Personnel Review

Using HR Analytics to predict employee turnover in tech startups: an evidence-based HRM approach

Patrícia Nogueira
School of Economics, Management and Political Science, University of Minho, Braga, Portugal, and
Ana Carvalho
School of Economics, Management and Political Science, University of Minho, Braga, Portugal and
Centre for Research in Economics and Management (NIPE), University of Minho, Campus de Gualtar, Braga, Portugal

Received 25 September 2025
Revised 10 December 2025
24 February 2026
13 April 2026
Accepted 14 May 2026

Abstract

Purpose – This study seeks to test the effectiveness of Evidence-Based Human Resource Management (EBHRM) in tech startups. It compares two approaches to HR Analytics: the use of generic Human Resource Information Systems (HRIS) and bespoke statistical models built on firm-specific data to predict employee turnover.
Design/methodology/approach – A comparative case study of a small tech startup was conducted. Turnover predictions from the firm's HRIS were contrasted with custom-built logistic regression (LR) and decision tree models using data from 95 employees, assessing predictive accuracy and the actionability of insights.
Findings – The HRIS predictions were no better than random chance. The LR model was significantly more accurate. Both tailored models, using local evidence, identified specific, context-sensitive drivers of turnover that were entirely missed by the HRIS.
Research limitations/implications – Findings are based on a single case study with a limited dataset, and only two analytical techniques are tested.
Practical implications – Startup managers should critically evaluate generic HRIS analytics. The study suggests that cultivating an evidence-based mindset and training HR professionals to use simple statistical models on internal data is as important as investing in technology when employing HR Analytics to inform effective retention strategies.
Originality/value – This study advances research on EBHRM, HRIS and employee dynamics in startups by (1) clarifying that the "best available evidence" relies on contextual fit, transparency and interpretability; (2) proposing context-fit is a boundary condition for standardised HRIS analytics; and (3) pointing to the contextual idiosyncracies of tech startups as an explanation for why commercial HRIS may fail, as well as proposing transparency and simplicity as conditions for HR Analytics effectiveness in those organisations.

Keywords Evidence-based HRM, HR Analytics, Tech startups, Employee turnover, Predictive models, HRIS

Paper type Research article

Introduction

There is growing enthusiasm for HR Analytics due to its promise to enable more rigorous evidence-based decision-making and optimise the effectiveness of HR policies and initiatives, contributing to improve business outcomes (McCartney *et al.*, 2021; Bayo-Moriones *et al.*, 2025; Di Lauro *et al.*, 2025). Organisations can use the data they generate to identify and better understand HR problems and gain insights that allow them to make more informed decisions and design effective policies (Fernandez and Gallardo-Gallardo, 2021), overcoming guesswork

Funding: This work was supported by Fundação para a Ciência e a Tecnologia (award id: UID/03182).

Personnel Review
Emerald Publishing Limited
e-ISSN: 1758-6933
p-ISSN: 0048-3486
DOI 10.1108/PR-09-2025-1088

and misconceptions (Rousseau and Barends, 2011; Cavanagh *et al.*, 2024; Shet, 2025). This conceptually aligns with Evidence-Based Human Resource Management (EBHRM), which advocates critically considering the best available evidence from multiple sources to make informed decisions (Rousseau and Barends, 2011; Kroon, 2022; Maertens *et al.*, 2025). By providing reliable and valid data, HR Analytics should constitute a key enabler of EBHRM.

This promise is particularly salient in startup organisations, and especially in technological (tech) startups, which have distinctive characteristics. They are young ventures, typically five years old or younger (Haase and Eberl, 2019; Domurath *et al.*, 2023), corresponding to an early stage of development. This temporal nature means that a startup's processes, strategies and even its core identity are in a constant state of flux (Haase and Eberl, 2019; Domurath *et al.*, 2023; Tai *et al.*, 2025). Tech startups are fundamentally driven by (sometimes disruptive) innovation as they are created to develop and commercialise innovative, emergent technologies (Roach and Sauermann, 2024; Tai *et al.*, 2025). They are characterised by dynamic, fast-paced, and informal work cultures and informal structures with minimal hierarchy (Roach and Sauermann, 2024; Varma and Dutta, 2023) that attract highly skilled employees based on the opportunities for autonomy, task variety and skills development rather than high pay and job security (Domurath *et al.*, 2023; Roach and Sauermann, 2024). They often operate under challenging conditions, facing high uncertainty and severe resource constraints (Haase and Eberl, 2019; Roach and Sauermann, 2024; Tai *et al.*, 2025).

Within this volatile environment, employee turnover presents a uniquely severe threat to survival. In tech startups, where value is disproportionately tied to the specialised knowledge and innovative capacity of key personnel, the departure of even a single core team member can disrupt product development cycles and deter crucial investments (Tai *et al.*, 2025). Furthermore, the financial and operational costs of turnover – such as the loss of firm-specific human capital, the erosion of organisational memory and the difficulty of replacing highly skilled workers in an extremely competitive labour market – are magnified by the startup's limited resources (Domurath *et al.*, 2023). This, along with the critical role of key personnel at these early stages, makes employee turnover not only a significant operational and financial challenge but also an existential risk for startup firms (Domurath *et al.*, 2023; Tai *et al.*, 2025). Furthermore, professional sources indicate that turnover rates in startups could be as high as double the general US business average (Armitage, 2025) and are highest during the early stages of employment, with roughly half of employees leaving within the first three years (Steinfeld, 2024). The capacity to accurately predict and proactively manage employee attrition is, therefore, not merely beneficial but a crucial driver of sustainability for these organisations (Domurath *et al.*, 2023; Varma and Dutta, 2023; Tai *et al.*, 2025).

Predicting employee turnover is one of the most commonly used applications of HR Analytics. Research on turnover prediction models (e.g. Akasheh *et al.*, 2024; Živković *et al.*, 2024) and recent studies on retention strategies (e.g. Domurath *et al.*, 2023; De Vos *et al.*, *2025*; Bayo-Moriones *et al.*, 2025) emphasise the importance of data-driven approaches to managing turnover, substantiating the role of HR Analytics.

However, young and small companies often resort to commercial Human Resource Information Systems (HRIS), which offer packaged solutions that promise to embed analytical capabilities into the HR function (Wang, 2024). Their efficacy in the unique context of tech startups remains underexamined. Recent research suggests the one-size-fits-all nature of these systems fails to capture organisational idiosyncrasies. Even if they deliver outputs based on data, an acritical use of HRIS predictions may not correspond to using "the best available evidence", as espoused by EBHRM (Rousseau and Barends, 2011; Lange, 2013). In particular, Marler and Boudreau (2017, p. 20) stress that "'push button' intuitive HR Analytics via e-HRM technology" should be examined very carefully and critically to ensure they deliver valid information. Yet we still know little about whether standardised HRIS analytics generate valid and actionable turnover evidence in the distinctive context of tech startups.

Despite the theoretical promise of EBHRM, a significant research gap persists regarding its practical application (Maertens *et al.*, 2025), particularly in the high-velocity ecosystem of tech startups, as the existing literature predominantly focuses on HR Analytics within large, resource-rich corporations (McCartney *et al.*, 2021; Shet *et al.*, 2021). There is insufficient evidence of HR Analytics' impact that might compel resource-constrained startups to invest in the technology and analytics skills required to implement evidence-based practices (McCartney and Fu, 2022; Rafiq *et al.*, 2025), especially when confronted with the proliferation of HR technologies that promise, but may not deliver, valid insights (Angrave *et al.*, 2016; Marler and Boudreau, 2017).

We seek to address this by employing an EBHRM lens to assess different approaches to HR Analytics in startup firms. Drawing on the case of MiX – a small technology startup operating in a highly dynamic, competitive environment where retaining talent is paramount and the cost of turnover is particularly high – we compare the turnover predictions generated by MiX's commercial HRIS with custom-built logistic regression (LR) and decision tree (DT) models based on internal organisation data.

LR and DTs are well-established predictive models (Marín Díaz *et al.*, 2023; De Vos *et al.*, 2025) noted for their relative simplicity, high interpretability and robust performance, making them well-suited for resource-constrained startups (Guerranti and Dimitri, 2023; Marín Díaz *et al.*, 2023; Živković *et al.*, 2024). These methods allow organisations to draw on scientific knowledge and use context-relevant data (Kroon, 2022; Maertens *et al.*, 2025) to make not only accurate but also informative predictions. Moreover, statistical analyses such as LR are essential to detect discrepancies that may reflect biases in algorithms used in HRIS (Wang, 2024).

Our results illustrate the potential shortcomings of the "push button", one-size-fits-all systems and suggest that tailored statistical models, using "the right data" (Rasmussen and Ulrich, 2015), can deliver significantly more accurate and, above all, more informative predictions of employees at risk of leaving. Crucially, unlike the HRIS *black box* recommendations, these methods deliver interpretable results based on the best available evidence (Rousseau and Barends, 2011) to uncover the underlying causes of turnover (Allen *et al.*, 2010; De Vos *et al.*, 2025), thus providing startup organisations with actionable intelligence for targeted and effective retention interventions (Allen *et al.*, 2010). The paper thus makes three main contributions. Firstly, it advances EBHRM by clarifying what the "best available evidence" can entail, emphasising that the *quality of evidence* relies on contextual fit, transparency and interpretability, all of which can be more relevant than access to technology. Secondly, it contributes to the HRIS literature by showing that context-fit is a *boundary condition* for standardised HRIS analytics: when workforce dynamics differ from the reference assumptions embedded in generic HRIS, predictive validity and actionability deteriorate. Thirdly, it advances research on employee dynamics in startups by highlighting that turnover in tech startups is structurally different from generic benchmarks, thus explaining *why* commercial HRIS may fail and organisation-specific analytics are needed to explain and manage it effectively. In addition, considering the typical challenges faced by startups, *transparency* and *simplicity* are proposed as conditions for HR Analytics effectiveness in those organisations.

EBHRM and HR Analytics

Based on the principles of Evidence-Based Management, EBHRM advocates for critically considering the best available evidence from multiple sources to make informed decisions (Rousseau and Barends, 2011; Kroon, 2022; Maertens *et al.*, 2025). This involves identifying relevant HR problems, gathering and critically assessing the quality of evidence from various sources, using insights derived from its analysis to formulate courses of action and evaluating the outcomes of those HR interventions (Rousseau and Barends, 2011; Kroon, 2022). Sources of evidence include scientific research, internal organisational data, practitioners' expertise and judgment, and stakeholders' concerns and expectations (Rousseau and Barends, 2011; Maertens *et al.*, 2025; Negt and Haunschild, 2025). Sophisticated data analyses relevant to the organisational context (Rousseau and Barends, 2011; Maertens *et al.*, 2025) are essential to uncover valid cause-and-effect relationships that have practical relevance (Allen *et al.*, 2010; Lange, 2013; Negt and Haunschild, 2025).

Although these principles are well supported, the implementation and conceptual underpinnings of EBHRM are subject to significant discussion. A persistent debate revolves around the gap between research and practice, where communication barriers, methodological issues and the accessibility of research findings make it hard to convince practitioners to base their practice on existing research evidence (Negt and Haunschild, 2025). The traditional conceptualisation of EBHRM itself may be insufficient for guiding practice, leading Maertens *et al.* (2025) to propose adding a "how-to" (required methods and skills) and a "why-to" (why it is relevant) dimensions that help clarify the practical application of EBHRM.

HR Analytics, in turn, involves applying statistical techniques to HR-related data to support effective decision-making (Marler and Boudreau, 2017; Fernandez and Gallardo-Gallardo, 2021). As such, it is a critical enabler of EBHRM, providing the evidence required for informed decision-making. It is a source of internal organisational data, uses advanced statistical analyses to produce meaningful insights (Rasmussen and Ulrich, 2015; Falletta and Combs, 2021; McCartney and Fu, 2022) and has the potential to demonstrate cause-and-effect relationships between HR interventions and business results (Angrave *et al.*, 2016; Falletta and Combs, 2021). HR Analytics can further contribute to institutionalise the EBHRM design of expanding evidence gathering and incorporating evidence into decision-making (Rousseau and Barends, 2011).

However, the availability of data and statistics, or indeed technology and software, does not automatically translate into useful information (Falletta and Combs, 2021; Fernandez and Gallardo-Gallardo, 2021). HR Analytics requires an EBHRM mindset to support strategically informed decisions (Cavanagh *et al.*, 2024; Negt and Haunschild, 2025; Wang, 2024) and avoid being "a mere fad" (Rasmussen and Ulrich, 2015).

HR Analytics and HRIS in tech startups

Although the adoption of HR Analytics by organisations is growing (Bayo-Moriones *et al.*, 2025), studies indicate that the pace of adoption has been slower than anticipated (Fernandez and Gallardo-Gallardo, 2021; Shet *et al.*, 2021). Many organisations, including large multinationals, struggle to fully integrate analytics into their HR functions (Shet *et al.*, 2021). Most organisations predominantly use only descriptive analytics for operational reporting (McCartney *et al.*, 2021; Di Lauro *et al.*, 2025), and the large amount of data collected by HR departments is not analysed (Rasmussen and Ulrich, 2015; McCartney and Fu, 2022; Cavanagh *et al.*, 2024; Rafiq *et al.*, 2025). Widespread strategic implementation remains limited (Shet *et al.*, 2021; Di Lauro *et al.*, 2025).

Numerous barriers hinder HR Analytics implementation in startup firms that typically operate under significant financial and human resource constraints (Haase and Eberl, 2019; Roach and Sauermann, 2024). A first key barrier is a critical skills gap. Startups are rarely able to hire dedicated data analysts (Varma and Dutta, 2023; Rafiq *et al.*, 2025), and many HR professionals lack the analytical, statistical and data-science competences required to derive strategic insights (Behl *et al.*, 2019; Haase and Eberl, 2019; McCartney *et al.*, 2021; Shet *et al.*, 2021). Lack of financial resources also curtails investments in sophisticated analytics software and infrastructure (Rafiq *et al.*, 2025), with some startups using third-party or shared IT infrastructure (Behl *et al.*, 2019). Organisational factors, such as insufficient top management support and lack of awareness of the potential benefits of HR Analytics and its impact on firm performance, further impede progress (Shet *et al.*, 2021; Di Lauro *et al.*, 2025; Rafiq *et al.*, 2025). Furthermore, the informality typical of startups (Varma and Dutta, 2023) also holds back the routinisation of systematic processes for data collection (Haase and Eberl, 2019). The lack of analytics solutions specifically tailored for startups further exacerbates these challenges (Rafiq *et al.*, 2025), preventing them from fully capitalising on data-driven insights.

Given these limitations, employing commercial HRIS packages can bring major benefits to the HR performance of startups (Varma and Dutta, 2023), including the ability to perform HR Analytics (McCartney and Fu, 2022; Wang, 2024).

HRIS are sophisticated IT systems designed for HR professionals to manage HR data. They automate repetitive tasks, freeing HR staff from administrative workload and generally improving the efficiency of HR functions (Angrave *et al.*, 2016; Jemine and Guillaume, 2022; Wang, 2024). Their primary contribution to HR Analytics lies in their capacity to gather, store and organise HR-related data (Angrave *et al.*, 2016; McCartney and Fu, 2022). By centralising employee information, HRIS can facilitate the extraction of data necessary for various analytical activities, from basic reporting to more complex analyses (McCartney *et al.*, 2021).

However, while HRIS are effective for operational reporting and *descriptive* analysis (i.e. analyses that reveal and describe current and historical data patterns and relationships) (De Vos *et al.*, 2025), many are not inherently designed to support more advanced analytics that link HR metrics to broader business outcomes (Angrave *et al.*, 2016; Fernandez and Gallardo-Gallardo, 2021) by enabling *predictive* (predictions about the future based on the current and historical data) or *prescriptive* modelling (analyses that offer guidance on recommended actions to optimise desired outcomes) (De Vos *et al.*, 2025).

The evolution of HRIS towards more sophisticated and cloud-based solutions has started to incorporate analytical capabilities (Angrave *et al.*, 2016). These newer systems aim to provide more user-friendly interfaces, dashboards and some level of analytical tools, providing insights into the workforce and supporting better decision-making (McCartney and Fu, 2022). However, the more recent and sophisticated HRIS, often supported by Artificial Intelligence (AI), raises new concerns. For one, algorithms used by HRIS can lack transparency, making it difficult for managers – let alone employees – to understand the outputs of analyses (De Vos *et al.*, 2025). Indeed, a comprehensive study of 18 HRIS vendors found that they are not particularly forthcoming about their practices nor effective at preventing algorithm bias (Raghavan *et al.*, 2020). Moreover, trained on historical data, these algorithms may reproduce pre-existing patterns and assumptions, thereby perpetuating historical biases (Wang, 2024).

Besides, the focus often remains on the technological implementation rather than on fostering the strategic and analytical mindset required to fully leverage the data these systems hold (Angrave *et al.*, 2016; Marler and Boudreau, 2017).

A key issue relates to a potential mismatch between HRIS predictions and the specific organisational context that may undermine the relevance of analyses (Rousseau and Barends, 2011; Maertens *et al.*, 2025). Angrave *et al.* (2016) point out that the professionals developing HRIS are not acquainted with the specific context of each organisation, while HR professionals lack the knowledge to adapt the software's standard models to their particular situation. Resource-strapped and less structured startups often rely on the support of external vendors (Behl *et al.*, 2019) and acquire standardised products with limited customisation (Jemine and Guillaume, 2022). In fact, Jemine and Guillaume (2022) allude to a "technological isomorphism" created by HRIS vendors and consultants exerting pressure towards standardisation based on what they consider to be best practices and their prior experience with work processes, promoting a technological uniformity among their clients. In those cases, HRIS can be an impediment rather than an enabler of effective HR Analytics (Marler and Boudreau, 2017; McCartney and Fu, 2022). This can be particularly problematic when predicting employee turnover, when we consider that it is highly susceptible to contextual and social influences, including peer relations, climate and leadership (Rubenstein *et al.*, 2018; Amarakoon and Colley, 2023; Tai *et al.*, 2025), which are idiosyncratic to each organisation and less likely to be captured by general benchmarks. It is therefore essential that turnover predictions be based on organisation-specific evidence if they are to produce valid and informative insights for effective decision-making.

Predicting employee turnover in tech startups

Voluntary employee turnover is a critical concern in any organisation due to its substantial impact on organisational functioning, performance and survival (Allen *et al.*, 2010; Hancock *et al.*, 2013; Rubenstein *et al.*, 2018). It carries extensive financial costs involved in re-hiring and training new employees (Allen *et al.*, 2010; Hancock *et al.*, 2013; Rubenstein *et al.*, 2018) as well as indirect costs associated with the loss of valuable firm-specific human capital (Rubenstein *et al.*, 2018), the erosion of social capital and loss of organisational memory (Allen *et al.*, 2010; Hancock *et al.*, 2013). Turnover also disrupts work processes, diminishes quality and safety and can potentially reduce customer satisfaction (Hancock *et al.*, 2013), affecting overall organisational performance (Allen *et al.*, 2010; Hancock *et al.*, 2013).

These challenges can be exacerbated in high-tech startups, where the "war for talent" is especially fierce (Marín Díaz *et al.*, 2023; Varma and Dutta, 2023; Tai *et al.*, 2025). While attracting and retaining highly skilled workers is crucial for the development and growth of knowledge-intensive startups (Roach and Sauermann, 2024; Tai *et al.*, 2025), these also pose a significant challenge as startups often experience higher levels of turnover (Amarakoon and Colley, 2023; Domurath *et al.*, 2023).

Startups' dynamic environment – characterised by high levels of uncertainty, rapid change, and evolving organisational structures and identities typical of young ventures (Haase and Eberl, 2019; Domurath *et al.*, 2023; Tai *et al.*, 2025) – coupled with often underdeveloped HR practices (Domurath *et al.*, 2023; Tai *et al.*, 2025), can lead to a misalignment between employees' initial expectations and their actual work experiences over time and subsequent higher turnover intentions (Domurath *et al.*, 2023). Startups also often operate with limited financial and human resources, which means they may not be able to offer the same level of pay and job stability as more established firms (Roach and Sauermann, 2024; Tai *et al.*, 2025). And the impact of losing key personnel will be felt more acutely (Domurath *et al.*, 2023; Tai *et al.*, 2025).

It is therefore of the utmost interest of startup organisations to predict and manage turnover proactively (Varma and Dutta, 2023) in order to mitigate the substantial direct and indirect costs associated with turnover (Allen *et al.*, 2010; Hancock *et al.*, 2013; Marín Díaz *et al.*, 2023). Evidence-based predictive analytics enable startups to identify at-risk employees and intervene before they decide to leave by addressing the underlying causes of employee dissatisfaction and implementing targeted retention strategies, such as improving the work environment, offering professional development or adjusting compensation (Amarakoon and Colley, 2023; Tai *et al.*, 2025).

Organisations can predict turnover intentions using various analytical approaches, predominantly employing machine learning and statistical modelling techniques on HR data (Lazzari *et al.*, 2022; Akasheh *et al.*, 2024; De Vos *et al.*, 2025). Common methods include LR, DTs, random forests, support vector machines and neural networks (Akasheh *et al.*, 2024; De Vos *et al.*, 2025), which analyse a wide array of organisation-specific data such as employee demographics, work-specific factors (e.g. seniority, pay), job satisfaction surveys and performance metrics (Rombaut and Guerry, 2018; Lazzari *et al.*, 2022; Živković *et al.*, 2024). These models identify patterns and correlations within the data to forecast the likelihood of an employee leaving (Edwards and Edwards, 2016; De Vos *et al.*, 2025), allowing HR professionals to understand the underlying mechanisms behind an employee's intention to leave and take preventative actions (Marín Díaz *et al.*, 2023; De Vos *et al.*, 2025).

Methodology

In order to assess different approaches to HR Analytics in tech startups within an EBHRM framework, we specify two testable propositions.

P1. Given the fast-evolving nature of tech startups, analytical models using local evidence will produce significantly more accurate predictions of employee turnover than an HRIS model based on generic benchmarks.

P2. Custom-built predictive models using local evidence will provide more interpretable insights regarding the factors influencing employee turnover in startups, facilitating managerial decision-making, compared to the outputs of a generic HRIS.

We therefore compare HRIS-generated predictions with custom-built predictive models in the case of MiX, a critical case (Yin, 2018), representing an ideal setting in which to test these propositions. MiX was selected based on three criteria: (1) it is a technology startup operating in a talent-intensive environment; (2) it was experiencing a high level of undesired turnover, with significant organisational side effects (e.g. knowledge loss and team disruption); and (3) it had implemented an HRIS with embedded HR Analytics functionalities. The combination of these conditions made it particularly suitable for examining predictive models, as detailed ahead. Consistent with critical case logic, this study can be useful to test our propositions derived from exploration research (Dul and Hak, 2008).

To illustrate custom-built predictive models based on organisational evidence, we employ LR and DTs, which are widely studied methods (De Vos *et al.*, 2025) and stand out as having advantages particularly relevant in the context of the challenges faced by startups. LR is a common and effective estimation technique for predicting employee turnover (Edwards and Edwards, 2016; Rombaut and Guerry, 2018; De Vos *et al.*, 2025). It is flexible, easy to use and provides meaningful estimates of model parameters, making it a well-known technique in many applications (Marín Díaz *et al.*, 2023; Akasheh *et al.*, 2024). In addition, LR is the kind of statistical analysis especially highlighted as fundamental in the detection of algorithm biases and enabling the necessary corrections (Wang, 2024).

DTs also offer distinct advantages. They are noted for their visual and clear representation of decisions and decision-making processes, and for capturing nonlinear relationships (Kang *et al.*, 2021; Živković *et al.*, 2024). One of their key strengths is the explainability of their results, alongside ease of use, computational speed and robustness to outliers (Guerranti and Dimitri, 2023; Akasheh *et al.*, 2024). Furthermore, DTs can help identify specific employee groups at risk for turnover, providing actionable insights for HR departments (Rombaut and Guerry, 2018; Kang *et al.*, 2021).

The case

Founded in 2019, MiX (a pseudonym) is a technology startup dedicated to developing innovative digital solutions, specialising in AI and B2B sales. Since its inception, the company has grown remarkably, amassing more than 450 clients in various global markets and consolidating its position as a leading name in the sector. In the last 2 years, it has tripled its revenues and number of employees and currently has 85 telecommuting workers, about half of whom are software engineers.

Nevertheless, MiX has recently faced significant challenges related to high voluntary employee turnover, which stands at nearly a quarter (24.18%) over the last two years, with an impact on team management, project deadlines and product quality, thus jeopardising the company's competitive advantage in a highly demanding market.

MiX uses a global market-leading HRIS system, developed by an external vendor, for various HRM-related functions, including storing and maintaining employees' personal and professional information and monitoring the organisation's turnover by analysing this data. The HRIS produces a qualitative output that associates different levels of risk of leaving with each employee – low, medium and high – based on voluntary turnover indicators. For example, one employee might be assigned 4 low-risk, 2 medium-risk and 5 high-risk factors, while another may be assigned 2 low-risk, 3 medium-risk and 2 high-risk factors. This quantification is not accompanied by the identification of specific factors, therefore not providing managers with a clear indication of what measures might prevent employees from leaving. Besides this equivocal output, the system's attrition risk forecasts also lack methodological transparency, as the underlying calculation method that delivers those results is not disclosed. Additionally, the benchmarking data is based on a study by the consulting firm Mercer focused on the financial services sector, raising concerns about its relevance to other industries.

This HRIS, therefore, constitutes a *black box* regarding employee turnover prediction (Marín Díaz *et al.*, 2023; De Vos *et al.*, 2025). Besides ignoring organisation-specific evidence, it does not explore the underlying causes of turnover, concealing cause–effect relationships that might inform retention strategies.

Sample and data

A number of variables were collected from the company's HRIS system. Apart from the description of the variable, Table 1 indicates which variables are used in each method under analysis: HRIS predictions, LR and DT. The variables used in the HRIS predictions are based on assumptions derived from the study by Mercer. For the custom-built analyses, an additional set of common turnover predictors identified in the scientific literature was included, as signalled in the table.

The dataset initially included records for 116 employees. However, 22 records were excluded for the following reasons: employees had not yet started their role at the time of analysis, had been with the company for less than one year, or had left involuntarily. The final sample consisted of 95 employees aged between 23 and 61 years (mean age = 32 ± 6 years), of whom 63% were male and 37% female. All procedures adhered to the company's anonymisation protocols to ensure data privacy and confidentiality.

Procedure

Human Resource Information Systems. The HRIS system considers the variables and assumptions indicated in Table 1 to deliver a combination of low-, medium- and high-risk factors for each employee. In the absence of a clearer indication, the number of high-risk factors attributed to each employee was considered the closest predictor of turnover. Its accuracy was tested using the Receiver Operating Characteristics (ROC) curve, interpreted through the Area Under the ROC Curve (AUC) (Lazzari *et al.*, 2022; De Vos *et al.*, 2025). This AUC is a comprehensive threshold-independent method commonly used to measure the performance of machine learning algorithms (see De Vos *et al.*, 2025 for a detailed explanation).

Logistic regression. LR is an established statistical method that estimates the probability of binary "response" variables (in our case, "leaves" or "stays") equalling 1 in the presence of a set of predictors (Edwards and Edwards, 2016). As such, it calculates the probability of an employee leaving based on individual and organisational factors.

A binary LR analysis was conducted on the IBM® SPSS® Statistics (v23) software using the variables displayed in Table 1. The results indicate the probability of employees voluntarily leaving the organisation based on a set of independent variables, including age, gender, satisfaction survey score, time since last raise and tenure (see Table 1). A 5% level of significance was considered.

Statistical tests were performed to assess the fit and quality of the model, namely Nagelkerke's R^2 and the ROC curve, interpreted through the AUC (De Vos *et al.*, 2025).

Decision tree. DTs classify data into classes (in our case, "leaves" or "stays") by combining a sequence of simple tests that compare attributes against a threshold value or a set of possible values (Živković *et al.*, 2024).

In this study, a DT was implemented using Python in Spyder (v6.0.1) with the Scikit-learn library. In contrast to the statistical tables produced by SPSS, this approach has the advantage of providing a visual outcome to capture complex interactions between variables, enhancing interpretability (Kang *et al.*, 2021; Akasheh *et al.*, 2024). The widely used Classification and Regression Tree (CART) algorithm was employed. To build a model, the CART algorithm examines all potential predictors, selects the one with the strongest relationship to the target outcome and uses it to recursively partition the dataset (Kang *et al.*, 2021). The division criterion is based on the Gini index to split the data into the two classes. The Gini index measures the impurity of a node in the DT and varies between 0 and 1, where a value close to

Table 1. Variables used in the three models

Variable name	Description	Expected impact on turnover	HRIS	Logistic regression	Decision tree
Age	Employee's age in years	Employees aged 35 or older are less likely to leave the company (HRIS assumption)	✓	✓	✓
Dependents	Indicates whether the employee has dependents (children, elderly)	Employees with dependents tend to have lower turnover risk (HRIS assumption)	✓	✓	✓
Peer Role Count	Number of colleagues sharing the same role	Employees with fewer than or equal to four peers in the same role show reduced turnover probability (HRIS assumption)	✓	X	✓
Managerial Responsibility	Whether the employee manages other employees	Managers are generally less likely to leave (HRIS assumption)	✓	✓	✓
Team Size Managed	Number of direct reports managed by the employee	Employees managing larger teams tend to have lower turnover rates (HRIS assumption)	✓	X	✓
Tenure in Current Role	Time (in years) in the current position	Longer tenure (>4 years) in the current role is associated with higher turnover risk (HRIS assumption)	✓	✓	✓
Time Since Last Raise	Time (in years) since the last salary increase	Employees without raises for over 2.5 years are more likely to leave (HRIS assumption)	✓	✓	✓
Gender	Employee's gender	Female employees tend to have lower turnover rates than males (Rombaut and Guerry, 2018; Rubenstein *et al.*, 2018)	X	✓	✓
Location	Geographic location of the employee	Turnover rates vary across countries (Hancock *et al.*, 2013; Lazzari *et al.*, 2022)	X	✓	✓
Tenure	Total time (in years) working for the company	Longer overall tenure is linked to lower turnover risk (Allen *et al.*, 2010; Guerranti and Dimitri, 2023; Rubenstein *et al.*, 2018)	X	✓	✓
Survey Score	Average satisfaction score by department (1–5 scale)	Higher satisfaction and engagement scores correlate with lower turnover (Allen *et al.*, 2010; Guerranti and Dimitri, 2023; Lazzari *et al.*, 2022; Rubenstein *et al.*, 2018; Živković *et al.*, 2024)	X	✓	✓
Leaver status	Dependent class variable, indicating the actual employee status 1 = left the company; 0 = stayed in the company		✓	✓	✓
HRIS Risk prediction	Compound variable delivered by the HRIS related to employee turnover risk - Low-risk: number of factors considered low-risk - Medium-risk: number of factors considered medium-risk - High-risk: number of factors considered high-risk		✓	X	X
Predictive Turnover	Outcome variable LR: Estimated probability of leaving (%) DT: binary variable (1 = leaves; 0 = stays)		X	✓	X

Source(s): Authors' own work, based on the references contained in the table

0 is ideal, indicating a pure node (all samples belong to the same class) (see Kang *et al.*, 2021 for a detailed explanation).

To prevent overfitting and ensure the model's generalisability, the dataset was randomly split into two subsets: one for training and another for testing purposes. The model was trained on the first subset and then validated on the second, which had not been used during training. Test scores for precision, recall and F1-score were obtained.

Results

Human Resource Information Systems

Turnover predictions delivered by the HRIS are based on the High-risk variable, corresponding to the number of high-risk factors assigned to each employee. The AUC for these predictions assessed against employees' actual turnover status (LeaverStatus) delivered a non-statistically significant value just above 0.5 (AUC = 0.588, SE = 0.061; $p = 0.149$), with a confidence interval between 0.469 and 0.707. The non-significant result indicates the model does not perform better than a random assumption (Rombaut and Guerry, 2018), which means it does not discriminate well between workers who will or will not leave the company.

Logistic regression

Table 2 has the results of the LR, showing that the variables that significantly influence the probability of turnover are "Gender", "Survey Score" on job satisfaction and "Tenure".

For the "Survey Score" on job satisfaction, the coefficient was $B = -1.51$ ($p = 0.02$), with odds ratio (exp(B)) = 0.22. So, greater job satisfaction reduces the odds of turnover by around 78%. This is largely supported by previous studies (e.g. Allen *et al.*, 2010; Marín Díaz *et al.*, 2023; Živković *et al.*, 2024).

Considering "Tenure", the coefficient was $B = -1.22$ ($p = 0.01$), with odds ratio (exp (B)) = 0.30, indicating that for each additional year of service, the odds of turnover are reduced by around 70%. This result is also in line with previous studies (e.g. Allen *et al.*, 2010; Rubenstein *et al.*, 2018; Kang *et al.*, 2021).

As for "Gender", the coefficient was $B = -1.55$ ($p = 0.02$), with odds ratio (exp (B)) = 0.21, showing that the odds of turnover are approximately 79% lower for male (0) than for female (1) employees. This contrasts with other studies on general populations (e.g. Rombaut and Guerry, 2018; Rubenstein *et al.*, 2018), which observed that women tend to exhibit a lower probability of turnover. In the technological sector, however, attrition tends to be significantly higher for women than for men, which is attributed to a predominant masculine culture (Accenture, 2020). This was not captured by the HRIS, which, in this case, considers a benchmark based on the financial industry, which further reinforces the importance of using internal data to attend to the specificities of each organisation.

The model coefficient tests indicate that the model is statistically significant ($\chi^2(9) = 54.74$, $p = 0.000$) and has a moderate ability to explain employee turnover (Nagelkerke's $R^2 = 0.61$). This is in line with real-world organisational contexts, where phenomena such as turnover can be influenced by multiple internal and external factors, some of which may not have been captured by the variables in the model.

Table 2. Results of the logistic regression

Variables	β	Exp (β)	p-value	SE	95% CI
Constant	5.15	171.92	0.13	3.35	Lower
Survey Score	−1.51	0.22	*0.02*	0.66	[0.06–0.81]
Tenure	−1.22	0.30	*0.01*	0.47	[0.12–0.74]
Gender	−1.55	0.21	*0.02*	0.65	[0.06–0.75
Age	0.09	1.10	0.10	0.06	[0.98–1.22]
Dependents	0.54	1.71	0.18	0.40	[0.78–3.75]
Managerial Responsibility	0.46	1.58	0.60	0.89	[0.28–8.98]
Tenure in Current Role	−1.16	0.31	0.051	0.59	[0.10–1.00]
Time Since Last Raise	0.9	2.47	0.33	0.93	[0.40–15.24]
Location	1.05	2.85	0.18	0.79	[0.61–13.28]

Note(s): χ^2 (9) = 54.74; $p = 0.00$; Nagelkerke's $R^2 = 0.61$

Source(s): Authors' own work

To further ensure the robustness of the analysis, we performed the inference using 5,000 bootstraps (Table 3). In this case, "Tenure in Current Role", which was previously borderline ($p = 0.051$), also became statistically significant ($p = 0.03$).

The model's predictive performance when compared to workers' actual turnover status indicates an overall success rate of 80.60%. Specifically, the percentage of hits for workers who remained in the organisation (LeaverStatus = 0) was 86.20%, while for those who left (LeaverStatus = 1), the hit rate was 71.40%. Thus, the model has a good and balanced performance in classifying both categories, contributing to its usefulness in predicting turnover in the organisational context. The AUC for this result delivered a value greater than 0.5 (AUC = 0.899, $p = 0.000$), indicating the model performs better than a random assumption, with a good ability to discriminate between workers who will or will not leave the company (Rombaut and Guerry, 2018; De Vos *et al.*, 2025). The confidence interval was between 0.839 and 0.959, which reinforces the reliability and accuracy of the results.

Decision tree

Figure 1 shows the results of the DT. Calculation based on the CART algorithm indicated only three variables are statistically significant in this model: "Tenure in Current Role", "Age" and "Time Since Last Raise". The model showed a total accuracy of 65.52%. However, the

Table 3. Inference results of the logistic regression with 5,000 bootstraps

Variables	β	Bootstrap bias	p-value	SE	95% CI
Constant	5.15	5.06	0.14	183.69	[−5.86 to 21.60]
Survey Score	−1.51	−2.46	0.02	50.31	[−4.99 to −0.13]
Tenure	−1.22	−2.90	0.01	65.67	[−4.44 to 0.35]
Gender	−1.55	−3.60	0.03	95.60	[−5.27 to −0.09]
Age	0.09	0.20	0.14	5.52	[−0.02 to 0.46]
Dependents	0.54	0.96	0.29	41.61	[−1.09 to 2.62]
Managerial Responsibility	0.46	0.11	0.65	46.15	[−2.84 to 3.70]
Tenure in Current Role	−1.16	−2.06	0.03	48.38	[−3.78 to −0.17]
Time Since Last Raise	0.90	4.58	0.39	143.79	[−1.61 to 5.61]
Location	1.05	2.53	0.31	108.05	[−1.26 to 5.71]

Note(s): χ^2 (9) = 54.74; $p = 0.00$; Nagelkerke's $R^2 = 0.61$
Source(s): Authors' own work

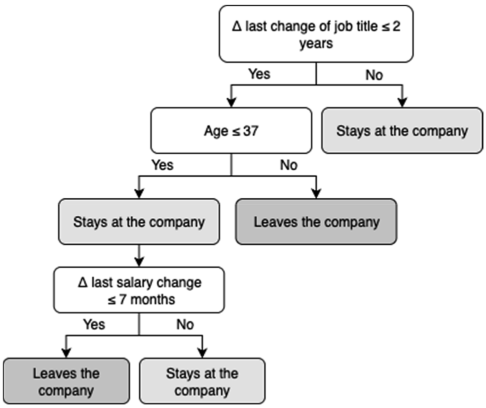

Figure 1. Decision Tree results. Source: Authors own work

evaluation metrics revealed greater effectiveness in predicting the workers who remained in the organisation compared to those who left the organisation (see Table 4). The identical values for precision and recall across both classes suggest that the model maintains a consistent trade-off between identifying true positives and avoiding false positives. These low classification results for the “Leaves” class are explained by the class imbalance that results from there being a much higher number of workers who stayed than of those who left (Kang *et al.*, 2021), causing a bias towards the majority class (in this case, “Stays”). This is frequently reported in real-world classification problems and is hard to overcome by common rebalancing techniques (De Vos *et al.*, 2025).

The variable located at the root node of the tree – “Tenure in Current Role” (number of years since the last change of professional position) – is the most important predictive factor. In other words, workers whose position has not changed in more than two years are more likely to stay with the organisation (“Stays” class). This result corroborates previous studies (Allen *et al.*, 2010; Rubenstein *et al.*, 2018).

“Age” was the second most relevant variable, with a cut-off at 36.5 years. Workers aged 36.5 or less were more likely to stay, while older workers were more prone to leaving. This contradicts the HRIS assumption that older workers are more likely to stay. One advantage of DT is that it considers predictors in conjunction. In this case, it tells us that workers who have been in their job for less than two years and are over 36.5 years old are more likely to leave the organisation.

The number of years since the last salary change (“Time Since Last Raise”) was the last determining factor in the model, with a threshold of 0.542 years equivalent to approximately 6.5 months. Salary changes made less than 6.5 months ago were associated with a higher risk of turnover. This is a particularly relevant variable when combined with age and time in the job. When the worker has been in the job for less than 2 years, is less than 36.5 years old and has had their last salary change more than 6.5 months ago, the likelihood of staying increases. Although studies mention the importance of salary increases and a higher salary (e.g. Rombaut and Guerry, 2018; Marín Díaz *et al.*, 2023; Tai *et al.*, 2025), workers at MiX may value career development and having a challenging job more than financial compensation, as this allows them to work with cutting-edge technology and expand their expertise. This is typical of tech startups (Domurath *et al.*, 2023; Roach and Sauermann, 2024) and is yet another idiosyncratic feature of this context that was not captured by the HRIS. On the other hand, for workers under 36.5 years old and in the same position for less than 2 years, the salary increases may have come too late in such a competitive environment as the software sector.

Discussion

These results partially confirm our first proposition, which states that, given the fast-evolving nature of tech startups, using local organisational evidence would demonstrate higher predictive accuracy than a generic HRIS. Indeed, the results suggest that the HRIS system may not be accurately capturing the risk factors that led to departures from this company. As for the custom-built analyses, the LR model in particular provided considerably more accurate estimates. However, the DT showed low accuracy in predicting the “Leaves” class, owing to

Table 4. Test scores for the Decision Tree

Variables	Precision	Recall	F1-score
0 (“Stays” Class)	0.77	0.77	0.77
1 (“Leaves” Class)	0.33	0.33	0.33

Note(s): Overall accuracy: 65.52%
Source(s): Authors’ own work

the class imbalance typical of turnover data (Kang *et al.*, 2021; Lazzari *et al.*, 2022). Despite this, the DT was still useful to identify specific factors that explain what makes employees stay in MiX, which is critical to inform retention strategies (Allen *et al.*, 2010; Tai *et al.*, 2025).

This leads us to our second proposition, which states that using custom-built predictive models based on organisation-specific evidence would identify context-specific drivers of turnover that might not be captured by the generic HRIS. This was fully confirmed. As described, while the HRIS assigns a number of risk factors (low, medium and high) to workers, its output does not inform managers on specific causes that enable action. Nor do the variables and assumptions considered align well with the turnover factors identified in the LR and DT models based on the data observed in this specific company. In contrast, by modelling attrition on MiX's own data, both the LR and the DT identified specific variables that are relevant to its context. In addition, the DT captured more of the complexity of turnover by showing how these factors work in combination (in this case, how tenure in the current role interacts with age and with time since the last raise). The visual output presents an intuitive representation of these key risk pathways, supporting clearer insights into employee turnover dynamics that allow managers of different levels of technical expertise to understand these results (Živković *et al.*, 2024) and take preventative measures. Applying these techniques to MiX's own data delivers interpretable results that allow managers to design effective retention strategies based on relevant evidence.

In summary, despite its essential role in storing and managing relevant employee data, the HR Analytics produced by MiX's HRIS constitutes a *black box* regarding turnover predictions that provides poor information for decision-making. Instead, using organisation-specific evidence, both LR and DT identify characteristics that can explain, with a high degree of confidence, why workers leave or remain in this organisation. This provides more accurate and actionable insights into the causes of turnover that allow MiX's managers to take targeted action. In addition, both models captured predictors that are specific to MiX and are being overlooked by the HRIS's standardised assumptions. These models showed that only some of the variables stored by the system were statistically significant in predicting turnover. In turn, they point to variables that were not even considered by the HRIS algorithm, and contradict its standardised assumptions relative to gender, age, tenure in current role and time since last raise. LR and DT predictions were able to consider MiX's idiosyncratic and context-specific features that are more likely to inform evidence-based, effective retention strategies tailored to the company (Allen *et al.*, 2010; Rubenstein *et al.*, 2018; De Vos *et al.*, 2025).

Conclusions

High levels of undesired turnover are particularly taxing for startup companies, leading them to incur onerous costs and lose critical human capital (Domurath *et al.*, 2023; Tai *et al.*, 2025). Predicting turnover in order to take preventative measures becomes vital (Marín Díaz *et al.*, 2023; Varma and Dutta, 2023). However, tech startups often face challenges such as limited financial resources and a lack of analytics expertise (Haase and Eberl, 2019; Roach and Sauermann, 2024; Rafiq *et al.*, 2025), which pushes them to rely on standardised HRIS predictions (Behl *et al.*, 2019; Jemine and Guillaume, 2022).

In this paper, we adopt an EBHRM lens to assess different approaches to HR Analytics. Drawing on the case of MiX, a tech startup, we compare the relative merits of the predictions delivered by the company's HRIS and by alternative predictive analyses based on context-specific evidence. Our results lead us to conclude that, despite the many benefits of HRIS to startups (Varma and Dutta, 2023; Wang, 2024), the HR Analytics produced by these systems can constitute *black box* forecasts that are hard to interpret and act upon, and do not consider the organisation's idiosyncrasies (Angrave *et al.*, 2016; McCartney *et al.*, 2021; Shet *et al.*, 2021). This is especially relevant in startup companies that, by their innovative and disruptive nature (Roach and Sauermann, 2024; Tai *et al.*, 2025), are unlikely to conform to the general benchmarks based on more established organisations used by commercial HRIS.

These findings are theoretically important because they speak not only to the effectiveness of different analytical tools but also to the conditions under which HR Analytics can genuinely support EBHRM in startup organisations.

Theoretical implications

The research has, therefore, three main theoretical implications. Firstly, we address the gap regarding the practical application of EBHRM by clarifying what the "best available evidence" can entail when it comes to HR Analytics, thus contributing to the "how-to" and "why" dimensions of EBHRM (Maertens *et al.*, 2025). The study shows that *having* HR Analytics (via an HRIS) is not equivalent to *evidence-based* HR decision-making. Bespoke predictive models using the company's own data emerged as better enablers of evidence-based HR Analytics by empowering managers with transparent, relevant information that allows them to make more informed and effective decisions, as espoused by EBHRM (Rousseau and Barends, 2011; Lange, 2013; Marler and Boudreau, 2017). They significantly outperformed the generic HRIS predictions based on non-contextual benchmarks, reinforcing the EBHRM principle that the source, relevance and quality of evidence matter at least as much as the sophistication of the tools used to analyse it.

Secondly, and relatedly, while the literature often positions technology as a primary *enabler* of EBHRM (McCartney and Fu, 2022), our findings challenge this assumption and suggest HRIS can just as much become a *barrier* to evidence-based practice in contexts where organisational idiosyncrasies diverge from generic benchmarks. We therefore contribute to the HRIS literature by providing empirical weight to the critique of "technological isomorphism" (Jemine and Guillaume, 2022). Existing research suggests that HRIS vendors exert pressure towards standardisation based on "best practices" from established industries like finance. Our study reveals the functional cost of this standardisation. It shows that the validity of the HRIS model degrades when organisational context diverges from the reference population. This emphasises the importance of *context* in HRM research, suggesting that it acts as a *boundary condition* for the effectiveness of HRIS-based analytics, particularly within unique ecosystems like that of technology startups (Domurath *et al.*, 2023; Roach and Sauermann, 2024; Tai *et al.*, 2025), and phenomena that are highly susceptible to contextual and social influences, like employee turnover (Rubenstein *et al.*, 2018; Amarakoon and Colley, 2023).

Thirdly, we address the research gap regarding HR Analytics in tech startups and contribute in two ways. One, our study advances research on employee dynamics in tech startups, showing that generic HRIS-based analytics can fail in these settings and, more importantly, explaining *why* they fail. The startup literature emphasises uncertainty, rapid change, underdeveloped HR practices and intense competition for talent as defining features of young ventures (Haase and Eberl, 2019; Domurath *et al.*, 2023; Roach and Sauermann, 2024; Tai *et al.*, 2025). Our findings show how these conditions translate into organisation-specific turnover drivers and interactions that differ structurally from established firms and that will not be detected by standardised models, making tech startups especially vulnerable to the above-mentioned "technological isomorphism". For instance, the higher attrition risk for female employees at MiX (captured by our bespoke models but missed by the HRIS) highlights the impact of the masculine culture prevalent in tech, a factor that general benchmarks overlook. Furthermore, autonomy and the timing of career shifts appear to be more relevant in predicting retention in tech startups than traditional retention strategies like salary increases. The contribution is therefore to position employee turnover in tech startups as a dynamic and locally contingent process, which requires organisation-specific analytics.

In addition, our findings point to the characteristics required of HR Analytics to be effective in tech startups facing the challenges described earlier. In these settings, the value of an analytical method depends not only on its predictive accuracy but also on satisfying two main conditions: the method must be both *transparent* enough for non-specialist HR practitioners to interpret and *simple* enough for them to build and maintain within constrained resources. This contributes to the HR Analytics adoption literature (Shet *et al*., 2021; Rafiq *et al*., 2025) in tech startups by proposing that interpretability and capability are not independent factors but interact: opaque methods require higher capability to evaluate critically, creating a double disadvantage for resource-constrained startups. Conversely, transparent methods reduce the capability threshold, making evidence-based practice more accessible. Methods such as LR and DTs satisfy both transparency and simplicity, while a generic HRIS satisfies neither. Their black box outputs can weaken the link between analytics and action, whereas results based on transparent models that managers can understand and interpret provide actionable insights that are more likely to be implemented and produce actual results (Marín Díaz *et al*., 2023; Živković *et al*., 2024).

Practical implications

A number of practical implications also ensue. Firstly, startup managers are urged to think critically about the data being used and seek methods that allow them to understand cause-and-effect relationships as the basis for decision-making (Allen *et al*., 2010; Lange, 2013), as well as detect existing biases in the algorithms used by generic HRIS (Wang, 2024). They should also opt for methods that best suit their data and organisational dynamics (De Vos *et al*., 2025), including the levels of analytics maturity and competence of HR professionals (McCartney and Fu, 2022; Shet, 2025). For example, LR and DTs are methods that require low levels of analytical expertise (Živković *et al*., 2024), which suits the circumstances typical of startups. Incorporating AI tools can further help tech startups manage more complex analyses and uncover patterns not easily detected through traditional methods, enabling more informed and affordable decision-making (Di Lauro *et al*., 2025; Shet, 2025).

Although simple and accessible, these methods can be used not only to forecast turnover but also for many other relevant predictions, including workload fluctuation, hiring needs and performance (Edwards and Edwards, 2016; Fernandez and Gallardo-Gallardo, 2021; Shet, 2025). They allow tech startups to extract value from data that they have, but often do not use (Rasmussen and Ulrich, 2015; McCartney and Fu, 2022) and enable a more strategic implementation of HR Analytics (Shet *et al*., 2021; Wang, 2024; Di Lauro *et al*., 2025). They serve as a starting point for a more evidence-based approach to decision-making within an integrated approach that needs to take into account the conditions and capabilities that enable effective HR Analytics in startup companies.

Our study corroborates the notion that HRIS are a valuable but insufficient component of the overall HR Analytics *architecture* (Shet, 2025). Capability deficits can be a binding constraint in tech startups, and the analytical skillset of internal staff and top management are key enablers of HR Analytics in startups (Behl *et al*., 2019). Startup managers should therefore prioritise equipping HR professionals with the analytical skills that allow them to reap the full benefits of investments made in HRIS technology (Wang, 2024; Rafiq *et al*., 2025). As evidenced by the methods used in this study, and advocated by some authors (e.g. Falletta and Combs, 2021), startups can start small in this journey, training staff to think critically and use simple statistical methods that, nonetheless, deliver reliable and informative results.

Furthermore, given the temporary nature of startups that evolve fast as they develop their processes, strategies and identities (Domurath *et al*., 2023), new evidence should be continually collected and analysed (Rafiq *et al*., 2025) in order to revise retention strategies. Moreover, following the EBHRM tenets (Rousseau and Barends, 2011), policy decisions based on analytical insights should themselves be evaluated and continuously refined. Managers should also strive to use multiple sources of evidence, relying not only on the organisational data stored in the HRIS but also their own informed expertise and stakeholder information (in particular, employee surveys and exit interviews) to further enhance the database used to predict turnover and allow them to inquire more deeply into its causes (Allen *et al*., 2010; Rousseau and Barends, 2011; Shet, 2025). The aim is to integrate HR Analytics as a powerful component of a broader evidence-based strategy without allowing it to dominate or replace the multi-faceted approach espoused by EBHRM.

Limitations and future research

Our study has limitations. For one, we rely on a single case study with a limited dataset and focus on the problems specific to the HRIS used by this company. However, the challenges faced by MiX are likely shared by most startups, given the widespread use of standardised HRIS output (Fernandez and Gallardo-Gallardo, 2021). Many companies may indeed be using this very HRIS platform, as it is a leader in the market. This makes our findings and recommendations relevant beyond this single case, although future multi-case or survey-based research would be needed to ensure generalisability.

Another limitation is that we employed only LR and DT analyses, while several other equally adequate techniques can be used (Akasheh *et al.*, 2024; De Vos *et al.*, 2025), some of which may deliver more accurate results. These methods were chosen for their ease of use and interpretable results, which are especially important in startups (Marín Díaz *et al.*, 2023; Živković *et al.*, 2024), but they serve only as an illustration of what organisations can achieve without prohibitive investment. Companies should experiment and opt for methods that best suit their data and organisational dynamics (De Vos *et al.*, 2025).

The small size of our sample is another limitation which may have compromised the accuracy of our results (especially the "Leaves" class in the DT analysis). To address this common constraint of small sample sizes in startups (Marín Díaz *et al.*, 2023; Živković *et al.*, 2024), companies should gradually build richer datasets and adopt analytical methods suited to limited data. Over time, organisations are encouraged to explore alternative techniques better aligned with their data structures and operational contexts (De Vos *et al.*, 2025) and use different sources of complementary evidence (Rousseau and Barends, 2011).

Despite these limitations, this study highlights the importance of taking into account the context in which HRM is practised (Allen *et al.*, 2010; McCartney *et al.*, 2021; Amarakoon and Colley, 2023) and provides much-needed empirical results required to support EBHRM strategies (Marler and Boudreau, 2017; Maertens *et al.*, 2025) and evidence-based HR Analytics (Falletta and Combs, 2021). Future studies that present empirical evidence of HR Analytics implementations in diverse contexts and conditions are vital to assess its effectiveness and role in sustaining evidence-based interventions.

References

Accenture (2020), "Resetting tech culture", available at: https://www.accenture.com/content/dam/accenture/final/a-com-migration/pdf/pdf-134/accenture-a4-gwc-report-final1.pdf#zoom=50

Akasheh, M.A., Malik, E.F., Hujran, O. and Zaki, N. (2024), "A decade of research on machine learning techniques for predicting employee turnover: a systematic literature review", *Expert Systems with Applications*, Vol. 238, 121794, doi: 10.1016/j.eswa.2023.121794.

Allen, D.G., Bryant, P.C. and Vardaman, J.M. (2010), "Retaining talent: replacing misconceptions with evidence-based strategies", *Academy of Management Perspectives*, Vol. 24 No. 2, pp. 48-64, doi: 10.5465/amp.24.2.48.

Amarakoon, U. and Colley, L. (2023), "Attraction and retention of employees in an Australian regional SME: the importance of place and scale in human resource management", *Personnel Review*, Vol. 52 No. 3, pp. 521-538, doi: 10.1108/pr-03-2021-0172.

Angrave, D., Charlwood, A., Kirkpatrick, I., Lawrence, M. and Stuart, M. (2016), "HR and analytics: why HR is set to fail the big data challenge", *Human Resource Management Journal*, Vol. 26 No. 1, pp. 1-11, doi: 10.1111/1748-8583.12090.

Armitage, G. (2025), "Startup turnover rates are high, onboarding support is lacking", available at: https://www.linkedin.com/posts/greggarmitage_the-average-startup-turnover-rate-is-25-activity-7394786480608399360-BfSX

Bayo-Moriones, A., Erro-Garcés, A. and Lera-López, F. (2025), "The role of technology, calculative and collaborative HRM and social trust in driving human resource analytics in European firms", *Personnel Review*, Vol. 55 No. 1, pp. 1-26, (ahead-of-print), doi: 10.1108/PR-06-2024-0561.

Behl, A., Dutta, P., Lessmann, S., Dwivedi, Y.K. and Kar, S. (2019), "A conceptual framework for the adoption of big data analytics by e-commerce startups: a case-based approach", *Information Systems and e-Business Management*, Vol. 17 Nos 2-4, pp. 285-318, doi: 10.1007/s10257-019-00452-5.

Cavanagh, J., Bartram, T., Walker, M., Pariona-Cabrera, P. and Halvorsen, B. (2024), "Health services in Australia and the impact of antiquated rostering practices on medical scientists: a case for HR analytics and evidenced-based human resource management", *Personnel Review*, Vol. 53 No. 1, pp. 18-33, doi: 10.1108/pr-09-2021-0690.

De Vos, S., Bockel-Rickermann, C., Van Belle, J. and Verbeke, W. (2025), "Predicting employee turnover: scoping and benchmarking the state-of-the-art", *Business and Information Systems Engineering*, Vol. 67 No. 5, pp. 733-752, doi: 10.1007/s12599-024-00898-z.

Di Lauro, S., Tursunbayeva, A., Antonelli, G. and Moschera, L. (2025), "Disrupting human resource management with people analytics: a study of applications, value, enablers and barriers in Italy", *Personnel Review*, Vol. 54 No. 2, pp. 697-721, doi: 10.1108/pr-11-2023-0927.

Domurath, A., Taggar, S. and Patzelt, H. (2023), "A contingency model of employees' turnover intent in young ventures", *Small Business Economics*, Vol. 60 No. 3, pp. 901-927, doi: 10.1007/s11187-022-00629-2.

Dul, J. and Hak, T. (2008), *Case Study Methodology in Business Research*, Routledge, London.

Edwards, M.R. and Edwards, K. (2016), *Predictive HR Analytics: Mastering the HR Metric*, Kogan Page, London.

Falletta, S.V. and Combs, W.L. (2021), "The HR analytics cycle: a seven-step process for building evidence-based and ethical HR analytics capabilities", *Journal of Work-Applied Management*, Vol. 13 No. 1, pp. 51-68, doi: 10.1108/jwam-03-2020-0020.

Fernandez, V. and Gallardo-Gallardo, E. (2021), "Tackling the HR digitalization challenge: key factors and barriers to HR analytics adoption", *Competitiveness Review*, Vol. 31 No. 1, pp. 162-187, doi: 10.1108/cr-12-2019-0163.

Guerranti, F. and Dimitri, G.M. (2023), "A comparison of machine learning approaches for predicting employee attrition", *Applied Sciences*, Vol. 13 No. 1, p. 267, doi: 10.3390/app13010267.

Haase, A. and Eberl, P. (2019), "The challenges of routinizing for building resilient startups", *Journal of Small Business Management*, Vol. 57 No. S2, pp. 579-597, doi: 10.1111/jsbm.12511.

Hancock, J.I., Allen, D.G., Bosco, F.A., McDaniel, K.R. and Pierce, C.A. (2013), "Meta-analytic review of employee turnover as a predictor of firm performance", *Journal of Management*, Vol. 39 No. 3, pp. 573-603, doi: 10.1177/0149206311424943.

Jemine, G. and Guillaume, K. (2022), "Lifting the veil on HRIS adoption: the role of vendors and consultants in the diffusion of HR innovations", *Information Technology and People*, Vol. 35 No. 7, pp. 2418-2440, doi: 10.1108/itp-02-2021-0114.

Kang, I.G., Croft, B. and Bichelmeyer, B.A. (2021), "Predictors of turnover intention in U.S. federal government workforce: machine learning evidence that perceived comprehensive HR practices predict turnover intention", *Public Personnel Management*, Vol. 50 No. 4, pp. 538-558, doi: 10.1177/0091026020977562.

Kroon, B. (2022), *Evidence Based HRM: what we Know About People in Workplaces*, Open Press Tilburg University, Tilburg.

Lange, T. (2013), "Evidence-based HRM: a scholarship perspective with a difference", *Evidence-based HRM: a Global Forum for Empirical Scholarship*, Vol. 1 No. 1, pp. 4-15, doi: 10.1108/20493981311318584.

Lazzari, M., Alvarez, J.M. and Ruggieri, S. (2022), "Predicting and explaining employee turnover intention", *International Journal of Data Science and Analytics*, Vol. 14 No. 3, pp. 279-292, doi: 10.1007/s41060-022-00329-w.

Maertens, L., Daniëls, E., Hondeghem, A. and Vandenabeele, W. (2025), "Towards a new conceptualisation of evidence-based human resource management", *Journal of Organisational Effectiveness: People and Performance*, Vol. 12 No. 3, pp. 559-584, doi: 10.1108/joepp-12-2023-0541.

Marín Díaz, G., Galán Hernández, J.J. and Galdón Salvador, J.L. (2023), "Analyzing employee attrition using explainable AI for strategic HR decision-making", *Mathematics*, Vol. 11 No. 22, p. 4677, doi: 10.3390/math11224677.

Marler, J.H. and Boudreau, J.W. (2017), "An evidence-based review of HR analytics", *International Journal of Human Resource Management*, Vol. 28 No. 1, pp. 3-26, doi: 10.1080/09585192.2016.1244699.

McCartney, S. and Fu, N. (2022), "Promise versus reality: a systematic review of the ongoing debates in people analytics", *Journal of Organisational Effectiveness: People and Performance*, Vol. 9 No. 2, pp. 281-311, doi: 10.1108/joepp-01-2021-0013.

McCartney, S., Murphy, C. and McCarthy, J. (2021), "21st century HR: a competency model for the emerging role of HR analysts", *Personnel Review*, Vol. 50 No. 6, pp. 1495-1513, doi: 10.1108/pr-12-2019-0670.

Negt, P. and Haunschild, A. (2025), "Exploring the gap between research and practice in human resource management (HRM): a scoping review and agenda for future research", *Management Review Quarterly*, Vol. 75 No. 1, pp. 837-879, doi: 10.1007/s11301-023-00397-7.

Rafiq, U., Wang, X. and Guerra, E. (2025), "Data analytics in software startups: understanding key concepts and critical challenges", *Information and Software Technology*, Vol. 180, 107652, doi: 10.1016/j.infsof.2024.107652.

Raghavan, M., Barocas, S., Kleinberg, J. and Levy, K. (2020), "Mitigating bias in algorithmic hiring: evaluating claims and practices", *Proceedings of the 2020 conference on fairness, accountability, and transparency (FAT* '20)*, Barcelona, January 27-30, 2020, pp. 469-481, doi: 10.1145/3351095.3372828.

Rasmussen, T. and Ulrich, D. (2015), "Learning from practice: how HR analytics avoids being a management fad", *Organisational Dynamics*, Vol. 44 No. 3, pp. 236-242, doi: 10.1016/j.orgdyn.2015.05.008.

Roach, M. and Sauermann, H. (2024), "Can technology startups hire talented early employees? Ability, preferences, and employee first job choice", *Management Science*, Vol. 70 No. 6, pp. 3619-3644, doi: 10.1287/mnsc.2023.4868.

Rombaut, E. and Guerry, M.-A. (2018), "Predicting voluntary turnover through human resources database analysis", *Management Research Review*, Vol. 41 No. 1, pp. 96-112, doi: 10.1108/mrr-04-2017-0098.

Rousseau, D.M. and Barends, E.G.R. (2011), "Becoming an evidence-based HR practitioner", *Human Resource Management Journal*, Vol. 21 No. 3, pp. 221-235, doi: 10.1111/j.1748-8583.2011.00173.x.

Rubenstein, A.L., Eberly, M.B., Lee, T.W. and Mitchell, T.R. (2018), "Surveying the forest: a meta-analysis, moderator investigation, and future-oriented discussion of the antecedents of voluntary employee turnover", *Personnel Psychology*, Vol. 71 No. 1, pp. 23-65, doi: 10.1111/peps.12226.

Shet, S.V. (2025), "Developing human resource analytics capability architecture in organisations", *Personnel Review*, Vol. 54 No. 9, pp. 2286-2310, doi: 10.1108/pr-01-2024-0009.

Shet, S., Poddar, T., Wamba, S.F. and Dwivedi, Y.K. (2021), "Examining the determinants of successful adoption of data analytics in human resource management – a framework for implications", *Journal of Business Research*, Vol. 131, pp. 311-326, doi: 10.1016/j.jbusres.2021.03.054.

Steinfeld, J. (2024), "Employee attrition", available at: https://carta.com/sg/en/learn/startups/compensation/employee-attrition/

Tai, C.-S., Uen, J.-F. and Lu, S.-H. (2025), "Effects of employer brand on employee retention in small startup high-tech companies: the moderation of agile value", *Employee Relations*, Vol. 47 No. 2, pp. 332-354, doi: 10.1108/er-12-2023-0655.

Varma, D. and Dutta, P. (2023), "Empowering human resource functions with data-driven decision-making in start-ups: a narrative inquiry approach", *International Journal of Organisational Analysis*, Vol. 31 No. 4, pp. 945-958, doi: 10.1108/ijoa-08-2021-2888.

Yin, R.K. (2018), *Case Study Research and Applications: Design and Methods*, 6th ed., Sage, Thousand Oaks, CA.

Živković, A., Šebalj, D. and Franjković, J. (2024), "Prediction of the employee turnover intention using decision trees", *Proceedings of the 26th International Conference on Enterprise Information Systems*, Anger, April 28-30 2024, pp. 325-336, doi: 10.5220/0012538400003690.

Wang, A. (2024), "Enhancing HR management through HRIS and data analytics", *Applied and Computational Engineering*, Vol. 64 No. 1, pp. 222-228, doi: 10.54254/2755-2721/64/20241394.

Corresponding author
Ana Carvalho can be contacted at: anac@eeg.uminho.pt