Global Journal of Economics and Business (GJEB)

Volume.14, Issue.4 pp: 387-408 Article 6
Aug (2024)

Exploring the Impact of Artificial Intelligence and Machine Learning on Predicting Outcomes in Regional Development Conventions: A Case Study of the Tanger-Tetouan-Al Hoceïma Region, Morocco

مساهمة الذكاء الاصطناعي والتعلم الآلي لتدبير اتفاقيات التنمية الجهوية: دراسة حالة جهة طنجة تطوان الحسيمة، المغرب

Asmaa Faris, Mostafa Elhachloufi

[1,2] University Hassan II of Casablanca, Morocco
[1] asmaa-faris-etu@etu.univh2c.ma, [2] elhachloufi@yahoo.fr

Accepted قبول البحث	Revised مراجعة البحث	Received استلام البحث
2024/5/8	2024/4/24	2024 /3/2

DOI: https://doi.org/10.31559/GJEB2024.14.4.6

REFAAD رفـــاد
www.refaad.com

Global Journal of Economics and Business (GJEB)

المجلة العالمية للاقتصاد والأعمال

Journal Homepage: https://www.refaad.com/Journal/Index/2

E-ISSN 2519-9293 | P-ISSN 2519-9285

Exploring the Impact of Artificial Intelligence and Machine Learning on Predicting Outcomes in Regional Development Conventions: A Case Study of the Tanger-Tetouan-Al Hoceïma Region, Morocco

مساهمة الذكاء الاصطناعي والتعلم الآلي لتدبير اتفاقيات التنمية الجهوية: دراسة حالة جهة طنجة تطوان الحسيمة، المغرب

Abstract:

Objectives: For over six years, Morocco has been engaged in an initiative aimed at promoting inclusive and sustainable territorial development, characterized by the implementation of an advanced regionalization experiment. However, the main challenge facing government leaders today is managing a growing number of regional development conventions.

Methods: To achieve these objectives, we developed a predictive model using artificial intelligence and machine learning aimed at anticipating the outcomes of regional development conventions. Through this method, we were able to identify the risks associated with their potential failure and forecast their impact. Simultaneously, various data extraction methods and classification algorithms were implemented. We collected and analyzed information from 310 past and current regional development conventions, spanning from 2016 to 2020, considering various factors influencing their success or failure. We conducted extensive experiments to assess the effectiveness of various predictive models.

Results: The results we obtained indicate that the classifiers Ada Boost, Random Forest, and Bernoulli NB are the three most effective algorithms for predicting the outcomes of regional development conventions.

Conclusion: This study contributes to shaping a broader discourse on the practical application of artificial intelligence in public policies and regional development, highlighting its potential to enhance resource allocation and alleviate the burden of repetitive administrative tasks for organizations with limited resources.

Keywords: *Artificial Intelligence; Machine Learning; Regional Development; Government challenges; Predictive modeling.*

الملخص:

الأهداف: يشارك المغرب منذ أكثر من ست سنوات في برنامج يهدف إلى تعزيز التنمية الجهوية الشاملة والمستدامة، والتي تتميز بتطبيق تجارب التقسيم الجهوي. لكن المسؤولين الحكوميين الحاليين يواجهون التحدي المتمثل في تدبير عدد متزايد من الاتفاقيات التنمية الجهوية.

المنهجية: لتحقيق هذه الأهداف، قمنا بتطوير نموذج تنبؤي باستخدام الذكاء الاصطناعي والتعلم الآلي للتنبؤ بنتائج الاتفاقيات التنمية الجهوية. أتاحت لنا هذه الطريقة تحديد المخاطر المرتبطة بالفشل والتنبؤ بعواقبه. وفي الوقت نفسه، تم تطبيق أساليب مختلفة لاستخراج البيانات وخوارزميات التصنيف. قمنا بجمع وتحليل بيانات من 310 الاتفاقية من عام 2016 إلى عام 2020، مع النظر في العوامل المختلفة التي تساهم في النجاح أو الفشل. أجرينا دراسة شاملة لتقييم أداء نماذج التنبؤ المختلفة.

النتائج: تظهر نتائجنا أن Ada Boost وRandom Forest وBernoulli NB هي الطرق الثلاثة الأكثر فعالية للتنبؤ بنتائج الاتفاقيات التنمية الجهوية.

الخلاصة: تساهم هذه الدراسة في مناقشة أوسع حول تطبيق استراتيجيات الذكاء الاصطناعي في السياسة العامة والتنمية الإقليمية، مما يدل على إمكانية تحسين تخصيص الموارد وتقليل عبء الأنشطة الإدارية المتكررة للمنظمات الغنية بالموارد.

الكلمات المفتاحية: *الذكاء الاصطناعي؛ التعلم الآلي؛ التنمية الجهوية؛ تحديات الحكومة؛ النمذجة التنبؤية.*

1 Introduction

This article focuses on the crucial role of Artificial Intelligence (AI) and Machine Learning (ML) in strengthening the anticipation and execution of regional development conventions, with a particular emphasis on the Tanger-Tetouan-Al Hoceïma (TTAH) region in Morocco. As emerging technologies continue to redefine the landscape of project management and planning, understanding their implications on regional development becomes imperative. The TTAH region, characterized by its diverse economic activities and unique geographical features, provides a captivating case study to explore the transformative potential of AI and ML.

Regional development conventions play a fundamental role in promoting economic growth, improving social well-being, and preserving environmental sustainability. However, their success relies on the ability to anticipate with precision, plan meticulously, and execute efficiently. The integration of AI and ML in this context opens new perspectives to optimize decision-making processes, mitigate risks, and enhance overall convention outcomes.

Having state-of-the-art innovation tools is essential for enabling public and political leaders when enhanced regionalization is implemented successfully. Artificial intelligence (AI) and machine learning (ML), which provide quick, accurate, and data-driven evaluations, have disrupted dynamic methods. With the help of these cutting-edge innovations, we can act even more quickly and accurately, which is especially important in the context of territorial development conventions (DR). There has been a steady rise in the quantity of DRs since 2016. It is alarming to see, nevertheless, that about 40% of these norms fall short and are never carried out, resulting in poor execution. These observations are concerning, particularly in light of the large out-of-pocket costs and lengthy processing times linked to these conventions' administrative procedures. These difficulties impede the region's overall growth and throw a shadow over regionalization's advancement. Through the use of ML and artificial intelligence to evaluate a DR's risks and determine its outcome, we can make highly informed decisions that maximize benefits and minimize anticipated drawbacks. Eventually, using computer-based intelligence and machine learning will allow DRs to operate even more profitably, efficiently, and productively.

In any case, the integration of AI and ML in regional development in provincial improvement isn't without challenges. Issues connected with information protection, algorithmic predispositions, and the requirement for straightforwardness in dynamic cycles arise as central issues, which will be investigated in this review.

In this article, an anticipation of the outcomes of development conventions in the Tanger-Tétouan-Al Hoceïma region will be elaborated following the methodology of the Cross Industry Standard Process for Data Mining (CRISP-DM) Figure 1. Advanced prediction approaches will allow for the precise classification of conventions, the identification of commonalities between them, and the application of learned lessons to improve the convention process as a whole. In actuality, the algorithm will use information from earlier conventions to recognize examples that are similar to new ones and adjust its predictions appropriately. We can improve decision-making, boost convention efficiency, and eventually increase success rates by utilizing these insights. Predictive models will enable us to leverage our historical experiences and create a more robust and knowledgeable base for conventions in the future.

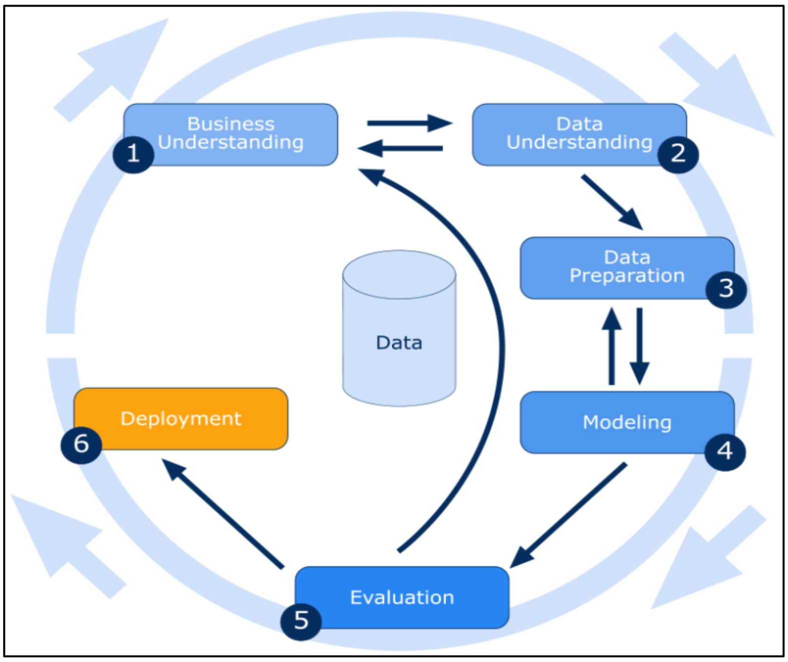

Figure 1: CRISP-DM Methodology

In this study, our primary objective is to conduct an in-depth analysis of how artificial intelligence calculations can be leveraged to identify patterns within the various development conventions initiated by the Local Board of Tanger-Tétouan-Al Hoceïma (TTAH). To achieve this, we have structured our paper as follows:

In the initial sections of our work, we engaged in a thorough theoretical exploration, focusing particularly on the theoretical and conceptual aspects related to the use of AI in the analysis of regional development data. This initial part aims to provide a solid conceptual framework for our subsequent analysis.

In the subsequent section, we developed our recommended strategy for applying AI calculations in identifying patterns within regional development conventions. This section of the paper highlights the specific methodological approaches we have adopted to achieve our research objective.

Following this, in the fourth part, we delve further into the construction of our recommended strategy. We detail the concrete steps we have taken to implement our AI methods in the analysis of development convention data.

The results of our analysis are presented comprehensively in the fifth section of the paper. We examine in detail the trends and patterns identified through our artificial intelligence methods, highlighting the key insights gleaned from our study.

Finally, in the sixth and final section, we conclude our paper by summarizing the main conclusions drawn from our study and discussing the implications of our research for understanding and managing regional development conventions. This conclusion also provides avenues for future research in this exciting field.

2 AI and ML Applications in Regional Development

The integration of artificial intelligence systems into public decision-making processes to optimize regional development is a relatively underexplored research area. Despite the crucial importance of regional improvement in national development strategies, previous studies have primarily focused on theoretical aspects or broader policy strategies, often overlooking the practical application of new technologies to enhance regional development outcomes.

For instance, (Domashova, 2020)examined the use of artificial intelligence techniques, particularly neural networks, to predict the outcomes of public procurement contracts in the pipeline industry. The author developed a Python application to identify contracts at high risk of non-execution using a classification approach for public contracts.

On the other hand, (Kripak, 2021)investigated the use of artificial intelligence methods to predict unfavorable outcomes in public procurement procedures related to building and land maintenance. Their study identified risks such as collusion among suppliers, collusion between clients and suppliers, and erroneous data in the unified information system. The authors employed associative rule extraction techniques and clustering methods to identify sustainable supplier groups and potential violations.

Furthermore, (J. Gallego, 2021) explored the use of artificial intelligence models to predict weaknesses and corruption in public acquisitions in Colombia. Additionally, (M. J. García Rodríguez, 2022) discussed the use of artificial intelligence algorithms to detect collusion in public procurement auctions. Their study compared the performance of eleven algorithms on six datasets from five countries and concluded that techniques such as extra trees, random forests, and AdaBoost are the most effective for detecting collusion, even with limited data.

These works highlight the growing importance of integrating artificial intelligence into public decision-making processes to enhance the transparency, efficiency, and effectiveness of regional development policies. They also offer valuable insights for professionals working in the field of public procurement oversight and pave the way for further research in this exciting area.

The study conducted by (T. Inan, 2022) presents an AI model based on Long Short-Term Memory (LSTM) to predict convention costs. This model utilizes a seven-layer feature vector, incorporating performance factors related to schedule and cost, along with their moving averages as indicators. The results of this study demonstrate that the LSTM model provides more accurate estimates compared to the traditional model based on the Earned Value Management (EVM) index. Furthermore, the study aims to demonstrate how decision-makers can utilize data from completed conventions using AI.

In a different approach, (U. Park, 2022)proposes a two-level heterogeneous ensemble algorithm to estimate the cost of construction conventions in the early stages. This method combines Random Forest (RF), Support Vector Machine (SVM), and CatBoosting algorithms and was evaluated using cost data disclosed by the public procurement administration in South Korea. The results indicate that the two-level ensemble model outperforms individual ensemble models, thus providing objective information for decision-making at the beginning of a construction project.

Lastly, a recent study conducted by (Jihad Satri, 2023) focuses on the BKM region and aims to determine the most effective algorithm for modeling the outcomes of development conventions. The researchers explore the fields of Artificial Intelligence and Machine Learning to identify the most suitable models for this task. This study highlights the growing importance of AI in modeling the outcomes of development conventions, thereby opening new avenues for research in this field.

Having examined the applications of AI and machine learning in regional development, it is now opportune to focus our attention on a thorough exploration of the Tanger-Tétouan-Al Hoceïma (TTAH) region, in order to comprehensively grasp its specificities and characteristics.

3 Exploring the Tanger-Tétouan-Al Hoceïma (TTAH) Region

Located in northern Morocco, the Tanger-Tétouan-Al Hoceïma (TTAH) region presents a splendid amalgamation of economic diversity, demographic richness, and singular challenges. This section provides a comprehensive analysis of the essential elements defining the region, thereby offering valuable insights into its distinctiveness.

3.1 Economic Sectors:

The Tanger-Tétouan-Al Hoceïma (TTAH) region exhibits a diversified economic structure, as evidenced by recent statistical data.

- **Manufacturing Sector**: The manufacturing sector in the region comprises over 2000 companies, representing approximately 15% of all local businesses. These enterprises provide employment opportunities for nearly 50,000 workers, thus contributing to about 20% of the regional GDP. The manufacturing industries in the region cover a diverse range of products, from textiles to pharmaceuticals to agri-food. They play a vital role in the economic diversification of the region and in the creation of skilled employment, thereby contributing to the region's industrial development.
- **Agriculture:** Agriculture in the TTAH region is also of paramount importance. The region cultivates a variety of agricultural products, including citrus fruits, olives, vegetables, and cereals. Data indicates that agriculture accounts for approximately 30% of total employment in the region, supporting over 100,000

families. In addition to providing livelihoods, agriculture in the region contributes to national food security and promotes rural development by stimulating local economies and preserving agricultural traditions.

- **Tourism**: Furthermore, tourism constitutes a major pillar of the regional economy. With over 2 million visitors each year, the region generates tourism revenues exceeding $500 million. This burgeoning tourism industry contributes to approximately 10% of the regional GDP, underscoring its growing importance in the economic development of the region. The region's tourist attractions, such as its beaches, mountains, historical sites, and rich culture, attract domestic and international travelers throughout the year, thereby stimulating economic growth and creating employment opportunities in the tourism sector and related industries.

Table 1: The Most Important Economic Indicators for The Tanger-Tetouan-Al Hoceima Region

Contribution to the Gross National Product (GNP)	9,4 %
Gross National Product (GNP) per capita	24,650 DH, ranked 6 nationally
The market value of marine capture	11% of the national value, 3% of the total volume
Contribution to National Industrial Production	10%
Contribution to National Industrial Exports	25%
The percentage of exports from traditional industries	0,8%
Industrial zones	1 820 ha
Workers in the industrial sector	80% affiliated with other regions
Tourism	30% employment rate
Air transport	Ibn Battuta ranks 6th

Source: HCP 2014

In essence, the TTAH region's economic vibrancy is a result of the synergies among its diverse sectors. Manufacturing, agriculture, and tourism collectively contribute to the region's multifaceted economic profile, ensuring a balanced and sustainable approach to development. Understanding the intricate dynamics of these sectors becomes essential for policymakers and stakeholders in crafting strategic initiatives that capitalize on the strengths and potential of each, fostering holistic economic development within the TTAH region.

3.2 Demographic Characteristics:

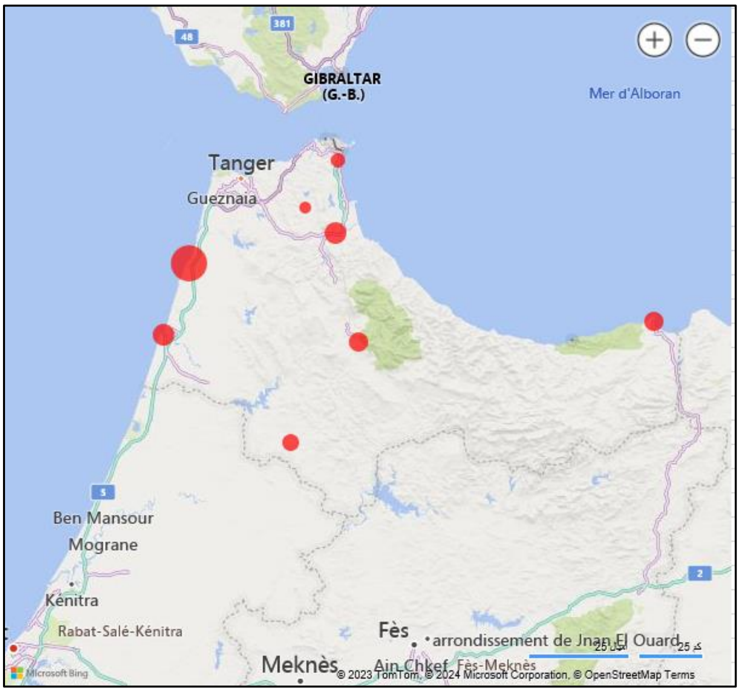

Figure 2: The Population Density of the Tanger-Tetouan-Al Hoceima Region
Source: Hcp 2014

Gaining insight into the demographic composition is indispensable for comprehending the social intricacies of the Tanger-Tétouan-Al Hoceïma (TTAH) region. The unique feature of this region lies in the diversity of its population, influenced by a rich blend of cultural factors. Key demographic elements, including age distribution, educational attainment, and the composition of the workforce, offer invaluable insights into the human dynamics shaping the region.

The age distribution within the population provides a nuanced understanding of the demographic pyramid, highlighting the proportions of various age groups. Educational levels reflect the region's intellectual capital, showcasing the knowledge and skills available within the community. Additionally, understanding the composition of the workforce sheds light on the region's economic activities and the skill sets contributing to its development.

These demographic factors collectively play a pivotal role in shaping the development needs and priorities of the TTAH community. Age demographics influence educational requirements, healthcare demands, and workforce development initiatives. Educational levels influence the region's capacity for innovation, research, and technological advancements. The composition of the workforce guides economic planning and development strategies, aligning them with the skills and expertise available locally.

In essence, a detailed comprehension of the demographic landscape enables policymakers, planners, and community leaders to tailor development initiatives that resonate with the specific characteristics and requirements of the TTAH region. It forms the foundation for creating targeted interventions that address the unique social fabric, fostering sustainable development and improving the overall quality of life for the residents.

Table 2: The most important social indicators for the Tanger-Tetouan-Al Hoceima region

Unemployment rate	11,2 %
Growth rate	1,42 %
Illiteracy rate	31,7 %
Poverty rate	10,08 %

Source: HCP 2014

3.3 Unique Challenges

The Tanger-Tétouan-Al Hoceïma (TTAH) region, abundant in potential, encounters distinct challenges that require meticulous attention in the developmental context. Geographic complexities, environmental considerations, and socio-economic disparities present critical challenges that call for thoughtful and strategic interventions. Delving into these challenges provides an in-depth understanding of the obstacles that regional development initiatives must confront to attain sustainable and inclusive growth.

Geographic intricacies, encompassing factors such as diverse landscapes and topographical variations, pose challenges in planning and executing development conventions. Environmental concerns, including issues like biodiversity conservation, sustainable resource management, and climate change impacts, necessitate environmentally conscious approaches. Moreover, socio-economic disparities, reflecting variations in income, education, and access to resources, demand targeted strategies to ensure equitable and inclusive development.

Exploring these challenges not only sheds light on the complexity of the regional development landscape but also emphasizes the need for tailored and strategic interventions. It underscores the importance of adopting holistic approaches that consider the interplay of geographical, environmental, and socio-economic factors. By navigating through these challenges with precision, regional development initiatives can pave the way for sustainable growth and foster inclusivity, addressing the unique needs and characteristics of the TTAH region.

3.4 SRAT and PDR

The establishment of the new regional councils signifies a pivotal moment in advancing the implementation of advanced regionalization. This involves an intricate process that requires time to materialize effectively. Implementing planning and programming processes, along with formalizing supporting tools like the Regional Spatial Planning Scheme (SRAT) and the Regional Development Program (PDR), should be grounded in innovative territorial planning and programming tools. These methods need to be capable of showcasing the region as a key player in development.

The SRAT serves as a benchmark document for realizing the vision of territorial development and outlining the region's directions over a 25-year span. It also lays out a comprehensive framework for sustainable and harmonious regional development in urban and rural areas, proposing regional and planning conventions.

This document aims to coordinate interventions by the State, local authorities, and private investors, guiding their strategic decisions regarding spatial planning and development. Achieving this involves extensive consultation during implementation, aligning with the guidelines of national spatial planning policy and integrating with strategies, programs, and sectoral plans at both local and national levels. The SRAT particularly strives to achieve consensus between the State and the Region on spatial planning measures and developments, guided by a strategic and forward-looking vision, facilitating the formulation of directions and decisions for territorial development.

Regarding the PDR, serving as the reference document for programming conventions and activities intended for implementation in the region, its objective is to facilitate coordinated and sustainable development. This includes enhancing the attractiveness of the regional environment and bolstering its economic competitiveness. Given its significance, the development and implementation of the PDR should adhere to a comprehensive and rigorous process, guided by legal texts governing this matter.

3.5 Conventions

The Tanger-Tétouan-Al Hoceïma (TTAH) provincial assembly convenes periodically for sessions to deliberate and make decisions, including the production of various large-scale programs focused on territorial development. These programs often result from conventions established through joint efforts by the chamber and partners in key sectors. The TTAH territory committee plays a crucial role in carefully considering these developments.

The committee should thoroughly review the execution of these arrangements because the outcome of these conventions is vital for the overall progress of local development. Winning these conventions is a critical step in implementing the Territorial Improvement Program (PDR), initiated by the TTAH local assembly. To ensure the program's effectiveness and alignment with long-term strategic goals, it should be continuously monitored, updated, and evaluated.

Therefore, the TTAH regional council must prioritize the effective implementation and realization of conventions. This focus is essential for achieving and sustaining the council's long-term strategic goals and ensuring the success of regional development initiatives.

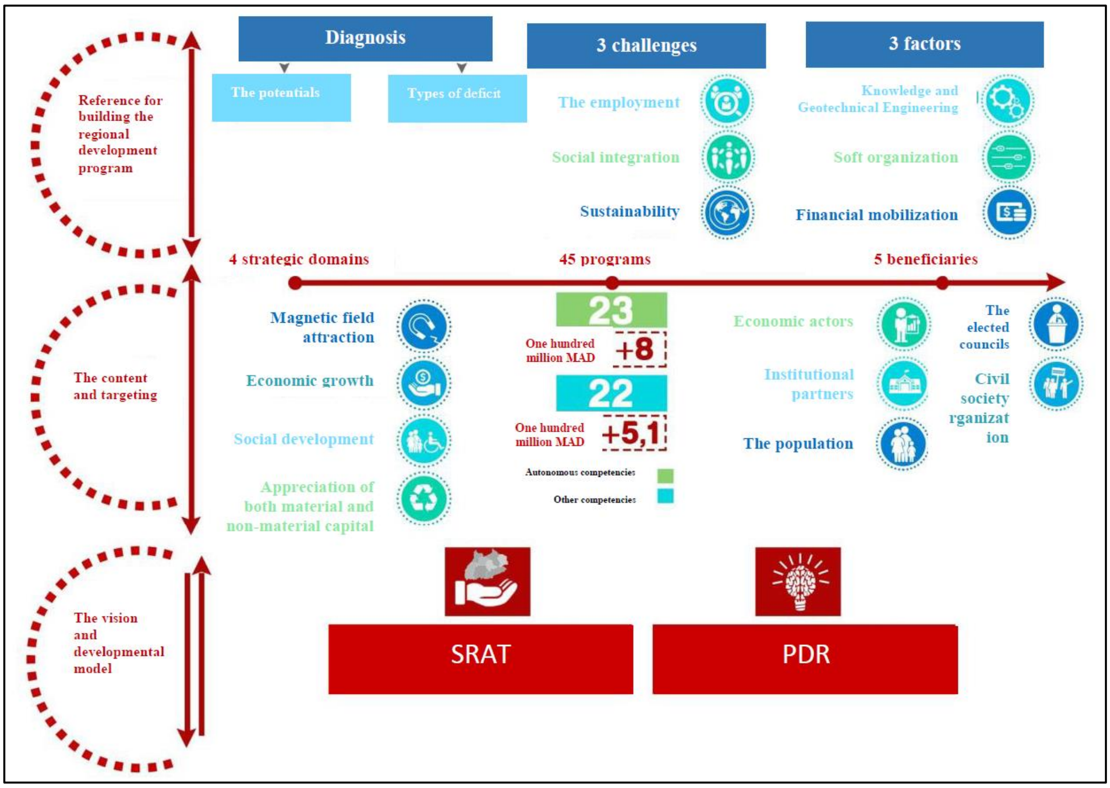

Figure 3: Structuring of The PDR And SRAT Source : Regional Council TTAH 2017 - 2022

3.7 Tailoring AI and ML to Regional Needs

Exploring the TTAH region forms the basis for tailoring applications of Artificial Intelligence (AI) and Machine Learning (ML) to meet its specific needs. By understanding the economic, demographic, and challenges, targeted solutions can be developed to enhance regional development conventions. This understanding ensures the seamless integration of AI and ML technologies into the regional context, optimizing their impact on the TTAH community.

Delving into the TTAH region provides stakeholders with a comprehensive perspective, allowing them to appreciate its diversity, leverage economic opportunities, address demographic nuances, and confront challenges effectively. This understanding serves as a cornerstone for implementing AI and ML applications, fostering sustainable development tailored to the region's unique characteristics.

Despite broad exploration, there is a shortage of AI (ML)-based prediction models in the ongoing literature on provincial development conventions (DCs). Our research focuses on examining public procurement contracts, a key component of our study. Our dataset includes information from public procurement contracts and other attributes related to committee decisions, conventions, and territorial DCs over a 4-year period. Insights from this study can contribute to broader research efforts aimed at developing a decision support tool to assess the risks associated with territorial DCs across all Moroccan provincial boards.

4 Methods

Dealing with huge proportions of data to help decision cycles is becoming extending thought in corporate IT frameworks (Krcmar, 2015; Laudon, 2010). Data Science is the discipline to get significant pieces of information from data by mathematical and logical models and applications. Data science exercises can profit from project the leaders and cycle frameworks. Suchsystems capability as progress factors (Saltz J. N., 2018; Cato, 2016). Regardless, following an endeavor situation thoroughly could be a test for data science gatherings. Process models like New DM (Wirth, 2000) and the Assessment Life Cycle (Long, 2015) should help and can be further developed by flexible procedures additionally (Grady, 2017) . In current data science projects, process models are not profoundly grounded, in spite of the way that cycle models exists for an onsiderable time span. As (Saltz J. N., 2018) shows, only 18% of data science bunches follow an express cycle model. Association in data science has similarly become great thought fairly as of late. Because of the own time of sending in New DM we should focus in on noticing how this stage is coordinated through the cycle models in research.There are a couple of assessments about adventure and cycle methods in data mining or data science projects. Nevertheless, another productive composing overview of the general course of New DM is at this point missing. The examinations of (Saltz J. S., 2016) also, (Cato, 2016) show the meaning of cycle strategies in tremendous data projects in general (Cato, 2016; Saltz J. S., 2016). (Saltz J. I., 2017) has pondered process procedures such New DM and swift methodologies by talking data science project bunches (Saltz J. I., 2017). We base in this paper on every time of New DM just and bar other cycle approaches explicitly since New DM is at this point a notable interaction eventually and in research. As such, we will sort out advantages and difficulties how New DM is driven in research studies. This paper adds to the current status of assessment by recognizing best practices and cycle stages, in which analysts can be more maintained.

In essence, the widespread adoption of the CRISP-DM is justified by its proven efficacy in delivering structured and high-quality solutions tailored to the specific needs of data mining, analytics, and data science endeavors. This emphasis on the early stages of the process serves as a guiding principle to optimize convention success and enhance the overall value derived from predictive data analytics initiatives.

In the context of this study, we endeavored to collect over 310 regional development conventions from the regional council. These conventions were launched in the Tanger-Tétouan-Al Hoceïma region between the years 2016 and 2020. This initiative reflects the active engagement of 146 local governments, each responsible for implementing a distinct set of development projects.

However, the sheer quantity of these conventions in a specific geographical area poses a significant challenge in terms of data collection. Given that Morocco is divided into 12 different regions, each with its unique administrative arrangement, this diversity adds an additional level of complexity to the task of data collection across these 12 regions.

With this in mind, we made the decision to specifically focus on examining the development conventions emanating from the Tanger-Tétouan-Al Hoceïma Regional Council (TTAH). This approach was adopted with the aim of effectively addressing our research objectives and thoroughly analyzing the development initiatives specific to this region.

4.1 Business and data understanding

Characterizing the degree and goals of an information science show is urgent for its prosperity. The most common way of understanding the issue includes a few basic advances. The show's goals, right off the bat, should be plainly characterized by recognizing what the business needs to accomplish and why. Understanding the business issue is a key stage in any information examination show as it decides the kind of experiences that a prescient investigation model can give to resolve the issue. This characterizes the investigation arrangement that the information researcher will expect to fabricate utilizing ML procedures. Characterizing the examination arrangement is the main errand in the business understanding period of the Fresh DM process. Subsequently, it is fundamental to allot adequate time and assets to precisely characterize the extension and goals of the show to guarantee its prosperity.

When the targets of an information science show have been characterized, the subsequent stage is to recognize the fitting examination approach, like grouping, relapse, bunching, or inconsistency recognition. It is additionally fundamental for outline the issue by deciding the best arrangement and reformulating the issue articulation as an investigation issue. This includes building a bank of inquiries and grasping the partners, as well as their assumptions and prerequisites. Thusly, we can guarantee that the investigation arrangement is lined up with the business targets and addresses the issues, everything being equal.

The Regional Council of TTAH holds conventions with its partners in order to form DCs, as was covered in section 3 (conventions). Two organizational operational processes were found following a thorough examination of the DCs and conversations with key players in the creation and execution process. Conventions are developed in the first process, and DCs are implemented through procurement contracts in the second process. A DC's ability to succeed depends on the results of conventions. Therefore, in order to perform predictive studies and precisely foresee the outcomes of the DC, it is required to combine the data from conventions and DCs. A classification model with a categorical target variable that indicates "FAIL" or "SUCCESS" is needed for this. Instead of resolving the main issues, the models depend on plans removed from obvious datasets and give experiences to assist the relationship with arriving at better conclusions about how to address business troubles.

In a broader context, to articulate the problem statement, It is possible to create a model that forecasts the chance of convention failure. Every new convention proposal might be given a failure probability using this model, enabling managers and decision-makers to identify and review those that have a higher chance of failing. By focusing on the convention proposals that are most likely to fail, this tailored approach to reviewing them could make the most of the limited time allotted for reviews. As a result, this would improve the detection of ineffective conventions and consequently shorten processing times.

After the problem has been fully understood, the next step is all about comprehending the data (Data Examination, Identifying Data Issues). In this phase, pertinent data is methodically gathered and examined in order to evaluate its completeness, quality, and organization. This procedure creates the framework for an extensive investigation, guaranteeing that the convention's latter stages are based on a strong basis of comprehended and verified data.

As-is State	**The Tanger-Tetouan-Al Hoceima Regional Council aims to assess the potential risks linked to the failure of a development project and proactively gauge their potential impact. The objective is not only to identify potential challenges and pitfalls but also to anticipate the consequences that may arise in the event of project failure. This forward-looking approach enables the regional council to develop strategies and contingency plans that can mitigate risks, enhance project resilience, and contribute to successful project outcomes. By thoroughly understanding and preparing for potential setbacks, the council can foster a more robust and adaptive approach to project development and ensure effective risk management throughout the project lifecycle.**
To Be State	**A machine learning-based model has been developed with the purpose of assessing the likelihood of project failure based on the information provided. This model utilizes advanced algorithms and statistical techniques to analyze the project data, identifying patterns and correlations that can indicate potential failure. By leveraging machine learning, the model has the capability to learn from historical project data, adapting and improving its accuracy over time.**
Succes	**Minimizing project failures is a multifaceted initiative that involves implementing various strategies and methodologies to enhance overall project success rates. This comprehensive approach aims to identify, mitigate, and prevent factors that contribute to project failure.**
Entity	**The regional council of Tanger Tetouan el Hociema**
Traget	**In the event of a project failure, it is necessary to activate an alert mechanism. This alert system is designed to promptly notify relevant stakeholders, project managers, and decision-makers about the critical status of the project. The purpose of implementing such an alert is toensure timely awareness of the failure, allowing for immediate intervention and corrective actions to be taken.**
Evaluation Matric	**accuracy (confusion matrix)**

Figure 4: Structure For Formulation (Produced By The Authors)

Here, we utilize a dataset of conventions and agreements for obtainment that was taken from an Excel data set to dissect convention recommendations exhaustively and figure out which ones are fruitful and which are not utilizing AI characterization calculations (see Table 3). To characterize information into preset classes or classes, characterization calculations a subset of AI strategies are utilized (Talingdan, 2019). These calculations gain from named preparing information and utilize that learned data to estimate the class or classification of newfound, inconspicuous information.

It is worth noting that there are multiple categorization algorithms, each with distinct advantages and disadvantages. The particular algorithm chosen will rely on the type of problem being solved as well as the properties of the data itself. Carefully weighing these variables guarantees that the selected method best fits the unique problems posed by the convention, enhancing the efficiency and precision of the analysis as a whole. The paper also highlights the significance of a careful algorithm selection procedure designed to handle the particular intricacies of the convention's goals and the features of the dataset being studied.

Table 3: Description of DC Dataset

Attribute	Type	Explanation
province	categorical	The province in which the DC is executed
specialization_region	boolean	Specialization of DC (self=0, shared=1)
sector	categorical	The sectors to which the DC belongs
cont_region	numeric	The financial contribution of the region in the DC
cont_stakeholders	numeric	The financial contribution of the stakeholders in the DC
Total_Fin_Env	numeric	The total amount of the DC
VISA	boolean	The signing of DC with the president of the regional council (yes=1, no=0)
delay_years_stakeholders	numeric	Number of years of commitment from the stakeholders in the DC
delay_years_region	numeric	Number of years of commitment from the region in the DC
Type_Territorial	categorical	Territorial type that will benefit from the DC
Progress_Percent	numeric	he percentage of the convention's progress
completion	**boolean**	**When the convention is fulfilled (no = 0, yes = 1), the target variable**

Source: produced by the authors

As of late, an information assortment was completed in which a name (target variable) was doled out as per previous outcomes. The conventions and acquirement contracts dataset will be used to look at this information from different points, figuring out the information, perceiving types, structures, and designs, and assessing the nature of the information. Moreover, a measurable investigation will be used to evaluate information quality according to a few points of view. An exploratory information examination (EDA) will be utilized to do this. EDA is a manufactured and scientific technique for working with datasets that looks to recognize and understand connections, examples, and patterns in the information (K. Sahoo, 2019). Acquiring information and cognizance of the information is the primary goal of exploratory information examination (EDA), which may then be utilized to illuminate more top to bottom investigation or dynamic systems.

Information changes, outline measurements, and representations are a couple of the numerous strategies that are remembered for EDA. Examples and linkages can be tracked down utilizing representations, and information appropriations and essential inclinations can be immediately studied with rundown measurements. Scaling and normalizing are two instances of information changes that can be utilized to further develop interpretability or prepare the information for more examination. By and large, is done first in quite a while examination process, before formal displaying or speculation testing. Prior to going on with the review, it helps in finding potential anomalies, missing qualities, or different issues that could should be fixed. Besides, EDA can help in figuring out speculations or coordinating the formation of progressively complicated models.

The dataset underwent processing through data mining tools, and any missing outliers were successfully addressed. Concerning missing values, several instances were identified and addressed in the course of the analysis. Initially, our dataset comprised 310 conventions. Following the data processing phase, we were left with a refined subset containing complete information for 170 conventions. It's noteworthy that during the data collection phase, limitations were encountered, preventing us from obtaining all the required information for every convention.

The accompanying Figure 5 visually demonstrates the effectiveness of our data cleaning efforts, highlighting the absence of missing values in the refined dataset.

```
In [358]: dataset.isnull().sum()

Out[358]: province                     0
          specialization_region        0
          sector                       0
          cont_region                  0
          cont_stakeholders            0
          Total_Fin_Env                0
          VISA                         0
          delay_years_stakeholders     0
          delay_years_region           0
          Type_Territorial             0
          Progress_Percent             0
          completion                   0
          dtype: int64
```

Figure 5: Detect And Fill The Missing Data (Python Output Produced By The Authors)

4.2 Data preparation

The checked on papers generally utilize standard information arrangement undertakings like information choice, change, and cleaning. Among these assignments, information change is oftentimes featured, given its intricacy inside the information readiness process. Following information change, information choice and it are addressed to clean errands. Eminently, creators of five investigations unequivocally notice the innovations used, including RapidMiner (Schnell, 2019), R (Poh, 2018), MS Access and WEKA (Kebede, 2017), and Microsoft SQL Server and RapidMiner (Putpuek, 2018). Besides, three examinations use different techniques for information arrangement, coming about in numerous datasets.

Except for one applied paper (Morais, 2017), the excess papers give nitty gritty clarifications of models customized to explicit use cases. Participating in preparing, testing, and examining various models and calculations demonstrates gainful for contrasting outcomes across various models during assessment. In addition, creators of nine examinations determine the advances utilized in this stage, for example, WEKA (Morais, 2017; Chiheb, 2017), Flash (Soliman, 2017), Orange Material, and RapidMiner (Oliveira, 2017).

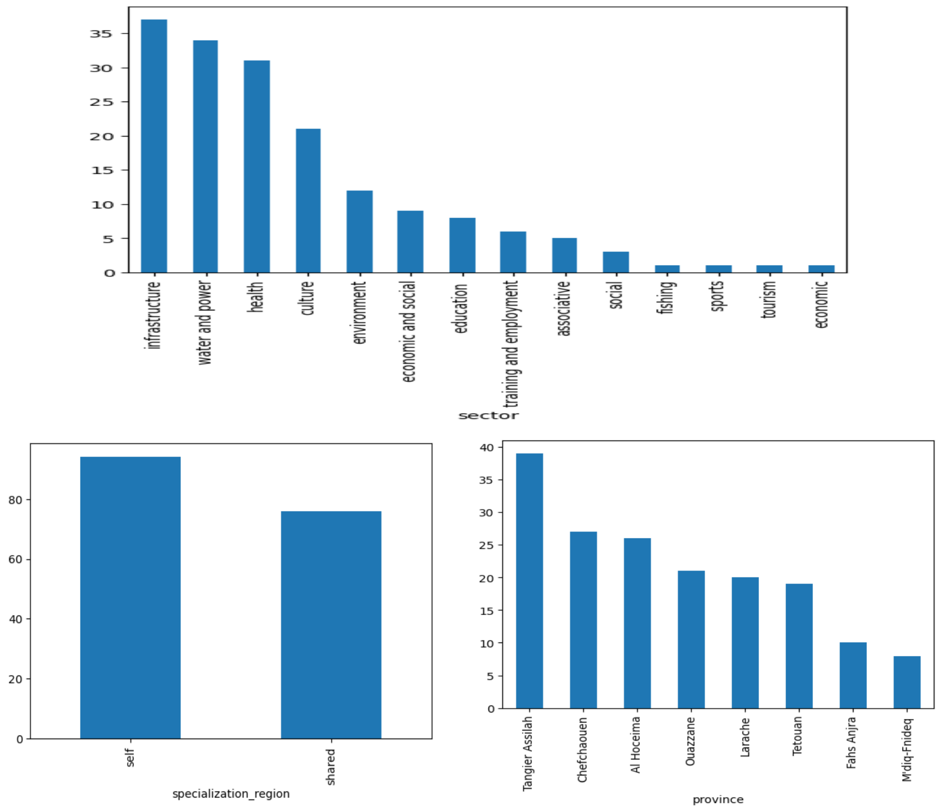

Figure 6: Number of Conventions by Province, Specialization of Region and Sector (Python Output Produced by the Authors)

Better grasp the distribution of development conventions across the various provinces, we can delve into the details of Figure 6. Our findings demonstrate that, in comparison to other provinces, the Tanger Assilah province has a noticeably greater number of DCs. This result is a good sign because it is in line with the goals of the Regional Development Plan (PDR). Notable benefits of Tanger Assilah's advantageous location as a gateway port between Europe and Africa are numerous. Consequently, in order to increase foreign trade in Morocco, it is essential to give it particular attention and bring it up to par with or better than other regions.

Moreover, it is important to remember that individual specialty agreements, not shared specialties, are the primary cause of most DCs. This pattern emphasizes the importance of concentrating on particular industries while adopting DCs throughout many provinces, most notably infrastructure, water and power, and healthcare. This observation emphasizes how crucial it is to support the energy and industry sectors in order to enable the effective implementation of DCs.

In contrast, a normal distribution is shown in the general distribution of the total convention costs per province. The fact that individual agreements come with a hefty price tag is especially noteworthy and emphasizes how important these agreements are to legislators.

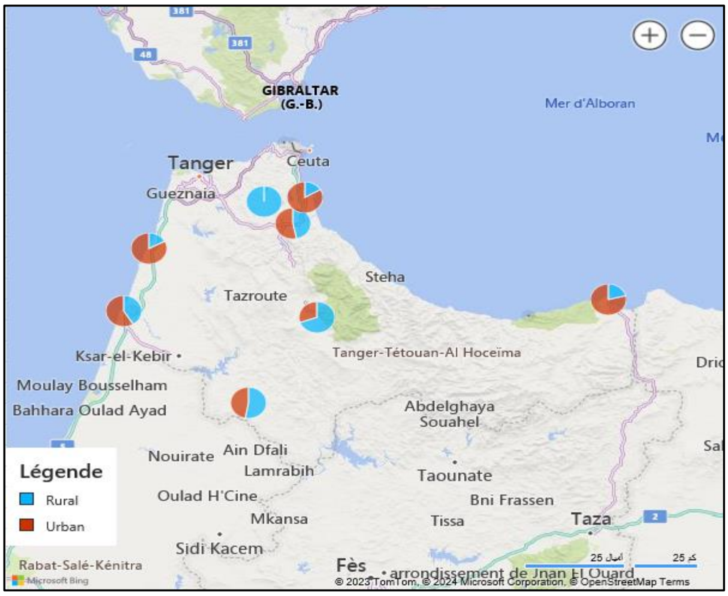

Figure 7: Maps of The Number of Conventions by Type Territorial (Produced by the Authors)

According to the map Figure 7, it is evident that investment in development conventions (DCs) is focused on rural areas in provinces that hold strategic positions and are recognized for tourism in Morocco, such as Chefchaouen and Fnideq.

This emphasis on rural development in strategically positioned provinces aligns with efforts to promote balanced regional growth and leverage the potential of areas with significant tourist attractions. Chefchaouen, renowned for its picturesque blue-washed buildings and natural beauty, and Fnideq, located near the Mediterranean coast and serving as a gateway to the Spanish enclave of Ceuta, are both prime examples of regions with considerable tourism potential.

Investing in rural development in these provinces not only enhances local infrastructure and living standards but also contributes to the overall tourism sector by improving accessibility, amenities, and attractions in these areas. This strategic focus underscores the government's commitment to promoting sustainable development and harnessing the economic potential of Morocco's diverse regions, particularly those with unique tourism assets.

4.2.1 Target variable

In this approach, information is sorted into groups or categories based on actual results obtained from information collecting, where names were assigned based on past performance. This tactic is frequently used with AI and vision display positions. Taking a look at Figure 8, where the goal variable is distributed evenly throughout the classes, may provide an idea of a fair dataset. This is often desired since the model may benefit from a comparable portrayal of each class. It is important to remember that the optimal balance may vary depending on the particular problem and industry.

In predictive modeling and machine learning, class imbalance can pose challenges. In situations where one class significantly outweighs the others, the model may become biased towards the majority class, leading to reduced accuracy and effectiveness in predicting the minority class. Techniques such as resampling, where the dataset is adjusted to create a more balanced distribution of classes, or using algorithms specifically designed to handle imbalanced data, can help address this issue.

Moreover, understanding the distribution of the target variable is crucial for selecting appropriate evaluation metrics and interpreting model performance. For example, accuracy may not be an appropriate metric for imbalanced datasets, as a high accuracy score can be misleading if the model is simply predicting the majority class. Instead, metrics such as precision, recall, and F1-score, which take into account the true positive, false positive, and false negative rates, provide a more comprehensive assessment of model performance in the context of class imbalance.

Overall, while balanced datasets are generally preferred for machine learning tasks, it is essential to consider the specific characteristics of the data and the problem at hand when determining the appropriate balance and selecting evaluation metrics. This ensures that the model effectively learns from the data and provides accurate predictions across all classes.

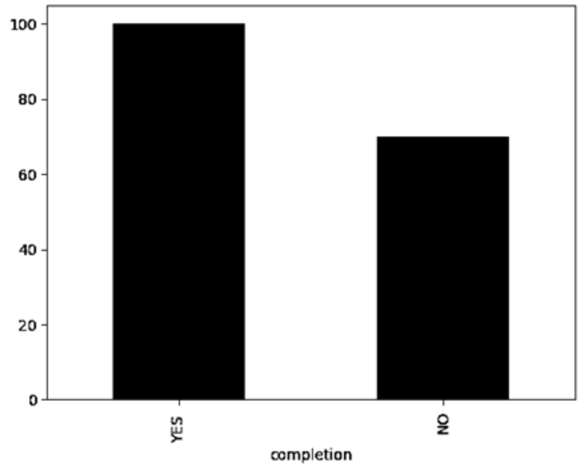

Figure 8: Target Variable Distribution (Python Output Produced by the Authors)

At this stage, our dummy model exhibits a distribution of 58.82% (Success) - 41.12% (Failure) Figure 8, making it a simple model that is used as a standard to assess how well more sophisticated models function. It functions as a vital point of comparison for evaluating the efficacy of other models. Usually, the dummy model makes predictions based just on random processes or basic principles, ignoring any relationships or patterns seen in the data.

By looking at measurements like exactness, accuracy, review, or the region under the beneficiary working trademark bend (ROC-AUC), we can decide whether the model essentially outflanks the spurious model. Assuming the model neglects to outperform the fake model, it demonstrates that more refined approaches or element designing are expected to catch huge examples in the information.

Accuracy measures the proportion of correct predictions made by the model, while precision quantifies the proportion of true positive predictions among all positive predictions. Recall, on the other hand, calculates the proportion of true positive predictions among all actual positive instances. The ROC-AUC metric evaluates the model's ability to distinguish between the positive and negative classes across various thresholds.

If the model achieves higher accuracy, precision, recall, or ROC-AUC values compared to the dummy model, it demonstrates improved performance in correctly predicting outcomes or classifying instances. However, if the model's performance metrics do not surpass those of the dummy model, it suggests that the model may not adequately capture the underlying patterns in the data. In such cases, further refinement of the model architecture, feature selection, or data preprocessing techniques may be necessary to enhance its predictive capabilities.

4.2.2 Data transformation

Information change is an essential interaction that modifies or converts unique information into an alternate portrayal, making it more reasonable for examination or demonstrating purposes. This includes using different procedures to upgrade information quality, remove important experiences, and meet the particular prerequisites of investigation or demonstrating errands. Information change incorporates a wide exhibit of procedures, including taking care of missing qualities, exceptions, and deviation, as well as performing scaling and standardization. The determination of procedures relies upon the information's qualities, the particular goals of investigation or displaying, and the algorithmic prerequisites. A definitive objective of information change is to upgrade information quality, interpretability, and execution for resulting investigations or demonstrating.

While managing all out factors in a dataset, a strategy known as "get_dummies" is frequently utilized as a component of the information change process. This procedure explicitly changes over all out factors into a mathematical organization by producing parallel pointer factors for every class. It actually integrates all out data into AI models, empowering them to utilize this sort of information. By utilizing the "get_dummies" capability, straight out factors are changed into a mathematical portrayal that can be effortlessly perceived and used by AI calculations. This works on the proficiency and precision of models in consolidating all out data during the examination or demonstrating process.

4.2.3 Feature selection

Highlight determination is an essential cycle in AI that includes picking a subset of important elements from a bigger arrangement of accessible highlights in a dataset. This assignment depends on utilizing stable earlier information and plans to streamline the exhibition of the AI model by zeroing in on the most useful highlights.

The component choice interaction principally involves eliminating identifiers, dates, information with high cardinality, and consistent qualities, as these may not contribute essentially to the prescient force of the model. After this underlying separating, relationship examination is utilized as a standard for highlight choice in AI.

During relationship examination, the connection between's each component (free factor) and the objective variable (subordinate variable) in the dataset is processed utilizing a connection coefficient Figure 9, like Pearson's relationship coefficient. This measurement evaluates the strength and heading of the direct connection between two factors. Highlights with high certain or negative connection values demonstrate a possibly critical relationship with the objective variable, recommending their significance for expectation undertakings.

Moreover, evaluating the relationship between's free factors themselves is fundamental. Exceptionally corresponded highlights might give repetitive or covering data, prompting multicollinearity issues. In such cases, it's fitting to choose just a single delegate highlight from each exceptionally related gathering to lessen overt repetitiveness and work on model interpretability and strength.

In the wake of assessing the relationship among's highlights and the objective variable, alongside taking into account connections among the actual elements, the last subset of highlights is chosen for consideration in the AI model. These chose elements ought to show a high relationship with the objective variable while limiting overt repetitiveness and multicollinearity, in this way upgrading the model's prescient presentation and interpretability.

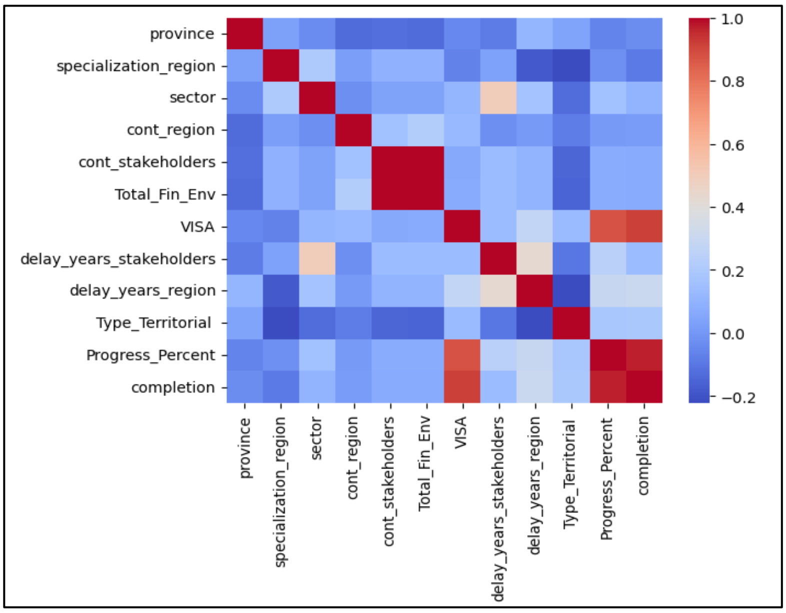

Figure 9: Correlation Heatmap (Python Output Produced by the Authors)

4.3 Modeling

The showing stage bases on making and planning judicious or expressive models considering the organized dataset. This dataset is isolated into planning and testing sets. The accompanying stage incorporates picking fitting models for planning. Taking into account that the issue incorporates game plan, sensible classifiers ought to be chosen to play out this endeavor. The portrayal computations used in this assessment include:

```
MLA = [
     #Ensemble Methods
    ensemble.AdaBoostClassifier(),
    ensemble.BaggingClassifier(),
    ensemble.ExtraTreesClassifier(),
    ensemble.GradientBoostingClassifier(),
    ensemble.RandomForestClassifier(),
    #Gaussian Processes
    gaussian_process.GaussianProcessClassifier(),
    #GLM
    linear_model.LogisticRegressionCV(),
    linear_model.PassiveAggressiveClassifier(),
    linear_model. RidgeClassifierCV(),
    linear_model.SGDClassifier(),
    linear_model.Perceptron(),
    #Navies Bayes
    naive_bayes.BernoulliNB(),
    naive_bayes.GaussianNB(),
    #Nearest Neighbor
    neighbors.KNeighborsClassifier(),
    #SVM
    svm.SVC(probability=True),
    svm.NuSVC(probability=True),
    svm.LinearSVC(),
    #Trees
    tree.DecisionTreeClassifier(),
    ]
```

Figure 10: The Classification Algorithms Used in this Study are Python Code Snippets

Increasing the information prior to setting our model is a vital stage in the information preprocessing stage. It guarantees that all parts are standardized to a comparative scale or reach, tending to the responsiveness of various artificial knowledge assessments to highlight scales. Expecting parts have differentiating scales, those with more prominent degrees have some control over the creating experience and excessively influence the model. By scaling the information, we spread out consistency among highlights, safeguarding against any single part overwhelming the learning assessment.

Then, the models are organized utilizing the coordinated dataset, with the objective variable and proper elements fittingly portrayed. The plan information is utilized to work on as far as possible and become familiar with the mystery models and relationship in the information. Exactly when the models are created, their show should be overviewed. This is finished utilizing evaluation assessments, for example, accuracy, precision, review, F1 score, subject to the sort of issue (course of action and break faith). The models are regularly outlined utilizing assessment philosophies like cross-underwriting or holdout support on a substitute underwriting dataset.

4.4 Evaluation

At this juncture of the project, you have developed one or more models that seem to exhibit high quality from a data analysis standpoint. Prior to advancing to the final deployment phase of the model, it is imperative to conduct a more comprehensive evaluation of the model and review the steps taken to construct it, ensuring that it effectively meets the business objectives. A primary goal is to ascertain whether any significant business issues have been overlooked. By the conclusion of this stage, a determination regarding the utilization of the data mining results should be made (Rüdiger Wirth, 2000).

4.5 Deployment

Typically, creating the model marks only a midpoint in the project. The acquired knowledge often necessitates organization and presentation in a manner accessible to the customer. Depending on specific requirements, the deployment phase may range from generating a report to implementing a systematic data mining procedure. Frequently, it is the user rather than the data analyst who oversees deployment tasks. Regardless, it is crucial to have a clear understanding from the outset of the necessary actions to effectively utilize the developed models. (Rüdiger Wirth, 2000)

5 Results and Discussion

Table 4 constitutes a comprehensive presentation of the results derived from the evaluation of the performance of 18 clustering algorithms applied in this study. These performances are assessed through several essential metrics, such as precision, accuracy, recall, and AUC (area under the ROC curve). These measures provide a detailed understanding of the predictive capability and reliability of each algorithm in data classification.

The predictive results of the classification model (DC) were generated using these 18 algorithms. This approach allows for a variety of perspectives on how data is classified and predicted by different models. This diversity of algorithms provides a valuable opportunity to explore and compare the performances of each model, highlighting their respective strengths and weaknesses.

A thorough evaluation was undertaken to analyze the performance of each algorithm in depth. The primary objective was to determine which model exhibits optimal performance for predicting the specific study data. This evaluation takes into account various aspects, such as the ability to accurately predict the classes of input data, as well as the ability to generalize these predictions to new datasets.

Ultimately, this rigorous analysis aims to identify the optimal predictive model that meets the needs and objectives of the study. This process of selecting the best model is essential to ensure reliable and accurate results in data-driven decision-making.

Exactness addresses the general accuracy of the model's expectations. AdaBoost Classifier displays the most noteworthy exactness (100 percent), while Gaussian NB has the least (44.19%). Accuracy estimates the precision of positive forecasts. AdaBoost Classifier, Random Forest Classifier, Bernoulli NB, Bagging Classifier, Decision Tree Classifier, Gradient Boosting Classifier, Extra Trees Classifier, and Linear SVC show high accuracy values, demonstrating a high level of accurately distinguished positive cases. Review centers around the model's capacity to recognize positive examples. AdaBoost Classifier, Random Forest Classifier, Bernoulli NB, Bagging Classifier, Decision Tree Classifier, Gradient Boosting Classifier, and Extra Trees Classifier all have the most noteworthy review (100 percent), demonstrating their ability to recognize all sure cases accurately. To additionally approve the presentation of the 18 calculations and decide the ideal classifier for the dataset, the AUC measure was used Figure 11. Taking into account the AUC metric, Ada Boost Classifier, Random Forest Classifier, Bernoulli NB, Bagging Classifier, Decision Tree Classifier, Gradient Boosting Classifier, and Extra Trees Classifier exhibit the best generally execution with an AUC score of 1, showing their viability in foreseeing the results of the DC.

Table 4: Classification Algorithms' Performance

	MLA Name	MLA Train Accuracy	MLA Test Accuracy	MLA Precission	MLA Recall	MLA AUC
0	AdaBoostClassifier	1.0000	1.0000	1.000000	1.000000	1.000000
4	RandomForestClassifier	1.0000	1.0000	1.000000	1.000000	1.000000
1	BaggingClassifier	1.0000	1.0000	1.000000	1.000000	1.000000
17	DecisionTreeClassifier	1.0000	1.0000	1.000000	1.000000	1.000000
3	GradientBoostingClassifier	1.0000	1.0000	1.000000	1.000000	1.000000
2	ExtraTreesClassifier	1.0000	1.0000	1.000000	1.000000	1.000000
11	BernoulliNB	0.9764	0.9767	0.958333	1.000000	0.975000
16	LinearSVC	0.9134	0.9070	0.851852	1.000000	0.900000
7	PassiveAggressiveClassifier	0.8425	0.7442	0.676471	1.000000	0.725000
5	GaussianProcessClassifier	1.0000	0.5814	1.000000	0.217391	0.608696
13	KNeighborsClassifier	0.7717	0.5814	0.592593	0.695652	0.572826
10	Perceptron	0.6693	0.5581	0.547619	1.000000	0.525000
6	LogisticRegressionCV	0.6063	0.5349	0.534884	1.000000	0.500000
14	SVC	0.6063	0.5349	0.534884	1.000000	0.500000
9	SGDClassifier	0.6142	0.5349	0.571429	0.521739	0.535870
15	NuSVC	0.5039	0.4884	0.545455	0.260870	0.505435
8	RidgeClassifierCV	0.4331	0.4651	0.500000	0.043478	0.496739
12	GaussianNB	0.4173	0.4419	0.000000	0.000000	0.475000

Source: Python output produced by the authors

AUC, accuracy, precision, recall, and Bagging are just a few of the measures that show the AdaBoost Classifier, Random Forest Classifier, Bernoulli NB, Bagging Classifier, Decision Tree Classifier, Gradient Boosting Classifier, Extra Trees Classifier, and Linear SVC algorithms to be highly performing. Both the Gaussian Process Classifier and the Passive Aggressive Classifier exhibit good precision and respectable accuracy. The classifiers with constant performance across all criteria are the Logistic Regression CV, Perceptron, SVC, and NuSVC. On the other hand, Gaussian NB performs less well than other classifiers. The top classifiers in this situation that are favored above the others are the AdaBoost Classifier, Random Forest Classifier, Bernoulli NB, Bagging Classifier, Decision Tree Classifier, Gradient Boosting Classifier, Extra Trees Classifier, and Linear SVC classifiers, according to the results. In contrast, the Random Forest classifier is demonstrated in the study by (Jihad Satri, 2023)to be an effective method for forecasting the results of regional development conventions.

The study's conclusions have important ramifications for the TTAH regional council since they offer insightful information on how to prioritize tasks, foresee the effects of convention outcomes, and reduce administrative costs. If decision-makers anticipate the potential failure of a development convention, they can take proactive measures early on by learning from the mistakes of past conventions. For instance, they can guarantee that the convention is included in the PDR and assess how well it fits the partner's skills. This entails looking over the convention's technical research, figuring out the land's characteristics and the convention's specified region, and taking the convention's legal standing into account. Furthermore, it is imperative to confirm that the convention's partners have the financial means to carry it out effectively. These discoveries present an opportunity increase convention effectiveness, streamline processes, and improve success rates over the long haul. The TTAH territorial committee can enhance convention results, smooth out tasks, and pursue all around informed key choices by using these ends.

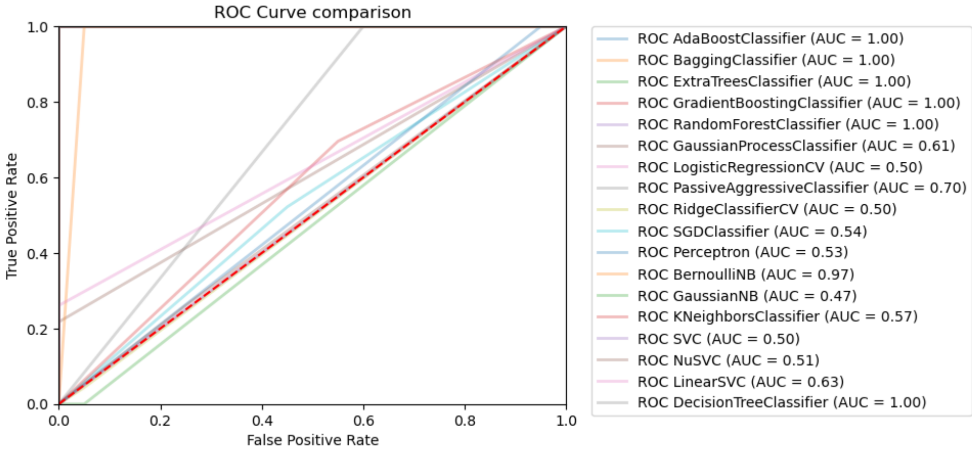

Figure 11: ROC Curve (Python Output Produced by the Authors)

6 Conclusion

The most common way of figuring out which algorithm is the most appropriate for a given information mining objective can be very troublesome. One proposed approach is to painstakingly contrast different calculations all together with recognize the one that creates the best exact results. This particular review included a near investigation that inspected seven different grouping calculations:

1) AdaBoost Classifier
2) Bagging Classifier
3) Extra Trees Classifier
4) Gradient Boosting Classifier
5) Random Forest Classifier
6) Gaussian Process Classifier
7) Logistic Regression CV
8) Passive Aggressive Classifier
9) Ridge Classifier CV
10) SGDC lassifier
11) Perceptron
12) Bernoulli NB
13) Gaussian NB
14) K Neighbors Classifier
15) SVC(probability=True)
16) NuSVC(probability=True)
17) Linear SVC
18) Decision Tree Classifier

This investigation was directed utilizing a dataset obtained from a Succeed information base, zeroing in explicitly on DC (success or failure) examination utilizing AI characterization calculations. To distinguish the most ideal prescient model, an intensive assessment of different characterization calculations was performed, considering pivotal execution measurements like exactness, accuracy, review, and AUC. The target of this assessment was to pinpoint the calculation reliably conveying prevalent outcomes across these presentation measurements. By examining the calculations in view of these measurements, an extensive comprehension of their viability in foreseeing results could be achieved, supporting the determination of the most trustworthy and proficient prescient model.

Upon a far-reaching examination of execution measurements in this review, the accompanying classifiers arose as exceptionally dependable and vigorous decisions for precisely foreseeing DC results: Ada Boost Classifier, Random Forest Classifier, Bernoulli NB, Bagging Classifier, Decision Tree Classifier, Gradient Boosting Classifier, Extra Trees Classifier, and Linear SVC.

This study enhances the continuous talk regarding the integration of artificial intelligence into public strategy and local improvement by presenting a commonsense point of view. It represents how artificial intelligence can be applied successfully to address true difficulties, presenting tangible solutions as opposed to staying bound to hypothetical conversations. Generally, the article fills in as a pragmatic exhibit of how artificial intelligence can possibly significantly upgrade public strategy and provincial turn of events. It exhibits the substantial advantages of artificial intelligence, for example, upgrading asset designation and mitigating regulatory weights. These benefits hold critical worth, especially for organizations operating under resource constraints.

As a suggestion for future research, there are opportunities for upgrading information assortment and recording processes, especially regarding data obtained from various local governments. This is vital in light of the fact that exact and complete information are basic for precise expectations of DC (success or failure). Also, investigating elective algorithmic models, like Neural Networks, Linear Regression, Fuzzy Logic, and Genetic Algorithm, among other high-level calculations, should be considered for further investigation. These elective methodologies might offer bits of knowledge and capacities past those given by the calculations analyzed in the ongoing review, potentially leading to enhanced predictive accuracy and effectiveness.

References:

Almasabha, G., et al. (2023). Structural performance of buried reinforced concrete pipelines under deep embankment soil. Construction Innovation. *Construction Innovation*. [CrossRef]

Alshboul, O., et al. (2022). Machine Learning-Based Model for Predicting the Shear Strength of Slender Reinforced Concrete Beams without Stirrups. *Buildings*, 12(8). [CrossRef]

Audu, A., et al. (2020). An Intelligent Predictive Analytics System for Transportation Analytics on Open Data Towards the Development of a Smart City. *Advances in Intelligent Systems and Computing*, 993, 224–236. [CrossRef]

Berryhill, J. et al. (2019). Hello, World: Artificial Intelligence and its Use in the Public Sector. [CrossRef]

C. Cath, S. W. (2018). Artificial intelligence and the 'good society': the US, EU, and UK approach. *SSRN Electronic Journal* 505-528. [CrossRef]

Cato, P. (2016). *Einflüsse auf den Implementierungserfolg von Big Data Systemen. Ergebnisse einer inhalts- und kausalanalytischen Untersuchung*. Dissertation. Friedrich-Alexander-Universität Erlangen-Nürnberg.

Chiheb, F. F. (2017). *Predicting students performance using decision trees: Case of an Algerian University*. International Conference on Mathematics and Information Technology (ICMIT), (pp. 113–121).

De Sousa W., et al. (2019). How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly*, 101-392. [CrossRef]

Domashova, A. O. (2020). Identification of public procurement contracts with a high risk of non-performance based on neural networks. *Procedia Computer Science*, 169, 795–799. [CrossRef]

Domashova, J. & Kripak, E. (2021). Aplication of machine learning methods for risk analysis of unfavorable outcome of government procurement procedure in building and grounds maintenance domain. *Procedia Computer Science*, 190, 171–177. [CrossRef]

El Haddadi, T., et al. (2021). Automatic analysis of the sustainability of public. *Cleaner and Responsible Consumption*, 3. [CrossRef]

El-Diraby, A. M. (2022). Application of Machine-Learning in Network-Level Road Maintenance Policy-Making: The Case of Iran. *Expert Systems with Applications*, 191. [CrossRef]

Erdelyi, O., et al. (2018). Regulating Artificial Intelligence: Proposal for a Global Solution. Social Science Research Network, Rochester. https://papers.ssrn.com/abstract=3263992

Gallego, J., Rivero, G. & Martínez, J. (2021). Preventing rather than punishing: An early warning model of malfeasance in public. *International Journal of Forecasting*, 37(1), 360–377. [CrossRef]

García Rodríguez, M., et al. (2022). Collusion detection in public procurement auctions with machine learning algorithms. *Automation in Construction*, 133. [CrossRef]

Grady, N. W., Payne, J. & Parker. H. (2017). Agile Big Data Analytics. AnalyticsOps for Data Science. *IEEE International Conference on Big Data (BIGDATA)* (17), (pp. 2331–2339). [CrossRef]

Guenduez, A., Mettler, T. & Schedler, K. (2020). Technological frames in public administration: what do public managers think of big data? *Government Information Quarterly*, 101-406. [CrossRef]

Inan, T., Narbaev, T. & Hazir, O. (2022). A Machine Learning Study to Enhance Project Cost Forecasting. *IFAC-PapersOnLine*, 55(10), 3286–3291. [CrossRef]

Ingrams, A., Kaufmann, K. & Jacobs, D. (2021). In AI we trust? Citizen perceptions of AI in government decision making. *Policy & Internet*, 1-20. [CrossRef]

J. Eager, M. W.-D. (2020). *Opportunities of Artificial Intelligence. Think Thank European Parliament*. https://www.europarl.europa.eu/thinktank/en/document/IPOL_STU(2020)652713

Kebede, M., Zegeye, D. & Zeleke, B. (2017). Predicting CD4 count changes among patients on antiretroviral treatment: Application of data mining techniques. *Computer Methods and Programs in Biomedicine*, 152, 149–157. [CrossRef]

Kolkman, D. (2020). The usefulness of algorithmic models in policy making. *Government Information Quarterly*, 37(3). [CrossRef]

König, P. & Wenzelburger, G. (2020). The legitimacy gap of algorithmic decision-making in the public sector: why it arises and how to address it. *Technology in Society*, [CrossRef]

Krcmar, H. (2015). *Informations management*. Wiesbaden: Springer Gabler. 6th ed.

Laudon, K. C. (2010). *Wirtschafts informatik. Eine Einführung. 2nd ed.* München: Pearson Deutschland. (IT).

Long, C. A. (2015). *Data science & big data analytics. Discovering, analyzing, visualizing and presenting data. Indianapolis*.

M. Janssen, G. K. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, 371-377. [CrossRef]

Martinez-Plumed, F., et al. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 33(8), 3048–3061. [CrossRef]

Mikhaylov, S., Esteve, M. & Campion, A. (2018). Artificial intelligence for the public sector: Opportunities and challenges of cross-sector collaboration. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2128). [CrossRef]

Morais, A., et al. (2017). Predicting the need of Neonatal Resuscitation using Data Mining. *Procedia Computer Science*, 113, 571–576. [CrossRef]

Oliveira, D. F. (2017). *Towards an Intelligent Systems to Predict Nosocomial Infections in Intensive Care*. 5th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), (pp. 150–155).

Park, U., et al. (2022). A Stacking Heterogeneous Ensemble Learning Method for the Prediction of Building. *Applied Sciences (Switzerland)*, 12(19). [CrossRef]

Poh, C., Ubeynarayana, C. & Goh, Y. (2018). Safety leading indicators for construction sites: A machine learning approach. *Automation in Construction*, 93, 375–386. [CrossRef]

Pourhomayoun, M. & Shakibi, M. (2020). Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart Health*, 20. [CrossRef]

Putpuek, N. (2018). *Comparative Study of Prediction Models for Final GPA Score: A Case Study of Rajabhat Rajanagarindra University*. 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), 92–97.

Rawat, S., et al. (2021). Application of machine learning and data visualization techniques for decision support in the insurance sector. *International Journal of Information Management Data Insights*, 1(2). [CrossRef]

Reis, J., et al. (2019). *Artificial intelligence in government services: a systematic literature review*. (H. A. Á. Rocha, Ed.) Springer International Publishing. [CrossRef]

Rüdiger Wirth, J. (2000). *CRISP-DM: Towards a Standard Process Model for Data*. https://www.cs.unibo.it/~danilo.montesi/CBD/Beatriz/10.1.1.198.5133.pdf

Sahoo, K., et al. (2019). Exploratory data analysis using python. *International Journal of Innovative Technology and Exploring Engineering*, 8(12), 4727–4735. [CrossRef]

Saltz, J. & Shamshurin, I. (2016). *Big data team process methodologies: A literature review and the identification of key factors for a project's success*. IEEE International Conference on Big Data (Big Data), (pp. 2872–2879). [CrossRef]

Saltz, J. (2017). *Comparing Data Science Project Management Methodologies via a Controlled Experiment*. Hawaii International Conference on System Sciences 2017 (HICSS-50). Hilton Waikoloa Village, Hawai.

Saltz, J. (2018). *Exploring Project Management Methodologies Used Within Data Science Teams Orleans*. 24th Americas Conference on Information Systems. New Orleans, LA, USA: Association for Information Systems.

Satri, J., El Mokhi, C. & Hachimi, H. (2023). Predicting the outcome of regional development projects using machine learning. *IAES International Journal of Artificial Intelligence*, 13(1), 863~875. [CrossRef]

Schnell, J., et al. (2019). Data mining in lithium-ion battery cell production. *Journal of Power Sources*, 413, 360–366. [CrossRef]

Shehadeh, A., et al. (2021). Machine learning models for predicting the residual value of heavy construction equipment: An evaluation of modified decision tree, Light GBM, and XG Boost regression. *Automation in Construction*, 129. [CrossRef]

Soliman, G., et al. (2017). *Predicting all star player in the national basketball association using random forest*. Intelligent Systems Conference (IntelliSys), (pp. 706–713). [CrossRef]

Soylu, A., et al. (2022). *Data Quality Barriers for Transparency in Public Procurement*. *Information (Switzerland)*, 13(2). [CrossRef]

Talingdan, J. (2019). *Performance comparison of different classification algorithms for household poverty classification*. 4th International Conference on Information Systems Engineering (pp. 11–15). ICISE. [CrossRef]

Van Calster, B., et al. (2019). Predictive analytics in health care: how can we know it works? *Journal of the American Medical Informatics Association*, 26(12), 1651–1654. [CrossRef]

Wilson, C. & Velden, M. (2022). Sustainable AI: An integrated model to guide public sector decision-making. *Technology in Society*, (68). [CrossRef]

Wirth, R. (2000). *CRISP-DM: Towards a Standard Process Model for Data Mining*. Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (4), (pp. 29–39).

Wirtz, B., Weyerer, J., Sturm, B. (2020). The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration. *International Journal of Public Administration*, 43(9), 818–829. [CrossRef]

Wirtz, B.W., Weyerer, J. & Sturm, B. (2020). The dark sides of artificial intelligence: an integrated AI governance framework for public administration. *International Journal of Public Administration* 818-829. [CrossRef]

Yeung, K., et al. (2019). *Algorithmic regulation*. (M. L. K. Yeung, Ed.) 1-18. [CrossRef]

Zekić-Sušac, M., Mitrović, S. & Has, A. (2021). Machine learning based system for managing energy efficiency of public sector as an approach towards smart cities. *International Journal of Information Management*, 58. [CrossRef]

Zuiderwijk, A., Chen, Y. & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, 38(3).