Research Plan

Reinforcement Learning for Early Educational Intervention: A Cost-Aware and Fairness-Aware Approach

Student's Name: Hai Nguyen

In my research plan, I will clearly outline the objectives, scope, research methodology, and schedule of different phases for a thesis topic focusing on the Early Educational Intervention using Reinforcement Learning.

I. Topic & Preliminary Title:

- **Preliminary Title**: *Reinforcement Learning for Early Educational Intervention: A Cost-Aware and Fairness-Aware Approach*

- **Scope of the study:** This thesis focuses on applying Reinforcement Learning (RL) to optimize early interventions for students at risk of dropout or poor performance. Unlike static early-warning systems, RL enables adaptive, sequential decision-making under resource and fairness constraints. The study will emphasize:
    - RL formulation for intervention planning.
    - Comparative analysis with heuristic and contextual bandit baselines.
    - Evaluation of effectiveness, efficiency, and fairness.
    - Ethical considerations in algorithmic decision-making.

II. Potential Commissioner and Field:

My thesis will be conducted independently in the field of Educational Technology and AI for Learning Analytics. This will ensure flexibility in research design and methodology, allowing deeper exploration of Reinforcement Learning applications in education.

III. Other actors:

To carry out my thesis successfully, I need to carefully consult:

- TUAS supervisors and lecturers for academic guidance.
- Educational data scientists and EdTech professionals for technical insights.
- Teachers and student-support staff for practical constraints and fairness perspectives.

IV. Reasons for choosing the topics:

There are two reasons leading to my decision of selecting this topic:

1) Having completed my bachelor’s and master’s degree programs in the field of education and intercultural communication (2018 and 2021), and currently studying Data Engineering & AI at TUAS, I consider this topic a meaningful combination of my previous and current studies.

2) Currently, most Early Warning Systems in education only stop at predicting the risk of students dropping out or declining in performance, but do not provide specific actions to improve the situation.

   Educational institutions are facing a major challenge: how to not only identify risks but also provide intervention recommendations that are appropriate, timely, cost-effective, and ensure fairness for all student groups.

   This requires an adaptive system capable of:

   - Optimizing sequences of intervention actions instead of just performing a single step.
   - Balancing learning effectiveness and resource costs, avoiding the overuse of costly measures.

- Monitoring and minimizing fairness disparities among student groups (e.g., by ability, economic conditions, region).

This thesis aims to address that urgent need by applying Reinforcement Learning (RL) to build a dynamic decision-making framework, helping schools and teachers make data-driven decisions and optimize long-term outcomes for learners.

V. Objective of the thesis:

- Formulate early intervention as a Markov Decision Process (MDP)/Partially Observable MDP (POMDP) with cost and fairness constraints.
- Implement Reinforcement Learning algorithms (Q-learning) and compare with heuristic and contextual bandit baselines.
- Evaluate policies in simulation and offline datasets using metrics for retention, performance, cost, and fairness.
- Provide practical guidelines and dashboards for responsible deployment.

VI. Methods and Materials:

a) Simulation Experiments: Design a student-behaviour simulator to test RL policies safely and perform ablations (reward weights, action sets, fairness penalties).

b) Offline Evaluation: Replay policies on historical learning analytics datasets (e.g., OULAD) to validate performance under real-world conditions.

c) Literature Review: Synthesize theoretical foundations from Reinforcement Learning, bandits, Partially Observable Markov Decision Process, Educational Data Mining, and fairness frameworks.

d) Expert Consultation: Interviews or short surveys with educators and EdTech experts to validate assumptions on cost models and fairness constraints.

VII. Components:

a) Literature Review:

I will synthesize theoretical and empirical foundations from Reinforcement Learning (RL), Contextual Bandits, and Educational Data Mining (EDM). The review will critically analyse existing Early Warning Systems to identify the gap between predicting risk and prescribing interventions, while also surveying fairness metrics (e.g., demographic parity) to ensure ethical algorithmic decision-making.

b) Methodology & Framework Formulation:

I will formulate the educational intervention problem mathematically as a Markov Decision Process (MDP) or POMDP. This involves defining:

- State Space (S): Student features (performance, attendance, behaviour).
- Action Space (A): Pedagogical interventions (e.g., email reminders, tutoring, counselling).
- Reward Function (R): A composite function balancing student success (retention/grades) against intervention costs and fairness constraints.

c) Implementation & Evaluation (Simulation):

I will develop a simulation environment (Virtual Classroom) to train and test the RL agent (Q-learning) since experimenting on real students is unethical. The evaluation will use Off-policy evaluation on historical datasets or simulated data. I will measure performance using quantitative metrics:

- Effectiveness: Improvement in student retention/grades.
- Efficiency: Cost reduction compared to heuristic baselines.
- Fairness: Analysing disparate impacts across different student groups.

VIII. Ethical Considerations:

Ethical integrity is focused through three key pillars:

1) Data Privacy is strictly enforced via GDPR compliance and data anonymization.
2) To mitigate Algorithmic Bias, the proposed "Fairness-Aware" framework ensures intervention recommendations are equitable across all demographics, preventing discrimination based on historical data patterns.
3) Human-in-the-loop and Beneficence: While the system is "Cost-Aware," student welfare prioritizes financial optimization. The AI agent acts solely as a decision-support tool, not a replacement for human judgment. Final intervention decisions remain with educators to prevent automation bias and ensure that recommendations are pedagogically sound and compassionate.

IX. Related Sources:

As this is literature-based and research, there will be a wide range of previous research works used as references. In this research plan, I list the first 5 materials that I am reading now to get the main ideas of the topic:

- Sutton & Barto (2018). *Reinforcement Learning: An Introduction.*
- Plaat (2022). *Deep Reinforcement Learning.*
- Lattimore & Szepesvári (2020). *Bandit Algorithms.*
- Romero et al. (2010). *Handbook of Educational Data Mining.*
- Barocas et al. (2023). *Fairness and Machine Learning.*

X. Schedule Plan:

This is the table of my schedule plan, from which I pre-plan the different phases to go through to complete my research process.

I hope to receive teachers' feedback and evaluation of the plan's possibility before I can officially start implementing it.

Time	Main activities
December 2025	• Finalize research proposal and thesis plan. • Conduct preliminary literature survey on RL and educational interventions.
January 2026	• Perform comprehensive literature review. • Define research methodology and formalize the MDP/POMDP framework. • Design simulation environment specifications.
February 2026	• Develop the simulation environment. • Preprocess offline datasets. • Implement baseline algorithms (Heuristics, Contextual Bandits) for comparison.
March 2026	• Implement the Reinforcement Learning agent(Q-learning). • Design and integrate reward functions with cost and fairness constraints.
April 2026	• Conduct extensive experiments and ablation studies. • Perform offline evaluation and analysis. • Generate data visualizations (Pareto plots, performance tables).

May 2026	• Draft core chapters: Methodology, Experiments, and Results. • Submit partial drafts for supervisor feedback and incorporate revisions.
June 2026	• Finalize Introduction and Conclusion chapters. • Compile references and appendices. • Complete the full thesis draft for final review.
July 2026	• Proofread and format according to Turku UAS guidelines. • Submit final thesis. • Prepare presentation slides for the thesis defence.