Employees recruitment: A prescriptive analytics approach via machine learning and mathematical programming

Dana Pessach[a],*, Gonen Singer[b], Dan Avrahami[a], Hila Chalutz Ben-Gal[c], Erez Shmueli[a], Irad Ben-Gal[a]

[a] *Department of Industrial Engineering, Tel-Aviv University, Israel*
[b] *Faculty of Engineering, Bar-Ilan University, Israel*
[c] *Department of Industrial Engineering and Management, Afeka College of Engineering, Israel*

A B S T R A C T

In this paper, we propose a comprehensive analytics framework that can serve as a decision support tool for HR recruiters in real-world settings in order to improve hiring and placement decisions. The proposed framework follows two main phases: a local prediction scheme for recruitments' success at the level of a single job placement, and a mathematical model that provides a global recruitment optimization scheme for the organization, taking into account multilevel considerations. In the first phase, a key property of the proposed prediction approach is the interpretability of the machine learning (ML) model, which in this case is obtained by applying the Variable-Order Bayesian Network (VOBN) model to the recruitment data. Specifically, we used a uniquely large dataset that contains recruitment records of hundreds of thousands of employees over a decade and represents a wide range of heterogeneous populations. Our analysis shows that the VOBN model can provide both high accuracy and interpretability insights to HR professionals. Moreover, we show that using the interpretable VOBN can lead to unexpected and sometimes counter-intuitive insights that might otherwise be overlooked by recruiters who rely on conventional methods.

We demonstrate that it is feasible to predict the successful placement of a candidate in a specific position at a pre-hire stage and utilize predictions to devise a global optimization model. Our results show that in comparison to actual recruitment decisions, the devised framework is capable of providing a balanced recruitment plan while improving both diversity and recruitment success rates, despite the inherent trade-off between the two.

1. Introduction

One of the most challenging and strategic organizational processes is to efficiently hire suitable workforce. A comprehensive study by the Boston Consulting Group has shown that the recruitment function has the most significant impact on companies' revenue growth and profit margins compared to any other function in the field of human resources (HR) [1]. Indeed, poor recruitment decisions may lead not only to low-performing employees but also to increased turnover. Turnover may have a direct impact stemming from employee replacement costs (e.g., interviews and rehiring costs, training and productivity loss, overtime of other employees), as well as indirect effects, such as poor service to clients or a decline in employee morale [2]. Thus, improving organizational recruitment processes by hiring the most suitable candidates

has a significant impact on organizational performance [3,4].

In this study, we propose a data analytics approach, which can be used as a decision support tool for recruiters in real-world settings to improve hiring decisions of candidates to specific positions or jobs. The proposed approach comprises two components: a local prediction model for recruitment success per candidate and job type, and a global optimization model of the recruitment process.

The first part of this study is based on interpretability ML modeling, which provides meaningful insights into the potential recruitments related to the candidate's background features as well as the planned job placement. The output of these models is the probabilities of successful recruitment per employee and job. The second part in this research is based on a mathematical modeling formulation at an organizational level that takes into account multi-objective considerations and

optimize the recruitment process over many candidates and jobs by using the success probability outputs of the ML models.

Previous efforts have been invested in trying to predict recruiters' decisions (e.g., [5,6]). Such prediction models, if accurate enough, may eventually replace the human recruiter and save a considerable amount of resources. Note, however, that recruiters' decisions are inherently subjective, and human intuition plays an important part in recruitments and placements. Hence, using interpretability modeling tools that can enrich and guide recruiters' decisions by insight seems to be a relevant approach, which recently gained popularity and is also known as *explainable artificial intelligence* (XAI) (see, for example, [7]). Another line of work has focused on the post-hire prediction of turnover or performance (e.g., [8]). While such measures are somewhat more objective, post-hire prediction efforts might be too late in certain cases to act upon. Therefore, in this paper, we focus on the pre-hire prediction of performance and turnover as a combined objective measure.

A key property of our approach is the interpretability of predictions, providing a useful explanation of how they are obtained. Apart from the accuracy of the prediction model, users' trust in the model is often directly impacted by how much they can understand and anticipate its behavior [9]. Understanding why the model behaves the way it does may increase users' trust and their potential to act upon its recommendations. This is especially true in decisions that involve human beings' intuition, such as in the case of employees' recruitment and job placement.

To address the prediction task described above, we propose applying the interpretable Variable-Order Bayesian Network (VOBN) model [10,11]. In contrast to other interpretable models such as decision trees, which often suffer from high variance and overfit to the training set, the VOBN model provides an inherent modeling flexibility that reduces such effects. Therefore, it often results in an improved generalization and predictive ability over various test sets. Finally, we show that the VOBN model is also flexible enough for mining significant patterns and insights in HR data.

Nevertheless, recruitment requires not only hiring the highest-potential workforce, but also meeting other organizational objectives. For example, there is a necessity to meet the demand for employees in different departments, the facilitation of diversity in teams and the allocation of the workforce among different departments in a balanced manner. Each of these dimensions may also include numerous points of view: the local point of view of each separate candidate-position pair, the positional point of view and the organizational or regulatory point of view. Given that there are requirements of various stakeholders in the organization, there is a need to balance the trade-offs in this multi-objective scenario. Hence, in the second part of this research, we address the recruitment problem with a global perspective by accounting for the various dimensions and points of view.

We evaluate the proposed method using a unique dataset obtained from a large nonprofit service organization that is highly diversified over roles, accountabilities and job descriptions, with heterogeneous population of employees with diverse backgrounds, geographic locations and levels of socioeconomic status.

The dataset includes a rich feature set of hundreds of thousands of employment cases collected over a decade and represents a wide range of heterogeneous populations. These characteristics enable us to test potentially biased recruitment policies and placement decisions that traditionally may not be tested due to the absence of sufficient data on such large groups in the population.

The results of our evaluation reveal that the proposed prediction approach can perform well in terms of both accuracy and interpretability, despite the inherent trade-off that often exists between the two [9,12]. In addition, we demonstrate how our interpretable approach can be used to extract meaningful insights that may support and benefit the recruiters' decision process. These extracted insights are sometimes counter-intuitive and shed light on the limitations of existing approaches and on the recruiters' intuition, which is limited and biased at times.

Moreover, we demonstrate that it is feasible to predict a successful placement of a candidate to a specific position at a pre-hire stage with a relatively high prediction performance (AUC = 0.73) and then utilize these predictions to devise a global optimization model. Our results show that using the proposed mathematical programming model, we are able to increase diversity (by 40%) while maintaining a high level of recruitment success (decreased by only 1%). Moreover, the results show an improvement of both diversity and recruitment success rates compared to recruiters' actual selections, although these objectives are generally found to be in conflict. The proposed approach can provide recruiters and organizations alike, with an applicable decision support tool for hiring successful candidates while improving organizational recruitment and placement processes and procedures.

This paper is structured as follows. Section 2 reviews the relevant literature. Section 3 describes the proposed analytics framework and the experimental settings. Section 4 describes the results, and finally, Section 5 summarizes and provides some concluding remarks.

2. Background and literature review

We organize the relevant literature review as follows. We first survey the related studies that address predictive analytics in HR and classify them along three core dimensions: functional, data and method. We then review the related topics from the HR literature.

2.1. Functional dimension

In recent years, several preliminary studies have focused on predicting recruiters' decisions [5,6,13–15]. However, imitating the recruiter's decision may not necessarily be the best approach, since they are often affected by highly subjective and potentially inaccurate judgments that preserve, rather than improve, hiring biases. Consequently, there is a need for an objective measure of the actual success of employee recruitment and performance, as well as providing meaningful insights to the recruiters themselves.

Other recent studies have focused on objective measures of successful recruitment based on employee past performance. Some of these studies examined the post-hire prediction of turnover or performance with predictors collected over the employment period [8,16–23]. Note that the prediction of turnover or performance using post-hire data (such as absenteeism, punctuality and performance reviews) may be useful as part of some retention activities but may lead to a late discovery of recruitment errors and may often be too late to act upon [8,24].

In contrast, the potential benefit of the early pre-hire foresight of longer-term employee success may be much higher, saving more financial and social costs. Few studies have addressed the pre-hire prediction of recruitment success using performance assessments [25–27] separately from turnover assessments [25,28].

Measuring performance may incorporate one aspect of the success of an employee; however, high-performers will not necessarily remain in the organization. Moreover, turnover alone may only partially indicate recruitment success — as often happens in practice, low-performers may not leave the organization due to organizational policies to minimize layoffs and promote high internal mobility.

No previous study has referenced the combination of turnover and performance into one measure that represents an objective measure of recruitment success (see Fig. 1 for a taxonomy of the functional dimension). Thus, in this study, we focus on the case of pre-hire predictions of recruitment success using a combined measure. In the rest of this review, we focus mostly on the case of pre-hire predictions of recruitment success. Note that our methodology approaches hiring from the point of view of recruiters, as opposed to other methodologies that examine the perspective of candidates (for example, how they browse or select relevant job positions [29–31]).
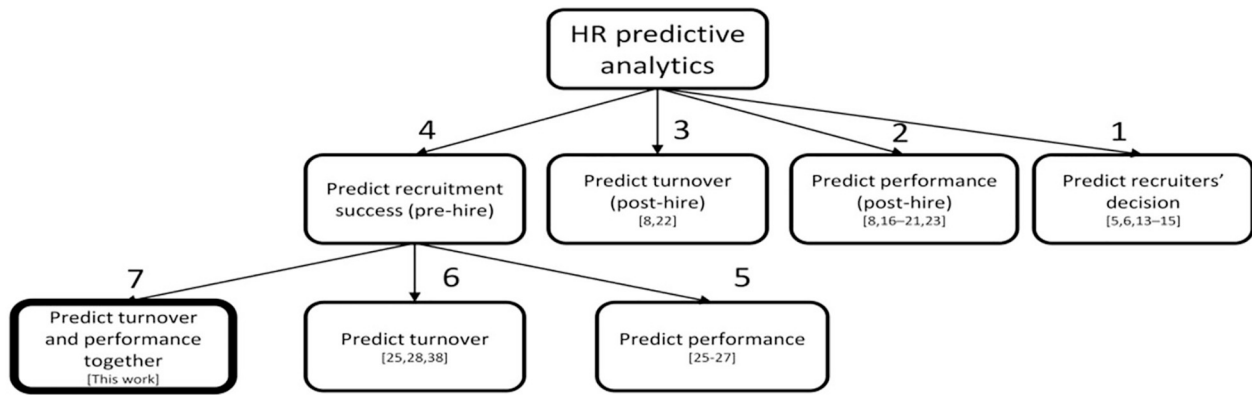
Fig. 1. Literature review based on the functional dimension.

2.2. Data dimension

One of the challenges of using machine learning (ML) techniques in HR is the deficiency of empirical data. A noticeable number of studies have examined rather small datasets, in terms of both the number of candidates, as well as the number of features (e.g. [8,15,23,26,32]).

Within the line of studies that have addressed pre-hire prediction, studies traditionally included a rather narrow set of samples (such as [25–27]). However, in most cases, a small dataset fails to adequately portray the characteristics of the population, yielding the challenge to adequately train a reliable model based on such a small dataset. Narrow datasets often result in low support values of subpopulations, meaning that very few samples are associated with each predicted (or rule-based) subpopulation, resulting in low statistical significance. This challenge is even more noticeable with the growth in the number of features.

Some studies have also involved a limited set of features. For example, Li et al. [26] and Bach et al. [27] use only psychological assessments of personality and cognitive abilities, whereas Mehta et al. [28] use resume data only. Chien and Chen [25] use only a few features, such as age, gender, marital status, educational background, work experience, and recruitment channels. Mehta et al. [28] conclude that features that capture candidate attributes, such as leadership, may contribute significantly to the analysis and that different models should be evaluated for different jobs. They indicate that a broader set of features and samples may enhance both prediction results and root cause analysis.

Lack of sufficient empirical data is reflected not only in the absolute amount of data (features, candidates) but also in the available data on populations that are usually not recruited and often are not even interviewed. It is evident that to extract significant insights using the potential of machine learning techniques on HR data, data should include a range of differing applicants [33]. Hence, data collected from a large organization that promotes a wide social diversity policy and hires a wide range of heterogeneous populations would be beneficial in showing new understandings and counter-intuitive results.

In contrast to many of the abovementioned papers, in our study, we use a large dataset with hundreds of thousands of employees from a wide range of heterogeneous populations, containing > 100 features. This unique dataset allows us to extract relatively deeper rules and insights based on a wider feature set and with high significance predictions of successful or unsuccessful recruitments.

2.3. Method dimension

Preliminary studies in HR analytics often used conventional statistical tools such as descriptive statistics, hypothesis testing, analysis of variance, regression and correlation analysis [27,34–37]. Bollinger et al. [37] used a *t-test* to determine the factors that affect recruiters'

decisions and integrated them into their aggregated score. Then, this single-score measure was used as a correlated measure to recruiters' surveyed opinions. Samuel and Chipunza [35] used the Chi-square test to identify which post-hire employment factors impact organizational turnover. Bach et al. [27] used multiple regression analysis to test which personality traits and cognitive ability features have an impact on employee performance. However, their regression models obtain a low fit ($R^2 = 0.054$, $R^2 = 0.088$).

More recent studies have started to use machine learning techniques for HR analytics. Some of them have implemented models that provide interpretable insights (e.g., [19,21,25,38]) and others have implemented non-interpretable models that provide solely the predictions or their ranked scores (e.g., [8,16,28,39]; further literature is detailed in recent surveys, e.g., [18,33]). In the rest of this section, we mainly focus on papers that addressed the pre-hire prediction of recruitment success using ML tools.

Chien and Chen [25] used the CHAID decision tree to extract rules for three different problems with separate classification targets: employee performance levels, turnover in the first three months of employment, and turnover in the first year of employment. They extracted several rules based on the demographic data of a rather moderately sized dataset of 3825 applicants, using all data as the training set (without using validation or test set, which can lead to overfitting). They suggested implementing some strategies based on the one-time findings from the obtained decision trees, such as recruiting from first-tier universities. However, they indicate that the HR staff found the extracted rules to be difficult to implement. The researches suggest performing an in-depth analysis to further clarify the root causes of turnover and implementing processes to effectively improve organizational retention rate. The small dataset used in their research could be the reason for the limitations of the extracted rules.

Li et al. [26] used a support vector machine (SVM) model to predict the performance of seven test candidates using a training set of 32 employees and focused on their personality test features. Mehta et al. [28] showed the results of a random forest classifier on a dataset containing resumes of candidates. However, they did not use an interpretable model to provide recruitment insights for the organization.

It should be noted that the suggested modeling approach in this study is intended to be used by HR professionals in order to facilitate improved interaction with candidates. Thus, there is significant importance to the provision of an interpretable model that can be well comprehended by HR professionals. The model evaluation should consider the interpretability as well as the accuracy of the model [9,12].

Another challenge that the proposed approach must take into account is complexity. In the recruitment-success classification problem under consideration, the complexity arises from a large set of features in the HR dataset (with > 150 features). Each feature has several or more possible values, resulting in a large combinatorial space of

potential feature interactions. Specifically, the dataset includes many categorical features, such as education certificates, test results, background details and potential assigned positions. In fact, extracting rules (i.e., patterns of feature values), even with a small number of features, may result in an extremely large space of potential combinations [10].

This study investigates several interpretable machine learning algorithms for predicting recruitment and placement success. The proposed method, which has not been used before for this objective, performs well in terms of both interpretability and accuracy, despite the inherent trade-off between the two [9,12]. The results of this research are expected to provide recruiters and organizations alike, with a useful modeling approach that generates insights for supporting recruitment and placement plans.

Moreover, the above reviewed studies provide local prediction scores, rankings or rules but do not provide a global prescriptive method that takes into account the position or the organizational point of view as a whole. To conclude, a prescriptive solution, rather than only a predictive methodology, is required for implementation in an actual organizational environment.

2.4. HR practices and HR analytics

Employees are considered one of the most important assets for modern organizations; hence, many efforts are invested in improving their success in the workplace. This has led to the rise of fields such as human resources (HR) analytics (which includes other related topics, such as "workforce analytics", "people analytics", and "human capital analytics" [40]). A recent review [40] maps the different tasks of HR practices to HR analytics tools and discusses how these tools can influence the organizational return on investment (ROI). The review shows that HR predictive analytics in workforce planning and recruitment have the highest effect on organizational ROI (similar conclusions are shown in a report by the Boston Consulting Group in [1]). Interestingly, as opposed to recruitment and workforce planning, other HR tasks, such as "industry analysis", "job analysis" and "performance management", have low expected ROI. Tasks such as "training", "compensation" and "retention" have medium expected ROI [40].

These findings correspond with our approach of a pre-hire in-advance design of the recruitment plan, which is expected to have more impact than a post-hoc approach. Post-hire information includes information such as: employee engagement, organizational commitment, organizational support and HR practices applied for retention [41–44]. This information surely affects employees' success and could improve the prediction accuracy if included in the model, but it may be too late to act upon this information while inducing much higher expenses. Nevertheless, there is already much hinted evidence in pre-recruitment information that can help predict success, even before it is known how the recruited individual engages with the organization. Hence, it is highly beneficial to focus on early pre-hire predictions that have the highest effect on organizational ROI.

An additional important organizational aspect to examine is diversity. A report by McKinsey & Company shows that diversity leads to better profits and that diverse companies may outperform others [45,46]. Therefore, there are economic incentives for enhancing diversity, not solely social or legal incentives.

Literature reveals that there is some criticism with regards to the use of HR analytics for business and commercial use [47–49]. Gelbard et al. (2017) [41] state that one of the main reasons for the rather scarce adoption of HR analytics approaches among organizations is the use of "black-box" methods and a lack of actionable items. As shown in [40], indeed, the focus of most human resources studies is mostly descriptive or predictive, and fewer are focused on prescriptive methodologies; however, a prescriptive solution can benefit organizations greatly [18]. For further information about the literature in the field of HR analytics, we refer the reader to recent reviews in [40–44].

In this paper, we aim to provide a prescriptive methodology that

includes interpretable insights and an optimization tool for recruitment planning and execution. This tool can be used as a decision support tool for HR professionals, since it not only provides actionable items but also allows for the incorporation of their valuable knowledge and experience into the model.

3. Methods and data

The goal of this study is to develop an analytic framework that can be implemented as a decision support tool for HR recruiters in real-world settings to efficiently hire suitable candidates and place them in the organization. The proposed methodology comprises two main components: i) a local prediction scheme for the recruitments' success with a technique for extracting meaningful insights based on the trained ML model and ii) a robust mathematical model that provides a global optimization of the recruitment process, taking into account multilevel considerations.

3.1. Local recruitment perspective

The first phase of this study is essentially aimed at predicting the fit of an employee to a specific position he or she is hired for. In this part of the study, we focus on using machine learning models for the pre-hire prediction of recruitment success and for the extraction of interpretable insights. The recruitment success measure is based on a combination of turnover and an objective performance indicator. This approach has several advantages in comparison to traditional methods: i) the target measure is objective; ii) it takes into account both turnover and performance; and iii) it focuses on the pre-hire prediction of recruitment success.

The use of an objective target measure, as opposed to other evaluations, allows for the examination of existing recruitment policies as well as the extraction of actionable and sometimes intriguing and unexpected insights. Objective performance is affected by the circumstances leading to a position change within the organization.

For classification and prediction of successful and unsuccessful recruitments and placements, as well as for mining significant patterns, we use a Variable-Order Bayesian Network model (VOBN) proposed by Ben-Gal et al. [10] and Singer and Ben-Gal [11]. Further details on the model used and its implementation in the recruitment process can be found in Appendix A. We evaluate the model against other interpretable and non-interpretable machine learning algorithms applied to the real-world recruitment dataset. We show that although the VOBN model has not been used before for the task of predicting recruitment success, it performs very well in terms of both interpretability and accuracy.

We use the trained VOBN model to identify context-based patterns that can support the organization in the recruitment process. As opposed to some black box models, the VOBN model can be used to extract rules and actions for the recruiters without any machine learning background, providing both scores and specific insights on factors and root causes that affect the success of recruitments.

In this phase, we focus on insights and interpretability (that are further discussed in Section 4 and Section 5), while in the second phase, we use the predicted probabilities for successful recruitments as inputs into a global recruitment optimization scheme that addresses more global parameters and objectives of the recruitment decisions at an organizational level.

3.2. Global recruitment optimization perspective

Recruitment success at an organizational level requires not only hiring the highest-potential workforce in a greedy manner but also optimizing the process to meet more general objectives. For example, a greedy allocation of candidates to jobs, such that the first candidates are allocated to the most promising job in terms of allocation success, can result in a sub-optimal situation in which certain jobs in the

organization will be poorly allocated. Other high-level goals that could be considered are meeting the need for employees at a certain proportion, facilitating the diversity of teams, or properly balancing the workforce among different departments. Each of these dimensions may also include numerous viewpoints, e.g., successful recruitment from the candidate viewpoint, successful allocation from the job viewpoint, and an overall organizational regulatory viewpoint. In this section, we mainly focus on a global optimization perspective that takes into consideration multiple goals of various organizational stakeholders.

In the first phase, we pursued interpretability via extracted patterns, through which HR professionals can locally act. In this phase, however, we aim at higher prediction accuracy rather than interpretability for the purpose of designing a more global optimization strategy. To this end, the model with the best prediction results (even if non-interpretable) can be used to predict the probability of success of each candidate for each of the intended positions. These predictions can then be used to address a more global recruitment plan that controls more parameters of the recruitment decisions.

3.2.1. Global optimization implementation

The considered problem spans multiple *dimensions*, satisfying different requirements as follows: i) **demand** – minimizing the difference between the required workforce demand and the actual number of recruited employees; ii) **accuracy** – maximizing the sum of the probabilities of the successful recruitment of employees in the organization; iii) **diversity** – balancing diverse groups of employees to maintain a heterogeneous work environment.

Note that when facing a recruitment challenge at an organizational level, it is important to ensure that each of the above dimensions is balanced across the various business units and positions in the organization. For example, when aiming to minimize the total number of non-filled open positions in the organization, the solution has to also account for fulfilling the demand over all the open positions in a balanced manner.

3.2.1.1. Mathematical programming formulation.

We consider the global recruitment task as an optimization problem and propose a mathematical programming formulation to solve it. The proposed formulation incorporates the objectives that were described above. We use the following parameters as input for the problem: the set of candidates E; the set of positions J; the binary qualification of candidate i to position j, represented by q_{ij} (it equals 1 if candidate i is qualified for position j and 0 otherwise); the predicted probability of candidate i to succeed in position j, denoted by P_{ij} which is the output of the learning model such as VOBN or GBM; and the number of open jobs in position j, denoted by N_j.

Since different positions may have different values associated with successful recruitment, our formulation introduces V_j as an input parameter that represents the value of successful recruitment to position j, it equals 1 if all the jobs are considered evenly, or can be set propositionally to the compensation value of that position relatively to other positions. To support diversity, this formulation includes, in addition, the following input parameters: T denotes different types or classes of candidates (T may represent, for example, the association with diverse groups of the population); the association of candidate i to a class of type t, denoted by b_{it} (it equals 1 if candidate i belongs to class t and 0 otherwise); and the minimal proportion of candidates of type t for position j, denoted by PR_{jt}. A summary of the notations, including the input parameters, the indices, and the model's decision variables, is presented in Table 1. Additionally, we use a more simplified and less constrained formulation for benchmark purposes.

The first formulation (Formulation 1) is used for benchmark purposes and is a rather simple adjustment to the assignment problem [50], in which the objective function (1.1) maximizes the sum of the predicted probabilities of assignments. Constraint set (1.2) ensures that no candidate is recruited to more than one position. Constraint set (1.3)

Table 1
Formulation notations.

Input parameters	
E	The set of candidates, $i \in 1, \dots |E|$
J	The set of open positions (jobs), $j \in 1, \dots |J|$
q_{ij}	Equals 1 if candidate i is qualified for position j, 0 otherwise
P_{ij}	Probability for candidate i to succeed in position j
N_j	Number of open jobs in position j
V_j	Value of a successful recruitment to position j
T	Set of candidate class types, $t \in 1, \dots |T|$
b_{it}	Equals 1 if candidate i belongs to class t, 0 otherwise
PR_{jt}	Required proportion of workers of class t in position j
B	A parameter that balances accuracy and demand objectives

Indices	
i	Candidate
j	Position
t	Class types of candidates

Decision variables	
X_{ij}	Recruitment of candidate i to position j
Y_j	The difference between required and recruited employees for position j
Y_{max}	Maximum allowed difference between required and recruited employees
Z_{jt}	Number of candidates of type t recruited to position j

requires that the number of recruitments for position j will not exceed N_j. Constraint set (1.4) ensures that only qualified candidates are recruited for positions. The next set of constraints (1.5) limits the set of possible values for X_{ij} (whether to assign candidate i to position j) to 0 or 1.

Formulation 1. A simple linear programming based on the classic assignment problem solution.

(1.1) $max(\Sigma_{i,\,j}[P_{ij} \cdot X_{ij}])$
Subject to the constraints

(1.2) $\Sigma_j X_{ij} \leq 1 \qquad , \forall i \in E$
(1.3) $\Sigma_i X_{ij} \leq N_j \qquad , \forall j \in J$
(1.4) $X_{ij} \leq q_{ij} \qquad , \forall i \in E, j \in J$
(1.5) $X_{ij} \in \{0,1\} \qquad , \forall i \in E, j \in J$

Formulation 1 raises several challenges that we wish to address. For example, positions that have a very low probability of succeeding might not receive any recruitments (hence, not considering the positional point of view of our demand requirement). Another challenge is that employees might not be evenly distributed among positions. In Formulation 2 below, we propose one way to address these requirements by adding a cost to the deviation from the recruitment demand (can be proportional to the loss due to this position staying unfulfilled).

Formulation 2 introduces the decision variable Y_j, which represents the difference between the required and recruited employees to the position while Y_{max}, is set the maximal allowed position shortage (constraints sets (2.5) and (2.6)). Accordingly, we then modify the objective function (2.1) to penalize the maximum deviation from the number of open positions ($B \cdot Y_{max}$), where B is a parameter that balances accuracy and demand objectives.

Hence, this penalty approach leads to a better distribution of the employee shortage among positions. Note that we choose to use the demand as a "soft" constraint and penalize shortages in the objective function, rather than forcing a specific level of demand satisfaction. This enables a larger feasible solution space and allows for achieving higher demand satisfaction by minimizing shortages in the objective function.

Formulation 2 also introduces diversity constraints into the model.

Table 2
Dimensions addressed by Formulations 1 and 2.

Dimension view	Demand	Accuracy	Diversity
Position	√(Formulations 2)		√(Formulation 2)
Organization - total value	√(Formulations 1, 2)	√(Formulations 1, 2)	√(Formulation 2)
Organization - balance across business units	√(Formulation 2)		√(Formulation 2)

Constraints (2.7) require that Z_{jt} will determine the number of candidates of type t that are assigned to position j. Constraints (2.8) require that the proportion of candidates of type t assigned to position j will be at least PR_{jt}. Table 2 presents the requirements that both Formulations 1 and 2 address in terms of the dimensions and viewpoints presented above.

Formulation 2. Proposed linear programming with diversity and penalty on maximal position shortage.

$$(2.1) \quad max(\Sigma_{i,\ j}[V_j P_{ij} \cdot X_{ij}] - B \cdot Y_{max})$$

Subject to the constraints

$$(2.2) \quad \Sigma_j X_{ij} \leq 1 \qquad\qquad , \forall i \in E$$
$$(2.3) \quad \Sigma_i X_{ij} \leq N_j \qquad\qquad , \forall j \in J$$
$$(2.4) \quad X_{ij} \leq q_{ij} \qquad\qquad , \forall i \in E, \forall j \in J$$
$$(2.5) \quad Y_j = N_j - \Sigma_i X_{ij} \qquad , \forall j \in J$$
$$(2.6) \quad Y_{max} \geq Y_j \qquad\qquad , \forall j \in J$$
$$(2.7) \quad Z_{jt} = \Sigma_i X_{ij} \cdot b_{it} \qquad , \forall j \in J, t \in T$$
$$(2.8) \quad Z_{jt} \geq PR_{jt} \cdot \Sigma_i X_{ij} \qquad , \forall j \in J, t \in T_{protected}$$
$$(2.9) \quad X_{ij} \in \{0,1\}, Z_{jt} \in \{0,1\} \quad , \forall i \in E, j \in J, t \in T$$
$$(2.10) \quad Y_j \in Integer \qquad\qquad , \forall j \in J$$

3.2.2. Motivating and illustrative example

We first solve the model over a sample of the real-world dataset as a motivating and illustrative example. We then continue to implement the solution using a larger real-world dataset and perform an analysis of the trade-off between the objectives, as shown in Section 4.4. As a first step, we use small sample data from the real-world dataset to illustrate the properties of the formulations. The data for the example are shown in Fig. 2. It includes four positions (columns), sixteen candidates (rows), two types of candidates that need to be balanced (e.g., based on their background), and the predicted success probability for each pair of candidate and position (shades of green represent high probability and shades of red represent low probabilities). In addition, we assume a demand of 6 employees for each position.

For example, it is clear from the table that if the only objective is to maximize the sum of success probabilities of candidates for each position separately, position 379 will be filled by candidates from the group of type 2 only.

Fig. 3 illustrates four different solutions to the problem: (1) solution to Formulation 1, (2) solution to Formulation 2 with PR = 0, (3) solution to Formulation 2 with PR = 0.1, and (4) solution to Formulation 2 with PR = 0.3. The rows represent the different candidates, and the columns represent the positions. Within each position, an assignment of a candidate to that position is marked with color. Table 3 shows several aggregated properties of the different solutions.

We observe the following: i) Accuracy: the predicted success probability decreases with more constrained models (i.e., models with more constraints) as a result of the shift from global to local objectives; ii) Demand: (1) formulations that penalize deviation from the required demand (Solutions 2–4) avoid cases of positions in shortage of assignments, and (2) solutions that incorporate the penalty on the maximum shortage (Solutions 2–4) manage to better balance the demand satisfaction among positions; and iii) Diversity: (1) adding diversification constraints to the formulations (Solutions 3 and 4) results in higher

diversity without significantly compromising accuracy, and (2) solutions that impose a high diversity requirement (Solution 4) may result in higher demand shortage. Similar results are expected over larger recruitment experiment.

3.3. Dataset description and target definition

The input dataset for this research includes hundreds of thousands of employment cases (approximately 700,000 cases) of employees who were recruited to the organization over the span of a decade (hired between the years 2000–2010). The pre-hire features in the dataset include age, gender, family and marital status, residence, nationality details, background record, education and grades, interviews and test scores (including leadership scores and language scores), professional preferences questionnaires, family details (when available), "lifestyle" data (when available), and details about the positions. Table 4 presents the main categories of the 164 features in the dataset.

In the preprocessing phase, 21 data tables were consolidated to mask sensitive private data and personal identification; on this dataset we also performed feature enrichment processes and addressed missing data and outliers. Specifically, in the feature enrichment process, we identified several interesting hierarchies of position groups and background data. In addition, we used residence-related data to deduce the socioeconomic levels of the candidates, using statistical data from the Central Bureau of Statistics. Missing values were tagged in the dataset by zeros, since these values mainly represented a lack of a specific test result or interview attribute. The reason to avoid a certain test or question for a specific candidate was not random nor uniform but rather based on the candidate's profile. For example, candidates who seemed to be less relevant to a specific job type were not asked to complete a related questionnaire or did not go through a specific interview segment. As such, these zeros indicate a specific categorical decision, which could be overlooked had we used the mean values (e.g., the mean of the results of certain tests, to impute them). The data records of candidates with many missing values were removed entirely; however, only < 1% of the records were removed in total. Additional dimensionality reduction procedures were performed in accordance with each of the applied machine learning algorithms (see details in Section 4).

The class feature definition for successful and unsuccessful recruitments was conducted by utilizing the following process: based on HR department records, the reasons for employee turnover were analyzed and accordingly divided into two groups: *successful recruitments* (e.g., the employee left for "natural" reasons, such as leaving the job after a sufficient time period) and *unsuccessful recruitments* (e.g., job termination after a short amount of time or due to poor performance). Position and placement changes were classified as negative (e.g., "misfit") or positive (e.g. "promotion" or "job enrichment processes").

To conclude, the combination of turnover and position changes was used as a combined measure for labeling successful vs. unsuccessful recruitments, as seen in Table 5. To clarify, the fifth row in the table represents instances that were excluded from the analysis as their period of employment was not long enough to determine if they were successful or not. To maintain consistency, the a-priori distributions of the target class in both the training and testing datasets include 30% of the unsuccessful recruits and 70% of the successful recruitments.

Recall that the dataset was acquired from a large nonprofit service

Candidate type	Candidate ID	Position 1409	Position 1509	Position 379	Position 40
type 1	Candidate 7317	0.67	0.68	0.68	0.48
	Candidate 8320	0.47	0.68	0.48	0.56
	Candidate 9346	0.53	0.68	0.39	0.56
	Candidate 3145	0.61	0.68	0.51	0.55
	Candidate 5438	0.63	0.68	0.48	0.57
	Candidate 0142	0.67	0.68	0.51	0.51
	Candidate 1617	0.55	0.68	0.45	0.55
type 2	Candidate 8610	0.66	0.68	0.85	0.58
	Candidate 2939	0.48	0.68	0.85	0.60
	Candidate 4349	0.56	0.73	0.85	0.60
	Candidate 4842	0.69	0.68	0.85	0.55
	Candidate 7983	0.69	0.58	0.86	0.59
	Candidate 6405	0.64	0.68	0.86	0.59
	Candidate 7882	0.59	0.62	0.86	0.48
	Candidate 5481	0.84	0.68	0.85	0.54
	Candidate 0226	0.62	0.69	0.83	0.54

Fig. 2. Predicted probabilities of success of assigning sixteen candidates of two types of populations to four positions. The entries are color-coded by the success probability values, green - high probability, red - low probability. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

organization that is highly diversified over roles, accountabilities and job descriptions with a heterogeneous population. These characteristics allow for testing potentially biased recruitment policies and decisions that traditionally may not be tested due to the absence of sufficient data on certain groups or lack of information on different personal properties. Specifically, it enables us to focus on various groups in the population and to show some counter-intuitive understandings based on data, which is not commonly available.

With respect to data selection, we aimed to focus on early pre-hire predictions; thus, the features that were integrated as predictors in the model included only the available pre-hire data, i.e., data from before the recruitment day. The motivation for such data selection was based

on several reasons. First, the recruitment day is an important decision point in which it is easier for the organization to take action —for example, the early identification of a possible misfit may save a great deal of financial and social costs. Second, such data selection enables the identification of actionable recommendations for preventive actions. For example, there is little interest in the revelation of turnover among employees who were absent for a long period of time immediately before they resigned (these causes are obvious and self-evident and also occur too late to be acted upon).

Note that although post-hire data was available (i.e., data about each employee through his employment period), we utilize only the pre-recruitment data. This approach allows us to achieve the goal of
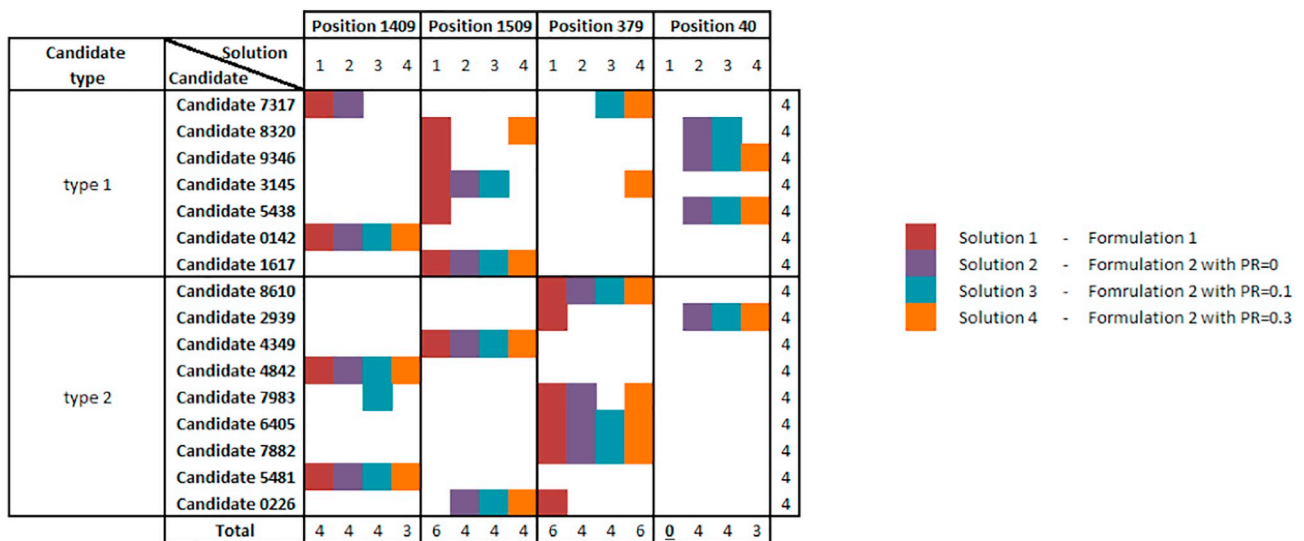

Fig. 3. Assignment of candidates to positions by four different solutions. For example, solution 1 (marked in red) suggests the following: i) recruiting 4 candidates to position 1409; ii) recruiting 6 candidates to position 1509; iii) recruiting 6 candidates to position 379 (note that none of them are of type 1); and iv) not recruiting any of the candidates to position 40 (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.).

Table 3

Illustrative example results. Entropy is used as a suitable measure for diversity in the case of more than two candidate type.

Solution #	Description	Demand		Diversity		Average success probability
		# of completely unassigned positions	Y_{max} (maximal position shortage)	Minimum proportion of type 1 population	Average position entropy	
1	Formulation 1	1	6	0	0.413	0.756
2	Formulation 2 with PR = 0	0	2	0	0.703	0.711
3	Formulation 2 with PR = 0.1	0	2	0.25	0.858	0.701
4	Formulation 2 with PR = 0.3	0	3	0.33	0.939	0.667

improving the recruitment process and providing insights that may be integrated within recruitment decision processes.

3.4. Prediction model and evaluation measure

The classification algorithms were trained on 70% of the candidates (first 8 years in the dataset). In the test stage, we used the trained classification models to predict the recruitment success of the remaining 30% of the candidates and validate our predictions with the ground truth. Note that we used time-dependent partitioning for training and testing to reassure the applicability of the model in the real world and show that the model can still be valid even when the organization changes.

In this process, we examined five interpretable machine learning algorithms and four non-interpretable algorithms. We evaluated the results of the prediction models by relying on the AUC (area under ROC curve) measure. According to the literature, e.g., Chawla [51], when the dataset is imbalanced (e.g., when the target variable includes large differences between the frequencies of different class values), an appropriate performance measure is the ROC curve and the AUC measure.

4. Results

4.1. Model evaluation

The study results are presented in Table 6 below. Comparing various interpretable and non-interpretable models, the best AUC score obtained by an interpretable model was obtained using the VOBN algorithm [10,11], with an AUC = 0.705 on the test set. The best results by a non-interpretable model, were obtained by the gradient boosting machine (GBM) algorithm with an AUC = 0.73. Thus, for interpretability purposes, we suggest selecting the VOBN model, whereas for solely aiming at prediction, we suggest using the GBM model.

A conventional approach to handle multiple (conflicting) objectives is to use a *Pareto-optimality* approach [52]. The model's AUC and its interpretability can be considered as two conflicting objectives that should be addressed by a Pareto-optimality approach. In this sense, the VOBN and the GBM algorithms are both "Pareto optimal". Specifically, the GBM should be selected if the objective is mainly prediction (although non-interpretable), while the VOBN model should be prioritized if the model interpretability is important, despite a relatively small decrease in the AUC score. Such interpretability not only enhances the understanding of key features in the prediction model but also provides root cause analysis and insights into the recruitment process. Following the evaluation of the different models, the VOBN and GBM models were used for further experimentation and analysis — the former for identifying interpretable patterns and the latter for a global optimization approach.

4.2. Identified patterns

Patterns in this use case can be thought of as regularities in the

dataset that characterize subpopulations of candidates with common characteristics. A pattern is often described by a set of rules that can be used to cluster subpopulations into different categories. The VOBN, as an interpretable descriptive model, enables the extraction of patterns that can be mapped into insights for the recruitment process, as seen in the next example. The VOBN model has generated more than a thousand patterns that went through a filtering process based on the following: their statistical validity (i.e., statistical significance and support set that indicates how many cases they refer to) and the change they imply on the recruitment's success probability with respect to other subpopulations. The final set of implemented patterns contained few dozens of patterns (a number that also depends on the ability of the recruiters to implement it in their routine procedures), including the ones used by the HR department and the ones presented in the next examples. These patterns were selected by a prioritization process that included the following steps: *i*) selecting patterns that contain at least one variable that can be controlled by the HR department, such as a threshold on a test result (otherwise the pattern is non-actionable); *ii*) selecting patterns in which the controlled variables separates well the population into subgroups resulting in different success probability outcomes; *iii*) prioritizing patterns that represent "counter-intuitive" phenomena that were not known to the recruiters; and *iv*) prioritizing patterns with larger number of instances in the leaves and with a larger turnover percentage.

The following are several examples of patterns, some of which are counter-intuitive and were extracted from the data by the VOBN algorithm.

Example 1. Correlation of a high analytical score in a pre-placement test with the dropout rate in a specific administrator position over different subpopulations.

As shown in Fig. 4, an interesting pattern is found related to the correlation of a high analytical score in a pre-placement test on the position dropout rate of certain administrator positions. As seen in the left figure, the position dropout rate falls only slightly (from 42.5% to 39.3%) when the candidate obtains a higher analytical score. However, as seen from the pattern in the right figure, for men with low leadership skills scores and low language scores, the dropout rate increases significantly (from 58.1% to 68.3%, with *p*-value < 0.001) if the candidate has a high analytical score. A possible explanation can be related to the fact that a high analytical ability has an over-qualifying effect on these specific candidates.

Skowronski [53] reviews the connections between over-qualification and turnover as well as performance. The paper proposes several practices for the pre-hire and post-hire management of overqualified employees and suggests considering perceived over-qualification rather than merely objective over-qualification. In the case of the considered pattern, it is likely for an employee to feel overqualified and less motivated if he or she is highly skilled but not able to demonstrate his or her competence due to language and communication gaps.

To overcome the above difficulties, recruiters should investigate which jobs' properties might decrease the probability of successful

Table 4
Feature summary (after data preparation procedures).

Feature cluster	Lifestyle	Family	Interview and test scores	Special interview scores	Education	Position	Nationality	Language	Residence	Culture	Background record	Gender	Age
# of Features	62	30	29	14	7	5	5	5	2	2	1	1	1
Avg. rank by GBM importance	12	11	2	9	3	1	7	8	5	6	4	10	13

recruitment and adjust the specific job requirements to accommodate for wider populations of employees. They may also devise unique programs for different populations that includes for example language, communication and technical training.

Example 2. The effect of competencies on the position dropout rate for a specific field-support position.

In general, the data show that candidates with high competencies are less likely to leave their position than are candidates with low competencies (15% vs. 30% position dropout rate, respectively, with *p*-value < 0.001). However, for specific field-support positions, this effect is reversed. Fig. 5 illustrates how candidates for specific support positions who have high competencies follow a significantly higher position dropout rate than do candidates with low competencies (43% vs. 21% position dropout rate, respectively, with p-value < 0.001). Here, the recruiters should again be aware of the reversed relation in the case of this field-support position.

This considered pattern also shows that the dropout rate for low-competency employees has decreased when they are assigned to a specific support position. This is somewhat unexpected since it implies that an organization should strive for the heterogeneity and diversity of its employees rather than recruiting only the most highly scored individuals. This notion is also supported in a report by McKinsey & Company that interestingly showed that diversity leads to better profits among organizations [45,46]. Let us note again that this output is due to the analysis of a unique dataset of a large nonprofit service organization that hires diverse populations with different backgrounds and skills.

Example 3. Correlation between low personal interview scores and low management skill levels with position dropout rates in specific business units for male candidates.

Table 5
Target definitions by HR department.

Employment status	Completed expected time in position (Position dependent)	Reason for leaving	HR term	Target feature label
Left the organization	Yes	"Natural" reasons	Turnover	Successful recruitment
Left the organization	Yes	Negative reasons	Turnover	Unsuccessful recruitment
Left the organization	No	Negative reasons	Turnover	Unsuccessful recruitment
Employed in the organization	Yes	–	Retention	Successful recruitment
Employed in the organization	No	–	–	–
Employed in the organization	No	Promotion or job enrichment	Position change - promotion	Successful recruitment
Employed in the organization	No	Negative reasons	Position change demotion	Unsuccessful recruitment

Table 6
Evaluation of models.[a]

	Explainability/interpretability[b]	AUC results over all test samples	AUC results by each position[c]	
			Number of positions with AUC > 0.7[d]	Average rank over all positions (1 - highest)[e]
GBM (gradient boosting)	No	0.730	174	3.07
RF (random forest)	No	0.719	164	3.39
VOBN (variable-order Bayesian networks)	Yes	0.705	145	3.78
LR (logistic regression)	Partial	0.700	129	4.30
SVM (support vector machine)	No	0.697	100	5.16
C45 (J48)	Yes	0.682	103	5.37
CHAID	Yes	0.681	105	5.12
Naive Bayes	Yes	0.677	80	5.81
CART	Yes	0.644	7	7.92

[a] Note that both the RF and GBM models and their implementations are generally robust to noisy and high dimensionality datasets, since they base their decisions on multiple permutations of the dataset (see [56,66–69]). For the logistic regression and decision tree models, we implemented a feature selection preprocess by using information gain analysis (see [70]). For the SVM model, we used the built-in model as implemented in [71], that can deal with high dimensionality by testing different subsets of the data. In the VOBN model, there is a built-in preprocess procedure that uses mutual information to identify the high-impact features (see the Appendix A for further details).
[b] We consider interpretable and non-interpretable models based on the classification presented in [72].
[c] These results show the AUC for each position in the organization. The AUC scores were calculated over all the candidates that were recruited and placed in specific positions.
[d] Out of 456 positions.
[e] For each position, the compared algorithms were ranked by the AUC score —the values in this column represent the average rank for each algorithm over all positions. A lower rank implies a better average AUC score.
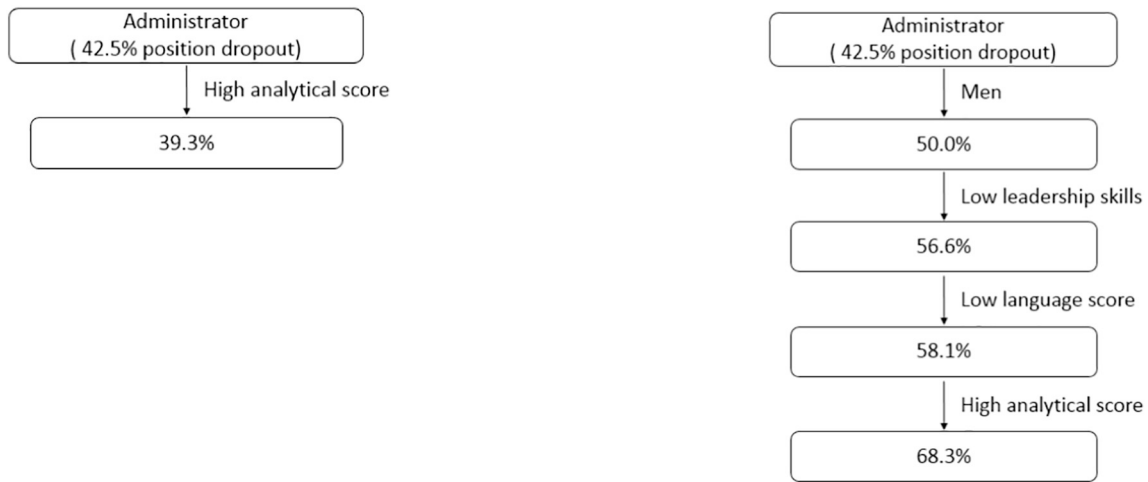
Fig. 4. High analytical score effect on administrator position dropout rate over various subpopulations.
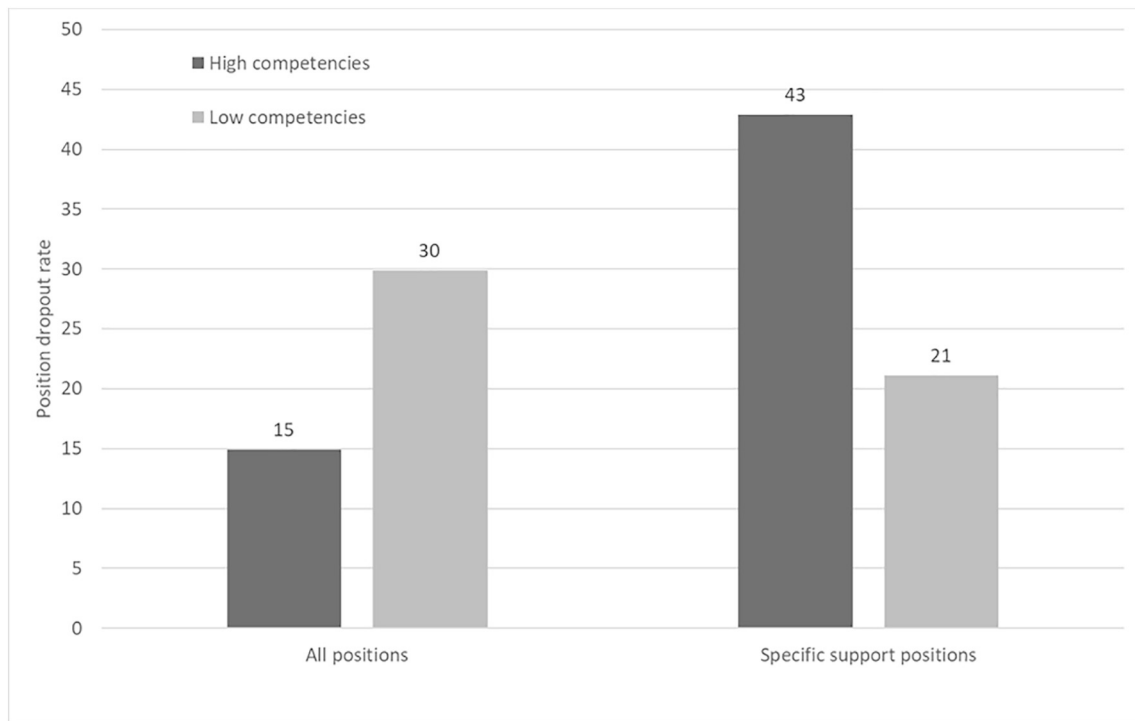

Fig. 5. The effect of competencies on the position dropout rate for all positions and for a specific field support position.

The pattern under consideration shows that the effect on the position dropout rate for male candidates with low scores in a specific section in the personal interview, combined with low management-skills score, is business-unit dependent, as shown in Fig. 6. For males, these low scores are associated with a position dropout rate of 37%, compared to an average dropout rate of 29% for all male candidates. However, this observation changes significantly among different business units, as seen in the figure.

In business unit A, the position dropout rate for all males is 39% (1928 out of 4916), while candidates with low scores have a considerably higher position dropout rate of 60% (394 out of 662). In business unit B, the opposite effect is observed: the dropout rate for male candidates with low scores is 23% (4113 out of 17,600), which is slightly lower than the rate for males, with an average score of 26% (7648 out of 29,049). All these differences have *p*-values lower than 0.001.

As mentioned above, these findings support previous observations

in the literature that call for the diversification and heterogeneity of workers [45,46]. Moreover, these findings emphasize the advantage of using data-driven methods to allocate people with diversified backgrounds and skills to specific positions (involving complex hidden patterns, related in this example to gender, business units, managerial and personal skills as well as specific test scores), in which they have a higher potential for success and good performance. Using the proposed approach, organizations should detect the characteristics of specific positions that are found to be statistically related to the allocation success of candidates from various backgrounds. These recruitment and allocation insights should be implemented accordingly, as long as they follow the required regulations for transparency, fairness and explainability (e.g., see GDPR: The EU's General Data Protection Regulation).

Example 4. Cultural background effect on position dropout for a specific office administrative position.

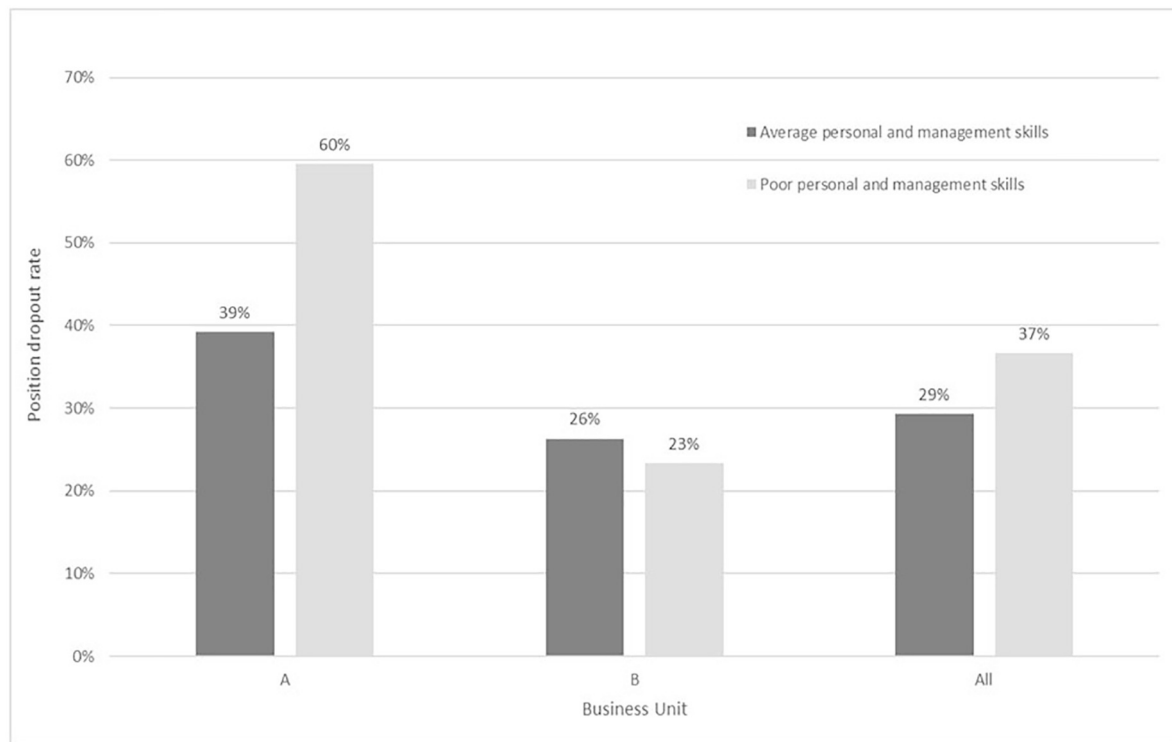The model identified a unique pattern that is related to a specific

Fig. 6. The relationship between poor-skill levels and position dropout rates for male candidates in different business units.
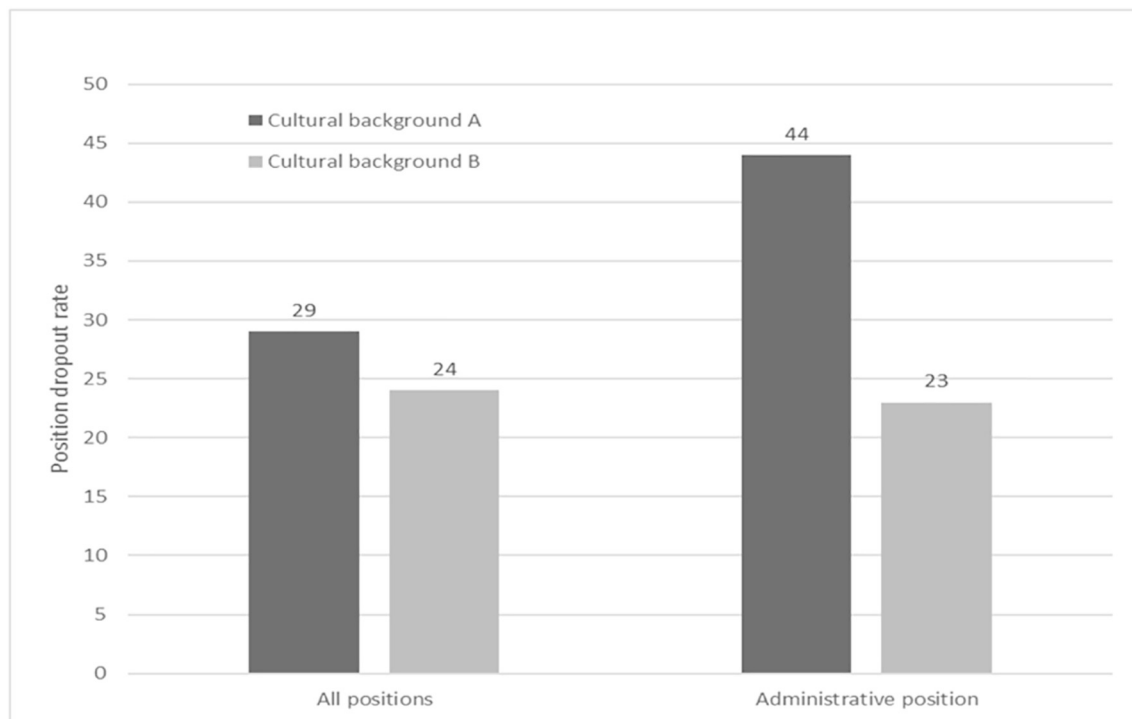

Fig. 7. The potential effect of the candidates' cultural background (A or B) on the dropout rate for all positions and for a specific administrative office position.

administrative office position. It turns out that for this office position, allocating a subpopulation of people with a specific common background results in a significantly lower dropout rate (23% instead of 44%, with p-value < 0.001).

Note that without a granular pattern-detection model, such as the one proposed here, it would be extremely difficult to identify such significant correlations between this office position and the specific

cultural background. As seen in Fig. 7, the average effect of the cultural background over all the positions is minor (indicating a 5% difference only). However, for this considered administrative office position, the effect on the dropout rate is marginal, i.e., more than four times greater (a 21% difference).

Organizations and researchers should investigate why some subpopulations of candidates who share common characteristics
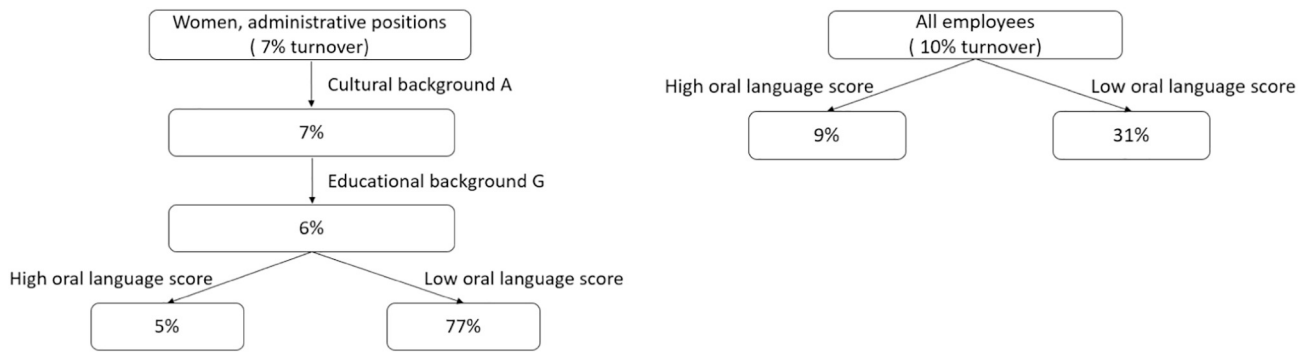
Fig. 8. The effect of the oral language score on turnover changes for specific subpopulations of candidates.

outperform or underperform in specific jobs or scenarios. Accordingly, they should find more opportunities to include (rather than exclude) specific populations as well as to adjust other organizational practices to support successful recruitment, considering the data-driven patterns discovered. In this context, it worth noting that the literature already recognized, for example, that some subpopulations of immigrants who share common cultural assets and social norms are sometimes better equipped than others to succeed in specific scenarios and vice versa (see [53,54]).

Example 5. The effect of oral language score on turnover differs by specific subpopulation.

The effect of an oral language score on turnover is heavily dependent on the chosen subpopulation. Fig. 8 shows that when analyzing this factor over all the employees in the organization, the turnover rate associated with a low oral language score results in a significantly higher position dropout rate (31% vs. 9%, with *p*-value < 0.001) and thus a lift of 3.2. However, when considering a subpopulation of women in administrative positions from a specific cultural background and a certain educational path, the turnover lift grows approximately to 15.5 (77% vs. 5%, with p-value < 0.001). This pattern addresses a rather privileged group of women according to their cultural and educational background, and although expected to succeed in their placement (with a 6% turnover only), there is a noticeable language deficiency that affects their ability to succeed in specific jobs.

It is interesting to compare the relative contribution of features when considering a large population to that of a specific subpopulation. Note that the feature importance of language according to Table 4 is relatively low; however, for a specific subpopulation, there is a greater impact of language skills. This notion is also closely related to *Simpson's Paradox* [55], which shows that an observed trend in subgroups may behave quite differently (even reversely) when these subgroups are aggregated and analyzed together.

4.3. Results application

When recruiters are looking for candidates to be placed in certain positions, they can take advantage of many patterns, such as those shown above. First, they can check that all the relevant data being used by the algorithm are collected and analyzed for all candidates. In addition, they can decide to send some of the candidates to undergo additional testing shown to be informatively correlated with the dropout rates. Then, they can apply the obtained patterns that were discovered by the algorithm to improve the recruitment and placement processes.

Finally, the obtained patterns can be used to reveal insights about factors that contribute to the recruitment success of specific positions. This in turn can provide feedback to the organization and can be used to adjust the position definition, such that it increases employment satisfaction and the overall recruitment success probability.

It is important to note that these insights must be considered in the proper ethical and legal contexts following specific regulations (e.g., GDPR). Organizations should investigate the reasons as to why some candidates underperform under certain scenarios and find opportunities to include, rather than exclude, diverse populations while adjusting organizational practices when the discovered patterns are considered. In the next section, we discuss a proposed global optimization formulation that can be used to enhance diversity in the organization while maintaining high success placement rates.

4.4. Data-driven global optimization results

In this section, we show an analysis of the proposed global optimization model. The analysis incorporates the data of real candidates, positions and demand and includes the predicted success probabilities derived from the prediction for a yearly planning program of our organization. We then perform a sensitivity analysis of the results and compare them to the recruiters' actual decisions.

The best prediction was obtained using the GBM algorithm [56], with AUC = 0.73 (see Table 6). We analyzed the robustness of the

Table 7
Results for a yearly plan.

| #	Method		Accuracy		Diversity		
	Solution	Diversity requirement (PR)	Average of Predicted success probability	Standard deviation over the mean probabilities of all positions	Minimum proportion of type 1 population	Average position entropy	Mean difference in accuracy between candidate croups
1.1	Actual selection	–	0.7087	0.1602	0	0.718	9.67%
2.1	Formulation 2	0	0.7654	0.1434	0.0057	0.619	7.43%
2.2	Formulation 2	0.1	0.7653	0.1431	0.1003	0.653	7.33%
2.3	Formulation 2	0.2	0.7648	0.1427	0.2	0.700	6.07%
2.4	Formulation 2	0.3	0.7638	0.1422	0.3005	0.780	4.86%
2.5	Formulation 2	0.4	0.7623	0.1414	0.4	0.810	3.69%
2.6	Formulation 2	0.5	0.7607	0.1407	0.5	0.864	2.33%
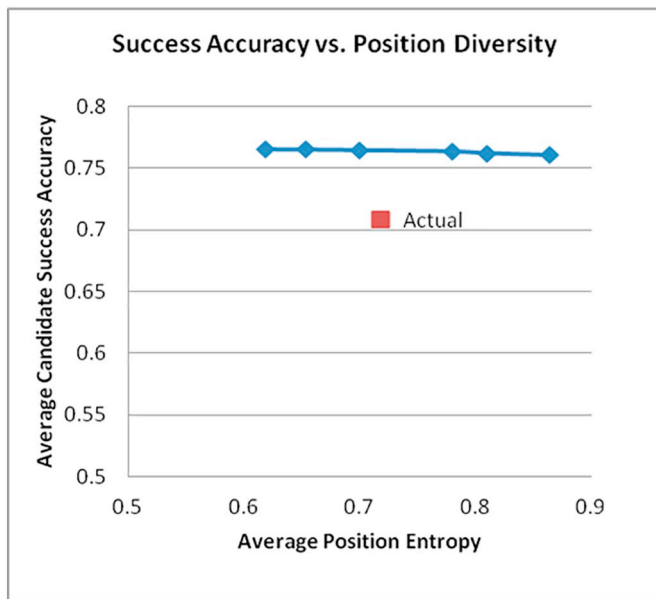
Fig. 9. Pareto efficiency for a yearly plan of the real-world scenario.

model by using different time-based partitions for training and testing and noted that the AUC value remained stable. Let us note that the results may be improved using post-hire data; however, as mentioned above, in this study, we focus on recruitment, as performing a prediction in a later post-hire phase may be too late to act upon, leading to much higher expenses.

The problem includes 30 position-types and all the candidates recruited to these position-types during a period of one year (10,329 candidates). As in the previous section, we compared different formulations of the problem to the actual assignment of the recruiters in the organization (see Table 7 and Fig. 9) and analyzed the trade-offs between different objectives. We expect to have similar results and to be able to show an improvement (in terms of both accuracy and diversity) compared to the actual allocation that was performed by the recruiters.

The results above indicate the following[1]:

- We expect that more complicated diversity requirements will lead to a reduced predicted probability of success. However, it can be observed that in the suggested formulations (Solutions 2.4–2.6 in Table 7), there was a significant improvement in both objectives in comparison to the actual assignment.
- Interestingly, requiring diversity at the position level also contributes to the organizational level:
  - o Average entropy - measuring individual positional diversity (the higher it is, the better).
  - o Mean difference - measuring balanced scoring between candidate groups (the lower it is, the better).
  - o Standard deviation of average probabilities for positions - measuring balanced average scoring between positions (the lower the standard deviation is, the higher the balance between positions).

To combine the local and global procedures, it is possible to integrate additional limitations or preferences in the model that were induced from interpretable insights or from other sources, such as legal or regulatory requirements. One approach for doing so is to incorporate

additional constraints to the model. Another possible approach is to indicate entries in the probability matrix, such as in the *Big M Method* [57]. For a future study, we suggest devising a model that will explicitly require the reduction of the imbalance between positions.

4.4.1. Arc deletion heuristic

We note that a more demanding and complex diversity requirement entails longer runtime, which ranges from a few minutes for PR = 0 to 51 min with PR = 0.1 up to several hours with a higher diversity requirement. However, the bottom line is that the size of a relevant assignment problem, even for a large organization with thousands of workers, is fully feasible with the proposed approach (specifically with the suggested heuristic described below).

In order to reduce the computational complexity of the model, we suggest a simple heuristic for the deletion of arcs. The heuristic deletes arcs that have a predicted success probability under a certain threshold (in our experimentation we use a threshold of $P_{ij} < 0.45$). We found that such a deletion allows one to reduce computational complexity while having only a minor effect on the obtained accuracy. In particular, after introducing arc deletion, the observed gaps to the best solution (without an arc deletion) were at most 1% but resulted in runtimes shorter by half or less with respect to the original runtimes. Note that deleting arcs, although improving runtimes, may compromise demand satisfaction in cases in which there are many candidates with an allocation probability which is lower than the threshold, for a certain position.

5. Conclusions

The objective of this study is to develop a hybrid decision support tool for HR professionals in the operations of recruitment and placement. The proposed methodology consists of two main components. The first is the definition of the problem as a machine learning problem with objective recruitment success as the target variable for specific candidate-position recruitments. The second is the development of a method based on mathematical modeling, which provides a global prescriptive hiring policy at an organizational level rather than a local one.

In the first phase, the machine learning model predicts the probabilities of successful recruitments and placements by taking into account various turnover scenarios and pre-recruitment data. The proposed approach is objective, based on an integrated performance indicator as opposed to some other evaluation schemes from the HR literature. It allows for an examination of current recruitment policies and the extraction of interpretable and actionable pattern-based insights.

In the second phase, the methodology considers the multi-stakeholder environment of the recruitment problem, including multisided balance and diversity in the process. We show that using the proposed mathematical programming model, even with the requirements of balanced demand and diversity, one is able to maintain a high level of accuracy and while improving the multiple objectives, compared to the actual selection of the recruiters. Implementing the presented approach as a decision support tool can increase the impact of recruiters and maximize organizational return on investment.

The utilized dataset in this study is unique and includes the data of hundreds of thousands of employees over a decade. The data represent a wide range of heterogeneous populations represented in a big-data repository. These characteristics allow us to analyze various recruitment policies and decisions that traditionally could not be tested due to the absence of proper data for such research studies.

The proposed methodology can be acted upon directly by HR professionals, without a need for deeper technical or machine learning knowledge, and can be implemented as a support software tool for recruiters and HR managers. A detailed study on the contribution of the proposed approach with respect to existing HR theories is beyond the

---

[1] The experiments were solved using the R *Rglpk* package (see [73,74]) for solution with the *presolver* option on (*presolve = True*) and were executed on a Windows Server based 64-bit with two 6-core CPU processors with 1.9GHz and 128 GB memory.

scope of this paper and can be found in [40,58].

We recognize that a prediction model that stands alone may be inherently biased; hence, in this work, we approach this potential bias through several measures: i) an objective target measure; ii) a large dataset incorporating a large range of differing applicants; iii) a mathematical programming model that enhances diversity and balance; and iv) a proposition to use a combined decision of both the recruiter and the used algorithm.

For future research, we suggest examining various directions of post-hire feature analysis and studying how these factors affect recruitment performance in comparison to the baseline literature as well as to a pre-hire analysis only. In light of the explainable patterns discovered in relevant candidate profiles, organizations may also devise and adjust personalized practices, such as specific training programs, awareness workshops, compensation and benefit plans, definitions of job duties, work-life balance policies, management and communication campaigns, and the overall organizational culture [44].

CRediT authorship contribution statement

All authors conceived of the presented ideas, developed the theory, performed the computations, discussed the results and took part in writing the paper. All authors read and approved the final manuscript.

Appendix A. Use of VOBN for recruitment success prediction

In this study, we propose to use flexible and generalized version of the Bayesian Network (BN) models [59], called *Variable Order Bayesian Networks* (VOBN) model as proposed by [10,11]. Similar to the BN it is an interpretable model that can be used to describe the relationship among various features, however, as opposed to BN possible connection between features does not imply necessarily that *all* the feature values of the conditioning features affect the conditioned feature. The possibility to construct such a flexible learning model that is not necessarily balanced over the entire feature space and at the same time can reveal those specific value-dependent patterns is found to be of outmost importance in the case of HR recruitment applications. For example, for a certain position, the probability of a successful recruitment might depend only on a specific language test score (e.g., a test score above 95) that is correlated with a specific managerial background, while all the other scores and background levels do not affect the recruitment success and should be therefore ignored or "lumped" together.

The following walk through example demonstrates the VOBN algorithm and implementation for predicting the turnover rate of female candidates who were hired to perform administrative roles. Detailed discussion on the VOBN algorithm can be found in [10].

Stage 1: Bayesian Network construction

First, the algorithm builds a Bayesian Network for the available features and the target variable, which in this case is the turnover rate. It uses the mutual information between the features as a dependence measure, and constructs the maximum likelihood graph structure, by placing feature with high mutual information next to each other. Next, it locates the target variable in the Bayesian Network and the features leading to it. In Fig. 10 one can see a portion of the Bayesian Network generated for the candidates' dataset. It shows that the conditioned distribution of the turnover rate, depends directly on the *Oral Language Score* feature, which depends on the feature *Educational Background*, which depends on the *Birth Country* etc.

For an alternative algorithm which uses a Bayesian network instead of the Bayesian tree see [10].

Stage 2: variable order Markov (VOM) context tree construction

After the Bayesian network (or a Bayesian tree in this example) is constructed, the algorithm constructs a complete and balanced tree of depth L – a fixed-order Markov tree of depth L, using the features from the Bayesian Network. It sets R to be the minimal frequency of samples in a leaf, for statistically significance evaluation. It chooses an initial depth L for the context tree, such that there are on average at least R examples in each leaf, to enable sufficient number of leaves with minimal frequency after the pruning stage (see stage 3). In the walk-through example the depth of the tree is set to $L = 3$ to obtain an average of $R = 100$ samples per leaf. We use the order found by the Bayesian network, as an input for the context tree construction.
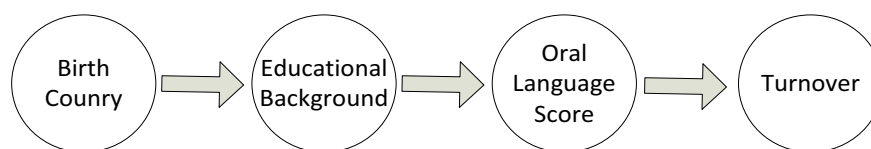
Dana Pessach: Conceptualization, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.**Gonen Singer:**Conceptualization, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.**Dan Avrahami:**Conceptualization, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.**Hila Chalutz Ben-Gal:**Conceptualization, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.**Erez Shmueli:**Conceptualization, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.**Irad Ben-Gal:**Conceptualization, Methodology, Formal analysis, Validation, Writing - original draft, Writing - review & editing.

Declaration of competing interest

None.

Fig. 10. Portion of the Bayesian Network constructed for the dataset.

Stage 3: context tree pruning

In order to obtain a minimal context tree, which capture most of the information in the features, and allows statistical significance, two pruning rules are applied as follows.

i) **Pruning rule 1** – leaf (i.e. the end node in the context tree) which has less examples than our minimal frequency of $R = 100$. Note that in Fig. 11 the pruned nodes/leaves are marked with a dashed border. Specifically, leaf {7} has only 48 examples and therefore it is pruned, since it has a smaller frequency than the minimal required (with 100 entities).

ii) **Pruning rule 2** – The algorithm compares the information obtained from the descendant leaf, defined by series of features *sb*, to the information obtained from the parent node, defined by series of features *s*. It then prunes the descendant node if the difference is smaller than a predefined penalty value for making the tree bigger – this penalty is called the *pruning threshold*. Hence, a node that has a turnover distribution similar to the distribution of the parent's node is pruned, since it doesn't add enough information. In this example, the algorithm estimates the turnover probability for each of the nodes, according to the frequencies of turnover cases.

In Eq. (1) the algorithm computes $\Delta N(sb)$ - the (ideal) code length difference between each descendant leaf, denoted by the pattern *sb* and its parent node, marked by the pattern *s*. *b* is the last split feature and its value of the descendent leaf, and *s* is the pattern defined by all previous split features and their values till the parent node. For example, in Fig. 11, descendant leaf can be *sb* = {*Low oral language score, Educational background A, Birth Country I*}, while its parent node is denoted by *s* = {*Low oral language score, Educational background A*}.

$$\Delta N(sb) = \sum_{x \in X} n(x \mid sb) log_2 \left(\frac{\widehat{P}(x \mid sb)}{\widehat{P}(x \mid s)} \right) \tag{1}$$

$X = \{turnoverTrue, turnoverFalse\}$

$\widehat{P}(x \mid sb)$ is the conditional probability for obtaining the value *x* in the descendant node *sb*, and $n(x|sb)$ denotes the number of samples with the value *x* in the descendant node *sb*, *X* is the finite set of values of the variable target. In our case, these are the turnoverTrue and the turnoverFalse values. If the difference is smaller than a pre-selected *pruning threshold*, the leaf is pruned, as defined in Eq. (2).

In order to reduce over-fit and simplify the context tree, without losing much information, the algorithm prunes the context tree, which leaves nodes that contributes significantly to the turnover classification task and contains enough samples to allow statistical significance. In order to achieve this requirement, it requires that $\Delta N(sb)$ will satisfy Eq. (2).

$$\Delta N(sb) > C(d + 1) \cdot log_2(t + 1) \tag{2}$$

where *C* is a pruning constant tuned to the considered process requirements (with default of $C = 2$ as suggested in [60]). *d* is the number of values the target variable can obtain, in our case $d = 2$ (since the target variable includes only two values: turnoverTrue, turnoverFalse) and *t* is the number of features defined by the pattern *sb* of the examined node.

We will now show an example for the calculations of Eqs. (1) and (2) using the tree shown on Fig. 11. When we calculate the (ideal) code length difference for the bottom descendent left leaf {6}, defined by the patterns *sb* = {Low oral language score, Educational background B, Birth country II}, compared to its parent node defined by *s* = {Low oral language score, Educational background B} using Eq. (1) we obtain the following result:

$$\Delta N(sb) = 12 \cdot log_2 \frac{0.12}{0.25} + 91 \cdot log_2 \frac{0.88}{0.75} = 8.27$$

In order for this descendent leaf not to be pruned, Eq. (2) must hold, i.e.,

$$\Delta N(sb) > 2 \cdot (2 + 1) \cdot log_2(3 + 1) = 12.$$

Since ΔN (8.27) is below the threshold in our case, 12, then leaf is pruned.

Similarly, the algorithm prunes leaf {2} in the tree shown on Fig. 11, with the pattern *High Oral Language Score* since its turnover rate (6%) is similar to the turnover rate of its parent node - in this case the root node {1} (7%).

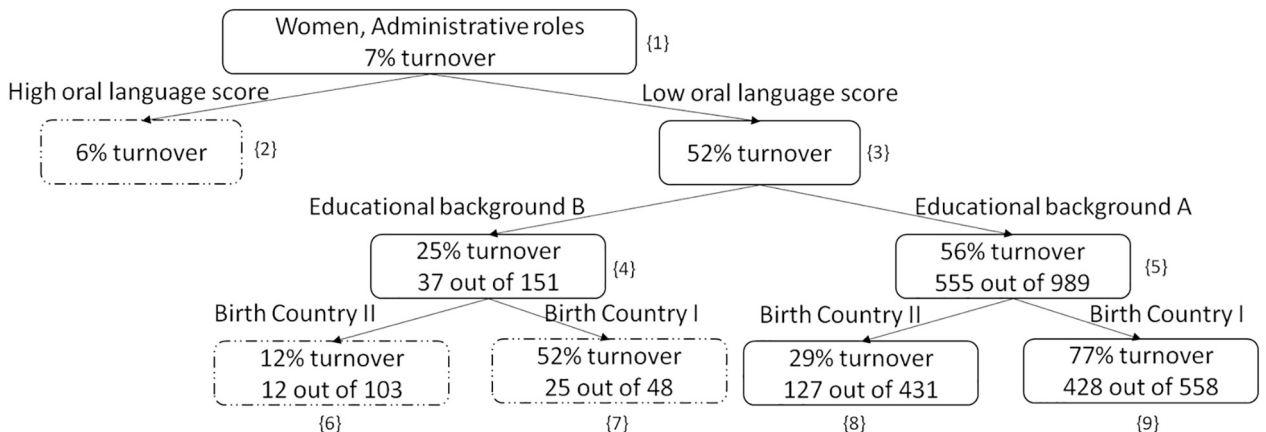The summary of the used notations in this section is presented in Table 8.

Fig. 11. VOM Context Tree of the walk thorough example Pruned nodes are marked with a dashed line.

Table 8
Notations.

Notation	
L	Depth of the complete and balanced tree
R	Minimal frequency of samples per leaf for statistically significance.
s	Pattern, define by series of variable of the parent node
sb	Pattern define the descendent leaf, by the series of the variables of the parent s, and addition split variable b.
x	The value of the target variable. In our case, $x \in X$ {turnoverFalse, turnoverTrue)
X	Finite set for the target variable X {turnoverFalse, turnoverTrue)
$n(x\|sb)$	Number of samples with the value x in the descendant node sb
$\Delta N(sb)$	The (ideal) code length difference between the descendent node sb and the parent node
$\widehat{P}(x \| sb)$	The estimated conditional probability for getting the value x in the descendant node
$\widehat{P}(x \| s)$	The estimated conditional probability for getting the symbol x (in the parent)
d	The size of the finite set X. In our case d = 2 since X = {turnoverFalse, turnoverTrue)
C	The pruning constant tuned to process requirements (with default $C = 2$)
t	The pattern size of an examined node (depth of leaf).

Stage 4: patterns identifications

The pruned context tree is left with a smaller number of leaves, each represents a pattern related to a specific sub-population, whose turnover rate is distinguishably different than the parent sub-population. In the context tree in Fig. 11, the following patterns are found for women in administrative roles with low oral language score:

1. Candidates from educational background B – 25% turnover rate.
2. Candidates from educational background A who were born in country II – 29% turnover rate.
3. Candidates from educational background A who were born in country I – 77% turnover rate.

Strength and Weaknesses of the VOBN Model in Recruitment Analysis

VOBN provides an important extension with respect to both Bayesian Network and Decision Tree models. In Decision Tree models, leaves represent class labels, nodes represent features and branches represent conjunctions of features that lead to those class labels. When the target variable takes a discrete set of values these trees are often called Classification Trees, while for a continuous target variable they are called Regression Trees. Decision Trees can generate a set of rules directing how to classify the target variable based on the associated features values, yet in a tree-like structure, thus where each node (feature) has only a single parent node. As opposed to decision trees, VOBN constructs a more general graph structure, where several nodes can be the parents of other nodes, thus representing a more general dependencies structure among different features in the model (these structures can then be mapped to a simpler tree-like rules, as done in this study). This generalization is important in the considered recruitment and placement application, since complex dependency patterns that involve several features and their interactions (e.g., background, performance, motivation etc.) can lead to different placements and recruitment recommendations that can result in a higher performance, as seen in Table 6.

The VOBN not only generalize Decision Trees but also generalizes the conventional Bayesian Network (BN) model. In BN modeling each variable (feature) depends on a fixed subset of random variables that are locally connected to it, however, in VOBN models these subsets may vary based on the specific realization of their observed variables. For example, a complex dependency between a Language Score and a Leadership Score features to the placement success in a specific position, might exists only for specific score values, while for other score values such a dependency is practically insignificant. This generalization lead to a reduction in the number of the model parameters, resulting in a better training and performance. The observed realizations in the VOBN are often called the contexts and, hence, VOBN models are also known as Context-Specific Bayesian networks.

In summary, compared to Decision Trees and conventional Bayesian Networks, often the classification performance of the VOBN is better, based on its higher flexibility in learning and expressing complex conditions and patterns among subsets of feature values. In the considered domain of HR analytics, this flexibility implies that the context dependency (based on the variable ordering) may be represented differently for each of the considered positions. Additionally, the VOBN handles better the variance-bias tradeoff, compared to decision trees, which often suffers from overfitting [10,61] and may cause high variance. The VOBN models have previously shown good performance in analyzing various datasets (some of which publicly available), including DNA sequence classification [10,62,63], transportation and production monitoring [11,64,65]. The VOBN has two main limitations. First, the dataset has to contain relatively large amount of data in order to construct the initial network structure. Second, the features introduced into the model should be discretized in a preprocess stage. In this study we used a large HR dataset, in which most of the features contain discrete values, hence yielding high performance of the VOBN.

It is worth noticing that this machine learning model has two main distinctions from traditional hypothesis-testing and regression models. First, the latter focuses on features that are highly correlated with trends across the entire aggregated sample, whereas the analysis by the VOBN model allows for identifying patterns in specific sub-groups. This notion is also closely related to Simpson's Paradox [55], which shows that an observed trend in subgroups may behave quite differently when these subgroups are aggregated and analyzed together. Second, hypothesis-testing requires in-advance assumptions about the interactions among features, whereas machine learning models do not require such assumptions, and allow for discovering insights that were not assumed ahead. For further mathematical and experimental details on the construction of the VOBN model, please see [10,11].

References

[1] J. Sullivan, News flash: recruiting has the highest business impact of any HR function, https://www.ere.net/news-flash-recruiting-has-the-highest-business-impact-of-any-hr-function/, (2012) , Accessed date: 14 July 2019.

[2] H. Boushey, S.J. Glynn, There are significant business costs to replacing employees, https://www.americanprogress.org/issues/economy/reports/2012/11/16/44464/there-are-significant-business-costs-to-replacing-employees/, (2012) , Accessed date: 14 July 2019.

[3] P.R. Bernthal, R.S. Wellins, Retaining talent: a benchmarking study, HR Benchmark Gr 2 (2001) 1–28.

[4] I. Lee, Modeling the benefit of e-recruiting process integration, Decis. Support. Syst. 51 (2011) 230–239.

[5] M.A. Shehu, F. Saeed, An adaptive personnel selection model for recruitment using domain-driven data mining, J. Theor. Appl. Inf. Technol. (2016) 91.

[6] A. Janusz, S. Stawicki, K. Drewniak Michałand Ciebiera, D. Ślkezak, K. Stencel, How to match jobs and candidates-a recruitment support system based on feature engineering and advanced analytics, Int. Conf. Inf. Process. Manag. Uncertain. Knowledge-Based Syst, 2018, pp. 503–514.

[7] P. Biecek, DALEX: explainers for complex predictive models in R, J. Mach. Learn. Res. 19 (2018) 3245–3249.

[8] E. Ribes, K. Touahri, B. Perthame, Employee turnover prediction and retention policies design: a case study, ArXiv Prepr, ArXiv1707 (2017) 01377.

[9] M.T. Ribeiro, S. Singh, C. Guestrin, Model-agnostic interpretability of machine learning, ArXiv Prepr, ArXiv1606 (2016) 05386.

[10] I. Ben-Gal, A. Shani, A. Gohr, J. Grau, S. Arviv, A. Shmilovici, S. Posch, I. Grosse, Identification of transcription factor binding sites with variable-order Bayesian networks, Bioinformatics 21 (2005) 2657–2666, https://doi.org/10.1093/bioinformatics/bti410.

[11] G. Singer, I. Ben-Gal, The funnel experiment: the Markov-based SPC approach, Qual. Reliab. Eng. Int. 23 (2007) 899–913.

[12] A.A. Freitas, Comprehensible classification models: a position paper, ACM SIGKDD Explor. Newsl. 15 (2014) 1–10.

[13] N. Sivaram, K. Ramar, Applicability of clustering and classification algorithms for recruitment data mining, Int. J. Comput. Appl. 4 (2010) 23–28.

[14] A. Barak, R. Gelbard, Classification by clustering decision tree-like classifier based on adjusted clusters, Expert Syst. Appl. 38 (2011) 8220–8228.

[15] E. Faliagka, L. Iliadis, I. Karydis, M. Rigou, S. Sioutas, A. Tsakalidis, G. Tzimas, On-line consistent ranking on e-recruitment: seeking the truth behind a well-formed CV, Artif. Intell. Rev. 42 (2014) 515–528.

[16] K.-Y. Wang, H.-Y. Shun, Applying back propagation neural networks in the prediction of management associate work retention for small and medium enterprises, Univers. J. Manag. 4 (2016) 223–227.

[17] X.-L. Qu, A decision tree applied to the grass-roots staffs' turnover problem—take CR Group as an example, Grey Syst. Intell. Serv. (GSIS), 2015 IEEE Int. Conf, 2015, pp. 378–382.

[18] R. Punnoose, P. Ajit, Prediction of employee turnover in organizations using machine learning algorithms, Int. J. Adv. Res. Artif. Intell. 5 (2016).

[19] E. Sikaroudi, A. Mohammad, R. Ghousi, A. Sikaroudi, A data mining approach to employee turnover prediction (case study: Arak automotive parts manufacturing), J. Ind. Syst. Eng. 8 (2015) 106–121.

[20] X. Gui, Z. Hu, J. Zhang, Y. Bao, Assessing personal performance with M-SVMs, Comput. Sci. Optim. (CSO), 2014 Seventh Int. Jt. Conf, 2014, pp. 598–601.

[21] C.-Y. Fan, P.-S. Fan, T.-Y. Chan, S.-H. Chang, Using hybrid data mining and machine learning clustering analysis to predict the turnover rate for technology professionals, Expert Syst. Appl. 39 (2012) 8844–8851.

[22] V.V. Saradhi, G.K. Palshikar, Employee churn prediction, Expert Syst. Appl. 38 (2011) 1999–2006.

[23] J.M. Kirimi, C.A. Moturi, Application of data mining classification in employee performance prediction, Int. J. Comput. Appl. 146 (2016).

[24] M.A. Valle, S. Varas, G.A. Ruz, Job performance prediction in a call center using a naive Bayes classifier, Expert Syst. Appl. 39 (2012) 9939–9945.

[25] C.-F. Chien, L.-F. Chen, Data mining to improve personnel selection and enhance human capital: a case study in high-technology industry, Expert Syst. Appl. 34 (2008) 280–290.

[26] Y.-M. Li, C.-Y. Lai, C.-P. Kao, Incorporate personality trait with support vector machine to acquire quality matching of personnel recruitment, 4th Int. Conf. Bus. Inf, 2008, pp. 1–11.

[27] M.P. Bach, N. Simic, M. Merkac, Forecasting employees' success at work in banking: could psychological testing be used as the crystal ball? Manag. Glob. Transitions. 11 (2013) 283.

[28] S. Mehta, R. Pimplikar, A. Singh, L.R. Varshney, K. Visweswariah, Efficient multi-faceted screening of job applicants, Proc. 16th Int. Conf. Extending Database Technol, 2013, pp. 661–671.

[29] S. Benabderrahmane, N. Mellouli, M. Lamolle, A. Janusz, S. Stawicki, K. Drewniak Michałand Ciebiera, D. Ślkezak, K. Stencel, On the predictive analysis of behavioral massive job data using embedded clustering and deep recurrent neural networks, Int. Conf. Inf. Process. Manag. Uncertain. Knowledge-based Syst, Elsevier, 2018, pp. 503–514.

[30] S. Yang, M. Korayem, K. AlJadda, T. Grainger, S. Natarajan, Combining content-based and collaborative filtering for job recommendation system: a cost-sensitive Statistical Relational Learning approach, Knowledge-Based Syst 136 (2017) 37–45.

[31] K.G. Bigsby, J.W. Ohlmann, K. Zhao, The turf is always greener: predicting de-commitments in college football recruiting using twitter data, Decis. Support. Syst. 116 (2019) 1–12.

[32] H. Jantan, A.R. Hamdan, Z.A. Othman, Towards applying data mining techniques for talent management, Int. Conf. Comput. Eng. Appl. IPCSIT, IACSIT Press, Singapore, 2011.

[33] S. Strohmeier, F. Piazza, Domain driven data mining in human resource management: a review of current research, Expert Syst. Appl. 40 (2013) 2410–2420.

[34] C.C. Hao, H.C. Jung, O. Yenhui, A study of the critical factors of the job involvement of financial service personnel after financial tsunami: take developing market (Taiwan) for example, African J. Bus. Manag. 3 (2009) 798.

[35] M.O. Samuel, C. Chipunza, Employee retention and turnover: using motivational variables as a panacea, African J. Bus. Manag. 3 (2009) 410.

[36] B. Coomber, K.L. Barriball, Impact of job satisfaction components on intent to leave and turnover for hospital-based nurses: a review of the research literature, Int. J. Nurs. Stud. 44 (2007) 297–314.

[37] J. Bollinger, D. Hardtke, B. Martin, Using social data for resume job matching, Proc. 2012 Work. Data-Driven User Behav. Model. Min. From Soc. Media, 2012, pp. 27–30.

[38] C.-F. Chien, L.-F. Chen, Using rough set theory to recruit and retain high-potential talents for semiconductor manufacturing, IEEE Trans. Semicond. Manuf. 20 (2007) 528–541.

[39] V.M. Menon, H.A. Rahulnath, A novel approach to evaluate and rank candidates in a recruitment process by estimating emotional intelligence through social media data, Next Gener. Intell. Syst. (ICNGIS), Int. Conf, 2016, pp. 1–6.

[40] H. Chalutz Ben-Gal, An ROI-based review of HR analytics: practical implementation tools, Pers. Rev. 48 (2019) 1429–1448, https://doi.org/10.1108/PR-11-2017-0362.

[41] R. Gelbard, R. Ramon-Gonen, A. Carmeli, R.M. Bittmann, R. Talyansky, Sentiment analysis in organizational work: towards an ontology of people analytics, Expert. Syst. 35 (2018) e12289.

[42] D. McIver, M.L. Lengnick-Hall, C.A. Lengnick-Hall, A strategic approach to workforce analytics: integrating science and agility, Bus. Horiz. 61 (2018) 397–407.

[43] A. Tursunbayeva, S. Di Lauro, C. Pagliari, People analytics: a scoping review of conceptual boundaries and value propositions, Int. J. Inf. Manag. 43 (2018) 224–247.

[44] T.H. Davenport, J. Harris, J. Shapiro, Competing on talent analytics, Harv. Bus. Rev. 88 (2010) 52–58.

[45] More Evidence That Company Diversity Leads To Better Profits, (n.d.). https://www.forbes.com/sites/karstenstrauss/2018/01/25/more-evidence-that-company-diversity-leads-to-better-profits/#2d3f7eff1bc7 (accessed November 23, 2019).

[46] Racially Diverse Companies Outperform Industry Norms by 35%, (n.d.). https://www.forbes.com/sites/ruchikatulshyan/2015/01/30/racially-diverse-companies-outperform-industry-norms-by-30/#5d4de9cf1132 (accessed November 23, 2019).

[47] A. Levenson, A. Fink, Human capital analytics: too much data and analysis, not enough models and business insights, J. Organ. Eff. People Perform. 4 (2017) 145–156.

[48] J. Boudreau, W. Cascio, Human capital analytics: why are we not there? J. Organ. Eff. People Perform. 4 (2017) 119–126.

[49] D. Minbaeva, Human capital analytics: why aren't we there? Introduction to the special issue, J. Organ. Eff. People Perform. 4 (2017) 110–118.

[50] D.W. Pentico, Assignment problems: a golden anniversary survey, Eur. J. Oper. Res. 176 (2007) 774–793.

[51] N.V. Chawla, Data mining for imbalanced datasets: an overview, Data Min. Knowl. Discov. Handb, Springer, 2009, pp. 875–886.

[52] A. Chinchuluun, P.M. Pardalos, A survey of recent developments in multiobjective optimization, Ann. Oper. Res. 154 (2007) 29–50.

[53] M. Skowronski, Overqualified employees: a review, a research agenda, and recommendations for practice, J. Appl. Bus. Econ. 21 (2019).

[54] D.C. Maynard, E.M. Brondolo, C.E. Connelly, C.E. Sauer, I'm too good for this job: Narcissism's role in the experience of overqualification, Appl. Psychol. 64 (2015) 208–232, https://doi.org/10.1111/apps.12031.

[55] C.R. Blyth, On simpson's paradox and the sure-thing principle, J. Am. Stat. Assoc. 67 (1972) 364–366, https://doi.org/10.1080/01621459.1972.10482387.

[56] J. Elith, J.R. Leathwick, T. Hastie, A working guide to boosted regression trees, J. Anim. Ecol. 77 (2008) 802–813.

[57] I. Griva, S.G. Nash, A. Sofer, Linear and Nonlinear Optimization, Siam, 2009.

[58] H. Chalutz Ben-Gal, D. Avrahami, D. Pessach, Singer Gonen, I. Ben-Gal, A Human Resources Analytics and Machine-Learning Examination of Turnover: Implications for Theory and Practice, (2020).

[59] I. Ben-Gal, Bayesian networks, Encycl. Stat. Qual. Reliab, 2007.

[60] M.J. Weinberger, J.J. Rissanen, M. Feder, A universal finite memory source, IEEE Trans. Inf. Theory 41 (1995) 643–652.

[61] I. Ben-Gal, G. Morag, A. Shmilovici, Context-based statistical process control, Technometrics 45 (2003) 293–311, https://doi.org/10.1198/004017003000000122.

[62] S. Posch, J. Grau, A. Gohr, I. Ben-Gal, A.E. Kel, I. Grosse, Recognition of cis-regulatory elements with vombat, J. Bioinforma. Comput. Biol. 5 (2007) 561–577.

[63] J. Grau, I. Ben-Gal, S. Posch, I. Grosse, VOMBAT: prediction of transcription factor binding sites using variable order Bayesian trees, Nucleic Acids Res. 34 (2006).

[64] I. Ben-Gal, G. Singer, Statistical process control via context modeling of finite-state processes: an application to production monitoring, IIE Trans. (Institute Ind. Eng.) 36 (2004) 401–415, https://doi.org/10.1080/07408170490426125.

[65] I. Ben-Gal, G. Morag, A. Shmilovici, Context-based statistical process control: a monitoring procedure for state-dependent processes, Technometrics 45 (2003) 293–311.

[66] M.N. Wright, A. Ziegler, {ranger}: a fast implementation of random forests for high dimensional data in {C + +} and {R}, J. Stat. Softw. 77 (2017) 1–17, https://doi.org/10.18637/jss.v077.i01.

[67] R.J. Hijmans, S. Phillips, J. Leathwick, J. Elith, dismo: Species Distribution

Modeling, https://cran.r-project.org/package=dismo, (2017).

[68] S. Janitza, E. Celik, A.-L. Boulesteix, A computationally fast variable importance test for random forests for high-dimensional data, Adv. Data Anal. Classif. 12 (2018) 885–915.

[69] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.

[70] P. Romanski, L. Kotthoff, FSelector: Selecting Attributes, (2016).

[71] I. Steinwart, P. Thomann, {liquidSVM}: A Fast and Versatile SVM Package, ArXiv E-Prints 1702.06899, http://www.isa.uni-stuttgart.de/software, (2017).

[72] S.B. Kotsiantis, I. Zaharakis, P. Pintelas, Supervised machine learning: a review of classification techniques, Emerg. Artif. Intell. Appl. Comput. Eng. 160 (2007) 3–24.

[73] S. Theussl, K. Hornik, Rglpk, Rglpk: R/GNU Linear Programming Kit Interface, https://cran.r-project.org/package=Rglpk, (2017).

[74] J.M. Sallan, O. Lordan, V. Fernandez, Modeling and Solving Linear Programming With R, OmniaScience, 2015.

Dana Pessach is a PhD candidate at the Big Data Lab at Tel-Aviv University. Her research focuses on applying computational methods to social sciences and contributing to the development of new methods to deal with real-life problems. She is additionally a member of the steering committee at LAMBDA, the artificial intelligence lab at Tel-Aviv University, and a freelance consultant to start-ups and entrepreneurs in the fields of Machine Learning and Artificial Intelligence. As part of her role she leads research projects in domains such as: HR analytics using Machine Learning methods, financial data technologies, health data analysis and smart-cities. Her teaching experience includes several topics for undergraduate programs such as: operations research, robotics, mobile application development, modeling and 3D printing, social network analysis and computer vision.

Gonen Singer, Ph.D., is a Senior Lecturer of Industrial Engineering and Information Systems at Bar-Ilan University. Before joining Bar Ilan, Gonen was a Senior Lecturer at AFEKA-Tel-Aviv Academic College of Engineering. He joined to the Department of Industrial Engineering and Management at AFEKA, shortly after its establishment at 2008 and was appointed as Head of the Department in the years 2009–2015. Gonen Singer has extensive expertise in machine learning techniques and stochastic optimal control and their application to real-world problems in different areas, such as retail, manufacturing and education and has around 50 Scientific and professional publications. Gonen Singer served as a principal investigator in 10 Research Grants since 2011 that were obtained from various sources.

Dan Avrahami holds an M.Sc. from LAMBDA, the artificial intelligence lab at Tel-Aviv University, where he also taught Information Systems Engineering. In his HR analytics research, he explored the application of machine learning in general, and a generalization of Bayesian Networks in particular, to improve employees' recruitment and placement.

Dan is a data scientist and machine learning researcher with experience implementing various machine learning methods on challenging data domains. He has a solid background as a software engineer and database administrator.

Hila Chalutz Ben-Gal is a Senior Lecturer at the Department of Industrial Engineering and Management, Tel Aviv Afeka College of Engineering, Israel. She received her BA degree (with honors) from the Hebrew University in Jerusalem, Master's Degree from Brandeis University, Boston, MA, USA, and her Ph.D from Haifa University, Israel. Her research interests include HR Analytics, and the changing nature of organizations in the digital era, with a focus on organizational design and the future of work. She conducts field research and consulting work in a variety of organizational settings. Some of her current projects include HR analytics and turnover forecasting, career moves detection in on-line labor markets and the changing nature of work due to COVID-19 pandemic. During 2015–17 she was a Visiting Scholar at the ChartLAB at the University of San Francisco in California, USA.

Erez Shmueli is a senior lecturer and the head of the Big Data Lab at the department of Industrial Engineering at Tel-Aviv University and a research affiliate at the MIT Media Lab. He received his BA degree (with honors) in Computer Science from the Open University of Israel and MSc and PhD degrees in Information Systems Engineering from Ben-Gurion University of the Negev and spent two years as a post-doctoral associate at the MIT Media Lab. His research interests include Big Data, Complex Networks, Computational Social Science, Machine Learning, Recommender Systems, Database Systems, Information Security and Privacy. His professional experience includes being a programmer and a team leader in the Israeli Air-Force and, a project manager in Deutsche Telekom Laboratories at Ben-Gurion University of the Negev, a co-founder of two startups (Babator and SafeMode), and a consultant (among the rest to Microsoft and the municipality of Ashdod).

Irad Ben-Gal is a full professor and the Head of the Laboratory for AI, Machine learning, Business & Data Analytics (LAMBDA) at the Engineering Faculty in Tel Aviv University. He received his M.Sc. (1996) and Ph.D. (1999) from the Boston University. His research interests include machine learning, applied probability and statistical control, involving R &D collaborations with companies such as Oracle, Intel, GM, AT&T, Applied Materials and Nokia. He wrote three books, published > 120 journal and peer-reviewed conference papers and patents, supervised dozens of graduate students and received numerous awards for his work. He co-heads the TAU/Stanford "Digital Living 2030" research program that focuses on data science challenges of modern digital life. The program involves scientists at TAU and Stanford, where he held a visiting professor position, teaching "analytics in action" to graduate students. Irad is the co-founder of CB4 ("See Before"), a startup backed by Sequoia Capital that provides predictive analytics solutions to retail organizations.