

Name of the project : **Employee attrition prediction**

Stage 1:

Input / output

- 1) LinkedIn profile -open to work
- 2) Performance appraisal from manager – under performer/ over performer
- 3) Employee satisfaction survey – how happy they are with their work

Input is Text in all the above cases

So stage 1 domain selection is NLP – **Natural language Processing**

The output is Boolean – Employee may leave / may not leave

Under NLP, it falls under **Supervised Learning** because the Requirement is clear and it has clear set of input and respective output

LinkedIn -Open to work	May resign soon
Performance appraisal – under performer	May resign
Performance appraisal – over performer	May resign if the hike is low
Employee satisfaction survey	May resign
Performance appraisal – promoted with good hike	Would not resign

Under supervised learning , it will come under **classification** as the output is categorical .

Dummy dataset to perform the above action:

Employee id	employee name	Department	Performance rating	Satisfaction score
14213	Ram	Engineer	4	9
15676	Priya	Solution	3	3