Boosting Retention through Predictive Modeling

Yuejia (Mandy) Teng, PhD

Agenda

Executive Summary

Purpose

- Leverage predictive analytics to identify at-risk employees.
- Develop retention strategies to reduce turnover rates.
- Propose methods and metrics to evaluate the effectiveness.

Analysis and Findings

- A Random Forest model selected to predict turnover and identified key drivers.
- Employees with lower belonging and engagement, limited growth, low recognition have higher risk.
- Specific groups (e.g., onsite) have higher turnover.
- Approximately 10% of employees at high risk.

Recommendations

- Take actions to improve employee experience such as belonging, engagement, growth, recognition.
- Pinpoint at-risk employees and subgroups to prioritize retention efforts.
- Evaluate the progress in real time and post-doc, and communicate to stakeholders and employees frequently.

Next Steps

- Balance quick wins vs. long term bets in retention strategies to achieve sustained results.
- Evaluate the effectiveness of strategy adoption using workforce, customer, and business outcomes.
- Maintain and refine the predictive model to identify at-risk employees in real-time.

Assumptions and Data Overview

Objective

- Leverage predictive modeling to identify turnover risk factors and propose data-driven retention recommendations.
- Propose methods to measure and evaluate retention strategies over time.

Assumptions

- A mid-sized multinational company with a turnover rate higher than the industry benchmark.
- Employee turnover and survey data representative of the entire organization.
- Longitudinal data collected at 2 timepoints with no missing values.
- Stable economic conditions, with no major industry shifts during the analysis period.

Data Sources

- 5000 data points from engagement surveys, performance management systems, HRIS.
- Incorporated external benchmarks from industry reports to contextualize findings.

Data at a Glance

Experience Surveys

- Job Satisfaction
- Engagement
- Belonging
- Recognition
- Growth Opportunity
- Work Life Balance
- Psychological Safety
- Culture
- Satisfaction with Supervisor

Demographics

- Tenure (years)
- Job Level*
- Department*
- Team Size
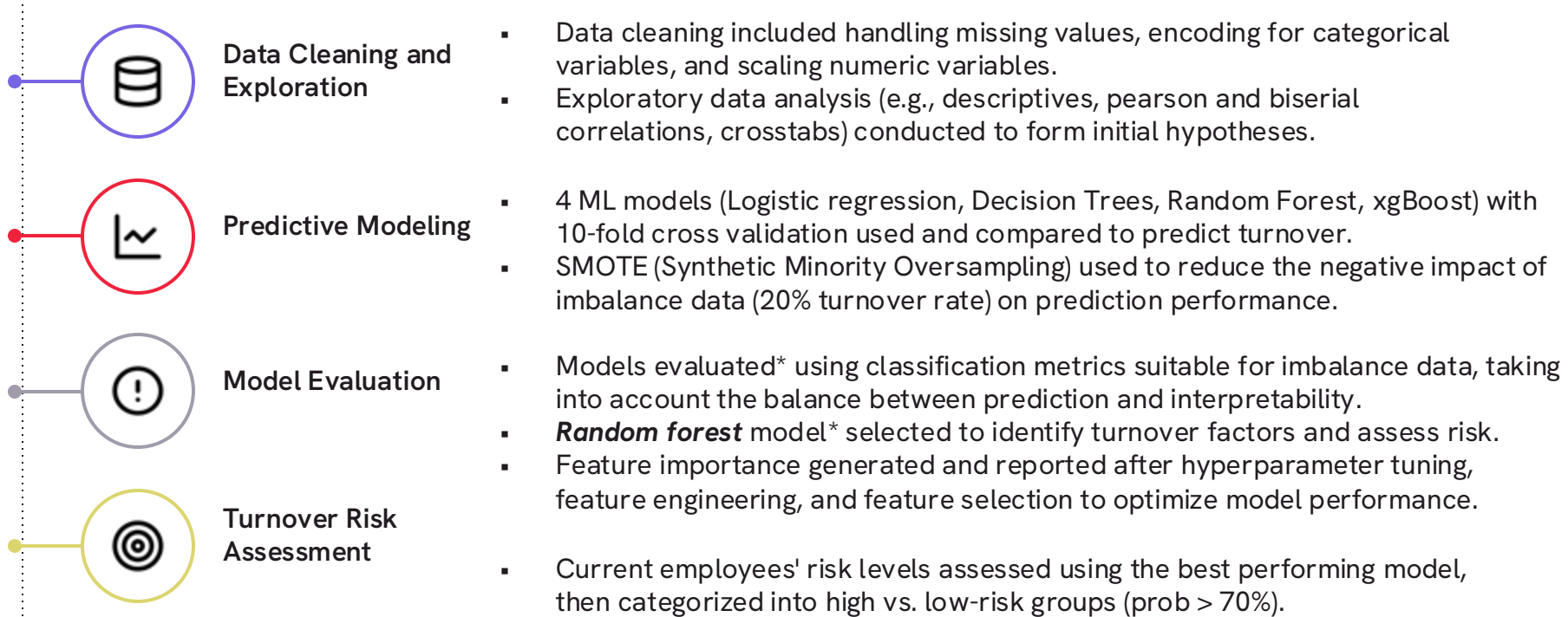- Race / Ethnicity*
- Region*

Other HRIS Data

- Work Arrangement (onsite vs. hybrid vs. remote)*
- Performance Rating
- Work hours per Week
- Compensation
- Months since Promotion
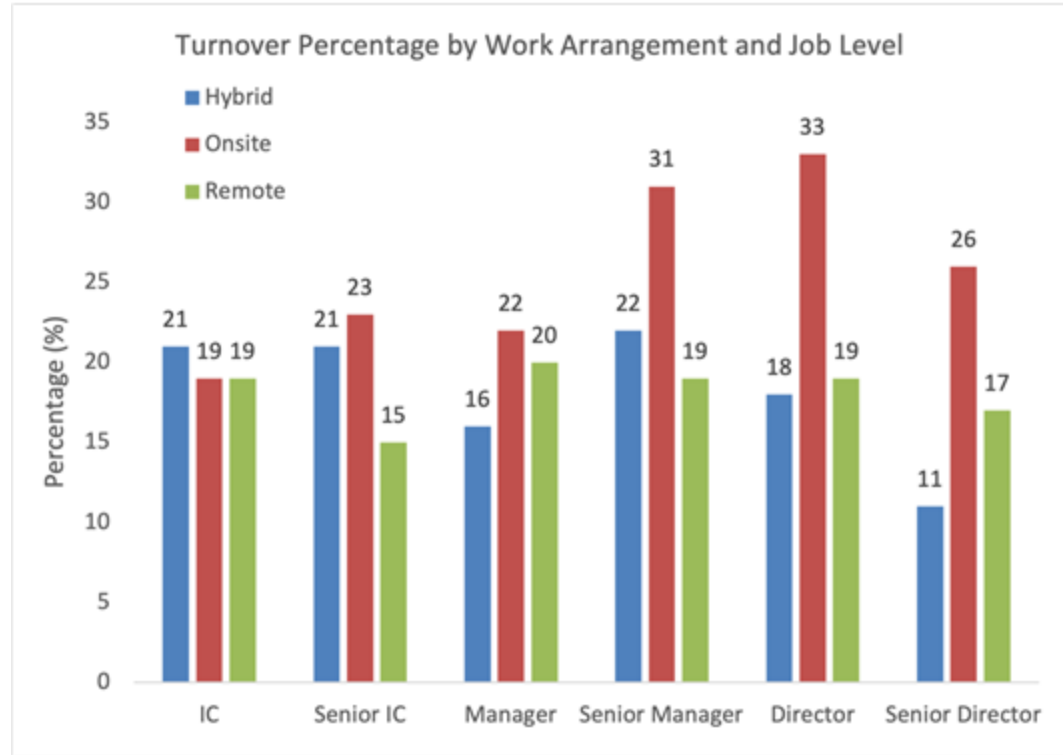
indicates categorical variables; the rest are numeric.

Data-Driven Approach to Predicting Turnover

Data Cleaning and Exploration

- Data cleaning included handling missing values, encoding for categorical variables, and scaling numeric variables.
- Exploratory data analysis (e.g., descriptives, pearson and biserial correlations, crosstabs) conducted to form initial hypotheses.

Predictive Modeling

- 4 ML models (Logistic regression, Decision Trees, Random Forest, xgBoost) with 10-fold cross validation used and compared to predict turnover.
- SMOTE (Synthetic Minority Oversampling) used to reduce the negative impact of imbalance data (20% turnover rate) on prediction performance.

Model Evaluation

- Models evaluated* using classification metrics suitable for imbalance data, taking into account the balance between prediction and interpretability.
- *Random forest* model* selected to identify turnover factors and assess risk.
- Feature importance generated and reported after hyperparameter tuning, feature engineering, and feature selection to optimize model performance.

Turnover Risk Assessment

- Current employees' risk levels assessed using the best performing model, then categorized into high vs. low-risk groups (prob > 70%).

* See the Appendix for model evaluation metrics and comparison.

Data Exploration and Initial Insights

- Employee experience variables and performance ratings lean to the positive end.

- All employee experience variables are moderately to strongly correlated.

- No statistically significant difference is found in demographic segmentations except for work arrangement.

- Specific groups have higher turnover, including senior managers, directors, Operations and Sales, Asians, Latin Americas, and onsite employees.

- Onsite workers have higher turnover; senior leaders (SM, D, SD) who work onsite have statistically significantly higher turnover.

- No clear relationship exists between team size and turnover ($r = -0.01$); a few larger teams have higher turnover as outliers.

Turnover Percentage by Work Arrangement and Job Level

■ Hybrid
■ Onsite
■ Remote

Job Level	Hybrid	Onsite	Remote
IC	21	19	19
Senior IC	21	23	15
Manager	16	22	20
Senior Manager	22	31	19
Director	18	33	19
Senior Director	11	26	17

Y-axis: Percentage (%) — scale 0, 5, 10, 15, 20, 25, 30, 35

Key Insights on Reducing Employee Turnover

Top Predictors for Employee Turnover

- **Employee Experience** factors are more significant predictors than pay equity, demographics, and other variables.
- **Belonging**: Lack of a sense of belonging significantly drives turnover ($r = -0.38$).
- **Team Size**: Even though team size predicts turnover, it is not correlated with turnover ($r = -0.01$).
- **Engagement**: Low engagement strongly indicates potential attrition ($r = -0.37$).
- **Growth Opportunity**: Limited career growth leads to higher turnover risk ($r = -0.39$).
- **Recognition**: Lack of recognition demotivates employees, leading to exits ($r = -0.29$).

Impact and Actionable Recommendations

Belonging

*Higher belonging is **50%** less likely to be at risk, **8x** more likely to be high performers.*

- Organize regular, informal onsite/virtual networking events (e.g., games, coffee chats, lunch and learn)
- Redesign onboarding by adding a 90-day onboarding journey with weekly check-ins and peer mentors
- Provide sponsorship and a dedicated budget for ERGs

Engagement

*Higher engagement is **2x** less likely to be at risk, **6x** more likely to be high performers.*

- Encourage taking PTOs and launch wellbeing initiatives to encourage work-life balance
- Align employees' work and growth with company values and mission to create a sense of purpose
- Assess current work arrangements to pinpoint groups and job levels for onsite work optimization.

Growth

*Higher growth is **25%** less likely to be at risk, **4x** more likely to be high performers.*

- Develop clear promotion criteria and career pathways
- Offer ongoing training and skill-building workshops and evaluate results
- Implement mentorship matching programs for both peer and top-down mentorship
- Create a 'gig work marketplace' to offer stretch projects across teams

Recognition

*Higher growth is **2x** less likely to be at risk, **4x** more likely to be high performers.*

- Implement a online recognition platform to provide shout-outs
- Introduce 5-minute 'kudos moments' in meetings for frequent formal appreciation
- Launch award initiatives to recognize stellar contributions
- Offer personalized rewards to make recognition meaningful and tangible

Measuring Impact and Next Steps

Action planning based on effort vs. impact, and strategic alignments with business goals and company values

Short Term Quick Wins (0-3 months)

- Organize informal networking events
- Introduce 'Kudos Moments' in meetings
- Encourage PTOs & wellbeing initiatives
- Create 'Buddy Systems' for new hire
- Support ERGs with budget & sponsors

Mid Term (3-6 months)

- Create a 90-day onboarding journey
- Use advanced analytics to pinpoint employees to work onsite
- Set up mentorship matching programs
- Offer ongoing training workshops

Long-term Big Bets (6-12 months)

- Implement Gig Work Marketplace
- Implement a recognition platform
- Maintain and refine the predictive model to identify at-risk employees in real-time.

Success Metrics & Expected Outcomes

- **Workforce**:
  - *20% increase in Employee engagement* (e.g., job satisfaction, eNPS)
  - 20% decrease in *turnover intention*
  - *10% YOY decrease in turnover*
- **Customer**:
  - *20% increase in customer satisfaction* (NPS)
- **Business**:
  - *10% reduction in recruitment costs*
  - *10% YOY increase in profits*

Evaluation Methods

- Quantify success of each initiative
- Monitor effects on the at-risk groups
- Use quarterly surveys to track EX over time
- Create dashboards to track real-time turnover
- Partner with HR, Sales, and Finance to obtain hiring, customer, and financial data and assess change