# Online Course Recommendation System Dataset V2

## Problem Statement

The goal of this dataset is to build an online course recommendation system that suggests relevant courses to learners based on their interests, past enrollments, and engagement levels. The dataset includes course ratings, instructor information, previous learning history, study material availability, and certification offerings, making it suitable for recommendation models using collaborative filtering, content-based filtering, or hybrid approaches.

## Variable Descriptions

|  |  |  |
| --- | --- | --- |
| Variable Name | Data Type | Description |
| user\_id | Integer | Unique identifier for each learner. |
| course\_id | Integer | Unique identifier for each online course. |
| course\_name | String | The name of the online course. |
| instructor | String | The name of the instructor teaching the course. |
| course\_duration\_hours | Float (5.0 - 100.0) | The duration of the course in hours. |
| certification\_offered | String (Yes/No) | Indicates whether the course provides a certification upon completion. |
| difficulty\_level | String | The difficulty level of the course (Beginner, Intermediate, Advanced). |
| rating | Float (1.0 - 5.0) | User-provided rating for the course. |
| enrollment\_numbers | Integer | The total number of students enrolled in the course. |
| course\_price | Float (20.0 - 500.0) | The price of the online course. |
| feedback\_score | Float (0.0 - 1.0) | A normalized score representing the feedback sentiment from students. |
| study\_material\_available | String (Yes/No) | Indicates whether additional study materials are available. |
| time\_spent\_hours | Float (1.0 - 100.0) | The average time spent by students in the course (in hours). |
| previous\_courses\_taken | Integer | The number of previous courses the learner has taken before enrolling in this one. |