Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 1-19

Siddiqui and Tariq

What is Predictive Analytics for Workforce Attrition? A Machine Learning Approach to Strategic Human Resource Management

[1] Dr. Farhan Siddiqui
[2] Zainab Tariq

How to cite this article:

Siddiqui, F., & Tariq, Z. (2020). What is Predictive Analytics for Workforce Attrition? A Machine Learning Approach to Strategic Human Resource Management. *Journal of Technological Information, Management & Engineering Sciences, 1(2), 1-19.*

Received	Accepted	Published
06 July 2020	10 August 2020	14 November 2020

Abstract

The loss of employees can have a significant impact on the organization's knowledge, service continuity, team dynamics, and workforce planning. People analytics is being used more and more in strategic HRM to uncover patterns with employee turnover and to target specific retention strategies. Machine learning can help with this work by leveraging information from the workforce, role, career, engagement and organization to predict attrition risk. Prediction, however, does not equal explanation, and a model that is successful with statistics could still cause harm if it is based on biased data, scores that are not transparent, or if it is used in a way that is inappropriate for employee monitoring. In this paper, a conceptual framework is developed for leveraging predictive analytics on workforce attrition in a responsible and strategically beneficial way. It summarizes some of the common statistical and machine learning methods, provides an overview of a data governance and modeling process, and suggests a framework for evaluating models based on predictive performance, calibration, fairness, interpretability, and business impact. The methodology includes problem definition, leakage-free validation, handling of class imbalance, comparing model performance, pilot deployment, and monitoring. A decision support process is suggested in which risk estimates are not used directly to decide the person's employment but are used to alert the organisation to review its position for support. Illustrative tables and a workflow diagram explain how HR teams can link outputs of the model to interventions including career development, coaching of managers, workload review and compensation analysis. The paper does not provide any results of a conducted experiment, nor does it include any organization-specific dataset, so that no empirical results are claimed, but rather a structured research and implementation blueprint is provided. It is concluded that by incorporating fair governance, employee support, and outcomes-based evaluation, attrition analytics can be a key component in strengthening strategic HRM.

Keywords: *employee; turnover; customer relations; transformers; text classification; customer experience.*

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Introduction

High employee turnover is a continual problem for any company looking to keep working well, retain institutional knowledge, and control payroll. This can lead to recruitment and onboarding costs, loss of customer relationships, pressure on the remaining staff and project delays. The impact on operations can be different depending on the role: a turnover in a highly sought-after technical or other leadership position will affect operations differently than a turnover in a volume-based role that has a clear succession plan. Therefore, HR attrition should be viewed as a strategic workforce problem and not just administrative HR metric.

Traditional retention strategies are typically based on exit interviews, annual employee engagement surveys, manager observations, and retrospective turnover data. These sources are valuable but can be constrained by a lack of timely feedback, participation, understanding or ability to observe trends in large populations. A complementary way is through predictive analytics, which predicts the chances of a certain outcome over a certain period of time based on past data. Machine learning models can detect the non-linear relationships and interactions among factors like, tenure, role change, career progression, compensation position, workload, and engagement.

However, there are some methodological and ethical issues arising from attrition prediction. Factors for employees to leave vary and many are not captured in HR systems. Historical labels may not be based on personal preference but organizational practice. It is not accurate to say that a high-risk score means an employee wants to quit, or for any other reason that they may leave. If managers believe predictions as reality, they could act differently to their staffs, cut down development opportunities or set up self-fulfilling scenarios. A responsible system must therefore be purposeful, minimize data use, be fair to test, supervised by human oversight, and distinguish between prediction and causal explanation.

This paper proposes a conceptual research model that brings predictive analytics into strategic human resource management (SHRM). It emphasizes the role that a model can play in the ongoing learning of the organization and in the retention

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

of the employee, while safeguarding employee rights and organizational trust, through employee-centred retention interventions.

Research aims, objectives, and questions.

This paper aims to present a scientific approach to modelling employee turnover and using predictions to inform HR strategies, in a responsible manner.

2.1 Research objectives

- Examine existing workforce prediction theories and machine learning techniques.
- Recognize appropriate data preparation, validation and evaluation procedures for HR data.

Create an overall model for the creation of risk assessment and facilitating retention measures.

Define fairness, privacy, explainability and governance protections for deployment.

Discuss methods to assess value of a business beyond the model's accuracy, such as retention results, employee experience, etc.

2.2 Research questions

RQ1: What types of data and model families can be used in workforce attrition prediction?

- RQ2: When attrition events are relatively rare, how can the predictive performance be evaluated?

How to include the principles of fairness, privacy and interpretability in model design and deployment?

When can predictions help in retaining employees without hurting their trust?

2.3 Research propositions

While a turnover baseline based on history alone may suffice, having all of these indicators combined should increase the prediction, if data quality is acceptable. P2: Leakage resistant validation and probability calibration will result in more useful decisions than accuracy only. P3: P3: Human-centred interventions and manager training will shape the potential for measurable retention value from predictive insight. P4: Fairness and privacy controls must be built into the design for sustainable adoption.

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

The analysis of relevant literature including theoretical backgrounds.

3.2 Staff morale and satisfaction with the workplace

Job satisfaction, organizational commitment, perceived alternatives, job embeddedness, and shocks that encourage rethinking about working, have been explored as factors to take into consideration when studying turnover. Mobley's process model illustrates how dissatisfaction can result in thoughts about leaving and withdrawal processes, as well as the fact that the relationship between dissatisfaction and leaving is not necessarily linear. Mitchell and associates' idea of job embeddedness focuses on three components that can make staying appealing: links, fit, and sacrifice. The above views indicate that personality factors and the employment situation play a part in the attrition process.

A machine learning model may be used to discover relationships between the observed variables, but it is not considered to be a complete psychological process as described in turnover theory. Knowledge, although theoretical, is still relevant in the selection of variables, the understanding of the patterns and the development of interventions. For instance, if there is an association between low internal mobility and high attrition risk, it might lead to the consideration of career pathways; however, it doesn't establish a causal link between not being promoted and why a particular employee has left.

3.2 Strategic human resource management

SHRM aligns workforce practices to long-term organizational goals. This view should help with workforce capability, succession planning, skills development, managers' effectiveness, and organisational design through the use of attrition analytics. Modelling is most helpful where outputs relate to decisions that HR can have a legitimate influence on. Simply identifying a set of employees with high predicted risk is not of strategic importance because of the value it brings to the organization when they are asked the question, "Does the appropriate intervention make a difference in outcomes?

3.3 People Analytics and Machine Learning

People analytics uses descriptive, diagnostic and predictive – and sometimes prescriptive – approaches to answer questions about the workforce. Logistic

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

regression provides an easy-to-interpret baseline and provides probability estimates based on the modeling assumptions. Decision trees and random forests are able to detect nonlinear patterns and interactions. Support vector machines and neural networks can be particularly suitable in specific scenarios, and gradient-boosted trees can work very well on tabular data. Don't be swayed by the novelty of the model, but by data size, data structure, operational requirements and governance needs.

3.4 Research gap

One weakness of applied attrition analysis is the focus on classification performance, with a lack of consideration to the deployment context. This can be deceptive when a large number of employees remain, as a model can look good only because it is picking up information about an employee's decision to leave. Also, there may be differences in performance between job families and/or demographic groups. Thus, a comprehensive system must include temporal validation, calibration, subgroup analysis, interpretability, intervention design, and assessment of the real-world effects of the intervention on an organization.

4. Conceptual Framework

A proposed framework integrates five layers: (1) strategic workforce question, (2) governed and relevant data, (3) validated risk estimation, (4) human reviewed interventions and (5) outcome monitoring. The target should be clearly stated, e.g. voluntary resignation with 6 months from start of prediction date for those who were active on the prediction date. It is important to differentiate between voluntary and involuntary departures as they imply different processes and may require different actions.

Apply the rules of the game Play the gameManage the game

Approval of scope by HR, legal, privacy, workforce stakeholders Clear definition of outcomes, population, horizon and decision purpose

Data and governanceGather necessary HR, role, career, and survey information

Limit data; keep data provenance, access, and usage restrictions documented

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 1-19

Siddiqui and Tariq

Model development Baseline model and candidate model development; temporal validation Prevent leakage; compare candidate and simple baseline models

Human review - no automatic adverse employment action Review risk patterns and identify supportive actions Decision and intervention

Review, monitor, audit, and monitor prediction quality, fairness, retention, and trust; track impact of interventions and unintended consequences;

4.1 Workflow diagram

Explain the difference between retention question and prediction horizon

↓

Collect relevant workforce data, while it is under the control of the collectee.

↓

Clean Data, Generate Time Valid Features, Partition by time

↓

Train baseline and candidate models → Calibrate and audit.

↓

Discuss patterns with HR and managers

↓

Provide suitable assistance → Assess results and equity

↓

Be on the lookout for drift, record changes, and re-evaluate

Figure 1. Proposed responsible attrition analytics workflow. Risk scores are not a reflection of intent and do not replace employee dialogue, but are intended to provide information for review and support.

5. Data Needs and Data Preparation

5.1 Potential data sources

Candidates are considered for being a potential predictor in the list below: Job family, Role level, Tenure, Contract type, Location at sufficiently high level, Internal move, Promotion history, Training participation, Manager span of control, Absence patterns (where legally and ethically appropriate), Compensation position, Working arrangements, Voluntary engagement survey

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

results. The significance and acceptability of each variable should be considered in the context. Sensitive attributes should not be collected simply because they are available unless the gathering of the information is lawful and appropriate in order to determine the fairness of the audit process.

Data should be linked to a specific forecast date. The information available for that date only should be used as a predictor. If one of these occurs after the prediction point, the result might escape into a termination workflow status, scheduling exit interviews, or a resignation notice that's recorded. Leakage can yield seemingly high validation scores which then break in real use.

5.2 Data quality

Eliminate duplicate employee records and inconsistent identifiers and maintain an audit-traceable history of transformations.

Review missingness – by variable, time period, role group, and outcome; missingness itself can be a manifestation of process differences.

Uniform categories and record job code changes as well as organizational changes and HR system changes.

Check that outcome labels differentiate voluntary resignation, dismissal, retirement, and completion of contracts and agreements from transfers from one position to another within a company.

Eliminate or modify irrelevant, intrusive fields, or fields which are available after the date of the prediction.

5.3 Feature engineering

Time-aware features can include tenure, time since promotion, number of role changes, recent compensation change, time since training and changes in engagement scores. Only data up to the prediction date should be used for rolling windows. The features should have an easily understood operational meaning, be stable enough to be reproducible. They are not automatically suitable indicators of digital activity, private communication or intrusive behavioural monitoring, and may lead to a loss of trust and introduce significant legal and ethical issues.

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 1-19

5.4 Class imbalance

In many organisations, employees that do not stay in for a period of time are in the minority. A model could predict a high accuracy, that everyone would stick around. Thus, precision, recall, area under curve, calibration and decision costs should all be considered for evaluation. Any resampling must be from the training data; weighting of classes (or other such changes to the model training) may be used to help model training. Final validation and test sets should include the natural attrition rate.

6. Machine Learning Methods

Testing Method Advantages Limitations and appropriate role

Logistic regression Efficient, may be difficult to interpret (baseline) if no feature engineering, can produce probability estimates.

Decision tree Easy to visualize, captures interactions, can be unstable without constraints, can overfit decision tree.

Single tree Simple to understand; good baseline for tabular data; probability calibration may be needed Random forest Captures nonlinearities; robust baseline for tabular data

Gradient-boosted trees Can be very effective on structured HR data, but requires tuning and careful explanation; can overfit small datasets

Can perform well in high dimensional problems Scaling & probability interpretation may not be as easy as SVM

Neural network Flexible with large and complex data sets, but may require more data and computing power; difficult to explain and control.

Survival analysis Models time to departure and takes into account censoring

Needs appropriate formulation of time to event and assumptions

6.1 Baseline-first model development

The first model should be a baseline model like the historical attrition rate by general job group or a logistic regression with a limited number of valid variables. More complex candidates are only to be retained where it is believed that they will enhance the out-of-time performance, or offer other measurable benefits, which outweigh the increased complexity. In addition to discrimination,

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

factors for model selection include: calibration, stability, ease of explanation, and implementation cost.

6.2 Model interpretation

Global feature importance can be used to determine which features most strongly influence the predictions from a dataset, whereas local techniques like SHAP can be used to explain features that are associated with a prediction. The explanations do not provide any causal explanation, nor do they provide any clues to the employee's actual reasons. Careful explanations should not be given as a way to accuse the person of being disloyal, uncommitted or likely to leave.

6.3 Probability calibration

A predicted probability should be similar to the frequency of the event. Validation can be performed by reliability plots, Brier score, calibration error, and the use of calibration methods like Platt scaling or isotonic regression can be considered on validation data. An accurate estimate helps with workforce-level planning, but **remains a best guess and shouldn't be considered a certainty about a person.**

7. Proposed Research Methodology

This paper introduces a quantitative, design oriented research process that can be followed by an organisation where the HR data has been suitably governed. Relevant ethics, privacy, legal and HR governance stakeholders must approve the study before data are extracted.

7.1 Research design and population

Identify the population for each prediction date and decide if the outcome is voluntary external departure, all external departure, or a role-specific outcome. Select a prediction horizon, depending on the intervention lead time, e.g., three-month, 6-month, or 12-month. Internal transfers should generally be kept separate from external attrition as they can be considered as an achievement of internal mobility.

7.2 Data splitting and validation

Train on past periods, tune on a later validation period, and evaluate on a later holdout. When snapshots are repeated, then the subsequent snapshots of the same employee should not be leaked in training that is not available at

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

deployment. When employees are limited in number, apply skillfully engineered rolling-original validation and report uncertainty. Any preprocessing / imputation / encoding / feature selection / resampling should be fitted in each fold of the training set.

7.3 Model comparison

Compare a simple baseline with logistic regression, tree ensemble and—a when appropriate—a survival model. Apply the same target definition, time horizon, data splits and evaluation rules. Hyperparameters must be tuned only during the training and validation phase. Choose a threshold rather than a decision threshold to maximize accuracy that is based not only on available support capacity, but also on the relative consequences of a false positive versus a false negative.

7.4 Proposed hypotheses

H1: A machine learning model validated for discriminating between future leavers and stayers will be superior to a prevalence-only baseline.

H2: Temporal validation will yield a more realistic prediction of future performance than a random split of data that does not consider time.

H3: Calibration and subgroup auditing will enhance usefulness and governance of model outputs.

H4: The effects of supportive retention interventions based on the model will only be considered as having an impact on retention outcomes if they are compared to a credible comparison group.

7.5 Pilot implementation

Start with retrospective testing and then have a shadow pilot where prediction are made but not given to decision makers or influence employees. Conduct a review of errors and subgroup performance prior to a limited, controlled decision support pilot. Employees in HR should be educated of uncertainty, limits, privacy and proper use. Before deployment a rollback plan and an incident-reporting system should be set up.

8. Evaluation Framework

Predictive, probability, fairness and operational value should all be considered when evaluating models. Results should be presented for the overall workforce

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

and for meaningful groups, if sample size and legal requirements allow. Caution should be used for small subgroup samples as metrics may be unstable.

Units Description Application to attrition analysis.

Correctness Fewer than 10 percent difference between number of predicted and observed leavers Predicts the amount of follow-up effort that may be wasted

Recall / sensitivity % of actual leavers correctly identified by the model Measures the coverage of observed leavers

F1-score Harmonic mean of precision and recall Summary metric when both Precision and Recall are important

PRIM (Precision at Recall level) Precision at a defined Recall point Useful when the departures are infrequent

ROC-AUC: Ranking ability across thresholds; Useful discrimination measure, but not enough on its own

Brier score Mean squared error of probability predictions Evaluates the accuracy and calibration of probability predictions

To check that the risk estimates can be interpreted as probabilities, a calibration plot is created.To check that the risk estimates can be interpreted as probabilities, a calibration plot is created.

Subgroup error analysis Precision, recall, calibration and error rates by subgroup – distinguishes between possible subgroup discrepancies that may need further investigation

Intervention lift: Difference in retention outcome versus a credible comparison: Assesses whether action creates value beyond prediction

There is trust and experience among employees There are feedback, perceived fairness and process complaints There are checks for unintended harm and adoption risk

8.1 Avoiding misleading conclusions

A model that has a high ROC-AUC can potentially yield a large number of false positives in the case of low prevalence of attrition. Reports should include base rate, confusion matrix, precision/recall results, calibration, confidence intervals, and performance by time period. The organisation should also evaluate the

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Technological Information, Management & Engineering Sciences (2020),1(2), 1-19

proposed intervention's effectiveness, affordability and employees' acceptance. Performance in prediction is a different question than effectiveness of intervention.

8.2 Illustrative evaluation table

Evaluation dimension Proposed acceptance question Evidence to record

Predictive value Is the model' predictive value better than the baseline in later period data? The four evaluation metrics are baseline comparison, PR-AUC, recall, and precision.There are four evaluation metrics: baseline comparison, PR-AUC, recall, and precision.

Calibration – Is the predicted risk band consistent with observed rates? A calibration plot, Brier score and subgroup checks are required.A calibration plot, Brier score and subgroup checks are necessary.

Equity Are errors or intervention accessible differently? Detailed data and reviewed documentation

Operational feasibility Is HR able to take action on the number of cases flagged? As the number of cases increases, the amount of time required for review grows and the capacity for intervention diminishes.

Business impact Does a pilot have a positive impact on retention or prevent avoidable turnover? We need to compare the design, costs and outcomes of each of these, and also, follow up for a period of time to see the results.

Are uses of data transparent, necessary and controlled? The following are to be used for impact assessment, access logs, complaints and staff feedback:

Evaluation questions, rather than findings not achieved, are identified in the table. The actual acceptance threshold should be negotiated prior to testing and should be specific to the organizational context.

10. Taking Action; Making Predictions

Predictive analytics adds value to an organization when it enables them to recognize factors they can change in their workplace and offer support that makes a difference. These risk factors should be followed up by a risk review process, not a resigned-to-the-fire attitude. Managers and HR partners should

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

take into account the context of the conversation, ask questions when relevant and not press an employee into disclosing their concerns.

The potential pattern to be investigated is being sought by repeating the process with a different pattern.The different pattern is being repeated to find the pattern to investigate.

Limited internal mobility Discuss career paths, skills, mentoring or internal vacancies without promising to promote or implying dissatisfaction without discussion

Absence data as a proxy for commitment Absence data not a measure of workload or scheduling strain Staffing, priorities, flexibility and recovery time reviewed

Compensation position concern Conduct equitable market and internal-pay review Do not make automatic pay decision based on risk scores

Limited engagement trend Provide confidential listening avenues and team-level improvement opportunities. Participation and score data in survey can be noisy or sensitive.

Manager or team turnover pattern Review leadership support, role clarity, and team conditions Don't blame individuals for a pattern of turnover by using aggregated analysis

Development is stalled Provide learning plans, challenges or coaching Make sure opportunities are fairly distributed

9.1 From individual scores to organizational insight

In some environments, a list of aggregated patterns might be more suitable than a list of risk factors at the individual level. When a specific job family experiences an increasing rate of turnover, the management can study the factors that make the job less attractive, such as salary competitiveness, job demands, career ladder, or organizational practices, rather than deeming specific individuals as less desirable. The aggregation analysis can also help minimize privacy risk and prompt leaders to focus on systemic causes instead of calling out on to individuals.

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

9.2 Measuring intervention impact

Piloting should involve specifying outcomes before the activity begins, including feelings of fairness by employees, absenteeism (if applicable), engagement, mobility (internally) and voluntary turnover. If possible and ethical, conduct a randomized or phased rollout; otherwise attempt the matched comparison group or a strong quasi-experimental design. A simple comparison before and after might be affected by seasonality, labour market transformations or other organizational changes. The analysis should take into account the costs of intervention and the unintended consequences.

Ethics, Fairness, Privacy and Governance 10

10.1 Fairness and discrimination risks

Historical HR information can be influenced by inequitable promotions opportunities, manager assessment and past discrimination. These patterns can be replicated by a model even if they are not associated with a protected characteristic, as other factors could be proxies for the protected characteristic. Data quality, target definition, error rates, calibration, access to interventions and downstream outcomes should be assessed for fairness. There is no one fairness measure to solve all the trade-offs; the organization should record the criteria that are important and why.

10.2 Privacy and data minimization.

Only data that is relevant for a valid and well-defined purpose should be collected by the organisation, access should be limited, there should be limits on retention and documentation of who can access scores. Understandable explanation of the purpose, categories of information in the system, and safeguards, shall be provided to employees in accordance with applicable law and organizational policy. There is a need to not reuse sensitive information just because it enhances prediction.

10.3 Human oversight & prohibited uses

A score should not be translated directly into hiring, promotion, salary, evaluations, reprimand, termination, or developmental opportunities. These applications can lead to negative feedback and could be illegal or immoral. HR should check if there are any limitations in the model and the context of the

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

action, and have clear accountability for decisions. Staff have a means of challenging or questioning uses of their data that have consequences.

10.4 Model governance

- Adopt owners for data quality, model performance, intervention policy and incident response.

Keep versioned datasets, code, model components, thresholds, documentation, approvals etc.

Check and review drift and performance regularly and following significant organizational and system changes.

- Need to validate and document an upgrade before putting it out.
- Access controls, audit trails, secure storage, and a deletion/correction request procedure.

11. Discussion

The framework places at the centre of SHRM, the capability to predict the attrition. It can be used to give a focus to inquiry and uncover patterns that a description might not show, but it cannot decide on the intent of an employee nor can it supplant a positive work environment. The organization should be able to tell the difference between statistical association and cause and should not be using risk estimates as labels. Among other things, recent changes in role or responses to engagement notifications could be related to turnover, and the changes in these roles are decisions that can be made by management.

An important decision in design is to make a single prediction, team prediction or both. Individual risk estimates can aid in personalized discussions but have increased privacy and stigma risks. Team level forecasts can be used to assist workforce planning and policy review and also can mitigate the risk of treating a score as a fact about an individual. The level of granularity should be determined by purpose, intervention capacity and governance protection.

It also underscores the fact that the value of retention cannot be established through prediction. An organization might be able to determine the people who are at risk to leave, and then provide them with interventions that they don't want.An organization may be correct in determining who will leave, and provide them with interventions which they are not receptive. On the other hand, a

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

somewhat predictive model can be applied, even at an aggregate level, to detect structural problems. Evaluation should then be based on predictive accuracy and effectiveness of intervention, cost, fairness, and employee experience.

Last but not least, it is important to have participation from HR, analytics, information security, privacy, legal, line managers and employee representatives (where applicable) when dealing with responsible adoption. Technical teams must be knowledgeable about the field to set reasonable goals, HR teams must be analytical to understand uncertainty and leaders must be willing to tackle systemic issues. If this organizational ability is missing, a complex model might be more complicated but not better at retaining.

Finally, the limitations and future directions of the study are discussed.

The paper is conceptual and lacks the use of a real organizational dataset or measures the model performance. The proposed methods and workflow should be proofed in a particular context. Definitions of atritions, legal requirements, size of the workforce, changes in the labour market and means of intervention differ between organizations. What works in one organization may not work in another and results can vary from one year to the next.

Further studies are needed to evaluate the usefulness of models across sectors and different workforce, to examine the implications of different prediction horizons, and to examine the effects of individual versus team-level decision support. There is still a need to finish the work on causal inference for retention interventions, privacy preserving analytics, strong calibration in presence of drift, and fairness assessment in case of small samples in subgroups. Longitudinal studies should evaluate the effects of analytics on employee experience, as well as organizational outcomes, over several years.

13. Conclusion

Predictive analytics can be used to assist HRM in predicting employee attrition rate, **detect patterns to investigate, and prioritize HRM retention resources. To be** effective, the target must be clearly specified, the data must be relevant and legally controlled, the validation must be watertight, comparisons must be made with simple baselines, probabilities must be calibrated, auditing checks must be performed on subgroups and the process of implementation must be

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

monitored continuously. Choosing machine learning techniques based on complexity alone is not sufficient – they should be based on evidence and operational requirements. Most importantly, risk assessments should not lead to automatic adverse decisions with regard to employees but rather to supportive review and systemic improvement. Attrition analytics can be a responsible tool in the long-term workforce planning, when combined with human judgment, transparent governance, and credible measures of the effectiveness of interventions.

References

Allen, D. G., Bryant, P. C., & Vardaman, J. M. (2010). Retaining talent: Moving from myth to fact-based solutions. Academy of Management Perspectives, Vol. 24, No. 2, pp. 48-64.

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). Classification and Regression Trees. Wadsworth.

Chapman, D. W., & Webster, J. (2003). Use of technology for recruiting, screening and selection of job applicants. International Journal of Selection and Assessment, 11(2–3), 113-120.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique is a technique used for over-sampling the minority class. Journal of Artificial Intelligence Research, 16, 321–357.

Fawcett, T. (2006). Introduction to ROC analysis. Pattern Recognition Letters, 27(8), 861–874.

Friedman, J. H. (2001). Greedy function approximation: Gradient Boosting machine. The Annals of Statistics, 29(5), 1189–1232.

Hardt, M., Price, E., & Srebro, N. (2016). Equal opportunity supervised learning. In Advances in Neural Information Processing Systems, 29.

Mobley, W. H. (1977). Intermediate links in the relationship between job satisfaction and employee turnover. Journal of Applied Psychology, 62(2), 237–240.

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C. J., & Erez, M. (2001). Why people stay: Voluntary turnover prediction using job embeddedness. Academy of Management Journal, 44(6), 1102–1121.

Molnar, C. (2022). Interpretable Machine Learning (2nd ed.). Christophm.github.io/interpretable-ml-book/.

Rosenbaum, P. R., & Rubin, D. B. (1983). The main role of the propensity score in observational studies of causal effects. Biometrika, 70(1), 41–55.

Ruhnau, B. (2000). Eigenvector-centrality—a node-centrality? Social Networks, 22(4), 357–365.

Shmueli, G. (2010). To explain or to predict? Statistical Science, 25(3), 289–310.

Van den Broeck, A., Ferris, D. L., Chang, C.-H., & Rosen, C. C. (2016). A brief overview of basic psychological needs at work from the perspective of self-determination theory. Journal of Management, 42(5), 1195–1229.

Vardaman, J. M., & Halter, A. P. (2017). Employee turnover: Review and directions for the future. In The Routledge Companion to Human Resource Management.

A Responsible Deployment Checklist is provided as Appendix A.

Define Voluntary Attrition, Prediction Horizon, Eligible Population, and Legitimate Decision Purpose.

Ensure necessary HR, privacy, legal, security and ethics approvals are obtained.

Record data provenance, permissions, data retention, and data missingness.

Avoid features that reveal information after the event or don't have a proper purpose.

Partition by time and avoid leakage of data down to the employee/batch level.

- Compare candidate models to a simple model and evaluate calibration.

Check subgroup error rates, access to intervention, and possible proxy discrimination.

Conduct a shadow pilot prior to any decision-support use.

Educate HR users on uncertainty, interpretation, privacy and prohibited uses.

Assess the consequences of intervention, employee trust, costs and unintended consequences.

Designate owners to monitor/respond to incidents, retrain and rollback.

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan
Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.

Example Model Evaluation Record (Appendix B)

Field	Record
No. of students	Number of Students
Target and horizon	Outcome definition and prediction period
R1 - 2019/11/15	All the source systems and the extraction date and the inclusion criteria are provided.
Split strategy	Divided time period / Employee overlap checking
Model	Algorithm, features, parameters, software version
Confidence intervals	PR-AUC, ROC-AUC, precision, recall, calibration
Decision to review based on fairness and metrics from subgroups, limitations	
Operational feasibility	Number of cases flagged / HR review capacity / cost
Intervention evaluation	Intervention offered, comparison approach, measured outcomes
Approval and limitations	Approver, un-resolved risks, next review date

[1] Department of Software Engineering, University of Engineering and Technology, Taxila, Pakistan
[2] Department of Computer Science, Fatima Jinnah Women University, Rawalpindi, Pakistan

Copyright (c) 2020 Author(s)
This work is licensed under a Creative Commons Attribution 4.0 International License.