# Service Intelligence Week 4.

# [Service Review Mining for Service Improvement]

Chiehyeon Lim

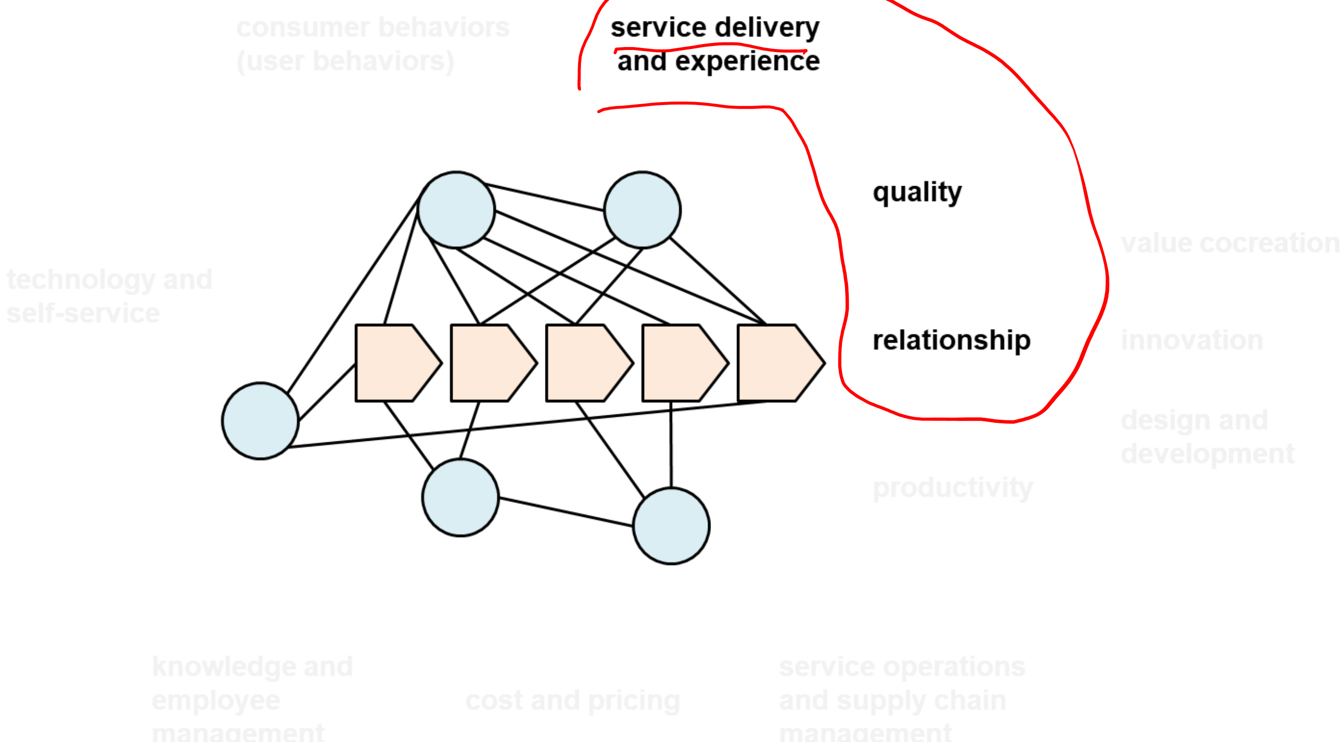
2022. 9. 19

# Service Quality Representation:

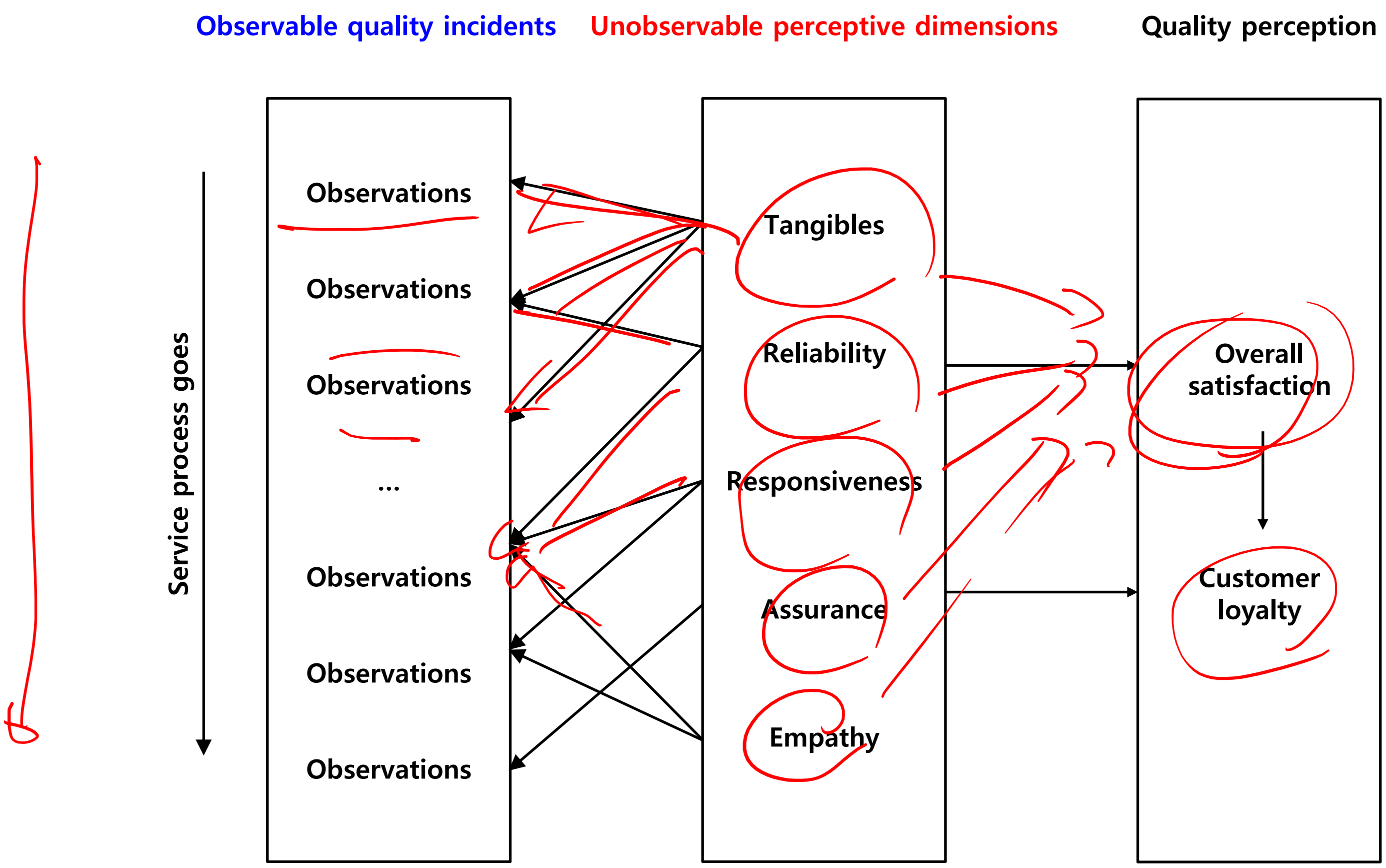
# A Review and Further Discussion

# Questions and Learning Objectives

- How can we “explicitly” evaluate and improve the “implicit” quality of services?
- How can we use data and learning intelligence for service quality representation?

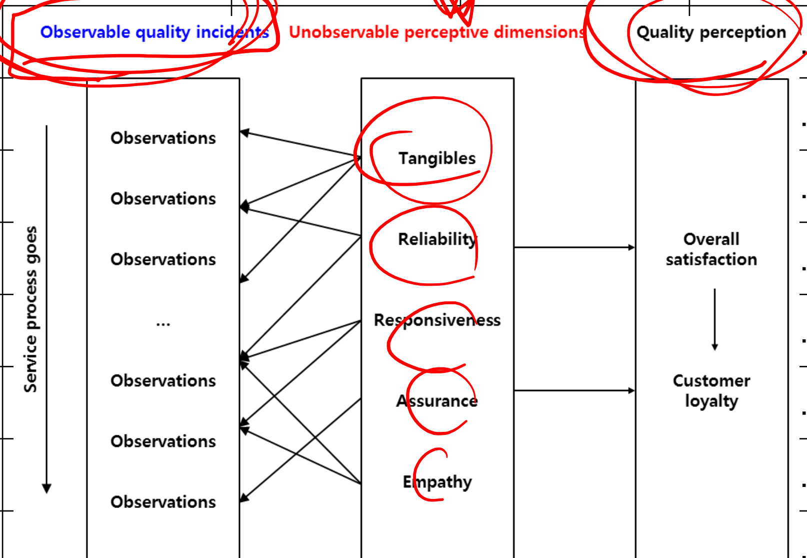


# Illustration of Service Quality Evaluation



# Service Quality “Representation”

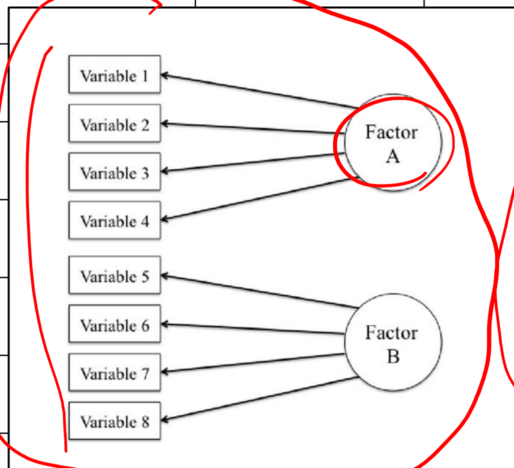
| | Feature 1 | Feature 2 | Feature 3 | … | Feature m-1 | Feature m |
|---|---|---|---|---|---|---|
| Customer 1 | … | … | … | … | … | … |
| Customer 2 | … | … | … | … | … | … |
| Customer 3 | … | | | | | … |
| … | … | | | | | … |
| … | … | | | | | … |
| … | … | | | | | … |
| … | … | | | | | … |
| … | … | | | | | … |
| … | … | | | | | … |
| … | … | | | | | … |
| Customer n-1 | … | … | … | … | … | … |
| Customer n | … | … | … | … | … | … |

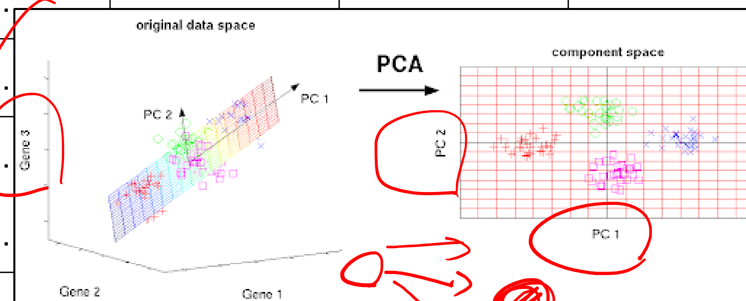


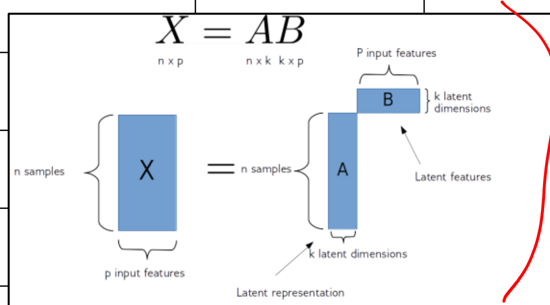
# Service Quality "Representation"

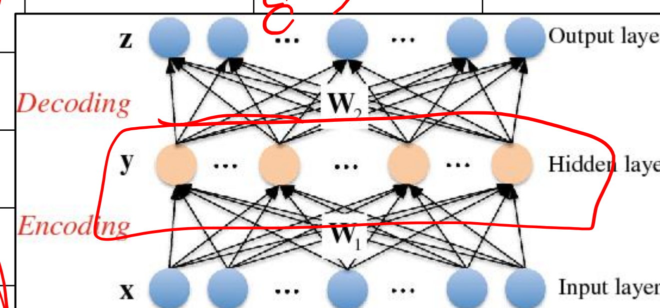
$PC_1 = \lambda_{11} x_1 + \lambda_2 x_2 +$

| | Feature 1 | Feature 2 | Feature 3 | … | Feature m-1 | Feature m |
|---|---|---|---|---|---|---|
| Customer 1 | | | … | … | … | … |
| Customer 2 | | | … | | | |
| Customer 3 | | | … | | | |
| … | | | … | | | |
| … | | | … | | | |
| … | | | … | … | … | … |
| … | … | … | … | … | | … |
| … | | | | | | |
| … | | | | | | |
| … | | | | | | |
| Customer n-1 | | | | | | |
| Customer n | … | … | … | … | … | … |



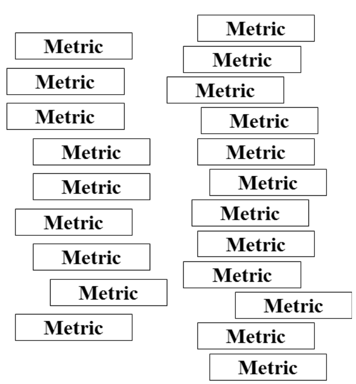



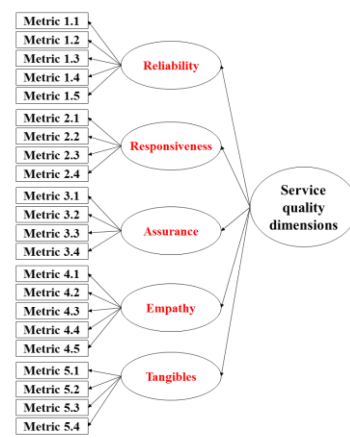
# Data Collection and Purification ($1^{st}$ & $2^{nd}$ stages)

**1st stage** | **2nd stage**

| 10 dimensions<br>A set of 97 items | 10 dimensions<br>A set of 54 items | 7 dimensions<br>A set of 34 items | 5 dimensions<br>A set of 22 items |
|---|---|---|---|

- tangibles
- reliability
- responsiveness
- communication
- credibility
- security
- competence
- courtesy
- understanding the customer
- access




- tangibles
- reliability*
- responsiveness*
- assurance
- empathy

# 1st Data Collection

10 dimensions
A set of 97 items

10 dimensions
A set of 54 items

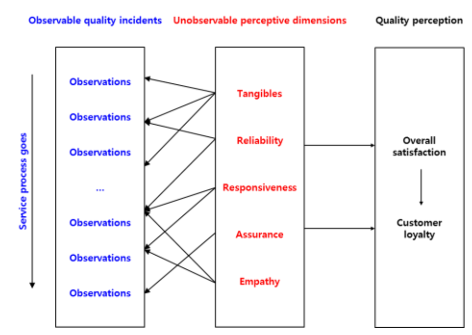
7 dimensions
A set of 34 items

5 dimensions
A set of 22 items

- 200 adult respondents in shopping mall
  - Used the services of the firm within the past three months

- Five different service categories
  - Appliance repair and maintenance
  - retail banking
  - long-distance telephone
  - securities brokerage
  - credit card

tangibles
reliability
responsiveness
communication
credibility
security
competence
courtesy
understanding the customer
access



# 1st Purification with the Reliability Analysis

- Coefficient alpha

10 dimensions
A set of 97 items

.55 ~ .78

→

10 dimensions
A set of 54 items

.72 ~ .83

→

7 dimensions
A set of 34 items

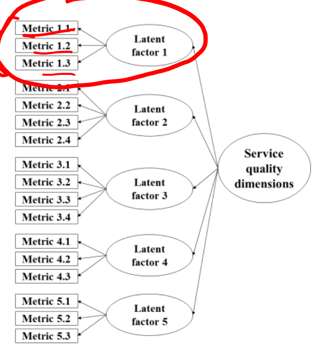
→

5 dimensions
A set of 22 items

- tangibles
- reliability
- responsiveness
- communication
- credibility
- security
- competence
- courtesy
- understanding the customer
- access

## Coefficient alpha (Cronbach's alpha) analysis

- Estimate of reliability and an indicator of internal consistency
- Deletion of certain items based on the corrected item-to-total correlations
- .55 ~ .78 → .72 ~.83 across 10 dimensions
- > .70 reliable variance




$$\rho_T = \frac{k^2\,\overline{\sigma_{ij}}}{\sigma_X^2}$$

$k$ = number of items

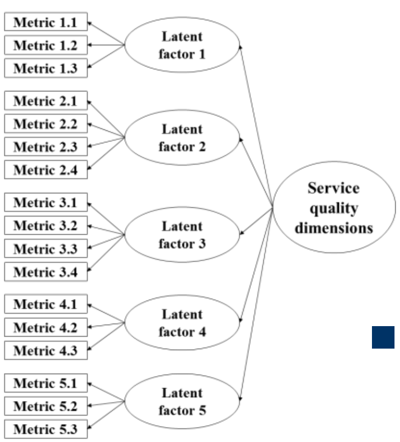
$\sigma_{ij}$ = covariance between Xi and Xj

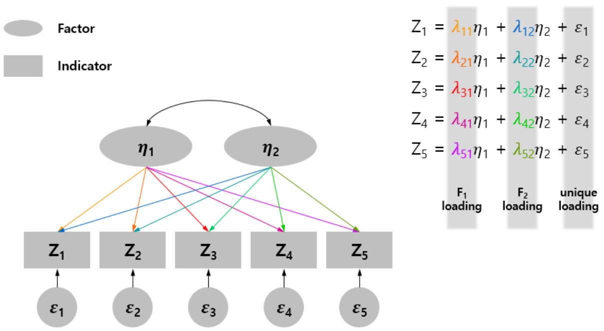
$\sigma_X^2$ = item variances and inter-item covariances

# 1st Purification with the 1st Factor Analysis

-Factor analysis
-Factor loading matrix

10 dimensions A set of 97 items → **10 dimensions A set of 54 items** → **7 dimensions A set of 34 items** → 5 dimensions A set of 22 items




- Factor analysis
  - Analyze interrelations among the dimensions and facilitate easy interpretation
  - 10 factors were assumed in this study initially
- Factor loading matrix interpretation
  - Items with high loadings on more than one factor → removed or integrated
  - Relatively meaningless factors → removed

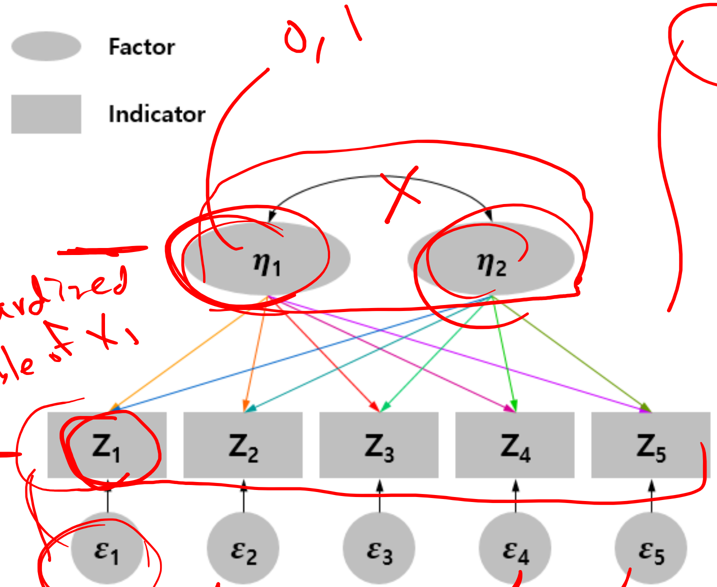
# Factor Analysis

- Statistical method to analyze observable variables in understanding a specific topic
- To identify the underling constructs (i.e., to represent latent variables)

$Z = \Lambda H + E$

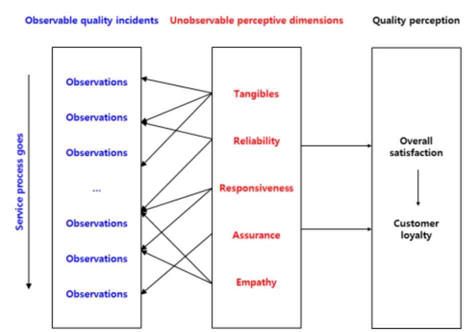
0, 1

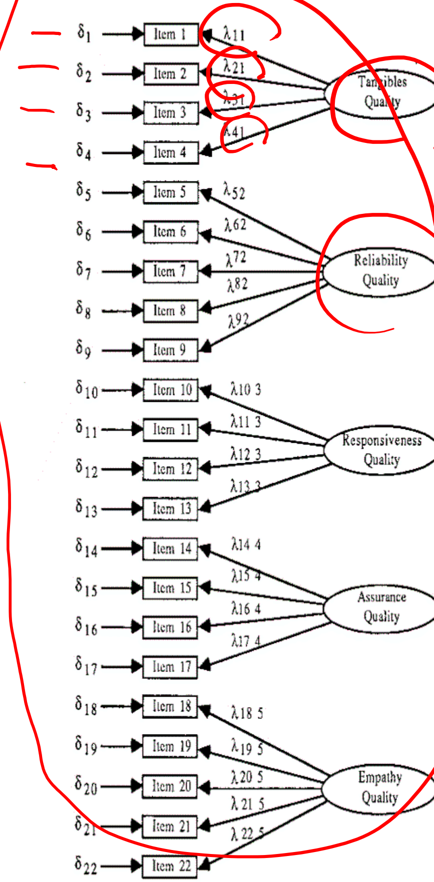
standardized variable of $X_1$




$$
\begin{aligned}
Z_1 &= \lambda_{11}\eta_1 + \lambda_{12}\eta_2 + \varepsilon_1 \\
Z_2 &= \lambda_{21}\eta_1 + \lambda_{22}\eta_2 + \varepsilon_2 \\
Z_3 &= \lambda_{31}\eta_1 + \lambda_{32}\eta_2 + \varepsilon_3 \\
Z_4 &= \lambda_{41}\eta_1 + \lambda_{42}\eta_2 + \varepsilon_4 \\
Z_5 &= \lambda_{51}\eta_1 + \lambda_{52}\eta_2 + \varepsilon_5
\end{aligned}
$$

$F_1$ loading | $F_2$ loading | unique loading

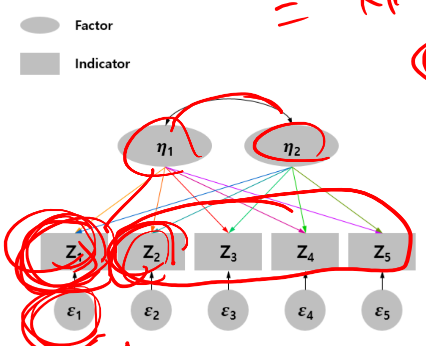


# Factor Analysis

$Var(Z_1) = Var(\lambda_{11}\eta_1) + Var(\lambda_{12}\eta_2) + Var(\varepsilon_1)$

$= \lambda_{11}^2 + \lambda_{12}^2 + \psi_1$




$$
\begin{aligned}
Z_1 &= \lambda_{11}\eta_1 + \lambda_{12}\eta_2 + \varepsilon_1 \\
Z_2 &= \lambda_{21}\eta_1 + \lambda_{22}\eta_2 + \varepsilon_2 \\
Z_3 &= \lambda_{31}\eta_1 + \lambda_{32}\eta_2 + \varepsilon_3 \\
Z_4 &= \lambda_{41}\eta_1 + \lambda_{42}\eta_2 + \varepsilon_4 \\
Z_5 &= \lambda_{51}\eta_1 + \lambda_{52}\eta_2 + \varepsilon_5
\end{aligned}
$$

$F_1$ loading | $F_2$ loading | unique loading

variance function ( $Z = \Lambda H + E$ )

$$\Sigma = \Lambda\Phi\Lambda' + \Psi$$

R

$$= \begin{pmatrix} Var(Z_1) & Cov(Z_1,Z_2) & \cdots \\ Cov(Z_2,Z_1) & Var(Z_2) & \\ & & \ddots \end{pmatrix} = \begin{pmatrix} \lambda_{11}^2 + \lambda_{12}^2 + \psi_1 & \lambda_{11}\lambda_{21} + \lambda_{12}\lambda_{22} & \cdots \\ \lambda_{21}\lambda_{11} + \lambda_{22}\lambda_{12} & & \ddots \end{pmatrix}$$

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$R - \Psi = \Lambda\Phi\Lambda^T$

$= \Lambda M M^T \Lambda^T$

$= \Lambda M (\Lambda M)^T$

$1 - \psi_1$, $1 - \psi_2$

$1 - \psi \approx f(z_2, z_3, z_4, z_5)$

$\sim \beta_0 + \beta_2 z_2 + \beta_3 z_3$

$R^2$

# 2nd Data Collection

| 10 dimensions<br>A set of 97 items | → | 10 dimensions<br>A set of 54 items | → | 7 dimensions<br>A set of 34 items | → | 5 dimensions<br>A set of 22 items |
|---|---|---|---|---|---|---|

- 200 adult respondents in shopping mall
  - Used the services of the firm within the past three months

- Four different service categories
  - Appliance repair and maintenance
  - Retail banking
  - Long-distance telephone
  - Credit card

# 2nd Data Collection

-Factor analysis
-Factor loading matrix

10 dimensions
A set of 97 items

10 dimensions
A set of 54 items

7 dimensions
A set of 34 items

| Dimension | Label |
|---|---|
| Tangibles | D1 |
| Reliability | D2 |
| Responsiveness | D3 |
| Communication Credibility Security Competence Courtesy | D4 D5 |
| Understanding/ Knowing Customers | D6 |
| Access | D7 |

- Two differences emerged
  - (1) Corrected item-to-total correlations for several items were lower and the alpha for the dimensions were lower (items making up the dimensions D4 and D7)
  - (2) Factor-loading matrices showed much greater overlap between dimensions D4 and D5, between dimensions D6 and D7
- Further purification of 34-item scale was necessary

# 2nd Purification after the 2nd Factor Analysis

10 dimensions
A set of 97 items
→
10 dimensions
A set of 54 items
→
7 dimensions
A set of 34 items
→
5 dimensions
A set of 22 items

tangibles
reliability*
responsiveness*
assurance
empathy

- 2nd stage of purifications
  - Items with low item-to-total correlations were deleted
  - As suggested by the second factor analysis, the items remaining in D4&D5 and D6&D7 were combined to form two separate dimensions
- This procedure resulted in SERVQUAL with 22 items spread among five dimensions

# 5 Dimensions and 22 Items of SERVQUAL

| Dimension | Label | Number of Items | Reliability Coefficients (Alphas) | Items | Factor Loadings of Items on Dimensions to Which They Belong[a] |
|---|---|---|---|---|---|
| Tangibles | F1 | 4 | .72 | Q1 | 69 |
| | | | | Q2 | 68 |
| | | | | Q3 | 64 |
| | | | | Q4 | 51 |
| Reliability | F2 | 5 | .83 | Q5 | 75 |
| | | | | Q6 | 63 |
| | | | | Q7 | 71 |
| | | | | Q8 | 75 |
| | | | | Q9 | 50 |
| Responsiveness | F3 | 4 | .82 | Q10 | 51 |
| | | | | Q11 | 77 |
| | | | | Q12 | 66 |
| | | | | Q13 | 86 |
| Assurance | F4 | 4 | .81 | Q14 | 38 |
| | | | | Q15 | 72 |
| | | | | Q16 | 80 |
| | | | | Q17 | 45 |
| Empathy | F5 | 5 | .86 | Q18 | 78 |
| | | | | Q19 | 81 |
| | | | | Q20 | 59 |
| | | | | Q21 | 71 |
| | | | | Q22 | 68 |

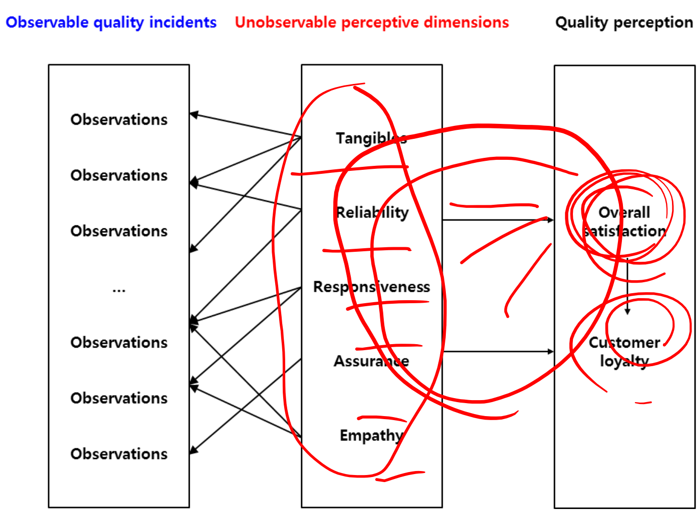
FACTOR LOADINGS

| Items | Bank | | | | | Credit Card Co. | | | | | Repair & Maintenance Co. | | | | | L-D Telephone Co. | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | F1 | F2 | F3 | F4 | F5 | F1 | F2 | F3 | F4 | F5 | F1 | F2 | F3 | F4 | F5 | F1 | F2 | F3 | F4 | F5 |
| Q1 | 34 | 28 | — | — | — | 36 | — | 35 | — | — | 34 | — | — | — | — | 42 | — | — | — | — |
| Q2 | 64 | — | — | — | — | 70 | — | — | — | — | 70 | — | — | — | — | 72 | — | — | — | — |
| Q3 | 39 | — | — | 28 | — | 52 | — | — | — | — | 53 | — | — | — | — | 51 | — | — | — | — |
| Q4 | 28 | — | — | 28 | — | 52 | — | — | — | — | 65 | — | — | — | — | 59 | — | — | 30 | — |
| Q5 | — | 72 | — | — | — | — | 54 | — | — | — | — | 73 | — | — | — | — | 52 | — | — | — |
| Q6 | — | 63 | — | — | — | — | 43 | 27 | — | — | — | 51 | — | — | — | — | 40 | — | — | — |
| Q7 | — | 71 | — | — | — | — | 87 | — | — | — | — | 84 | — | — | — | — | 79 | — | — | — |
| Q8 | — | 80 | — | — | — | — | 83 | — | — | — | — | 88 | — | — | — | — | 59 | — | — | — |
| Q9 | — | 39 | — | — | — | — | 49 | — | — | — | — | 29 | — | 30 | — | — | 54 | — | — | — |
| Q10 | — | — | 37 | — | — | — | — | 43 | — | 26 | — | — | 56 | — | — | — | — | 39 | — | — |
| Q11 | — | — | 55 | — | — | — | — | 48 | — | — | — | — | 52 | — | — | — | — | 43 | — | — |
| Q12 | — | — | 62 | — | — | — | — | 54 | — | — | — | — | 74 | — | — | — | — | 92 | — | — |
| Q13 | — | — | 69 | — | — | — | — | 33 | — | — | — | — | 71 | — | — | — | — | 53 | — | — |
| Q14 | — | — | — | 68 | — | — | — | — | 65 | — | — | — | — | 86 | — | — | — | — | 69 | — |
| Q15 | — | — | — | 84 | — | — | — | — | 76 | — | — | — | — | 89 | — | — | — | — | 81 | — |
| Q16 | — | — | — | 72 | — | — | — | — | 73 | — | — | — | — | 65 | — | — | — | — | 61 | — |
| Q17 | — | — | — | 64 | — | — | — | — | 61 | — | — | — | — | 64 | — | — | — | — | 66 | — |
| Q18 | — | — | — | — | 37 | — | — | — | — | 64 | — | — | — | — | 42 | — | — | — | — | 59 |
| Q19 | — | — | — | — | 48 | — | — | — | — | 72 | — | — | — | — | 61 | — | — | — | — | 79 |
| Q20 | — | — | — | — | 41 | — | — | — | — | 63 | — | 28 | 34 | — | 46 | — | — | — | — | 55 |
| Q21 | — | — | — | — | 33 | — | — | — | — | 59 | — | — | — | — | 32 | — | — | — | — | 36 |
| Q22 | — | — | — | — | 68 | — | — | — | — | 64 | — | — | — | — | 61 | — | — | — | — | 59 |

[a] All numbers in the table are magnitudes of the factor loadings multiplied by 100. Loadings that are .25 or less are not shown. The percentage of variance extracted by the five factors in the bank, credit card, repair and maintenance, and long-distance telephone samples were 56.0%, 57.5%, 61.6%, and 56.2%, respectively.

# 5 Dimensions and 22 Items of SERVQUAL

- Relative importance of the five dimensions in validation
  - 1st : **Reliability**
  - 2nd : **Assurance**
  - 5th **: Empathy** (the least important)




**Relative Importance of the Five Dimensions in Predicting Overall Quality**

| Dimension | Standardized Slope Coefficient | Significance Level of Slope[a] | Adusted $R^2$ |
|---|---|---|---|
| Bank | | | |
| Tangibles | .13 | .07 | .28 ($p < .00$) |
| Reliability | .39 | .00 | |
| Responsiveness | .07 | [illegible] | |
| Assurance | .13 | .09 | |
| Empathy | .01 | .89 | |
| Credit Card Co. | | | |
| Tangibles | .07 | .26 | .27 ($p < .00$) |
| Reliability | .33 | .00 | |
| Responsiveness | .12 | .11 | |
| Assurance | .17 | .02 | |
| Empathy | .04 | .58 | |
| Repair & Maintenance Co. | | | |
| Tangibles | .04 | .48 | .52 ($p < .00$) |
| Reliability | .54 | .00 | |
| Responsiveness | .11 | .09 | |
| Assurance | .16 | .02 | |
| Empathy | .01 | .81 | |
| L-D Telephone Co. | | | |
| Tangibles | .08 | .17 | .37 ($p < .00$) |
| Reliability | .45 | .00 | |
| Responsiveness | .12 | .09 | |
| Assurance | .15 | .03 | |
| Empathy | .02 | .78 | |

[a] Significance levels are for two-tailed tests.

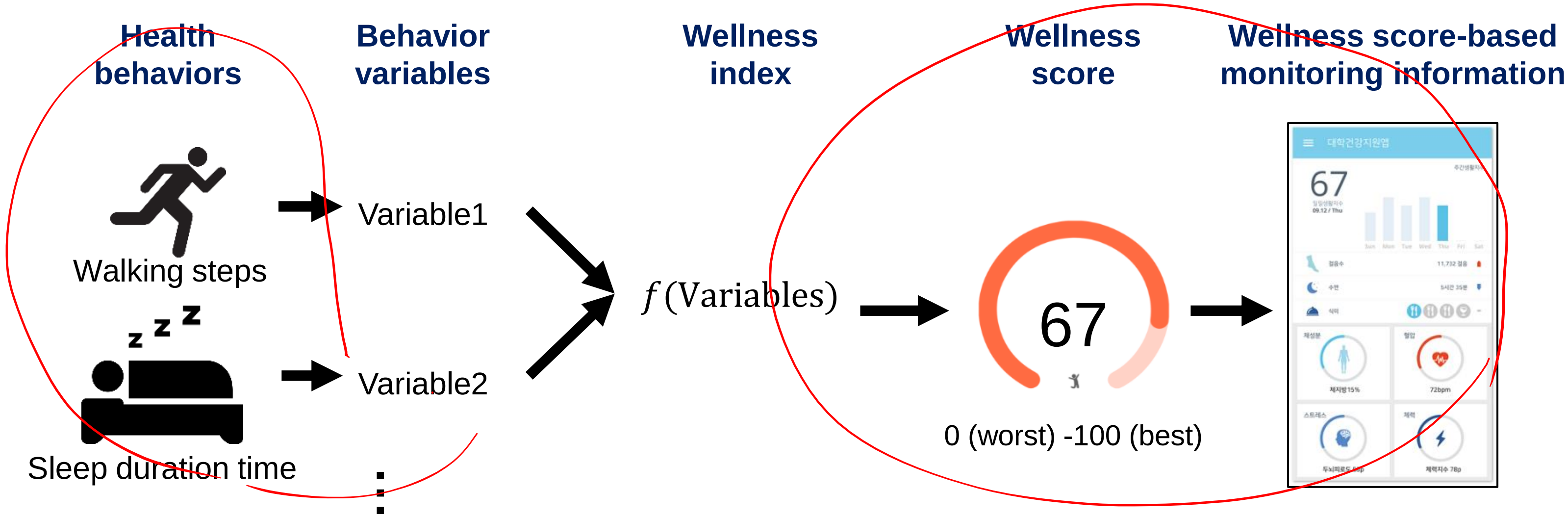
# Assignment 3 (by 09.23 11:59 pm)

- Read Article 3. Summarize the article (one paragraph) with your own comments to the article (one paragraph) (i.e., two paragraphs in total).
- By yourself, complete the identification of the latent factors indicating the quality of the Onecare mobile healthcare service based on the practice demonstrated by the TA. Use the user-question matrix data provided. Do it all by yourself, and describe the identification process and outcome in detail. Interpret the outcome (i.e., interpret the quality dimensions you identified).
- Discuss the “quality representation and measurement of service systems from a customer/user perspective” (focus on the service system you are interested or concerned). What other data and learning methods can be used for the identification of the latent factors indicating the quality of a service? Describe your thoughts/ideas on learning service quality dimensions with data about customer perception, behaviors, etc., in detail.
- What dimensions do you think we should consider for the evaluation of quality of AI-based services? i.e., As a user/customer of AI services (or as an undergraduate student researcher), what are the requirements of AI service you think important? And why do you think so? Describe the rationale or reasons for your suggestion.
- Furthermore, assume that you actually need to represent and evaluate the quality of an AI-based service in your company or institute. How would you develop a quality representation and measurement method for the service? What AI-based service are you going to focus on? What kinds of data and methods are you going to collect, analyze, and learn? Describe your research plan in detail. If possible, visualize your research framework clearly (e.g., draw image, mathematical model).
- Upload your code and a several paragraph essay in the Blackboard.

# Read Article 3: Onecare Service Overview

- Necessity of wellness management support for college students
  - Wellness means lifestyle to maintain good mental and physical health (Myers et al. 2000)
  - Many students have unhealthy activity, sleep, and diet behaviors (Small et al. 2012)
  - Students need overall support to manage various behaviors

- Concept of Onecare




Source: Kim, K., Kim, K., Lim, C., and Heo, J., "Development of a Lifelogs-based Daily Wellness Score to Advance a Smart Wellness Service," *Service Science*, Vol. 10, No. 4, 408–422, 2018.

# Read Article 3: Onecare Service Overview

## Onecare pilot run

191 college students → Use Onecare over 3 weeks → Respond to a survey → 191 survey responses

**Quality dimensions of mHealth services**

| | |
| --- | --- |
| Content quality | Quality of provided service content (Stoyanov et al. 2015) |
| Engagement | Ability to increase user engagement with a service (Platt et al. 2016) |
| Privacy | Ability to protect the privacy of users (Akter et al. 2013) |
| Reliability | Ability to perform the promised service dependably and accurately (Akter et al. 2013) |
| Usability | Ease of service use (Zapata et al. 2015) |
| Continuance Intention (CI) | User intention to continuously use (Limayem et al. 2007) |



Source: Kim, K., Kim, K., Lee, D., and Kim, M., "Identification of Critical Quality Dimensions for mHealth Services: A Case Study of Onecare Service," International Journal of Information Management, Vol. 46, 187–197, 2019.

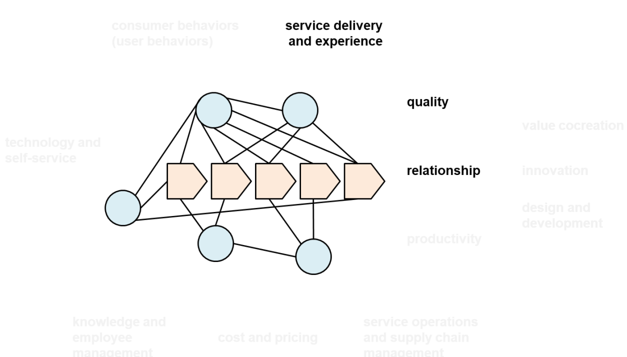
# Revisiting the Questions and Learning Objectives

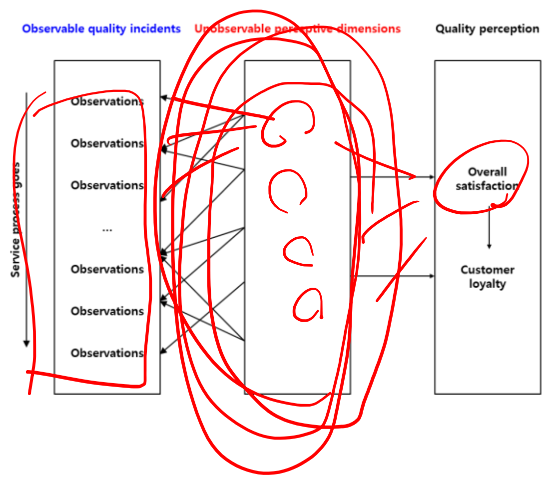
- How can we “explicitly” evaluate and improve the “implicit” quality of services?

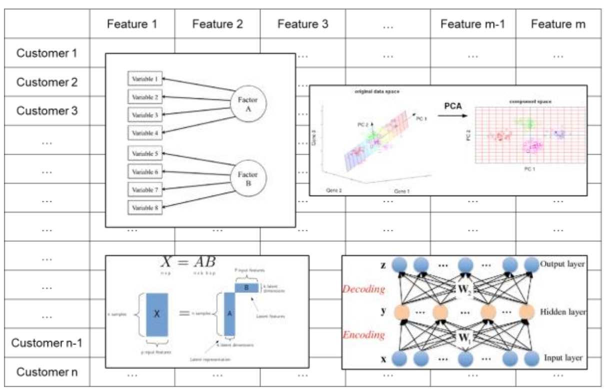
  → The service representation (quality dimension identification) is essential

- How can we use data and learning intelligence for service quality representation?

  → Observed data are varied. Estimations can be different depending upon the assumptions and models

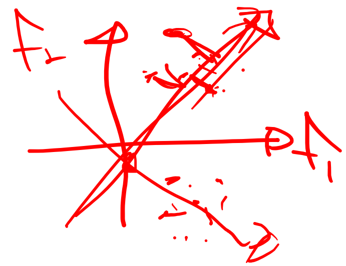


# Discussion

- Think about
  - Fundamental reasons to use the latent factors rather than the observed variables
  - Directions from the factors to the observed variables
  - Implications of the unique loadings about the reliability of observations (i.e., raw data collection)
  - Factor Analysis vs. Principal Component Analysis
  - Use of covariance matrix vs. Use of correlation matrix
  - Estimation of the parameters depending upon the model assumptions
  - Pros and Cons coming from the ambiguity of factor rotations (i.e., $R = \Lambda\Lambda' + \Psi = (\Lambda M)(\Lambda M)' + \Psi$)
  - Validation of the factor analysis outcomes (i.e., Validity of the identified factors)
  - …



# Further Readings Recommended

- Parasuraman, A., Zeithaml, V. A., & Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *1988*, *64*(1), 12-40. Article 4

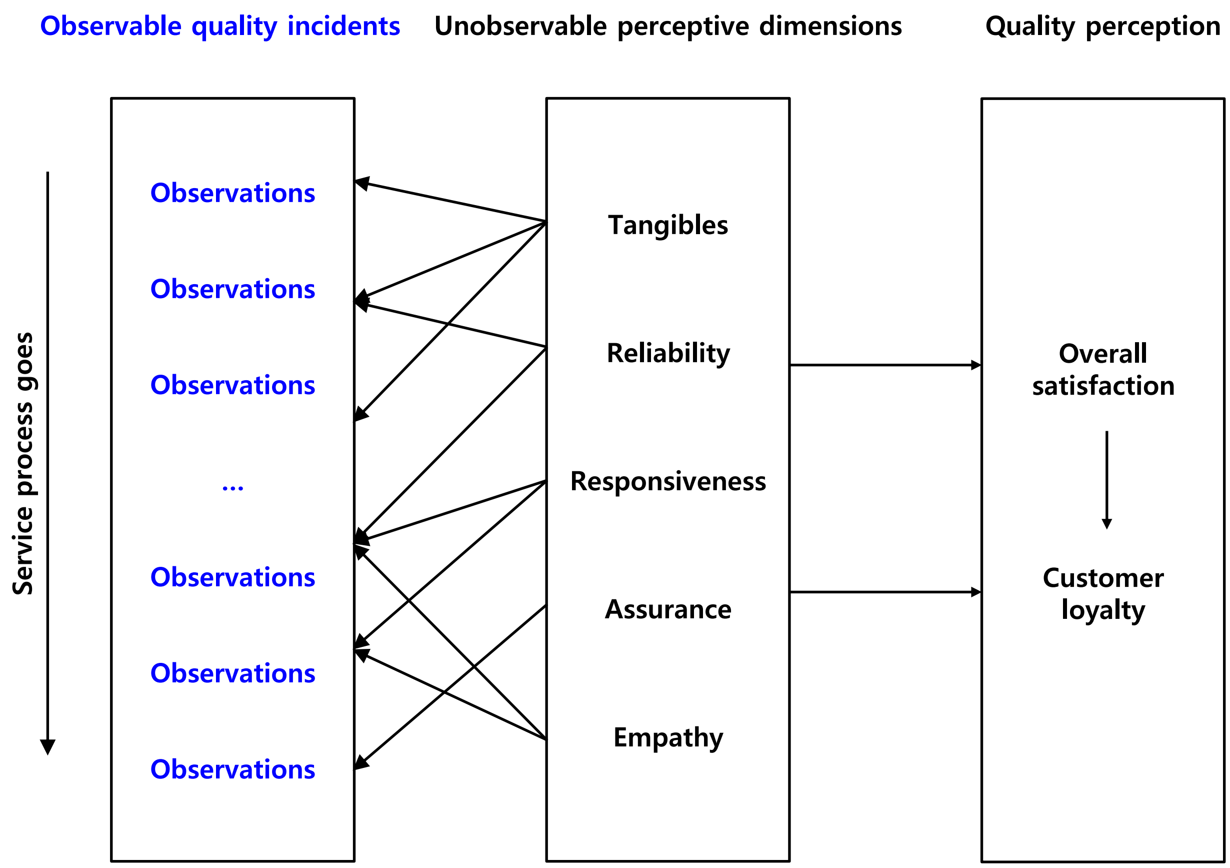
- Mejia, J., Mankad, S., & Gopal, A. (2020). Service quality using text mining: Measurement and consequences. *Manufacturing & Service Operations Management*.
  Due to the limited access, Article 5 is uploaded in the Blackboard

# Service Intelligence Week 4.

# [Service Review Mining for Service Improvement]

# Illustration of Service Quality Evaluation

- What other observations are available these days?



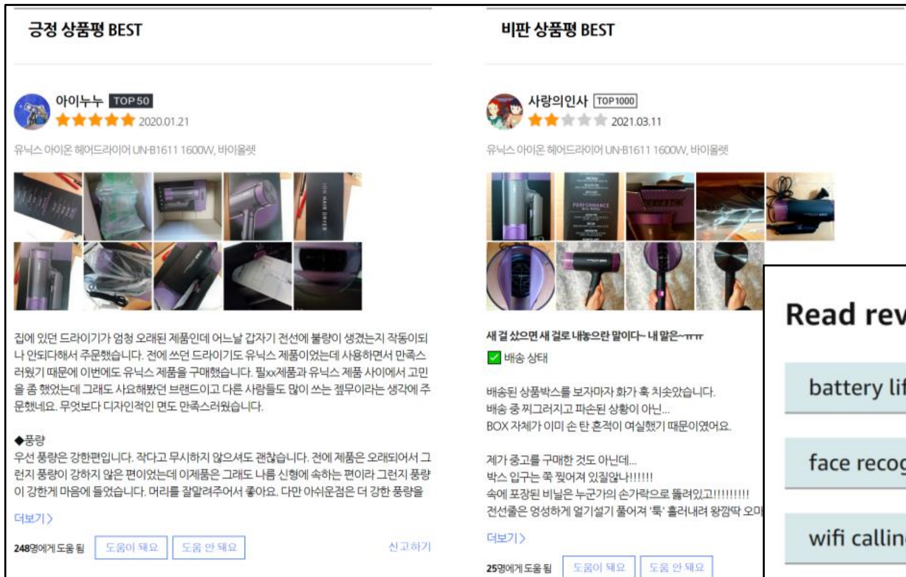
# Online Review Mining for Service Improvement

- What other observations are available these days?

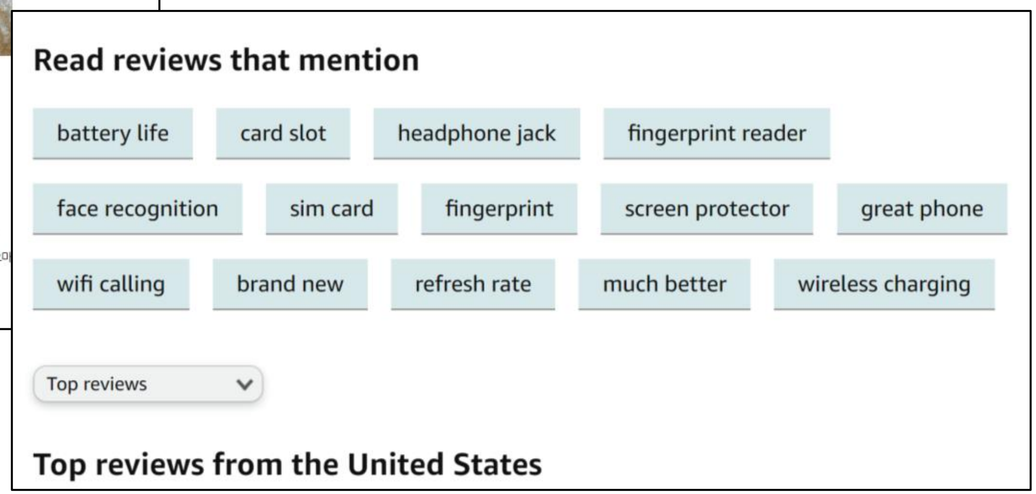

vs

# What is Online Review Mining?

- Review analysis for customers
  - Best and worst review retrieval, top keywords, and keyword-based search, etc



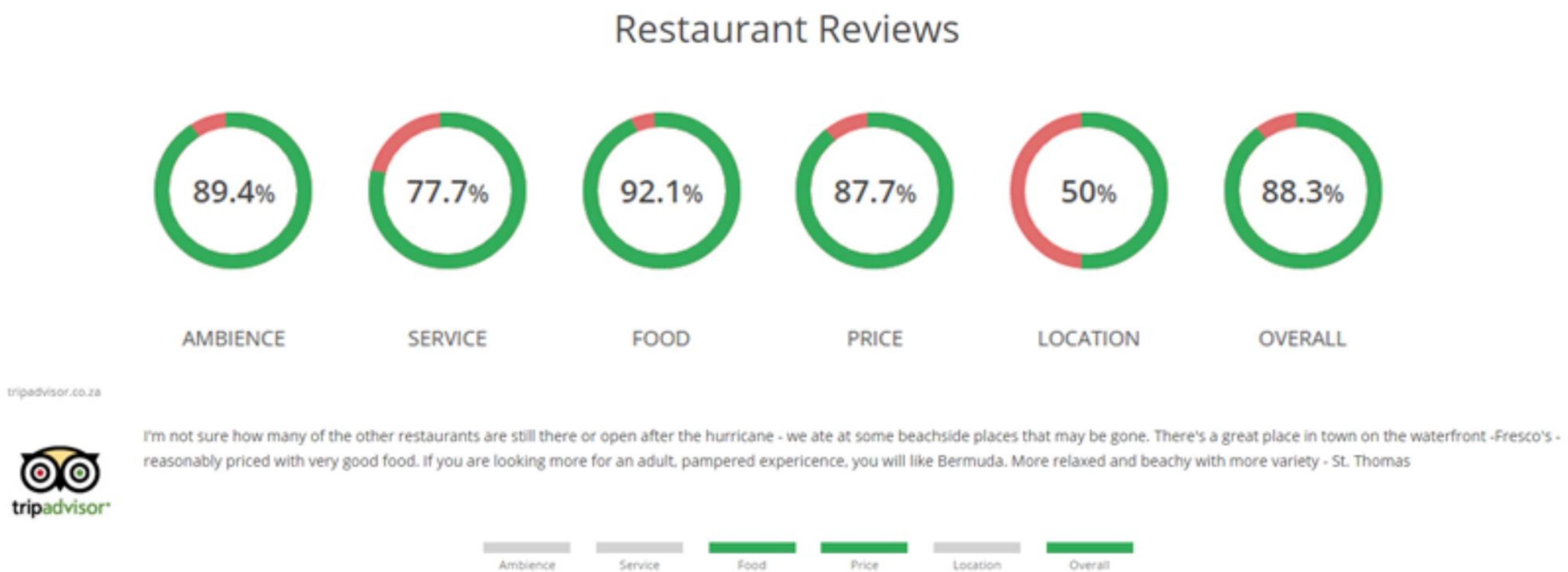

Best and worst reviews in Coupang




Top keywords and keyword-based search in Amazon

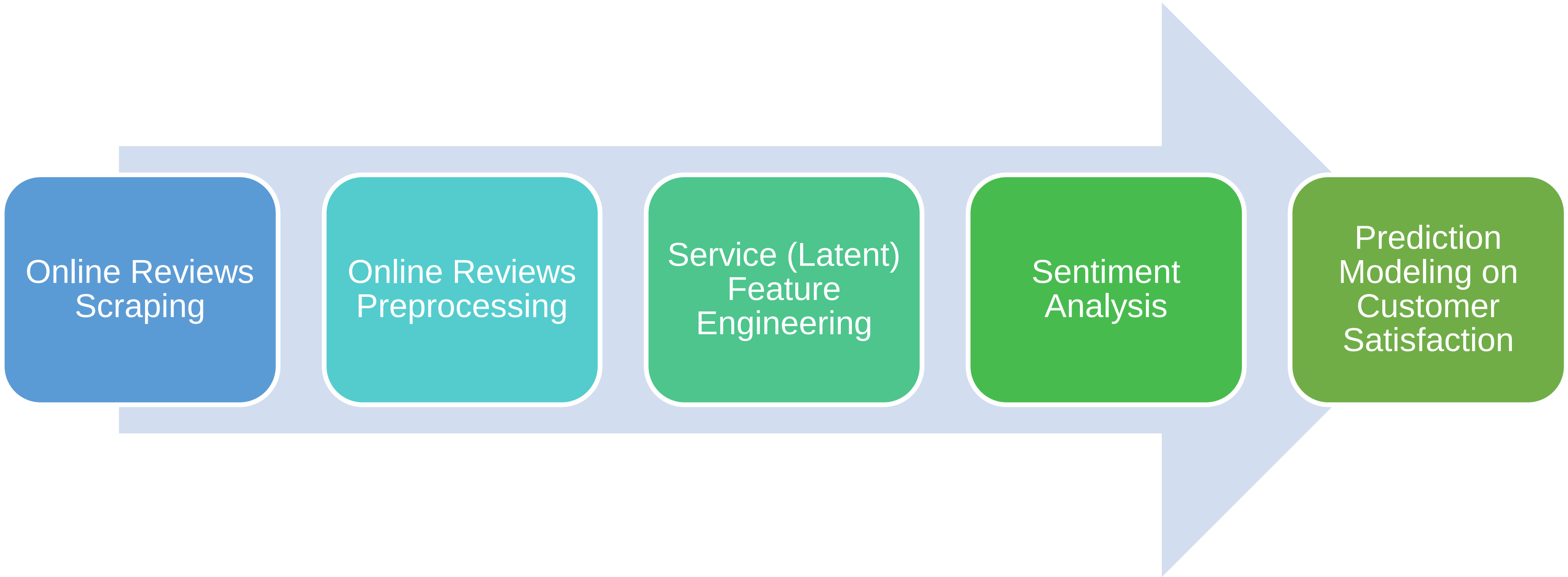
# What is Online Review Mining?

- Review analysis for companies
  - Topic analysis, sentiment analysis (i.e., performance analysis), and importance analysis, etc




Restaurant reviews analysis

# Online Review Mining Framework for Service Improvement



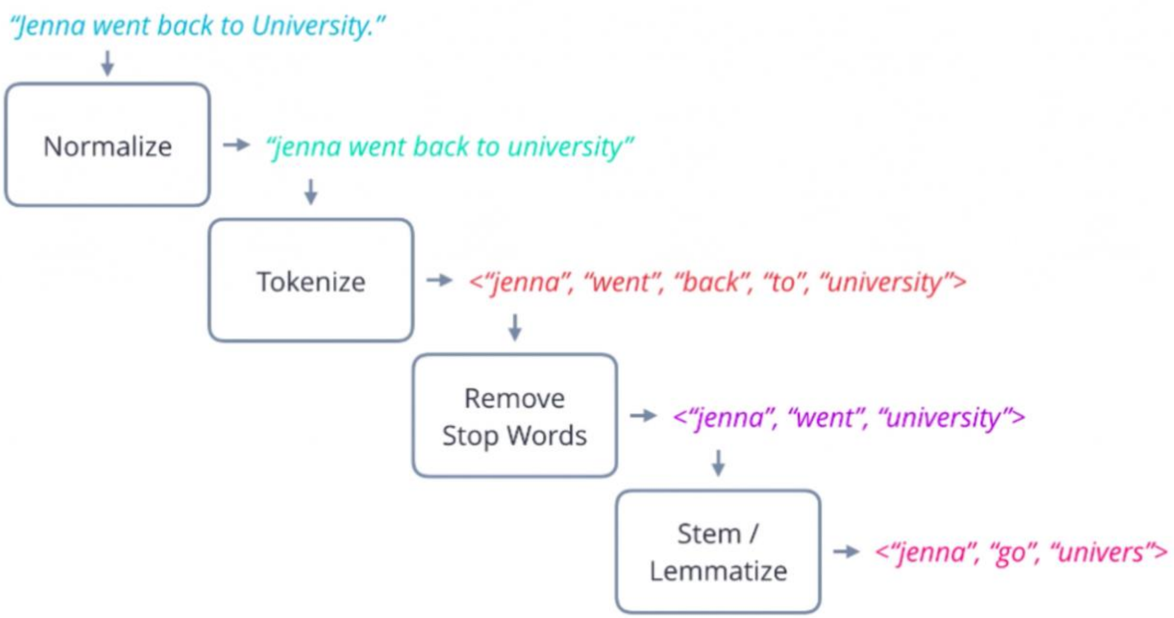
# Online Reviews Scraping

- What is online reviews scraping?
  - Web scraper goes to review web platforms or websites and extracts all the data that interests you


Web scraping

# Online Reviews Preprocessing

- What is text preprocessing?
  - Method to clean the text data and make it ready to feed data to the model
    - Lowercasing
    - Tokenization
    - Removal of stop words
    - Lemmatization
    - POS (Part-Of-Speech) tagging



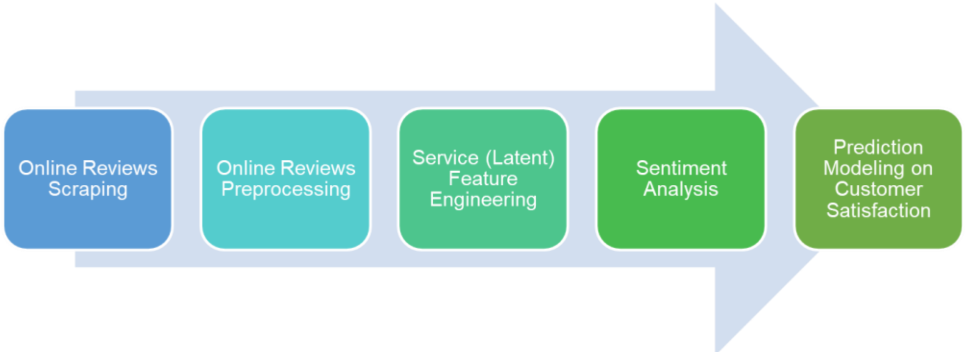
# On the Customer Review – Word Feature Matrix

| | **Word 1** | **Word 2** | **Word 3** | **…** | **Word m-1** | **Word m** | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | 1 | 2 | 0 | 3 | 4 | 0 | 5 |
| Review 2 | - | - | - | - | - | - | 5 |
| Review 3 | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 3 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 1 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 5 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 3 |
| Review n-1 | - | - | - | - | - | - | 2 |
| Review n | - | - | - | - | - | - | 4 |

# On the Customer Review – Service Feature Matrix

| | **Feature 1** | **Feature 2** | **Feature 3** | **…** | **Feature k-1** | **Feature k** | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | ? | ? | ? | ? | ? | ? | 5 |
| Review 2 | ? | ? | ? | ? | ? | ? | 5 |
| Review 3 | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 1 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 5 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| Review n-1 | ? | ? | ? | ? | ? | ? | 2 |
| Review n | ? | ? | ? | ? | ? | ? | 4 |

# Sentiment Analysis and Satisfaction Prediction Modeling




| | Word 1 | Word 2 | Word 3 | … | Word m-1 | Word m | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | 1 | 2 | 0 | 3 | 4 | 0 | 5 |
| Review 2 | - | - | - | - | - | - | 5 |
| Review 3 | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 3 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 1 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 5 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 3 |
| Review n-1 | - | - | - | - | - | - | 2 |
| Review n | - | - | - | - | - | - | 4 |

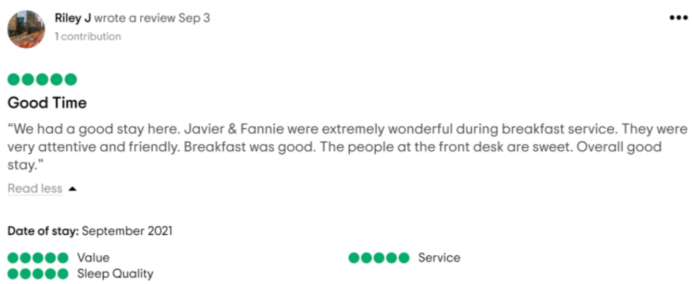
| | Feature 1 | Feature 2 | Feature 3 | … | Feature k-1 | Feature k | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | ? | ? | ? | ? | ? | ? | 5 |
| Review 2 | ? | ? | ? | ? | ? | ? | 5 |
| Review 3 | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 1 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 5 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| Review n-1 | ? | ? | ? | ? | ? | ? | 2 |
| Review n | ? | ? | ? | ? | ? | ? | 4 |

# Case Study

- Hotels in Singapore
  - Hotels requiring improvement due to low customer satisfaction
  - Time: 2010.01-2019.12
  - Customer reviews in 2020 were not considered due to Covid-19

# Case Study: Online Reviews Scraping and Preprocessing

- Remove duplicate reviews
- Collect reviews with the information of customer segments and hotel classes
- Selenium web crawler library of Python was used to collect customer reviews from TripAdvisor
- NLTK package of Python was used to structure each review into preprocessed word tokens

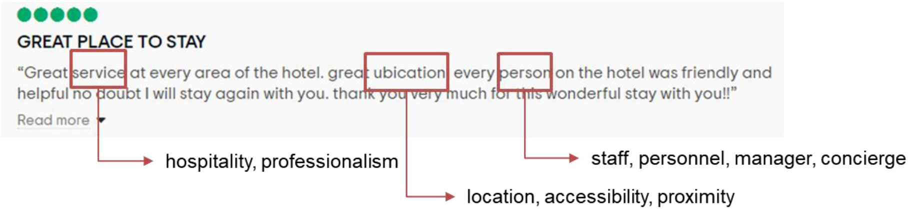



Collected data information based on customer segments

| Information | All | Hotel classes | | | Travel types | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Class 2 | Class 3 | Class 4 | Solo | Couple | Friends | Family | Business |
| Sample | 32,044 | 4,571 | 12,528 | 14,945 | 3,318 | 9,853 | 3,593 | 7,812 | 4,892 |
| Ratio | 0.56 | 0.51 | 0.54 | 0.60 | 0.59 | 0.56 | 0.58 | 0.60 | 0.50 |

# Service Feature Engineering

- What is service feature engineering in online review mining?





**Review*word matrix**

| | $word_1$ | … | $word_m$ |
|---|---|---|---|
| $Review_1$ | value | | |
| … | | | |
| $Review_n$ | | | |

Dimension reduction

How to estimate the value?

**Review*feature matrix**

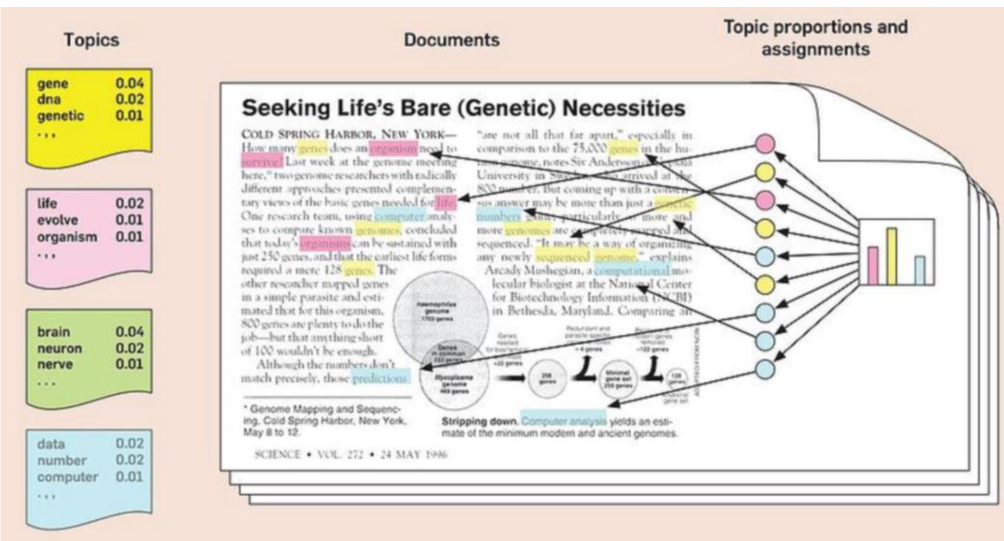
| | $feature_1$ | … | $feature_i$ |
|---|---|---|---|
| $Review_1$ | | | |
| … | | | |
| $Review_n$ | | | |

| | about | bird | heard | is | the | word | you |
|---|---|---|---|---|---|---|---|
| About the bird, the bird, bird bird bird | 1 | 5 | 0 | 0 | 2 | 0 | 0 |
| You heard about the bird | 1 | 1 | 1 | 0 | 1 | 0 | 1 |
| The bird is the word | 0 | 1 | 0 | 1 | 2 | 1 | 0 |

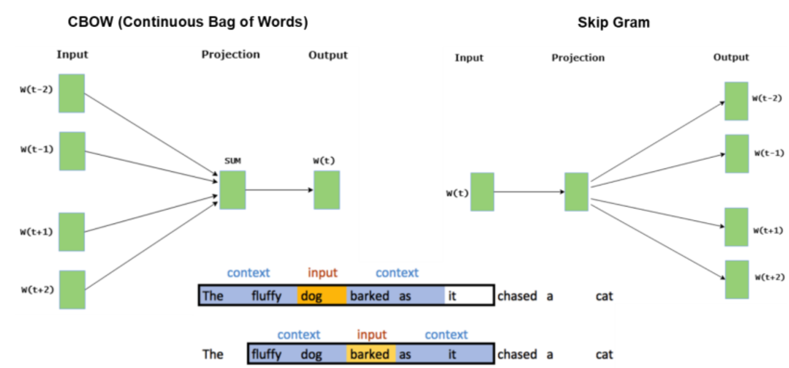
Example of hotel service features: location, breakfast, room, service, sleep quality, value, etc.

# Service Feature Engineering

- How to perform feature engineering from online reviews?
  - Latent topic feature extraction: LDA (Latent Dirichlet Allocation), NMF (Nonnegative Matrix Factorization)
  - Word embedding (vectorization): Word2Vec, BERT (Bidirectional Encoder Representations from Transformers)




Topic analysis




word embedding

# Service Feature Engineering

- NMF (Lee and Seung, 1999) and LDA (Blei et al, 2003)
  - Idea: each review can be described by a distribution of topics and each topic can be described by a distribution of words

**Review*word matrix (input)**

| | $word_1$ | $\dots$ | $word_m$ |
|---|---|---|---|
| $Review_1$ | | | |
| $\dots$ | | | |
| $Review_n$ | | | |

**Review*topic matrix (output)**

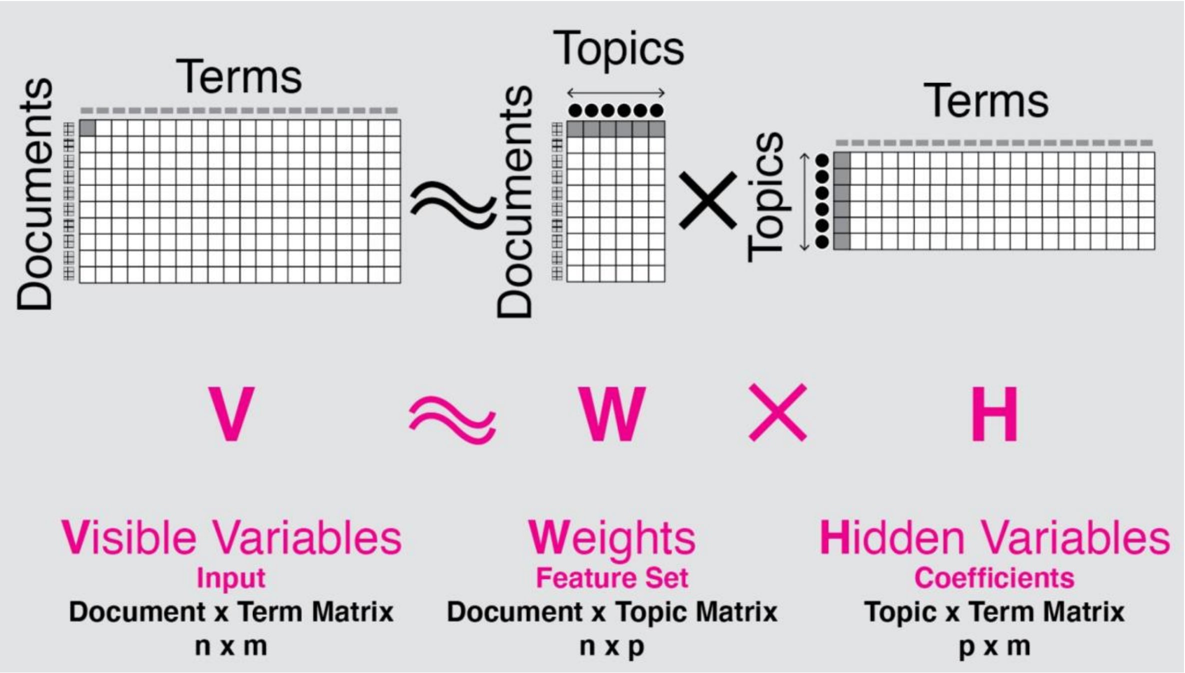
| | $Topic_1$ | $\dots$ | $Topic_k$ |
|---|---|---|---|
| $Review_1$ | | | |
| $\dots$ | | | |
| $Review_n$ | | | |

**Topic*keyword matrix (output)**

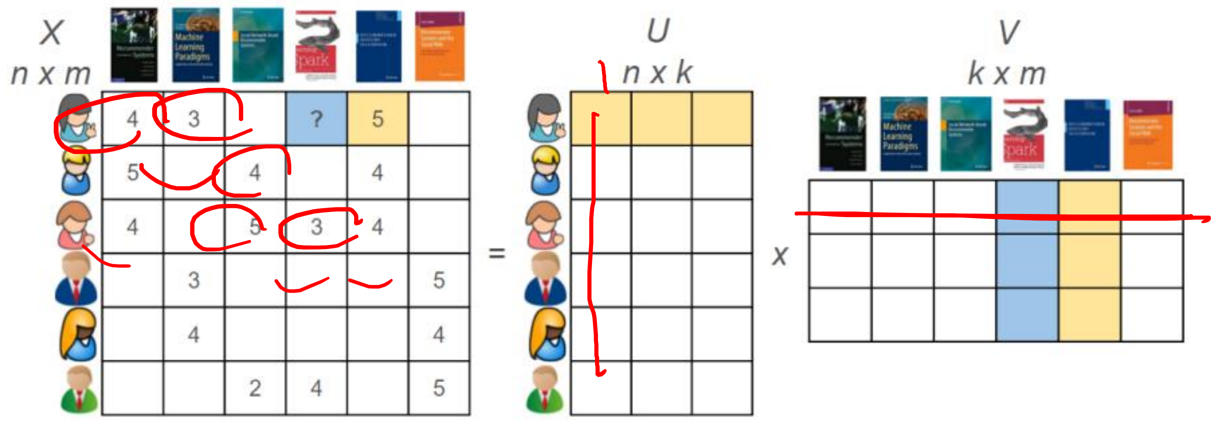
| | $word_1$ | $\dots$ | $word_m$ |
|---|---|---|---|
| $Topic_1$ | | | |
| $\dots$ | | | |
| $Topic_k$ | | | |

# Service Feature Engineering

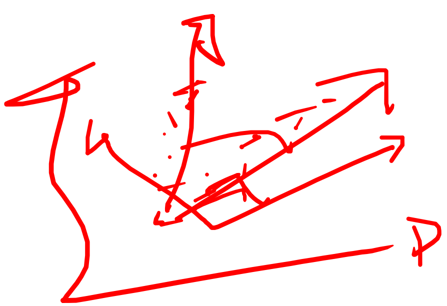
- NMF (Lee and Seung, 1999)
  - NMF is a linear-algeabreic model, that factors high-dimensional vector into a low-dimensionality representation similar to principal component analysis

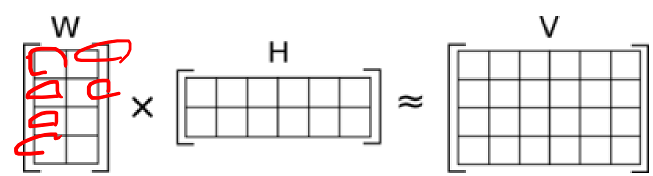


# Recall the Collaborative Filtering with Matrix Factorization




$$\|V - WH\|_F^2$$






initialize: $\mathbf{W}$ and $\mathbf{H}$ non negative.

Then update the values in $\mathbf{W}$ and $\mathbf{H}$ by computing the following, with $n$ as an index of the iteration.

$$\mathbf{H}^{n+1}_{[i,j]} \leftarrow \mathbf{H}^{n}_{[i,j]} \frac{((\mathbf{W}^n)^T\mathbf{V})_{[i,j]}}{((\mathbf{W}^n)^T\mathbf{W}^n\mathbf{H}^n)_{[i,j]}}$$

and

$$\mathbf{W}^{n+1}_{[i,j]} \leftarrow \mathbf{W}^{n}_{[i,j]} \frac{(\mathbf{V}(\mathbf{H}^{n+1})^T)_{[i,j]}}{(\mathbf{W}^n\mathbf{H}^{n+1}(\mathbf{H}^{n+1})^T)_{[i,j]}}$$

Until $\mathbf{W}$ and $\mathbf{H}$ are stable.

**Target rating matrix $R$**

| user \ item | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 5 | ? | 4 | ? | 3 | 3 |
| 2 | ? | 4 | ? | ? | 3 | ? |
| 3 | 3 | ? | ? | 3 | ? | ? |
| 4 | ? | 4 | ? | 5 | ? | 1 |

**User feature matrix $P$ (initial state)**

| user \ dimension | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 0.11 | 0.07 | 0.19 |
| 2 | 0.09 | 0.16 | 0.19 |
| 3 | 0.09 | 0.05 | 0.04 |
| 4 | 0.03 | 0.13 | 0.18 |

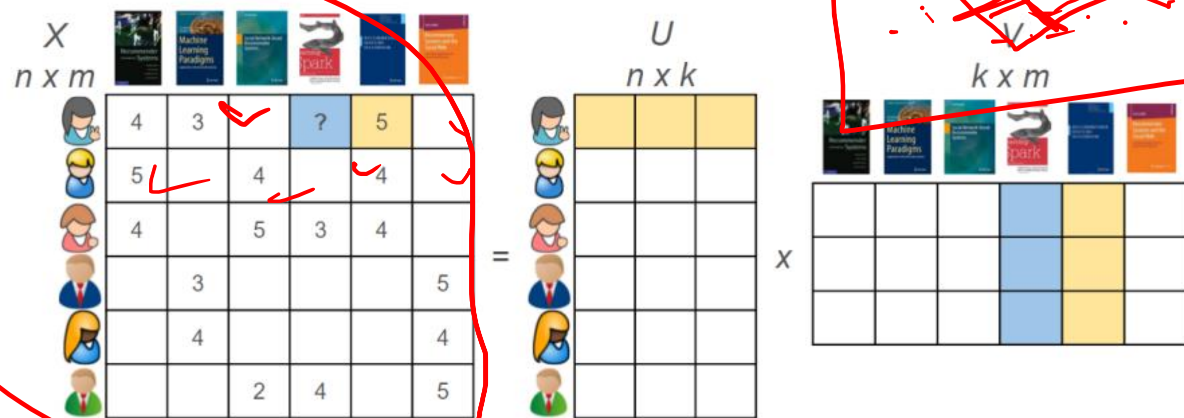
**Item feature matrix $Q$ (initial state)**

| dimension \ item | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 0.16 | 0.01 | 0.07 | 0.17 | 0.02 | 0.20 |
| 2 | 0.18 | 0.19 | 0.10 | 0.05 | 0.18 | 0.15 |
| 3 | 0.02 | 0.18 | 0.03 | 0.14 | 0.17 | 0.06 |



Image source: https://medium.com/@lineyin/the-singular-value-decomposition-and-pca-cae825ff28fc
https://www.dataminingapps.com/2020/02/singular-value-decomposition-in-recommender-systems/
https://en.wikipedia.org/wiki/Non-negative_matrix_factorization

# Recall the Collaborative Filtering with Matrix Factorization

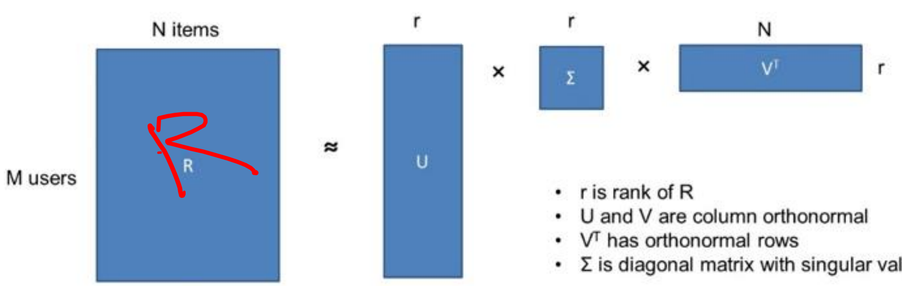



$R \neq U\Sigma V^T$

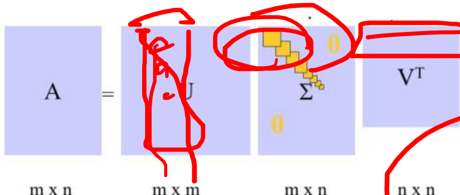
$(RV = U\Sigma)$

orthogonal Scaler

vector




- r is rank of R
- U and V are column orthonormal
- $V^T$ has orthonormal rows
- Σ is diagonal matrix with singular values



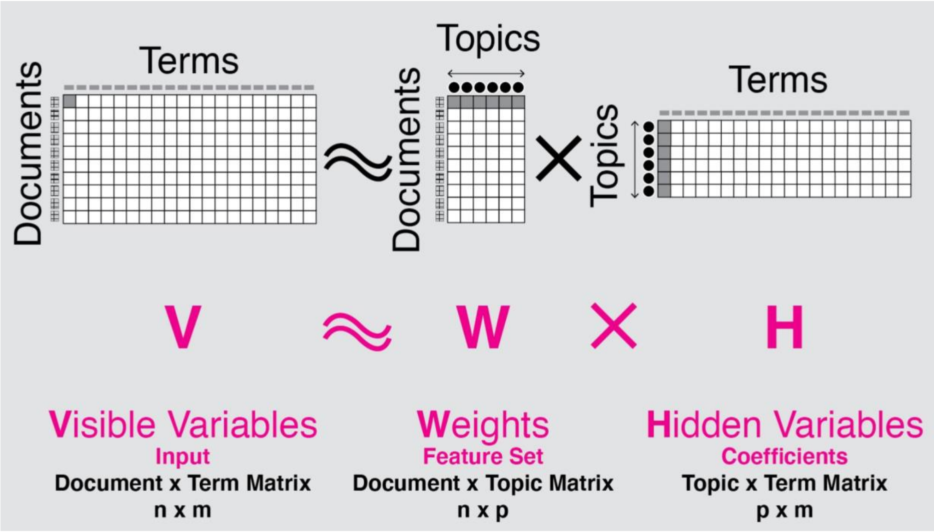

$R = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots$

Image source: https://medium.com/@lineyin/the-singular-value-decomposition-and-pca-cae825ff28fc
https://www.dataminingapps.com/2020/02/singular-value-decomposition-in-recommender-systems/
https://en.wikipedia.org/wiki/Non-negative_matrix_factorization

# Service Feature Engineering

- NMF (Lee and Seung, 1999)
  - NMF operates by starting with a guess of values for W and H, and iteratively minimizing the loss function
  - Typically it is implemented by updating one matrix (either W or H) for each iteration, and continuing to minimize the error function

$$Minimize \left\| V - WH \right\|, W, H \geq 0$$



# Recall the Attempt to Estimate the Mechanism Behind the Simple Dot Product

- Deep-learning-based nonlinearity consideration complements the traditional approaches

## Neural Collaborative Filtering


Xiangnan He
National University of Singapore, Singapore
xiangnanhe@gmail.com

Lizi Liao
National University of Singapore, Singapore
liaolizi.llz@gmail.com

Hanwang Zhang
Columbia University
USA
hanwangzhang@gmail.com

Liqiang Nie
Shandong University
China
nieliqiang@gmail.com

Xia Hu
Texas A&M University
USA
hu@cse.tamu.edu

Tat-Seng Chua
National University of Singapore, Singapore
dcscts@nus.edu.sg



### ABSTRACT

In recent years, deep neural networks have yielded immense success on speech recognition, computer vision and natural language processing. However, the exploration of deep neural networks on recommender systems has received relatively less scrutiny. In this work, we strive to develop techniques based on neural networks to tackle the key problem in recommendation — collaborative filtering — on the basis of implicit feedback.

Although some recent work has employed deep learning for recommendation, they primarily used it to model auxiliary information, such as textual descriptions of items and acoustic features of musics. When it comes to model the key factor in collaborative filtering — the interaction between user and item features, they still resorted to matrix factorization and applied an inner product on the latent features of users and items.

By replacing the inner product with a neural architecture that can learn an arbitrary function from data, we present a general framework named NCF, short for *Neural network-based Collaborative Filtering*. NCF is generic and can express and generalize matrix factorization under its framework. To supercharge NCF modelling with non-linearities, we propose to leverage a multi-layer perceptron to learn the user–item interaction function. Extensive experiments on two real-world datasets show significant improvements of our proposed NCF framework over the state-of-the-art methods. Empirical evidence shows that using deeper layers of neural networks offers better recommendation performance.



### Keywords

Collaborative Filtering, Neural Networks, Deep Learning, Matrix Factorization, Implicit Feedback

*NExT research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IRC@SG Funding Initiative.





*WWW 2017*, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4913-0/17/04.
http://dx.doi.org/10.1145/3038912.3052569

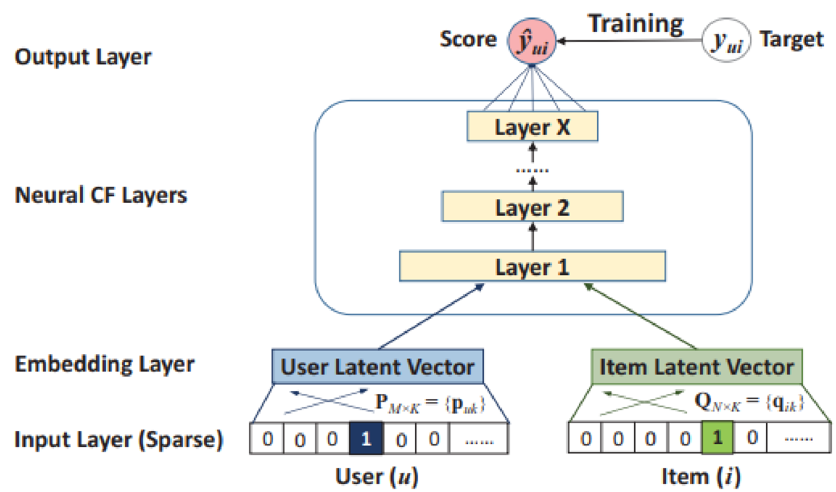

### 1. INTRODUCTION

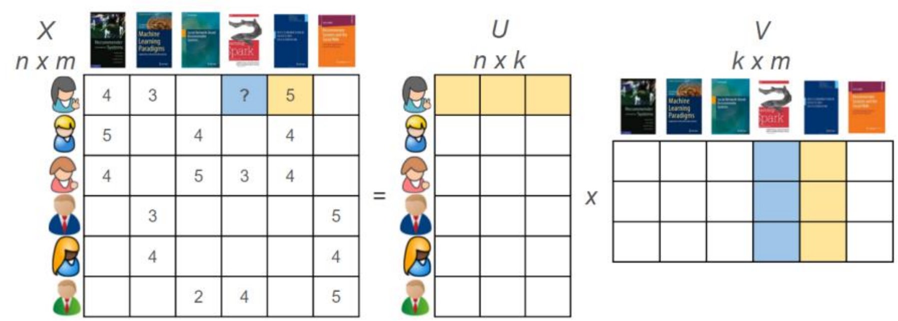
In the era of information explosion, recommender systems play a pivotal role in alleviating information overload, having been widely adopted by many online services, including E-commerce, online news and social media sites. The key to a personalized recommender system is in modelling users' preference on items based on their past interactions (*e.g.*, ratings and clicks), known as collaborative filtering [31, 46]. Among the various collaborative filtering techniques, matrix factorization (MF) [14, 21] is the most popular one, which projects users and items into a shared latent space, using a vector of latent features to represent a user or an item. Thereafter a user's interaction on an item is modelled as the inner product of their latent vectors.

Popularized by the Netflix Prize, MF has become the *de facto* approach to latent factor model-based recommendation. Much research effort has been devoted to enhancing MF, such as integrating it with neighbor-based models [21], combining it with topic models of item content [38], and extending it to factorization machines [26] for a generic modelling of features. Despite the effectiveness of MF for collaborative filtering, it is well-known that its performance can be hindered by the simple choice of the interaction function — inner product. For example, for the task of rating prediction on explicit feedback, it is well known that the performance of the MF model can be improved by incorporating user and item bias terms into the interaction function[1]. While it seems to be just a trivial tweak for the inner product operator [14], it points to the positive effect of designing a better, dedicated interaction function for modelling the latent feature interactions between users and items. The inner product, which simply combines the multiplication of latent features linearly, may not be sufficient to capture the complex structure of user interaction data.

This paper explores the use of deep neural networks for learning the interaction function from data, rather than a handcraft that has been done by many previous work [18, 21]. The neural network has been proven to be capable of approximating any continuous function [17], and more recently deep neural networks (DNNs) have been found to be effective in several domains, ranging from computer vision, speech recognition, to text processing [5, 10, 15, 47]. However, there is relatively little work on employing DNNs for recommendation in contrast to the vast amount of literature

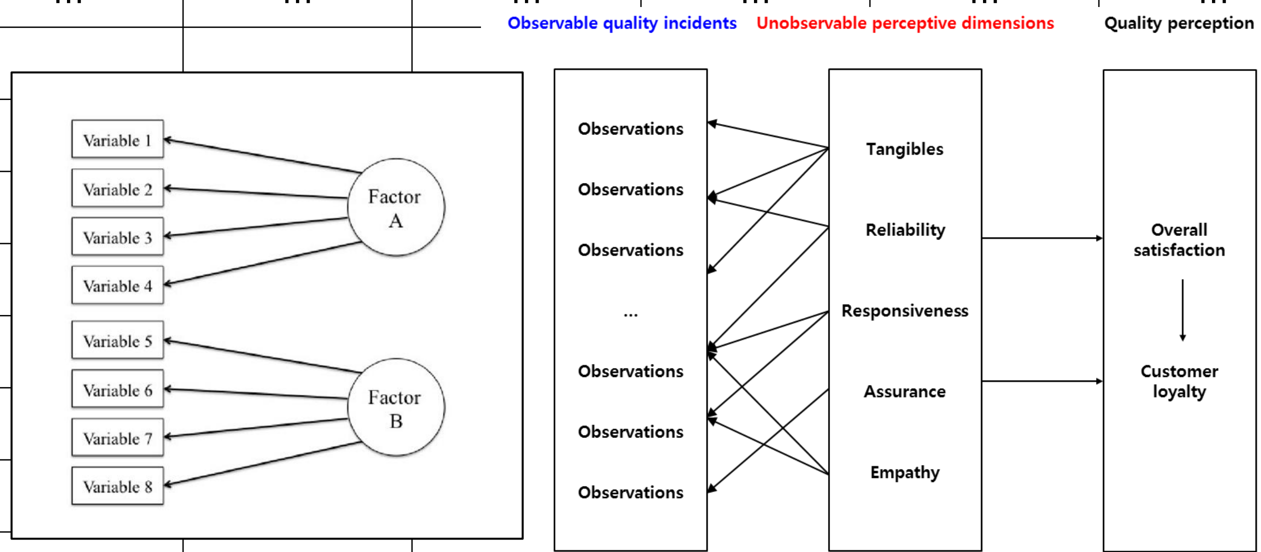
[1] http://alex.smola.org/teaching/berkeley2012/slides/8_Recommender.pdf



# Recall the Attempt to Estimate the Latent Factors Generating the Observed Data

| | metric 1 | metric 2 | metric 3 | … | metric m-1 | metric m |
|---|---|---|---|---|---|---|
| Customer 1 | … | … | … | … | … | … |
| Customer 2 | … | … | … | … | … | … |
| Customer 3 | | | | | | |
| … | | | | | | |
| … | | | | | | |
| … | | | | | | |
| … | | | | | | |
| … | | | | | | |
| … | | | | | | |
| … | … | … | … | … | … | … |
| Customer n-1 | … | … | … | … | … | … |
| Customer n | … | … | … | … | … | … |



# Service Feature Engineering

- Assume that there is a certain process of generating review words with latent service features

| | **Word 1** | **Word 2** | **Word 3** | **…** | **Word m-1** | **Word m** | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | 1 | 2 | 0 | 3 | 4 | 0 | 5 |
| Review 2 | - | - | - | - | - | - | 5 |
| Review 3 | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 3 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 1 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 5 |
| … | - | - | - | - | - | - | 4 |
| … | - | - | - | - | - | - | 3 |
| Review n-1 | - | - | - | - | - | - | 2 |
| Review n | - | - | - | - | - | - | 4 |

# Service Feature Engineering

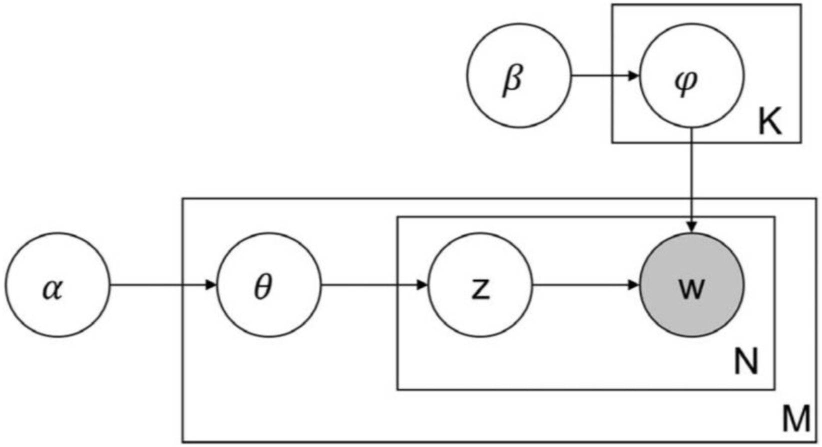
- Assume that there is a certain process of generating review words with latent service features

# Service Feature Engineering

- Latent Dirichlet Allocation (LDA) (Blei et al, 2003)

Step 1: Choose $\theta_i \backsim \text{Dir}(\alpha)$, where $i \in \{1, \ldots, M\}$
Step 2: Choose $\varphi_k \backsim \text{Dir}(\beta)$, where $k \in \{1, \ldots, K\}$
Step 3: For each word position $i, j$, where $i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, N_i\}$
Choose a topic $z_{ij} \backsim \text{Multinomial}\ (\theta_i)$
Choose a word $w \backsim \text{Multinomial}\ (\varphi_{z_{ij}})$

P1=P(topic k / document D)
P2=P(word w / topic k)




Graphical model representation of LDA

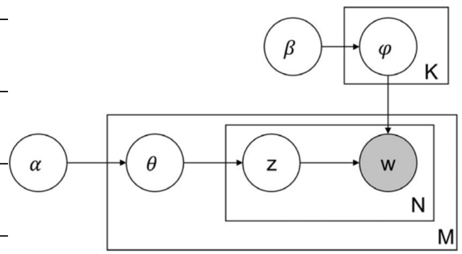
$M$ = number of customer reviews
$N$ = number of words in a review
$K$ = number of topics
$\alpha$ = parameter of the Dirichlet prior on the per-review topic distribution
$\beta$ = parameter of the Dirichlet prior on the per-topic word distribution
$\theta_i$ = topic distribution for review i (the sum of $\theta_i$ is 1)
$\varphi_k$ = word distribution for topic $k$
$z_{ij}$ = topic for the $j^{th}$ word in review i
$w$ = specific word

# On the Customer Review – Word Matrix

| | Word 1 | Word 2 | Word 3 | … | Word m-1 | Word m | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | 1 | 2 | 0 | 3 | 4 | 0 | 5 |
| Review 2 | - | - | - | - | - | - | 5 |
| Review 3 | - | | | | | - | 4 |
| … | - | | | | | - | 3 |
| … | - | | | | | - | 4 |
| … | - | | | | | - | 1 |
| … | - | | | | | - | 4 |
| … | - | | | | | - | 5 |
| … | - | | | | | - | 4 |
| … | - | - | - | - | - | - | 3 |
| Review n-1 | - | - | - | - | - | - | 2 |
| Review n | - | - | - | - | - | - | 4 |

Step 1: Choose $\theta_i \backsim \mathrm{Dir}(\alpha)$, where $i \in \{1, \ldots, M\}$
Step 2: Choose $\varphi_k \backsim \mathrm{Dir}(\beta)$, where $k \in \{1, \ldots, K\}$
Step 3: For each word position $i, j$, where $i \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, N_i\}$
Choose a topic $z_{ij} \backsim \mathrm{Multinomial}\,(\theta_i)$
Choose a word $w \backsim \mathrm{Multinomial}\,(\varphi_{z_{ij}})$

P1=P(topic k / document D)
P2=P(word w / topic k)




$M$ = number of customer reviews
$N$ = number of words in a review
$K$ = number of topics
$\alpha$ = parameter of the Dirichlet prior on the per-review topic distribution
$\beta$ = parameter of the Dirichlet prior on the per-topic word distribution
$\theta_i$ = topic distribution for review i (the sum of $\theta_i$ is 1)
$\varphi_k$ = word distribution for topic $k$
$z_{ij}$ = topic for the $j^{th}$ word in review i
$w$ = specific word

Graphical model representation of LDA

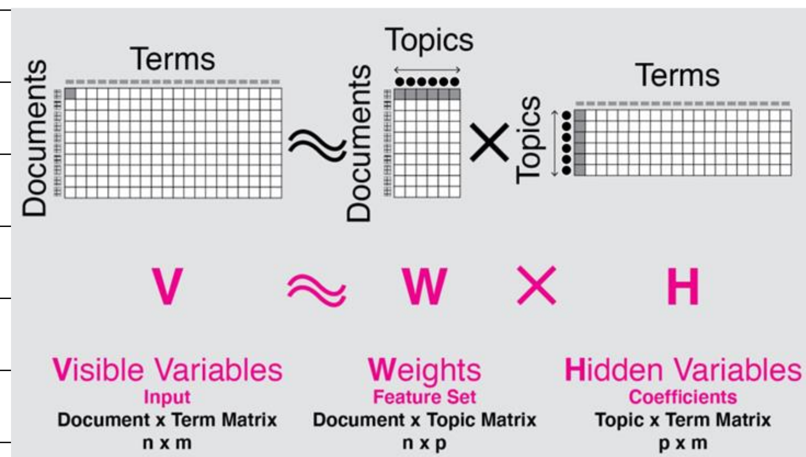
# On the Customer Review – Word Matrix

$tf\text{-}idf$ *value after normalization, when*

$$tf\text{-}idf(t,d) = tf(t,d) \times (\log \frac{1+n_d}{1+df(d,f)} + 1)$$

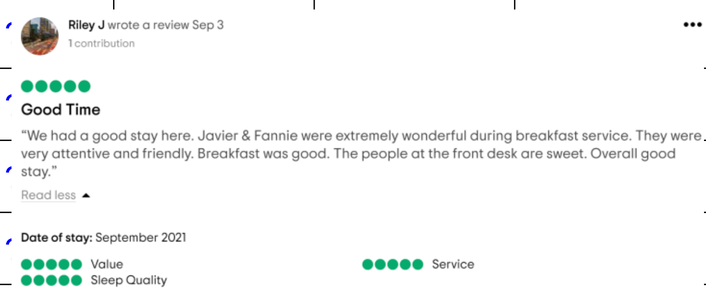
| | Word 1 | Word 2 | Word 3 | … | Word m-1 | Word m | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | 0.04 | 0.18 | 0 | 0.23 | 0.10 | 0 | 5 |
| Review 2 | - | - | - | - | - | - | 5 |
| Review 3 | - | | | | | - | 4 |
| … | - | | | | | - | 3 |
| … | - | | | | | - | 4 |
| … | - | | | | | - | 1 |
| … | - | | | | | - | 4 |
| … | - | | | | | - | 5 |
| … | - | | | | | - | 4 |
| … | - | | | | | - | 3 |
| Review n-1 | - | - | - | - | - | - | 2 |
| Review n | - | - | - | - | - | - | 4 |

$$Minimize\ ||V - WH||, W, H \geq 0$$



# On the Customer Review – Word Matrix

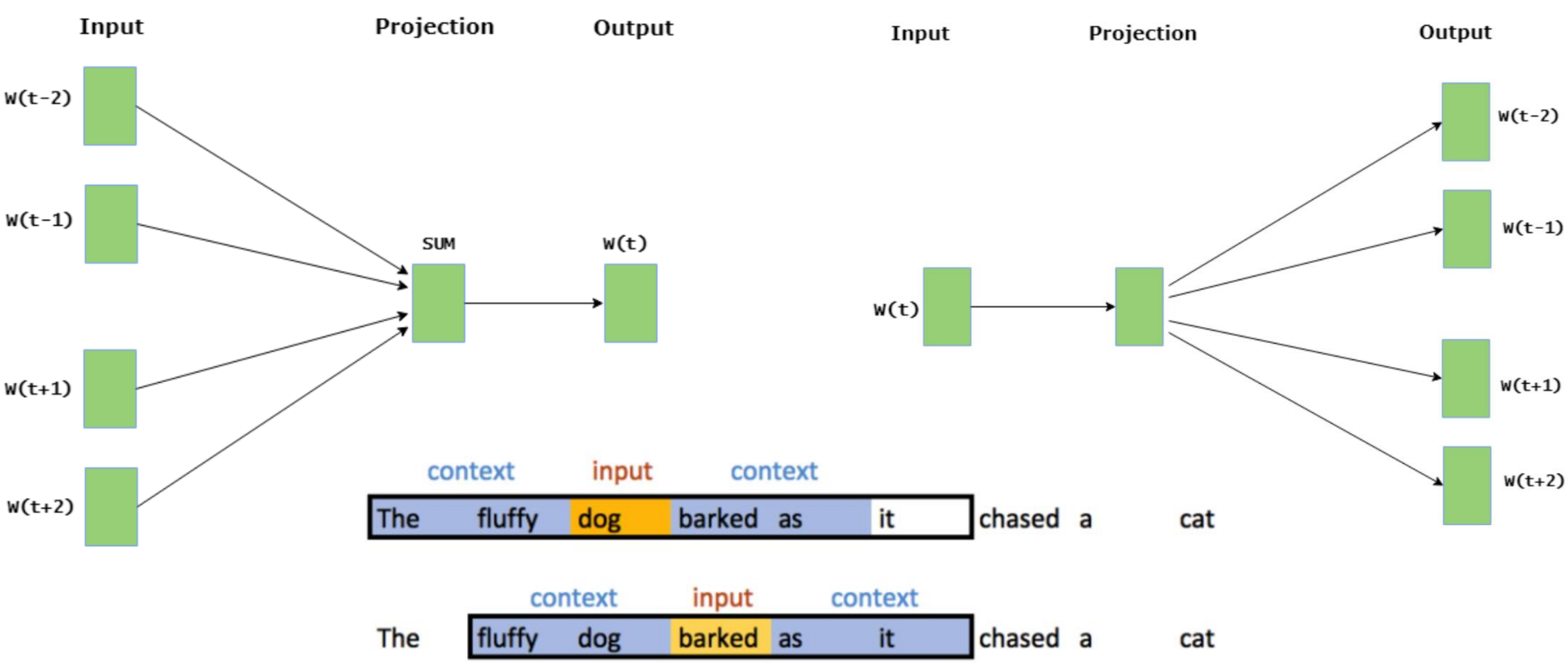
| | **Word 1** | **Word 2** | **Word 3** | **…** | **Word m-1** | **Word m** | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | ? | ? | ? | ? | ? | ? | 5 |
| Review 2 | ? | ? | ? | ? | ? | ? | 5 |
| Review 3 | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | | | | | ? | 3 |
| … | ? | | | | | ? | 4 |
| … | ? | | | | | ? | 1 |
| … | ? | | | | | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 5 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| Review n-1 | ? | ? | ? | ? | ? | ? | 2 |
| Review n | ? | ? | ? | ? | ? | ? | 4 |



# Word Vectorization for Service Feature Engineering

- Word2vec (Mikolov et al., 2013)
  - Word2Vec uses shallow two layer neural networks having one input layer, one hidden layer and one output layer

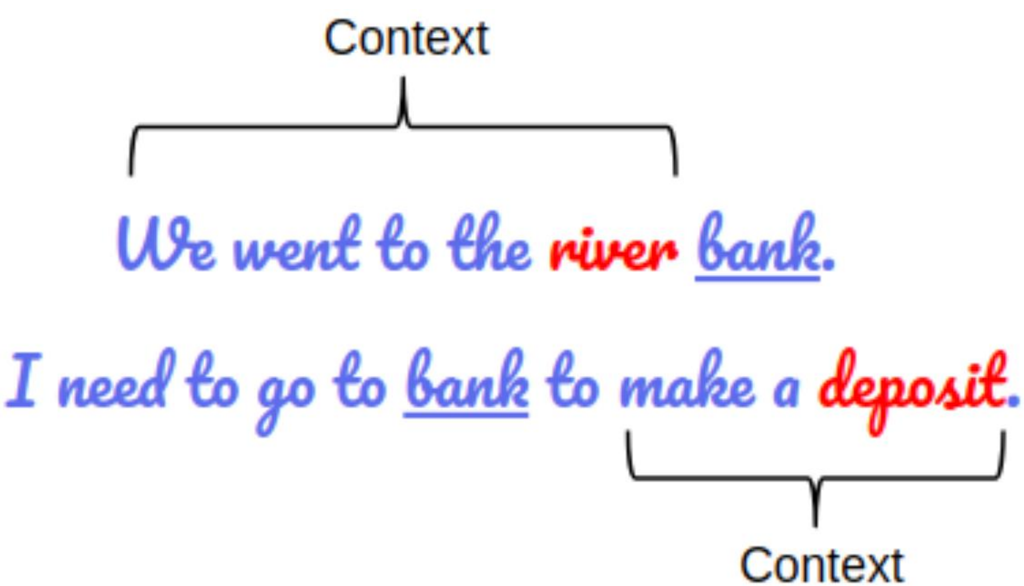
**CBOW (Continuous Bag of Words)** **Skip Gram**




Window size: 5

# Word Vectorization for Service Feature Engineering

- Word2vec (Mikolov et al., 2013) and BERT (Devlin et al., 2019)
  - Word2vec will generate the same single vector for the word bank for both the sentences
  - BERT will generate different vectors for the word bank being used in different contexts



# Case Study: Service Feature Engineering

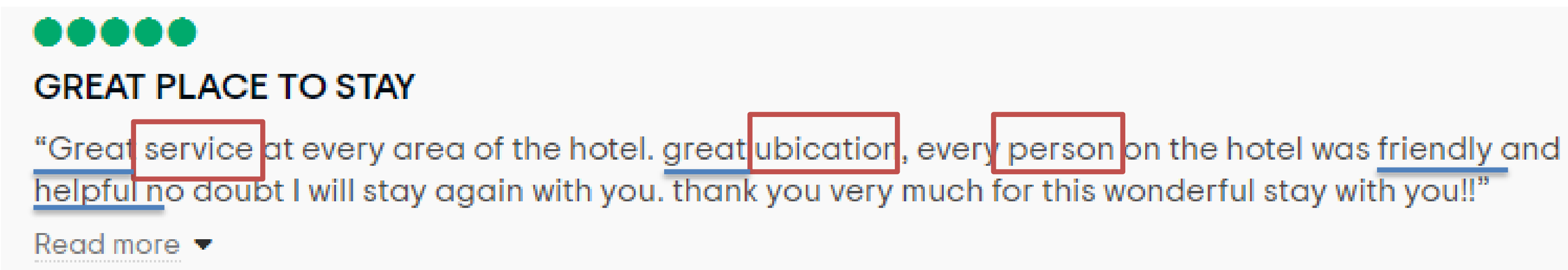
- Use LDA to identify latent service features and their related words
- For example, the Gensim library of Python, which applies variation expectation maximization algorithm (Blei et al. 2003) can be used to execute LDA

Hotel service features in Singapore

| | Feature | Frequent word | # of words | # of reviews |
|---|---|---|---|---|
| $f_1$ | Location | location, $\cdots$ | 63 | 26,700 |
| $f_2$ | View | view, outlook,$\cdots$ | 15 | 6,527 |
| $f_3$ | Breakfast | breakfast, buffet,$\cdots$ | 24 | 13,484 |
| $f_4$ | Sleep quality | bed, mattress,$\cdots$ | 20 | 10,707 |
| $f_5$ | Bathroom | bathroom, toilet,$\cdots$ | 24 | 11,466 |
| $f_6$ | Service | service, staff,$\cdots$ | 32 | 20,864 |
| $f_7$ | Check | check, checkin,$\cdots$ | 19 | 12,651 |
| $f_8$ | Value | value, price,$\cdots$ | 6 | 11,477 |
| $f_9$ | Internet | internet, wifi,$\cdots$ | 32 | 6,137 |

# Sentiment Analysis

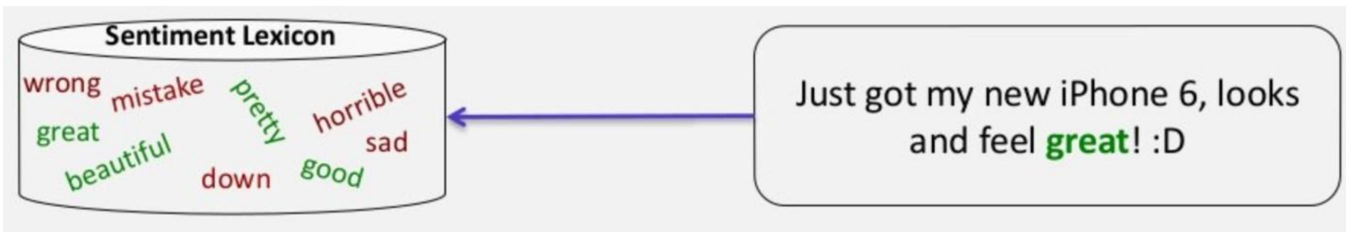
- What is sentiment analysis?
  - Actual performance of features can be determined from a customer perspective
  - The sentiments of the identified features are obtained at each review



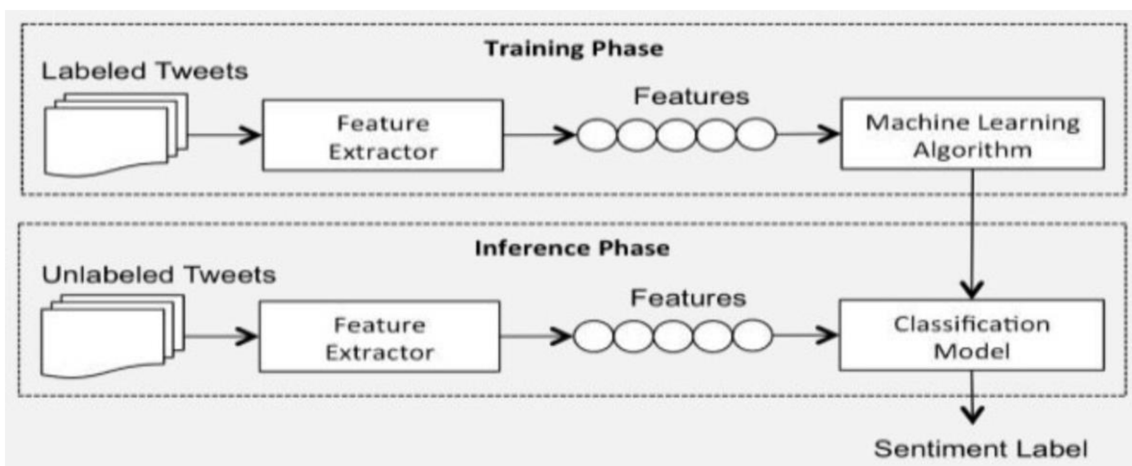

Sentiment analysis of features: Service (very positive), Location (very positive), Staff (very positive)

# Sentiment Analysis

- How to perform sentiment analysis?
  - Lexicon-based approach: e.g., VADER (Valence Aware Dictionary for Sentiment Reasoning)




  - Machine learning approach



# Sentiment Analysis

- VADER sentiment analysis (Gilbert et al., 2014) and SentiStrength (Thelwall et al., 2012)
  - A lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments in social media
  - Available regardless of the domain

| Lexicon | Positive Words | Negative Words |
|---|---|---|
| Simplest (SM) | good | bad |
| Simple List (SL) | good, awesome, great, fantastic, wonderful | bad, terrible, worst, sucks, awful, dumb |
| Simple List Plus (SL+) | good, awesome, great, fantastic, wonderful, best, love, excellent | bad, terrible, worst, sucks, awful, dumb, waist, boring, worse |
| Past and Future (PF) | will, has, must, is | was, would, had, were |
| Past and Future Plus (PF+) | will, has, must, is, good, awesome, great, fantastic, wonderful, best, love, excellent | was, would, had, were, bad, terrible, worst, sucks, awful, dumb, waist, boring, worse |
| Bing Liu | 2006 words | 4783 words |
| AFINN-96 | 516 words | 965 words |
| AFINN-111 | 878 words | 1599 words |
| enchantedlearning.com | 266 words | 225 words |
| MPAA | 2721 words | 4915 words |
| NRC Emotion | 2312 words | 3324 words |

# Case Study: Sentiment Analysis

| | $A_1$ | $A_2$ | … | $A_9$ |
|---|---|---|---|---|
| 1 | **Location** was great | Awesome **view** | … | Terrible **internet** speed |
| 2 | | City **view** was not good | … | |
| 3 | | | … | |
| … | … | … | … | … |
| 32044 | | | … | |

<u>Sentiment intensity score</u>
<u>from VADER sentiment analyzer</u>
'Location': 0.6249
'View': 0.6249, -0.3412
'Internet', -0.4767

$$S_{im} = \begin{cases} 4, & if\ 0.525 \leq \text{Sentiment intensity} \leq 1 \\ 3, & if\ 0.05 \leq \text{Sentiment intensity} < 0.525 \\ 0, & if\ -0.05 < \text{Sentiment intensity} < 0.05 \\ 2, & if\ -0.525 < \text{Sentiment intensity} \leq -0.05 \\ 1, & if\ -1 \leq \text{Sentiment intensity} \leq -0.525 \end{cases}$$

**Encoding for input variable**

| | $A_1$ | $A_2$ | … | $A_9$ | Star ratings |
|---|---|---|---|---|---|
| 1 | 4 | 4 | … | 2 | 0 |
| 2 | 0 | 2 | … | 0 | 0 |
| 3 | 0 | 0 | … | 0 | 1 |
| … | … | … | … | … | … |
| 32044 | 0 | 0 | … | 0 | 1 |

Negative label (1, 2, 3 ratings)

Positive label (4, 5 ratings)

Input variables

Output variables

# Prediction Modeling on Customer Satisfaction

- What is prediction modeling on customer satisfaction?
  - Prediction modeling for identifying the effect or importance values of performance of service features on overall customer satisfaction
  - For example, linear logit classification or other models with interpretation can be applied

| | $f_1$ | $f_2$ | … | $f_i$ | Star ratings |
|---|---|---|---|---|---|
| 1 | 5 | 5 | … | 2 | 0 |
| 2 | 0 | 2 | … | 0 | 0 |
| 3 | 0 | 0 | … | 0 | 1 |
| … | … | … | … | … | … |
| $M$ | 0 | 0 | … | 0 | 1 |

Negative label (1, 2, 3 ratings)

Positive label (4, 5 ratings)

Input variables

Output variables

Dataset for prediction modeling on customer satisfaction

# Case Study: Prediction Modeling on Customer Satisfaction

- Logit model
  - $y = 0.227 \cdot location + 0.085 \cdot view + 0.095 \cdot breakfast + 0.009 \cdot sleep\ quality + 0.119 \cdot bathroom + 0.328 \cdot service + 0.01 \cdot check + 0.087 \cdot value + 0.04 \cdot internet$

$$Y_i = \beta_0 + \beta_1 A_1 + \beta_2 A_2 + \cdots \beta_i A_i$$

| | $f_1$ | $f_2$ | … | $f_9$ | Star ratings |
|---|---|---|---|---|---|
| 1 | 5 | 5 | … | 2 | 0 |
| 2 | 0 | 2 | … | 0 | 0 |
| 3 | 0 | 0 | … | 0 | 1 |
| … | … | … | … | … | … |
| 32044 | 0 | 0 | … | 0 | 1 |

Negative label (1, 2, 3 ratings)

Positive label (4, 5 ratings)

Input variables

Output variables

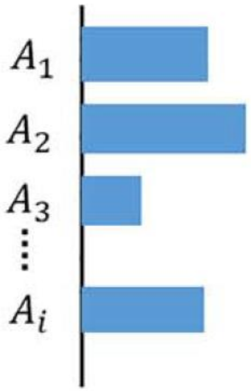
# Prediction Modeling on Customer Satisfaction

## Feature ablation

- Technique for calculating feature importance values that works for any machine learning models
  - 1) Train the model on your train set and calculate a score on the test set
  - 2) For each of the *i* features, remove it from the training data and train the model. Then, calculate the score on the test set.
  - 3) Importance of each feature is the difference between the original score and the score by removing a feature

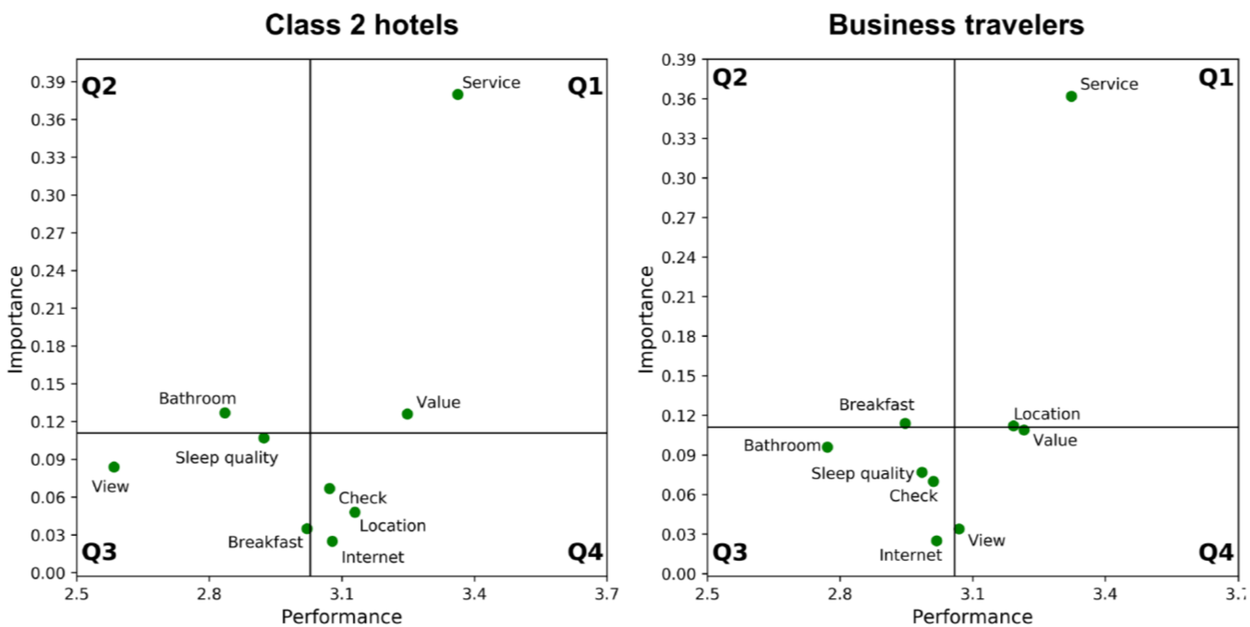
i.e., What about changes in model performance after removing each feature?

Machine learning models (ex. Neural network, random forest, naïve bayes, SVM)

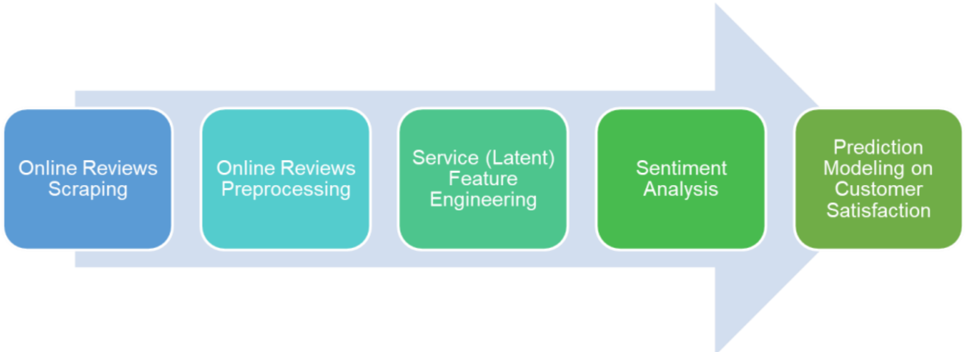


# Service Improvement Implications from Online Review Mining

- Importance-Performance Analysis (IPA) for service improvement
  - Q1: “Keep up the good work” : major strengths
  - Q2: “Concentrate here” : immediate action for improvement
  - Q3: “Low priority” : minor weaknesses
  - Q4: Possible overkill” : minor strengths



# Online Review Mining Framework for Service Improvement




| | Word 1 | Word 2 | Word 3 | … | Word m-1 | Word m | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | ? | ? | ? | ? | ? | ? | 5 |
| Review 2 | ? | ? | ? | ? | ? | ? | 5 |
| Review 3 | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 1 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 5 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| Review n-1 | ? | ? | ? | ? | ? | ? | 2 |
| Review n | ? | ? | ? | ? | ? | ? | 4 |

| | Feature 1 | Feature 2 | Feature 3 | … | Feature k-1 | Feature k | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | ? | ? | ? | ? | ? | ? | 5 |
| Review 2 | ? | ? | ? | ? | ? | ? | 5 |
| Review 3 | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 1 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 5 |
| … | ? | ? | ? | ? | ? | ? | 4 |
| … | ? | ? | ? | ? | ? | ? | 3 |
| Review n-1 | ? | ? | ? | ? | ? | ? | 2 |
| Review n | ? | ? | ? | ? | ? | ? | 4 |

# Assignment 4 (by 09.30 11:59 pm)

- By yourself, (1) complete the construction of the review-feature dataset on service quality of hotels in Singapore based on the practice demonstrated by the TA. Then, (2) using the review-feature matrix you constructed, develop a service quality prediction model for the hotels in Singapore (i.e., predict the customer's quality evaluation with review data). Do it all by yourself, and describe the analysis process and outcome in detail. Interpret the outcome (e.g., name the service features you identified, interpret the coefficient/importance values of service features to the quality ratings).
- (3) What other interesting machines can be developed using the review-feature matrix dataset you constructed? Describe your ideas in detail (e.g., describe the learning objective and process). Try to think your own creative, unique ideas! You have completed the basic review mining activities (tasks 1 and 2) as well as your own idea generation (task 3). Then, (4) describe how you can use your machine(s) to automate the monitoring, evaluation, and improvement of hotel service quality? Imagine you are working for a real hotel.
- Using a similar approach that you have practiced so far, (5) what other services can be improved using review mining machines or another intelligent machine that learns other types of raw data traces of service quality (e.g., customer behavior data)? Assume that you actually manage the quality of service in question. (6) How would you conduct this job in your own creative, unique way? What kinds of data and methods are you going to collect, analyze, and learn? Describe your service intelligence development plan in detail. If possible, visualize your plan clearly (e.g., draw an image, construct a mathematical model).
- Upload your code and a several paragraph essay on the tasks (1)~(6) in the Blackboard.

# Notice on the Survey Next Class

- This course is operated under the "AI-adopted course development project" of UNIST. Thus, the UNIST Education center would like to support the operations of this course.
  - The "AI-adopted course" is a type of interdisciplinary course where the students learn AI-related knowledge and problem-solving skills on the topics of the major (i.e., industrial engineering) through practice-oriented learning and group-project-based learning. This way, the students can integrate the traditional topics of the major with the modern AI knowledge and skills, so that they can broaden and deepen their perspectives of the major.
  - In the next class, there will be a survey to identify the needs of students taking the AI-adopted courses in UNIST, so that the lecturers and Education center ensure the courses operate well throughout this semester.

# Term Project Announcement:

# "Develop Your Own Service Intelligence"

# Topics of the Service Intelligence Course

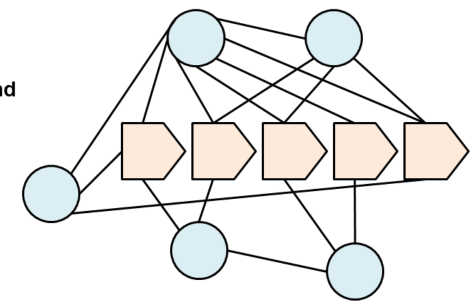
**Service Review Mining**

**Behavioral Data Mining**

**Recommender Systems Development**

**Service Quality Evaluation and Improvement**

**All Weeks**




**Customer Segmentation and Service Customization**

**Special Lectures on Service Intelligence**

**Service Process Assessment and Improvement**

**Service Optimization**

**Industrial Service Intelligence**

# Previous Term Project Topics

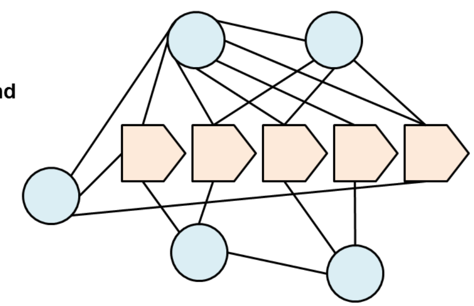
Review-based food recommendation

Behavior-and-symptom -based health diagnosis

Animation recommender system
Book recommender system

Metainfo-and-review-based restaurant recommendation

All Projects




Customized service design for air conditioning machine users and clients

Allergy-free diet planning for children

Optimal routing for COVID-19 vaccine distribution

Industrial service solution development for shale gas productivity prediction and investment

# Assignments, Grading, and Policies

- Assignments
  - You need to complete an assignment as you follow along with the required class
  - Each assignment will require you to answer questions, solve problems, and/or write a report
  - Assignments must be done in the MS Word format and submitted with the filename “Assignment#_ID_Name.docx” (e.g., Assignment1_20201200_ChiehyeonLim.docx)
- Grading

| Item | Portion | Criteria |
| --- | --- | --- |
| Class Participation | 10% | Attendance and In-Class-Presentation for Discussion |
| Assignments | 40% | Comprehension, Completeness, and Creativity |
| Term Project | 50% | Completeness, Adherence to the Course Material, and Creativity |

- Other policies: Late work with penalty and NO cheating

# Customer Complaints Monitoring with ML

# Shouldn't We Consider the Service Quality Dynamics in Time?

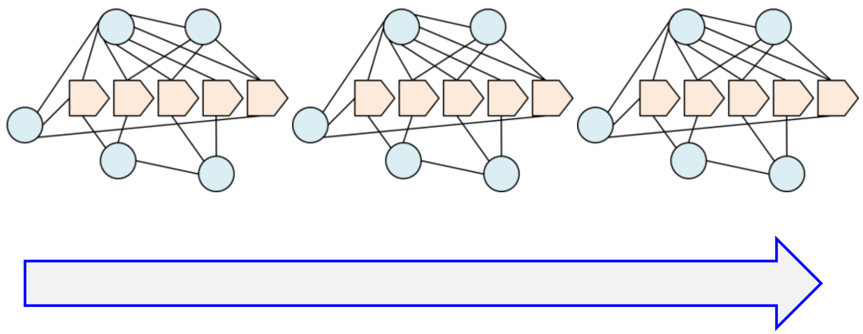
consumer behaviors
(user behaviors)

service delivery
and experience

quality

value cocreation

technology and
self-service



relationship

innovation

design and
development

productivity

knowledge and
employee
management

cost and pricing

service operations
and supply chain
management

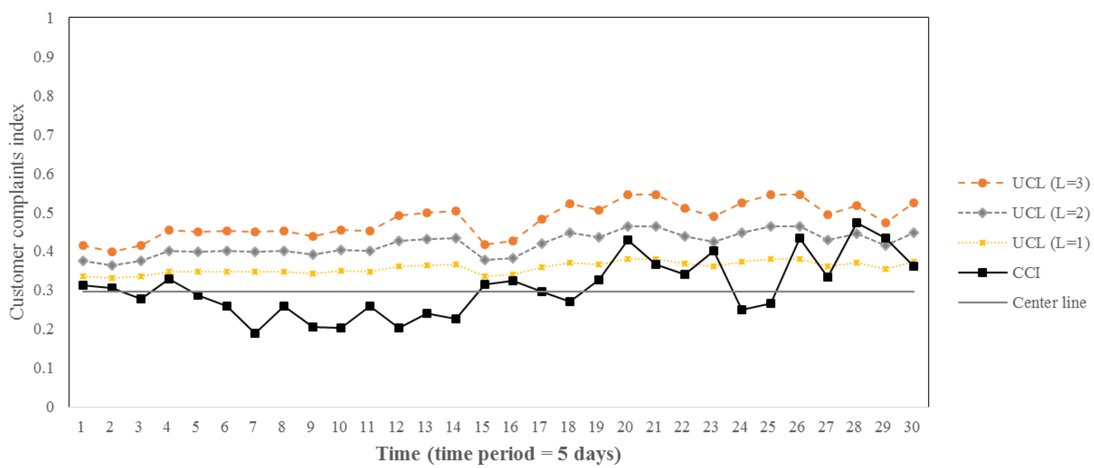
# On the Time Dynamics of Customer Review – Service Feature

| | Feature 1 | Feature 2 | Feature 3 | … | Feature m-1 | Feature m | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | … | … | … | … | … | … | 5 |
| Review 2 | … | … | … | … | … | … | 5 |
| Review 3 | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 3 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 1 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 5 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 3 |
| Review n-1 | … | … | … | … | … | … | 2 |
| Review n | … | … | … | … | … | … | 4 |

| | Feature 1 | Feature 2 | Feature 3 | … | Feature m-1 | Feature m | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | … | … | … | … | … | … | 5 |
| Review 2 | … | … | … | … | … | … | 5 |
| Review 3 | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 3 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 1 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 5 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 3 |
| Review n-1 | … | … | … | … | … | … | 2 |
| Review n | … | … | … | … | … | … | 4 |

| | Feature 1 | Feature 2 | Feature 3 | … | Feature m-1 | Feature m | Rating |
|---|---|---|---|---|---|---|---|
| Review 1 | … | … | … | … | … | … | 5 |
| Review 2 | … | … | … | … | … | … | 5 |
| Review 3 | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 3 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 1 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 5 |
| … | … | … | … | … | … | … | 4 |
| … | … | … | … | … | … | … | 3 |
| Review n-1 | … | … | … | … | … | … | 2 |
| Review n | … | … | … | … | … | … | 4 |

**Time**



# Case Study on the Mobile Game Service Quality Monitoring

- Case study of mobile game service
  - Representing 58% of all downloads in the mobile markets
  - One of the fastest-changing areas with fierce competition
  - Need to manage service quality and update their game services by adding new attributes or fixing bugs


- Data summary
  - Database: Apple app store
  - Application: Angry bird 2
  - Data variables: Date, reviewer ID, rating, title, review content, version
    - Pre-processing the review contents via NLTK in python
  - Period: 2017-07-31 ~ 2017-12-25
  - Country: United States
  - Total number of reviews: 2,010

# Case Study on the Mobile Game Service Quality Monitoring

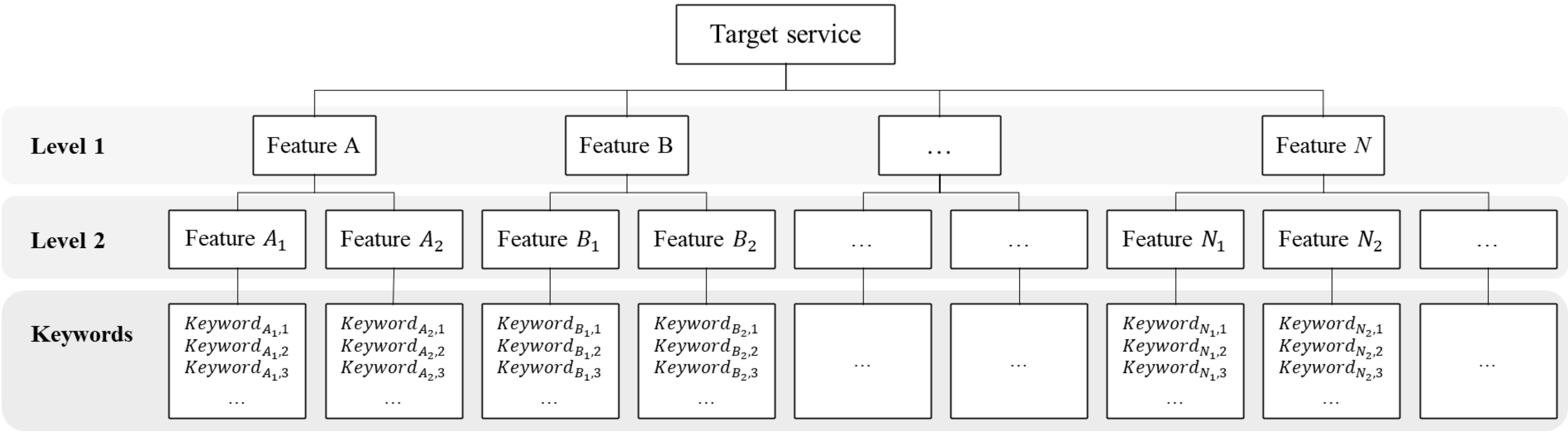
- Step 1: Data collection and pre-processing
  - Collecting the review date and review content
  - Pre-processing the review content via tokenization, stop-words removal, POS tagging, and lemmatization
  - Part of the customer review database

| No. | Date | Reviewer ID | Rate | Title | Review content (raw data) | Review content (after pre-processing) | Version |
|---|---|---|---|---|---|---|---|
| 1 | 2017-07-31 | CPJ*** | 3 | A good pastime | It's Angry Birds, but on steroids. The graphics and sound effect make it fun to pass the time playing. | [(angry, JJ), (bird, NN), (steroid, JJ), (graphic, JJ), (sound, NN), effect, NN), (make, VBP), (fun, NN), (pas, NN), (time, NN), (play, NN)] | 2.14.0 |
| … | … | … | … | … | … | … | … |
| 2010 | 2017-12-25 | Frogit to*** | 5 | Frogit | It's addictive I can't stop playing I have so much fun great game! | [(addictive, JJ), (cant, NN), (stop, VB), (play, NN), (much, JJ), (fun, NN), (great, JJ), (game, NN)] | 2.17.2 |

** Part of speech tags and descriptions: CD=cardinal number, DT=determiner, JJ=adjective, NN=noun (singular), PRP=personal pronoun, RB=adverb, RBR=adverb (comparative), VB=verb (base form), VBD=verb (past tense), VBP=verb (non-3rd person singular present), VBZ=verb (3rd person singular present), WRB=wh-adverb

# Case Study on the Mobile Game Service Quality Monitoring

- Step 2: Construction of a service feature hierarchy with keyword dictionary
  - Service feature hierarchy



# Case Study on the Mobile Game Service Quality Monitoring

- Step 2: Construction of a service feature hierarchy with keyword dictionary
  - Service feature hierarchy employed in this study

| Service feature (level 1) | Description | Service feature (level 2) | Description | Ciurumelea et al. (2017) | McIlroy et al. (2016) | Khalid et al. (2015) | Maalej and Nabil (2015) | Fu et al. (2013) |
|---|---|---|---|---|---|---|---|---|
| Compatibility | Issues related to version of the OS or the specific phone device | Version | Issues related to update or mobile app version | ✓ | ✓ | ✓ | | |
| | | Hardware | Issues related to a specific mobile phone device of OS | ✓ | ✓ | ✓ | | ✓ |
| Usage | Reports the things that are uncomfortable to use and things that user want to improve | Attribute requests | Issues related to additional attribute(s) or modification | ✓ | ✓ | ✓ | ✓ | |
| | | Bug reporting | Issues related to unexpected bug | | ✓ | ✓ | ✓ | |
| | | Difficulty of game | Issues related to difficulty of mobile game | Added by the authors | | | | |
| | | Spam | Issues related to advertisement | | | | | ✓ |
| Resources | Mentions the memory or battery usage | Battery | Issues related to battery usage | ✓ | ✓ | ✓ | | |
| | | Memory | Issues related to memory usage | ✓ | ✓ | ✓ | | |
| Pricing | Refers the licensing model, price of the app, or in-app purchase issues | Price | Issues related to the licensing model, price of the app, or in-app purchase | ✓ | ✓ | ✓ | | ✓ |
| Protection | States the security issues or user privacy | Security | Issues related to security or lack of it | ✓ | | | | |
| | | Privacy | Issues related to permissions and privacy | ✓ | ✓ | ✓ | | |

# Case Study on the Mobile Game Service Quality Monitoring

- Step 3: Identification of customer complaints

  - Part of the results of sentiment analysis

| Review contents | Positive | Neutral | Negative | Compound | Classification |
|---|---|---|---|---|---|
| It's Angry Birds, but on steroids. The graphics and sound effect make it fun to pass the time playing. | 0.268 | 0.645 | 0.087 | **0.671** | Positive |
| Amazing and fun game I love it | 0.79 | 0.21 | 0.0 | **0.906** | Positive |
| I love playing the game it is so much fun and it is challenging it make your mind work thank you so much | 0.446 | 0.554 | 0. | **0.913** | Positive |
| It's cute and the upgrades are nice | 0.537 | 0.463 | 0.0 | **0.7** | Positive |
| The graphics on this latest version are amazing! Even on my iPad 2. | 0.254 | 0.746 | 0.0 | **0.624** | Positive |
| … | … | … | … | **…** | … |
| App keeps freezing when you are playing, do you lose your archived levels. Very frustrating! | 0.088 | 0.54 | 0.372 | **-0.699** | Negative |
| Too hard to reach boss - run out of birds before getting to final room! | 0.074 | 0.811 | 0.114 | **-0.151** | Negative |
| Game has started freezing in the middle of the game. Is there a bug? | 0.0 | 0.896 | 0.104 | **-0.103** | Negative |

  - Out of 2,010 customer reviews, a total of 640 customer reviews identified as customer complaints

# Case Study on the Mobile Game Service Quality Monitoring

- Step 4: Development of a customer complaints chart via SPC
  - Three issues for developing the customer complaints chart using SPC
    - In terms of the use of statistic
      - Interpreting each customer complaint as non-conformity
      - Defining customer complaints index (CCI) by modifying the index used for measuring service performance (Chen and Yang, 2000; Rasouli and Zarei, 2016; Yang and Chen, 2000)

$$CCI_{i,t} = \frac{The\ number\ of\ negative\ customer\ reviews_{i,t}}{The\ total\ number\ of\ customer\ reviews_{i,t}} = \frac{N(NCR_{i,t})}{N(CR_{i,t})}$$

where $NCR_{i,t}$ represents the number of negative customer reviews for the *i* th service feature at the time period *t*

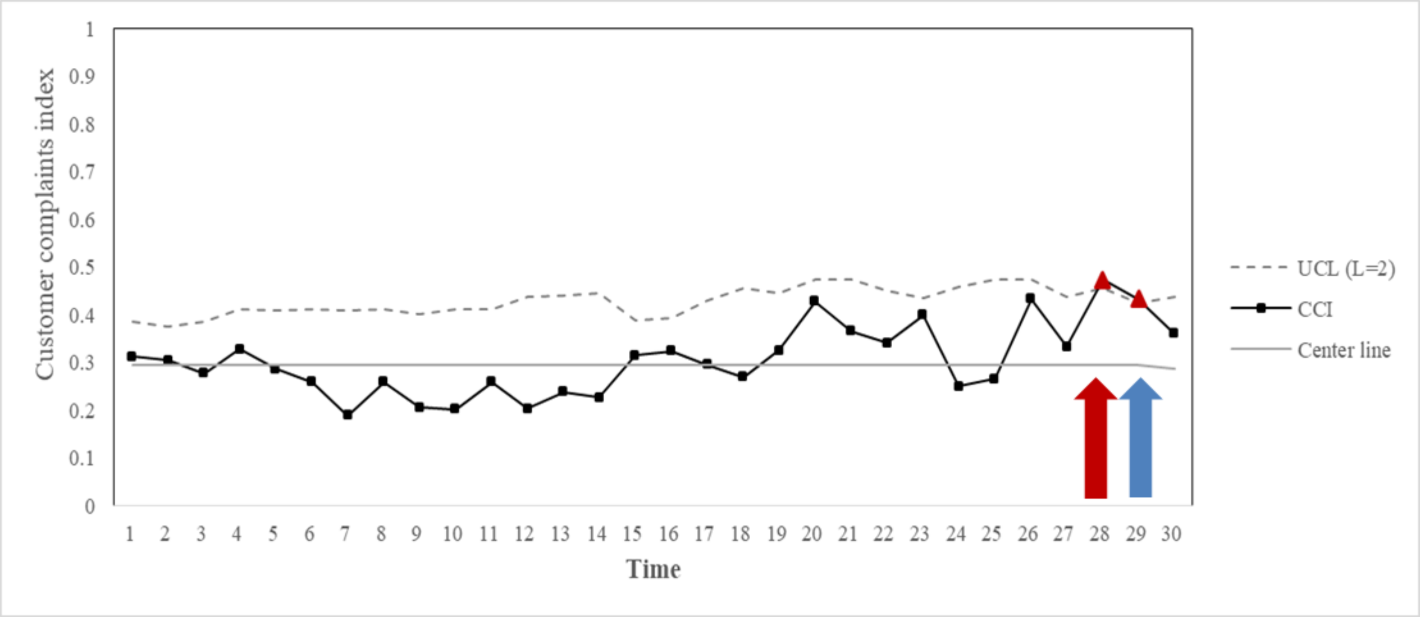
- Defining the center line(CL) and upper control limit (UCL) as:

$$CL_i = \frac{\sum_{t=1}^{T} NCR_{i,t}}{\sum_{t=1}^{T} CR_{i,t}} \qquad UCL_{i,t} = CL_i + L\sqrt{\frac{CL_i(1-CL_i)}{CR_{i,t}}}$$

where *T* and *L* denote the number of time periods and sensitivity parameter, respectively

# Case Study on the Mobile Game Service Quality Monitoring

- Step 4: Development of a customer complaints chart via SPC
  - Customer complaints chart for the overall status (time period=5days, L=2)




Scheduled update to add a new attribute on December 11th, which caused the compatibility issue

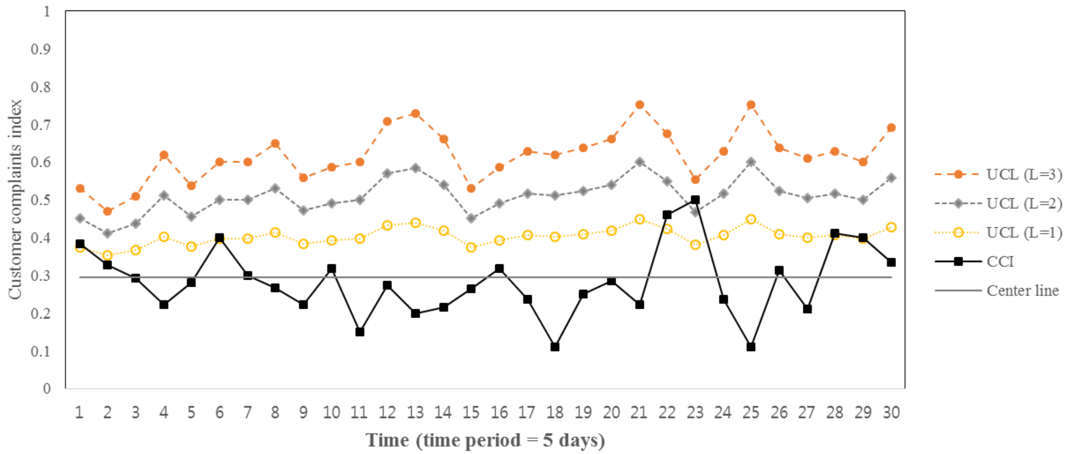
Out-of-control signals on December 13th-17th

Minor update to solve the compatibility problem on December 20th

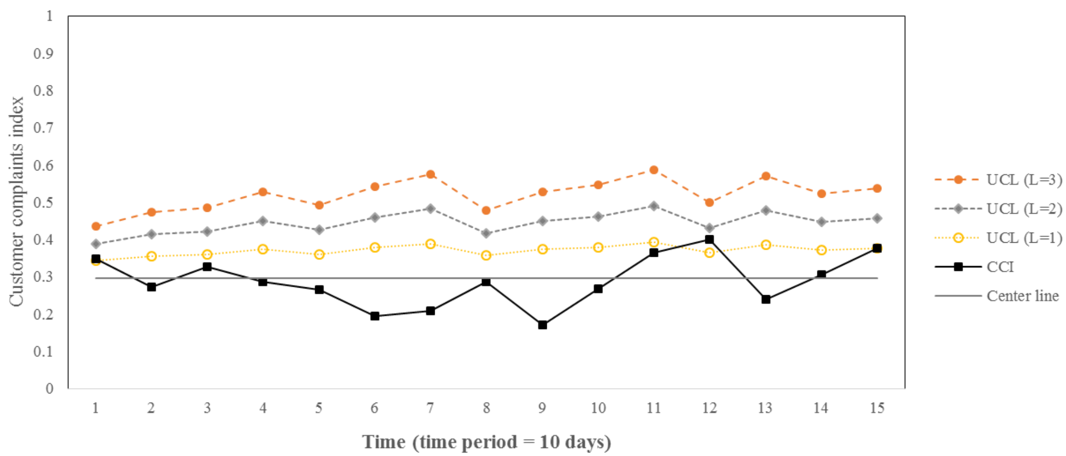
# Case Study on the Mobile Game Service Quality Monitoring

- Customer complaints chart for the attribute request feature with different value of control parameters

(a)




(b)




| Service feature (level 1) | Description | Service feature (level 2) |
|---|---|---|
| Compatibility | Issues related to version of the OS or the specific phone device | Version<br>Hardware |
| **Usage** | **Reports the things that are uncomfortable to use and things that user want to improve** | **Attribute requests**<br>Bug reporting<br>Difficulty of game<br>Spam |
| Resources | Mentions the memory or battery usage | Battery<br>Memory |
| Pricing | Refers the licensing model, price of the app, or in-app purchase issues | Price |
| Protection | States the security issues or user privacy | Security<br>Privacy |

# Further Readings Recommended

- Shin, J., Joung, J., and Lim, C., "Online Review Mining Meets Interpretable Machine Learning for Customer-oriented Service Quality Management," 2022. Article 6
- Kim, J. and Lim, C., "Customer Complaints Monitoring with Customer Review Data Analytics: An Integrated Method of Sentiment And Statistical Process Control Analyses," Advanced Engineering Informatics, Vol. 49, 101304, 2021. Article 7