v. 4, n. 26, 2024

Journal of
Engineering
Research

OCR-BASED FRAMEWORK FOR CALCULATING THE ROI OF RESIDENTIAL SOLAR ENERGY INSTALLATIONS

Acceptance date: 04/11/2024

Christiano Talhate
Instituto Federal do Espirito Santo, Serra, Brasil

Renner Sartório Camargo
Instituto Federal do Espirito Santo, Serra, Brasil

Abstract: This project proposes the development of a framework based on Optical Character Recognition (OCR) to calculate the Return on Investment (ROI) of residential solar energy installations. The importance of such a system stems from the growing adoption of renewable energy solutions in homes, aimed not only at environmental sustainability, but also at economic efficiency. In this context, the calculation of ROI emerges as a critical factor in assessing the viability and performance of solar energy investments. The core of this project is the combination of advanced image processing and machine learning techniques. This dual approach makes it possible to automate the collection and analysis of the data needed to calculate ROI, overcoming traditional methods that are often manual, time-consuming and prone to errors, because through OCR, the system is able to extract relevant information from documents and images, such as energy bills. This data, once scanned and processed, provides a solid basis for subsequent analysis. The next stage involves the use of machine learning algorithms, which are trained to interpret the extracted data and perform ROI-related calculations.
Keywords: ROI; Solar Energy; Residential Installation; OCR; Return On Investment;

INTRODUCTION

Technological advances and growing awareness of environmental issues have driven the search for more sustainable and economically viable energy alternatives. Within this context, residential solar energy is emerging as a promising solution, offering a way of reducing dependence on fossil fuels and minimizing low environmental impact.

However, the adoption of this technology depends not only on its environmental viability, but also on its economic efficiency. It is therefore essential to assess the Return on Investment (ROI) to ensure that the financial benefits justify the initial investment.

Recognizing this need, this project aims to develop a framework based on Optical Character Recognition (OCR) and machine learning techniques to automate the collection and analysis of data needed to calculate the ROI of residential solar energy installations.

This system allows an approach to how solar energy investments are evaluated, but also facilitates informed decision-making by homeowners, thus encouraging the wider adoption of renewable energy solutions.

The motivation behind this project lies in the growing need for sustainable and economically viable energy solutions in the context of the global climate and the depletion of natural resources.

Solar energy, one of the most promising and abundant renewable sources, presents itself as a viable alternative for reducing dependence on fossil fuels and mitigating greenhouse gas emissions. However, large-scale adoption of residential solar systems is often hindered by the complexity and uncertainty associated with calculating Return on Investment (ROI).

This challenge highlights the need for tools capable of simplifying and automating the financial evaluation of solar energy installations, making it more accessible and understandable for homeowners and investors.

Therefore, its development comes as a direct response to this demand, seeking not only to promote environmental sustainability, but also to guarantee the economic viability of investments in renewable energy, thus encouraging a faster and more effective transition to clean energy sources.

This research project aims to fill several important knowledge gaps in the field of residential solar energy and the financial analysis of renewable energy projects, bringing significant contributions in several areas, such as automating the financial analysis of solar energy installations, as currently, many of the processes involved in evaluating ROI are manual, time-consuming and prone to errors.

This project can significantly simplify these processes, increasing the accuracy and efficiency of data collection and analysis.

The integration of diverse data also addresses the difficulty of integrating and analyzing data from diverse sources and in varied formats, a common issue in the evaluation of solar energy projects. By developing advanced image processing techniques and machine learning models, the proposed system can extract and synthesize information from a wide range of documents, filling the knowledge gap related to heterogeneous data management.

We also identified adaptation to different regulatory and market contexts, which is why the ability to adapt financial analysis to different market contexts is another knowledge gap that the project aims to overcome.

Solar energy systems operate in a complex environment, subject to significant variations in terms of government policies, tax incentives and market conditions. The framework developed seeks to incorporate flexibility and adaptability, allowing for accurate ROI calculations under different scenarios.

And finally, we can say that the project in question allows the democratization of access to the financial analysis of solar energy installations, since the complexity and lack of accessible tools limit the ability of homeowners and small investors to adequately assess the financial potential of investments in solar energy.

In this way, by providing an intuitive interface and clear, accurate results, the proposed system has the potential to empower a wider audience, promoting wider adoption of renewable energy solutions.

In summary, this project has the potential to fill critical gaps in knowledge related to automation and accuracy in the financial analysis of solar energy installations, management of diverse data, adaptability to changing contexts and democratizing access to financial information, thus helping to speed up the energy transition to more sustainable sources.

MATERIAL AND METHODS

The following methods were used to carry out this research:

Data Collection and Preparation: We began by collecting a wide range of documents related to the solar energy installation, including energy bills, obtained through OCR, and technical specifications, obtained from a graphical interface where the user enters the values of the respective variables.

OCR application: We used OCR tools to extract text from the prepared documents. This step was crucial for transforming unstructured information into analyzable data. The accuracy of the extraction was validated through manual review and adjustments, guaranteeing the quality of the data for subsequent analysis.

USER INTERFACE IMPLEMENTATION:

We have developed an intuitive and accessible interface, allowing users to easily enter data and access ROI analyses. This interface has been designed to be usable by a wide audience, without the need for advanced technical knowledge.

Through this methodological approach, the project aims not only to develop a tool, but also to ensure that it is practical and accessible to facilitate decision-making in residential solar energy investments.

The most common way to calculate the Return on Investment (ROI) of a residential solar energy installation involves a few basic steps and takes into account various factors, such as the total cost of the solar system, savings on the electricity bill, tax incentives and subsidies, as well as the depreciation of the system over time. We have therefore defined the following variables for calculating ROI:

Calculating the Total Cost of the System: Includes the purchase price of the solar panels, inverters, mounting materials, labor for the installation, and any other additional costs, such as permit or inspection fees.

Estimated Annual Energy Savings: This can be done by comparing your current energy bills with the estimated energy production of the solar system over the course of a year. These savings can vary based on the location, orientation and inclination of the panels, the efficiency of the system and the price of energy in your area.

Consideration of Incentives and Subsidies: Many governments offer tax incentives, subsidies or other forms of financial support for the installation of solar energy systems. These benefits should be subtracted from the total cost of the system to obtain the net cost of the investment.

ROI calculation: ROI is calculated by subtracting the net investment cost (after incentives and subsidies) from the total estimated energy savings over the expected lifetime of the solar system. The result is then divided by the net investment cost, and the value obtained is multiplied by 100 to be expressed as a percentage.

This is how we arrived at the following formula for calculating ROI:

$$ROI = \frac{\text{Total Energy Savings - Net Investment Cost}}{\text{Net Investment Cost}} \times 100$$

Consideration of the Payback Period: The payback period is the time needed for the savings generated by the solar installation to cover the initial investment. This gives you an idea of how long it will take to start "profiting" from your investment.

It is important to consider that the accuracy of these calculations depends on the quality of the cost and savings estimates. In addition, the value of the property and the reduction in carbon emissions, although not directly quantified in the ROI calculation, are additional benefits that many owners take into account when evaluating the investment in solar energy.

To define the software architecture for this project, we used the C4 framework, adapted from Brown (2006), which is an effective approach for visualizing and communicating software architecture.

It provides a set of simple but powerful diagrams that allow software architects to represent the various perspectives of a system in a clear and concise way, starting with a high-level context and delving into increasingly specific details, helping to organize and structure complex information, allowing a quick and efficient understanding of the system.

The container diagram in Figure 1 shows the high-level form of the software architecture and how responsibilities are distributed across it. It also shows the main technology options and how containers communicate with each other. It is a simple, high-level technology-focused diagram that is useful for software developers and support/operations staff.

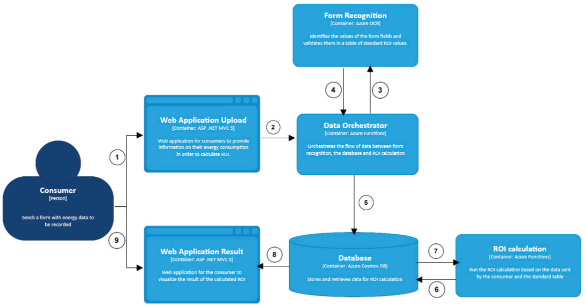

Figure 1- Project software architecture

The concept of framework was used in this project, and according to Gimenes & Huzita (2005), one of the main objectives of Software Engineering is reuse. By reusing software, it is possible to increase quality and reduce development effort.

By adopting a solution based on the framework model, it is possible to rely on pre-developed functionalities and modules, such as machine learning algorithms, data pre-processing methods, visualization of results, among others.

This can significantly speed up project development, allowing developers to focus more on the specifics of the problem at hand, since the framework presented seeks to encapsulate a series of complex techniques and concepts, providing interfaces and tools that simplify the implementation of these techniques, allowing developers with different levels of experience to work on advanced projects.

In addition to these forms of reuse, framework technology allows a family of products to be generated from a single structure that captures the most general concepts of the family of applications (Pinto, 2000).

Frameworks can be classified in various ways. Initially, they are classified into two main groups: object-oriented application frameworks (Fayad et al., 1999a, b; Fayad & Johnson, 2000) and component frameworks (Szyperski, 1997), which by definition, this project will be classified as a component framework.

The main benefits of using frameworks, according to Fayad et al. (1999b) and Pinto (2000), come from the modularity, reusability, extensibility and inversion of control that frameworks provide.

According to Mattsson (1996, 2000), a framework is an architecture developed with the aim of achieving maximum reuse, represented as a set of abstract and concrete classes with great potential for specialization.

Another technology involved is cloud computing, which is an approach that involves the design and organization of computing, storage, network and service resources in an online infrastructure on the Internet, as well as providing the advantage of ease of integration and collaboration.

Cloud architecture allows interoperability between different systems and services, facilitating the integration of applications and the sharing of data between them, boosting collaboration between teams, suppliers and partners, promoting innovation and business agility.

To build this project we used the following components identified below in the Microsoft Azure cloud computing environment, adapted from Altaiar(2022):

App Service: is a PaaS (platform as a service) offering in Azure. You can use App Service to host web applications that can be scaled or horizontally scaled manually or automatically. The service supports various languages and frameworks, such as ASP.NET, ASP.NET Core, Java, Ruby, Node.js, PHP and Python.

Azure Functions: is a serverless computing platform that you can use to create applications. With Functions, you can use triggers and associations to react to changes in Azure services, such as Blob Storage and Azure Cosmos DB. Functions can execute scheduled tasks, process data in real time and process message queues.

Forms Recognition: part of Azure Applied AI Services. Forms Recognition offers a collection of pre-built endpoints for extracting data from invoices, documents, receipts, ID cards and business cards. This service maps each piece of extracted data to a field as a key-value pair. Form Recognition also extracts the content and structure of the table. The output format is JSON.

Azure Storage: cloud storage solution that includes object, blob, file, disk, queue and table storage. Blob Storage is a service that is part of Azure Storage. Blob Storage offers optimized cloud object storage for large amounts of unstructured data.

Once the components have been defined, the architecture of the solution is as follows, shown in figure 2.

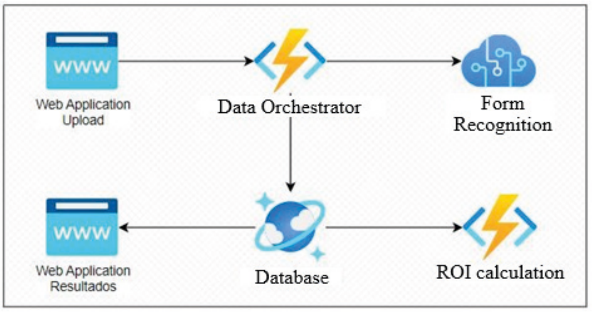

Figure 2-Project implemented in Microsoft Azure

RESULTS AND DISCUSSION

The study achieved considerable results, demonstrating significant progress in optimizing the financial evaluation of renewable energy projects, highlighting both technical innovations and practical impacts.

As a result, we will have created an effective, automated system for extracting data from documents related to the installation of solar energy, such as electricity bills, contracts and technical specifications for the equipment.

Using OCR and cloud computing technologies, it has shown excellent accuracy and efficiency in converting physical and digital documents into analyzable data. This success represents a significant step towards reducing human error and increasing efficiency in data collection, for example for financial analysis.

The development of this framework presents significant technical challenges, especially with regard to the accuracy of the OCR in dealing with the diversity and complexity of the data. However, the

Overcoming this challenge has the potential to transform the way homeowners and investors assess the financial impact of solar energy installations.

Furthermore, the successful implementation of this project could set a new standard for the analysis of renewable energy investments, offering valuable insights for stakeholders at different levels, from individual consumers to policymakers and large-scale investors.

By providing an automated tool to calculate ROI, the framework facilitates informed decisions, thus promoting the adoption of solar energy and contributing to the global transition towards more sustainable energy sources.

This project aims not only to optimize the financial evaluation of solar energy projects, but also to encourage their adoption by making investment analysis more accessible and reliable. As a result, such innovation is expected to play a crucial role in promoting environmental sustainability and maximizing the economic benefits of renewable energy for consumers.

From a practical point of view, the framework developed has had an impact on the way solar energy investments can be evaluated by homeowners and investors. Through an intuitive and accessible user interface, the system has made it possible to quickly and accurately analyze ROI, making the financial evaluation of solar energy projects more transparent and accessible.

This advance has made it significantly easier to make informed decisions about adopting solar energy, and could contribute to an increase in the implementation of residential solar systems in a more reliable way.

By providing a reliable tool for calculating ROI, the project has helped to demystify investment in solar energy, encouraging more people to consider this sustainable energy generation option.

In terms of contributions to sustainability, the project plays a vital role in promoting solar energy as a viable and financially attractive alternative to fossil fuels.

By facilitating the calculation of ROI, the system not only promotes the adoption of renewable energies, but also supports the transition to a low-carbon economy by reducing dependence on non-renewable energy sources and mitigating the environmental impact associated with energy production.

Journal of Engineering Research ISSN 2764-1317 DOI https://doi.org/10.22533/at.ed.31742624011110

CONCLUSION

The completion of this project marks a significant advance in the intersection of technology and sustainability, demonstrating the transformative potential of automated tools in evaluating investments in renewable energy.

The development of this project represents an important milestone in facilitating the energy transition to more sustainable sources. This work not only attests to the technical and economic viability of residential solar energy solutions, but also highlights the importance of technological innovations in overcoming barriers to the adoption of renewable energies.

The accuracy achieved in data extraction and subsequent ROI calculations demonstrates that it is possible to provide reliable financial assessments, essential for informed decision-making by homeowners and investors.

In addition, the system's adaptability to different regulatory and market contexts highlights its applicability in a variety of geographical regions, significantly increasing its reach and impact.

This flexibility ensures that the framework can contribute to the expansion of solar energy in different jurisdictions, effectively tackling the various economic and regulatory barriers.

This usability is key to democratizing access to solar energy investment analyses, allowing a broader spectrum of society to participate in the energy transition.

This project has clearly opened up avenues for future research and development. Integrating real-time data on energy tariffs and weather conditions, for example, could offer even more accurate and personalized ROI estimates. In addition, expanding the framework to cover other types of renewable energy investments, such as wind or biomass energy systems, could significantly increase its impact.

In this way, this project not only achieved its initial objectives of developing an automated system to calculate the ROI of residential solar energy installations, but also demonstrated the potential of advanced technologies to drive the adoption of sustainable energy solutions. By simplifying and making more accessible the financial analysis of such investments, the framework makes a significant contribution to accelerating the global energy transition.

As we move towards a more sustainable future, continued technological innovation, together with the adoption of favorable policies, will be crucial to overcoming the remaining challenges and seizing the opportunities that renewable energies present for society and the environment.

ACKNOWLEDGMENTS

To Professor Dr. Renner Sartório Camargo, whose expertise, patience and guidance were fundamental at every stage of this study.

REFERENCES

Brown, Simon (2006). **C4 Model for Visualizing Software Architecture.** Leanpub. Disponível em: https://s3.amazonaws.com/static.codingthearchite cture.com/presentations/aotb2019-visualisingsoftware-architecture-with-the-c4-model.pdf . Acesso em: 24 de fevereiro de 2024.

Altaiar, H. Garverick, J., Babe, I., Toroman, M.,Lee J., Minasyan, V (2022).. **The Developer's Guide to Azure**. Microsoft E-Book series. Disponível em: https://clouddamcdnprodep.azureedge.net/gdc/gd cz5VglF/original?ocid=AID3042111 . Acesso em: 24 de fevereiro de 2024.

Gimenes, I. M. S., Huzita, E.H.M. **Desenvolvimento Baseado em Componentes: Conceitos e Técnicas**. Ciência Moderna, Rio de Janeiro, 2005.

Pinto, S. C. C. S. **Composição em Web Frameworks**. Tese (Doutorado). Departamento de Informática PUC-Rio, Rio de Janeiro, 2000.

Fayad, M.E., Johnson, R.E. **Domain-specific application frameworks: frameworks experience by industry**. John. Wiley, New York, 2000.

Fayad, M.E., Schimidt, D.C., Johson, R.E. **Building application frameworks: object-oriented foundations of framework desing**. John Wiley, New York, 1999a.

Fayad, M.E., Schimidt, D.C., Johson, R.E. **Implementing Application Frameworks: Object-Oriented Frameworks at Work**. John Wiley, New York, 1999b.

Szyperski, Clemens. **Component Software. Beyond Object-Oriented Programming**. Pearson, Londres, 2011.

Mattsson, M. **Evolution and Composition ObjectOriented Frameworks**. Tese (Pós-Doutorado). University of Karskrona/ Ronneby, Department of Software Engineering and Computer Science, Suécia, 2000.

Mattsson, M. **Object-Oriented Frameworks – A Survey of Methodological Issues**. Tese (Mestrado). Department of Computer Science, Lund University, Suécia, 1996.

Junior H.Q.P. **Economia da Energia – Fundamentos Econômicos, Evolução Histórica e Organização Industrial**. LTC, rio de Janeiro, 2007.