Analyzing Cloud Business Services with Choquet Fuzzy Integrals and Support Vector Machines

Jose L. Salmeron and Pedro Palos

Abstract Cloud computing poses both opportunities and challenges for companies and IT professionals. Some of these are technical challenges that can be solved over time, while others are related to uncertainties arising from the commitment to a recent innovation. The objective of this research is to identify some of the uncertainties that IT professionals may have and can discourage them from adopting cloud computing. In fact, this paper is focused on predicting the perceived easy-of-use of cloud business services. For that purpose, we use Choquet Fuzzy Integral and Support Vector Machines.

Keywords Cloud services · Choquet fuzzy integrals · Support vector machines

1 Introduction

Cloud computing poses both opportunities and challenges for companies and IT professionals. Some of these are technical challenges that can be solved over time, while others are related to uncertainties arising from the commitment to a recent innovation. The objective of this research is to identify some of the uncertainties that IT professionals may have and can discourage them from adopting cloud computing.

An innovation is an idea, practice or object that is perceived as new [17]. Although the newness of cloud computing is certainly debatable, there is no doubt that its introduction challenges our conventional understanding of the location and management of IT infrastructure, the nature of products and services, business processes and practice of its services (both for IT professionals and consumers).

The factors which are potentially affecting the intention of IT professionals in the use of cloud computing to deliver products and services to their customers are taken mainly from the theory of diffusion of innovation [17].

J.L. Salmeron (✉) · P. Palos
Data Science Lab, Universidad Pablo de Olavide, Km. 1 Utrera Road, 41013 Seville, Spain
e-mail: salmeron@acm.org

I. Czarnowski et al. (eds.), *Intelligent Decision Technologies 2016*,
Smart Innovation, Systems and Technologies 57,
DOI 10.1007/978-3-319-39627-9_23

The theory identifies five variables that have a profound influence on the rate of adoption of innovation including: perceived attributes of innovation, the type of decision in innovation, communication channels, the nature of the social system, and changing the promotional efforts of agents. The perceived attributes of innovation are an important predictor of intent in adopting innovations [17].

There are many benefits in taking the services offered by a cloud provider. Its application will depend on the nature, size and needs of the company. The decision to choose the option that best meets customer needs is a complex task due to the appearance of different suppliers.

Each provider has its own pricing policy, a degree of flexibility in offering services and a technical support appropriate to the service that he thinks he should supply. The offer revolves around these three important pillars. Many companies like Amazon, Google, Microsoft and Salesforce have become cloud computing providers.

The open source community is also present as a provider within the business model that offers, looking very active in the area of cloud computing with numerous contributions, especially in virtualization technologies [15].

Virtualization is a key technology for the cloud which allows a more efficient and flexible use of resources. Virtualization is a key element of the cloud for its advantages such as flexibility, isolation and utilization rate of resources. Building a cloud environment often initially involves choosing a management solution for the cloud. Often, this decision is difficult, because each solution has its specific characteristics [6].

2 Theoretical Background

There are many variables that can influence the adoption of cloud computing. Research in this field is still scarce due to its recent emergence and adoption by businesses. Often, technological literature focuses on addressing issues and challenges related to adoption, such as service availability, performance, lack of interoperability standards and difficulty of integration and customization [7, 8, 10].

Moreover it is possible to find several studies that emphasize the importance of confidence, both in the adoption of cloud technology, as in the privacy in data storage. This paper focuses on the importance of various especially significant aspects in the literature of adoption of technological innovations, like Cooperation, Complexity Technology, Training, Top Management Support and Communication [16, 21].

- Cooperation, either in its internal or external aspects, provides synergies, it reassures users and helps achieve expectations. Internal cooperation is the exerted between the different functional areas of a company [12, 13]. Furthermore, external cooperation refers to the links that the organization maintains with the cloud provider.

- Top Management Support. It is defined as the active involvement of those responsible for the management in the successful implementation of technology [2]. This active participation materializes mainly through leadership and continuous contact with those who are directly linked to the IT planning [19]. Through these measures, users tend to assimilate the expectations of management, in addition to perceive that those responsible for an organization support its implementation. This increases the employees' favourable attitudes regarding IT.
- Training. Training is described as the degree to which a company instructs its employees in the use of a tool in terms of quality and quantity. In a complex information system such as cloud computing, the organization needs to train employees and develop skills for effective use in the future [9]. This reduces the potential stress of staff and provides greater motivation and a better understanding of the benefits of the cloud.
- Communication. Quality communication occurs when members employ a certain amount of time exchanging information and views, either formally or informally [11]. This type of communication increases the distribution of ideas and improves knowledge transfer, especially when the information transmitted is credible and from reliable sources.
- Technological complexity. It is defined as the degree by which an innovation is perceived as relatively difficult to understand and use. In the case of cloud computing, the aspects related to the complexity could be the time needed for the development of tasks, application integration with cloud infrastructure, interface design or efficiency in data transfer, etc. [9]. In short, the technological complexity can influence the adoption of a cloud solution.

3 Methodological Framework

3.1 Choquet Fuzzy Integral

Non-additive measures are known in literature as fuzzy measures, monotonic measures and capacities between others. Some additive operators such as simple weighted average, ordered weighted average, quasi arithmetic means, weighted min and weighted max are usually used for aggregation purpose. These operators assume that the attributes are always independent between them. This assumption is not correct in scenarios where in many cases, the attributes are strongly interrelated. According to this, aggregation should not be always carried out using common additive operators instead, Fuzzy Choquet Integrals are useful to aggregate with interrelated attributes [5].

Let's define a fuzzy measure μ on a finite set $N = \{1, 2, \ldots, n\}$ as a function $\mu : P(N) \rightarrow [0, 1]$ (where $P(N)$ is the power set of N) satisfying the following conditions:

$$\mu(\phi) = 0 \tag{1}$$

$$\mu(N) = 1 \tag{2}$$

$$A \subseteq B \text{ implies that } \mu(A) \leq \mu(B) \tag{3}$$

The third condition allows measures that do not satisfy the strong condition of additivity. For our purposes, this means that we can model systems where the high value of an attribute of the system in itself does not indicate deviations unless a set of other attributes show deviations from their usual values at the same time [14].

Fuzzy Choquet Integral is one of the most general formulations when using monotone measures as the basis of aggregation [18]. To formulate the definition, we assume n attributes measures $(c_1, \ldots, c_n)$ that it generate the corresponding $(s_1, \ldots, s_n)$ values after an evaluation performed.

The basic properties of the operator are determined by the monotone measure, such as symmetry, additivity and linearity. A discrete Fuzzy Choquet Integral with respect to a monotone measure μ is defined as

$$C_\mu(s_1, \ldots, s_n) = \sum_{i=1}^{n} (s_{(i)} - s_{(i-1)}) \mu(C_{(i)}) \tag{4}$$

where $s(i)$ denotes a permutation of the s_i values such that $s_{(1)} \leq s_{(2)} \leq \ldots \leq s_{(n)}$ and $C_{(i)} = \{c_{(i)}, c_{(i+1)}, \ldots, c_{(n)}\}$.

3.2 Support Vector Machines

Let us start from a binary classification

$$\{x_i, y_i\}, i = 1, ..., n, y_i \in \{-1, 1\}, x_i \in R^n \tag{5}$$

where x_i are data points, and y_i are labels. The data points are separated with a hyperplane given by $w^T x + b = 0$, where w is a n-dimensional coefficient vector that is normal to the hyperplane, and b is the offset from the origin (Fig. 1).

The linear SVM obtains an optimal separating margin by solving an optimization problem [20] as follows

$$\begin{aligned} \min(w, \xi) = \tfrac{1}{2} ||\mathbf{w}||^2 + C \textstyle\sum_{i=1}^{n} \xi_i \\ \text{s.t. } y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i, \xi_i \geq 0 \end{aligned} \tag{6}$$

where $\xi_i = \max(0, 1 - y_i(w \cdot x_i + b))$ if and only if ξ_i is the smallest non-negative value satisfying $y_i(w \cdot x_i + b) \geq 1 - \xi_i$.

Figure 2 shows an example of a linear SVM, where the solid line h in the figure is the final SVM solution.

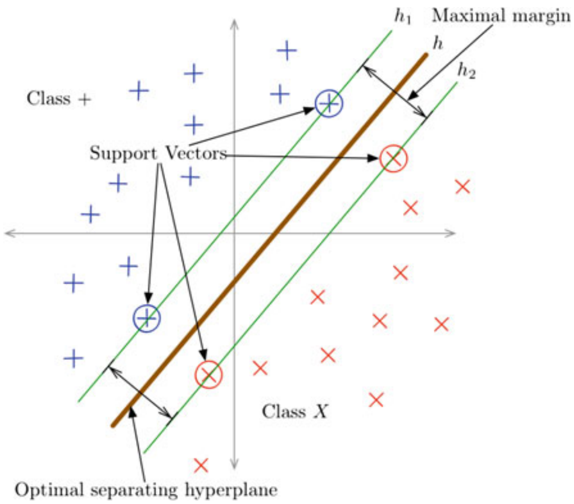

Fig. 1 A linear support vector machine for binary classification

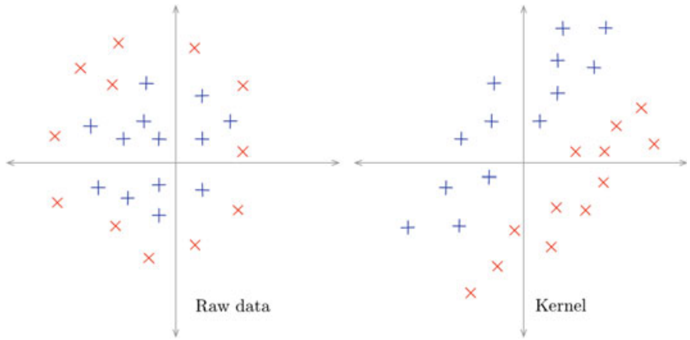

Fig. 2 From raw data to higher dimensional with kernels

Usually, the classification cannot be done linearly. In order for the linear classification to work well in non-linear data, kernels are introduced [20]. The original input space can be mapped into some higher-dimensional feature space where the training set is linearly separable [4]. With this kind of mapping, the decision function can be expressed as

$$g(\mathbf{x}) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i \cdot y_i \cdot K(\mathbf{x}_i, \mathbf{x}) + b \right) \tag{7}$$

Table 1 Kernel types and functions

Kernels	Functions
Linear kernel	$K(x, x_i) = (x^T x_i)$
Polynomial kernel	$K(x, x_i) = \left((x^T x_i) + 1\right)^n$
Radial Based Kernel (RBF)	$K(x, x_i) = exp(-\gamma \vert\vert x - x_i \vert\vert^2)$
Sigmoid kernel	$K(x, x_i) = \frac{e^{2\cdot(x^T x_i)+b} - 1}{e^{2\cdot(x^T x_i)+b} + 1}$

where the kernel function is $K(\mathbf{x}_i, \mathbf{x}) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$, $(\mathbf{x}_i)^T \mathbf{x}$ in the input space is represented as the form $\phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ in the feature space. It is not needed to know the functional form of the mapping $\phi(\mathbf{x}_i)$ since it is implicitly defined by one selected kernel (Table 1) [1].

4 Experimental Results

To evaluate our proposed theoretical approach we performed a survey with data from 128 respondents. All of them are CIO from companies located in Spain.

This paper is going to predict the perceived easy-of-use in cloud business services. For that purpose, the items used for measuring the Perceived easy-of-use construct are the following [2, 3]:

- Interaction with the cloud computing services is clear and friendly
- Working with the cloud services do not demand a mental effort
- Cloud services are simple to use
- It is straightforward to find some stuff in the cloud service

The construct Top management support is measured by the following items:

- Top management is interested in cloud services
- Top management understands the importance of cloud services
- Top management sponsors cloud services
- Top management understands the opportunities of cloud services

The items used for measuring the Technological complexity construct are the following:

- It is hard to understand what cloud services is doing
- Working with cloud services takes too long
- Working with cloud services needs a hard training
- In general terms, working with cloud services is so hard

The construct Communication is measured by the following items:

- Communication about the cloud services is fluid
- There is not constraints about cloud services' communication
- The information about cloud services is correct

Table 2 Error measurement

Measures	Results
MSE	0.08317709
RMSE	0.28840439
SMAPE	1.83722632

After run the proposal hybrid methodology we checked the results with three common error measures. The first one is the Mean Squared Error (MSE). It measures the average of the squares of the errors and is computed as follows

$$\mathrm{MSE} = \frac{1}{n} \cdot \sum_{i=1}^{n} (\hat{x}_i - x_i)^2 \tag{8}$$

where n is the number of experiments, x_i is the real value and $\hat{x}_i$ is the estimated one. The second error measure computed is the Root-Mean-Square Error (RMSE). It measures the differences between the estimated values and the real observed values. It is computed as follows

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (\hat{x}_i - x_i)^2} \tag{9}$$

The third error measures is the Symmetric Mean Absolute Percentage Error (SMAPE) is based on relative errors. It is usually defined as follows

$$\mathrm{SMAPE} = \frac{1}{n} \cdot \sum_{i=1}^{n} \frac{|\hat{x}_i - x_i|}{\left(|\hat{x}_i| + |x_i|\right)/2} \cdot 100 \tag{10}$$

The results are detailed in Table 2.

We consider that the results confirm that our proposal is a worthy endeavour.

5 Conclusions

Cloud computing is challenging for companies and IT professionals. The research proposes a hybrid machine learning methodology for predicting the perceived easy-of-use of cloud business services.

For that purpose, we use Choquet Fuzzy Integral and Support Vector Machines. The results confirm that this proposal is a worthy endeavour.

References

1. Akay, M.F.: Support vector machines combined with feature selection for breast cancer diagnosis. Expert Syst. Appl. **36**(2), 3240–3247 (2009)
2. Bueno, S., Salmeron, J.L.: Fuzzy modeling enterprise resource planning tool selection. Comput. Stan. Interfaces **30**(3), 137–147 (2008)
3. Burda, D., Teuteberg, F.: Exploring consumer preferences in cloud archiving—a student's perspective. Behav. Inf. Technol. **35**(2), 89–105 (2016)
4. Chen, H.L.: A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis. Expert Syst. Appl. **38**(9), 11796–11803 (2011)
5. Choquet, G.: Theory of capacities. Ann. Inst. Stat. Fourier **5**, 131–295 (1953/54)
6. Endo, P.T., Gonçalves, G.E., Kelner, J., Sadok, D.: A survey on open-source cloud computing solutions. In: Brazilian Symposium on Computer Networks and Distributed Systems, pp. 3–16 (2010)
7. Feuerlicht, G., Govardhan, S.: Impact of cloud computing: beyond a technology trend. Syst. Integr. **20010**, 2 (2010)
8. Feuerlicht, G., Burkon, L., Sebesta, M.: Cloud computing adoption: what are the issues. Systémová integrace, pp. 187–192 (2011)
9. Gangwar, H., Date, H., Ramaswamy, R.: Understanding determinants of cloud computing adoption using an integrated TAM-TOE model. J. Enterp. Inf. Manage. **28**(1), 107–130 (2015)
10. Géczy, P., Izumi, N., Hasida, K.: Cloudsourcing: managing cloud adoption. Glob. J. Bus. Res. **6**(2), 57–70 (2012)
11. Hsu, J. S.-C., Shih, S.P., Chiang, J.C., Liu, J.Y.: The impact of transactive memory systems on IS development teams' coordination, communication, and performance. Int. J. Proj. Manage. **30**(3), 329–340 (2012)
12. Joo, J., Sang, Y.: Exploring Koreans smartphone usage: an integrated model of the technology acceptance model and uses and gratifications theory. Comput. Hum. Behav. **29**(6), 2512–2518 (2013)
13. Lee, D.Y., Lehto, M.R.: User acceptance of YouTube for procedural learning: an extension of the technology acceptance model. Comput. Educ. **61**, 193–208 (2013)
14. Llamazares, B.: Constructing choquet integral-based operators that generalize weighted means and OWA operators. Inf. Fusion **23**, 131–138 (2015)
15. Mahjoub, M., Mdhaffar, A., Halima, R.B., Jmaiel, M.: A comparative study of the current cloud computing technologies and offers. In: 2011 First International Symposium on Network Cloud Computing and Applications (NCCA), pp. 131–134. IEEE (2011)
16. Pearson, S.: Toward accountability in the cloud. IEEE Internet Comput. **15**(4), 64 (2011)
17. Rogers, E.M.: Diffusion of Innovations: modifications of a model for telecommunications. In: *Die Diffusion von Innovationen in der Telekommunikation*, vol. 17, pp. 25–38. Springer, Berlin Heidelberg (1995)
18. Sanz, J., Lopez-Molina, C., Cerrn, J., Mesiar, R., Bustince, H.: A new fuzzy reasoning method based on the use of the Choquet integral. In: Proceedings of EUSFLAT, pp. 691–698 (2013)
19. Son, H., Park, Y., Kim, C., Chou, J.-S.: Toward an understanding of construction professionals' acceptance of mobile computing devices in South Korea: an extension of the technology acceptance model. Autom. Constr. **28**, 82–90 (2012)
20. Vapnik, V.N.: The Nature of Statistical Learning Theory. Springer, New York (1995)
21. Walterbusch, M., Martens, B., Teuteberg, F.: Evaluating cloud computing services from a total cost of ownership perspective. Manage. Res. Rev. **36** (6), 613–638 (2013)