1. To achieve this I must understand the problem and bring solution. Decide which domain is preferable according to the input type then to fix whether the process comes under supervised or unsupervised or semi supervised learning.
2. Stage-1 (Domain Selection)

As I go through the problem , the company might have the performance, engagement in work, feedback sessions, timely arrival to work etc of an employee in the form of score so our dataset is numbers and our domain selection is Machine learning

Stage-2 (Type of Learning)

This scenario has a clear requirement and input that is the score of an employee to predict his/her decision and clear output that is whether the employee would continue or drop out so it comes under Supervised learning.

Stage-3 (Classification/Regression)

The prediction is categorized that is whether the employee would continue or not so comes under classification.

1. Project Name : Employee Drop out Prediction
2. Dataset

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Name of employee | Performance  Out of 10 | Engagement at work  Out of 10 | Health condition  Out of 10 | Overall score  Average | Result  Staying/Dropping out |
| Remi | 9 | 8.5 | 7.5 | 8.3 | staying |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |