**Problem Description**

**Project Overview**

Employee turnover, also known as "employee churn," poses a considerable financial challenge for businesses. The actual cost associated with replacing an employee can be substantial. Particularly for IT service organizations, the rate at which employees leave, referred to as the churn rate, often falls within the range of 12% to 15%. This rate is notably high. Even if we conservatively assume a lower churn rate of 5%, the financial ramifications of an employee departing from the organization are significant. It can be approximated that the cost incurred when an employee leaves a company equals approximately 1.5 times their annual salary (Saradhi and Palshikar 2011).

This financial burden is undeniably unwelcome news for organizations contending with a high employee churn rate, also referred to as attrition. In essence, the cost of replacing employees remains a substantial burden for most employers due to factors such as the time and resources invested in recruiting and selecting replacements, sign-on bonuses, and the subsequent productivity loss during the transition period as the new employee adjusts to their role.

**Problem Statement**

Gaining insight into the reasons behind employee departures and identifying when they are most likely to occur can inform strategies to enhance employee retention and facilitate proactive hiring planning. I will employ a systematic, step-by-step approach that can be applied to a variety of machine-learning problems. This project aligns with the field commonly known as "HR Analytics" or "People Analytics."

**In this research, we aim to address the following problem statement:**

* How can we predict if an active employee leaving the company?
* What are the primary indicators that an employee is on the verge of leaving the company?
* What policies or strategies can be implemented based on the findings to enhance employee retention?
* Since we have data on former employees, this project entails a standard supervised classification problem in which the label is a binary variable: 0 (active employee) and 1 (former employee). In this study, our target variable, denoted as Y, represents the likelihood of an employee leaving the company.