

Predicting Employee Retention and Turnover in Last-Mile Delivery Logistics

Original Project Plan 10-20-2024

Research Question:

How can we predict employee retention and turnover, and what factors significantly influence the high turnover rates in the last-mile delivery logistics industry?

Expected Data Sources:

- **Internal Data:** The primary source will be internal company data, including employee demographics, job details, employment history, performance metrics, attendance records, salary, and reasons for termination. This will offer a comprehensive view of employee characteristics and their potential relationship to retention or turnover.
- **External Data (Optional):** External datasets, such as labor market trends and economic indicators, may also be integrated to enhance predictive power. Potential sources could include the Bureau of Labor Statistics and other relevant industry data.

Techniques Expected to Use:

The project will combine traditional survival analysis and machine learning techniques to model employee turnover, name key risk factors, and predict tenure:

- **Kaplan-Meier Estimator:** This technique will estimate employee retention rates over time, providing insights into the survival probabilities of different employee cohorts.
- **Cox Proportional Hazards Model:** This will assess the impact of multiple factors (e.g., job type, salary, and performance) on employee turnover risk while accounting for time-dependent variables.
- **Random Survival Forest (RSF):** A machine learning-based survival analysis technique to capture non-linear interactions between features, enhancing prediction accuracy.
- **Ensemble Learning Models** (e.g., Gradient Boosting): These models will improve the prediction of turnover risks by exploring complex relationships in the data.
- **Explainable AI Models** (e.g., SHAP or LIME): These will interpret the results from the models, explaining the contribution of each feature (e.g., job satisfaction or salary) to employee retention or turnover. This step ensures that business decisions are supported by transparent and interpretable insights.

Expected Results:

The expected outcome of this analysis is the development of predictive models that accurately forecast employee tenure and find key factors contributing to turnover. The models should help uncover patterns such as the influence of job roles, performance scores, and compensation on employee retention. By finding employees at high risk of turnover, the company can take preemptive measures to retain them, ultimately lowering hiring costs, improving operational efficiency, and enhancing employee engagement.

Why This Question is Important:

In the fast-paced, high-turnover environment of last-mile delivery logistics, predicting employee

retention is critical. High turnover in this industry leads to frequent recruitment cycles, increased costs, reduced productivity, and lower operational efficiency. The ability to identify employees who are likely to leave before they actually do will help streamline hiring processes, improve employee retention, and reduce costs.

This question is crucial not only to the company's current staffing challenges but also to future consulting opportunities. By successfully predicting turnover and improving employee retention, the company can enhance workforce stability and operational efficiency. Additionally, the insights gained from this analysis can be commercialized and offered as a consulting service to external clients, helping them manage their workforce challenges. Without addressing this question, the company risks continued high turnover, diminished productivity, and poor customer service due to staffing shortages.

Conclusion:

By using the tools provided by Python, such as the Kaplan-Meier estimator, Cox Proportional Hazards model, and Random Survival Forest, this project aims to develop a comprehensive, data-driven approach to predicting employee retention. Through the integration of Explainable AI techniques, the findings will not only predict turnover but also offer actionable insights into the factors influencing retention. Addressing turnover proactively can result in a more engaged workforce, lower recruitment costs, and higher operational efficiency for the company and its clients.