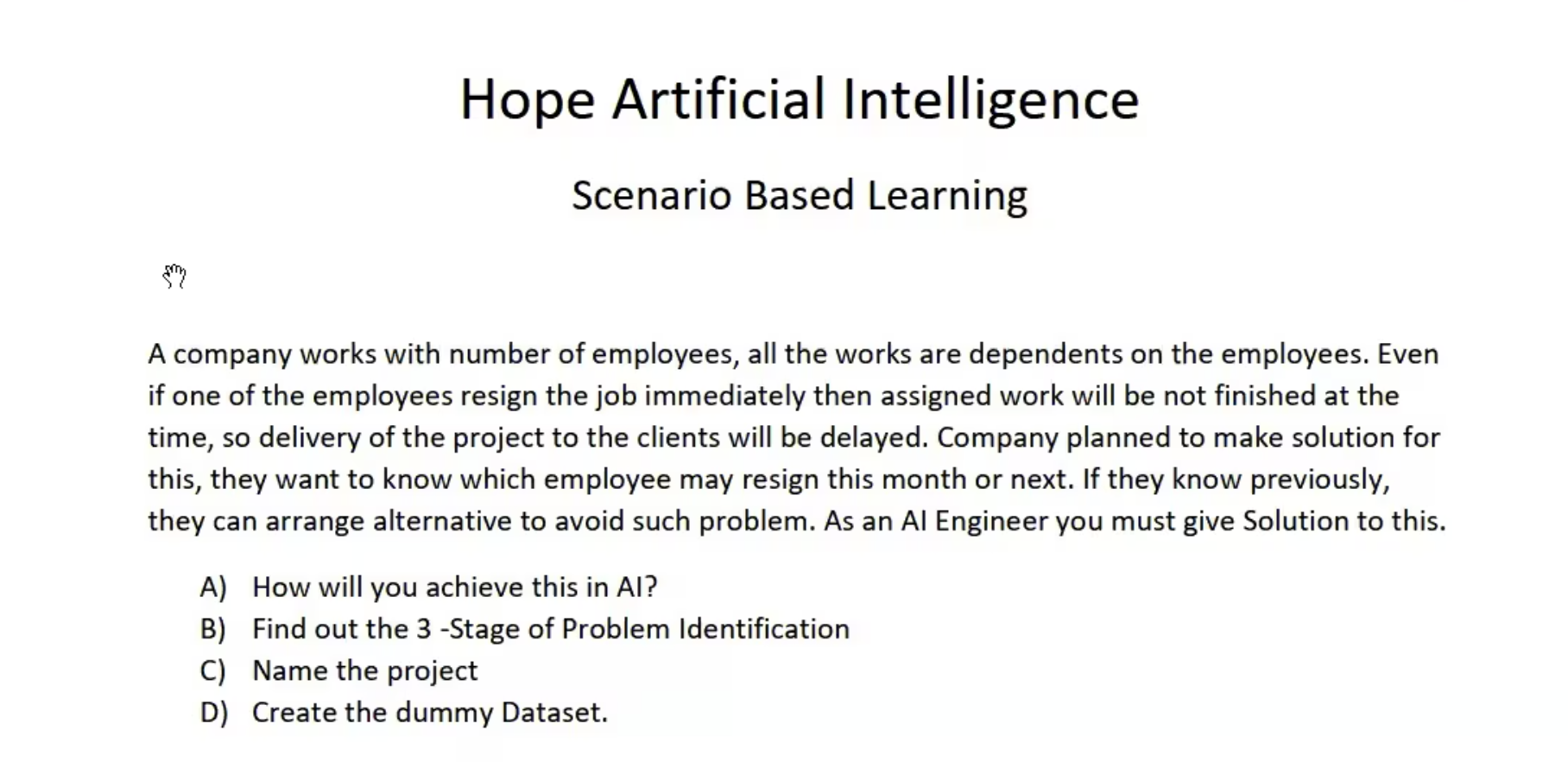
WEEK 3

Problem :



1. HOW will you achieve this in AI

Machine learning with supervised learning classification

1. Find out 3-stage of problem identification

Stage 1 : input, domain selection

* Input will be dataset
* Domain – Machine learning ( I feel like it can be combined machine learning and NLP based on input data)

Stage 2:

* Supervised selection

Stage 3:

* Classification

1. Name the Project

* Project Health Analyzer (PHA)
* Employee Risk Predictor (ERP)
* Project Delivery Risk Finder (PDRF)

1. Create Dummy Dataset

What is the dataset?

1. If the company has clear dataset like notice period , then it is very straight approach.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Employee name | Resigned | Resignation accepted | Current project | Is he in notice period | How many days left fro LWD | Output |
| Emp1 | yes | yes | Project name | yes | 30 | quit |
| Emp2 | no | NA | Project name | no | NA | Not quit |

1. If the company has to collect dataset from survey or process like intimation before leave and notice period.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Employee name | Num of task assigned in a week | Num of task completed in a week | Working / online availability in a week | Employee is able to complete the task on time | Employee is able to understand the requirement | Feedback about company from employee (rating) | Feedback about manger from employee (rating) | Is he/she willing to continue  in the same project | He needs change in technology career |
| Emp1 | 10 | 4 | 20 hrs | 4 | 4 | 4 | 4 | no | yes |
| EMP2 | 8 | 6 | 30 hrs | 8 | 8 | 7 | 6 | yes | no |

1. If input is not clear, we need to collect the data and audit data based on Browsing history, official profile activity data from third-party vendor.

Collect dataset based on browsing data using NLP supervised learning.  
if it matches job search keywords – then it is job searching

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Employee name | Is he active in job portal | Is he attending interviews via job portal | Browsing data matches job searching(outout from NLP) |  | Employee is able to complete the task on time | Employee is able to understand the requirement | Feedback about company from employee (rating) | Feedback about manger from employee (rating) | Is he/she willing to continue  in the same project | He needs change in technology career |
| Emp1 | 8 | 8 | yes |  | 4 | 4 | 4 | 4 | no | yes |
| EMP2 | 2 | 0 | no |  | 8 | 8 | 7 | 6 | yes | no |

My Queries : (I am in Week 3 )

1. Can we combine two/multiple domains to solve a problem ?
2. Can we use learning to create input dataset itself?

* For example NLP -> output -> input -> ML -> supervised learning

1. It looks like we need to preprocess the data with some automation before feed to learning. Is that correct?