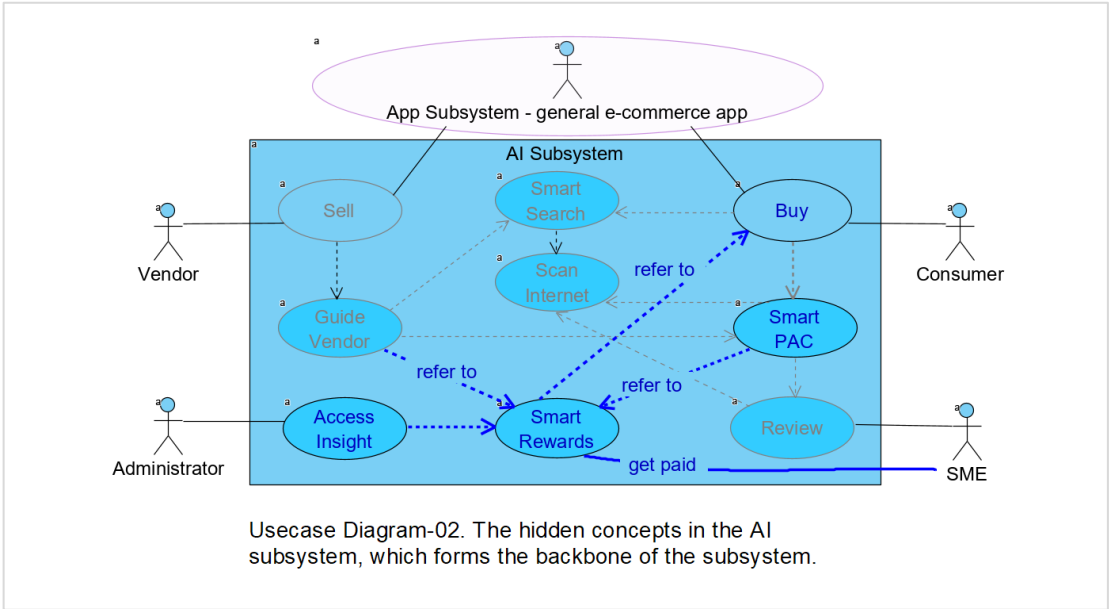


Usecase Diagram-02. The hidden concepts in the AI subsystem, which forms the backbone of the subsystem.

- STEP-3. PAC-by-Review will be enriched with behavioral parameters. The behavioral parameters will be a handle for Smart Rewards to guide PAC-by-Review, and will be the leverage that the Sponsor and Administrator have in order to influence the buying behavior of Consumer as well as the selling behavior of Vendor. It means that the Assessment &

Comparison will not be *purely objective and on its own* as it cannot be, and will be influenced by behavior of Consumer, Vendor, and SME (Subject Matter Expert) as well as the Sponsor.

- Right PAC for Right Customer/Vendor/Product will be implemented based on the intelligence gathered by Machine Learning techniques. PAC will refer to Smart Rewards for its behavioral parameters when it makes decision of, for examples:
  - What will be the best strategy of persuading a particular Consumer to buy a particular product?
  - Which Vendor's product to influence Consumer to buy, when there are multiple eligible products available for them?
  - Even, what is the best order of lookup/comparison list that will be provided to a specific Consumer who seems to want a specific category of product, if we dare not to hide or uphold information?



Usecase Diagram-02. The hidden concepts in the AI subsystem, which forms the backbone of the subsystem.

- AI aspect/opinion extraction technique will be integrated to help
  - customers to not only realize fashionable aspect/opinion on a given branded product, but also search for branded products within a neighborhood of a given aspect/opinion,
  - vendors to learn what aspects/opinions are arising on their categories of products in social networking sites.

- For automatic assessment of product, we will develop
  - AI models that can assess the money value of given aspect/opinion for a category of product. For example, the AI model will say "The reviews add 50$ value to the product."
  - additional techniques, like personalized assessment of product, when the platform gain substantial customer base and we are invited again.

See Appendices-2 for more.

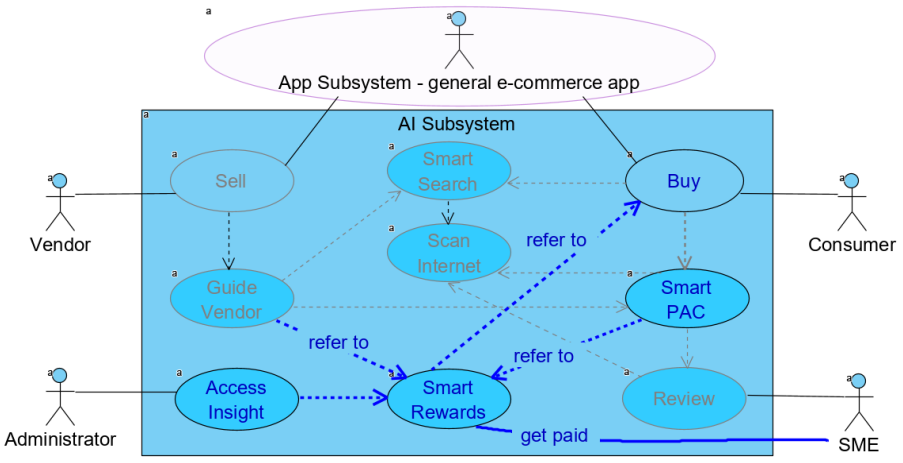
Appendix-2 Rewards & Retention - strategy

The fact that Vendor, Consumer, and SME have no way but co-exist with each other does not mean that they will adhere to a single e-commerce platform. There are many competing platforms out there.

In this section, we see that the [Smart Rewards, U.8] use-case, and its background modules, forms the backbone of the system for a feedback flow. In the form of circular loop, it will be both a source of long-term sophisticated performance, yet, a source of instability if not properly designed.

Challenges

While image detection and recognition in product search will become a commodity soon, the Rewards & Retention mechanism will be the ever-lasting foundation that enables the long-term health of the platform behind the scenes.



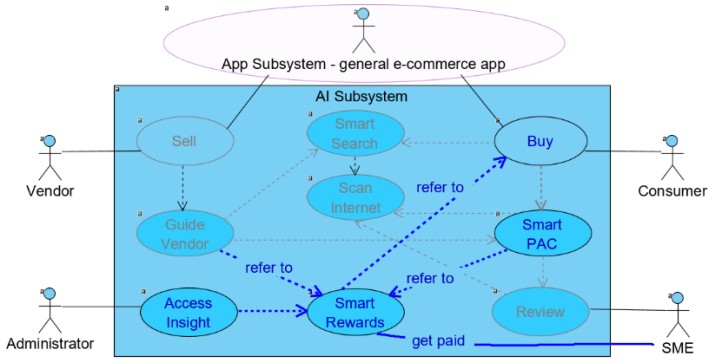
Usecase Diagram-02. The hidden concepts in the AI subsystem, which forms the backbone of the subsystem.

Before we can talk about algorithms of Smart Rewards, it would be helpful to identify the independent key factors that have influence on Smart Rewords. Vendor's satisfaction (so their revenue and cost), Consumer's satisfaction, and the SME's clearly belong to the key factors, although they are almost impossible to measure. It's even more so when we want to compare these factors with those on competitors' platforms. We would have to dive deep into the domain knowledge of e-commerce, in vain, to find the key factors if there was no modern AI technology.

Let us first try to walk a conventional thinking to find what are challenges of Rewards & Retention.

- Challenges in Rewards & Retention
  - Short-sighted Vendor tend to result in lower Value for Money products.
  - Consumer desires for higher Value for Money.
  - Value for Money is hard to represent numerically.
  - SMEs tend to alternate between their popularity among Consumer and payment from Vendor.

- Ways to reward/punish actors on a macro basis
  - Raise/lower Value for Money index to reward/punish Consumer.
  - Lower/raise Value for Money index to reward/punish Vendor.
  - Pay higher/lower to reward/punish SME.
  - Promote/degrade SME's popularity on the platform to reward/punish SME.



Usecase Diagram-02. The hidden concepts in the AI subsystem, which forms the backbone of the subsystem.

- Ways to raise/lower Value for Money index on the platform.
  - Bias/censorship in the process of automatic assessment of products.
    Ways: selective product data to feed to the process.
    Ways: selective reviews to feed to the process.
    Problem: We need a precise numerical representations of product data and review.

    Solutions to numerical representation/measure of product data and review:
    - Create a mathematical/symbolic/logical/ semantical model (formula).
      (let's call it an MSLS formula)
      ---→ We reject this. It would give faster result, but the accuracy is limited and not competitive today. There are too many features. We cannot understand the meaning of many of them.

- Create a statistical model, which does not care of the meaning of features.
  ---→ Better than MSLS.
- Create a Deep Learning model.
  ---→ Best so far. Enhance DL with other AI techniques.

This is a general common-sense reason for choosing modern AI, and not our invention.

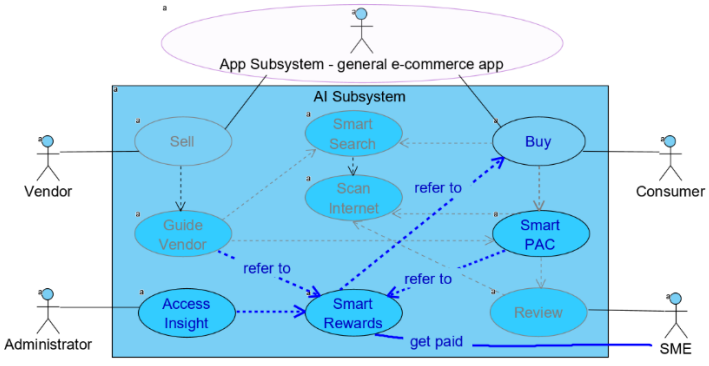
Surroundings

We are going to discuss the environment that surrounds Smart Rewards use-case.

- Smart Rewards use-case is guided by Administrator

Smart Rewards use-case is expected to be guided by Sponsor. Administrator will guide Rewords use-case by using a dashboard tool, represented as Access Insight in the diagram, to their desired state. They will, for example, drag the slider of Value for Money index (VMI below) rightward and leftward to change the index up and down, so that the local (platform-wide, micro, not national) VMI will be slowly decreased or increased. Do not go into how for the moment.

Administrator will also regulate how Smart Rewards takes care of SME.



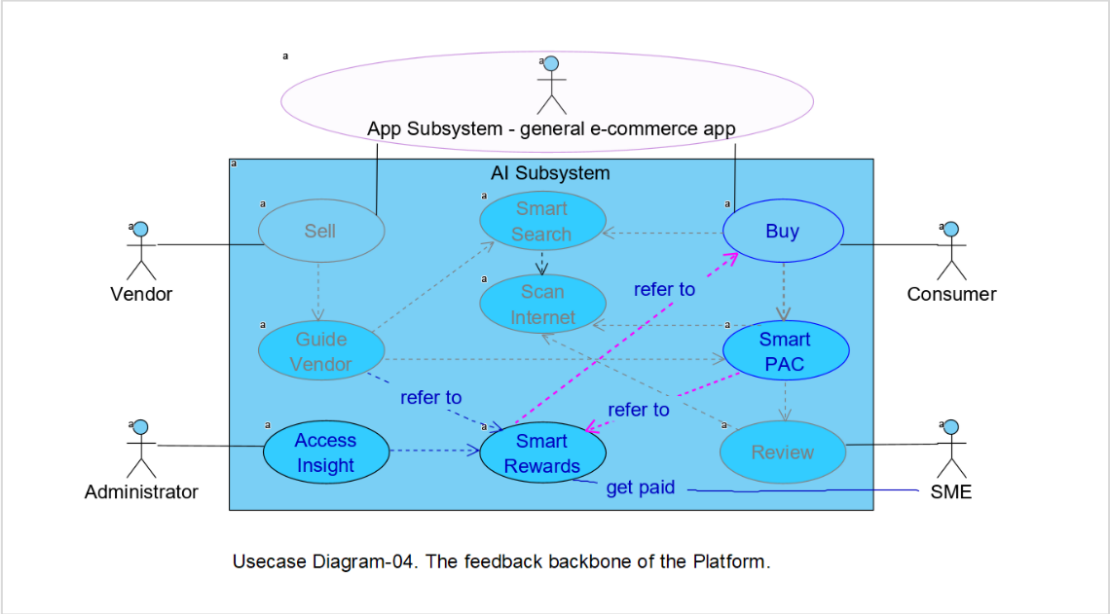
Usecase Diagram-02. The hidden concepts in the AI subsystem, which forms the backbone of the subsystem.

- Smart Rewards use-case taps into Buy use-case

Smart Rewards use-case will be intelligently self-adjusted based on feeds from Buy use-case. Buy is the place where the interests and performance of all actors/stakeholders interact. All aspects of performance are directly or indirectly measured there. Examples include:

- Which Vendor's which category of product is soled how much by which Consumer and at what seasons?
- What journeys, like product assessment/comparison and reviews, Consumer walked before they finally chose the product?
- What are the typical journeys Consumer walk for a specific category of product?
- What branches of journeys lead Consumer giving up buying a product or expand to cross-shopping?
- Which SME's which review affected the behavior of Consumer how?
- How does Consumer rate an SME's product review, though Consumers would rarely click at the rate report panels?
- How high is the satisfaction of Consumer, though they seldom report it and it is difficult to measure?
- What are the long-term tendencies of specific or average SME in their product review?

What we see through these examples, first of all, is that most of the features that reflect the operation of the platform is non-numerical, vague, qualitative features. Smart Rewards use-case will be utilizing these features completely by using Machine Learning techniques.



Usecase Diagram-04. The feedback backbone of the Platform.

What's more that justify Smart Rewards use-case tapping into Buy use-case: The health and retention rate of the platform in much affected by SMEs, who tend notoriously to get paid by Vendor backstage. Looking into this phenomenon, we see that the advantage they take to get paid is their popularity, which we plan is decided by Consumer. Rewards & Retention will refer to Buy Product to decide the performance of SME, yielding chances to Consumer to feedback about SME directly or indirectly.

Feedback Backbone

- Smart Rewards use-case is tapped into by Smart Assess use-case, closing the feedback loop.

  Influenced by Administrator and Buy use-cases, Smart Rewards will this time influence Smart Assess use-case, which, in turn, influences Buy Guide Vendor use-cases.

  This tap will form the backbone of feedback flow in the platform.

  The most important feedback flow in the platform is:

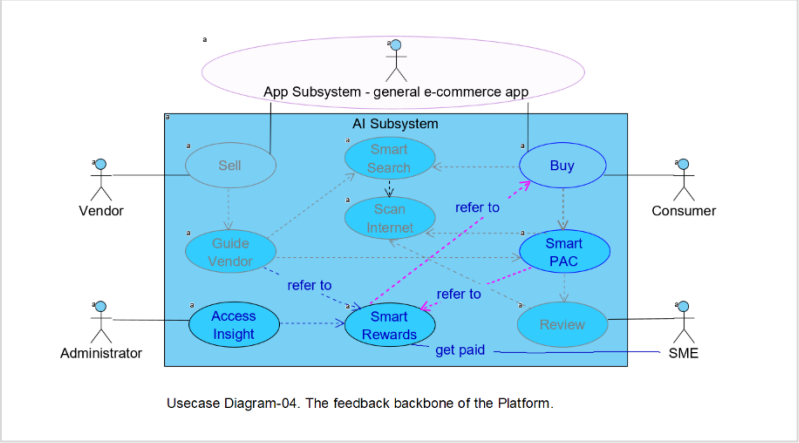
  Buy <- Smart Assess <- Smart Rewards <- Buy

  Note: the line shows the flow of feedback, which is the opposite of dependency flow.

  This way, Smart Rewards use-case stands as the canal of feedback flow, and is integrated to the key modules identified by the Sponsor. The importance of the connection between Smart Rewards and Smart Assess use-cases is that it un-obviously links its left and right chains that are obvious links. Let's call this flow of feedback the backbone.

  The backbone will integrate the remaining elements to generate sophisticated long-term performance of the platform. We will make it an AI model that can learn from the supervision of Sponsor, from trial and error, from historical data, and even from the behavior of competitors' platforms if possible.

  We are aware that the circular loop of backbone, and the hidden parameters inside it, on the other hand, may cause instability of the platform, because there may be spontaneous resonance between them.



Usecase Diagram-04. The feedback backbone of the Platform.

Machine Learning

- Smart Rewards decides payment to SME

  Actually, Rewards & Retention will decide payment not only to SME but also virtually to all actors, if only it contributes to the goal of Smart Rewards. Payment here is not limited to dollar, but may include more as we can imagine, including Value for Money biases mentioned earlier, promotion of specific SME or Vendor, and sort of premium, credit, and coupons. The algorithm will be in the form of symbolic logic for initial delivery, and will be replaced with Machin Learning algorithm if enough data have been accumulated.

- Machine Learning techniques can be used to find:
  - The Value for Money index running currently on the Platform. Auto-Encoding
  - running fashions, aspects/opinions of products.
  - dollar value of aspects/opinions. Auto-Encoding
  - the popularity index of SME. Auto-Encoding
  - the influences of the above derived factors on Rewards & Reward.
  - an algorithm to regulates the Value for Money index as well as the popularity of SME.
  - an algorithm that stabilizes the backbone on an overall basis.

Appendix-3 Opinion Analysis in general

(Under construction)

Appendix-4 Opinion Analysis techniques demonstrated

(On specific request)

Appendix-5 Image Detection/Recognition (IDR) techniques

(Under construction)

Appendix-6 Reinforcement Learning techniques

(Under construction)

Appendix-7 Market Research

(Under construction)

Appendix-8 Marketing Strategy

(Under construction)