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

- Archival Data Analysis
- Literature Review
- In-depth Interviews
- Competitive Analysis
- Affinity Diagram
- Dashboard Sketches
- Mid-Term
- Feature Selection
- Model Development
- Data Cleaning
- Data Analysis
- UX Design
- Concept User Testing
- Finalize Dashboard Design
- Final
- AI Model Explanation Iteration
- Persona Data Support
- Model Fariness
- Model Card
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.
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
Logistics Regression has the best accuracy score of 87.9% with the best hyperparameter.
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)
Excluded employee sensitive demographic features as they are illegal to collect at workspaces

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

Behavior
- Absence (days): 0 (down)
- Overtime (hours): 28 (up)
- Business Trip (days): 2 (down)

In Past 2 Months

Annual Performance
- Team Avg. Score: 76
- Leadership:
- Communication:
- Professional Skills:
- Innovation:

Income:
- Salary: $2,250/mo
- Stock: Level 0
- Bonus: $580

Career Progression
- 04.2020 - 11.2020 (7 months)
- Sales Trainer, Healthline Assist Team
- 12.2018 - 04.2020 (8 months)
- Sales Trainer, Healthline Product Team

Job Satisfaction
- Job Involvement: 3
- Relationship Satisfaction: 4

AI Attrition Prediction
- 85% likely to leave

View Full Report >>

View AI Model Explanation >>
Most of the employees' attrition risk are lower than 20%
Attrition Explanation

AI Attrition Prediction Explanation

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

- **Over Time**: 23%
  - **Yes**: If overtime is no, the attrition prediction score would decrease to 26.33%

- **Stock Option Level**: 8%
  - **Level 0**: If the stock level increase to level 1, the attrition prediction score would decrease to 52.68%

- **Environment Satisfaction**: 7%
  - **Satisfied**: If the job level increase to level 2, the attrition prediction score would decrease to 55.03%

- **Num Companies Worked**: 7%
  - **1**:

- **Job Level**: 6%
  - **Level 1**

- **Yrs in Current Role**: 6%
  - **0**:

**Department**

- **Sales**: 7%
  - Department Sales has the highest attrition rate in the company

**65% likely to leave**

- **Low Risk**: Low Risk
- **Medium Risk**: Medium Risk
- **High Risk**: High Risk

*High Impact* | *Medium Impact* | *Low Impact* | *Negative Impact*
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.
**Attrition Explanation**

**AI Attrition Prediction Explanation**

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

<table>
<thead>
<tr>
<th>Feature</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Over Time</td>
<td>23%</td>
</tr>
<tr>
<td>Num Companies Worked</td>
<td>7%</td>
</tr>
<tr>
<td>Stock Option Level</td>
<td>8%</td>
</tr>
<tr>
<td>Environment Satisfaction</td>
<td>7%</td>
</tr>
<tr>
<td>Department</td>
<td>7%</td>
</tr>
<tr>
<td>Job Level</td>
<td>6%</td>
</tr>
<tr>
<td>Yrs in Current Role</td>
<td>6%</td>
</tr>
</tbody>
</table>

- **Over Time**: Yes
  - If overtime is no, the attrition prediction score would decrease to 18.23%.

- **Stock Option Level**:
  - Level 0
  - Level 1: Attrition rate would decrease to 47.23%.
  - Level 2
  - Level 3

- **Environment Satisfaction**:
  - Satisfied

- **Department**:
  - Sales: Department Sales has the highest attrition rate in the company.

48% Likely to leave

48% Likely to leave

Low Risk | Medium Risk | High Risk
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 for Attrition
1. Sales Representatives (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% ●
Relationship Satisfaction 8% ●
Work Life Balance 7% ●

Current Employees

<table>
<thead>
<tr>
<th>Name</th>
<th>Job Role</th>
<th>Yrs in the Team</th>
<th>Performance Rate</th>
<th>Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robert Fox</td>
<td>President of Sales</td>
<td>5 Yrs 2 mos</td>
<td>87</td>
<td>Low Risk</td>
</tr>
<tr>
<td>Cameron Williamson</td>
<td>Marketing Coordinator</td>
<td>5 Yrs 2 mos</td>
<td>78</td>
<td>Medium Risk</td>
</tr>
<tr>
<td>Brooklyn Simmons</td>
<td>Marketing Coordinator</td>
<td>4 Yrs 8 mos</td>
<td>83</td>
<td>Low Risk</td>
</tr>
<tr>
<td>Doriene Robertson</td>
<td>Software Development Manager</td>
<td>3 Yrs 2 mos</td>
<td>90</td>
<td>Low Risk</td>
</tr>
<tr>
<td>Craig Donovan</td>
<td>Sales Representative</td>
<td>1 Yrs 2 mos</td>
<td>87</td>
<td>High Risk</td>
</tr>
</tbody>
</table>

Filter
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?