IES 2022 Innovation & Society 5.0: Statistical and Economic Methodologies for Quality Assessment

BOOK OF SHORT PAPERS

Editors: Rosaria Lombardo, Ida Camminatiello and Violetta Simonacci

Book of Short papers
10th International Conference **IES 2022**
Innovation and Society 5.0: Statistical and Economic Methodologies for Quality Assessment

Department of Economics, University of Campania "L. Vanvitelli",
January 27th - 28th 2022

Scientific Committee of the group of the Italian Statistical Society on Statistics for the Evaluation and Quality of Services – SVQS

Pietro Amenta -University of Sannio
Matilde Bini -European University of Roma
Luigi D'Ambra -University of Naples "Federico II"
Maurizio Carpita -University of Brescia
Paolo Mariani -University of Milan "Bicocca"
Marica Manisera -University of Brescia
Monica Palma -University of Salento
Pasquale Sarnacchiaro -University of Rome Unitelma Sapienza

Program Committee of the conference IES 2022

Chair: Rosaria Lombardo -University of Campania "L. Vanvitelli"
Fabio Bacchini -ISTAT
Laura Baraldi -University of Campania "L. Vanvitelli"
Eric Beh -University of Newcastle, Australia
Wicher Bergsma -The London School of Economic and Political Science, UK
Enrico Bonetti -University of Campania "L. Vanvitelli"
Eugenio Brentari -University of Brescia
Clelia Buccico -University of Campania "L. Vanvitelli"
Rosalia Castellano -University of Naples "Parthenope"
Ida Camminatiello -University of Campania "L. Vanvitelli"
Carlo Cavicchia -University of Rotterdam, The Netherlands
Enrico Ciavolino -University of Salento
Corrado Crocetta -University of Foggia
Claudio Conversano -University of Cagliari
Antonello D'Ambra -University of Campania "L. Vanvitelli"
Antonio D'Ambrosio -University of Naples "Federico II"
Alfonso Iodice D'Enza -University of Naples "Federico II"
Tonio Di Battista -University of Chieti "G. D'Annunzio"
Michele Gallo -University of Naples "L'Orientale"
Francesco Gangi -University of Campania "L. Vanvitelli"
Michele La Rocca -University of Salerno
Amedeo Lepore -University of Campania "L. Vanvitelli"
Riccardo Macchioni -University of Campania "L. Vanvitelli"
Filomena Maggino -University of Rome "La Sapienza"
Angelos Markos -Demokritus University of Thrace, Greece
Lucio Masserini -University of Pisa
Stefania Mignani -University of Bologna
Nicola Moscariello -University of Campania "L. Vanvitelli"
Francesco Palumbo -University of Naples "Federico II"
Alessandra Petrucci -University of Florence
Alessio Pollice -University of Bari

Donato Posa -University of Salento
Luca Secondi -University of Tuscia
Amalia Vanacore -University of Naples "Federico II"
Michel van de Velden -University of Rotterdam, The Netherlands
Rosanna Verde -University of Campania "L. Vanvitelli"
Donatella Vicari -University of Rome "La Sapienza"
Grazia Vicario -Polytechnic University of Turin
Maurizio Vichi -University of Rome "La Sapienza"

Organizing Committee

Ida Camminatiello -University of Campania "L. Vanvitelli"
Antonello D'Ambra -University of Campania "L. Vanvitelli"
Rosaria Lombardo -University of Campania "L. Vanvitelli"
Elvira Romano -University of Campania "L. Vanvitelli"
Luca Rossi -University Cusano
Violetta Simonacci -University of Naples "L'Orientale"

Editors

Rosaria Lombardo - University of Campania “L. Vanvitelli”, Italy
Ida Camminatiello - University of Campania “L. Vanvitelli”, Italy
Violetta Simonacci - University of Naples “L’Orientale”, Italy

PKE - Professional Knowledge Empowerment s.r.l.
Sede legale: Villa Marelli - Viale Thomas Alva Edison, 45 - 20099 Sesto San Giovanni (MI)
Sede operativa: Villa Marelli - Viale Thomas Alva Edison, 45 - 20099 Sesto San Giovanni (MI)
Ufficio Di Rappresentanza: Via Giacomo Peroni, 400 - 00131 Roma (RM)
CF / P.I. 03167830920 — www.pke.it; e-mail info@pke.it — Privacy

February 2022 PKE s.r.l.
ISBN 978-88-94593-35-8 on print
ISBN 978-88-94593-36-5 online

PKE Publisher

Monitoring consumer sentiment using control charts

Monitorare le opinioni dei consumatori con le carte di controllo

Marina Marino, Rocco Mazza, Michelangelo Misuraca and Agostino Stavolo

Abstract Product and service reviews written by consumers as free texts are more and more a fundamental data source to evaluate customer satisfaction. In this work, we propose a strategy to draw a control chart based on the semantic orientation of the opinions conveyed in the reviews. This tool can be used by firms as a visual analytic to monitor consumer experiences and detect which items have to be pushed or withdrawn. A case study based on a set of Amazon reviews is briefly discussed.

Abstract *Le recensioni di prodotti e servizi scritte dai consumatori come testi liberi sono sempre più una fondamentale fonte di dati per valutare la soddisfazione dei clienti. In questo lavoro si propone una strategia per disegnare una carta di controllo basata sull'orientamento semantico delle opinioni contenute nelle recensioni. Tale strumento può essere usato dalle aziende per monitorare le esperienze dei consumatori e individuare quali articoli spingere o eliminare. Un caso studio su recensioni di Amazon è qui brevemente discusso.*

Key words: opinion mining, polarity score, statistical process control

1 Introduction

The technological progresses and the increasing diffusion of Web 2.0 during the last few years changed the way people communicate feelings, opinions and experiences about each facet of everyday-life. More and more, social networks, online forums and review sites and weblogs, are used to share personal ideas and viewpoints and exchange information. This new habit impacted also on purchase intention, since people often decide to buy a product or a service after evaluating the positive and

Marina Marino, Rocco Mazza, Agostino Stavolo
University of Naples Federico II, email: marina.marino@unina.it; rocco.mazza@unina.it; agostino.stavolo@unina.it

Michelangelo Misuraca
University of Calabria, email: michelangelo.misuraca@unical.it

negative statements made by those that already bought the same product or service under consideration. The so-called *electronic word-of-mouth* became then a primary source of information at an individual level but also at a firm level [2,7].

Analysing people comments and reviews is essential for any kind of business, since dissatisfied consumers are potentially dangerous to the firms, triggering a vicious circle of bad reputation [19] and considerable financial losses [16]. Thus, in a managerial perspective, it is necessary to monitor what consumers say about their consumption experiences and feelings [17]. Since comments and reviews are written as free texts, information concerning the opinions are encoded in a form difficult to process automatically. In a text mining framework, it is possible to pre-treat the textual body in order to transform the unstructured data into structured data, and then perform statistical analyses to extract and manage the underlying knowledge base. Dealing with opinions, it is particularly interesting to consider the semantic orientation of the texts, expressing in a numerical form the so-called *sentiment*. Several alternative approaches have been proposed concerning how to calculate a score expressing the negative/positive orientation embodied in the texts and use these scores in an opinion mining strategy [6,15]. Nevertheless, the visualisation of sentiment is still an open research topic, with very few innovative contributions [13], and more oriented towards topic extraction [21,26] or classification [12]. Aiming at monitoring the orientation and the intensity of consumer sentiment, an interesting framework is offered by Statistical Process Control (SPC) techniques and by control charts in particular. Control charts have been widely used in industry as a tool to monitor product features [23]. Typically, they are implemented to control ongoing processes, predict the expected range of outcomes from a process, determining whether a process is stable in statistical control, analysing patterns of process variation from non-routine events or built-in events [24]. The sentiment emerging from the reviews written by consumers for a given product/service, can be seen as a quantitative characteristic of a process involving their purchases as well as their feelings about product quality, usability and functionality. For this reason, trying to build an easily readable visual analytic, here we propose a strategy to monitor the sentiment by using control charts. After briefly reviewing the reference literature and showing in details the adopted methodology, we present the preliminary results of a case study based on a set of reviews concerning the purchase of cell phones and related accessories on Amazon.

2 A brief literature review

SPC methods were initially developed to monitor the quality of products, later extending their use also to services. Control charts are generally classified into two groups. If the quality characteristic is measured on a continuous scale, we have control charts for variables. When instead the quality characteristic is classified as conforming or not-conforming considering the possess of given characteristics, we have control charts for attributes. Control charts have been applied in a large number of diverse areas. Control charts have been proposed, for example, in livestock farming to monitor

and manage animal production systems [3]. In a health-care domain, control charts have been used for public health surveillance [4,22] and surgery quality [8]. In survey design, some authors proposed the use of control charts to check data collection quality [5,9]. In a business domain, aside from manufacturing and product quality control, SPC and control charts have been also used to monitor customer satisfaction and perceived quality [10]. Customer survey data are typically collected considering both qualitative and quantitative characteristics.

As stated above, the use of textual data can offer in this field valuable insights by analysing directly what consumers say or think about a given product or service. In a more general framework concerning the statistical analysis of document collections, [1] suggested the application of control charts to evaluate the solution of a text mining strategy based on a Latent Semantic Analysis. The authors, in particular, evaluated topics explaining the sources of customer satisfaction or dissatisfaction, considering each case of non-conformity as an element requiring an in-depth investigation. This work represented an important crossroads in the reference literature body, since it focuses on unstructured textual feedback and establishes a mechanism for transitioning the textual analysis into an actionable SPC. The assumption underlying this approach is that "out-of-control" comments are negative and express a complain. The analysis of negative and positive comments is the core of opinion mining. [14] extended the SPC approach to the sentiment analysis of consumer reviews, detecting shifts in topic-sentiment combinations. Similarly, [11] proposed a sentiment-based SPC to systematically identify critical complaints within customer review data. Also in the latter case, the analysis specifically focused on negative comments, adopting a complaints index based on the measure used by [25] for service quality control. Here we propose an SPC strategy to monitor customer satisfaction at a product level, monitoring the polarity score attributed to consumer reviews written in natural language. The rationale is to check at the same time if a product or a service has a negative or positive opinion, offering to firms with a huge product/service catalogue a visual tool able to easily identify which items is necessary to push or withdraw.

3 A sentiment-based control chart

The analytical strategy based on the SPC of consumer reviews' sentiment is carried on in three different steps:

1. consumer reviews are retrieved and stored in a repository, together with available metadata (e.g., item ID, item rating, publishing date), then the textual body is pre-treated by cleansing and normalising the embodied terms;
2. for each comment in the repository, a polarity score is calculated to express the negativity/positivity of the underlying opinion about the item;
3. a control chart is drawn to detect the most relevant variations within the sentiment distribution over the item under control.

In the first step, after retrieving and storing the *n* reviews under investigation, texts are "tokenised" to obtain sets of distinct strings (namely, *tokens*) representing the

different terms embodied in the texts, separated by blanks and punctuation marks. All non-alphabetic characters and symbols – like numbers or emoticons – are removed, to consider in the following steps only content-bearing terms. Tokens are then reduced to lowercase. At the end of the pre-process, each text can be encoded as a *p*-dimensional vector, where *p* is the total number of terms listed in the collection.

In the second step, a polarity score is calculated for each review. In particular, each text is split into its composing sentences to take into account the sentiment associated with the different aspects concerning the item review made by the consumers. Given a review d_i, with $i=[1,n]$, its q_i sentences $\{s_{i1}, ..., s_{ik}, ..., s_{iq_i}\}$ are identified by considering as separators strong punctuation marks like full stops, question marks and exclamation marks. The *k*-th sentence is represented as a sequence of its p_k terms $\{w_{ik1}, ..., w_{ikj}, ..., w_{ikp_k}\}$, preserving the order of terms into the sentence. Each term w_{ikj} in the *k*-th sentence of the *i*-th review is compared with a dictionary of polarised terms, giving a score $PS_{w_{ikj}}$ of -1 to negative terms and a score of +1 to positive terms, respectively. Terms not listed in the dictionary are considered neutrals and scored with a 0. The polarity of each term is weighted considering negator terms (e.g., *never*, *none*), amplifier and de-amplifier terms (e.g., *very*, *few*), adversative and contrasting terms (e.g., *but*, *however*). This weighting scheme allows emphasising or dampening the negativity or positivity of each term, leading to a more effective measure of the polarity [20]. The $PS_{s_{ik}}$ polarity score of each sentence is obtained as the sum of its weighted term scores $PS_{w_{ikj}}$, on the square-root of the sentence length:

$$PS_{s_{ik}} = \frac{\sum_{j=1}^{p_k} PS_{w_{ikj}}}{\sqrt{p_k}} \tag{1}$$

Since we are interested in obtaining a polarity score at a review level, we calculate an overall score PS_{d_i} for each text by a down-weighted zeros average of sentence polarities, giving a minor weight to sentences conveying a neutral sentiment:

$$PS_{d_i} = \frac{\sum_{k=1}^{q_i} PS_{s_{ik}}}{q_i^{NEG}+q_i^{POS}+\sqrt{\log(1+q_i^{NEU})}} \tag{2}$$

where q_i^{NEG}, q_i^{POS} and q_i^{NEU} are the number of sentences in d_i with a negative, positive, or neutral polarity, respectively. The score PS_{d_i} assumes values in a $]-\infty,+\infty[$ interval.

In the third step of the strategy, the process to be statistically controlled is defined and then the resulting control chart is drawn and analysed. The logic beneath the process is controlling the distribution of sentiments expressed by consumers in the reviews of the monitored items and detecting out-of-control points. The distance from the central line of the chart is measured in terms of standard deviation, as the maximum and minimum extent of the divergence to label a point as out-of-control. The use of 3σ to plot upper and lower warnings is well established in the reference

literature [18] when SPC is applied to industrial mass production. Since we refer to the description of a process of opinion creation within a community of consumers, the sensitivity parameter can be set on values lower than 3.

4 Some preliminary results

We collected the data from an online repository[1] including 142.8 million reviews published on Amazon between May 1996 and July 2018. The dataset includes reviews (ratings, texts, helpfulness votes), product metadata (descriptions, category, price, brand, and image features), and links ("also viewed"/"also bought" images). In particular, we focused only on reviews about cell phones and accessories (3.4 mln), and considered the products with at least10 reviews and with verified purchases (842,000). Here we considered only the textual information, the date of the review and the so-called ASIN, a code used to identify uniquely a given product.

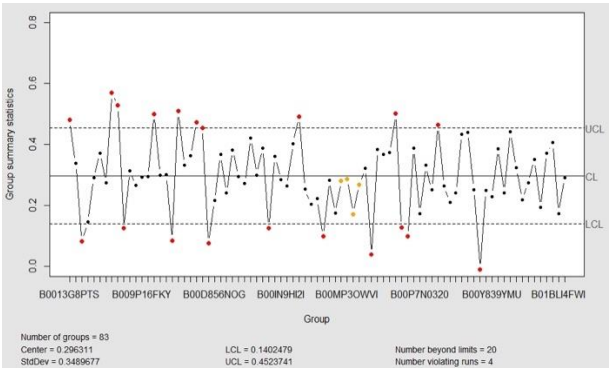

Figure 1: Sentiment-based control chart on products

After calculating the polarity for each comment, we generated a sentiment matrix in which for each ASIN (rows) we selected the 20 most recent comments (columns). This choice aims at avoiding variations in the process caused by modifications or updating of the products made by the sellers. The control line was defined as the average sentiment on each ASIN and the warnings considered a sensitivity parameter equal to 2. The chart in Fig. 1 highlights possible anomalies on the distribution. Products with a sentiment above the upper limit line can be described as strongly recommended products. Basically, the buying community highly recommended its purchase. On the other side, products below the lower limit line can be defined as disappointing products. Although these products do not showed an overall negative sentiment, they are the ones that deviate the most from the control limit, so they deserve more attention. Results will be discussed more in details elsewhere.

[1] See https://nijianmo.github.io/amazon

References

1. Ashton, T., Evangelopoulos, N., Prybutok, V.: Quantitative quality control from qualitative data: control charts with latent semantic analysis. Qual. Quant. **49**, 1081–1099. (2015)
2. Cheung, C.M.K., Thadani, D.R.: The impact of electronic word-of-mouth communication: A literature analysis and integrative model. Decis. Support Syst. **54**, 461–470 (2012)
3. De Vries, A., Reneau, J.K.: Application of statistical process control charts to monitor changes in animal production systems. J. Anim. Sci. **88**, E11–E24 (2010)
4. Finison, L.J., Finison, K.S, Bliersbacl, C.M.: Use of control chart to improve healthcare quality, J. Healthc. Qual. **15**, 9–23 (1993)
5. Gonzalez, Y., Oliver, B.: Producing Control Charts to Monitor Response Rates for Business Surveys in the Economic Directorate of the U.S. Census Bureau. In: 2012 FCSM Research & Policy Conference, pp. 1–12. https://nces.ed.gov/FCSM/pdf/Gonzalez_2012FCSM_V-C.pdf (2012)
6. Hemmatian, F., Sohrabi, M.K.: A survey on classification techniques for opinion mining and sentiment analysis. Artif. Intell. Rev. **52**, 1495–1545 (2019)
7. Ismagilova, E., Slade, E.L., Rana, N.P., Dwivedi, Y.K.: The Effect of Electronic Word of Mouth Communications on Intention to Buy: A Meta-Analysis. Inf. Syst. Front. **22**, 1203–1226 (2020)
8. Jaffray, B.: Am I out of control? The application of statistical process control charts to children's surgery. J. Pediatr. Surg. **55**, 1691–1698 (2020)
9. Jin, J., Loosveldt, G.: Assessing Response Quality by Using Multivariate Control Charts for Numerical and Categorical Response Quality Indicators. J. Surv. Stat. Methodol. **9**, 674–700 (2021)
10. Kenett, R.S., Deldossi, L., Zappa, D.: Quality standards and control charts applied to customer surveys. In: Kenett, R.S., Salini, S. (eds) Modern Analysis of Customer Satisfaction Surveys, pp. 413–438. Wiley, Chichester (2012)
11. Kim, J., Lim, C.: Customer complaints monitoring with customer review data analytics. An integrated method of sentiment and statistical process control analyses. Adv. Eng. Inf. **49**, 101304 (2021)
12. Kim, K., Lee, J.: Sentiment visualization and classification via semi-supervised nonlinear dimensionality reduction. Pattern Recogn. **47**, 758–768 (2014)
13. Kurcher, K., Paradis, C., Keren, A.: The state of the art in sentiment visualization. Comput. Graph. Forum. **37**, 71–96 (2018)
14. Liang, Q., Wang, K.: Monitoring of user-generated reviews via a sequential reverse joint sentiment-topic model. Qual. Reliab. Eng. Int. **35**, 1180–1199 (2019)
15. Liu, B., Zhang, L.: A Survey of Opinion Mining and Sentiment Analysis. In: Aggarwal, C., Zhai, C. (eds.) Mining Text Data, pp. 415–463. Springer, Boston (2012)
16. Luo, X.: Quantifying the Long-Term Impact of Negative Word of Mouth on Cash Flows and Stock Prices. Market. Sci. 28, 148–165 (2008)
17. Nambisan, P., Watt, J.H.: Managing customer experiences in online product communities. J. Bus. Res. **64**, 889–895 (2011)
18. Nelson, L.S.: When should the limits on a Shewhart control chart be other than a center line ±3-sigma? J. Qual. Technol. **35**, 424–425 (2003)
19. Shirdastian, H., Laroche, M., Richard M.-O.: Using big data analytics to study brand authenticity sentiments: The case of Starbucks on Twitter. Int. J. Inform. Manag. **48**, 291–307 (2019)
20. Vechtomova, O.: Disambiguating context-dependent polarity of words: An information retrieval approach. Inf. Process. Manag. **53**, 1062–1079 (2017)
21. Wang, C., Xiao, Z., Liu, Y., Xu, Y., Zhou, A., Zhang, K.: SentiView: Sentiment Analysis and Visualization for Internet Popular Topics. IEEE Trans. Hum. Mach. Syst. **43**, 620–630 (2013)
22. Woodall, W.H.: The Use of Control Charts in Health-Care and Public-Health Surveillance. J. Qual. Technol. **38**, 89–104 (2006)
23. Woodall, W.H., Spitzner, D.J., Montgomery, D.C., Gupta, S.: Using Control Charts to Monitor Process and Product Quality Profiles. J. Qual. Tech. **36**, 309–320 (2004)
24. Xie M., Goh T.N., Kuralmani V.: Statistical models and control charts for high-quality processes. Springer, New York (2002)
25. Yang, H.H., Chen, K.S.: A performance index approach to managing service quality. Manag. Serv. Qual. **10**, 273–278 (2000)
26. Zhao, Y., Qin, B., Liu, T., Tang, D.: Social sentiment sensor: a visualization system for topic detection and topic sentiment analysis on microblog. Multimed. Tools Appl. **75**, 8843–8860 (2016)