Managing employee turnover: machine learning to the rescue

Owen P. Hall
Graziadio School of Business,
Pepperdine University,
Malibu, CA 90263, USA
Email: ohall@pepperdine.edu

Abstract: Organisations continue to face ongoing employee retention and recruiting challenges, which have become even more acute due to the COVID-19 pandemic. In today’s unstable economy, employee retention is still one of the hot button issues facing many HR managers. Employee turnover has cost organisations billions of dollars each year. The empirical results from the current study, which included employee demographic, preference, and performance data, suggests that machine learning-based predictive models can provide automatic and timely employee assessments, which allow for both the identification of employees that may be planning to leave and the implementation of appropriate amelioration initiatives. Job engagement, work satisfaction, experience, and compensation are but four of the factors found to be closely aligned with an employee’s decision to leave. The primary purpose of this paper is to highlight how machine learning can reduce employee turnover through early detection and intervention.

Keywords: machine learning; human resource management; employee turnover; actionable knowledge discovery; intervention strategies; cost optimisation; market churn; decision trees.

Reference to this paper should be made as follows: Hall, O.P. (2021) ‘Managing employee turnover: machine learning to the rescue’, Int. J. Data Science, Vol. 6, No. 1, pp.57–82.

Biographical note: Owen P. Hall holds the Corwin D. Denney Academic Chair and is a Full Professor of Decision Sciences at Pepperdine Graziadio Business School. He is a Julian Virtue Professor and a Rothschild Applied Research Fellow. He has more than 35 years of academic and industry experience in computer decision systems and learning systems technology. He received the Harriet and Charles Luckman Distinguished Teaching Fellow in 1993, the Sloan-C Effective Teaching Practice Award in 2013, and the Howard A. White Teaching Excellence Award in 2009 and 2017. His current area of research includes the application of artificial intelligent agents to integrated learning systems and cloud-based computing. He is the vice chairman of the INFORMS University Analytics Program Committee. He is a registered professional engineer, State of California and is a member of the Beta Gamma Sigma Honor Society. He received his PhD from the University of Southern California and undertook post-doctoral studies at the Center for Futures Research.

This paper is a revised and expanded version of a presentation entitled ‘Predicting employee turnover using machine learning’ given at the INFORMS Annual Virtual Conference, October, 2020.
1 Introduction

Identifying employees planning to leave the organisation and implementing appropriate retention interventions has become a contentious topic throughout the HR universe. Even as COVID-19 has slowed the economy down many organisations will continue to face increased competition, more demanding customers, and intensified recruiting and onboarding challenges. This combination of forces tends to drive up the costs associated with employee acquisition and retention. The cost of voluntary employee churn can range as high as five times an employee’s annual salary depending on the specific position and market conditions (Sesil, 2014). Until recently, the processes used by most HR organisations have focused on data warehousing, which tends to provide a backward perspective, for example, assessing employee performance after the fact. In contrast, a forward-looking approach based on the application of machine learning can be extremely helpful in detecting employees that may be considering leaving and delivering customised interventions in a timely manner (Bekken, 2019). More specifically, this approach provides a vehicle to incorporate both quantitative and non-quantitative measurements (e.g., Tweets, blogs, surveys) in the turnover assessment process (Kumar et al., 2018). Machine learning models have found increased application in the study of employee turnover (De Winne et al., 2019; Saranya and Devi, 2018; Sikaroudi et al., 2015; Tambde and Motwani, 2019). However, in most instances they are limited to an assessment of model performance and do not address the equally important issue of designing successful interventions.

Machine learning can be used to both identify employees that are planning to leave and design specific implementation amelioration strategies. The expression machine learning was initially coined by Samuel in the late 1950s and is usually defined as a computer’s ability to automatically learn and improve from experience without being explicitly programmed by humans (Samuel, 1959). Machine learning is often considered a subset of artificial intelligence, which is receiving increased interest throughout the HR community of practice (Yano, 2017). The ‘semi-automated’ capability associated with machine learning algorithms is essential considering the large number of risk factors that need to be assessed. Today, the state-of-the-art in machine learning as applied to human resource management has advanced significantly compared to even a decade ago (Alom et al., 2019; Endert et al., 2017). Specific HR domains of interest include ensuring the functional relevance and success of data mining, confirming the provision of suitable data and information systems, certifying compliance with ethical and legal standards, elaborating a systematic overview of functional HR application areas, substantiating the relevance and characteristics of the problems (Strohmeier and Piazza, 2013).

Employee churn, a knockoff expression from customer churn used throughout marketing, provides a useful context for better understanding turnover, where employees, in many cases, are an organisation’s biggest asset (Dolatabadi, 2017). In churn marketing studies, machine learning has seen considerable technological advances, including the classification of data based on the near level of certainty (Amin et al., 2019; Garcia et al., 2017; Zhu et al., 2017). These same techniques are now being applied to the analysis of employee churn (Ekawati, 2019). This analytical approach can help reduce bias from the turnover assessment, recruiting and training processes and can improve cultural climate and organisational diversity (Altemeyer, 2019). Nevertheless, the use of the employee churn paradigm in HR applications is not without its critics. Listed in the following are some specific issues (Tambe et al., 2019; Van Den Heuvel and Bondarouk, 2017):
Managing employee turnover: machine learning to the rescue

- ethical and legal concerns, which generally require that decisions have explanations – a significant problem associated with some machine learning algorithms
- relatively small sample sizes at any given point in time, which can limit the number of key assessment factors
- employee blowback associated with the idea of being evaluated, even partially, by a machine.

However, for this study the focus is exclusively on the process for identifying employees at risk and formulating customised intervention strategies, which mitigates some of the issues listed above.

This paper is organised as follows: a review of the employee turnover literature, an introduction to machine learning, an example application, and an overview of retention interventions based on cost minimisation optimisation strategies. This paper’s primary contribution to the HR universe is the integration of machine learning-based employee detection models with the design of effective retention interventions. The expression ‘employees at risk’ will be used to characterise employees that are considering leaving their current employer.

In my experience, people don’t leave their organizations; they leave their managers. – Maureen Swick, CEO of the American Organization of Nurse Executives.

2 Detecting employees at risk

Understanding why employees leave or stay continues to be one of the most important issues facing organisations of all sizes and shapes. Over the years, this topic has received considerable attention from the academic community (Hom et al., 2017; Lee et al., 2017). These concerns are not limited to the USA but are seeing increasing interest on a worldwide basis. In Europe there appears to be a long-term trend toward shorter job tenure for certain age groups and that tenure varies strongly according to job characteristics (Bachmann and Felder, 2018). Furthermore, employees higher in extraversion, a measure of the extent to which an employee engages with the external world are more likely to change jobs (Bartolec, 2018). In Asia, job embeddedness, the collection of factors that influence employee retention, was found to be a stronger predictor of voluntary employee turnover compared to job attitudes and perceived ease of movement (Chunjiang et al., 2011; Xuran and Hongbo, 2019). Additionally, in a related study the data showed that organisational identity and employee satisfaction was significantly correlated with turnover intention (Si, 2019).

Recently, there has been a growing interest in the use of analytics to help address the voluntary employee turnover challenge (O’Keefe, 2016; Omar, 2019). This awakening is being driven by the immediate need for effective retention strategies coupled with machine learning’s ability to identify, with increasing accuracy, specific employees that are planning to leave the organisation (Choi and Fernandez-Blanco, 2017). Identifying attrition risk calls for advanced pattern recognition in surveying an array of candidate predictor variables. Advances in machine learning, which include the ability to process large amounts of unstructured data, can lead to the characterisation of employee emotional activity via, for example, social networking. The task of discovering
employees at risk has a strong association with the goal of identifying students at academic risk early in the matriculation process (Seidman, 1996). Specifically, Seidman proposed the following relationship, which links student retention to both early detection and continuous intervention.

\[ \text{Student Retention} = \text{Early Detection} + \text{Continuous Intervention} \]

This construct emphasises that early identification of students at risk as well as continuous intervention can be key to student retention. This same paradigm can be applied to the task at hand, namely employee retention. The impact of employee attrition should be viewed from a holistic perspective, including financial, developmental, and future advancement. Continuous interventions can take on many forms, for example, training and compensation. Today, HR managers wish to possess the ability to (Chapman et al., 2005; Khera, 2019):

- identify employees preparing to leave
- detect employees engaging in anti-organisational behaviour (e.g., fraud)
- locate employees exhibiting low motivation and find alternate means of reaching them
- organise employees into groups with similar characteristics
- predict probable employee performance outcomes
- design and deploy optimal recruiting strategies
- prepare organisational succession plans.

According to recent findings, when compensation, performance appraisal, career promotion, and training and development programs are satisfactory to employees and reflect the intrinsic and extrinsic aspects of career satisfaction, then employees are less likely to leave (Aburumman et al., 2020). Additional assessment data reveals that there is a significant relationship between both employees’ job satisfaction and organisational commitment on turnover intentions (Azeez et al., 2016).

Work-life balance has been one of the guiding core factors in HR orthodoxy for some time. Traditionally, work-life balance has been characterised as the amount of time spent working compared with the amount of time engaged in other pursuits (e.g., family). This definition leaves much to be desired since the notion of quality is not evident (i.e., quality-time). This is an area where data mining can be applied. Specifically, through data mining work strategies can be developed to relieve employees of monotonous and repetitive tasks, thus allowing more opportunities for self-realisation and increased engagement. These developments, in turn, could lead to an enhancement of employees’ work-life balance (Pawlicka et al., 2020). The onslaught of COVID-19 has further exacerbated the natural tensions surrounding the work-life balance construct. Today, a more common theme is work-life integration, which creates synergies between various life constructs, such as job, faith, family, health, community, and leisure. This approach emphasises gentle pivots rather than hard boundaries between the different ingredients of life (Shanafelt et al., 2019).

Table 1 provides a list of risk factors, which were identified based on the above discussion and a seminal meta-analysis employee turnover paper from the mid-1980s.
Managing employee turnover: machine learning to the rescue (Cotton and Tuttle, 1986). Cotton’s original taxonomy consisted of three broad correlates of turnover categories:

- environmental
- personal
- work related.

For the current study, a fourth category, performance, was added, which is designed to characterise the accomplishments of both the individual employee as well as the organisation. Additionally, government policy, political climate, and technology were added to the environmental category since they are playing an ever-increasing role in both society and employment. To that end, the impact of the COVID-19 pandemic is another factor that should be considered in future research. Cotton found that age, tenure, pay, and job satisfaction were both stable and reliable correlates with turnover.

Table 1  Candidate employee risk detection categories and factors

<table>
<thead>
<tr>
<th>Environmental</th>
<th>Personal</th>
<th>Work perspective</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>State of economy</td>
<td>Tenure</td>
<td>Work satisfaction</td>
<td>Assessment rating</td>
</tr>
<tr>
<td>Accession rate</td>
<td>Gender</td>
<td>Job satisfaction</td>
<td>Salary increase</td>
</tr>
<tr>
<td>Government policy</td>
<td>Marital status</td>
<td>Work life balance</td>
<td>Net profit margin*</td>
</tr>
<tr>
<td>Political</td>
<td>Age</td>
<td>Job involvement</td>
<td>Last promoted</td>
</tr>
<tr>
<td>Technology</td>
<td>Education level</td>
<td>Relationship satisfaction</td>
<td>Stock price*</td>
</tr>
</tbody>
</table>

*For non-profit organisations, fundraising ROI and landing page conversion rate could be substitutes.

“Many experts are of the view that employees are the staying power of any organization so organizations necessitate taking initiative to implement the employees’ motivation process, thereby enhancing the overall employees’ performance by providing quality products and offering excellent services.” (Mamun and Hasan, 2017)

3 Machine learning

Machine learning is the science of discovering and communicating meaningful patterns in data and supporting the development of actionable plans. Today, machine learning is seeing increased use throughout the HR community of practice. Two of the more popular models in this regard are neural nets (NN) and decision trees (DT). NN have seen considerable use in employee data mining studies (Fazlollahtabar et al., 2016; Jantan et al., 2009; Panchal et al., 2010; Sexton et al., 2005). There is actually a family of DT algorithms which include classification and regression trees (CART), Random Forest trees (RFT), and extreme gradient boosting trees (XGB). RFT utilise an ensemble strategy that combines weaker learners into stronger learners, which represents a technical improvement over the basic CART model. XGB takes the RFT process one step farther where a series of trees are constructed sequentially, and each tree attempts to correct the errors of the previous tree in the series. XGB often tends to outperform other machine learning models in many classification applications (Dinh et al., 2019;
Mihăiță et al., 2019). Like neural nets the downside of both RFT and XGB, compared to the basic CART model, is the difficulty in visualising and interpreting the results. The DT family has also seen considerable use in human resource analysis (Alao and Adeyemo, 2013; Chen and Guestrin, 2016; Rombaut and Guerry, 2015). Again, the purpose of this study is not to evaluate the effectiveness of various machine learning algorithms, per se, but to illustrate how these systems can be used in a HR setting, specifically to detect employees at risk.

The confusion matrix, illustrated in Table 2 for a binary classifier, is one of the standard formats for assessing model performance. The confusion matrix compares predictions with the actual observed conditions. For example, the statistic sensitivity, also called recall, measures the proportion of actual positives that are correctly identified, while specificity reports the proportion of actual negatives that are correctly characterised.

<table>
<thead>
<tr>
<th>Actual condition</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict positive</td>
<td>TP</td>
<td>FP</td>
</tr>
<tr>
<td>Predict negative</td>
<td>FN</td>
<td>TN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP/(TP+FN)</td>
<td>TP/(TP+FN)</td>
<td>TN/(FP+TN)</td>
<td>(TP+TN)/ T</td>
</tr>
</tbody>
</table>

1TP = True positive, TN = True negative, FP = False positive, FN = False negative, T = TP+TN+FP+FN; ²PPV = Positive predictive value; ³NPV = Negative predictive value.

In the context of this study, a positive predictive value is the probability that an employee classified at risk is actually at risk. In contrast, a negative predictive value is the probability that an employee was classified as not at risk when they actually are not at risk. In many studies, the metric Accuracy is used to both judge the relative performance of the various candidate classification models and as a standard for selecting the ‘best’ model for subsequent usage. The receiver operating characteristic (ROC) curve, which is illustrated in Figure 1, represents another method to evaluate model performance. The ROC graph, featuring sensitivity (true positive rate) on the vertical axis and one minus specificity (false positive rate) on the horizontal axis, is designed to illustrate the diagnostic ability of a binary classifier as its discrimination threshold is varied. The random line represents the so-called ‘line of no discrimination’, which is equivalent to coin flipping. Points above the random line indicated better classification performance compared to the random process. Specifically, the curve at the upper left reports the performance of an example classifier model. Clearly, this model has outperformed the random line in correctly classifying the two categories. The ideal performance case is when the model line rises vertically from the origin to the top of the Y-axis and then horizontally to the end of the X-axis. Typically, the classifier generates a statistic called the area under curve (AUC), which provides a numerical estimate of model classification performance. This statistic ranges in value between zero and one. An AUC value of one suggests perfect discrimination, while a value of 0.5 indicates a random process (i.e., coin flipping). The AUC graphic can also be used to compare the performance of several
Managing employee turnover: machine learning to the rescue

classification models by plotting their ROC curves on the same diagram and comparing the corresponding AUCs.

In some classification applications the importance of one category can be much higher than the importance of other categories. For example, in the current pandemic the capability of detecting individuals with COVID-19 is much more important than identifying those individuals that are not infected. The standard approach in these instances is to assign different costs to the various errors in proportion to their significance. This technique is called cost-sensitive classification (Elkan, 2001; Fatlawi, 2017). In this context of COVID-19, ‘high’ costs are assigned to misidentifying the minority class while ‘low’ costs are allocated to misclassifying the majority class.

“IBM artificial intelligence can predict with 95% accuracy which workers are about to quit their jobs.” (Rosenbaum, 2019)

Figure 1  Example ROC curve

4 Illustrative example

To illustrate the process outlined above, personal, work related, and performance data on 1470 employee records associated with a high-tech, healthcare-oriented organisation were assessed (Subhash, 2017). This database has already been evaluated using a variety of machine learning techniques (Zhao et al., 2019). The findings indicated that the neural nets and extreme gradient boost trees were the top two performers. Table 3 highlights the various model variable mnemonics and corresponding descriptive statistics for the assembled database. The environmental category highlighted in Table 1 was not included since information on these factors was not part of the original database. For this study, a binary target scheme was employed, which characterised whether an employee left the organisation over the course of the given timeframe. The data shows that approximately 16% of the employees left the organisation during this period. This relatively small proportion signals that there is an imbalance challenge associated with the current database (Li et al., 2013). There are a variety of different approaches that can be used to
help address the imbalance problem (Krawczyk, 2016; Longadge et al., 2013; Mahajan et al., 2020; Wang, 2018). The over-sampling method was adopted for the current study, where the proportion of the minority class was increased to 48%. Additionally, both the continuous and ordinal variables were normalised using the min-max procedure. In some instances, the continuous variables can be rescaled using the standardisation method, which retains more information. However, for this database each of the continuous variables were highly skewed as measured by the Kolmogorov-Smirnov test (e.g., employee age). The candidate study variables reported in Table 3 were grouped by the corresponding measurement scales: continuous, ordinal, and binary (nominal).

For example, the average employee age was 38, the average work environment satisfaction score was 2.74 (1 to 4 scale), and 40% of employees were women.

Table 3  Selected variable mnemonics and descriptive statistics ($N = 1470$)

<table>
<thead>
<tr>
<th>Mnemonic</th>
<th>Definition</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Continuous</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE</td>
<td>Chronological age</td>
<td>36.92</td>
<td>9.14</td>
</tr>
<tr>
<td>YAC</td>
<td>Years at firm</td>
<td>7.01</td>
<td>6.13</td>
</tr>
<tr>
<td>DFH</td>
<td>Distance from home</td>
<td>9.19</td>
<td>8.18</td>
</tr>
<tr>
<td>MIC</td>
<td>Monthly income</td>
<td>6502</td>
<td>4707</td>
</tr>
<tr>
<td>NCW</td>
<td>Number of companies worked</td>
<td>2.69</td>
<td>2.50</td>
</tr>
<tr>
<td>PSH</td>
<td>Percent salary increased</td>
<td>15.21</td>
<td>3.66</td>
</tr>
<tr>
<td>TWY</td>
<td>Years of working experience</td>
<td>11.28</td>
<td>7.78</td>
</tr>
<tr>
<td>TTY</td>
<td>Number of training events last year</td>
<td>2.80</td>
<td>1.29</td>
</tr>
<tr>
<td>YCR</td>
<td>Years in current role</td>
<td>4.23</td>
<td>3.62</td>
</tr>
<tr>
<td>YLP</td>
<td>Years since last promotion</td>
<td>2.19</td>
<td>3.22</td>
</tr>
<tr>
<td>YCM</td>
<td>Years with current manager</td>
<td>4.12</td>
<td>3.57</td>
</tr>
<tr>
<td><strong>Ordinal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WES</td>
<td>Work environment Satisfaction (1–4)</td>
<td>2.72</td>
<td>–</td>
</tr>
<tr>
<td>JIN</td>
<td>Job involvement (1–4)</td>
<td>2.73</td>
<td>–</td>
</tr>
<tr>
<td>JOS</td>
<td>Job satisfaction (1–4)</td>
<td>2.73</td>
<td>–</td>
</tr>
<tr>
<td>PER</td>
<td>Performance rating (1–4)</td>
<td>3.15</td>
<td>–</td>
</tr>
<tr>
<td>RES</td>
<td>Relationship satisfaction (1–4)</td>
<td>2.71</td>
<td>–</td>
</tr>
<tr>
<td>WLB</td>
<td>Work life balance (1–4)</td>
<td>2.76</td>
<td>–</td>
</tr>
<tr>
<td><strong>Binary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRF</td>
<td>Travel (frequent = 0, infrequent = 1)</td>
<td>0.81</td>
<td>–</td>
</tr>
<tr>
<td>GND</td>
<td>Gender (female = 1, male = 0)</td>
<td>0.40</td>
<td>–</td>
</tr>
<tr>
<td>OPT</td>
<td>Stock option plan (yes = 1, no = 0)</td>
<td>0.57</td>
<td>–</td>
</tr>
<tr>
<td>MAR</td>
<td>Marital status (married = 1, no = 0)</td>
<td>0.46</td>
<td>–</td>
</tr>
<tr>
<td>MAN</td>
<td>Manager (yes = 1, no = 0)</td>
<td>0.07</td>
<td>–</td>
</tr>
<tr>
<td>MTB</td>
<td>Medical education (yes = 1, no = 0)</td>
<td>0.32</td>
<td>–</td>
</tr>
<tr>
<td>OVT</td>
<td>Overtime (yes = 1, no = 0)</td>
<td>0.28</td>
<td>–</td>
</tr>
<tr>
<td>ELO$^1$</td>
<td>Employee left organisation (yes = 1, no=0)</td>
<td>0.16</td>
<td>–</td>
</tr>
</tbody>
</table>

$^1$Original proportion (minority class).
|       | AGE  | YAC  | DFH  | MIC  | NCW  | PSH  | WRE  | TTY  | YCR  | YLP  | YCM  | ELO  |
|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| AGE   | 1.0  |      |      |      |      |      |      |      |      |      |      |      |      |
| YAC   | 0.311** | 1.0  |      |      |      |      |      |      |      |      |      |      |      |
| DFH   | -0.002 | 0.010 | 1.0  |      |      |      |      |      |      |      |      |      |      |
| MIC   | 0.498** | 0.514** | -0.017 | 1.0  |      |      |      |      |      |      |      |      |      |
| NCW   | 0.300** | -0.118** | -0.029 | 0.150** | 1.0  |      |      |      |      |      |      |      |      |
| PSH   | 0.004 | -0.036 | 0.040 | -0.027 | -0.010 | 1.0  |      |      |      |      |      |      |      |
| WRE   | 0.680** | 0.628** | 0.005 | 0.773** | 0.238** | -0.021 | 1.0  |      |      |      |      |      |      |
| TTY   | -0.020 | 0.004 | -0.037 | -0.022 | -0.066* | -0.005 | -0.036 | 1.0  |      |      |      |      |      |
| YCR   | 0.213** | 0.759** | 0.019 | 0.364** | -0.091** | -0.002 | 0.460** | -0.006 | 1.0  |      |      |      |      |
| YLP   | 0.217** | 0.618** | 0.010 | 0.345** | -0.037 | -0.022 | 0.405** | -0.002 | 0.548** | 1.0  |      |      |      |
| YCM   | 0.202** | 0.769** | 0.014 | 0.344** | -0.110** | -0.012 | 0.459** | -0.004 | 0.714** | 0.510** | 1.0  |      |      |
| ELO   | -0.150** | -0.134** | 0.078** | -0.160** | 0.043 | -0.013 | -0.171** | -0.059* | -0.161** | -0.033 | -0.156** | 1.0  |

Correlation significant at the *0.05, **0.01 level
Table 4 reports the Pearson correlation coefficients for the continuous variables (non-normalised) along with the binary target variable ELO. For example, there is a small, inverse association between the length of time at the current firm (YAC) and ELO ($r = -0.134, p < 0.01$). This suggests that employees with more tenure at the current firm tend to be less inclined to leave. Interestingly, there is a strong positive association between years with the current firm and years in their current role YCR ($r = 0.759, P < 0.01$).

Table 5 reports the Kendall-Tau correlation coefficients for the ordinal predictor variable set (non-normalised) and the binary target variable ELO. For example, there was a weak, inverse correlation between job involvement (JIN) and employee left firm (ELO), which was statistically significant ($\tau = -0.114, p \leq 0.001$). The remaining correlations with ELO were less than 0.10, although three were significant at the 0.01 level (WES, JOS, WLB).

**Table 5** Correlation matrix (Kendall-Tau)

<table>
<thead>
<tr>
<th></th>
<th>WES</th>
<th>JIN</th>
<th>JOS</th>
<th>PER</th>
<th>RES</th>
<th>WLB</th>
<th>ELO</th>
</tr>
</thead>
<tbody>
<tr>
<td>WES</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JIN</td>
<td>-0.013</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOS</td>
<td>-0.003</td>
<td>-0.011</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PER</td>
<td>-0.027</td>
<td>-0.023</td>
<td>0.006</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RES</td>
<td>0.005</td>
<td>0.033</td>
<td>-0.012</td>
<td>-0.030</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WLB</td>
<td>0.024</td>
<td>-0.018</td>
<td>-0.026</td>
<td>0.006</td>
<td>0.015</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>ELO</td>
<td>-0.088**</td>
<td>-0.114**</td>
<td>-0.094**</td>
<td>0.003</td>
<td>-0.039</td>
<td>-0.049**</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Correlation significant at the *0.05, **0.01 level.

**Figure 2** Work environment satisfaction vs. turnover

![Work environment satisfaction vs. turnover](image)

Figure 2 characterises the relationship between work environment satisfaction and turnover. This data shows that approximately 25% of the employees that reported low
work environment satisfaction left the organisation compared to approximately 15% for the three groups reporting higher work environment satisfaction. A more dramatic trend featuring job involvement and turnover is given in Figure 3. Here, about one third of the employees reporting low job involvement left compared to under 10% for those reporting a very high level of job interest.

![Figure 3 Job involvement vs. turnover](image)

The standard approach in applying machine learning models is to divide the database into two sets (training and testing). Typically, a minimum sample size of 50 to 100 observations per predictor variable is suggested to support this two-stage method. Often, 70% of the data is used for training the model and the remaining 30% to test the model (Dobbin and Simon, 2007; Korjus et al., 2016). This was the approach taken in the current study. This technique tends to ameliorate the impact of overfitting, which often results in over-optimistic model performance (Bleidorn and Hopwood, 2019; Kosinski et al., 2016).

The modelling approach used in the present study was to characterise the target variable (ELO) as binary. However, an even more instructive approach is to identify employees at various levels of risk by, for example, introducing an ordinal target variable with three categories consisting of: 0 = low risk, 1 = moderate risk, and 2 = high risk. The benefit of this method would permit more targeted interventions based on the risk level. However, this expanded model design would require an even larger databases and the interpretation of the confusion matrix becomes somewhat problematic (Indria et al., 2015; Jianfeng et al., 2020; Kanaris et al., 2016). The sample size of 1470 associated with the current study was adequate given the number of candidate variables and the binary nature of the target variable (Park and Yu, 2018; Zavorka and Perrett, 2014). The two machine learning models highlighted above, neural nets and extreme gradient boosting trees, were employed to evaluate the database. Table 5 provides a comparison of the performance of the two classifiers. The numbers in the body of the table represent frequency count, for example, XGT correctly classified 210 employees, out of a total of
219 cases, as leaving the organisation that actually left. The corresponding sensitivity estimate is nearly 96%.

Table 5: Comparison of XGT and NN classification results (XGT/NN) – testing data

<table>
<thead>
<tr>
<th>Predict ELO</th>
<th>Actual ELO</th>
<th>Actual ERO</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>210/201</td>
<td>37/53</td>
<td>247/254</td>
<td>85.0/79.1</td>
<td>PPV³</td>
</tr>
<tr>
<td>Predict ERO</td>
<td>9/6</td>
<td>185/181</td>
<td>194/187</td>
<td>95.4/96.8</td>
</tr>
<tr>
<td>Total</td>
<td>219/207</td>
<td>222/234</td>
<td>441/441</td>
<td>89.6/86.6</td>
</tr>
<tr>
<td>%</td>
<td>95.9/97.1</td>
<td>83.3/77.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹Employee left organisation (ELO), ²Employee remained with organisation (ERO), ³Positive predictive value, ⁴Negative predictive value.

Table 6 summarises the two classifier’s performance based on overall Accuracy and AUC. These results suggest that the extreme boost model somewhat outperformed the neural net model based on both the Accuracy and AUC statistics using the same variable sets as reported in Table 7.

Table 6: Summary of classification accuracy by model (XGT/NN)

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>XGT</td>
<td>89.6</td>
<td>0.94</td>
</tr>
<tr>
<td>NN</td>
<td>86.6</td>
<td>0.90</td>
</tr>
</tbody>
</table>

Table 7: Top predictor variables relative importance

<table>
<thead>
<tr>
<th>Factor</th>
<th>XGT Weight</th>
<th>Factor</th>
<th>NN Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSA</td>
<td>100</td>
<td>OVT</td>
<td>100</td>
</tr>
<tr>
<td>AGE</td>
<td>73</td>
<td>TWY</td>
<td>96</td>
</tr>
<tr>
<td>TWY</td>
<td>53</td>
<td>YCM</td>
<td>92</td>
</tr>
<tr>
<td>YAC</td>
<td>52</td>
<td>YLP</td>
<td>74</td>
</tr>
<tr>
<td>OVT</td>
<td>52</td>
<td>AGE</td>
<td>71</td>
</tr>
<tr>
<td>YCM</td>
<td>49</td>
<td>OPT</td>
<td>64</td>
</tr>
<tr>
<td>YLP</td>
<td>49</td>
<td>YAC</td>
<td>48</td>
</tr>
<tr>
<td>WES</td>
<td>46</td>
<td>JIN</td>
<td>47</td>
</tr>
<tr>
<td>JOS</td>
<td>46</td>
<td>MSA</td>
<td>43</td>
</tr>
<tr>
<td>OPT</td>
<td>43</td>
<td>JOS</td>
<td>38</td>
</tr>
<tr>
<td>JIN</td>
<td>38</td>
<td>WES</td>
<td>12</td>
</tr>
</tbody>
</table>

Determining the number of predictor variables to incorporate in the final model is often based on the principle of parsimony (Occam’s razor). This principle, as applied to machine learning, suggests that both overfitting and complexity can be minimised by the judicious selection of the variable set (Dresp-Langley et al., 2019). To that end,
Table 7 identifies the relative importance of the top 11 predictor variables for the two classification models.

Monthly salary showed the biggest shift in relative importance between the two models, where it is ranked first in the XGT model and ninth in the NN model. The fact that both job satisfaction and involvement appeared in the top 11 factors listed in Table 7 while co-worker satisfaction did not, is consistent with the Cotton findings reported earlier. A Spearman test of the two variable ranks (XGT and NN) revealed no association ($r = 0.31, p$-value $= 0.355$), which indicates that the two variable rankings are different. The dataset was further explored using the stepwise feature of XGT model. The incorporation of additional variables did not significantly increase the performance accuracy of the model.

The above results are consistent with the results gleaned from similar studies (Chan et al., 2016; Huang and Su, 2016; Hongxi, 2019). Specifically, career adaptability (JIN), promotability (YLP) and career satisfaction (JOS) were significantly associated with turnover intention. Compensation also plays a role in an employee’s decision to remain or leave the organisation. This role is reflected in the results reported in Table 7. Today, compensation is being used directly as a retention technique and as a vehicle for communicating to the workforce that they are the partners in the growth of the organisation (Sarkar, 2018). The age of the employee (AGE) also contributed to explaining turnover. As employees age the chances of them leaving the organisation tend to decrease (Rombaut and Guerry, 2018). Employees engaged in overtime tend to be more inclined to leave as a result of work-life imbalances and job stress (Razig and Maulabakhsh, 2015; Sin, 2018). This observation is borne out in Table 7, where OVT was ranked high using both modelling systems. Working environment satisfaction (WES), while not ranked as high, does suggest a link to turnover. A recent study on this topic indicated that relationship conflict is negatively related to task performance, contextual performance, and turnover intentions (Shaukat et al., 2017). Further results indicate that supervisory behaviour, job characteristics and work-life balance have a significant negative relationship with turnover intention (Surienty et al., 2014). The modelling performance levels achieved in the present study were found to be comparable with earlier research involving predicting employee turnover using a database of similar size and characteristics (Gao et al., 2019). Specifically, the AUC and Accuracy for this study were 0.85 and 92.7%, respectively. Similar levels of performance using XGT (AUC $= 0.85$) were found when processing a much larger database ($N = 73,115$) involving employee turnover (Punnoose and Ajit, 2016). Once employees have been identified as potential turnover candidates, the next step is to design and implement specific intervention strategies, which is the subject of the next section.

“Many of these technological advancements are bringing greater insight into the manager-employee connection, which fosters authentic connection and better understanding for this crucial workforce relationship.” (Rogers, 2018)

5 Intervention strategies

The above example illustrates how machine learning can be used to identify employees at risk. Discovering workers that are considering leaving is one thing; taking steps to mitigate turnover is quite another. The next step is to design early-on intervention strategies that can ameliorate an employee’s desire to leave. Often, many machine
learning exercises end after the development of the detection model along with the selected predictor variables. One of the factors contributing to this situation is that considerable human expertise has frequently been required in the post machine learning phase as a vehicle for identifying specific rules and corresponding interventions. As previously discussed, machine learning has seen increased use in the customer relationship management field to identify specific actions that would transform a fickle customer to one loyal to the organisation (Karim and Rahman, 2013; Subramani and Balasubramaniam, 2016). Typically, the goal of these post-data mining decision support systems in a market churn context is to identify actions that optimise the chances of retaining a customer. Usually, net profit is calculated for each viable action and the alternative with the largest net profit value is selected for implementation. A classic example, in this regard, is to increase the customer’s level of service, which typically would have very little cost impact. Actionable knowledge discovery, an expression used for describing the use of machine algorithms to assist in identifying cost-effective interventions, was initially introduced in the late 2000s (Yang et al., 2008).

In employee turnover applications a more appropriate objective is to minimise total cost. This is because it is much more challenging to associate an employee or class of employees to profit. The cost approach has the added benefit that it can be applied to non-profit organisations. Basically, the two costs involved are replacement and retention. In its simplest form the task at hand turns out to be an unconstrained cost minimisation problem, where the trade-off is to balance replacement costs with retention costs. Normally, as the number of interventions grow, the replacement costs decrease, while retention costs increase as illustrated in Figure 4. Model attributes can be characterised as either being stable or flexible. Stable attributes are ones which cannot be changed at any reasonable cost (e.g., gender), while flexible attributes can be changed, (e.g., pay raise). The actionable knowledge discovery decision tree process requires at least one flexible attribute. Unconstrained cost minimisation problems have been studied extensively (Hayashi et al., 2020; Zhou et al., 2017).

Figure 4  Cost trade-off relationships
The objective function can be expressed as follows:

$$
\min F = \sum U_i \cdot (P_i - Q_i) + \sum \sum T_{ij} \cdot X_{ij}
$$

(1)

where

- $P_i$: probability of $i$th employee leaving pre-intervention
- $U_i$: cost of replacing $i$th employee
- $Q_i$: difference in probability of $i$th employee leaving between pre and post intervention(s)
- $T_{ij}$: cost of retaining $i$th employee with $j$th intervention
- $X_{ij}$: $j$th intervention assigned to $i$th employee.

Obviously, from this formulation an employee can be assigned more than one intervention. Constraints can be added to this basic model in the form of budget restraints and limits on the number of specific interventions. For example, there may be restrictions on the number of employees that can be taken off overtime due to production/service demand requirements.

To illustrate the rule generation process, a basic CART analysis of the database was performed using a subset of the classification predictor variables reported in Table 7. Figure 5 illustrates the resultant actionable decision tree. In this small-scale example, notice that there are both flexible attributes (working overtime, received training, receiving a raise, and being engaged in the stock purchase plan) and stable attributes (years at the firm). The numerical values in the diagram (i.e., leaves) represent the probability that the employee is planning to leave the organisation. Clearly, employees that are working overtime, that have received less than three training sessions per year, and that have been with the organisation for less than three years are at most risk (i.e., there is a 54% chance that individuals that meet these specifications are planning to leave). More specifically, they are twice as likely to leave than employees with three or more years of service with the organisation. At this point HR has several actionable (flexible) options: remove the employee from overtime, provide the employee with additional training, increase the size of the employee's raise, and offer the employee membership into the stock purchase plan.

For example, increasing the number of training sessions from less than three per year to three or more per year reduces the chances of the employee leaving from 54% to 27%.

One of the challenges associated with utilising actionable discovery decision trees is the conversion of continuous attributes to discrete intervals (e.g., tenure). The CART analysis was used to determine the resultant actionable decision tree. In this small-scale example, notice that there are both flexible attributes (working overtime, received training, receiving a raise, and being engaged in the stock purchase plan) and stable attributes (years at the firm). The numerical values in the diagram (i.e., leaves) represent the probability that the employee is planning to leave the organisation. Clearly, employees that are working overtime, that have received less than three training sessions per year, and that have been with the organisation for less than three years are at most risk (i.e., there is a 54% chance that individuals that meet these specifications are planning to leave). More specifically, they are twice as likely to leave than employees with three or more years of service with the organisation. At this point HR has several actionable (flexible) options: remove the employee from overtime, provide the employee with additional training, increase the size of the employee's raise, and offer the employee membership into the stock purchase plan.

For example, increasing the number of training sessions from less than three per year to three or more per year reduces the chances of the employee leaving from 54% to 27%.

One of the challenges associated with utilising actionable discovery decision trees is the conversion of continuous attributes to discrete intervals (e.g., tenure). The CART analysis was used to determine the resultant actionable decision tree. In this small-scale example, notice that there are both flexible attributes (working overtime, received training, receiving a raise, and being engaged in the stock purchase plan) and stable attributes (years at the firm). The numerical values in the diagram (i.e., leaves) represent the probability that the employee is planning to leave the organisation. Clearly, employees that are working overtime, that have received less than three training sessions per year, and that have been with the organisation for less than three years are at most risk (i.e., there is a 54% chance that individuals that meet these specifications are planning to leave). More specifically, they are twice as likely to leave than employees with three or more years of service with the organisation. At this point HR has several actionable (flexible) options: remove the employee from overtime, provide the employee with additional training, increase the size of the employee's raise, and offer the employee membership into the stock purchase plan.

For example, increasing the number of training sessions from less than three per year to three or more per year reduces the chances of the employee leaving from 54% to 27%.

One of the challenges associated with utilising actionable discovery decision trees is the conversion of continuous attributes to discrete intervals (e.g., tenure). The CART analysis was used to determine the resultant actionable decision tree. In this small-scale example, notice that there are both flexible attributes (working overtime, received training, receiving a raise, and being engaged in the stock purchase plan) and stable attributes (years at the firm). The numerical values in the diagram (i.e., leaves) represent the probability that the employee is planning to leave the organisation. Clearly, employees that are working overtime, that have received less than three training sessions per year, and that have been with the organisation for less than three years are at most risk (i.e., there is a 54% chance that individuals that meet these specifications are planning to leave). More specifically, they are twice as likely to leave than employees with three or more years of service with the organisation. At this point HR has several actionable (flexible) options: remove the employee from overtime, provide the employee with additional training, increase the size of the employee's raise, and offer the employee membership into the stock purchase plan.
different expected utilities. To help resolve these conflicts several different techniques have been proposed, including an evaluation of the utility prediction accuracies of conflicting rules (Su et al., 2017; Zeng et al., 2014). Furthermore, risks associated with incomplete rules, those incorporating a subset of the total predictor set, can be minimised through the application of association rules, which among other things do not require the user to identify minimum support thresholds (Sim et al., 2010).

**Table 8**  Selected attribute cost matrices

<table>
<thead>
<tr>
<th>Overtime</th>
<th>Training</th>
<th>Raise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0</td>
<td>3 0</td>
</tr>
<tr>
<td>No</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tenure</th>
<th>Stock</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes 0</td>
<td>No -50</td>
</tr>
<tr>
<td>≥Ave 0</td>
<td>50 0</td>
</tr>
</tbody>
</table>

At this point the economic costs associated with the various attributes can be analysed as a basis for developing the most cost-effective intervention plan. Table 8 highlights some examples for the attributes delineated in Figure 5. The numerical cost values are relative.

**Figure 5**  Example actionable knowledge discovery decision tree

The cost is five for switching an employee off or on to overtime. Notice that several of the attributes are asymmetric. For example, the cost for reducing the number of training sessions that an employee has already experienced is non-sensible and therefore has been assigned a cost of infinity. The situation for the raise attribute is somewhat more complicated since technically an organisation could reduce or completely revoke the pay raise; however, this could be in violation of certain laws and regulations. A negative value in the cost matrix indicates a savings (e.g., taking an employee off the stock option plan saves the organisation money). Armed with this cost data, which normally would be provided from domain experts, and the probability data from Figure 5, the optimisation model highlighted in equation (1) can be solved (Liu et al., 2020).
Managing employee turnover: machine learning to the rescue

This modelling approach can specify an array of interventions based on the characteristics and situation of a given employee. For example, if the model determined that one of the major factors in classifying an employee at risk is excessive work stress, then preventive counselling along with a reduction in workload (e.g., overtime) could be initiated. Excessive work stress can be the result of many factors, including unpredictable schedules, excessive overtime, heavy workloads, and a lack of flexibility in the workplace. In a recent study, work stress was conceptualised through three inter-related dimensions: workload, low rewards, and over commitment. The results show that all three dimensions of work stress are positively related to turnover intentions, and that employees’ perceptions of low rewards and their over commitment negatively affect their trust in managers (Uriesi, 2019).

Employee engagement is another factor that can influence a worker’s desire to change jobs. Historically, engagement was measured after the fact through annual surveys. Like work stress, engagement is related to a variety of constructs, including decision-making authority to perform assigned tasks, training to improve skills and capability, and ability to resolve customer concerns (Smith and Macko, 2014).

A growing body of evidence suggests that nonmonetary factors are potentially important for motivation and productivity and ultimately as a source of meaning (Cassar and Meier, 2018; Martela and Riketti, 2018). Autonomy in decision-making, a feeling of competence, feelings of relatedness, and beneficence are four basic components of nonmonetary factors that contribute to employee wellbeing and job satisfaction. Moreover, some employees have demonstrated a penchant for trading off lower compensation for more meaningful work. Non-profit organisations provide a useful example. Measuring employee engagement in near real time turns out to be a critical task in the retention process (Burnett and Lisk, 2018). A myriad of employee transitional data can now be collected and processed using the state-of-the-art in web-based technology. Specific metrics comprise customer interactions, production times, computer keystrokes, social media, and network tracking.

Management support can also play a role in ameliorating turnover. Specifically, positive supervisory engagement was found to have a constructive impact on employees’ organisational commitment and career satisfaction, which in turn led to reduced turnover (Kang et al., 2015). Some examples of specific leadership actions include (Carter et al., 2019):

- recognising excellence in employee performance
- celebrating teamwork
- communicating and managing change effectively
- providing competitive compensation
- investing in education and training.

Identifying employees with a proactive personality is also helpful in the retention process. Recent results support the assumption that proactive employees carry with them a higher risk of leaving (Lang et al., 2016). One approach to address proactive employees is through high levels of job embeddedness. Having the organisation provide enhanced career opportunities should make staying more attractive than leaving. To this end, some specific countermeasures that an organisation can evoke in combating the turnover challenge include (Zhang, 2016)
improving employee engagement and achievement in work by optimising job design
• establishing training and career development systems
• creating an employee-oriented corporate culture, which regards employees as the most important and most creative resources of the organisation.

Good communication between leadership and teams lies at the heart of this mindset. When leadership effectively communicates expectations and organisational goals, teams are better able to prioritise tasks and complete them in a timely manner. Some specific interventions for enhancing team performance revolve around team training, leadership training, team building, and team debriefing (Lacerenza et al., 2018).

Collective turnover is another facet of the problem that requires HR’s attention. Typically, collective turnover is defined as “the aggregate levels of employee departures that occur within groups, work units, or organizations” (Hausknecht and Trevor, 2011). Interest in the causes and consequences of collective turnover arose from the recognition that rising turnover rates – and the mounting human and social capital losses associated therein – portend a range of negative outcomes, such as poor customer service, reduced sales and profits, higher accident rates, and lower efficiencies. A number of metrics have been used for measuring and monitoring collective turnover, including turnover rate, instability rate, capacity index, and quality (Heavey et al., 2013). In a recent meta-analysis, the results revealed the following key findings (Hancock et al., 2017):

• collective attitudes and perceptions, along with personnel changes, have the most impact on collective turnover
• collective turnover has a negative relationship with overall organisational performance
• managerial turnover strengthens the negative relationship among collective turnover and performance
• contagion effects can accelerate employee turnover.

Figure 6 presents an overview of the proposed machine learning-driven intervention paradigm. The process consists of five integrated elements.

Specific employee qualitative data can include threaded discussion posts, chat message, social media, employee surveys, and employee interviews (Wu et al., 2018).
Managing employee turnover: machine learning to the rescue

Presented below are some general principles for building out the proposed intervention system.

- **Early detection:** Identifying employees early on who are contemplating leaving the organisation lies at the heart of the early detection-intervention paradigm. This approach is consistent with Seidman’s paradigm as applied to business (Ballenger et al., 2011).

- **Employee feedback:** Ongoing employee feedback is an essential feature of the intervention process. Employees should also be surveyed to obtain their views on the organisation, including their job interests (Knesek, 2015).

- **Employer branding:** Enhancing the organisation’s image as ‘best-in-class’ is a strategy that can be used to acquire and retain qualified talent (Tanwar and Prasad, 2016; Yi, 2019).

- **Work schedule flexibility:** Providing more flexibility in the workplace is essential, especially in a post-COVID-19 world. Regulations that promote more employee-centred adjustable working time can help workers ease work-life time conflicts, but also promote worker well-being generally (Golden et al., 2012; Hassan et al., 2019).

- **Personal development:** Opportunities for personal development should be made available to all employees. Training and development generally increase employee performance, which, in turn, enhances the growth and success of the organisation (Oluwaseun, 2018).

  “Transformational leadership not only directly prevents employees from forming intentions to leave but also indirectly does so by cultivating a collaborative culture.” (Sun and Wang, 2017)

6 Conclusions

While the full impact of COVID-19 on the HR community remains cloudy, the challenges associated with employee retention and recruiting continue unabated. This paper outlines how machine learning can be used to identify employees at risk and to develop employee-specific intervention strategies. The proposed risk assessment template, based on the Seidman formula (i.e., early detection plus continuous intervention), combines the following three components: dynamic database, machine learning algorithms, and specific mediation actions. The acquisition and maintenance of detailed employee data (demographic, behavioural, and performance) is a key ingredient in this process. To that end, HR managers can then decide how to enhance retention by mapping these contributions onto a desired process focused on specific outcomes. This process not only builds organisational capacity, but is also highly measurable, and alerts can be generated in near real time, which is essential when dealing with employees at risk.

The results of a machine learning analysis featuring extreme gradient boost trees and neural nets of a representative employee database yielded classification accuracy levels on the order of 90%. The performance of the models, as measured by AUC, revealed similar outcomes. Key predictive factors derived from both models for detecting
employees at risk included: job involvement, job satisfaction, working environment satisfaction, and time with the organisation. Naturally, the challenges associated with false positives and negatives remains an important issue. Also, addressing every employee who may be planning to leave could be extremely costly, therefore assigning a weighting factor, based on performance assessment and nature of the position, for example, could help mitigate this challenge. The second stage of the analysis employed actionable knowledge discovery decision trees in concert with an unconstrained optimisation model as a vehicle for identifying cost-effective interventions.

Engaging senior leadership in the employee risk mitigation paradigm is essential for ensuring success. Collaboration networks can help facilitate the proposed strategy through access to community best practices. Collaboration networks provide the HR community with the opportunity to converge, share, and exchange ideas to drive innovation in employee turnover mitigation. For some HR professionals, employee privacy might be a concern; however, most organisations already have the data, and it is simply a matter of organising and processing the information to achieve the goal of reducing turnover. The methodology outlined in this paper is also applicable to other aspects associated with employees at risk, for example, behavioural and ethical issues. The overall goal is to appreciate that every employee is unique and that no single intervention approach is optimal for every employee. To that end, the leadership can take on concrete steps, such as creating employee success centres and developing outreach initiatives, to improve retention and employee satisfaction.

“A key benefit in using artificial intelligence in HR is that it can help reduce bias, an issue that is becoming increasingly important in business today.” (Gikopulos, 2019)

References


Managing employee turnover: machine learning to the rescue


Managing employee turnover: machine learning to the rescue


Rosenbaum, E. (2019) *IBM Artificial Intelligence can Predict with 95% Accuracy which Workers are about to Quit their Jobs*, Assessed from: https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent-accuracy-which-employees-will-quit.html


Managing employee turnover: machine learning to the rescue


