E-BUSINESS AND CLOUD COMPUTING SERVICES IN GREEK COMPANIES DURING ECONOMIC RECESSION

Constantinos Halkiopoulos[1], Konstantinos Giotopoulos[1], Hera Antonopoulou[1], Leonidas Theodorakopoulos[1]
[1]*Entrepreneurship & Digital Innovation Laboratory, University of Patras, Greece*
*Corresponding author: E-mail: halkion@upatras.gr

ABSTRACT: *The Cloud Computing is a technology that will totally dominate in the future. Companies such as Google, Microsoft, Amazon and others have invested heavily in new technological structure while increasingly focus their services to take advantage of the revolutionary Cloud Computing capabilities. There are many ways of how cloud computing is changing the business and hence the digital economy. Consumers also have preceded the business to use services based on the logic of the cloud, which allows better cooperation with blogs, wikis and social networks (Facebook, Twitter). Flexibility is the ability of an enterprise to change in order to react better and faster to the needs in order to provide updated and secure services to its consumers. The purpose of this paper is to study the impact of cloud computing services in E-Business during the economic recession of Greek market. Through this research there is an effort to make an approach in order to determine if it is cloud computing widely used in enterprises and how many of the companies in the field of IT and non have adopted it. The tool for recording and evaluating of the impact between the Cloud Computing Services and E-Business was a closed-ended questionnaire. In particular, companies were asked to reply to a targeted questionnaire, with questions like, whether cloud computing is functional, easy to use, and economical of course whether they find it a safe tool for storing their business data. The research was conducted in 2019 during economic recession and the sample involved was Greek companies. The information was collected by the use of social media. To extract useful results we have made statistical analysis of collected data, which held at the middle of Economical year of 2019. We have also implemented appropriate Machine Learning techniques through the use of classification algorithms, clustering and correlation analysis in order to produce inference rules. For the implementation of the Machine Learning techniques we have used the open source software Machine Learning and Data Mining Code R. Finally we came to the fact that businesses are adopting cloud computing services as backbone of their business strategy they have great potential in this way to sustainable development in the cloud era.*

KEYWORDS*: cloud computing services, e-business, social networks, r, economic crisis*

INTRODUCTION

The cloud computing services are considered a new and not fully established trend in IT. But there is no doubt the advantages it can offer to businesses and organizations and can give a great boost to the development of small and medium-sized enterprises.

There will be many ways in which Cloud Computing will change business and the economy and most are difficult to predict, but an image is already emerging. Consumers have preceded businesses in using Cloud-based services, which allows better collaboration, such as blogs, wikis and social networks (Facebook, Twitter). The first generation of people who have grown up using all of these tools now enters into the workforce.

LITERATURE REVIEW

Although in modern times concepts such as cloud or cloud computing are quite familiar to people involved in IT in their business, there are several factors that need to be taken into account in order cloud computing to contribute significantly to the development of each enterprise. The factors that must be taken into account by IT managers for cloud computing to benefit the business are the following:

- **Determination of the goals that business can be able to achieve and the reasons for this purpose**
  Ensuring the smooth operation of the company's technology resources and the use of modern and optimized services contribute effectively to the success of a business.
- **Clarifying the objectives and motivation of the enterprise as well as the objectives of the IT section**
  Switching to a cloud computing model can be hacked by the IT department as a result of pressure to diminish costs and spare resources.
- **Careful planning and preparation for possible failure**
  It may seem obvious, but business executives must of course have a careful planning and have to study how the cloud method chosen will be embedded in IT, how it will be managed and by whom, and how it will be monitored.
- **Reducing complexity is just as important as reducing costs**
  Compared to the management of information systems only internally, cloud computing may not always be a more economical option. The additional cost of access to cloud computing and the cost of re-training the staff that may be needed may prove to be unprofitable for a business.
- **Calculation of risks**
  Businesses should investigate how their data will be protected and securely maintained in the cloud.
- **Choosing the right partner**
  Collaboration with vendors specializing in the provision of cloud services is essential to ensure that they meet the requirements of each business.
- **The supplier's service contract is appropriate for the business**
  The company should check that the supplier has the experience and the ability to act promptly and quickly if an application important to the operation of the business is not temporarily accessible so that it can be reopened as soon as possible.

Research in Greek Companies

Having clarified the concept of cloud computing, we will attempt through this research to make an approach to the widespread use of cloud computing in business as well as its application to them. More specifically, businesses have been asked to answer questions such as whether cloud computing is functional, easy to use, cost-effective, and of course, whether they are safe to store their business data.

The survey was conducted over a six months period (September - December) in Greece in 2019 in a period of intense economic crisis, and the population we addressed was Greek companies. The number of businesses tending to respond was 250 and of those responded 100 Greek companies (40%).

The type of this research is ranked in qualitative terms, and this derives from the fact that the answers are not based on numerical data but are defined and shaped each time based on the data of the particular business. This research is subjective and basically relies on the researcher.

Finally, by doing the research, the limitations we encountered were that we had a fairly small sample and that there was a non-response to several of the questions because of the specificity of the subject.

METHODOLOGY

In this research paper we use a quantitative research methodology. The process of a quantitative research follows two distinct phases. Firstly, the phase of the design that we mention the objectives of the research and the assumptions may be considered. In the implementation phase is involved in the collection of data, analysis, and the interpretation of results. According to the existing research and the issue of usability we use the questionnaire as a research tool for measuring the impact of cloud computing Services in E-Business. In this paper a online questionnaire has been completed by professionals through social networks. The purpose of the survey is to assess the impact of Cloud Computing Services in Greek Companies.

Data Mining is an emerging knowledge discovery process of extracting previously unknown, actionable information from very large scientific and commercial databases. It is imposed by the explosive growth of such databases. Usually, a data mining process extracts rules by processing high dimensional categorical and/or numerical data. Classification, clustering and association are the most well known data mining tasks. Classification is one of the most popular data mining tasks. Classification aims at extracting knowledge which can be used to classify data into predefined classes, described by a set of attributes. The extracted knowledge can be represented using various schemas. Decision trees, "if-then" rules and neural networks are the most popular such schemas.

A lot of algorithms have been proposed in the literature for extracting classification rules from large relational databases, such as symbolic learning algorithms including decision trees

algorithms (e.g. C4.5) and rule based algorithms (e.g. CN2), connectionist learning algorithms (e.g. back{propagation networks), instance-based algorithms (e.g. PEBLS) and hybrid algorithms. Association rules can be used to represent frequent patterns in data, in the form of dependencies among concepts attributes. In this paper, we consider the special case, that is known as the market basket problem, where concepts-attributes represent products and the initial database is a set of customer purchases (transactions).

RESULTS

Classification methods aim to identify the classes from some descriptive traits. They find utility in a wide range of human activities and particularly in automated decision making. Decision trees are a very effective method of supervised learning. It aims is the partition of a dataset into groups as homogeneous as possible in terms of the variable to be predicted. It takes as input a set of classified data, and outputs a tree that resembles to an orientation diagram where each end node (leaf) is a decision (a class) and each non- final node (internal) represents a test. Each leaf represents the decision of belonging to a class of data verifying all tests path from the root to the leaf. The tree is simpler, and technically it seems easy to use. In fact, it is more interesting to get a tree that is adapted to the probabilities of variables to be tested. Mostly balanced tree will be a good result. If a sub-tree can only lead to a unique solution, then all sub-tree can be reduced to the simple conclusion, this simplifies the process and does not change the final result. Ross Quinlan worked on this kind of decision trees.

Decision trees are built in "ctree (Conditional Inference Trees)" by using a set of training data or data sets. At each node of the tree, "ctee" chooses one attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. Its criterion is the normalized information gain (difference in entropy) that results from choosing an attribute for splitting the data. The attribute with the highest normalized information gain is chosen to make the decision. During the construction of the decision tree, it is possible to manage data for which some attributes have an unknown value by evaluating the gain or the gain ratio for such an attribute considering only the records for which this attribute is defined. Using a decision tree, it is possible to classify the records that have unknown values by estimating the probabilities of different outcomes. Ctree builds decision trees from a set of training data in the same way as ID3 or C4.5, using the concept of information entropy.

The training data is a set $S = s_1, s_2, \cdots$ of already classified samples. Each sample s_i consists of a p-dimensional vector $(x_{1,i}, x_{2,i}, ..., x_{p,i})$, where the x_j represent attribute values or features of the sample, as well as the class in which s_i falls. At each node of the tree, "ctree" chooses the attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. The splitting criterion is the normalized information gain (difference in entropy). The attribute with the highest normalized information gain is chosen to make the decision.

The "ctree" algorithm then recurs on the smaller sublists. In order to specify the best result, it was necessary to fit the data to the model in a proper way. This task was carried away by changing and testing the controls of "ctree".

The parameters in the control function that were altered are:

- mincriterion: The value of the test statistic (for testtype == "Teststatistic"), or 1 - p-value (for other values of testtype) that must be exceeded in order to implement a split.
- minsplit: The minimum sum of weights in a node in order to be considered for splitting.
- mtry: The number of input variables randomly sampled as candidates at each node for random forest like algorithms.
- maxdepth: The maximum depth of the tree.

Tree Parameters

- ✓ Depended variable: "Business Type"
- ✓ Independed variables: "How satisfied are you from the cloud" ,"Employees", "Why do you use the cloud.", "Years Active"
  - ➢ 'mincriterion' value: 0.01
  - ➢ 'minsplit' value: 20L
  - ➢ 'mtry' value: Inf (Infinite)
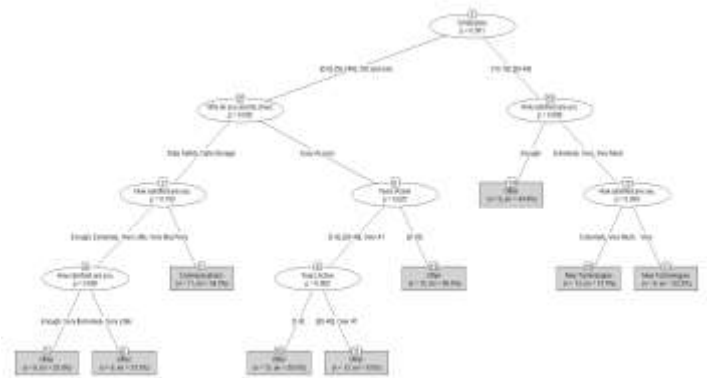 - ➢ 'maxdepht' value: Inf (Infinite)

Figure 1

Mining Association Rules

Association Rule Mining is a common technique used to find associations between many variables. In Data Mining, Apriori is a classic algorithm for learning association rules. Apriori is designed to operate on databases containing transactions (for example data collected from surveys in this case). As is common in association rule mining, given a set of item sets, the algorithm attempts to find subsets which are common to at least a minimum number C of the itemsets.

Apriori uses a "bottom up" approach, where frequent subsets are extended one item at a time, and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found. Apriori uses breadth-first search and a tree structure to count candidate item sets efficiently. It generates candidate item sets of length k from item sets of length k − 1. Then it prunes the candidates which have an infrequent sub pattern. According to the downward closure lemma, the candidate set contains all frequent k-length item sets. After that, it scans the transaction database to determine frequent item sets among the candidates.

Association rules present association or correlation between item sets. An association rule has the form of A →B, where A and B are two disjoint item sets.

The Goal: studies whether the occurrence of one feature is related to the occurrence of others. Three most widely used measures for selecting interesting rules are:

- ✓ **Support** is the percentage of cases in the data that contains both A and B,

- ✓ **Confidence** is the percentage of cases containing A that also contain B, and

- ✓ **Lift** is the ratio of confidence to the percentage of cases containing B.

Apriori rules visualization

- **Scatterplot**

This visualization method draws a two dimensional scatterplot with different measures of interestingness (parameter "measure") on the axes and a third measure (parameter "shading") is represented by the color of the points. There is a special value for shading called "order" which produces a two-key plot where the color of the points represents the length (order) of the rule.
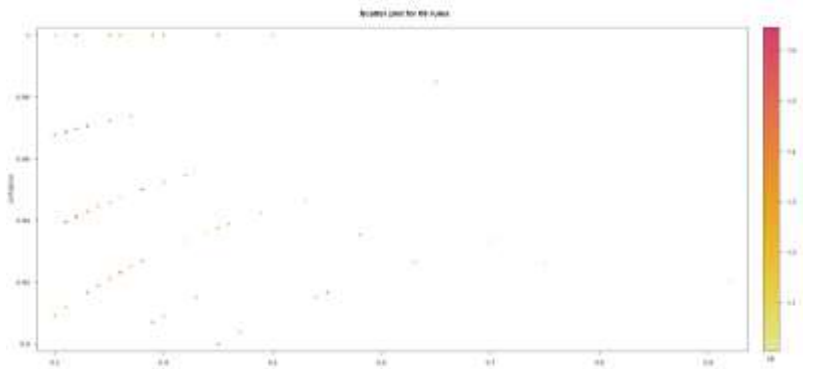

Figure 2

- **Matrix3D**

Arranges the association rules as a matrix with the item sets in the antecedents on one axis, and the item sets in the consequent on the other. The interest measure is either visualized by a color (darker means a higher value for the measure) or as the height of a bar (method "matrix3D"). Specifically of our use, the parameters that were altered are:

- ➤ measure = "lift"
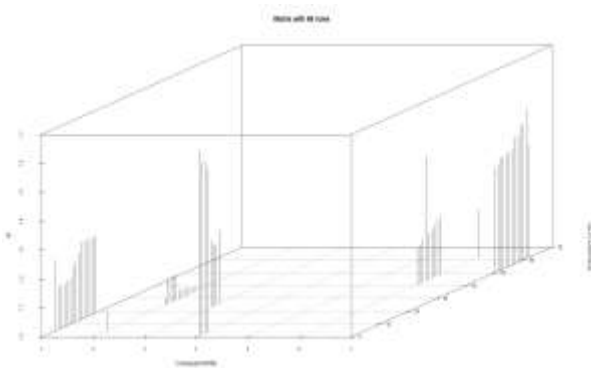
- ➤ control = list(reorder = TRUE)

Figure 3

- **Grouped Matrix plot**

Figure 4

Antecedents (columns) in the matrix are grouped using clustering. Groups are represented as balloons in the matrix.

Figure 5

Paracoord

- **Graph**

Represents the rules (or itemsets) as a graph. Specifically of our use, the parameters that were altered are:
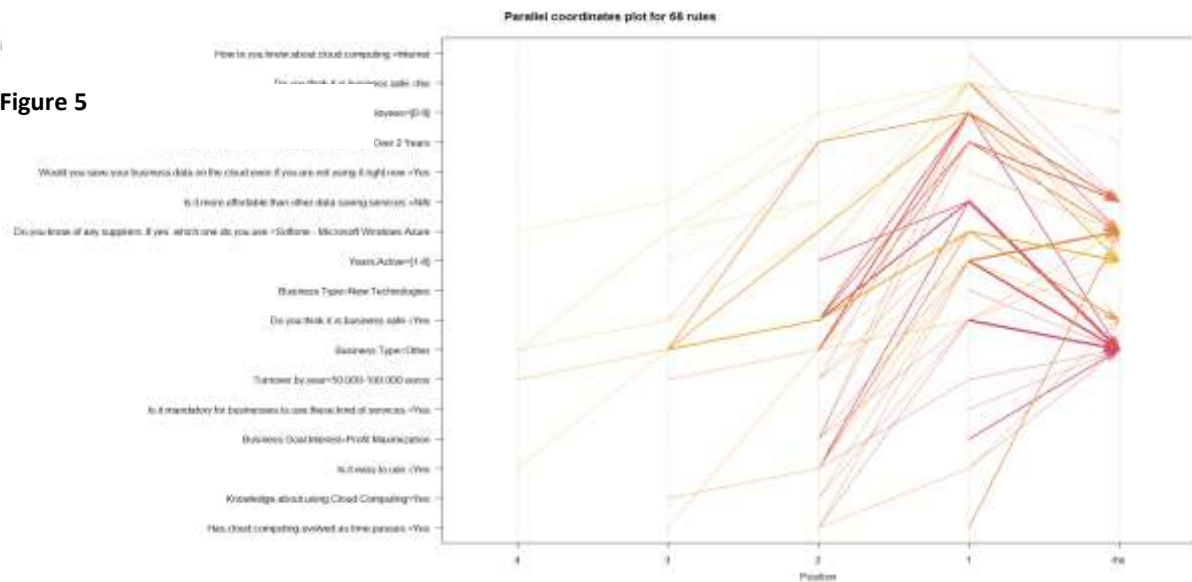
> control=list(type="items)

Figure 6

Parallel coordinate charts are a visualization that consists of N amount of vertical axes, each representing a unique data set of 82 rules, with lines drawn across the axes. The lines show the relationship between the axes, much like scatter plots, and the patterns that the lines form indicates the relationship. We can also gather details about the relationships between the axes when you see the clustering of lines. Let's take a look at this using the chart below as an example. Specifically of our use, the parameters that were altered are:

> control=list(reorder=TRUE)

Apriori rules

For the top 68 rules that were extracted from the apriori the following parameters were altered:

- support: A numeric value for the minimal support of an item set

- confidence: A numeric value for the minimal confidence of rules/association hyperedges

Specifically of our use: Support: 40%, Confidence: 90%

Here are the top 7 extracted rules

Table 1

rhs	Lhs support / confidence
1 {Knowledge.about.using.Cloud.Computing=Yes, Do.you.think.it.is.business.safe.=Yes, How.to.you.know.about.cloud.computing.=Internet} => {Is.it.easy.to.use.=Yes}	0.31 / 0.9687500
2 {Knowledge.about.using.Cloud.Computing=Yes, Would.you.save.your.business.data.on.the.cloud.even.if.you.are.not.using.it.right.now.=Yes, How.to.you.know.about.cloud.computing.=Internet, Has.cloud.computing.evolved.as.time.passes.=Yes} => {Is.it.easy.to.use.=Yes}	0.32 / 0.9411765
3 {Do.you.think.it.is.business.safe.=Yes, How.to.you.know.about.cloud.computing.=Internet, Has.cloud.computing.evolved.as.time.passes.=Yes} => {Is.it.easy.to.use.=Yes}	0.31 / 0.9393939
4 {Knowledge.about.using.Cloud.Computing=Yes, Would.you.save.your.business.data.on.the.cloud.even.if.you.are.not.using.it.right.now.=Yes, How.to.you.know.about.cloud.computing.=Internet} => {Is.it.easy.to.use.=Yes}	0.36 / 0.9230769
5 {Would.you.save.your.business.data.on.the.cloud.even.if.you.are.not.using.it.right.now.=Yes, Is.it.more.affordable.than.other.data.saving.services.=Yes, Is.it.mandatory.for.businesses.to.use.these.kind.of.services.=Yes, How.to.you.know.about.cloud.computing.=Internet} => {Do.you.think.it.is.business.safe.=Yes}	0.32 / 1.0000000
6 {Would.you.save.your.business.data.on.the.cloud.even.if.you.are.not.using.it.right.now.=Yes, Is.it.more.affordable.than.other.data.saving.services.=Yes, How.to.you.know.about.cloud.computing.=Internet, Is.it.easy.to.use.=Yes} => {Do.you.think.it.is.business.safe.=Yes}	0.33 / 0.9705882
7 {Is.it.more.affordable.than.other.data.saving.services.=Yes, Is.it.mandatory.for.businesses.to.use.these.kind.of.services.=Yes, How.to.you.know.about.cloud.computing.=Internet} => {Do.you.think.it.is.business.safe.=Yes}	0.32 0.9696970

CONCLUSION

To conclude, cloud computing is like working on a large-scale network with computers. Its main characteristic is that it is depended on various remote servers which they do the work of hosting space that it is used to save data from all over the world. Every day more and more businesses take advantage of cloud computing in order to lower their cost and boost their productivity at the same time. Furthermore, as cloud computing is becoming more mobile by

the day, business individuals use it to reach goals that effect at the same time both business and personal matters.

Most respondents were quite satisfied, and most of the sample was very satisfied. As far as the most familiar cloud systems are concerned, the icloud, Dropbox and Google Drive are known. At the same time, it is worth mentioning that the majority of respondents believe that cloud computing is a safe choice of data storage for the business, while most of the respondents have said that they think that this way of saving is cheaper.

Additionally, an important finding is that more than ¾ of the sample responded that they felt it was necessary for a company to use cloud computing services, and the most important reasons for its use by businessmen are data security and easy accessibility from everywhere, ease of storage their data. Businesses have highlighted through their responses that businesses benefit from the use of cloud computing with the most important being to store its data and security.

Finally, as far as the ease of use of cloud computing is concerned, all respondents were positive about it, while much of the sample agreed that there has been improvement in recent years with emphasis on security and speed.

References

[1] Babar, M.A., and Chauhan, M.A., (2011). A tale of migration to cloud computing for sharing experiences and observations, ACM.

[2] Buyya, R., Yeo, C.S., and Venugopal, S., (2008). Market – oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities, IEEE.

[3] Dikaiakos, M.D., Pallis, G., Katsaros, D., Mehra, P., Vakali, A. (2009). Cloud computing: distributed internet computing for IT and scientific research. IEEE Trans Internet Comput, 13:10–13.

[4] Halkiopoulos, C., Antonopoulou, H., Panagiotarou, A., Togias, P., (2016). "E-marketing and Online Advertising: Case Study in Greek companies during the economic crisis", International Conference on Contemporary Marketing Issues, 22-24 June, 2016, Heraklion, Greece.

[5] Maimon, Oded, and Rokach, Lior, (2010). "Data Mining and Knowledge Discovery Handbook", 2nd ed., Springer.

[6] Mann, R., et al., (2008). B2B E-Commerce: Why it business should incorporate e-commerce.

[7] Mell, P., and Grance, T., (2010). "The NIST Definition of Cloud Computing,", http://www.blogjava.net/zamber/archive

[8] Meng, K., (2008). A Walk in the Cloud: Uncovering Cloud Computing. China Network World, pp.12-14, 2008.6.16

[9] Pang-Ning Tan, Steinbach, Michael, and Kumar, Vipin, (2006). "Introduction to Data Mining", Addison-Wesley.

[10] R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

[11] Wu, Z. H., (2011). Cloud Computing: Analysis of the Core Technology. Posts & Telecom Press.

[12] Zheng, W. M., (2010). Opportunities and Challenges to Cloud Computing.