Universidade do Minho
Escola de Engenharia

Diogo Gonçalves Graçoeiro

The Importance of Workforce Planning and Turnover Prediction in Retail: Leveraging People Analytics for Strategic Decision Making

PREVIEW

Abril de 2025

Universidade do Minho
Escola de Engenharia

Diogo Gonçalves Graçoeiro

The Importance of Workforce Planning and Turnover Prediction in Retail: Leveraging People Analytics for Strategic Decision Making

PREVIEW

Dissertação de Mestrado
Mestrado em Engenharia e Gestão Industrial

Trabalho efetuado sob a orientação da
Professora Maria do Sameiro Carvalho
Professor João Nuno Gonçalves

Abril de 2025

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho

Atribuição

CC BY

https://creativecommons.org/licenses/by/4.0/

PREVIEW

ACKNOWLEDGMENTS

As I reach the culmination of my academic path, I am overwhelmed with a profound sense of accomplishment and gratitude. Completing this thesis means the end of a remarkable chapter in my life and represents the fulfillment of a long-held aspiration. This transformative journey has enriched me with invaluable experiences.

A first word to the University of Minho, where I pursued my degree, for providing me with a comprehensive education and fostering an environment of academic excellence. The knowledge and skills I acquired during my studies have been pivotal in shaping my career path. I would like to express my heartfelt thanks to Professor Maria Carvalho and Professor João Gonçalves, my supervisors at the University. Their unwavering support, valuable guidance and insightful feedback have significantly influenced the direction and quality of my research.

Foremost, I would like to express my deepest gratitude to LTPlabs, the company where I had the privilege to complete my internship. Their unwavering support and guidance throughout my internship were instrumental in shaping my skills and knowledge in the field. I am immensely thankful to all the LTPeers for their warm welcome, collaboration and shared expertise, which greatly contributed to my professional growth.

PREVIEW

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

PREVIEW

A Importância do Planeamento da Força de Trabalho e da Previsão de Rotatividade no Retalho: Alavancar People Analytics para a Tomada de Decisão Estratégica

RESUMO

O presente trabalho propõe uma abordagem inovadora para o planeamento estratégico de recursos humanos no setor do retalho, através do desenvolvimento de modelos preditivos de *turnover* com base em técnicas de People Analytics. A elevada taxa de turnover neste setor, aliada à necessidade de uma gestão mais eficiente da força de trabalho, justificou a criação de duas soluções distintas, ambas baseadas no algoritmo XGBoost: um modelo de previsão a um ano e outro a cinco anos.

O modelo de um ano tem como objetivo estimar a probabilidade de saída voluntária de cada colaborador com base nas suas características individuais. Já o modelo de cinco anos é desenvolvido a partir de dados agregados, considerando combinações de características como função, localização e tipo de loja, bem como perfis-tipo de colaboradores com atributos semelhantes. Esta estrutura permite realizar projeções estratégicas de longo prazo em contextos de maior incerteza. Para assegurar a aplicabilidade prática das previsões, o modelo utiliza apenas variáveis e categorias cujo valor é conhecido ou estimável à partida.

Ambos os modelos demonstraram capacidades preditivas relevantes, destacando-se a importância de variáveis como remuneração fixa, senioridade e idade. No entanto, verificaram-se enviesamentos sistemáticos em alguns segmentos — especialmente nos grupos com menor senioridade — e uma utilização limitada das avaliações de performance e potencial, fatores que foram analisados criticamente. As previsões desenvolvidas nesta dissertação destinam-se a alimentar um modelo de otimização já existente, reforçando o seu potencial para apoiar decisões estratégicas de recursos humanos.

Com base nos resultados obtidos, o estudo apresenta também um conjunto de propostas para investigação e desenvolvimento futuro, que incluem a revisão da forma como os perfis-tipo de colaboradores são definidos, a incorporação da mobilidade interna e o reforço da interpretabilidade local do modelo. Este projeto representa um contributo prático e relevante para a adoção de abordagens *data-driven* na gestão de talento, aproximando os recursos humanos da estratégia organizacional.

PREVIEW

PALAVRAS-CHAVE

People Analytics; Retalho; Turnover; Machine Learning; Planeamento Estratégico de Recursos Humanos

PREVIEW

ABSTRACT

This dissertation proposes an innovative approach to strategic workforce planning in the retail sector through the development of predictive turnover models based on People Analytics techniques. The high turnover rate in this sector, combined with the need for more efficient workforce management, justified the creation of two distinct solutions, both using the XGBoost algorithm: a one-year prediction model and a five-year forecast model.

The one-year model aims to estimate the probability of voluntary employee exit based on individual characteristics. The five-year model, on the other hand, is built from aggregated data, considering combinations of features such as role, location and store type, as well as employee-type profiles with similar attributes. This structure enables long-term strategic projections in contexts of higher uncertainty. To ensure practical applicability, the model relies only on variables and categories whose future values are known or can be reliably estimated in advance.

Both models demonstrated relevant predictive capabilities, with fixed compensation, seniority and age standing out as key drivers. However, systematic biases were identified in some segments — particularly among early-career groups — and performance and potential evaluations showed limited influence, which were critically assessed. These predictions are designed to serve as input for an existing workforce optimization model, supporting more informed hiring and retention planning over time.

Based on the results, the study also presents a set of proposals for future research and development, including a revision of how employee groups are defined, the incorporation of internal mobility and enhanced local interpretability of the model. This project provides a practical and relevant contribution to the adoption of data-driven approaches in talent management, bridging human resources and organizational strategy.

PREVIEW

KEYWORDS

People Analytics; Retail; Turnover; Machine Learning; Strategic Workforce Planning

PREVIEW

CONTENTS

Acknowledgments ... iii

Resumo ... v

Abstract ... vii

Keywords ... vii

Contents ... ix

List of Figures ... xii

List of Tables ... xiv

Acronyms ... xv

1. Introduction ... 1
   - 1.1 Motivation ... 1
   - 1.2 Project Background ... 2
   - 1.3 Project Objectives and Expected Results ... 3
   - 1.4 Thesis Outline ... 4
2. Literature Review ... 5
   - 2.1 Introduction to Employee Turnover ... 5
   - 2.2 Traditional Models for Employee Turnover ... 6
   - 2.3 Factors affecting Employee Turnover ... 8
   - 2.4 Workforce Planning and the Role of Predictive Analytics in Turnover Management ... 10
     - 2.4.1 The Impact of Worforce Planning on Turnover Management ... 11
     - 2.4.2 The Role of Predictive Analytics and Machine Learning ... 11
     - 2.4.3 Challenges in Implementing Workforce Planning Strategies ... 14
     - 2.4.4 Workforce Planning as a Competitive Advantage ... 14
   - 2.5 Machine Learning Models for Turnover Prediction ... 15
     - 2.5.1 Comparative Analysis of Machine Learning Models in Turnover Prediction ... 15
     - 2.5.2 XGBoost: A Benchmark Model for Turnover Prediction ... 19
   - 2.6 Summary of Key Findings ... 20
3. The Problem ... 21
   - 3.1 Problem Description ... 21

PREVIEW

3.1.1 As-Is 21
3.1.2 Point for Improvement 22
3.2 Data Overview and Exploratory Analysis 23
3.2.1 Data Overview 23
3.2.2 Data Analysis 24
3.3 Final Remarks 34
4. Methodology for Turnover Prediction 35
4.1 Data Preparation 37
4.1.1 Data Sources 37
4.1.2 Data Integration and Consolidation 37
4.1.3 Variable Structuring 38
4.1.4 Variable Classification 42
4.2 Data Cleaning and Processing 42
4.3 Predictive Models Design and Evaluation Methodology 43
4.3.1 One-year Model 43
4.3.2 Five-year Model 46
4.3.3 Evaluation 50
4.4 Final Remarks 52
5. Results Analysis and Discussion 54
5.1 One-year Model 54
5.1.1 Assessment of Predictive Performance 54
5.1.2 Drivers of Turnover Prediction 64
5.2 Five-year model 65
5.2.1 Assessment of Predictive Performance 65
5.2.2 Drivers of Turnover Prediction 77
5.3 Final Remarks 79
6. Conclusion 81
6.1 Critical Analysis Strategic Assessment and Critical Reflection 81

PREVIEW

6.2 Future Research and Development 83

6.2.1 Short-Term Improvements 84

6.2.2 Medium/Long-Term Improvements 87

References 90

Appendix 96

Appendix A – Supporting Data and Extended Analysis 96

PREVIEW

LIST OF FIGURES

Figure 1 - Annual Workforce Evolution and Turnover Rate 24
Figure 2 - Percentage of Employees with and without Turnover 25
Figure 3 - Workforce Distribution and Turnover Rate by District (Top 10) 26
Figure 4 - Workforce Distribution and Turnover Rate By Store Type 26
Figure 5 - Workforce Distribution and Turnover Rate by Role Cluster 27
Figure 6 - Workforce Distribution and Turnover Rate by Seniority 28
Figure 7 - Workforce Distribution and Turnover Rate by Age 29
Figure 8 - Workforce Distribution and Turnover Rate by Contract Type 29
Figure 9 - Workforce Distribution and Turnover Rate by Standardized Wage 31
Figure 10 - Workforce Distribution and Turnover Rate by Academic Qualifications 31
Figure 11 - Workforce Distribution and Turnover Rate by Workload Type 32
Figure 12 - Workforce Distribution and Turnover Rate by Employee Performance 33
Figure 13 - Workforce Distribution and Turnover Rate by Employee Potential 34
Figure 14 - Actual Turnover vs Turnover Prediction (One-year model) 55
Figure 15 - Actual Turnover vs Turnover Prediction by Top 10 District (One-year model) 57
Figure 16 – Actual Turnover vs Turnover Prediction by Store Type (One-year model) 58
Figure 17 - Actual Turnover vs Turnover Prediction by Top 10 Role Cluster (One-year model) 59
Figure 18 – Actual Turnover vs Turnover Prediction by Academic Qualifications (One-year model) 59
Figure 19 - Actual Turnover vs Turnover Prediction by Contract Relationship (One-year model) 60
Figure 20 - Actual Turnover vs Turnover Prediction by Workload Type (One-year model) 60
Figure 21 - Actual Turnover vs Turnover Prediction by Age (One-year model) 61
Figure 22 - Actual Turnover vs Turnover Prediction by Seniority (One-year model) 62
Figure 23 - Actual Turnover vs Turnover Prediction by Performance (One-year model) 63
Figure 24 - Actual Turnover vs Turnover Prediction by Potential (One-year model) 63
Figure 25 - Actual Turnover vs Turnover Prediction (Five-year model) 66
Figure 26 - Actual Turnover vs Turnover Prediction by Top 10 District (Five-year model) 69
Figure 27 - Actual Turnover vs Turnover Prediction by Store Type (Five-year model) 70
Figure 28 - Actual Turnover vs Turnover Prediction by Role Cluster (Five-year model) 71
Figure 29 - Actual Turnover vs Turnover Prediction by Academic Qualifications (Five-year model) 72
Figure 30 - Actual Turnover vs Turnover Prediction by Contract Relationship (Five-year model) 72

PREVIEW

Figure 31 - Actual Turnover vs Turnover Prediction by Workload Type (Five-year model) 73
Figure 32 - Actual Turnover vs Turnover Prediction by Age (Five-year model) 74
Figure 33 - Actual Turnover vs Turnover Prediction by Seniority (Five-year model) 75
Figure 34 - Actual Turnover vs Turnover Prediction by Performance (Five-year model) 76
Figure 35 - Actual Turnover vs Turnover Prediction by Potential (Five-year model) 76
Figure 36 - Workforce Distribution and Turnover Rate by Marital Status 96
Figure 37 - Store Type Distribution by Seniority 97
Figure 38 - Age Distribution by Seniority 98
Figure 39 - Contract Type Distribution by Age 99
Figure 40 - Contract Type Distribution by Workload Type 100
Figure 41 - Workforce Distribution and Standardized Wage by Role Cluster 101
Figure 42 - Workforce Distribution and Standardized Wage by Role Cluster (Others) 102

PREVIEW

LIST OF TABLES

Table 1 - Overview of Turnover Prediction Studies: Models, Metrics and Applications 18
Table 2 - Dataset Structure for Turnover Prediction 37
Table 3 - Variables Classification for Turnover Prediction Model 42
Table 4 - Dataset Structure for Five-year Model 46
Table 5 - Variables Classification for Turnover Five-year Prediction Model 48
Table 6 - Overall Performance Metrics for the One-Year Turnover Prediction Model 54
Table 7 – Turnover Prediction Model Performance at Aggregated Segment Levels (One-year model) .. 56
Table 8 - Top 10 Most Important Variables (One-year Model) 64
Table 9 - Overall Performance Metrics for the Five-Year Turnover Prediction Model 65
Table 10 - Turnover Prediction Model Performance at Aggregated Segment Levels (Five-year model) .. 68
Table 11 - Top 10 Most Important Variables (Five-year Model) 77

PREVIEW

ACRONYMS

AI – Artificial Intelligence.

AUC – Area Under the Curve.

AUC-ROC – Area Under the Receiver Operating Characteristic Curve.

FTE – Full-Time Equivalent.

GBM – Gradient Boosting Machine.

GBT – Gradient Boosting Tree.

GLM – Generalized Linear Model.

HR – Human Resources.

HRIS – Human Resource Information Systems.

HRM – Human Resource Management.

KNN – K-Nearest Neighbors.

LDA – Linear Discriminant Analysis.

LR – Logistic Regression.

MAE – Mean Absolute Error.

ML – Machine Learning.

NN – Neural Network.

RF – Random Forest.

RNN – Recurrent Neural Network.

RQ – Research Question.

SHAP - SHapley Additive exPlanations.

SMOTE – Synthetic Minority Oversampling Technique.

SVM – Support Vector Machine.

VBM – Value Based Management.

XGBoost – Extreme Gradient Boosting.

PREVIEW

1. INTRODUCTION

The present dissertation is within the realm of the Master in Industrial Engineering and Management at the University of Minho. The thesis was developed during an internship at LTPLabs, an analytical management consultancy company, as part of a project with one of its clients. The intent of this chapter is to give a general overview of the dissertation's subject and extent, problems that are handled, project's objectives and structure of the entire document.

1.1 Motivation

Employee turnover remains a persistent challenge in the retail sector, affecting operational efficiency, customer service quality and overall business performance. High turnover rates lead to increased recruitment and training costs while also disrupting store operations, particularly in frontline roles where experience and continuity are critical for maintaining service standards (Olubiyi et al., 2019). Despite advancements in Human Resources (HR) technologies — such as workforce analytics tools and talent management platforms — workforce planning in many organizations continues to rely heavily on empirical knowledge rather than data-driven decision-making, limiting the ability to anticipate workforce needs effectively (Majam & Jarbandhan, 2022).

Traditional workforce planning is often reactive, adjusting staffing levels only after turnover has occurred. In contrast, People Analytics enables a proactive approach, leveraging historical workforce data — such as tenure, role transitions and demographic indicators — to develop predictive models capable of forecasting turnover patterns (Polzer, 2022). By integrating predictive analytics into workforce planning, companies can improve long-term staffing strategies, optimize hiring processes and ensure business continuity (Majam & Jarbandhan, 2022).

Beyond workforce planning, organizational culture also influences turnover dynamics. Olubiyi et al. (2019) emphasize that Person-Organization Fit significantly impacts job satisfaction and employee commitment, which in turn affects retention. Employees who align with company values and perceive a strong sense of belonging are more likely to remain, particularly in customer-facing environments where engagement directly influences service quality. Additionally, their study highlights that turnover is driven not only by financial incentives but also by factors such as leadership quality, career development opportunities and workplace culture.

PREVIEW

While existing research explores various factors influencing turnover, fewer studies have examined how predictive analytics can be systematically applied to workforce planning. Machine learning models are increasingly being adopted in HR decision-making, offering greater accuracy in identifying workforce trends and informing strategic planning (Heidemann et al., 2024). However, the ethical implications of Artificial Intelligence (AI)-driven HR analytics must also be considered to ensure that predictive models are implemented in a responsible and transparent manner that supports both organizational and employee interests (Giermindl et al., 2022).

This dissertation addresses this gap by developing a predictive turnover model tailored to the retail sector. Rather than focusing on individual retention strategies, the study aims to support strategic workforce planning by forecasting employee departures using machine learning techniques. The predictive model provides annual turnover estimates, allowing for a more data-driven approach in staffing decisions. By identifying key workforce variables and analyzing turnover trends, this research offers a structured methodology for integrating predictive analytics into HR decision-making. The findings will help HR professionals anticipate workforce gaps, optimize hiring needs and improve cost efficiency in workforce management.

PREVIEW

1.2 Project Background

This dissertation was developed during a consulting project for a major Portuguese company operating in the retail sector, specifically in the consumer goods segment. The company is a key player in both the national and international market, managing a large portfolio of stores and offering a diverse range of products, including groceries, household items and personal care products. With an extensive store network and a high transaction volume, the company serves millions of customers annually across different locations, requiring careful workforce management to maintain operational efficiency.

Given the dynamic nature of consumer demand and the complexity of managing a large workforce, ensuring strategic workforce planning is a key challenge. For this study, the focus was placed on three primary store formats that constitute the core of the company's operations:

- Large hypermarkets, characterized by high transaction volumes and requiring a large, stable workforce;
- Medium-sized supermarkets, which balance product variety and efficiency, necessitating precise workforce planning;
- Neighborhood convenience stores, where agility and workforce flexibility are critical for operations.

Despite structured workforce planning efforts, employee turnover remains a major challenge, affecting workforce stability, service quality and operational costs. Currently, turnover predictions are based on empirical methods, relying on historical workforce variations rather than predictive analytics. This approach lacks analytical depth and limits the company's ability to anticipate workforce gaps proactively. Addressing this gap in turnover forecasting is essential for enhancing workforce planning and optimizing resource allocation, ensuring that stores remain well-staffed to meet business demands.

1.3 Project Objectives and Expected Results

The primary objective of this dissertation is to develop a turnover prediction model using advanced machine learning techniques to support strategic workforce planning. Given the company's reliance on historical workforce variations to estimate turnover, this model seeks to introduce a data-driven approach that incorporates both intrinsic (e.g., employee attributes) and extrinsic (e.g., market conditions) factors.

The model will generate year-by-year turnover predictions over a five-year period, providing granular insights into employee departure probabilities. Unlike the existing empirical approach, which primarily uses Full-Time Equivalent (FTE)-based estimations, this model will allow the company to proactively address workforce gaps, particularly in key operational functions.

From the second year onwards, the model does not aim to predict individual exits but rather forecasts the expected number of voluntary departures aggregated by employee profiles, role cluster, store type and region — supporting group-level workforce planning instead of employee-level forecasting.

The expected outcome of this dissertation is the development of a high-precision turnover prediction model, capable of identifying patterns in employee attrition and providing actionable insights to support HR planning and decision-making. This predictive capability will not only enable HR and management teams to develop targeted retention strategies and adjust recruitment plans, but will also serve as a key input for an already existing workforce optimization model developed by the company or project team. Through this integration, the predicted turnover rates will support more efficient workforce structuring, helping to balance hiring, internal mobility and promotion decisions across different employee segments and geographical locations.

In addition to model development, this dissertation also aims to reflect on how predictive tools and People Analytics can address the challenges of managing voluntary turnover in high-variability environments, such as retail and how such tools can support more effective, data-driven workforce strategies.

PREVIEW

To guide the investigation and structure the narrative across the different chapters, the following research questions (RQs) are proposed:

- RQ1: What are the main individual and contextual factors that influence voluntary employee turnover in the retail sector?
- RQ2: How can predictive analytics be used to forecast voluntary turnover and support strategic workforce planning in high-turnover environments?

1.4 Thesis Outline

The present dissertation is organized into six main chapters, each contributing to a structured understanding of the turnover prediction problem in the retail sector and the development of a data-driven solution. Chapter 1 introduces the thesis by presenting the motivation, project background, objectives and expected results, as well as outlining the document structure. Chapter 2 provides a literature review that contextualizes the issue of employee turnover and highlights the relevance of workforce planning and predictive analytics. It also explores traditional and modern approaches to turnover prediction, including machine learning methods, with special emphasis on the XGBoost algorithm. Chapter 3 describes the problem addressed in this project, beginning with an analysis of the current decision-making context, identifying the limitations of the existing approach and providing a detailed overview of the available data and exploratory analysis. Chapter 4 outlines the methodological approach adopted to develop the predictive models. It includes details on data preparation, cleaning, processing and structuring, followed by the deployment of two models: a one-year individual-level model and a five-year aggregate-level model. Chapter 5 presents the results obtained from each model. It evaluates their predictive performance and interprets the most important variables driving turnover, based on their relevance in the XGBoost framework. Chapter 6 delivers a critical analysis of the models' strengths and limitations, followed by a set of short-term and medium/long-term recommendations for future research and development. The dissertation concludes with references and an appendix that includes additional analyses and data used to support the work.

PREVIEW

2. LITERATURE REVIEW

This chapter provides a structured literature review to support the development of a predictive model for employee turnover in the retail sector. It begins by clarifying the concept of turnover and its organizational implications, followed by a review of traditional approaches and the key intrinsic and extrinsic factors that influence voluntary turnover. The chapter then explores how workforce planning can benefit from predictive analytics, highlighting both its strategic relevance and practical challenges. Finally, various machine learning techniques are analyzed, with a special focus on decision tree models such as XGBoost, to evaluate their suitability for turnover prediction in high-variability environments. This review lays the theoretical foundation for the methodology adopted in this dissertation.

2.1 Introduction to Employee Turnover

Employee turnover refers to the rate at which employees leave an organization and must be replaced. It is a critical workforce metric that significantly impacts organizational stability, financial performance and operational efficiency (Olubiyi et al., 2019). Turnover can occur for various reasons, ranging from job dissatisfaction and career progression to organizational restructuring and layoffs, making it a key concern in Human Resource Management (HRM) (Al-Suraihi et al., 2021).

There are two primary types of turnover: voluntary and involuntary. Voluntary turnover occurs when employees choose to leave their jobs, often in search of better career opportunities, higher salaries, improved work-life balance or greater job satisfaction. On the other hand, involuntary turnover happens when employees are terminated due to factors such as layoffs, underperformance or structural changes within the company (Al-Suraihi et al., 2021). Additionally, turnover can be classified as functional or dysfunctional. Functional turnover refers to the departure of underperforming employees, which can benefit the organization by improving team efficiency, whereas dysfunctional turnover involves the loss of high-performing employees, leading to potential skill shortages and disruptions in business operations (Wallace & Gaylor, 2012). Another important distinction is between avoidable and unavoidable turnover. Avoidable turnover results from factors that organizations can control, such as poor management, lack of career growth or inadequate compensation, while unavoidable turnover is driven by external factors, such as retirement, personal relocation or health issues (Barrick & Zimmerman, 2005).

Taking turnover into account is essential because of its financial, operational and cultural implications. Financially, replacing an employee is costly, with expenses related to recruitment, onboarding and training. The Work Institute (2024) estimates that replacing a single employee costs

PREVIEW

approximately 33% of their annual salary, making turnover a substantial financial burden, particularly in industries with high turnover rates such as retail. Operationally, frequent employee departures disrupt workflow and reduce productivity, as new hires require time to adapt and reach peak performance levels (Rahaman & Bari, 2024). In customer-facing industries like retail, turnover negatively impacts service quality and customer satisfaction, as experienced employees are more effective in handling customer interactions and operational demands (Bar-Gil et al., 2024).

Beyond financial and operational concerns, turnover also has psychological and cultural effects within organizations. A high turnover rate can lower employee morale, increase workplace stress and reduce overall job engagement. The departure of experienced employees leads to the loss of institutional knowledge, further weakening long-term strategic goals and business continuity (Gamba et al., 2024). Consequently, organizations are increasingly focusing on proactive strategies to manage turnover effectively.

One of the most effective approaches to mitigating turnover is the use of People Analytics, which enables HR departments to analyze key workforce indicators — such as job satisfaction, career progression and workplace sentiment — to anticipate employee departures and develop targeted retention strategies. Predictive analytics in HR allows organizations to shift from reactive turnover management to proactive workforce planning, ensuring business continuity and optimizing talent retention (Rahaman & Bari, 2024). Companies that strategically incorporate predictive modeling into their workforce planning processes can enhance workforce stability, reduce hiring costs and create a more resilient and engaged workforce (Levenson, 2018).

PREVIEW

2.2 Traditional Models for Employee Turnover

The study of employee turnover has gained significant attention over the years, leading to the development of several theoretical models that explain why employees decide to leave their organizations. Among the most influential contributions in this field are the works of Mobley (1977), Lee (1988), Lee & Mitchell (1994), Morrell et al. (2008) and Wöcke & Heymann (2012), each offering different perspectives on turnover decision-making.

One of the earliest and most foundational models is Mobley's framework, which emphasizes the relationship between job dissatisfaction and voluntary turnover. According to this model, employees who experience dissatisfaction begin evaluating alternative job opportunities, comparing offers and assessing the potential costs and benefits of leaving their current role. This perspective suggests that the decision to leave an organization is largely driven by perceived job alternatives and the weighing of risks associated with the transition (G. J. Lee & Rwigema, 2005; Wöcke & Heymann, 2012).

Building upon Mobley's findings, Lee introduced a refined approach that shifted the focus from job satisfaction to job commitment and involvement as the key determinants of voluntary turnover. His research highlights that employees' level of attachment to the organization plays a crucial role in influencing their decision to stay or leave, reinforcing the idea that commitment, rather than satisfaction alone, dictates turnover behavior (Lee, 1988).

A more dynamic and comprehensive approach was later proposed by Lee & Mitchell (1994) with the development of the *Unfolding Model of Turnover*. This framework suggests that turnover decisions are often triggered by external or internal shocks that lead employees to reassess their employment situation. The model outlines different decision-making paths, considering scenarios in which employees undergo a reevaluation process either due to a significant external event, an unexpected career opportunity or a gradual reassessment of their current role. In some cases, employees might be prompted to leave by a sudden career disruption, while in others, turnover results from a long-term decline in job commitment without an external trigger.

Seeking to refine and improve the predictive accuracy of turnover models, Morrell et al. introduced the *Mapping the Decision to Quit* framework, which builds upon the *Unfolding Model* by emphasizing that turnover is rarely an impulsive decision but rather a gradual and structured process. This model argues that employees typically go through multiple stages before making a final decision to leave, reinforcing the idea that turnover should be seen as a series of evolving considerations rather than a single event. Compared to its predecessors, this model is considered to provide a more accurate and structured depiction of turnover behavior, allowing organizations to anticipate employee departures with greater precision (Morrell et al., 2008).

These theoretical models have significantly shaped turnover research by providing insights into the various factors that drive employee departures. While early models primarily linked turnover to dissatisfaction and commitment, later frameworks integrated external influences, career shocks and long-term psychological processes, emphasizing the complex and multi-faceted nature of turnover. Understanding these models remains essential for organizations seeking to develop effective retention strategies and minimize workforce disruptions.

Despite the significant contributions of traditional models in explaining employee turnover, they also exhibit limitations that reduce their predictive power and practical applicability. One of the primary shortcomings is their limited predictive accuracy, as these models often focus on individual psychological

PREVIEW

factors such as job satisfaction and organizational commitment but fail to incorporate a comprehensive set of variables that can enhance prediction accuracy (Morrell et al., 2008). Research has shown that turnover decisions are influenced by a wide array of factors, many of which are not adequately captured by traditional frameworks. Additionally, traditional models fail to consider external influences, such as macroeconomic conditions, labor market trends and organizational restructuring, which can significantly impact an employee's decision to leave. While the *Unfolding Model* introduced the concept of external "shocks" triggering turnover, earlier models primarily treated turnover as a gradual process based on dissatisfaction, overlooking abrupt decision-making scenarios (Lee & Mitchell, 1994).

Another critical limitation of many traditional turnover models is their tendency to assume linear relationships between variables, whereas real-world turnover decisions often exhibit nonlinear patterns influenced by multiple interacting factors (Lee & Mitchell, 1994). These models tend to simplify turnover decision-making into predefined stages, which may not fully reflect the dynamic and evolving nature of employee exit decisions. Additionally, conventional statistical approaches, such as logistic regression, dominate turnover research but face limitations when applied to real-world workforce data. These methods struggle with imbalanced datasets, where the number of employees who leave is significantly smaller than those who stay, leading to biased predictions and reduced generalizability. Furthermore, traditional approaches often rely on predefined assumptions about variable relationships, which limits their ability to detect hidden correlations and capture the interplay between multiple factors that influence turnover, particularly when these interactions evolve over time (Park et al., 2024).

Modern approaches leveraging machine learning techniques, such as decision trees and ensemble models, have demonstrated superior predictive performance by automatically learning from data patterns rather than relying on fixed assumptions. Unlike traditional methods, machine learning models can process high-dimensional datasets, capturing intricate relationships between employee attributes, workplace conditions and external labor market factors (Park et al., 2024). Consequently, as workforce planning becomes increasingly data-driven, traditional turnover models face growing challenges in delivering actionable and precise insights, reinforcing the need for more sophisticated predictive methodologies.

2.3 Factors affecting Employee Turnover

Employee turnover is influenced by multiple factors, which can be categorized into financial, organizational, demographic and economic aspects. Understanding these determinants is crucial for developing predictive models and implementing retention strategies. While no single factor can fully

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.

PREVIEW