Human-AI Interaction | 2020 Fall

Exploring Employee Attrition for Talent Retention

Final Presentation

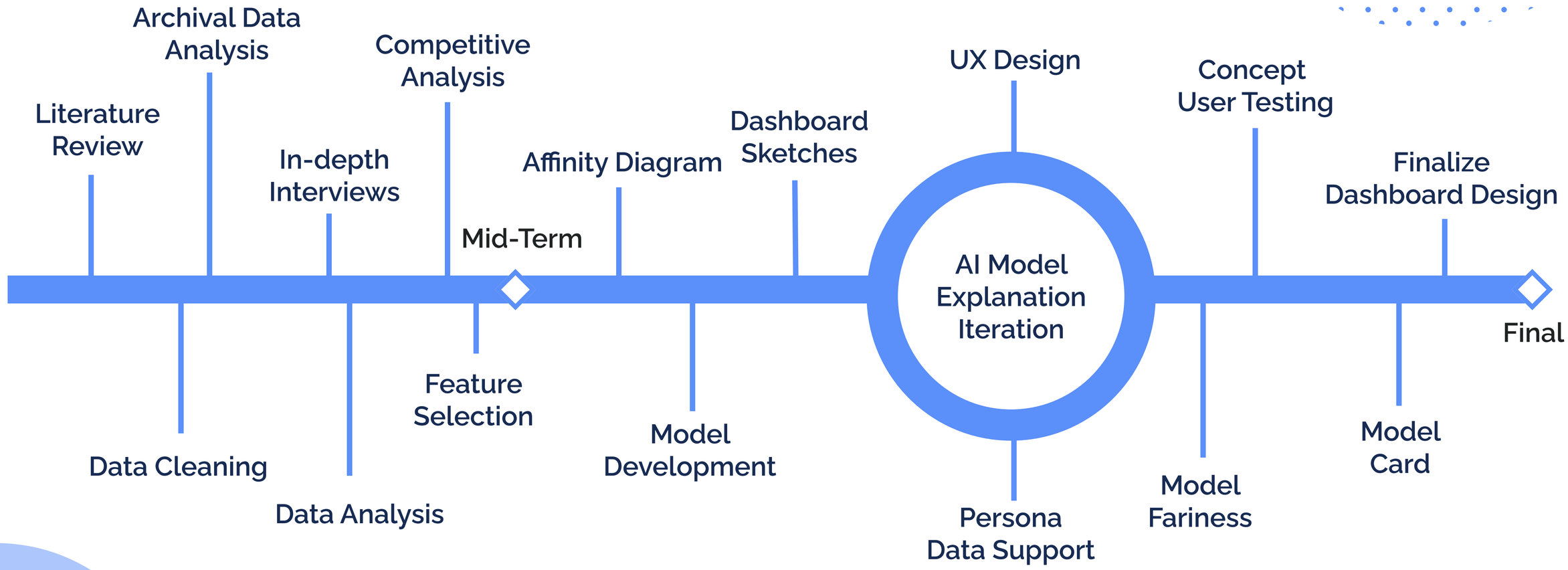
Janet Dai, Yi Wang, Dianne Kim

Abstract/Overview

Attrition is a widespread and expensive problem for companies
Costs of not only recruiting, interviewing, onboarding, training but also loss of productivity and output when an employee leaves unexpectedly
Cost of replacement = 20% salary of the lost employee

How might we create a solution that could help companies predict when employees are likely to leave...with the goal of providing a strategy to retain them?

Project Timeline



Literature Review

AI applications have been developed to try to tackle workers who are planning to leave their jobs, with companies like IBM. Armed with this knowledge, managers can take strategic actions to encourage that employee to stay. By focusing specifically on high performers with in-demand skills, IBM says it has saved $300 million in retention costs.

https://business.linkedin.com/talent-solutions/blog/artificial-intelligence/2019/IBM-predicts-95-percent-of-turnover-using-AI-and-data

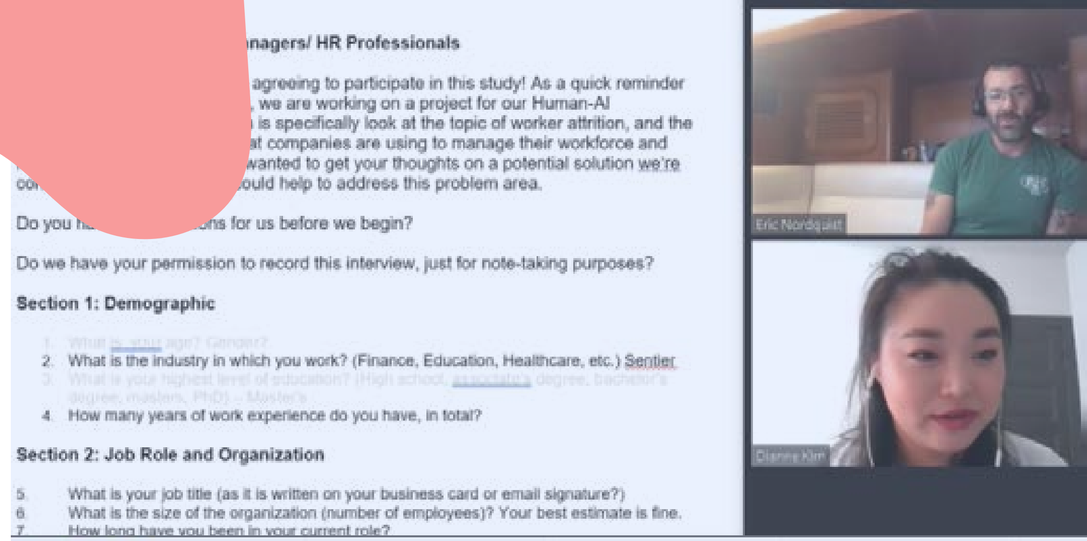
In-depth Interview

8 Participants (3 managers, 1 HR , 4 Employees)

- The choice to leave a job can be very complex and subjective, but the relationship with managers and co-workers is the most common one based on the interviews

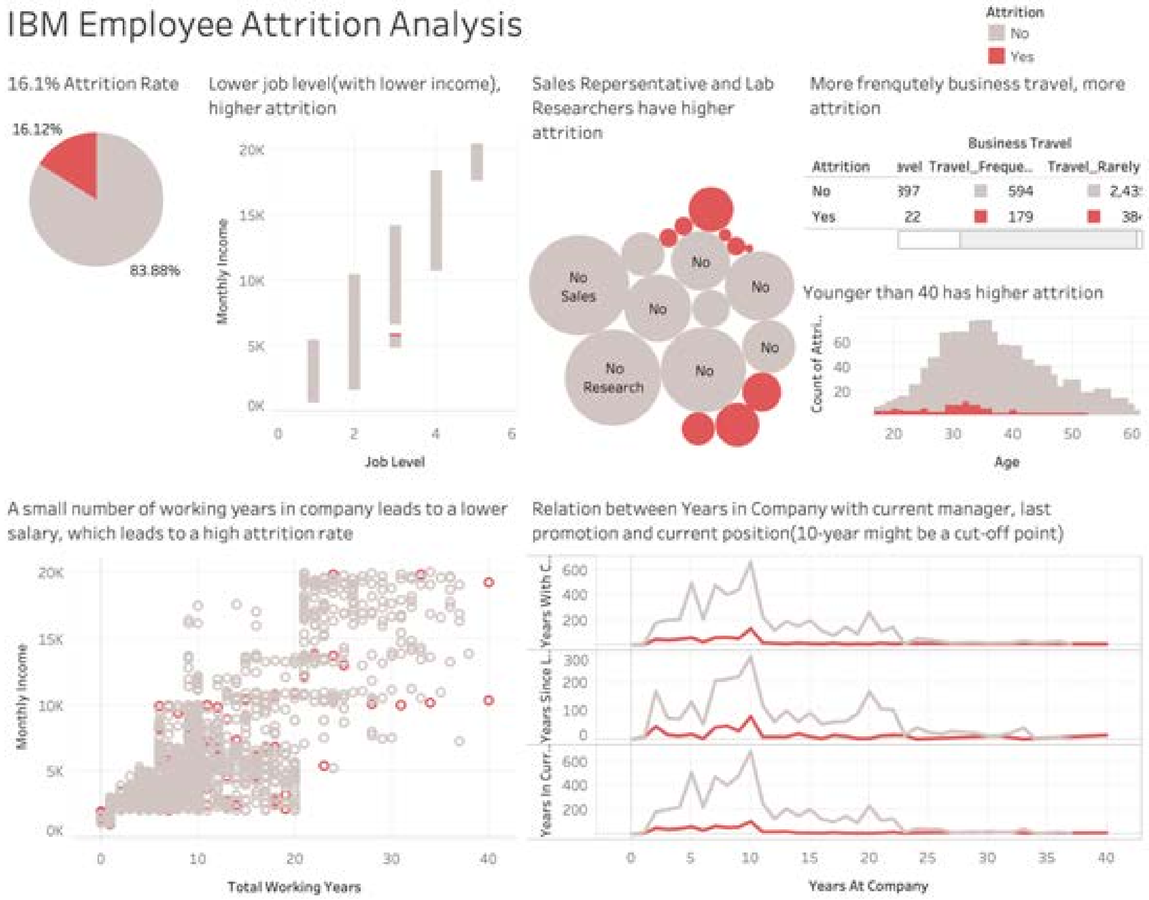
- Trust and explanation is super important for this kind of tool

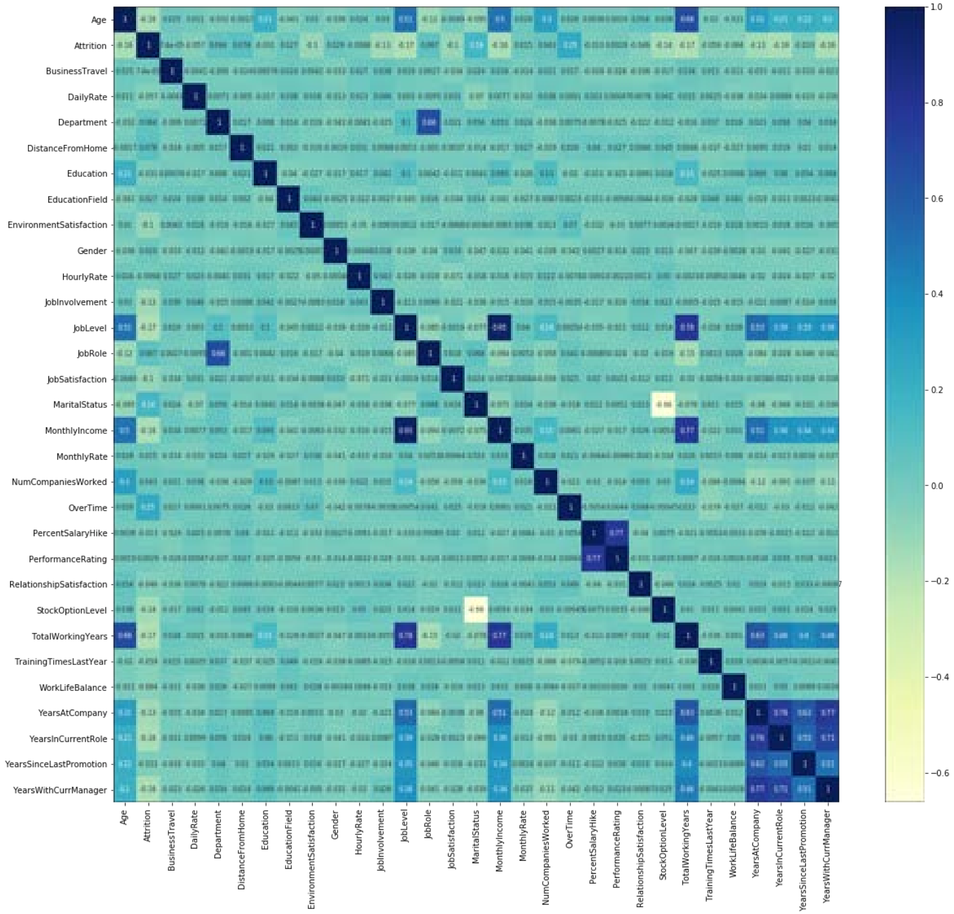
- Attrition prediction tools are ideal for large companies where HR are not familiar to all the employees



Exploratory Data Analysis

IBM Kaggle dataset: 35 features, included 1427 employees





Model Card

Logistics Regression has the best accuracy score is 87.9% with the best hyperparameter

Employee Attrition Prediction Model

The model predicts the risk of employee attrition based on their demograhic and working information and returns a probability score to assist Human Resouse Managers to evaluate their employees.
In this model card, you can understand about how the model perfroms on 28 features and learn about how to provide meaningful recommendations to HR based on the feature importances calculate from each input.

MODEL DESCRIPTION

Input: Employee's basic working information (IBM dataset features)

OverTime, JobLevel, MonthlyIncome, StockOptionLevel, TrainingTimesLastYear, BusinessTravel, DailyRate,EnvironmentSatisfaction, JobInvolvement,JobSatisfaction,WorkLifeBalance, Department, DistanceFromHome, PerformanceRating, EducationField, JobRole, NumCompaniesWorked, PercentSalaryHike, RelationshipSatisfaction, TotalWorkingYears, YearsAtCompany, YearsInCurrentRole, YearsSinceLastPromotion, YearsWithCurrManager

Output: The evaluation for the employee attrition decision. For each persona, the model will provide a leaving probability score.
- Attrition final results
- Probability
- Feature Importance
- Confidence score

LIMITATIONS

This model is established on the IBM data set, which limits the input of HR to be within the range of features included in the dataset. Although these features are splendid, it cannot contain all relevant factors. Different enterprises are hindered by different factors, resulting in unstable forecast results.

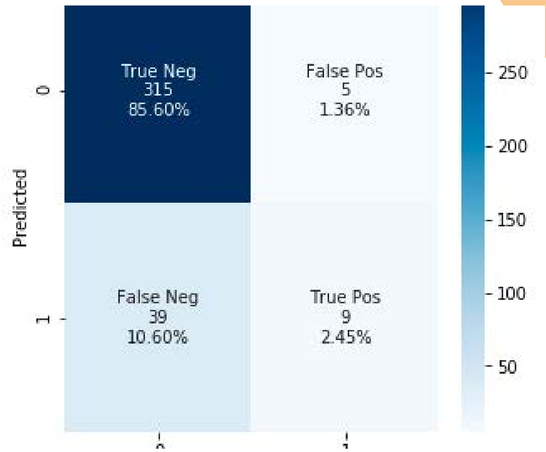
Meanwhile, the limitation of specific dataset will lead users to provide corresponding suggestions based on the conclusion of the importance of features, which requires us to further collect relevant data to make up the corresponding threaholds for each feature.

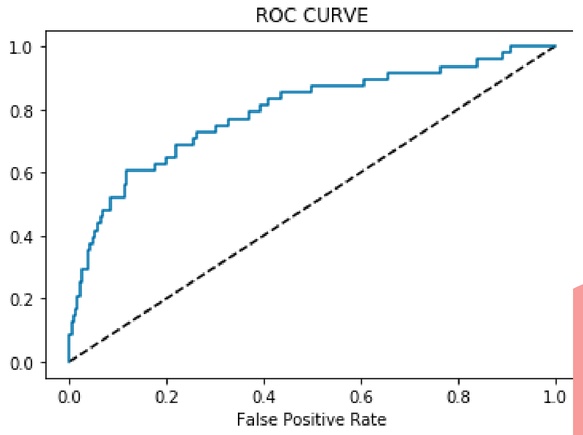
TRADE-OFF

There are some personal demographic information in the dataset, including **Age, Maritial Status and Gender**.

AI Model

- Convert Binary Classification into Probability score in order to present the risk of employee attrition
- Confusion Matrix analysis - Precision, Recall & ROC
- SMOTE Method for Imbalanced Data
- Limitations: The dataset itself has some limitations. (Feature dissatisfaction, lack of thresholds for recommendation, unclear features categories)

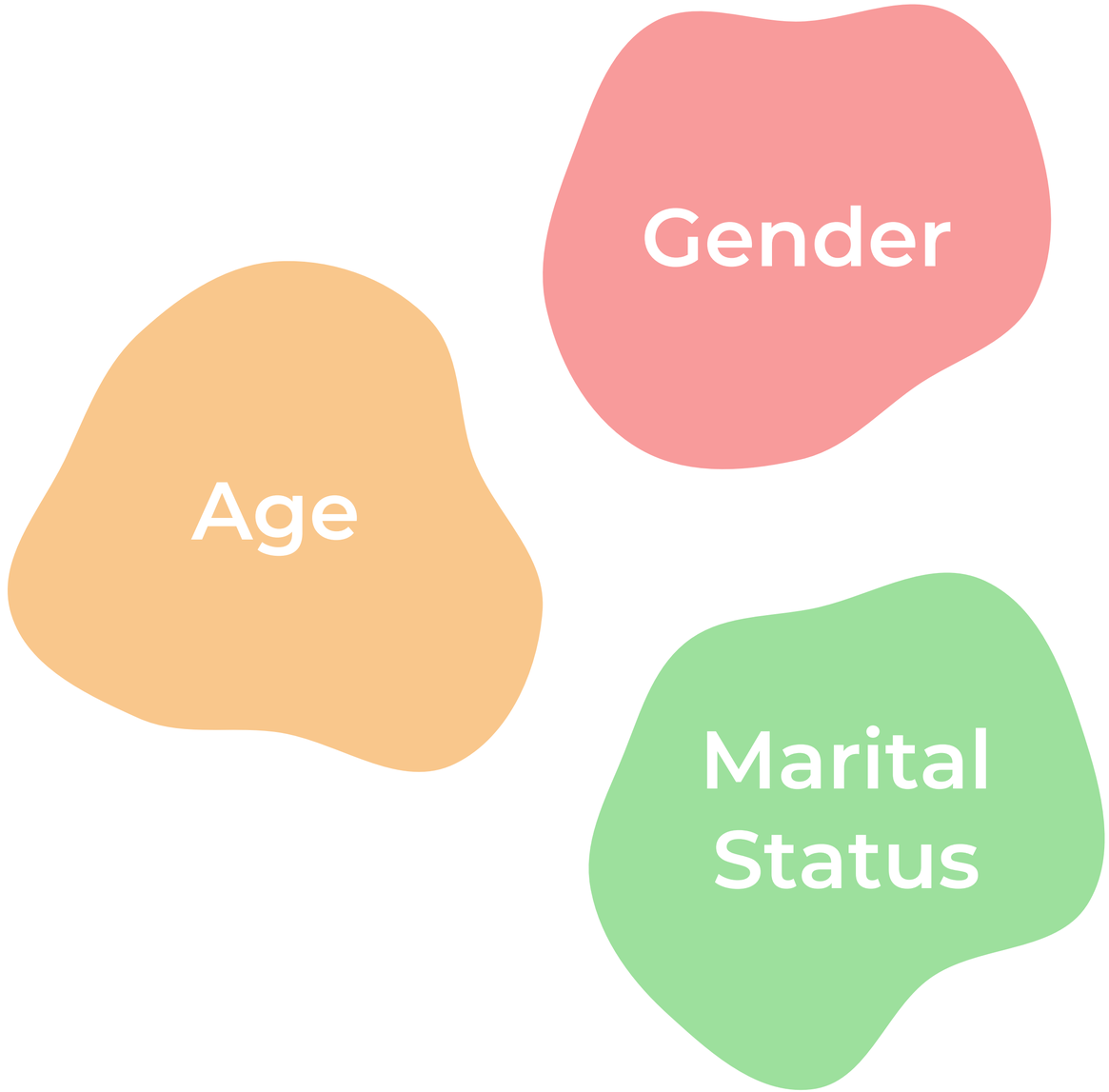


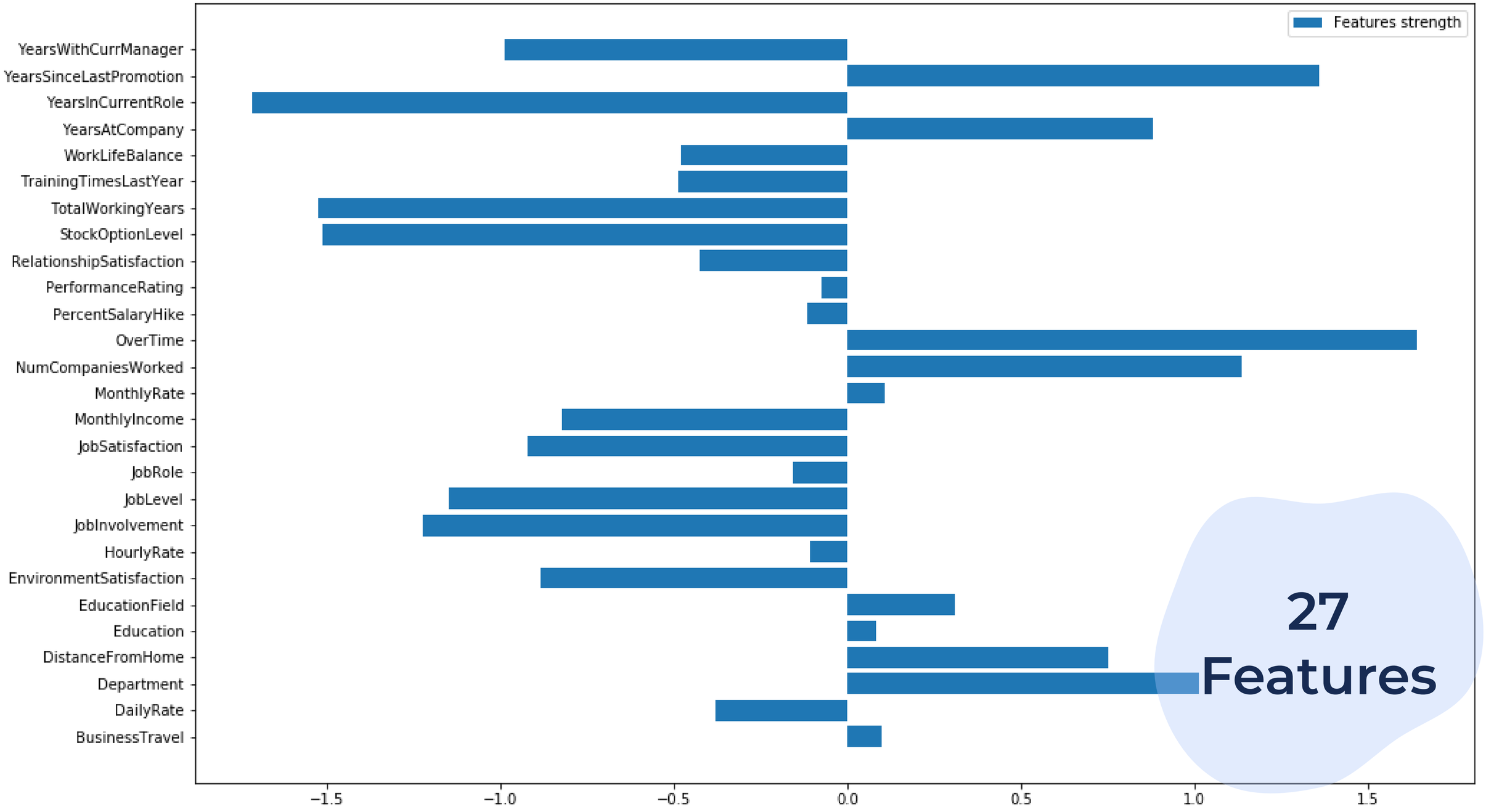


Fairness/Bias

Excluded employee sensitive demographic features as they are illegal to collect at workspaces

AI fairness metrics
- Equalized odds
- Test Fairness
- Counterfactual Fairness

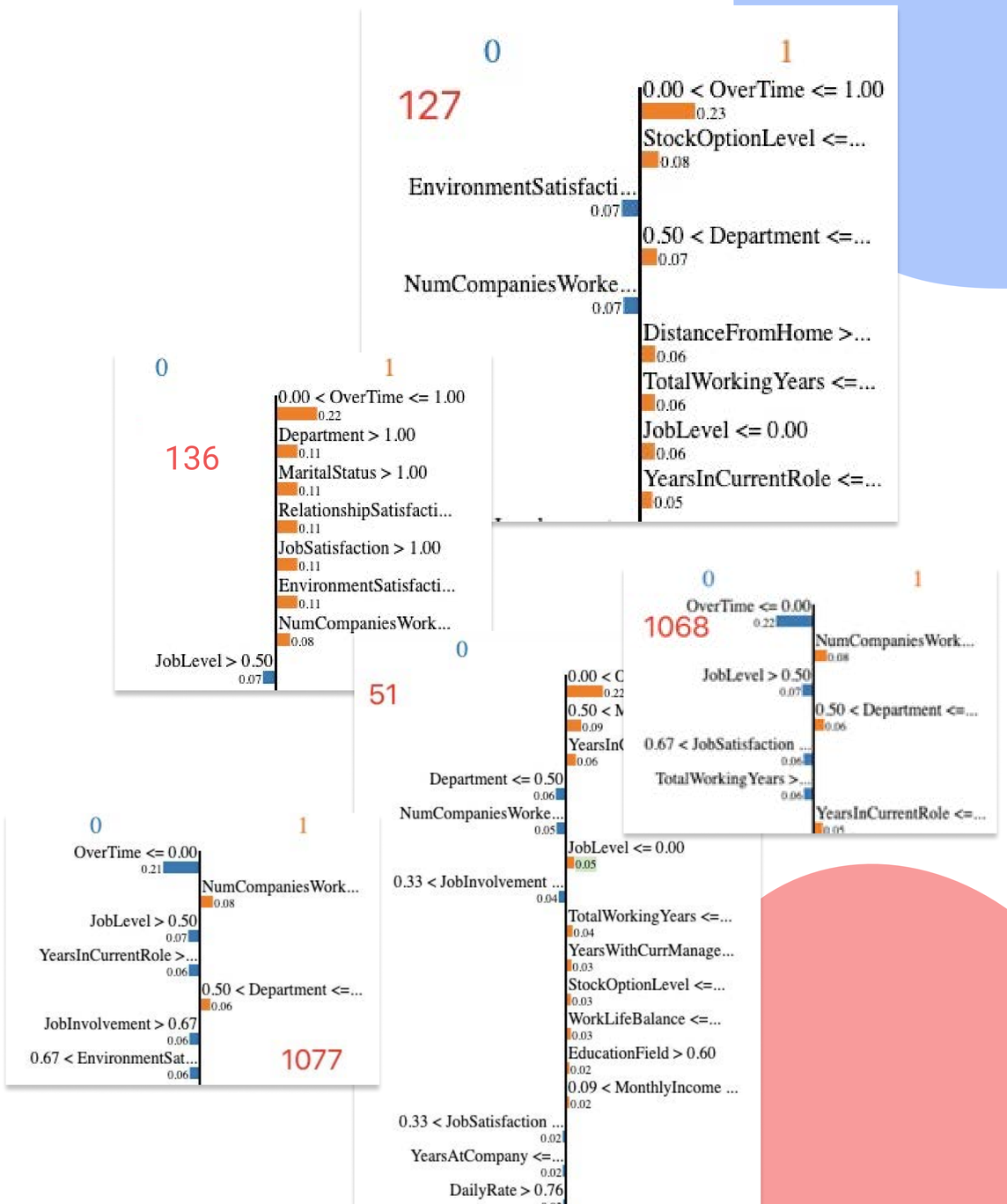



Features strength
YearsWithCurrManager
YearsSinceLastPromotion
YearsInCurrentRole
YearsAtCompany
WorkLifeBalance
TrainingTimesLastYear
TotalWorkingYears
StockOptionLevel
RelationshipSatisfaction
PerformanceRating
PercentSalaryHike
OverTime
NumCompaniesWorked
MonthlyRate
MonthlyIncome
JobSatisfaction
JobRole
JobLevel
JobInvolvement
HourlyRate
EnvironmentSatisfaction
EducationField
Education
DistanceFromHome
Department
DailyRate
BusinessTravel
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5

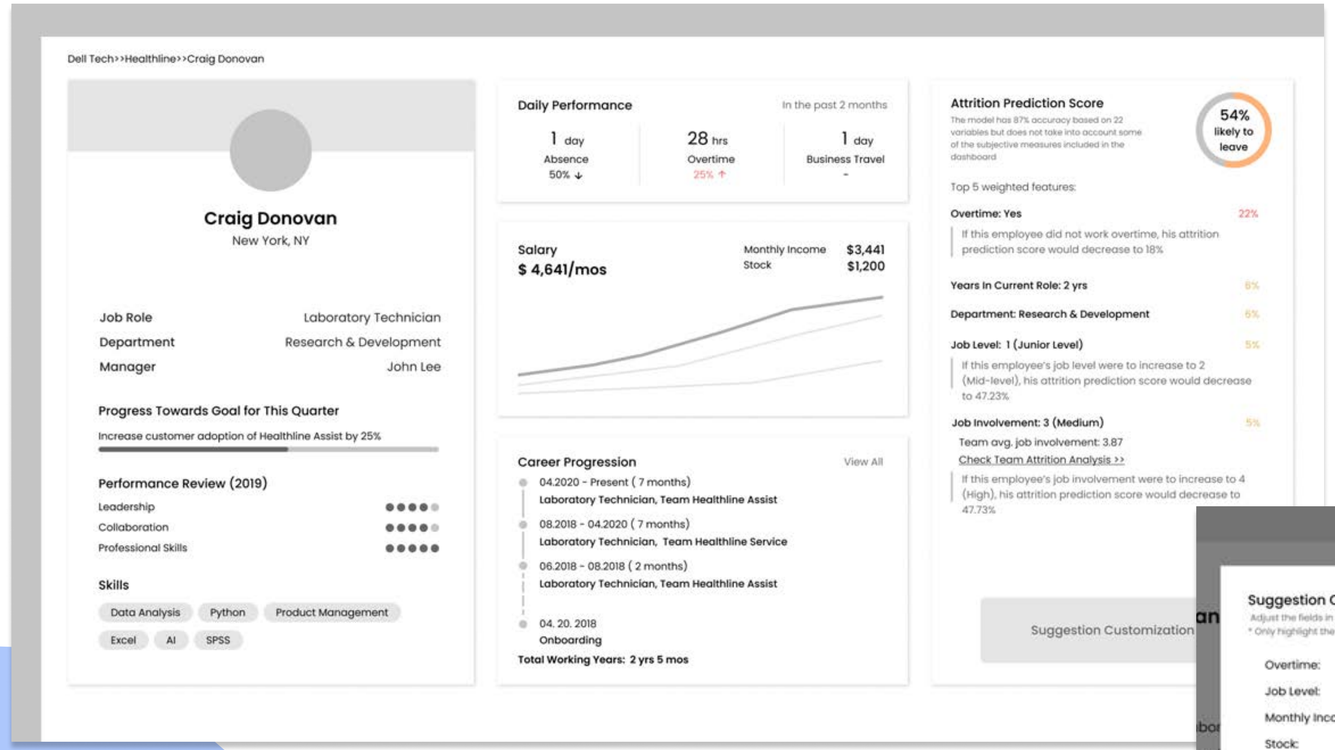
27 Features

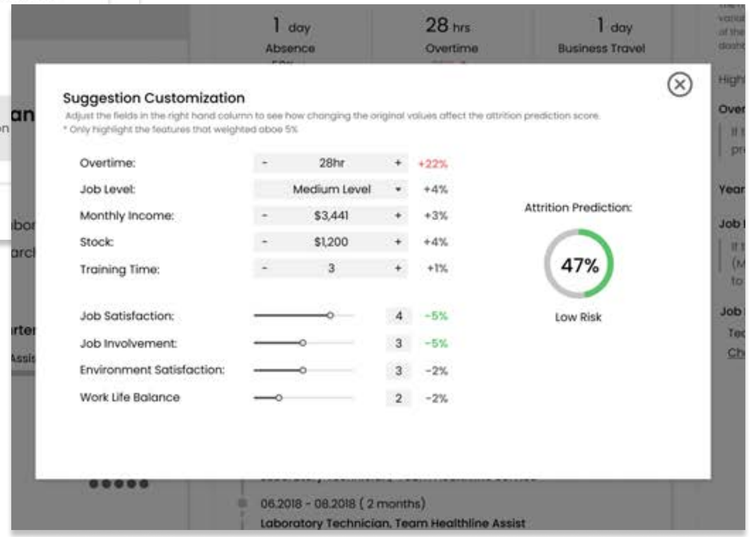
Persona

- Overtime is always the top reason for attrition
- Salary ranks much lower than expected, and has a low impact on attrition
- Easy to have stereotypes on unactionable features
- Most features weigh lower than 5%



Dashboard Design



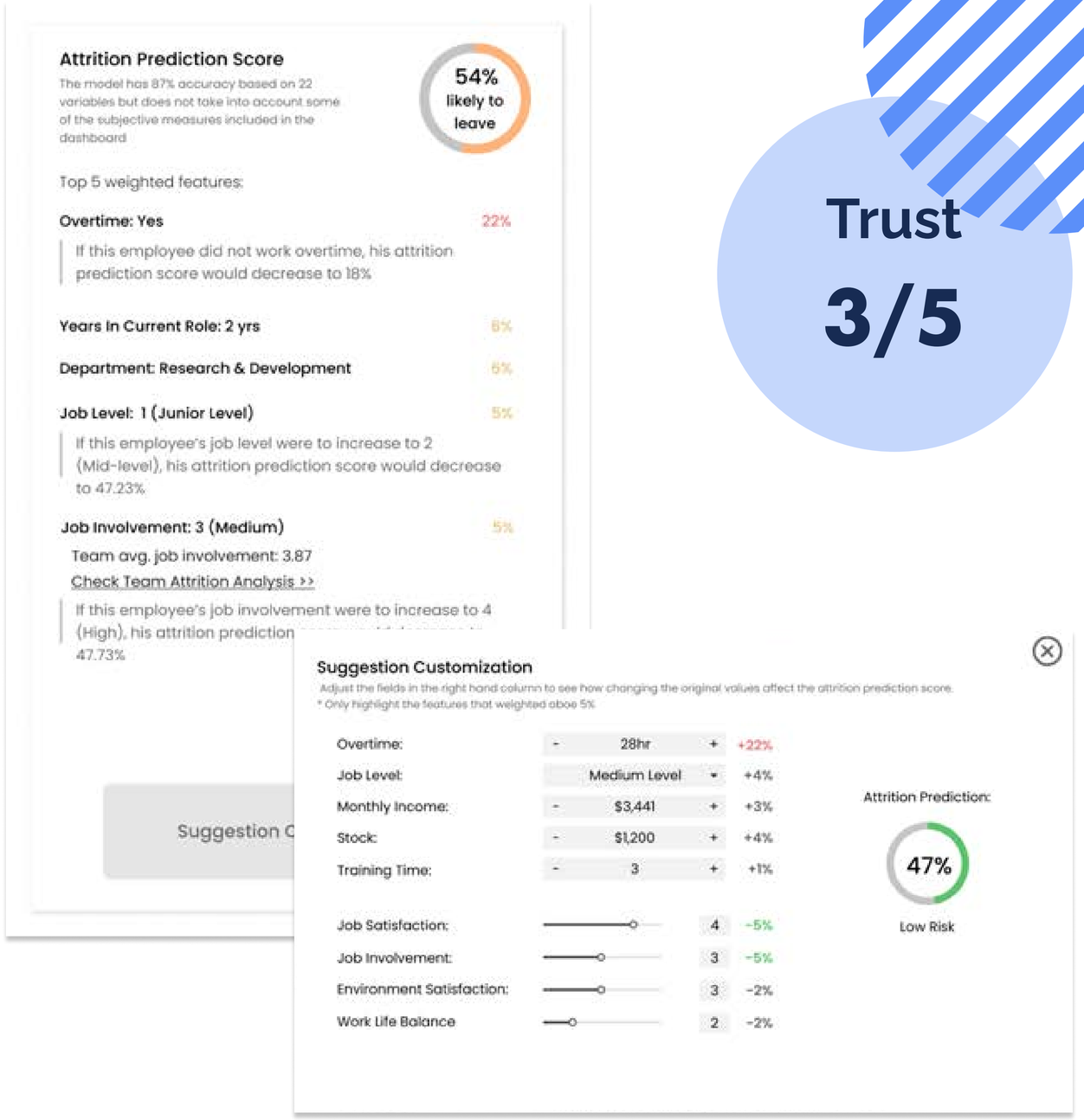


- Performance vs. Attrition
- Attrition Explanation
  - Percentage Risk
  - Highly Weighted Features
  - Understandable Language
  - Suggestion Customization

User Testing

3 users: 1 manager, 1 HR, 1 Employee

- Like the performance metric
- Outcome explanations themselves were clear but not understandable
- Customization was too granular and overwhelming to be useful
- Employee had a negative reaction



Trust

3/5

Personal Dashboard

Healthline Product Team>>Craig Donovan

Craig Donovan
New York, NY

Department	Sales
Job Title	Sales Representative
Manager	John Lee
Current Team	Healthline Product Team

Total Working Yrs	1 yrs 8mos
Yrs in Current Team	8 mos

Career Progression

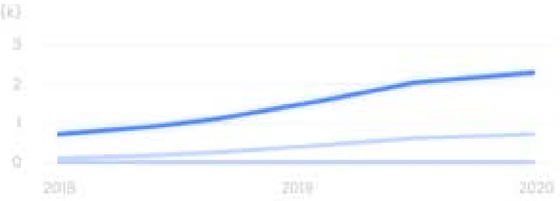
- 04.2020 - 11. 2020 (7 months)
  Sales Trainer , Healthline Assist Team
- 12.2019 - 04.2020 (5 months)
  Sales Trainer , Healthline Product Team

Behavior — In Past 2 Months

Absense (days)	Overtime (hours)	Business Trip (days)
0 ↓	28 ↑	2 -

Income

Salary	Stock	Bonus
$2,260/mo	Level 0	$568



Job Satisfaction

Annual Performance — Team Avg. Score: 76 — 87

Leadership	●●●●○
Communication	●●●●○
Professional Skills	●●●●○
Innovation	●●●○○

View Full Report>>

AI Attrition Prediction

*The model has 87% accuracy based on 22 quantitive variables but the attrition rate may also be related to other subjective reasons

65% likely to leave

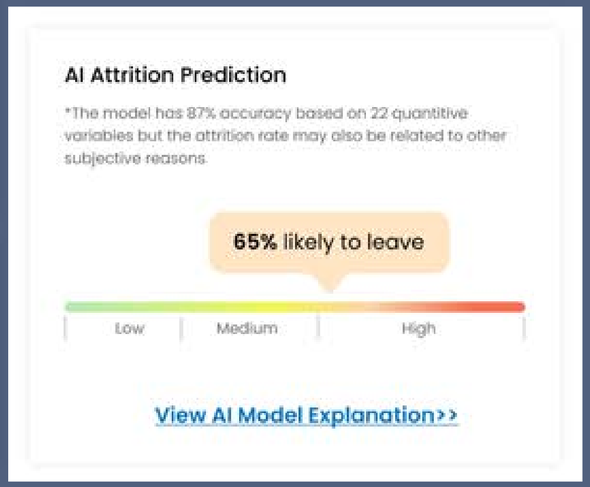
View AI Model Explanation>>

Personal Dashboard

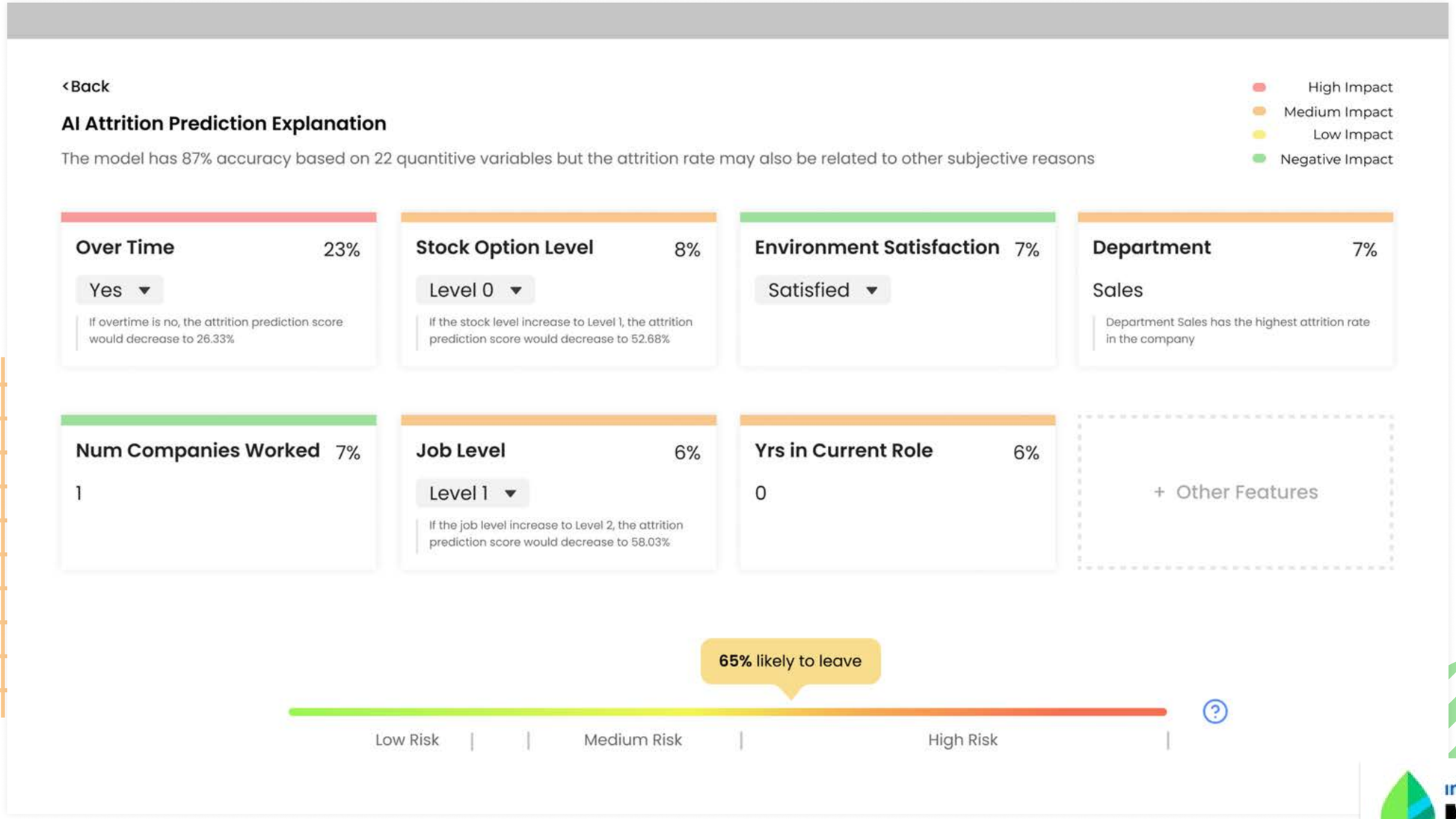
"I don't even know how to make sense of the 65% because everybody in the company has a 50/50 chance of leaving. "

Minimize the impact of attrition prediction
More as a support for the performance

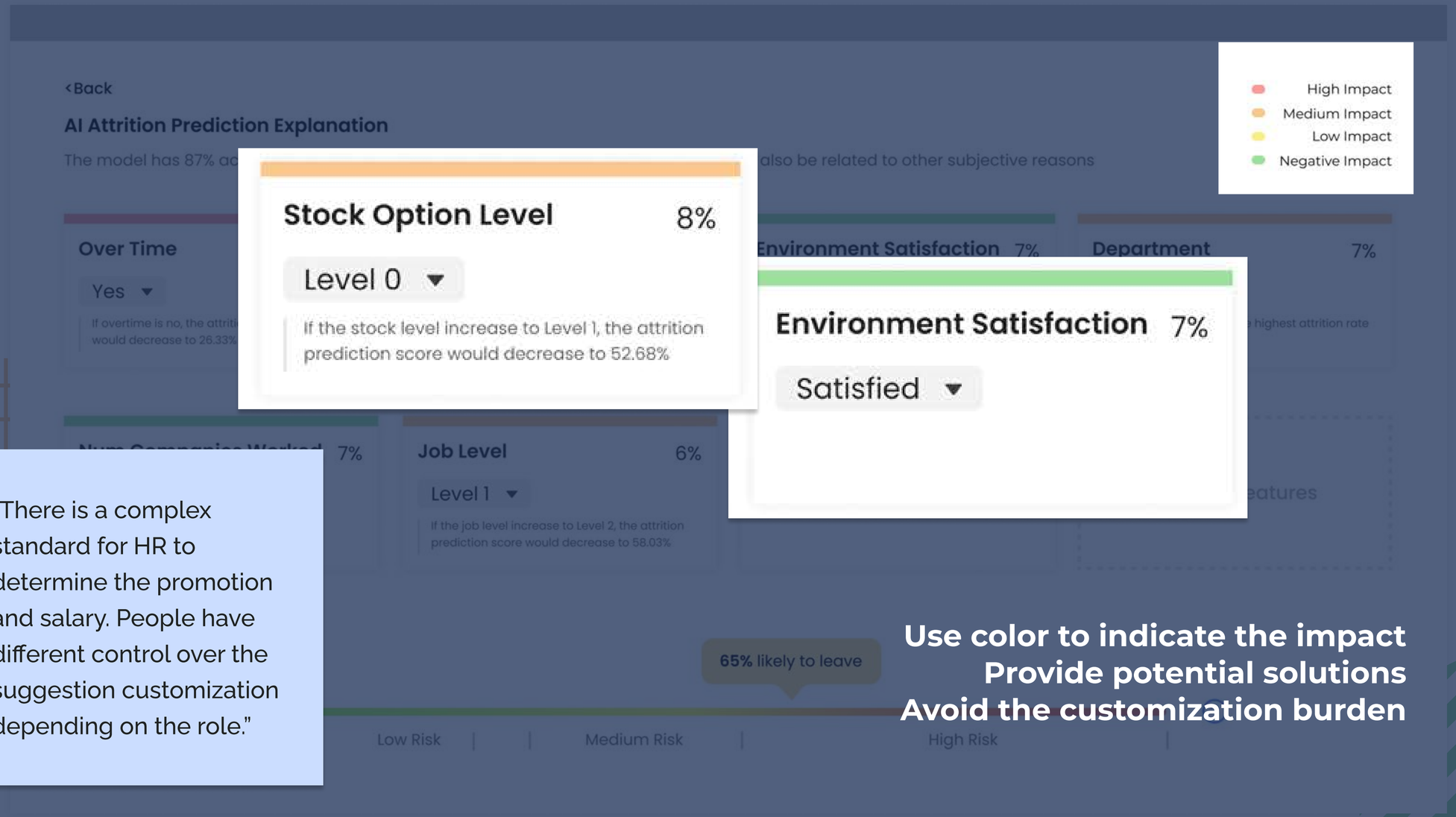
Most of the employees' attrition risk are lower than 20%



Attrition Explanation



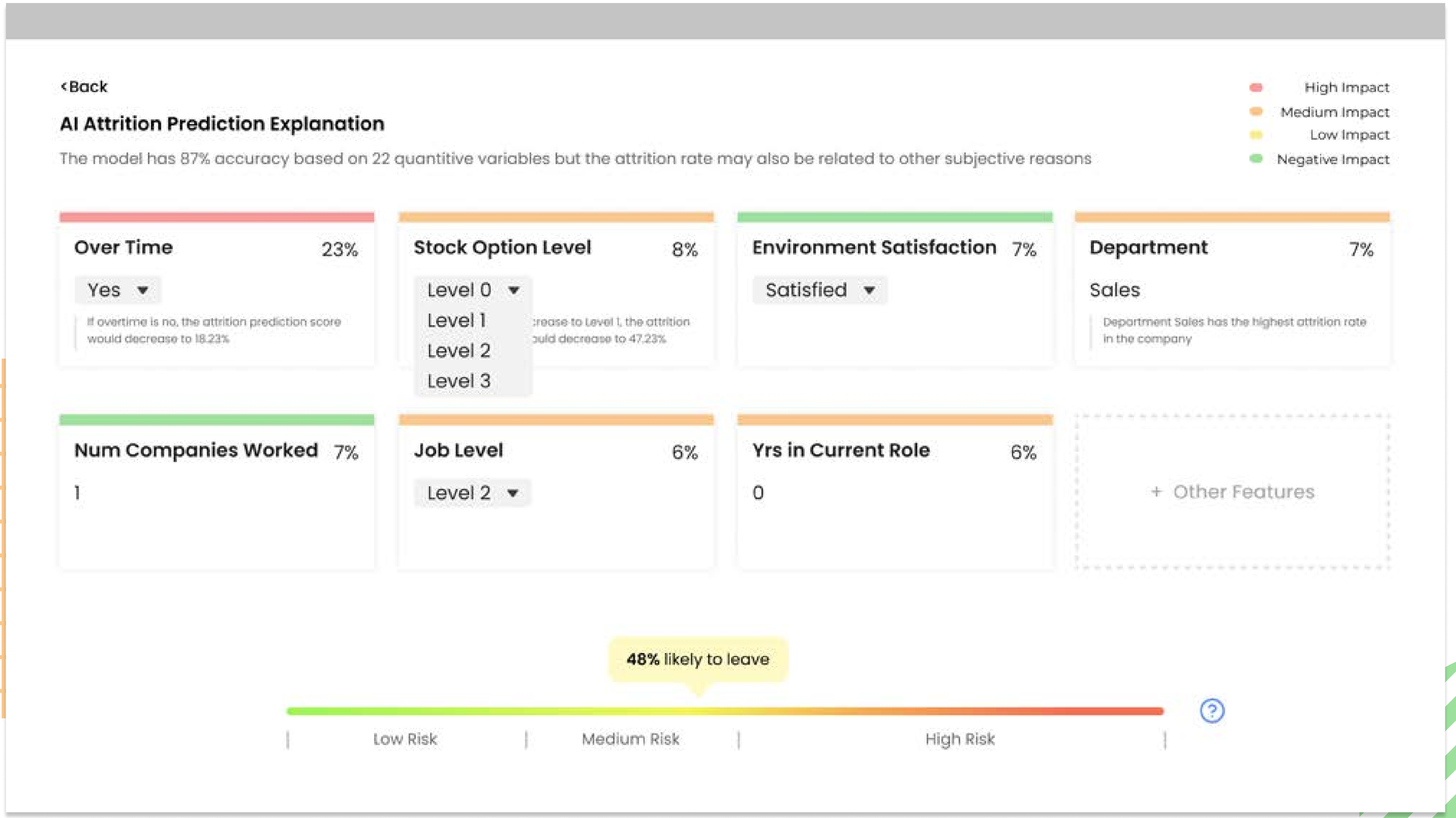
Attrition Explanation



"There is a complex standard for HR to determine the promotion and salary. People have different control over the suggestion customization depending on the role."

Use color to indicate the impact
Provide potential solutions
Avoid the customization burden

Attrition Explanation



Team Dashboard

Healthline Product Team

This team is responsible for connecting doctors, hospitals, researchers and patients on a shared technology platform to improve the patient experience

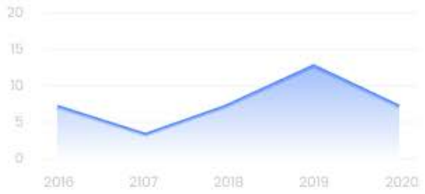
Current Employees
32

Attritioned Employees 2020
4 12%↓

Performance Rate 2019-2020
86 25%↑

Avg. Yrs in the Team
4.78

Attrition History 2016-2020



Top Reasons for Attrition

1. Monthly Income(28)
2. Years with Current Managers(21)
3. Job Level(18)
4. Total Years(17)
5. Job Satisfaction(15)

Top Jobs Roles for Attrition

1. Sales Represetitves(8)
2. Software Developer(7)
3. Product Manager(4)
4. Sales Manager(3)
5. Senior Product Manager(3)

Satisfaction Survey Rate 2019-2020

Job Involvement 12% ↗
3.87
0 — 5

Relationship Satisfaction 8% ↗
3.21
0 — 5

Work Life Balance 7% ↘
3.37
0 — 5

Current Employees | Attritioned Employees | Filter

Name	Job Role	Yrs in the Team	Performance Rate	Risk
Robert Fox	President of Sales	5 Yrs 2 mos	87	Low Risk
Cameron Williamson	Marketing Coordinator	5 Yrs 2 mos	78	Medium Risk
Brooklyn Simmons	Marketing Coordinator	4 Yrs 8 mos	83	Low Risk
Darlene Robertson	Software Development Manager	3 Yrs 2 mos	90	Low Risk
Craig Donovan	Sales Representative	1 Yrs 2 mos	87	High Risk

Guidelines for Future Use

Guidelines for HR

- Ideal for large companies where HR is a “black box” and HR does not have a 1:1 relationship with every employee or see dynamics within various teams
- Team-level view of attrition may be more valuable than individual-level
- Best for flagging flight risk for senior-level or above
- Neither employees nor managers would not have access to the platform (but do employees have a right to know?)

“I would keep it to HR because we have this responsibility to be responsible with the data. As much as you can train your managers, they’re people and they’re fallible.”
-HR professional

Implications for Future Use

False positive/negative scenarios

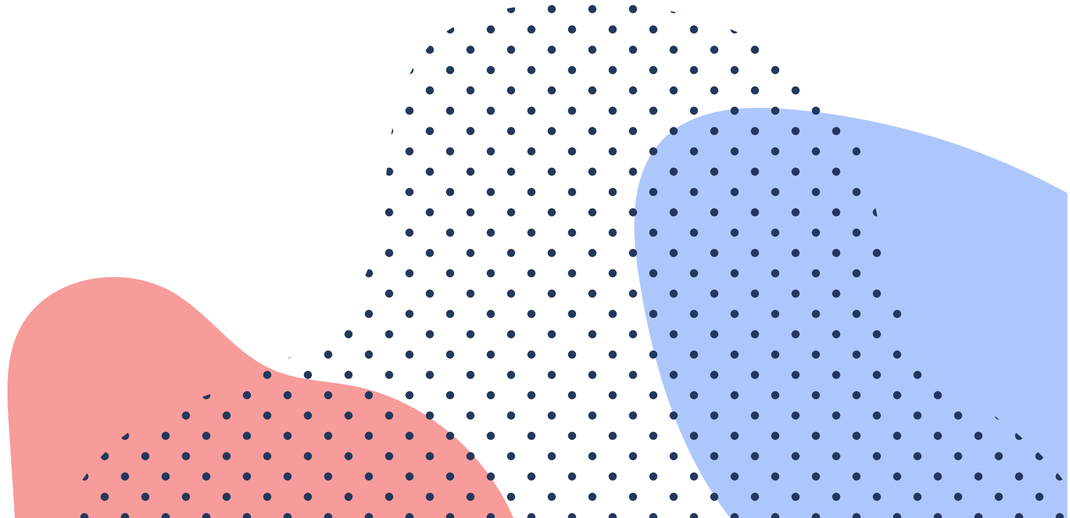
"How do you explain the idea of an AI thing that's judging you behind the books without showing them so they don't freak out? Also, you show it to them and they'll know what they can do to game it. They'll know the measures and then it becomes irrelevant."
-Manager

Challenges & Lessons Learned

- Limitations with the dataset
- Communication with technical/non-technical
- How to design for data

Thank you!

Questions?



Human-AI Interaction | 2020 Fall

Janet Dai, Yi Wang, Dianne Kim