**CHAPTER 3**

**PROBLEM STATEMENT**

**3.1. EXISTING SYSTEM**

Traditionally, companies have relied on subjective assessments or manual analysis to gauge their work culture. However, by implementing a random forest algorithm, an automated and data-driven approach can be adopted. This involves collecting data on various aspects such as employee satisfaction, turnover rates, diversity statistics, leadership styles, and company policies. These data points are then used to train the random forest model, which learns to identify patterns and relationships between different variables and work culture. Once trained, the model can provide recommendations for improving work culture based on the analysis of input data. This approach not only reduces bias and human error but also allows for continuous refinement and adaptation as new data becomes available. Ultimately, leveraging machine learning algorithms like random forest can lead to more informed decisions and foster a healthier and more productive work environment.

**3.2. DISADVANTAGES**

* Interpretability Challenges: Random Forest models, while powerful, are often considered "black box" models, meaning it can be challenging to interpret how individual variables contribute to the model's predictions.
* Overfitting Risks: Overfitting occurs when the model learns to capture noise in the data rather than the underlying patterns, leading to decreased generalization performance and potentially inaccurate insights about work culture dynamics.
* Data Dependency: Random Forest models require large amounts of high-quality data for training to achieve optimal performance. However, obtaining and maintaining such datasets, especially sensitive employee-related data, may pose privacy concerns and logistical challenges.