NOVA IMS Information Management School

MGI

Master Degree Program in
Information Management

ARTIFICIAL INTELLIGENCE FOR E-GOVERNMENT

A View on Children' Welfare

Tânia Filipa Medeiros Caria

Dissertation

presented as partial requirement for obtaining the Master Degree Program in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

ARTIFICIAL INTELLIGENGE FOR E-GOVERNMENT

A View on Children' Welfare

by

Tânia Filipa Medeiros Caria

Master Thesis presented as partial requirement for obtaining the Master's degree in Information Management, with a specialization in Information Systems and Technologies Management

Supervised by:

Prof. Doutor Vítor Manuel Pereira Duarte dos Santos

July, 2023

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor, Prof. Vítor Santos, for his invaluable guidance and support throughout my dissertation journey. Prof. Vítor always demonstrated great encouragement, availability, and expertise, which was crucial to complete and write this thesis.

Furthermore, I am grateful to Dora Alvarez and Rúben Cardoso from the National Commission and Promotion of Rights for Children (CNPDPCJ) for providing me with the resources, fundamental insight, and willingness to always support.

Finally, I would like to thank my dear friends and family for all the love and unconditional support. Without them, this journey would not have been possible.

ABSTRACT

This thesis aims to perform a holistic investigation concerning the use of artificial intelligence advance techniques in predicting children in danger of abuse or neglect. A dataset containing over 55322 cases from the Portuguese National Commission for the Promotion of the Rights and Protection of Children and Youth (CNPCJ) was collected and trained to uncover the patterns and additional findings. This research uses machine learning classification models to unveil the model with highest accuracy and robustness in the use predictive analytics in the children's' welfare field. This approach will allow social security services to understand the impact of underlying factors for further improvement of their services. Finally, this study develops a random forest predictive model to forecast children in risk, with an accuracy of 84.5%.

KEYWORDS

Artificial Intelligence; *Machine Learning*; *e-Government*; *Child maltreatment; Predictive Analytics*

Sustainable Development Goals (SGD):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. CONTEXT

In today's society, we are not only facing economic and social crises but also an alarming epidemic of violence against children, which thrives on silence and fear (Waterston, 2007). Shockingly, according to UNICEF, a child dies every five minutes due to violence, leaving millions of children unsafe in their homes, schools, and communities (fusion, n.d.). The profound impact of this crisis on the well-being of European children and the future of our society cannot be understated (Hillis et al., 2017).

Efforts to protect children are complex, primarily because assessing a child's risk of maltreatment is challenging, and no single factor can accurately predict it (Vaithianathan et al., 2013). This paper aims to provide an overview of the Portuguese child and family welfare system, acknowledging the significant role that data plays in driving innovation in Artificial Intelligence (AI). The latest advancements in AI have the potential to transform e-governments by enhancing productivity, improving accuracy, and facilitating localized planning (Ahn & Chen, 2022).

Artificial Intelligence, as a branch of computer science, seeks to develop computer systems and programs that can tackle complex problems by emulating human-like reasoning processes and creating intelligent machines and software. Neural Networks, a well-known area within Artificial Intelligence, have demonstrated real-world success in language processing and speech recognition (Ren et al., 2017). With its capabilities in reasoning, planning, learning, communication, perception, as well as manipulating and interacting with objects, Artificial Intelligence has become indispensable in various industries (Singh, 2017).

The objective of Artificial Intelligence is to create programs that replicate cognitive skills, such as learning and understanding rules, to effectively utilize data for organizational purposes. The algorithms developed, commonly referred to as rules, contain specific instructions for each step to complete tasks and achieve results with the highest possible accuracy (Haton, 2006).

By exploring the intersection of child welfare and Artificial Intelligence, this paper aims to harness the potential of AI to address the multifaceted challenges in child protection. The utilization of AI-driven algorithms and technologies can offer valuable insights, improve decision-making processes, and contribute to more effective and proactive interventions to safeguard the well-being of children at risk.

1.2. MOTIVATION

In recent years, there has been a significant increase in investment in Artificial Intelligence (AI), driven by emerging technologies, Big Data, and the Internet of Things (IoT). These advancements have led to the creation and growth of new applications and services, making AI a driving force for future economic growth across various fields.

The rise in the number of children reporting maltreatment to public authorities has brought forth pressing questions regarding the detection and response to child abuse. To address this humanitarian crisis, one potential solution is the utilization of machine learning methods to propose a comprehensive, multidisciplinary national framework aimed at preventing and effectively responding to all forms of violence against children.

This interrogation, lead to the formulation of the following Research Question:
RQ - "How can Data Science Models be used on service of a more trusted Social Welfare system?"

By analyzing the philosophy, structure, and operation of the gathered data, this research seeks to introduce an expanded perspective on policy options designed to protect children and establish a reliable forecasting procedure. Through the examination of patterns and historical data, this study aims to provide social security services with timely and informed decision support regarding potential cases of child abuse.

The contributions and findings of this thesis are intended to enhance the understanding of child protection and serve as a valuable resource for social security services. By leveraging AI techniques and historical data, this research offers the potential for improved child welfare outcomes and the ability to predict and prevent instances of child abuse.

Ultimately, this research endeavors to establish a stronger foundation for protecting children and promote community-driven action and welfare support. By combining AI capabilities, data analysis, and informed decision-making, this work strives to create a safer environment for children and empower communities to take proactive measures against child abuse.

1.3. RESEARCH GAP AND OBJECTIVES

The goal of the paper is to build a predictive risk model for a more trusted Social Welfare system and ensure that national governance can take the highest advantages from the use of AI technologies to protect children from violence, exploitation, and abuse. Additionally, it will take into account a collaborative approach to ensure that both social workers and the investigation performed are clear about what the data represents.

To achieve this goal, the following intermediate objectives were defined:

- Make a comprehensive study on the issue of children in danger and its mitigations strategies.
- Analyze the Portuguese children in dangerous' protection systems.
- Apply machine learning analytical tools that determine patterns of any kind of danger to a child based on historical data and known cases of violence in the environment.
- Employ these tools to support prevention and response to protection-related risks.

1.4. STUDY IMPORTANCE AND RELEVANCE

By incorporating best practices in machine learning, the government and Social Security can enhance their systems with increased flexibility, automation, and ultimately, foster broad public trust and support. This project goes beyond these objectives and aims to strengthen AI governance. Through exploratory analysis, new features will be derived, enabling scientists to develop a more effective public sector and establish a robust safety environment for children.

As a result, this research provides a comprehensive overview of the techniques used to analyze social security data, specifically concerning children. It focuses on anomaly detection within social security management, systematically analyzing the gathered information to establish a scientific and rational social safety management mechanism. This endeavor holds not only theoretical significance but also practical value in improving police efficiency and enhancing social governance capabilities.

Furthermore, this study enables the identification of socio-demographic variables that are associated with varying levels of child abuse potential among parents. By understanding child neglect and abuse patterns, future parents can take necessary precautions to safeguard their children and respond appropriately when faced with such circumstances.

Additionally, there are several avenues for further investigation in this research. Firstly, smart policies and educational programs can be developed and implemented to support the prevention of child abuse. Secondly, technology-based crime prevention strategies can be explored in other areas related to social security. Lastly, with increased awareness and legal analysis of child abuse, effective strategies for strengthening child protection can be identified, focusing on placing children, families, and communities at the forefront.

The ultimate aim of this paper is to contribute to a more trustworthy Social Security Administration system by proactively leveraging AI technologies to protect children from violence, exploitation, and abuse. By analyzing the specific information gathered, a scientific and rational social safety management mechanism will be established, delivering not only theoretical significance but also practical value in enhancing police efficiency and strengthening social governance capabilities (Sansone & Zhu, 2021). Moreover, this study will shed light on socio-demographic variables associated with varying levels of child abuse potential among parents, equipping future parents with knowledge of child neglect and abuse patterns to better protect their children.

2. LITERATURE REVIEW

2.1. THE CHILDREN IN DANGER ISSUE

Childhood is a crucial time of human development and children are particularly vulnerable to harm during this period.

The assessment of the child's needs, taking into account the child's developmental stage, as well as the risk and risk factors and protective factors present, in a holistic and systemic perspective, family support and positive parenting family and positive parenting are basic aspects to be taken into account in the promotion of the rights and protection of children and youth.

Contrary to the popular belief that children were thought of as being tangential and disconnected to the violence, and commonly labelled “silent witnesses”, recent qualitative research findings clarify that children make sense of their experiences, navigate through the complexity and terror intrinsic to domestic violence, and may be significantly affected by the experience of violence in their lives (Holt et al., 2008).

2.1.1. Background

Accurately predicting which children are at the highest risk of maltreatment is a challenge that the recent developments in digital technology have facilitated the recording and retrieval of administrative data from multiple sources about children and their families.

Combined with new ways to mine such data using algorithms that can ‘learn’, it has been claimed that it is possible to develop tools that can predict which individual children within a population are most likely to be maltreated. The proposed benefit is that interventions can then be targeted at the most vulnerable children and their families to prevent maltreatment from occurring. As expertise in predictive modelling increases, the approach may also be applied in other areas of social work to predict and prevent adverse outcomes for vulnerable service users.

Nevertheless, the application of PRM raises a number of moral and ethical concerns, and it is recommended a full ethical evaluation before its implementation (Vaithianathan et al., 2013). Thus, the first ethical lens to apply is the duty to respect individual autonomy and then evaluate the consequences of establishing intervention (Keddell., 2014).

The use of predictive modelling as a tool is used to lead to the identification of people within a child protection orientation, rather than as one aspect of a system predicated on the assumptions of the child welfare orientation. Therefore, it is important to recognize that the majority of families who become abusive are ordinary families who require a non-stigmatizing service to avoid the decline in abuse.

Furthermore, it is possible to find more challenges in using administrative data to develop an algorithm such as not seeking clarification from child protection agencies about how data about the substantiation of child maltreatment has been produced, which can result in unreliable statistics and misleading decision-making conclusions (Bromfield and Higgins., 2004).

On the other hand, PRM can help detect fraud and threats, mitigate the likelihood of risk, optimize policy campaigns, improve operational efficiency, and be the decision support system (Kumar & L., 2018).

Therefore, the key to developing predictive models is selecting reliable and valid outcome variables and ensuring that these are recorded consistently within the information systems (Gillingham, 2016).

2.1.2. Mitigation Strategies

In the past, child maltreatment prevention and intervention strategies focused on eliminating risk factors—conditions, events, or circumstances that increase a family's chances for poor outcomes, including child abuse and neglect. This emphasis on family risks (e.g., maternal depression, family violence, history of maltreatment) often left families feeling stigmatized or unfairly judged. In addition, focusing on risk factors does not always point the way toward solutions (Protective Factors Approaches in Child Welfare, n.d.).

On the other hand, the protective factors approach to the prevention of child maltreatment focuses on positive ways to engage families by emphasizing their strengths and what parents and caregivers are doing well, as well as identifying areas where families have room to grow with support. A protective factors approach can help agencies build capacity and collaborative partnerships with other service providers, such as early-childhood and youth-service systems, that are likely to enhance cross-systems collaboration to support children and families and promote their well-being. The protective factors approach also helps children, youth, and families build resilience and develop skills, characteristics, knowledge, and relationships that offset risk exposure and contribute to both short- and long-term positive outcomes.

Furthermore, studies of Head Start, a compensatory early education program for low-income children since 1965, and other childcare programs suggest that childcare services can help reduce maltreatment. The study speculated that having a child attend center-based childcare may reduce parents' use of physical discipline by relieving parent stress, exposing parents to alternative forms of discipline, and making the children more visible to potential reporters who would be aware of any maltreatment (Waldfogel, 2009).

A successful risk management strategy includes several crucial steps that should be taken in consultation with a holistic approach involving all stakeholders and, especially, children, and reviewed periodically. In this way, the Council of Europe and the rights of the child came up with a Strategy of a six-year implementation (2022-2027) – Table 1 - on the basis of an Action Plan prepared with other Council of Europe bodies and international partner organizations with a gender-sensitive, anti-discrimination, and child participation approach (Strategy for the Rights of the Child (2022-2027) - Children's Rights - Publi.Coe.Int, n.d.).

Table 1 – Mitigation Strategies [Council of Europe, 2022]

Challenges to be addressed urgently	– Prevent emotional or psychological violence, gender-based violence and neglect. – Create opportunities for children to speak out, including at the legislative level (e.g., in Parliaments) and through complaints procedures, and treat the voices of children and adults equally.
Possible action to be taken	– Create child-friendly care proceedings that are easier for children to understand, allow them to form and express their opinions and participate in the proceedings, without being fully dependent on adults. -Develop a model for prevention strategies at the national level and by addressing hate speech (including sexist hate speech) and the risk of children falling victim to violent radicalization. -Devise universal key definitions of violence, assessment tools with common indicators and provide guidance for a harmonized process of disaggregated data collection to obtain regular, specific, and reliable information. –Make child-friendly reporting and complaint mechanisms available and accessible for children at a low threshold, thus preventing (further) violence before it happens. – Add psychological check-ups to regular medical check-ups to assess the mental health

	of children and be able to identify and respond to any concerns. – Establish a “European day of the child’s voice” to raise awareness of the importance of each child’s voice. – Conduct workshops and educational programs in schools to respond to and prevent peer violence.
How children can be involved	– Train students as mediators or “peace agents” to make it easier for child victims of violence to ask for help, since confiding in peers is often easier. – Engage children in the design and evaluation of services for children, parents, and families and in the training of service providers to ensure services are meaningful for children and delivered in a child-centred way. – Create opportunities and structures for children to continue advising the Council of Europe on the implementation of the Strategy in the area of violence (e.g., by investing in campaigns, the creation of visibility material, the development of guidance with the support of experts).

2.1.3. Portuguese children in dangerous – Protection Systems

With the Childhood Protection Law in 1911, Portugal became the European pioneer in the creation of a legal framework for the protection of children. It remained in force until 1962, with the publication of the Organization for Guardianship of Minors (OTM), which reinforced the protective nature of the law for minors. It would be revised in 1978, remaining in force until 2001, when Law 147/99 came into effect, on September 1 - Law for the Protection of Children and Youngsters in Danger (LPCJP).

This new legal framework, which is currently the guiding instrument for all intervention in the area of protection of children and youth and the promotion of their rights, has abandoned the paternalistic view of the child, replaced the term "minors" with "children and youth", recognizing them as social actors in and subjects of rights, giving them the right to be heard, starting at 12 years of age or younger, as long as they have the right to be heard and the capacity to understand the meaning of the intervention.

The substitution of the term "risk" for the term "danger" in the definition of the scope of the legal diploma, which regulates intervention with children and young people, is also of note, insofar as it presupposes that the child is already is already facing one or several circumstances that make him/her vulnerable (typified in art. 3, paragraph 2) and not yet subject to a mere eventuality.

Noting the advantages of community intervention in the protection of children and youth at risk, the LPCJP consolidate the path begun with the then Commissions for the Protection of Minors, regulating the creation, competencies, and functioning of the Commissions for the Protection of Children and Young People (CPCJ), integrating attributions that are no longer only protection, but also aiming to act in the prevention of danger situations.

The CPCJ are, therefore, responsible for the promotion of rights and the prevention of dangerous situations, in its extended modality, making the local communities, the formal and informal networks established in the different territories, the families and the citizens, equally responsible for the proximity of eventual dangerous situations that they may be aware of.

In fact, the Portuguese system of protection is based on a holistic vision of preventive and protective intervention of the children, through the complementarity of the work developed by the informal networks and the specialized work done by the formal protection networks and the dynamic and reciprocal interactions and relations between the different actors different actors that make up the protection system.

With the committed and resilient work, the National Commission for the Promotion of the Rights and Protection of Children and Youth (CNPCJ) continued to be the safe haven for children in distress in Portugal. When analysing their 2021 Annual Report, we see that 69,727 (+4.8% than 2020) children were accompanied by the CPCJ, and the number of reports of danger has peaked to 73,241, increasing approximately 4.94% YoY (Relatórios de Atividade - Comissão Nacional de Promoção Dos Direitos e Proteção Das Crianças e Jovens, n.d.).

2.2. PREDICTIVE ANALYTICS

Predictive analytics is the use of data, statistical algorithm, and machine-learning techniques to identify the likelihood of future outcomes based on historical data. Predictive models use know results to develop a model that can be used to predict values for different or new data. The modelling results in predictions that represent a probability value for different or new data. Predictive analytics is used to predict trends, improve performance, drive decision making, and predict behaviour.

Thus, developing predictive analytics techniques and tools for modelling and simulation of historical data analysis is one of the main challenges that needs to be addressed to use big data effectively (Long et al., 2021).

‘Predictive Analytics is the art and science of using data to make better informed decisions. Predictive analytics helps you uncover hidden patterns and relationships in your data that can help you predict with greater confidence what may happen in the future, and provide you with valuable, actionable insights for your organization.’ (Introduction - Predictive Analytics For Dummies [Book], n.d.)

Additionally, industrial standards such as KDD, SEMMA and CRISP-DM define a set of sequential steps to guide the implementation of data mining applications (Azevedo & Santos, n.d.).

The KDD (Knowledge Discovery in Database) process is the process of using data mining methods to extract knowledge according to measures and thresholds while using a database along with a required preprocessing, sub sampling and transformation of the database. Its main steps are as follows:

1. Selection – Creating a target data set or focusing on a subset of variables or data samples, where discovery will be implemented.
2. Preprocessing – Target data cleaning to obtain consistency.
3. Transformation – Dimensionality reduction or transformation methods.
4. Data Mining – Searching for patterns usually for prediction purposes.
5. Interpretation/Evaluation – Understanding of the mined patterns discovered.

(Azevedo & Santos, n.d.)

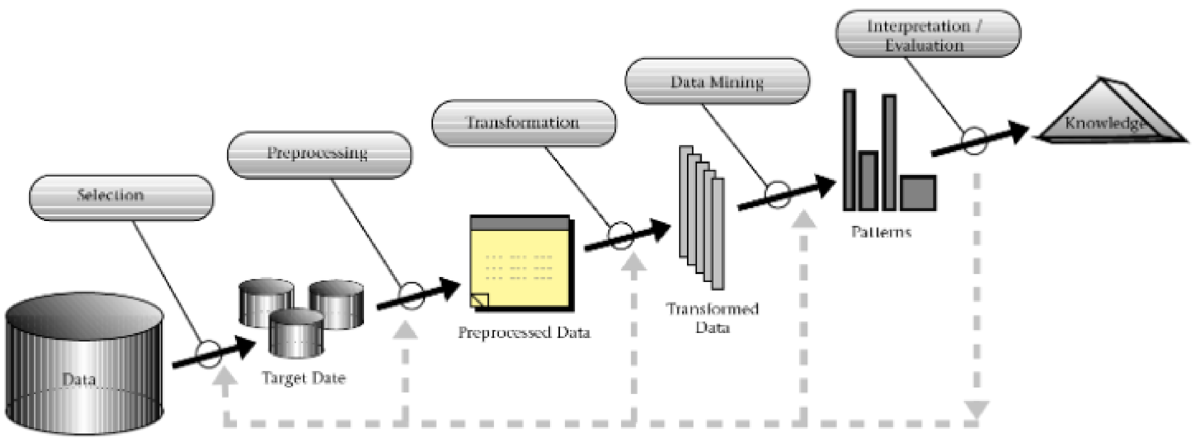

Figure 1 – Overview of the Steps That Compose the KDD Process [Fayyad at all., 1996]

The SEMMA (Sample, Explore, Modify, Model, Assess) process developed by the SAS Institute considers a cycle with 5 stages:

Sample – Optional stage based on sampling the data to contain the significant information and being set to be manipulated.
Explore – Searching unanticipated trends and anomalies.
Modify – Creating, selection, and transforming the variables to focus the model selection process.
Model – Allowing the software to search for a combination of data that accurately predicts a certain outcome.
Assess – Evaluating the usefulness and reliability of the findings.
(Azevedo & Santos, n.d.)

The CRISP-DM process (Cross-Industry Standard Process for Data Mining), illustrated in figure 2, consists of a cycle that comprises six stages:

Business understanding – Understanding the project objectives and requirements from a business perspective, then converting this knowledge into a data mining problem definition and a preliminary plan designed to achieve the objectives.

Data understanding – Initial data collection and familiarity with the data, to identify data quality problems, first insights and interesting subsets to form hypotheses for hidden information.
Data preparation – All activities to construct the final dataset from the initial raw data.

Modelling – In this phase, various modelling techniques are selected and applied, and their parameters are calibrated to optimal values.

Evaluation – At this stage the model (or models) obtained are more thoroughly evaluated and the steps executed to construct the model are reviewed to be certain it properly achieves the business objectives.

Deployment – Creation of the model is generally not the end of the project. Even if the purpose of the model is to increase knowledge of the data, the knowledge gained will need to be organized and presented in a way that the customer can use it.
(Azevedo & Santos, n.d.)

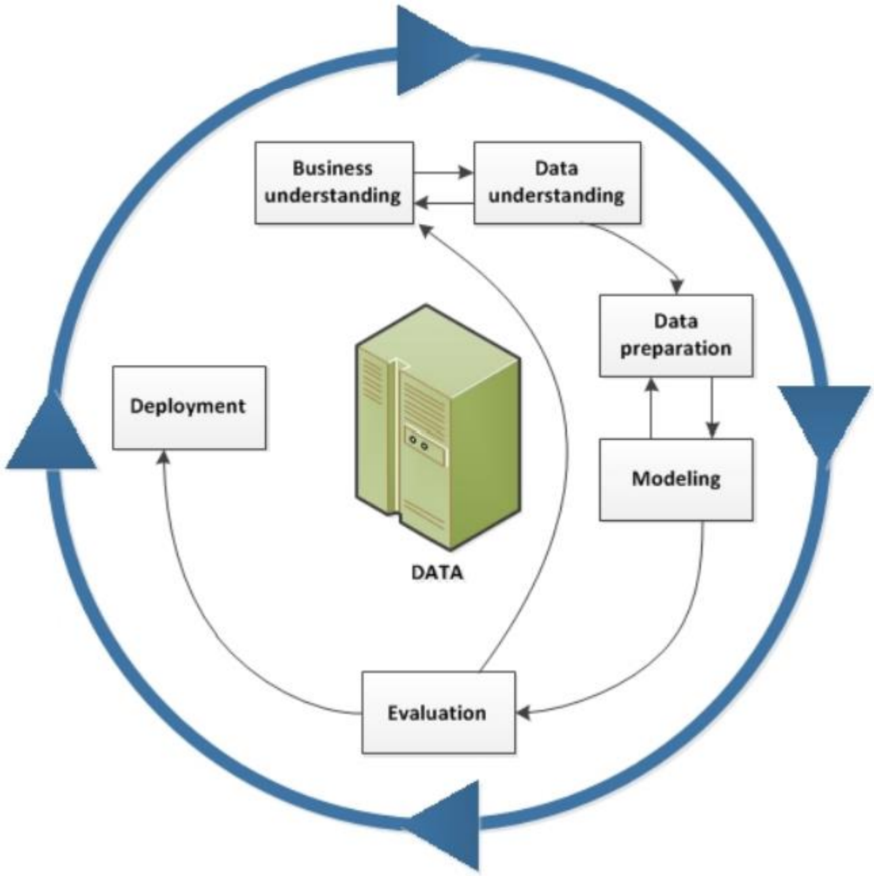

Figure 2 – Diagram showing the relationship between the different phases of CRISP-DM [IBM, 2023]

2.2.1. Methods, tools, and algorithms

Predictive analytics uses methods, tools, and algorithms of Data Mining, Machine Learning, Artificial Intelligence, and Statistics to develop a predictive model based on discovered patterns in past predictions. A score is given by mostly predictive analytics models. A higher score indicates the higher likelihood of occurrence of an event and a lower score indicates the lower likelihood of occurrence of the event (Kumar & L., 2018).

The process of predictive analytics passes through the following steps: requirement collection, data collection, data analysis and massaging, statistics/machine learning, predictive modelling, predictions & monitoring.

Requirement collection – In this stage, it will be identified which data is required in developing the model.

Data collection – After the requirements, the analyst will collect the datasets. This data may be in the structured or unstructured form and can be from different sources.

Data analysis and massaging – At this point, the data, in case it is unstructured, is converted into the structured form, where quality checks will then be applied, and the raw data will be formatted to deliver information.

Statistics, Machine Learning – Techniques will be applied such as probability theory, regression analysis, artificial neural networks, decision tree, support vector machines. Important to note that machine learning techniques have an advantage over conventional statistical ones, although statistical use must be involved in developing any predictive model (Kumar & L., 2018).

Predictive Modeling – In this phase, a model is developed based on statistical and machine learning techniques and the example dataset. After the development, it is tested on the test dataset which a part of the main collected dataset to check the validity of the model and if successful, the model is said to be fit. Once fitted, the model can make accurate predictions on the new data entered as input to the system.

Prediction and Monitoring – After the predictive modeling tests, the model is deployed and monitored to ensure accuracy.

2.2.2. Predictive Analytics Models

There are many types of machine-learning algorithms: supervised learning (the computer receives a set of inputs and desired outputs and aims to find the map between them, and returns a Boolean value), unsupervised learning (the computer receives a set of inputs but no desired outputs. The goal is to find the structure of data (probability distribution, groups), and reinforcement learning (aims to learn the behavior of software agents or robots based on feedback from the environment).

In supervised learning, the learner (a computer program) is provided with two sets of data, a training set (label for each data point) and a test set (without labels). The aim is to develop a rule, procedure that classifies new examples in the rest set by analyzing examples it has been given that already have a class label (Learned-Miller, n.d.).

Contrarily, in unsupervised learning, the model does not receive the labels of the examples. Instead, its goal is to find significant patterns in the input data that will enable the model to conclude its outcome. The most popular form of unsupervised learning is cluster analysis and dimensionality reduction (Alzubi et al., 2018).

In predictive analytics, supervised learning models are grouped into classification and regression models. If the output y is a scalar ($y \in R$ or R^q) the problem is known as a regression problem, such as weather forecasting, estimating life expectancy, population growth prediction, advertising popularity prediction, etc. If the output y is a label (categorical/discrete variable) $y \in \Omega$ $\Omega = \{\omega0, . . . , \omega K-1\}$ the problem is known as a classification problem, such has diagnostics, identity fraud detection, customer retention, image classification, etc. This project will focus on classification models.

2.2.2.1. Predictive Analytics Techniques – Algorithms for Classification

Logistic Regression

Consider a binary classification problem in which y ∈ {0, 1}. Logistic regression proposes a model for the a posteriori probability of the classes. This model guarantees that P(y = 1|x), P(y = 0|x) ∈ [0, 1] and P(y = 0|x) + P(y = 1|x) = 1.

Furthermore, logistic regression – figure 3 - usually states where the boundary between the classes exists, also states the class probabilities depend on distance from the boundary, in a specific approach. This moves towards the extremes (0 and 1) more rapidly when the data set is larger (Akinsola, 2017).

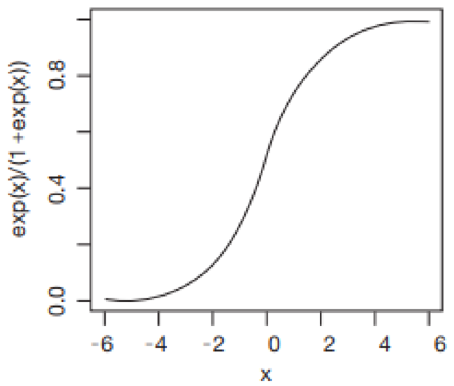

Figure 3 – Graph of logistic curve where α=0 and β=1 [J Koren Acad Nurs Vol.43 No.2, April 2013]

Decision Trees

As illustrated in figure 4, Decision Trees (DT) are trees that classify instances by sorting them based on feature values. Each node in a decision tree represents a feature in an instance to be classified, and each branch represents a value that the node can assume. Instances are classified starting at the root node and sorted based on their feature values (Akinsola, 2017).

Furthermore, DT split the training data into smaller subsets, in such a way that the label variables in each subset are as homogeneous as possible. The predictor (classifier) f (x) is assumed to be constant in each subset and defined as the most voted class.

The impurity of the tree drops or remains constant every time a node is split. In the limit we can grow the tree until each leaf is pure or all the attributes along that path have been used. This approach leads to overfitting. A second approach consists of using a validation technique. The tree is grown using a subset of training data (70%) and evaluated using the remaining patterns (30%) (validation set).

Thus, DT have the following drawbacks: early stop is not a good strategy to train the model because it suffers from lack of sufficient look ahead, it is better growing the tree until the leaves are pure or all attributes have been used and then prune the tree., and, finally, instability: a small change in the training patterns may lead to big changes of the decision boundaries.

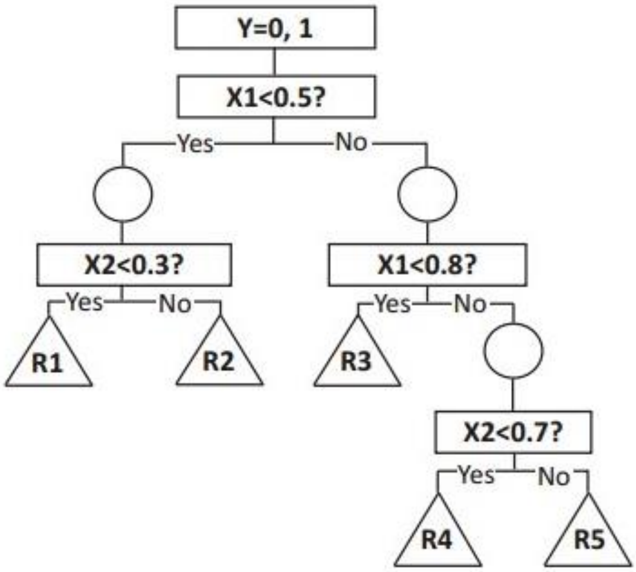

Figure 4 – Sample decision tree based on binary target variable Y [Shanghai Arch Psychiatry, 2015]

Random Forest

Random forest is a very simple and yet a very powerful classifier. Achieving state-of-the-art results in many problems, the algorithm is based on an ensemble of tree classifiers trained with bagging (Bagging and Random Forest in Machine Learning, n.d.).

In other words, it uses the bagging approach to create a bunch of decision trees with random subset of data. As shown in figure 5, the output of all decision trees in the random forest is combined to make the final decision trees. Additionally, there are two stages in Random Forest Algorithm, one is to create random forest, and the other is to make a prediction from the random forest classifier created in the first stage (Alzubi et al., 2018).

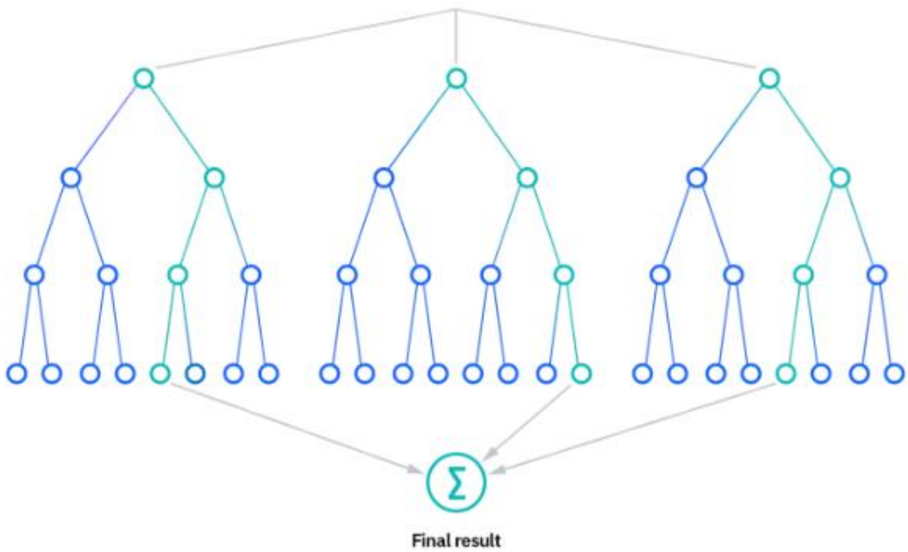

Figure 5 – Random Forest simplified [IBM, 2023]

Support Vector Machine

Support Vector Machines (SVMs) – figure 6 - were proposed by Vapnik and Chervonenkis in 1963 for binary classification problems with linear decision boundaries. They were extended later for nonlinear decision boundaries and for regression problems. The main objective is maximizing the margin and thereby creating the largest possible distance between the separating hyperplane and the instances on either side of it has been proven to reduce an upper bound on the expected generalization error (Akinsola, 2017).

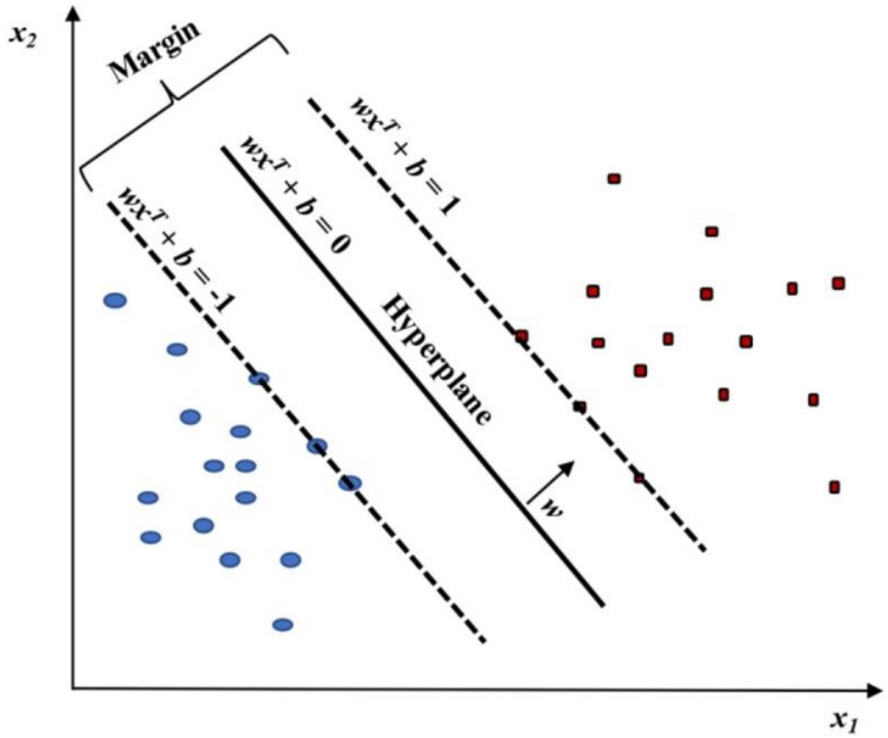

Figure 6 – Linear SVM model. Two classes (red versus blue) were classified [Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W., 2018]

k-Nearest Neighbor

Suppose we wish to predict a variable y knowing an input vector x ∈ Rp. Suppose we also know a collection of training examples (training set). A simple strategy to predict y for new values of x consists of finding the training pattern x(i) nearest to x and approximating y by y(i). The nearest neighbor (NN) method assigns x to the outcome of the nearest neighbor. In this way, the k-Nearest Neighbor, as shown below in figure 7, takes into account not one but k nearest neighbors of x. In classification problems, the predicted class is chosen as the most voted class in the sequence (y(1), . . . , y(k)).

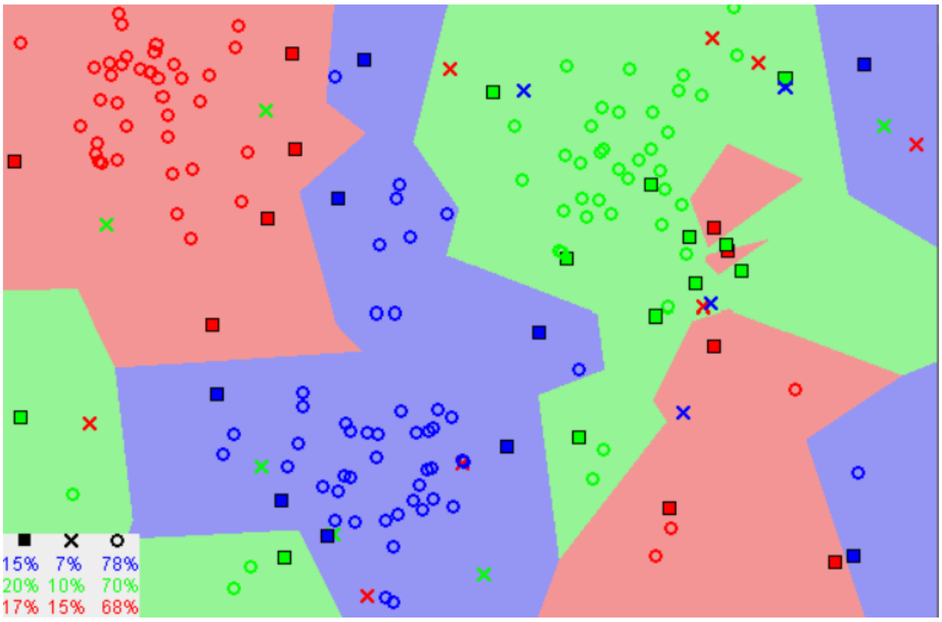

Figure 7 – Imagine showing how similar data points typically exist close to each other [Medium, 2018]

Naïve Bayes

Composed of directed acyclic graphs with only one parent (representing the unobserved node) and several children (corresponding to observed nodes) with a strong assumption of independence among child nodes in the context of their parent. Thus, the independence model (Naive Bayes) is based on estimating. Bayes classifiers are usually less accurate than other more sophisticated learning algorithms (such as ANNs). However, performed a large-scale comparison of the Naive Bayes classifier with state-of-the-art algorithms for decision tree induction, instance-based learning, and rule induction on standard benchmark datasets, and found it to be sometimes superior to the other learning schemes, even on datasets with substantial feature dependencies (Akinsola, 2017).

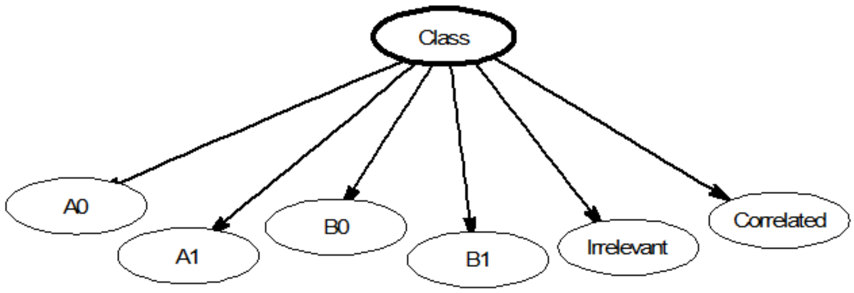

Figure 8 – Example of a Naïve Bayes classifier depicted as a Bayesian Network [IBM, 2023]

2.2.3. Applying Data Science in Social Welfare

Computational social welfare seems like a science well suited for solving modern social challenges. However, it has not yet been widely embraced and tested by social welfare scholars (Computational Social Welfare, n.d.).

The use of Data Science approaches in child protection systems can be seen in New Zealand, with the creation of a national database for vulnerable children and the application of PRM as being one means to select children for inclusion in it. This PRM has attracted academic attention suggesting that it will become more crucial in the provision of welfare services (Macchione et al., 2013, p. 374).

Furthermore, although still early in development, pediatric-specific Big Data projects – such as PEDSnet (PEDSnet Clinical Research Network, 2020), PhysioNet (PhysioNet Databases, n.d.), Children's Data Network (Children's Data Network – Harnessing the Scientific Potential of Linked, Administrative Data to Inform Children's Programs and Policies., n.d.) - have begun to emerge that include data commons for integration and interrogation of data from inpatient and ambulatory cohorts of children and electronic health record- or vital sign-based data for development of disease risk scores, while considering and maintaining high privacy and security standards (Vesoulis et al., 2022).

Indeed, ethical dilemmas arise from data science approach. Biases in machine learning applications seem to conflict with the open data movement. New standards for data and information are needed within organizations. Ultimately, the challenge for social work management is acquiring new skills based on new technologies. Research on social work and new technologies is increasing; however, knowledge of data science approach in specific is limited.

Therefore, full understanding of data science as a part of new communication technologies and as a tool in social work practice is required ("Web 2.0 in Social Work Macro Practice: Ethical Considerations and Ques" by Katharine M. Hill and Sarah M. Ferguson, n.d.). In addition, it is important to identify social workers who are more able in getting into technology-based practice (Electronic Information Systems and Social Work: Who Are We Designing For? Practice: Vol 26, No 5, n.d.).

The urgency of the COVID-19 pandemic has demonstrated the potential for rapid application of Big Data and data science to integrate and analyze data across information systems and these studies suggest the feasibility of the application of data science to child protection questions and the potential impact of such studies on prediction and mitigation of risk over decades of life.

2.3. PREDICTIVE ANALYTICS & CHILDREN' WELFARE – A SYSTEMATIC LITERATURE REVIEW

After an in-depth literature review, a systematic literature review was implemented with the aim of viewing the "lay of the land" with the current and peered review papers.

2.3.1. PRISMA Methodology

PRISMA stands for Preferred Reporting Items for Systematic Reviews and Meta-Analyses. The aim is to help authors improve the reporting of different types of research, and in turn to improv the quality of research used in decision-making (Emily Jones, n.d.).

Developed in 2005 by a group of 29 people, between scientists, authors, and other specialists from several areas, after an extended consensus they agreed on a 27-item checklist and a four-phase flow diagram, which resulted in a powerful template for researchers to use (Liberati et al., 2009).

The entire process is based on four phases that follow a simple workflow:

- Identification - Identify relevant articles, papers, and other scientific documentation based on a certain search strategy using the most common databases.
- Screening – Use a very well-defined criteria to include relevant papers and exclude those that bring no value to the research.
- Eligibility – Assess included articles for eligibility and exclude others for well justified reasons.
- Included – Result list of papers that will be used as full source for the research.

2.3.2. PRISMA Execution

The objective is to achieve a comprehensive understanding of the state of the art regarding the utilization of predictive analytics in children's welfare.

This research proposes to perform the following questions:

Table 2 – Systematic Literature Review's Research Questions

SLRQ1	What's the importance of predicting children in danger?
SLRQ2	What would the public sector AI look like?
SLRQ3	What are the major issues of predictive analytics in children's welfare?
SLRQ4	What are the advantages and disadvantages of applying predictive analytics techniques in this field?

To answer these questions, and according to the PRISMA definition, one has selected the most relevant studies in this field. To conduct the search, a set of keywords that one has considered to be more relevant amongst the several concepts analyzed in the theoretical background has been chosen. One has opted to use only English words, and therefore the outcome of the search was mostly articles written in English. The ones written in other languages were excluded from the selection, according to the criteria defined in one's PRISMA execution.

Table 3 – Systematic Review's Keywords

Keywords	Predictive Analytics	Child maltreatment
	Machine Learning	Child abuse

A specific search string was built to include the above words or terms with the objective of finding them in abstracts, titles or keywords of articles and other scientific papers. This choice of words assured that the results of the search would only retrieve data that's relevant for the topics in study.

One was interested only in scientific documents that are recent, as this technology evolves very fast, and as such, only the most recent articles can ensure up-to-date and relevant information. For this, one has set a filter to show only articles between 2018 and 2023, aiming to obtain accurate information about the current state of the art on the utilization of predictive analytics in the children's welfare field. The search string used was: ("Machine Learning" OR "Predictive Analytics") AND ("Child maltreatment" OR "Child abuse").

The search was conducted in January 2023 on the following scientific information resource databases:

Table 4 – Systematic Review's Resource Databases

Resource Database	**Resource URL**
Science Direct	https://www.sciencedirect.com/
IEEE Xplore	https://ieeexplore.ieee.org/
JSTOR	https://www.jstor.org/

The next step is to define the inclusion and exclusion criteria for the articles from the mentioned search.

Table 5 – Systematic Review's inclusion and exclusion criteria

<table>
<tr><th>Inclusion Criteria</th><th>Exclusion Criteria</th></tr>
<tr><td>Paper is published between 2018 and 2023</td><td>Articles before 2018</td></tr>
<tr><td rowspan="4">Paper must be a peered reviewed conference or journal paper written in English</td><td>Articles not in English and duplicate papers</td></tr>
<tr><td>Non-academic or non-scientific papers (e.g., websites, magazines reports, newspapers, consulting articles, books, citations)</td></tr>
<tr><td>Themes not relevant to the investigation</td></tr>
<tr><td>Lack of access to the article</td></tr>
</table>

After inserting the search string in sources websites, as output one has got all the identified records through database search, which resulted in a total of n=15994 with JSTOR (n=624), ScienceDirect (n=15349), IEEE(n=21) i.e., we're in the identification phase of the PRISMA workflow.

When moving to the screening phase, the first step is to remove articles before 2018 (n=5569) and then the inaccessible ones (n=1065). Here, (n=14929) articles have been removed, moving to the second step of the screening phase a total of (n=1065) records. In this second step of the screening phase, the inclusion and exclusion criteria have been applied: non-academic or non-scientific papers and duplicated records (n=153). Moving on to exclusion criteria of not relevant to the investigation which excluded n= articles. Thus, we move to the eligibility phase with a total of (n=17) articles.

In the eligibility phase, full-text articles were further analyzed and the ones that didn't have direct relevance to the study were excluded. Thus, resulting in a total of (n=) articles included in the meta-analysis. This process is represented in the following workflow picture (Figure 9):

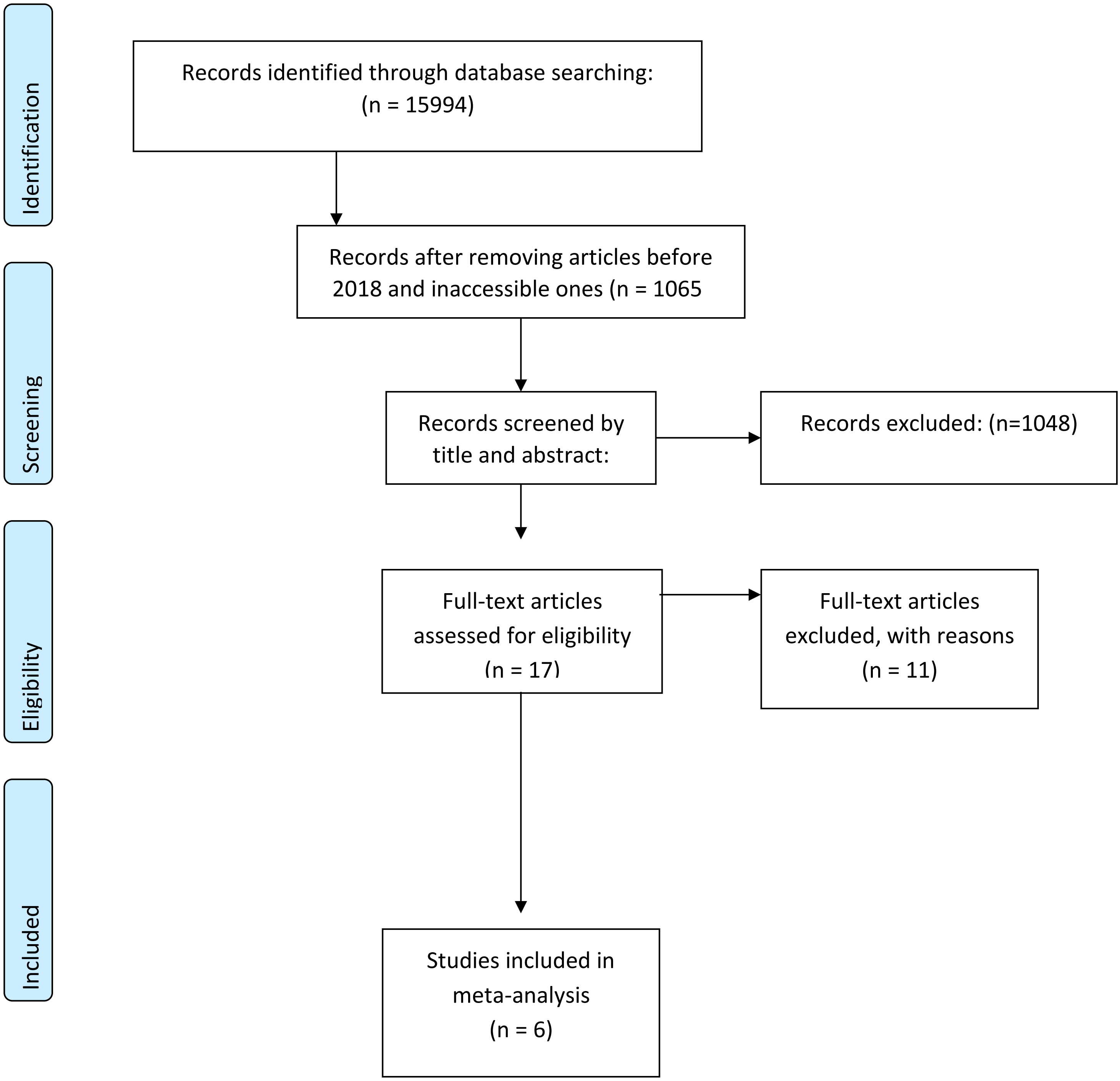

Figure 9 – PRISMA Execution

The output of this research is listed in the following table, along with a description of its contribution and conclusions.

Table 6 - PRISMA results table – included articles

#	Authors	Article	Main conclusions
[1]	(Margetts, 2022)	Rethinking AI for Good Governance	This paper revealed that AI can mitigate some structural biases and be used to tackle some profound inequalities in the distribution of resources and the design and the delivery of public services such as education, social welfare, and health care. In addition, this would lead to a resilient policy-making that would involve building data flows and using agent computing, machine learning, and other AI methodologies to create integrative models to both recover from the current crisis and face future shocks in a more anti fragile way.
[2]	(Penner & Dodge, 2019)	Using Administrative Data for Social Science and Policy	The lack of data infrastructure has human costs and puts science at a disadvantage. Recent efforts to create administrative data infrastructure have great promise to rectify the situation, making it

			an exciting time to be an administrative data researcher. In moving forward, coordinating efforts to ensure that we build the best data infrastructure possible, and that our data can benefit the public as much as possible is crucial.
[3]	(Oswald, 2018)	Algorithm-assisted decision-making in the public sector	Carefully considering exactly what the algorithm is or is not predicting and explaining to the decision-maker at the point results are displayed, is key to ensuring this fairness. The incorporation of an algorithm into a decision-making process may come with the risk of creating 'substantial' or 'genuine' doubt as to why decisions were made and what conclusions were reached, both for the subject of the decision and for the decision-maker themselves.
[4]	(Wesarg et al., 2022)	Childhood adversity and vagal regulation: A systematic review and meta-analysis	Childhood adversity, including exposure to abuse, neglect, poverty, and neighborhood violence, is among the most robust risk factors

			for the development of chronic health problems and psychopathology.
[5]	(Hamaker et al., 2020)	Description, prediction, and causation: Methodological challenges of studying child and adolescent development	Research in the social and behavioral sciences can be divided into having a descriptive, predictive, or explanatory goal. The aim of this paper was to shed light on the connection between research goals and methodology. It concluded that a major threat in prediction is overfitting, which is almost guaranteed to happen when no specific actions are taken to avoid it. Overfitting occurs when variables are selected that contribute to the prediction in the current sample but worsen prediction in other samples. To avoid the risk of overfitting, researchers can use cross-validation techniques. This technique is based on splitting the data into a training set and a validation set (also known as the holdout or test set or sample). The concept of cross-validation lies at the

			heart of supervised learning.
[6]	(D’arcy-Bewick et al., 2022)	Childhood abuse and neglect, and mortality risk in adulthood: A systematic review and meta-analysis	Those with histories of abuse and neglect in childhood appear to be at increased risk of mortality in adulthood. Analysis of possible sex differences identified an association between physical and emotional abuse, and greater mortality risk among women.

2.3.3. PRISMA Results Analysis

Among the results identified, it was possible to provide answers to the initial systematic literature review questions.

SLRQ1 - What’s the importance of predicting children in danger?

Research findings indicate that children with histories of abuse and neglect in childhood appear to be at a higher risk of mortality in adulthood, especially among women (D’arcy-Bewick et al., 2022).
Supplementary, these factors will also lead to the development of chronic health problems and psychopathology (Wesarg et al., 2022).

Thus, it is now more crucial than ever to protect future generations and gain an integrated understanding of the mechanisms underlying childhood adversity (Wesarg et al., 2022).
What would the public sector AI look like?
The pace of AI development in government seems to be accelerating. Governments would be able to prioritize innovation and overcome some of the recurring challenges. Furthermore, AI is able to help with core tasks of the government since it can enable real-time, transactional data to improve and enhance detecting tools to then build predictive models and support decision-making that will align with the intended outcomes (Margetts, 2022).

What are the major issues of predictive analytics in children’s welfare?

A sensitive topic in children’s welfare predictive analytics is the fact that the responsible decision-maker must be the one to determine whether the database used by the algorithm is indeed identical, i.e., that it represents all the factors that should be taken into account to build the predictive risk model. In addition, this person must determine whether the decision under consideration matches the one for which the algorithm was developed—for instance, an assessment of ‘risk’ may encompass much more than the forecast of a particular behavior by an algorithm—and whether the data on which the algorithm was trained match the circumstances of the current situation. In other words, a machine can make predictions based on the assumptions previously programmed, but it is the person who will indeed check these same assumptions (Oswald, 2018).

What are the advantages and disadvantages of applying predictive analytics techniques in this field?
Perhaps the biggest benefit of the use of AI would be to tackle issues of equality and fairness in governmental systems in a profound and transformative way, since it would be able to identify and reform long-standing biases in resource allocation, decision-making, the administering of justice, and the delivery of services. Moreover, more research findings on the field help to inform public opinion around these social questions and also play an educational role for future generations.

On the other hand, our current legal system is ill equipped to consider issues beyond an individualistic framework, so that harm to a collective group may not be recognized. Furthermore, another challenge faced by predictive analytics concerns the public skepticism about the limits of confidentiality and data protection threatens public support for the use of administrative data, especially taking into account the more recurring hacking events and misuse of large private data files (Penner & Dodge, 2019).

In a more technical side, a major threat in predictive analytics is overfitting, which is almost guaranteed to happen when no specific actions are taken to avoid it. A good practice is to use cross-validation techniques in these situations (Hamaker et al., 2020).

3. METHODOLOGY

3.1. OUTLINE

After a comprehensive literature review, it is time to review and improve the methodology for the conducted project. To achieve the goal of this study a methodological path will be followed divided into three different phases: exploration phase; analytical phase; and conclusive phase. Each phase is divided into specific steps as identified in figure 4. The analytical phase will use the Cross Industry Standard Process for Data-Mining also known as CRISP-DM – figure 5, which is an industry-independent process model developed in 1999 with funding from the European Union to develop a universal processes for data mining and consists of six iterative phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment (Schröer et al., 2021).

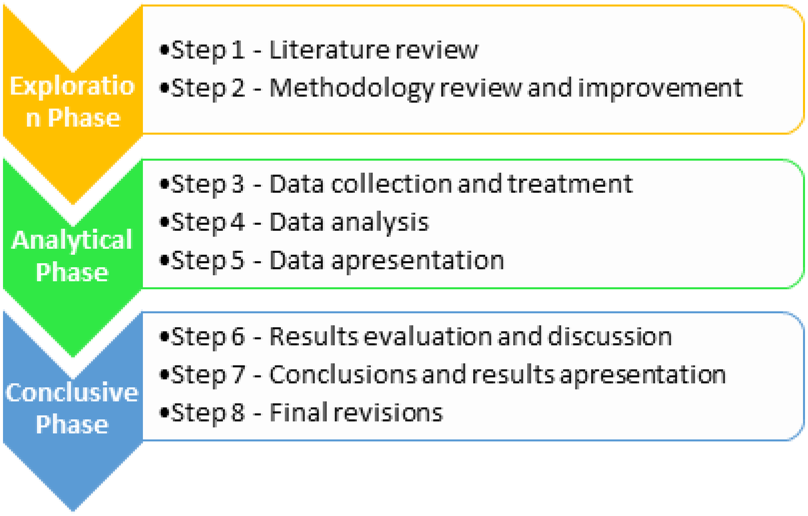

Figure 10 – Methodological approach

The CRISP-DM approach provides the necessary structure for the project since being a cross-industry standard, it can be implemented in any data science project irrespective of its domain. The following sections describe the process completed for this thesis.

The hereby table 7 describes the main idea, tasks, and outputs of the CRISP-DM phases.

Table 7 – Phase and description of CRISP-DM phases [Schröer et al., 2021]

Phase	**Brief description**
Business understanding	The business situation and the objectives are assessed and addressed to get a holistic overview of the available and required

	resources. Thus, it is crucial to define a compulsory project plan, context, and motivation.
Data understanding	Collection and exploration of the raw data. In addition, this phase incorporates data quality checks using statistical analysis and determining attributes and their collations.
Data preparation	In this stage, data cleaning (missing variables, outlier treatment, if applicable, feature selection & engineering) is an important step to fine-tune the dataset.
Modeling	The data modeling phase determines the modeling technique, the model, specific parameters, and the test case. For accurately assessing the model it is appropriate to cross-validate it.
Evaluation	Results are checked against the defined business objectives and the overall project is critically reviewed.
Deployment	The deployment phase consists of planning the deployment, monitoring, and maintenance.

3.2. DATA UNDERSTANDING (COLLECTION AND DESCRIPTION)

The dataset was initially collected from CNCPCJ with 55322 rows and five worksheets: Processes (26 columns), Signalization (4 columns), Main Caregiver (4 columns), Risk Factor (5 columns), Protection Measure (4 columns).

Hereby is described the name and description of each column.

Table 8 – Column and description of CNCPCJ dataset

Column	Description
ID_PROCESSO	The ID of each process that was registered in their system from 2018 and 2022.
COD_CPCJ_DETENTORA	Code of each CPCJ location.
DSC_CPCJ	Name description of each CPCJ.
REABERTURA_PROCESSO	Indication if the process was reopened.
DATA_NASCIMENTO	Birth date of the registered child.

COD_IDADE_REABERTURA	Age of the child.
COD_SEXO	Gender of the child.
PAIS_NACIONALIDADE	Country of birth.
DSC_ENQUAD_SOCIO_EDUCATIVO	Characterization if the child is already a part of any socio-educational framework.
DSC_FREQUENCIA_ESCOLAR	Information regarding if the child is frequenting school.
DSC_GRAU_ESCOLARIDADE	Information about the education level.
DSC_MODALIDADE_ENSINO	Description of the education modality.
DSC_PROFISSAO	Information about professional status.
DSC_TIPO_HABITACAO	House type of child.
DSC_ESTADO_CONSERVACAO	Status of the house where the child lives.
DSC_ELEMENTO_VIVE_CRIANCA	With whom the child lives with.
DSC_PODER_PATERNAL	Parental responsibility.
DSC_TIPO_AGREGADO	Type of household.
EXISTEM_FRATRIAS	Information regarding if the child has siblings.
COD_LUGAR_OCUPA	If child is first-born.
EXISTEM_FRATRIAS_SISTEMA_PP	If siblings are already registered in the system.
QTD_IRMAOS	Number of siblings.
EXISTEM_DESCENDENTES	Existence of descendants.
DESCENDENTES_SISTEMA_PP	Presence of descendants in the system.
QTD_FILHOS	Number of descendants.
ASCENDENDES_SISTEMA_PP	Information about the existence of ascendents in the system.
DSC_TIPO_SITUACAO_PERIGO	The type of danger situation faced by the child.
DSC_LEGITIMIDADE_INTERVENCAO	Information regarding the entitlement of intervention by the CPCJ.
DSC_DELIBERACAO	Deliberation action from CPCJ.
PRINCIPAL_CUIDADOR	Main caregiver.
DSC_CARACTERIZACAO_RENDIMENTO	Economic and financial information of the main caregiver.
TIPO_PROBLEMA_SAUDE	Health status of the main caregiver.
DSC_DIMENSAO_FACTOR	Dimension of the risk factor.
DSC_DOMINIO_FACTOR	Field where the action of CPCJ needs to be.
DSC_SUBDOMINIO	Subfield of the action of CPCJ.
DSC_TIPO_FACTOR	Information if the situation encountered is a risk of a protection factor.

DSC_TIPO_MEDIDA	Type of measure that needs to be implemented.
MEDIDA_EXECUTADA	Information regarding if measure was executed.
MEDIDA_CONCRETIZADA	Information regarding success of the measure implemented and executed.

3.3. DATA PREPARATION

Phase 1 - Exclusion of Variables

The exclusion of variables was based on the following criteria: not relevant for the study, over 90% missing values in a certain column since it wouldn't allow to search insights and reach conclusions. Further in the study we will apply the Pearson correlation to investigate additional dataset changes.

Thus, the following columns were taken from the investigation.

Table 9 - Column and description of CNCPCJ dataset after phase 1

Column name	Justification
ID_PROCESSO	The ID of each process that was registered in their system from 2018 and 2022.
DATA_NASCIMENTO	This study focuses on every child under the age of 18 years old, thus age mapping will not be relevant.
DSC_CPCJ	Redundant information since we already have the code of each CPCJ location in column COD_CPCJ_DETENTORA
REABERTURA_PROCESSO	Not relevant for the study as it is aimed to read each row as a unique case and not historical evolution.
COD_IDADE_REABERTURA	Aligned with the above variable, it is not intended to characterize historical information.
EXISTEM_FRATRIAS	Redundant information since in column "QTD_IRMAOS" it is reflected if the child has siblings.
COD_LUGAR_OCUPA	Not relevant for the study.

EXISTEM_DESCENDENTES	Redundant information since in column “QTD_FILHOS” it is reflected the number of descendants.

Additionally, since the column DSC_DELIBERACAO defines if CPCJ will pursue with the case and later in time implement according to measures, rows without the value “Avaliação Diagnóstica” were eliminated, since the rest don’t indicate that the child is in danger and suggest that either the case was archived or CPCJ didn’t reach a conclusion.

Furthermore, for this study it is crucial that in the success measure, in other words, the column MEDIDA_CONCRETIZADA has a value and not an unknown information value. Due to this, rows with “Unknown” value were deleted from the database.

After this process, we were left with 47088 rows and 33 columns. It was evaluated the ratio of missing values in the columns and the ones with over 90% of ‘NA’ were deleted. Thus, column ‘DSC_PROFISSAO’, ‘TIPO_PROBLEMA_SAUDE’, and ‘DESCENDENTES_SISTEMA_PP’ were removed.

Phase 2 - Conversion of data

In order to effectively handle the dataset containing text values, a dictionary was created to facilitate the transformation of textual information into numerical representations. This process of encoding text data into numerical values is crucial for many machine learning algorithms that require numeric input. By mapping each unique text value to a specific numerical code, the dictionary allows for a seamless integration of the textual features into the analysis pipeline.

The creation of this dictionary enables the conversion of qualitative information into a quantitative format, facilitating further analysis and modeling. This approach ensures that the underlying patterns and relationships within the text data can be effectively captured and leveraged by machine learning algorithms. Moreover, it provides a standardized and consistent representation of the textual features, enhancing the interpretability and comparability of the results.

By employing this dictionary-based approach, the dataset's text values can be seamlessly integrated into various machine learning techniques, allowing for comprehensive analysis and insightful decision-making based on the transformed numerical representations.

Phase 3 – Pearson Correlation

During the analysis of Pearson Correlation, it was observed that three variables exhibited a correlation exceeding 70%: DSC_DIMENSAO_FACTOR, DSC_DOMINIO_FACTOR, and DSC_TIPO_FACTOR. This strong correlation suggests a significant relationship between these variables, indicating that changes in one variable may be associated with similar changes in the others.

Thus, since the three variables are highly similar or redundant in nature, DSC_DIMENSAO_FACTOR was removed to mitigate multicollinearity.

Hereby is described the name and description of each column of the final table.

Table 10 - Column and description of final CNCPCJ dataset

Column	Description
COD_CPCJ_DETENTORA	Code of each CPCJ location.
COD_SEXO	Gender of the child.
PAIS_NACIONALIDADE	Country of birth.
DSC_ENQUAD_SOCIO_EDUCATIVO	Characterization if the child is already a part of any socio-educational framework.
DSC_FREQUENCIA_ESCOLAR	Information regarding if the child is frequenting school.
DSC_GRAU_ESCOLARIDADE	Information about the education level.
DSC_MODALIDADE_ENSINO	Description of the education modality.
DSC_PROFISSAO	Information about professional status.
DSC_TIPO_HABITACAO	House type of child.
DSC_ESTADO_CONSERVACAO	Status of the house where the child lives.
DSC_ELEMENTO_VIVE_CRIANCA	With whom the child lives with.
DSC_PODER_PATERNAL	Parental responsibility.
DSC_TIPO_AGREGADO	Type of household.
EXISTEM_FRATRIAS_SISTEMA_PP	If siblings are already registered in the system.
QTD_IRMAOS	Number of siblings.
DESCENDENTES_SISTEMA_PP	Presence of descendants in the system.
QTD_FILHOS	Number of descendants.
ASCENDENDES_SISTEMA_PP	Information about the existence of ascendents in the system.
DSC_TIPO_SITUACAO_PERIGO	The type of danger situation faced by the child.

DSC_LEGITIMIDADE_INTERVENCAO	Information regarding the entitlement of intervention by the CPCJ.
DSC_DELIBERACAO	Deliberation action from CPCJ.
PRINCIPAL_CUIDADOR	Main caregiver.
DSC_CARACTERIZACAO_RENDIMENTO	Economic and financial information of the main caregiver.
TIPO_PROBLEMA_SAUDE	Health status of the main caregiver.
DSC_DOMINIO_FACTOR	Field where the action of CPCJ needs to be.
DSC_SUBDOMINIO	Subfield of the action of CPCJ.
DSC_TIPO_FACTOR	Information if the situation encountered is a risk of a protection factor.
DSC_TIPO_MEDIDA	Type of measure that needs to be implemented.
MEDIDA_EXECUTADA	Information regarding if measure was executed.
MEDIDA_CONCRETIZADA	Information regarding success of the measure implemented and executed.

4. EVALUATION AND DISCUSSION

In this work, six different classification machine learning algorithms— Support Vector Machine (SVM), Naïve Bayes (NB), K-nearest neighbor (KNN), Logistic Regression (LR), Random Forest (FR), and Decision Tree (DT) —were used to detect the accuracy and to find the best-suited algorithm which can efficiently learn the pattern of children in danger. The data gathered from the feature set comparison was then applied as input as data feeds to train the system for future prediction and analysis using the best-fit algorithm chosen from the above three algorithms based on the performance metrics found. Also, the classification reports (Precision, Recall, and F1-score) and confusion matrix were generated and compared to finalize the support-validation status found throughout the testing phase of the model used in this approach.

A confusion matrix was used in this project to compare the number of predicted instances in each class and their correct classification (Tiwari, 2022). It is a straightforward representation of the model results, and a starting point to understand other metrics. The confusion matrix for the binary problem can be described as below:

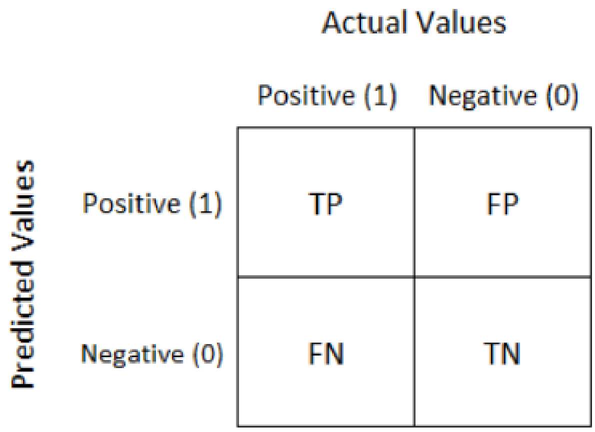

Figure 11 – Confusion Matrix

- “true positive” for correctly predicted event values.
- “false positive” for incorrectly predicted event values.
- “true negative” for correctly predicted no-event values.
- “false negative” for incorrectly predicted no-event values.

In addition, the classification report, which was also used and analyzed, is a visualization draft that displays the four base-parameters of a classification model—Precision, Recall, F1-score, and Support to generate the accuracy level because of model fitting. It helps in more straightforward interpretation and detection by integrating numerical scores with the help of a color-coded heat-map (Agarwal et al., 2021).

Accuracy is the most important performance metric that defines how capable the classification model (classifier) is. It means how accurately the algorithm is learning the data patterns in the dataset and how accurately it can predict unseen data.

Precision is an important performance metric that needs to be considered. It is the ratio of correctly observed positive results to all observed positive results.

The recall is the ratio of correctly observed positive results to the total observations in a class. It gives the result as the ratio of positive observations.

F1-score is an important performance metric to consider. In some cases, F1-score has more importance than accuracy. Sometimes in a large dataset, the cost of false positives and false negatives are not the same. When they are the same, accuracy is a better option. But, when they are not the same, we need to investigate F1-score.

Support is the number of true results observed that occurs in a class. It indicates the number of true results where predicted and actual results are the same.

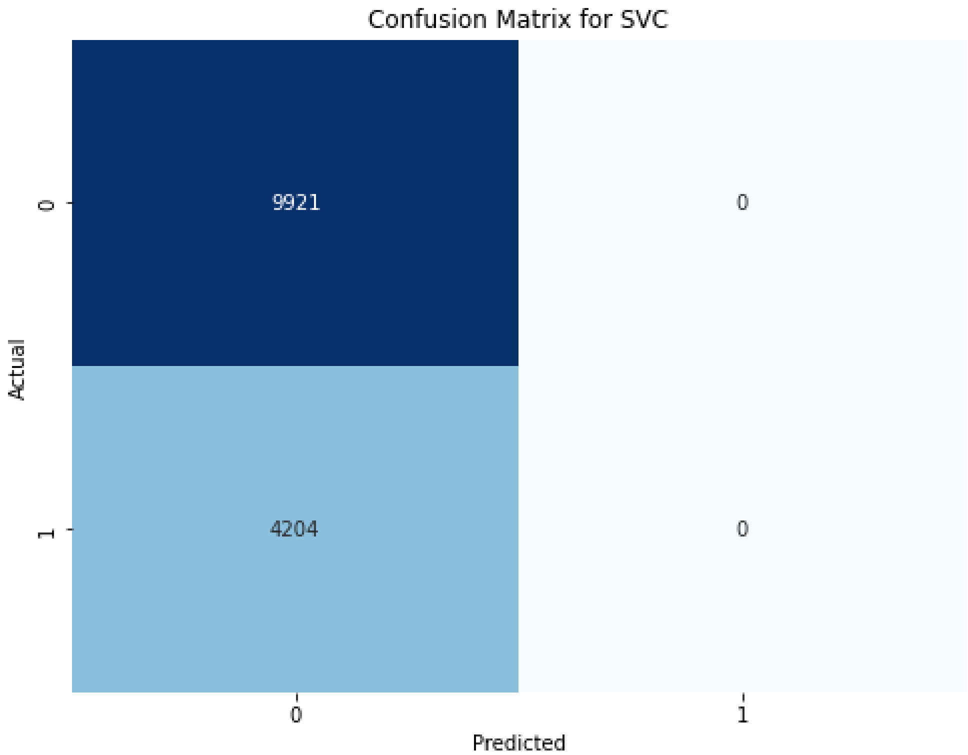

Figure 12 – Confusion matrix for SVM

Table 11 – Classification report for SVM

	Precision	Recall	F1-Score	Support
0	**0.7**	**1**	**0.83**	**9921**
1	**0**	**0**	**0**	**4204**
Accuracy			**0.7**	**14125**
Macro avg	**0.35**	**0.5**	**0.41**	**14125**
Weighted avg	**0.49**	**0.7**	**0.58**	**14125**

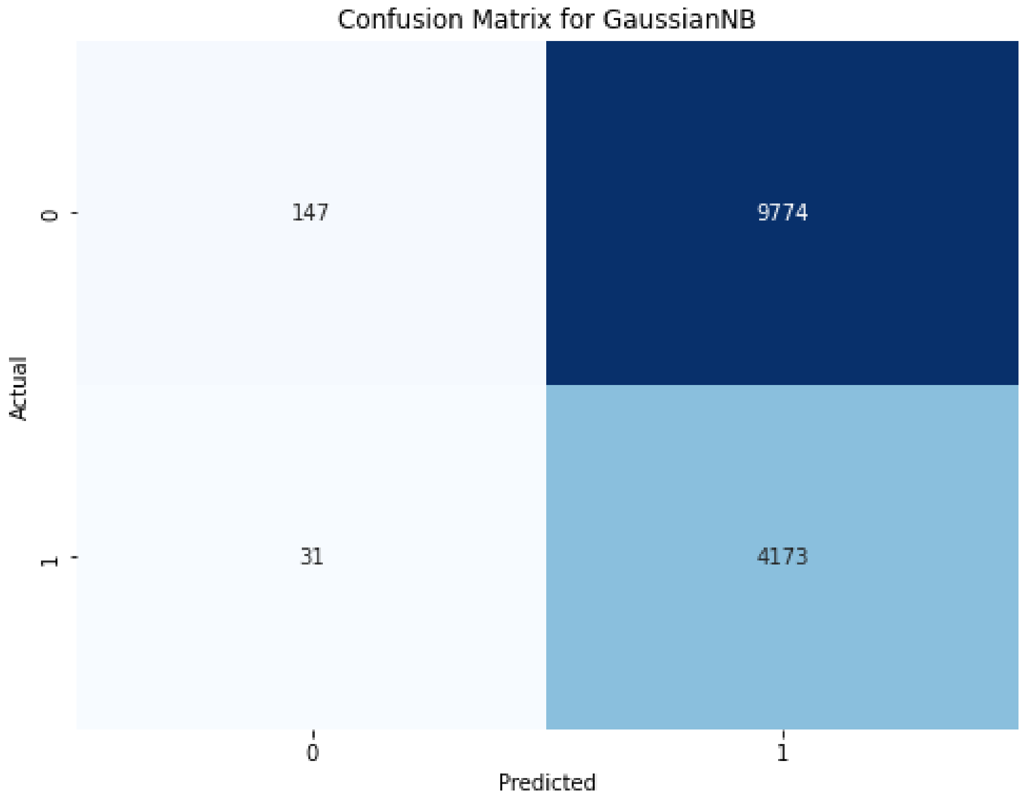

Figure 13 – Confusion matrix for NB

Table 12 – Classification report for NB

	Precision	Recall	F1-Score	Support
0	**0.83**	**0.01**	**0.03**	**9921**
1	**0.3**	**0.99**	**0.46**	**4204**
Accuracy			**0.31**	**14125**
Macro avg	**0.56**	**0.5**	**0.24**	**14125**
Weighted avg	**0.67**	**0.31**	**0.16**	**14125**

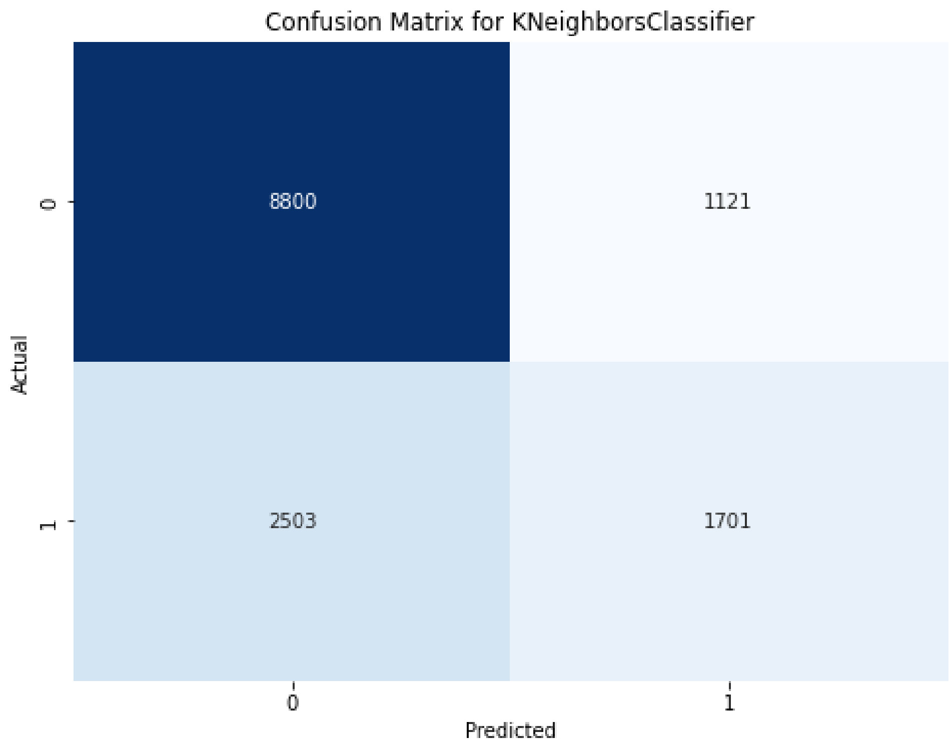

Figure 14 – Confusion matrix for KNN

Table 13 – Classification report for KNN

	Precision	Recall	F1-Score	Support
0	0.78	0.89	0.83	9921
1	0.60	0.40	0.48	4204
Accuracy			0.74	14125
Macro avg	0.69	0.65	0.66	14125
Weighted avg	0.73	0.74	0.73	14125

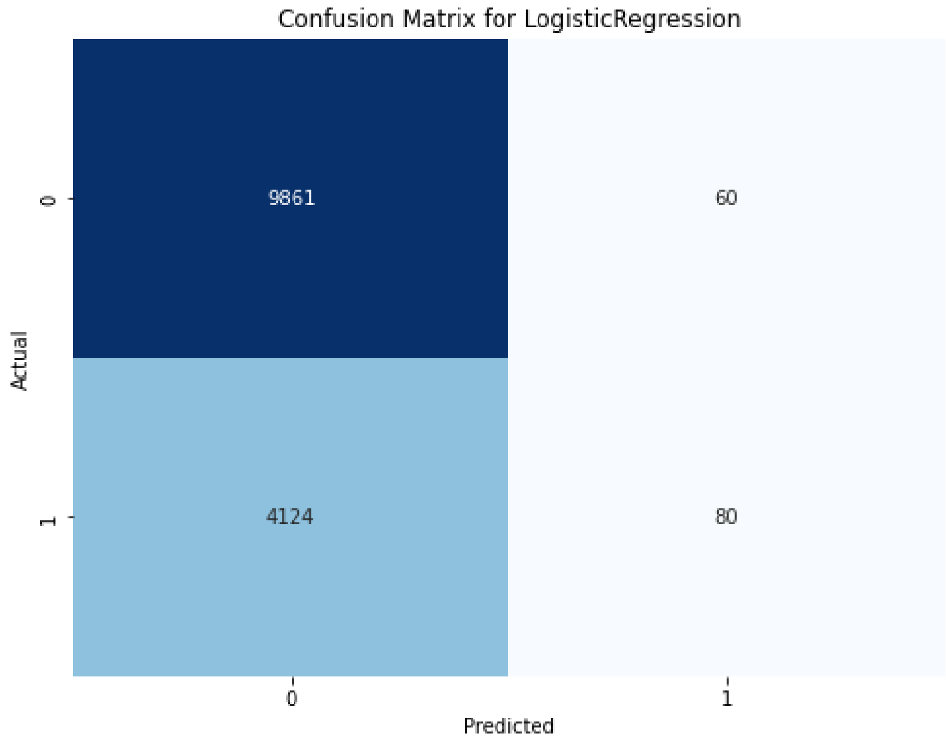

Figure 15 – Confusion matrix for LR

Table 14 – Classification report for LR

	Precision	Recall	F1-Score	Support
0	0.71	0.99	0.82	9921
1	0.57	0.02	0.04	4204
Accuracy			0.7	14125
Macro avg	0.64	0.51	0.43	14125
Weighted avg	0.67	0.7	0.59	14125

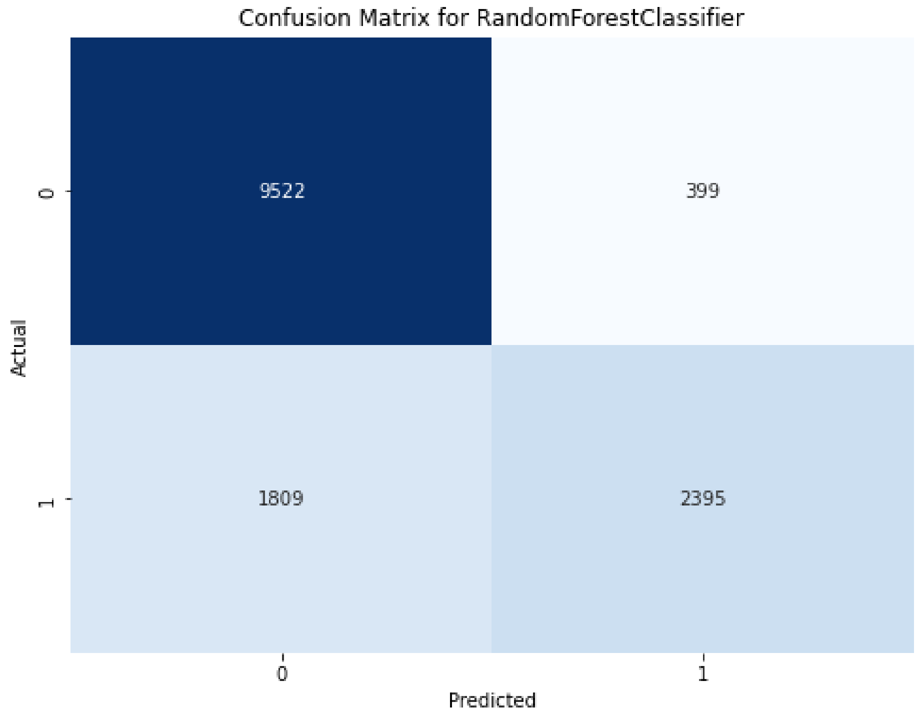

Figure 16 – Confusion matrix for RF

Table 15 – Classification report for RF

	Precision	Recall	F1-Score	Support
0	0.84	0.96	0.9	9921
1	0.86	0.57	0.68	4204
Accuracy			0.84	14125
Macro avg	0.85	0.76	0.79	14125
Weighted avg	0.85	0.84	0.83	14125

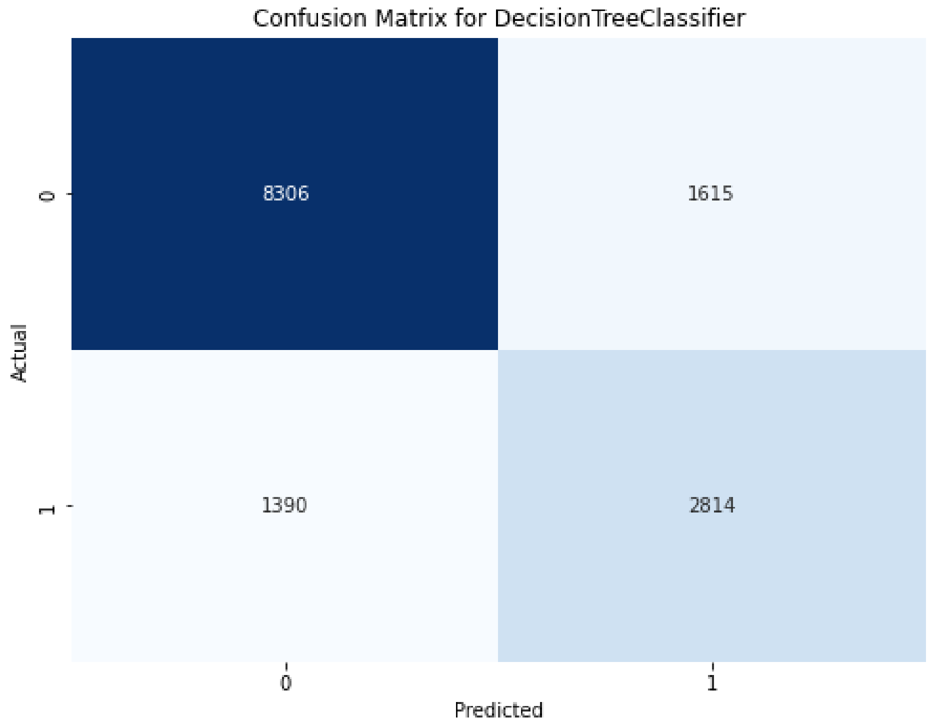

Figure 17 – Confusion matrix for DT

Table 16 – Classification report for DT

	Precision	Recall	F1-Score	Support
0	**0.86**	**0.84**	**0.85**	**9921**
1	**0.64**	**0.67**	**0.65**	**4204**
Accuracy			**0.79**	**14125**
Macro avg	**0.75**	**0.75**	**0.75**	**14125**
Weighted avg	**0.79**	**0.79**	**0.79**	**14125**

Based on the obtained results, it can be concluded that the Random Forest Classifier demonstrates higher accuracy on the utilized dataset, making it the most suitable choice for our model. Upon closer examination, among the six algorithms used in the proposed model, Random Forest achieved the most significant improvement in accuracy, with an approximate rate of 84.4% when evaluated using 30% of the test dataset.

Moreover, to gain further insights into the discussed models, several techniques can be employed to extract information about the importance of features in each classifier's training for predicting the target variables.

Considering that Random Forest outperformed the other models in terms of balanced accuracy, an investigation was conducted to determine the features that played a more crucial role in the prediction results. The top 10 features in terms of importance primarily consisted of historical features, including DSC_GRAU_ESCOLARIDADE, QTD_IRMAOS,

DSC_TIPO_AGREGADO, DSC_CARATERIZACAO_RENDIMENTO, and DSC_TIPO_SITUACAO_PERIGO. These features exhibited high importance rankings, as shown in Figure 18.

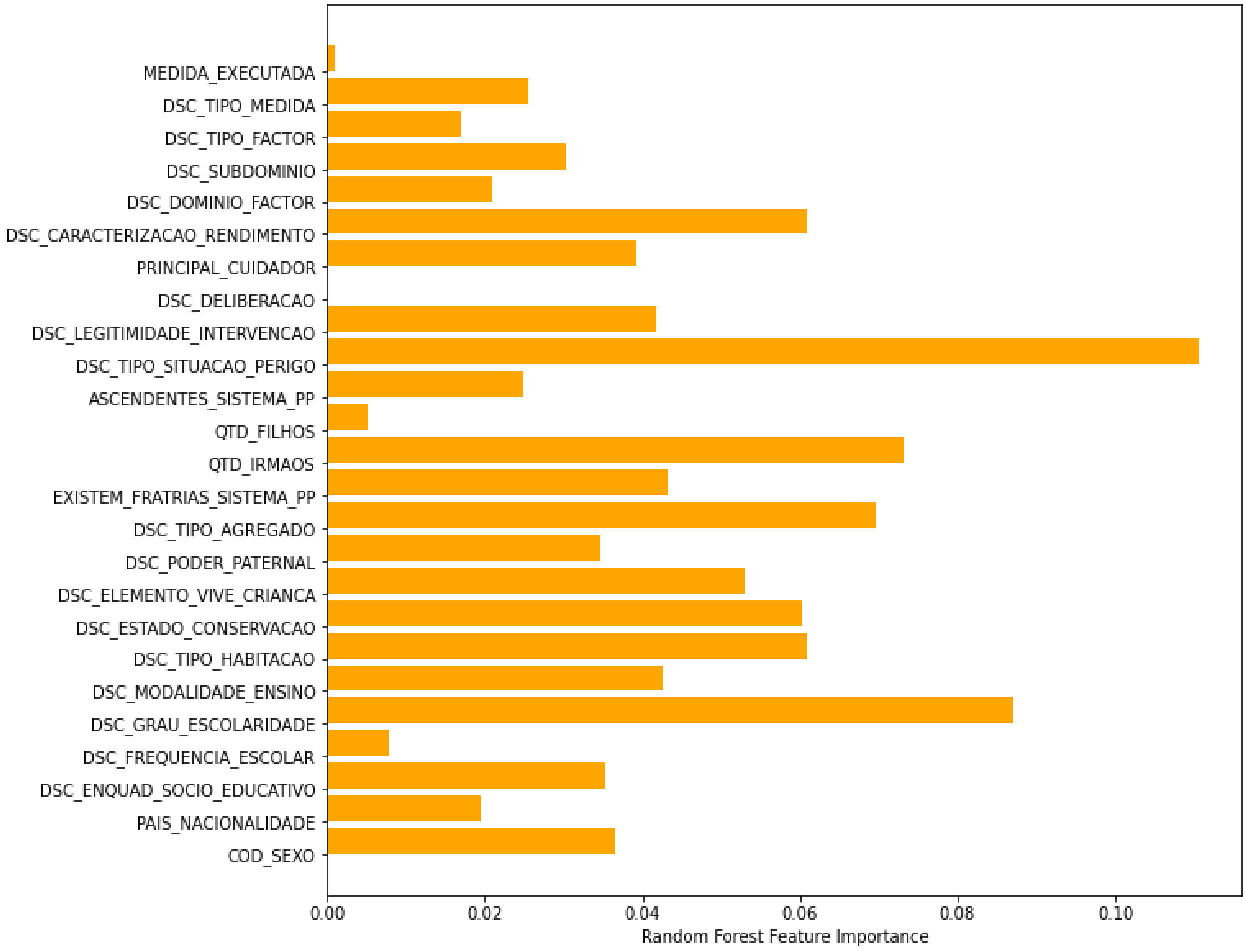

Figure 18 – Random Forest Feature Importance

Lastly, to enhance machine learning (ML) models through hyperparameter tuning, it is crucial to identify the key hyperparameters that need to be adjusted to tailor the ML models to specific problems or datasets (Yang & Shami, 2020).

Supervised learning algorithms encompass a set of ML algorithms that map input features to a target variable by training on labeled data. Some commonly used methods for hyperparameter tuning include:

1. Babysitting (also known as 'Trial and Error'): This is a basic hyperparameter tuning method widely employed by students and researchers. It involves manually tuning the hyperparameters by testing various values based on experience, intuition, or analysis of previously evaluated results.

2. Grid search (GS), Random search (RS), Bayesian Search (BS): These are simple hyperparameter optimization methods. However, they can be time-consuming and susceptible to the curse of dimensionality, making them less suitable for a large number of hyperparameters. Additionally, they may struggle to identify the global optimum for continuous parameters since they require a predefined, finite set of hyperparameter values.

To mitigate the steep learning curve associated with intuitively setting hyperparameters, this project utilized Bayesian Search. Bayesian Search explores every possible combination of hyperparameters within a specified set of values. It randomly generates combinations of parameters from a bounded domain and returns the best result after exploring a certain number of combinations (n).

Here is an overview of the optimal parameters used for each specified classification model:

Support Vector Machine (SVM) Classifier (SVC):
Best hyperparameters for SVC:
OrderedDict([('C', 2.256709025416922), ('kernel', 'linear')])

Gaussian Naive Bayes Classifier (GaussianNB):
No hyperparameters to be tuned.

K-Nearest Neighbors Classifier (KNeighborsClassifier):
Best hyperparameters for KNeighborsClassifier:
{'weights': 'uniform', 'n_neighbors': 7}

Logistic Regression Classifier (LogisticRegression):
Best hyperparameters for LogisticRegression:
OrderedDict([('C', 1.469245029096944), ('solver', 'newton-cg')])

Random Forest Classifier (RandomForestClassifier):
Best hyperparameters for RandomForestClassifier:
OrderedDict([('max_depth', None), ('n_estimators', 134)])

Decision Tree Classifier (DecisionTreeClassifier):
Best hyperparameters for DecisionTreeClassifier:
OrderedDict([('max_depth', 10), ('min_samples_split', 9)])

Subsequently, the models were evaluated using the optimal parameters, and the Random Forest classifier consistently demonstrated the highest accuracy, achieving a score of approximately 84.5%, as indicated by the results below.

1. SVC(C=2.256709025416922, kernel='linear')
    - TRAIN SCORE: 0.7069729336084476
    - TEST SCORE: 0.7023716814159292
2. GaussianNB()
    - TRAIN SCORE: 0.3016446170651778
    - TEST SCORE: 0.3058407079646018
3. KNeighborsClassifier(n_neighbors=7)
    - TRAIN SCORE: 0.811051098434276
    - TEST SCORE: 0.7411681415929203
4. LogisticRegression(C=1.469245029096944, solver='newton-cg')
    - TRAIN SCORE: 0.7081563296516568
    - TEST SCORE: 0.704070796460177
5. RandomForestClassifier(n_estimators=134)
    - TRAIN SCORE: 0.9890460007282437
    - TEST SCORE: 0.8445309734513274
6. DecisionTreeClassifier(max_depth=10, min_samples_split=9)
    - TRAIN SCORE: 0.7505158393008861
    - TEST SCORE: 0.7224070796460177

The evaluation of the classifiers provided valuable insights into their performance on the dataset.

Among the tested classifiers, the Random Forest classifier stood out with an impressive training score of 98.9% and a commendable testing score of 84.5%. This indicates that the Random Forest model effectively learned the training patterns and demonstrated strong generalization capabilities when applied to unseen data, showcasing its robustness.

On the contrary, GaussianNB exhibited the lowest training and testing scores among all the classifiers, with scores around 30.2%. This suggests that the GaussianNB model is not well-suited for the dataset, or the dataset may not adhere to the assumptions of the Naive Bayes algorithm.

KNN achieved a relatively high training score of 81.1% but had a slightly lower testing score of 74.1%. This discrepancy suggests a potential issue of overfitting, where the model memorizes the training data instead of effectively generalizing to new samples.

SVC and Logistic Regression produced similar training and testing scores, around 70.7% and 70.4% respectively. These models demonstrated moderate performance on the dataset, indicating reasonable generalization capabilities.

The Decision Tree classifier achieved a training score of 75.1% and a testing score of 72.2%. The relatively lower scores imply that the model may not capture the complexity of the dataset effectively.

Lastly, cross-validation was employed to obtain a more accurate estimate of the model's performance. This technique involves dividing the data into multiple subsets, training the model on different subsets, and averaging the performance scores to derive a more reliable estimate of the model's overall performance.

1. SVC(C=2.256709025416922, kernel='linear')
Mean cross-validation score: 0.7069729352912844

2. GaussianNB()
Mean cross-validation score: 0.30179634114186304

3. KNeighborsClassifier(n_neighbors=7)
Mean cross-validation score: 0.7363758574873356

4. LogisticRegression(C=1.469245029096944, solver='newton-cg')
Mean cross-validation score: 0.7084294439092437

5. RandomForestClassifier(n_estimators=134)
Mean cross-validation score: 0.8347798345566844

6. DecisionTreeClassifier(max_depth=10, min_samples_split=9)
Mean cross-validation score: 0.7262410135990089

In summary, the Random Forest classifier outperformed the other models with the highest mean cross-validation score of 83.48%. This indicates its superior ability to generalize well to unseen data. On the other hand, SVC, Logistic Regression, and KNN exhibited moderate performance, showcasing decent generalization capabilities. However, GaussianNB demonstrated poor performance, suggesting that it is not well-suited for the dataset. While the Decision Tree classifier showed acceptable performance, it may not capture the dataset's complexity as effectively as the other models.

5. CONCLUSION

5.1. SYNTHESIS OF THE DEVELOPED WORK

The objective of this project was to develop an effective methodology for accurately predicting children at risk of neglect or abuse. Hence, contributing to a more trusted Social Welfare system.

To achieve this goal, a variety of machine learning techniques were employed to create and transform features, aiming to optimize the performance of the predictive model. A dataset comprising 47,081 individual cases was collected from CNPCJ, spanning the years 2018 to 2022. The study focused on several categories defined by CNPCJ to evaluate dangerousness, including domestic violence, negligence, dangerous behaviour during childhood and youth, the right to education, physical maltreatment, psychological maltreatment, sexual abuse, abandonment, and child exploitation.

Subsequently, six algorithms were thoroughly tested. Among them, the Random Forest algorithm emerged as the most effective, exhibiting the highest performance in terms of the evaluated metrics and accuracy score. Additionally, simpler models such as LR, SVC, KNN, DT, and GNB were also considered.

Hyperparameter tuning was applied to the six algorithms using the Bayesian Optimization technique. This approach aimed to identify the optimal combination of hyperparameters that could further enhance the model's performance. Notably, the technique demonstrated success by improving the accuracy of the Random Forest algorithm from 84.4% to 84.5%. Moreover, the results of this study exceeded the baseline conditions, and additional aspects like Bayesian optimization and resampling techniques added value to the final work, offering valuable insights for future research.

Overall, this project successfully developed a methodology for predicting children in danger of neglect or abuse, utilizing various machine learning techniques, extensive data, and a comprehensive evaluation of algorithms and hyperparameters. The findings of this study contribute to the existing body of knowledge and present opportunities for further improvements and exploration in this critical domain.

5.2. LIMITATIONS AND RECOMMENDATIONS FOR FUTURE WORK

The present study highlights the potential for further improvement in predicting children in danger. Although the study is not exhaustive, there are various techniques that can be explored to enhance the accuracy of predictions.

One recommendation is to explore alternative machine learning techniques, particularly from a time-series perspective. By considering the temporal aspect of the data, it is possible to uncover patterns and trends that may contribute to more accurate predictions. Additionally, it would be worthwhile to investigate different configurations of fine-tuning that can not only enhance the results but also optimize the time efficiency of the process.

Furthermore, expanding the application of the model to other countries would be an intriguing avenue for exploration. Due to the time-limited collection of data from CNPCJ, it is essential to assess how the model performs in different contexts. This will enable a better understanding of its generalizability and the potential need for adjustments based on varying data quality standards in different regions.

While beyond the scope of the current project, another option worth considering is the development of a comprehensive dashboard to provide valuable insights. Such a dashboard could be customized to offer visualizations of individual and aggregated cases, as well as present the predictions generated by the model. Additionally, incorporating a monthly monitoring page into the dashboard would promote transparency among users and provide actionable insights to the development team for continuous improvement.

BIBLIOGRAPHICAL REFERENCES

garwal, A., Sharma, P., Alshehri, M., Mohamed, A. A., & Alfarraj, O. (2021). Classification model for accuracy and intrusion detection using machine learning approach. PeerJ Computer Science, 7, e437. https://doi.org/10.7717/peerj-cs.437

Ahn, M. J., & Chen, Y.-C. (2022). Digital transformation toward AI-augmented public administration: The perception of government employees and the willingness to use AI in government. Government Information Quarterly, 39(2), 101664. https://doi.org/10.1016/j.giq.2021.101664

Akinsola, J. E. T. (2017). Supervised Machine Learning Algorithms: Classification and Comparison. International Journal of Computer Trends and Technology (IJCTT), 48, 128–138. https://doi.org/10.14445/22312803/IJCTT-V48P126

Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. Journal of Physics: Conference Series, 1142, 012012. https://doi.org/10.1088/1742-6596/1142/1/012012

Azevedo, A., & Santos, M. F. (n.d.). KDD, SEMMA AND CRISP-DM: A PARALLEL OVERVIEW.

Bagging and Random Forest in Machine Learning: How do they work? (n.d.). Retrieved January 1, 2023, from https://www.knowledgehut.com/blog/data-science/bagging-and-random-forest-in-machine-learning

Children’s Data Network – Harnessing the scientific potential of linked, administrative data to inform children’s programs and policies. (n.d.). Retrieved January 1, 2023, from https://www.datanetwork.org/

Computational Social Welfare: Applying Data Science in Social Work. (n.d.). Obo. Retrieved January 1, 2023, from https://www.oxfordbibliographies.com/display/document/obo-9780195389678/obo-9780195389678-0286.xml

D’arcy-Bewick, S., Terracciano, A., Turiano, N., Sutin, A. R., Long, R., & O’Súilleabháin, P. S. (2022). Childhood abuse and neglect, and mortality risk in adulthood: A systematic review and meta-analysis. Child Abuse & Neglect, 134, 105922. https://doi.org/10.1016/j.chiabu.2022.105922

Electronic Information Systems and Social Work: Who are We Designing For?: Practice: Vol 26, No 5. (n.d.). Retrieved January 1, 2023, from https://www.tandfonline.com/doi/abs/10.1080/09503153.2014.958454?casa_token=2ZlUQG41o_wAAAAA:05EJBtzI8owWhuPtS7dPeAuPZKEtWpx-UdMbjNVb-oLZs_TAE2PTKGTPmNvZdEDr1H9zFKHPPEZq

Emily Jones, M. (n.d.). LibGuides: Creating a PRISMA flow diagram: PRISMA 2020. Retrieved January 15, 2023, from https://guides.lib.unc.edu/prisma/step-by-step

fusio. (n.d.). Every five minutes a child dies as a result of violence. UNICEF Ireland. Retrieved October 20, 2022, from https://www.unicef.ie/stories/every-five-minutes-a-child-dies-as-a-result-of-violence/

Gillingham, P. (2016). Predictive Risk Modelling to Prevent Child Maltreatment and Other Adverse Outcomes for Service Users: Inside the 'Black Box' of Machine Learning. British Journal of Social Work, 46(4), 1044–1058. https://doi.org/10.1093/bjsw/bcv031

Hamaker, E. L., Mulder, J. D., & van IJzendoorn, M. H. (2020). Description, prediction and causation: Methodological challenges of studying child and adolescent development. Developmental Cognitive Neuroscience, 46, 100867. https://doi.org/10.1016/j.dcn.2020.100867

Hillis, S. D., Mercy, J. A., & Saul, J. R. (2017). The enduring impact of violence against children. Psychology, Health & Medicine, 22(4), 393–405. https://doi.org/10.1080/13548506.2016.1153679

Holt, S., Buckley, H., & Whelan, S. (2008). The impact of exposure to domestic violence on children and young people: A review of the literature. Child Abuse & Neglect, 32(8), 797–810. https://doi.org/10.1016/j.chiabu.2008.02.004

Introduction—Predictive Analytics For Dummies [Book]. (n.d.). Retrieved December 30, 2022, from https://www.oreilly.com/library/view/predictive-analytics-for/9781118729410/xhtml/03_9781118728963-intro.xhtml

Kumar, V., & L., M. (2018). Predictive Analytics: A Review of Trends and Techniques. International Journal of Computer Applications, 182, 31–37. https://doi.org/10.5120/ijca2018917434

Learned-Miller, E. G. (n.d.). Introduction to Supervised Learning.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. Journal of Clinical Epidemiology, 62(10), e1–e34. https://doi.org/10.1016/j.jclinepi.2009.06.006

Long, C. K., Agrawal, R., Trung, H. Q., & Pham, H. V. (2021). A big data framework for E-Government in Industry 4.0. Open Computer Science, 11(1), 461–479. https://doi.org/10.1515/comp-2020-0191

Margetts, H. (2022). Rethinking AI for Good Governance. Daedalus, 151(2), 360–371. JSTOR.

Oswald, M. (2018). Algorithm-assisted decision-making in the public sector. Philosophical Transactions: Mathematical, Physical and Engineering Sciences, 376(2128), 1–20. JSTOR.

PEDSnet Clinical Research Network. (2020, April 30). PEDSnet Clinical Research Network | PCORI. https://www.pcori.org/research-results/2020/pedsnet-clinical-research-network

Penner, A. M., & Dodge, K. A. (2019). Using Administrative Data for Social Science and Policy. RSF: The Russell Sage Foundation Journal of the Social Sciences, 5(2), 1–18. JSTOR. https://doi.org/10.7758/rsf.2019.5.2.01

PhysioNet Databases. (n.d.). Retrieved January 1, 2023, from https://physionet.org/about/database/

Protective Factors Approaches in Child Welfare. (n.d.).

Relatórios de Atividade—Comissão Nacional de Promoção dos Direitos e Proteção das Crianças e Jovens. (n.d.). Retrieved December 25, 2022, from https://www.cnpdpcj.gov.pt/relatorio-atividades

Ren, Q., Cheng, H., & Han, H. (2017). Research on machine learning framework based on random forest algorithm. AIP Conference Proceedings, 1820(1), 080020. https://doi.org/10.1063/1.4977376

Sansone, D., & Zhu, A. (2021). Using Machine Learning to Create an Early Warning System for Welfare Recipients (arXiv:2011.12057). arXiv. https://doi.org/10.48550/arXiv.2011.12057

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. Procedia Computer Science, 181, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Strategy for the Rights of the Child (2022-2027)—Children’s Rights—Publi.coe.int. (n.d.). Children’s Rights. Retrieved January 15, 2023, from https://www.coe.int/en/web/children/strategy-for-the-rights-of-the-child

Tiwari, A. (2022). Chapter 2 - Supervised learning: From theory to applications. In R. Pandey, S. K. Khatri, N. kumar Singh, & P. Verma (Eds.), Artificial Intelligence and Machine Learning for EDGE Computing (pp. 23–32). Academic Press. https://doi.org/10.1016/B978-0-12-824054-0.00026-5

Vaithianathan, R., Maloney, T., Putnam-Hornstein, E., & Jiang, N. (2013). Children in the Public Benefit System at Risk of Maltreatment: Identification Via Predictive Modeling. American Journal of Preventive Medicine, 45(3), 354–359. https://doi.org/10.1016/j.amepre.2013.04.022

Vesoulis, Z. A., Husain, A. N., & Cole, F. S. (2022). Improving child health through Big Data and data science. Pediatric Research, 1–8. https://doi.org/10.1038/s41390-022-02264-9

Waldfogel, J. (2009). Prevention and the Child Protection System. The Future of Children, 19(2), 195–210.

Waterston, T. (2007). The Epidemic of Violence Against children. Journal of Tropical Pediatrics, 53(5), 295–297. https://doi.org/10.1093/tropej/fmm104

"Web 2.0 in Social Work Macro Practice: Ethical Considerations and Ques" by Katharine M. Hill and Sarah M. Ferguson. (n.d.). Retrieved January 1, 2023, from https://ir.stthomas.edu/ssw_pub/43/

Wesarg, C., Van den Akker, A. L., Oei, N. Y. L., Wiers, R. W., Staaks, J., Thayer, J. F., Williams, D. P., & Hoeve, M. (2022). Childhood adversity and vagal regulation: A systematic review and meta-analysis. Neuroscience & Biobehavioral Reviews, 143, 104920. https://doi.org/10.1016/j.neubiorev.2022.104920

Yang, L., & Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing, 415, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa