

Learner profile

- Learner ID
- Educational background (degree, year)
- Existing certificates
- Entry skill level (based on placement test
→ accurate (?))
- Target skill level (mapped to SFIA)

Focus: distribution of learner backgrounds vs target skills, skill level imbalance

Learning activity + interaction

- Session duration
- Timestamped (start, complete)
- Learning speed (content length vs completion time)
- Revisit content frequency
- Amount of time spent per module
- Clickable action
- Skipped content
- Exploration of alternative modules

Focus: time spent vs completion rate, time pattern long sessions → lower success or not, fast vs slow learners across content difficulty level, engagement vs mastery correlation, passive vs active learner, burned out signals

Assessment + performance

Assessment specific:

- Assessment ID
- Skill mapped (SFIA code + level + tag)
- Attempt number
- Score (raw & normalized) + quartiles percentage
- Time taken

Question specific:

- Question ID
- Concept tag
- Difficulty level
- True/false

- Type of mistakes (conceptual/careless/difficulty)

Focus: score distributions per skill and level, learning curves (attempt vs accuracy), error patterns per concept, and mastery insight

Content

- Module ID
- Skill mapping (SFIA)
- Difficulty
- Estimated completion time
- Pre-requisite module/skill level
- Learning objectives
- Content format
- Skill mastery achieved
- Time to mastery
- Flagged

Focus: skill coverage gaps, content difficulty vs learner performance, and format effectiveness

Temporal + longitudinal

- Timeline per learner
- Skill mastery over time
- Skill regression
- Gaps between learning session

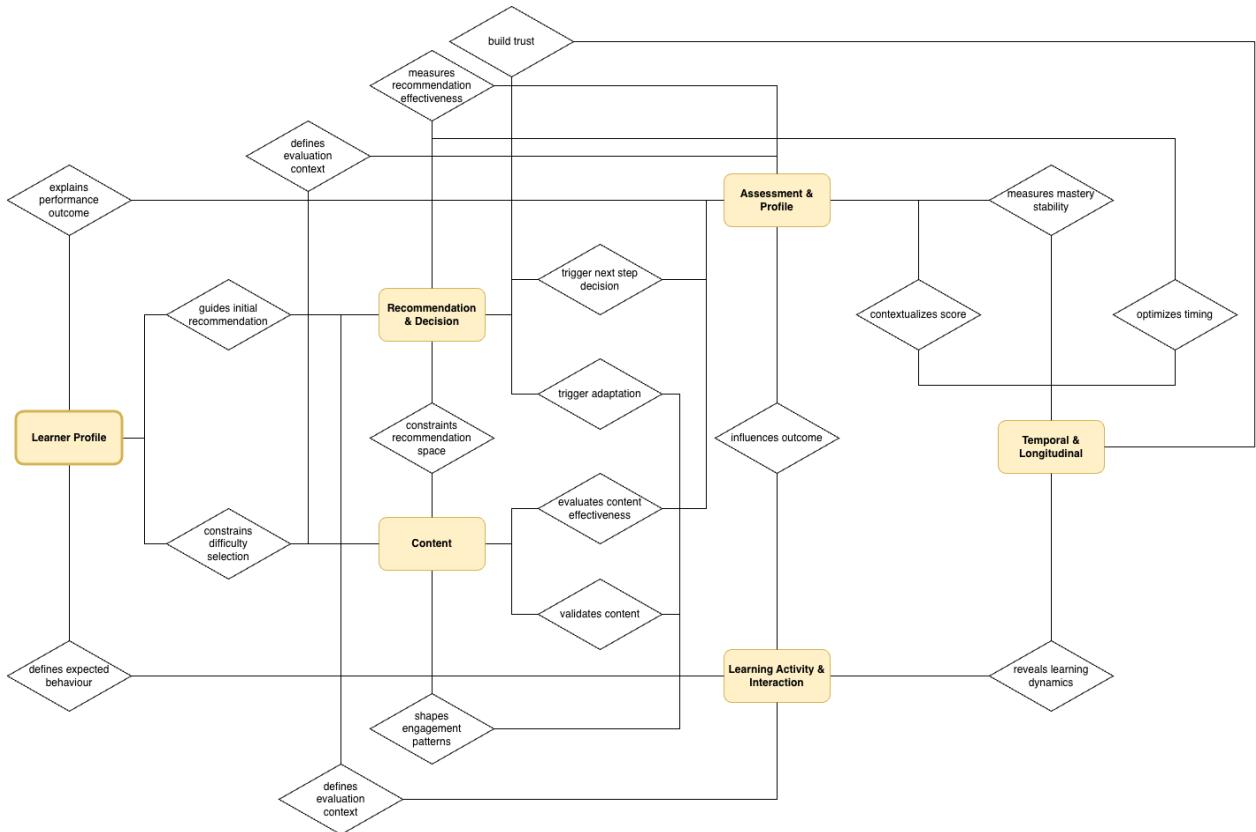
Focus: forgetting pattern, optimal review intervals, mastery stability, long term vs short term learners

Recommendation + decision

- Recommendation ID
- Recommended module list
- Chosen module
- Ignored recommendations
- Alternative path explored

Focus: acceptance rate per recommendation type, exploration patterns, recommendation diversity, feedback from user (bad recs → ignored)

Relations idea:



Feature mapping (based on algorithm):

- Rule-Based → learning structure

Purpose:

- Ensure learners follow the correct learning order
- Prevent skipping required prerequisites
- Provide reasonable recommendations when data is still limited

Related input:

- Entry skill level (if exists)
- Target skill level
- Latest assessment score
- Number of attempts
- Module difficulty
- Prerequisite rules
- Time since last assessment

Output:

- List of modules the learner can take next
- Decision to review/continue/move forward

- Bayesian Knowledge Tracking (BKT) → track understanding

Purpose:

- Track what concepts a learner understands over time
- Estimate learning progress and forgetting
- Provide an interpretable view of learner mastery

Related input:

- Learner ID
- Concept tag
- Correct or incorrect answer
- Attempt order
- Timestamp
- Time gap between attempts

Output:

- Mastery probability for each concept
- Mastery changes across learning sessions

- Multilayer Perceptron (MLP) → predict success

Purpose:

- Combine performance, behaviour, and time signals
- Predict whether a learner is likely to succeed or struggle
- Capture complex patterns (if exists)

Related input:

- Entry skill level (if exists)
- Educational background (encoded)
- Average time spent per module
- Completion rate
- Revisit frequency
- Skip ratio
- Recent assessment scores
- Mistake type distribution
- Content difficulty
- Content format
- Session gaps and learning streaks

Output:

- Predicted mastery level
- Risk of failure or disengagement

- K-Means Clustering → reveals behaviour patterns

Purpose:

- Group learners with similar learning behaviours

- Identify common learning patterns across the platform
- Support cohort-level personalization

Related input:

- Average session duration
- Modules completed per week
- Completion ratio
- Revisit frequency
- Average assessment score
- Score improvement rate
- Session regularity

Output:

- Learner behaviour cluster label
- Description of learner group characteristics

- DBSCAN → detect unusual cases

Purpose:

- Identify learners with unusual or irregular learning behaviour
- Detect extreme cases that do not fit common learner groups
- Highlight learners who may need special handling

Related input:

- Extremely long or short session durations (if exists)
- High skip ratios
- Large score fluctuations
- Irregular learning gaps
- Sudden drops/increases in performance

Output:

- Outlier or noise label
- Flag for irregular learning behaviour

- Reinforcement Learning → continuously improves recommendations

Purpose:

- Improve learning recommendations over time
- Learn which learning paths lead to better outcomes
- Balance progress, engagement, and exploration

Related input:

- Current mastery level (from BKT)
- Predicted mastery scores (from MLP)
- Engagement indicators
- Learning speed
- Learner cluster (from K-Means)

- Outlier flag (from DBSCAN)
- Time since last learning activity
- Recommendation accepted or not
- Module completion
- Score improvement
- Skill mastery

Output:

- Ranked list of recommended modules
- Alternative learning paths for exploration