ANALYSIS OF THE RELATIONSHIP BETWEEN SELF-ESTEEM AND DEPRESSION

Rosalba Cuapa C.[1], Adair Ponce[2], Meliza Contreras[3], Mireya Tovar[4] and Fernando Zacarias F.[5]

[1]Faculty of Architecture, Universidad Autónoma de Puebla, Puebla, México
[2,3,4,5]Computer Science, Universidad Autónoma de Puebla, Puebla, México

ABSTRACT

This article presents a comprehensive and in-depth analysis of the intricate relationship that underlies self-esteem and depression, two crucial aspects in the field of mental health. Study aims not only to scrutinize the existing relationship between self-esteem and depression but also to do so innovatively by leveraging the analytical and data visualization capabilities offered by two cutting-edge tools: Power BI [3], [4] and Weka [1], [2]. These platforms, widely recognized in the realm of data science, allow for not only exploring the relationship between these two fundamental variables but also presenting their findings in an accessible and enlightening manner for both specialized and general audiences. Thus, our study represents a comprehensive and multidisciplinary effort to unravel the complex relationship between self-esteem and depression. Through the synergy between advanced data analysis carried out with Weka and the powerful visualization provided by Power BI, the aim is not only to broaden the understanding of this crucial relationship but also to lay the groundwork for future research and approaches in the field of mental health. In a world where mental health takes center stage in socio-health concerns, this study aspires to contribute valuable insights that can inform and enrich the care, treatment, and policies related to depression and self-esteem.

1. INTRODUCTION

Depression is a common mental disorder that affects millions of people worldwide. It is characterized by a persistent feeling of sadness, loss of interest in activities that were once pleasurable, changes in weight, trouble sleeping, fatigue, difficulty thinking and concentrating, and thoughts of death or suicide. Self-esteem, on the other hand, is the way a person sees and values themselves. It is an important dimension of mental health and can have a significant impact on a person's well-being. People with low self-esteem often have a negative view of themselves, feel insecure, and do not feel worthy of love or happiness. Several studies have shown that there is a relationship between self-esteem and depression. People with low self-esteem are more likely to develop depression than people with healthy self-esteem. This relationship has been observed in adults, adolescents, and children. The underlying mechanisms of the relationship between self-esteem and depression are not fully understood. However, some researchers believe that low self-esteem can make people more vulnerable to stressors and negative life events. Additionally, low self-esteem can make it difficult for people to cope with life's challenges and to seek help when they need it [5].

This research seeks to further explore the relationship between self-esteem and depression. The primary goal of this study is to identify the underlying mechanisms of this relationship. The researchers will use modern tools such as Power BI and Weka to conduct complex data analysis. The results of this research could have important implications for clinical practice. If the researchers can identify the underlying mechanisms of the relationship between self-esteem and depression, they could develop interventions to help people with low self-esteem reduce their risk of developing depression. Here are some specific examples of how the results of this research could be useful in clinical practice:

- Physicians could use the results of this research to identify people with low self-esteem who are at an increased risk of developing depression.
- Therapists could use the results of this research to develop interventions to help people with low self-esteem improve their self-concept and self-esteem.
- Mental health organizations could use the results of this research to develop depression prevention programs targeted at people with low self-esteem.
- The use of modern tools such as Power BI and Weka is a strength of this research. These tools can help to conduct complex data analysis in an efficient and effective manner. The use of these tools could help the researchers to identify patterns and trends that might not be visible using more traditional analysis methods.

Overall, this research is an important effort that has the potential to improve our understanding of the relationship between self-esteem and depression. The results of this research could have important implications for clinical practice and help to reduce the burden of depression in society. Given that depression affects a significant proportion of the global population, understanding its links with self-esteem stands as an objective of paramount relevance. To achieve this goal, the study collects and utilizes data from diverse sources, enabling a comprehensive and encompassing perspective. These data undergo rigorous analysis involving both advanced statistical techniques and machine learning approaches, with the central support of Weka.

However, the mere acquisition and analysis of data are not sufficient to generate substantial impact. This is where Power BI comes into play, playing a pivotal role in transforming the analysis results into clear and effective visualizations. The platform not only allows the creation of interactive charts and dynamic dashboards but also facilitates the effective communication of findings to diverse audiences, including mental health professionals, researchers, policymakers, and the public.

2. PROBLEM DESCRIPTION AND FORMULATION

To carry out this comprehensive analysis, the researchers collected data related to self-esteem and depression from a variety of sources[6], including:

Clinical studies: The researchers reviewed the scientific literature to identify clinical studies that investigated the relationship between self-esteem and depression. These studies provided data on self-esteem and depression levels in groups of people with different demographic and psychological characteristics.

Research databases: The researchers also obtained data from research databases that collect information on people's mental health. These databases provided data on self-esteem and depression levels in a broader population, including people who have not participated in clinical studies.

The researchers used advanced data analysis and visualization techniques to analyse the collected data. These techniques included:

- Statistical analysis: The researchers used statistical analysis to identify relationships between self-esteem and depression levels. These analyses included correlation analysis, regression analysis, and cluster analysis.
- Data visualization: The researchers used data visualization to represent the results of the statistical analysis in an easy-to-understand way. The researchers created graphs, charts, and maps to show the relationships between self-esteem and depression levels.
- The combination of Power BI and Weka provided a solid platform for exploring relationships and patterns within the data. Power BI is a data analysis and visualization tool that allows researchers to explore data interactively. Weka is a machine learning tool that allows researchers to identify patterns in data.

2.1. Tools: Weka

Weka is a suite of software tools used for machine learning and data mining. The name "Weka" comes from the acronym for "Waikato Environment for Knowledge Analysis", referring to the University of Waikato in New Zealand, where it was originally developed by the university's Data Mining Group. Weka provides a graphical user interface and a set of pre-implemented machine learning algorithms that allow users to perform tasks such as classification, regression, clustering, and association rule mining, among others. In addition to its graphical user interface, Weka also offers a Java API (Application Programming Interface), which makes it easy to integrate its functionality into broader applications. Weka is a powerful and versatile tool that can be used to solve a wide range of data mining and machine learning problems. It is widely used by researchers, academics, and practitioners in a variety of fields, including healthcare, finance, and marketing.

Here are some of the key features of Weka [1]:

- Data Handling: Weka supports various data formats, including CSV, ARFF (Attribute-Relation File Format), and more. It provides tools for data preprocessing, cleansing, and transformation, making it easier to prepare data for analysis.
- Graphical User Interface (GUI): Weka offers a user-friendly GUI that allows users to interact with the software without needing to write code. This GUI provides access to various machine learning algorithms, preprocessing tools, and data visualization capabilities.
- Core Libraries: At the heart of Weka are its core libraries, which contain the implementation of numerous machine learning algorithms, data preprocessing techniques, and evaluation methods. These libraries are the foundation upon which Weka's functionality is built.
- Extensibility: Weka's architecture is highly extensible, allowing users to incorporate their own custom machine learning algorithms, data filters, and evaluation metrics. This extensibility is achieved through a well-defined API and plugin system.
- Cross-Validation and Evaluation: Weka includes functionality for cross-validation and model evaluation, allowing users to assess the performance of their machine learning models. It provides metrics like accuracy, precision, recall, F1-score, and others for model assessment.
- Integration with Other Tools: Weka can be integrated with other data analysis and visualization tools, enhancing its capabilities, and enabling seamless workflows.

Weka is a valuable tool for anyone who needs to analyse data and extract insights. It is a powerful and versatile tool that can be used to solve a wide range of problems.

2.2. Tools: Power BI

Power BI is a unified and scalable self-service business intelligence (BI) platform suitable for large enterprises. Connect to data, visualize it, and seamlessly embed visuals into the apps you use every day.

Power BI is a set of tools that makes knowledge available to everyone and gives us access to our data safely and quickly, generating great benefits for us and for our company. It is a predictive, intelligent, and highly supportive system, capable of translating data (simple or complex) into graphs, panels or reports due to its qualities such as the graphical capacity of presenting information, or the integration of Power Query: the search engine extract, transform, and load (ETL) software included in Excel.

Here are some of the key features of Power BI:

- **Easy handling:** you will not need to have advanced knowledge to create your own dashboards and reports. Simply build, edit, publish and share to collaborate with whoever you need.
- **Integration capacity:** Take advantage of the integration capacity of data sources as different as those that Salesforce, MailChimp, SAP BW, SAP HANA, MySQL, SQL Server, Teradata, Oracle, Google, Twitter, Facebook, Dynamics 365 can provide you.
- **Attractive graphics:** Visualize your data with amazing, animated infographics or if you prefer something more traditional, use the usual graphics.
- **Seguridad:** Con power BI podrás establecer perfiles o roles de acceso a la información para cada miembro de tu empresa.
- **Información en tiempo real:** Si es necesario podrás programar la descarga de datos para tus KPI´s tanto como necesites y acceder a información actualizada.

2.3. Comparative Analysis of Business Intelligence platforms

Based on comparative analyzes presented on BI tools such as: Joaquim Lapa, et.al. [12], Nuno Leite, et.al. [13], etc. and considering the criterion used in Gartner [14] we can paraphrase it as follows: The analysis of BI platforms should be considered from three broad categories: integration, information delivery and analysis. These criteria allow organizations to have clear classification and measurement systems to support decision making. In addition, we must consider the facilities and benefits that these tools provide us.

3. EXPERIMENTAL STUDY

the data were obtained from multiple sources revealed a notable negative correlation between self-esteem and depression. Individuals with low self-esteem exhibited a higher incidence of depressive symptoms compared to their counterparts with higher self-esteem. Additionally, a mediating role played by self-esteem was observed in the relationship between certain risk factors and depression. This finding underscores the importance of self-esteem as an influential variable in the development of depression.

3.1. Database Design

The database was designed for a type of archive native for weka, the type of this archive is ARFF:

- **@relation:** This command defines the name or title of the dataset. It specifies the dataset's overall name or a brief description.
- **@attribute:** This command defines the attributes or columns of the dataset. It specifies the names of the attributes, their types (such as numeric, string, or nominal), and their possible values or value ranges.
- **@data:** This command marks the beginning of the data section of the ARFF file. In this section, actual data instances are listed, with each instance's values corresponding to the attributes defined earlier.

The database was recovered from the official page of "The INEGI" (Instituto Nacional de Estadística y Geografía) is the Mexican National Institute of Statistics and Geography. It is a government agency responsible for collecting, analyzing, and disseminating statistical information and geospatial data for Mexico. INEGI plays a crucial role in providing accurate and up-to-date information on various aspects of the country, including population, economics, geography, and demographics.

3.2. First Source of Information

Methods: a cross-sectional study was carried out in 1,104 freshman students, 40.3% men and 59.7% women, at a public university in Mexico City. The 20-item depression scale (CES-D) and Food Frequency Questionnaire were applied to measure depressive symptoms and food consumption. Logistic regression analysis was carried out for food consumption frequency and CES-D depression score grouped in quartiles [7].

Results: the prevalence of depression symptoms was 18.2% in men and 27.5% in women (p < 0.001). A considerable proportion of the students reported poor eating habits: consumption of fried food (30.3%), sweetened drinks (49.0%) and sugary food (51.8%) 2-7 times/week; and less than half the students practiced vigorous physical activity (39.7%). In women, a higher depression score was associated with a higher frequency of consumption of fast food (OR = 2.08, p = 0.018), fried food (OR = 1.92, p = 0.01) and sugary food (OR = 2.16, p = 0.001), and a lower frequency of physical exercise (< 75 min/week; OR = 1.80, p = 0.017). In men, no association was observed between depression score and food consumption variables. An association was observed between depression and low exercise frequency (OR = 2.22, p = 0.006).

3.3. Second Source of Information

Data collected by the National Institute of Statistics and Geography (INEGI) of Mexico also provided crucial information for this analysis. The numbers and statistics provided by this entity enriched the understanding of the prevalence of depression in the population and how it may be related to self-esteem.

Questions asked in the survey:

1. I couldn't shake the sadness.
2. Never or rarely: I never felt this way.
3. A few or sometimes: I felt this way a few times.
4. A considerable number of times: I felt this way a lot of times.
5. All the time or most of the time: I felt this way all the time or most of the time.
6. Difficulty concentrating.
7. Never or rarely: I never had difficulty concentrating.
8. A few or sometimes: I had difficulty concentrating a few times.

9. A considerable number of times: I had difficulty concentrating a lot of times.
10. All the time or most of the time: I had difficulty concentrating all the time or most of the time.
11. Feeling depressed.
12. Never or rarely: I never felt depressed.
13. A few or sometimes: I felt depressed a few times.
14. A considerable number of times: I felt depressed a lot of times.
15. All the time or most of the time: I felt depressed all the time or most of the time.
16. Everything felt like an effort.
17. Never or rarely: I never felt like everything was an effort.
18. A few or sometimes: I felt like everything was an effort a few times.
19. A considerable number of times: I felt like everything was an effort a lot of times.
20. All the time or most of the time: I felt like everything was an effort all the time or most of the time.
21. Didn't sleep well.
22. Never or rarely: I always slept well.
23. A few or sometimes: I had trouble sleeping a few times.
24. A considerable number of times: I had trouble sleeping a lot of times.
25. All the time or most of the time: I had trouble sleeping all the time or most of the time.
26. Enjoying life.
27. Never or rarely: I always enjoyed life.
28. A few or sometimes: I enjoyed life a few times.
29. A considerable number of times: I enjoyed life a lot of times.
30. All the time or most of the time: I enjoyed life all the time or most of the time.
31. Feeling sad.
32. Never or rarely: I never felt sad.
33. A few or sometimes: I felt sad a few times.
34. A considerable number of times: I felt sad a lot of times.
35. All the time or most of the time: I felt sad all the time or most of the time.

3.4. Third Source of Information

Another important study, conducted by Sánchez-Rojas et al. (2022), examined children and adolescents with and without overweight or obesity. The study highlights the correlation between self-image, self-esteem, and depression. The results show that self-image and self-esteem influence the presence of depressive symptoms in these groups. This study provides valuable insights into how body image and self-evaluation can influence mental health [9], [10].

Methods: A cross-sectional comparative study of 295 children: 116 children with overweight or obesity (group 1) and 179 without obesity (group 2). Body mass index, schooling, academic achievement, school conflicts, socioeconomic status, self-image (current, desired), satisfaction, self-esteem, and presence of depression were recorded. Descriptive statistics, Spearman's rho, and Mann-Whitney U were used; $p \leq 0.05$ was considered significant.

Results: In group 1, 53.4% of the children perceived themselves as normal weight and 77.6% had a desired self-image of normal weight; 67.2% wanted to be thinner; 53.4% had high self-esteem; 75.9% were observed without depression. In group 2, 79.3% had a current self-image of normal weight and the desired self-image in 85.5% was normal weight; 35.2% wanted to be thinner; self-esteem was high in 49.7% and 77.1% did not present depression. Significant correlations were identified in self-esteem-depression ($r = 0.228$) and self-esteem-socioeconomic status ($r = 0.130$). The data obtained from multiple sources revealed a notable negative correlation between self-esteem and depression. Individuals with low self-esteem exhibited a higher incidence of

depressive symptoms compared to their counterparts with higher self-esteem. Additionally, a mediating role played by self-esteem was observed in the relationship between certain risk factors and depression. This finding underscores the importance of self-esteem as an influential variable in the development of depression.

Table 1. Data Within Weka, showing the minimum, maximum, mean, and standard deviation table at the top.
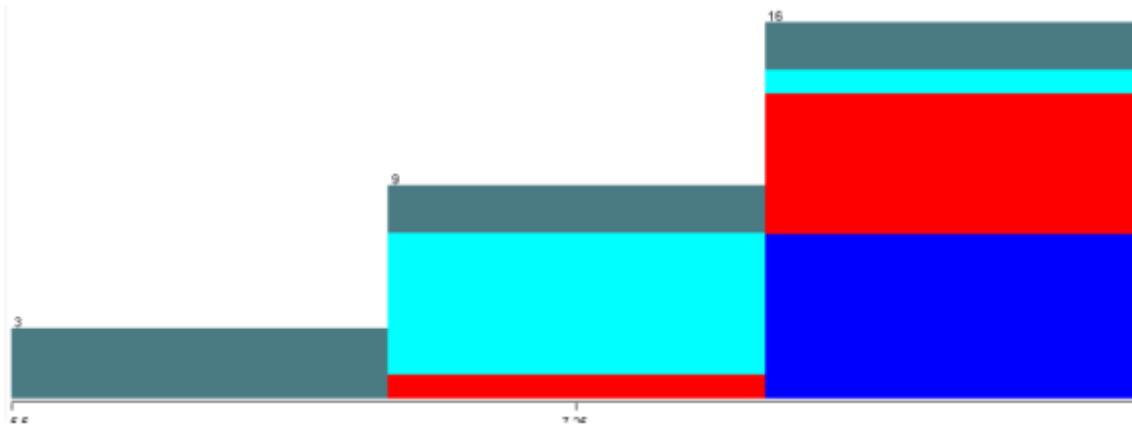
Figure 1. Attribute "Mujer"

Table 2. Data Within Weka, showing the minimum, maximum, mean, and standard deviation table at the top.
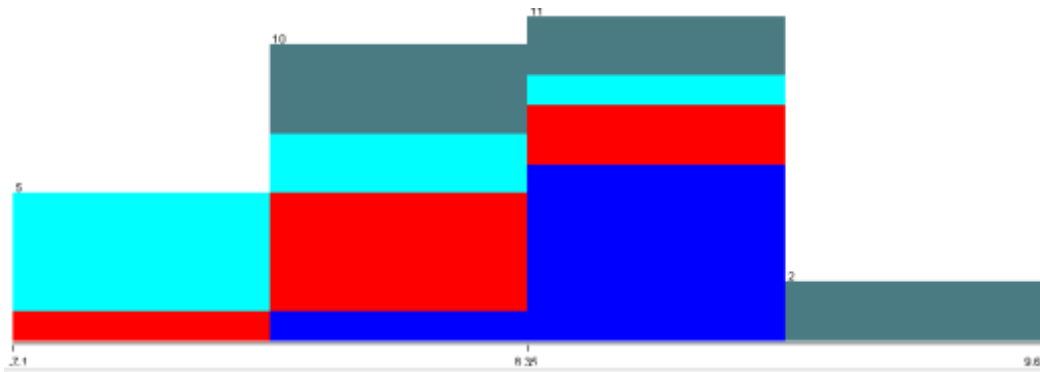
Figure 2. Attribute "Hombre"

Figure 3. File ARFF

Average life satisfaction (1) of the literate population aged 18 and over (2), by symptoms of depression (3) according to sex and age groups (18 to 29 years old).

(1) To measure the variable, a card with a scale of 0 to 10 is used, in which 0 means "totally dissatisfied" and 10 "totally satisfied".

(2) The ENBIARE interview process requires Spanish reading and writing skills on the part of the informant.

(3) To capture this topic, a card with a textual and numerical scale is used, presenting the numbers from 0 to 3 and the meaning of each one: 0 = rarely or never (less than 1 day), 1 = few or sometimes (1-2 days), 2 = a considerable number of times (3-4 days), and 3 = all the time or most of the time (5-7 days).

The dependent variable in this case is named "Value", it depends on the variable "Man" and "Woman", in addition to the variable "SymptomsDepresion" and "Frequency", of these depends "Man" and "Woman".

The INEGI data collection methodology involves a variety of sources and techniques. These include household surveys, establishment surveys, population and economic censuses, among others. Each of these sources provides a unique perspective on the reality of Mexico and its inhabitants. Through the collection of data in diverse contexts, INEGI seeks to guarantee the accuracy and representativeness of its statistical information. The data sets generated by INEGI have a significant impact on decision-making at the government, economic, and social levels. The results of surveys and statistical analyzes provide crucial information for understanding demographic, economic, and social trends in Mexico. These data are used to inform public policies, guide investments, evaluate programs and projects, and make informed business decisions. That is, these data that it provides us in themselves are the DataSets that INEGI generates, they are a source to obtain DataSets, in this case they are not processed because it shares them in a way that is difficult to analyze, however, in this case what was done was to adapt them to an ARFF file to be able to be analyzed with the Weka tool. The importance of INEGI's data sets also lies in the transparency and access to the information they provide. The availability of detailed and up-to-date statistical data allows researchers, academics, and citizens in general to analyze the information for themselves and generate new knowledge from it. This contributes to a greater understanding of the problems and challenges facing Mexico and promotes a dialogue based on solid data.

4. CONCLUSIONS

This comprehensive analysis underscores the need to consider self-esteem as a critical element in depression research. Self-esteem not only appears to play an essential role in the onset and development of depression, but it can also influence the effectiveness of therapeutic interventions. The results highlight the importance of taking self-esteem into account in the design of depression prevention and treatment strategies. It is essential to recognize that this analysis has inherent limitations. The data were collected from various sources with different methodologies, which can affect the quality and consistency of the results. In addition, the correlational nature of the data prevents definitive causal relationships from being established. Longitudinal and controlled research is suggested to delve into the relationship between self-esteem and depression.

5. FUTURE DIRECTIONS

This report provides a solid perspective on the relationship between self-esteem and depression and highlights the usefulness of visualization tools such as Power BI and Weka to better understand data patterns. However, to achieve a deeper and more accurate understanding of this relationship, a continuous focus on further research and the improvement of data analysis techniques is required. A full understanding of how self-esteem impacts depression could lead to significant advances in the diagnosis and treatment of this disorder [11].

REFERENCES

[1] ¿Qué es Weka y qué tiene que ver con Big Data? (2020, julio 9). Agenciab12.mx. https://agenciab12.mx/noticia/que-es-weka-que-tiene-que-ver-big-data.

[2] Donatella Merlini, Martina Rossini, Text categorization with WEKA: A survey,Machine Learning with Applications,Volume 4,2021,100033,ISSN 2666-8270, https://doi.org/10.1016/j.mlwa.2021.100033. (https://www.sciencedirect.com/science/article/pii/S2666827021000141)

[3] Gurpreet Singh, Ankul Kumar, Jaspreet Singh, Jagdeep Kaur. (2023) Data Visualization for Developing Effective Performance Dashboard with Power BI. 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), pages 968-973.

[4] Cree informes sofisticados y comparta conocimientos que impulsen los resultados. (s/f). Microsoft.com. Recuperado el 24 de agosto de 2023, de https://powerbi.microsoft.com/es-mx/landing/free-account/

[5] Arévalo García, E., Castillo-Jimenez, D. A., Cepeda, I., López Pacheco, J., & Pacheco López, R. (2019). Anxiety and depression in university students: relationship with academic performance. Interdisciplinary Journal of Epidemiology and Public Health, 2(1), e–022. https://doi.org/10.18041/2665-427X/ijeph.1.5342 (Original work published July 15, 2020)1.

[6] Mishra P, Pandey CM, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. Ann Card Anaesth. 2019 Jan-Mar;22(1):67-72. doi: 10.4103/aca.ACA_157_18. PMID: 30648682; PMCID: PMC6350423.

[7] Sánchez-Rojas, Alma A., García-Galicia, Arturo, Vázquez-Cruz, Eduardo, Montiel-Jarquín, Álvaro J., & Aréchiga-Santamaría, Alejandra. (2022). Autoimagen, autoestima y depresión en escolares y adolescentes con y sin obesidad. Gaceta médica de México, 158(3), 124-129. Epub 28 de septiembre de 2022. https://doi.org/10.24875/gmm.21000817.

[8] Lazarevich I, Irigoyen Camacho ME, Velázquez-Alva MC, Flores NL, Nájera Medina O, Zepeda Zepeda MA. Depression and food consumption in Mexican college students. Nutr Hosp. 2018 May 10;35(3):620-626. English. doi: 10.20960/nh.1500. PMID: 29974771. https://pubmed.ncbi.nlm.nih.gov/29974771/

[9] Instituto Nacional de Estadística y Geografía (INEGI). (2022). Buscador INEGI. https://www.inegi.org.mx/app/buscador/default.html?q=Depresión.

[10] Instituto Nacional de Estadística y Geografía (INEGI). (s/f). Quiénes somos. Recuperado el 19 de agosto de 2023, de https://www.inegi.org.mx/inegi/quienes_somos.html.

[11] Fan, C., Chen, M., Wang, X., Wang, J., & Huang, B. (2021). A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. Frontiers in Energy Research, 9, 652801

[12] Joaquim Lapa, Jorge Bernardino and Ana Figueiredo. A comparative analysis of open-source business intelligence platforms. Conference ISDOC: Information Systems and Design of Communication 2014. DOI: 10.1145/2618168.2618182

[13] Nuno Leite, I. Pedrosa, Jorge Bernardino. Comparative evaluation of open source business intelligence platforms for SME Published 2018 13th Iberian Conference on Information Systems and Technologies (CISTI).

[14] Group, Gartner. Magic Quadrant for Business Intelligence andAnalytics Platforms. http://www.gartner.com. [Online] 2013.

AUTHORS

Fernando Zacarias-Flores is researcher and professor of Computer Science at the Autonomous University of Puebla. He is a researcher in practical and theoretical Computer Science and Mobile Technologies and has conducted R&D projects in this area since 2000. Results from these projects have been reported in more than 80 national and international publications.