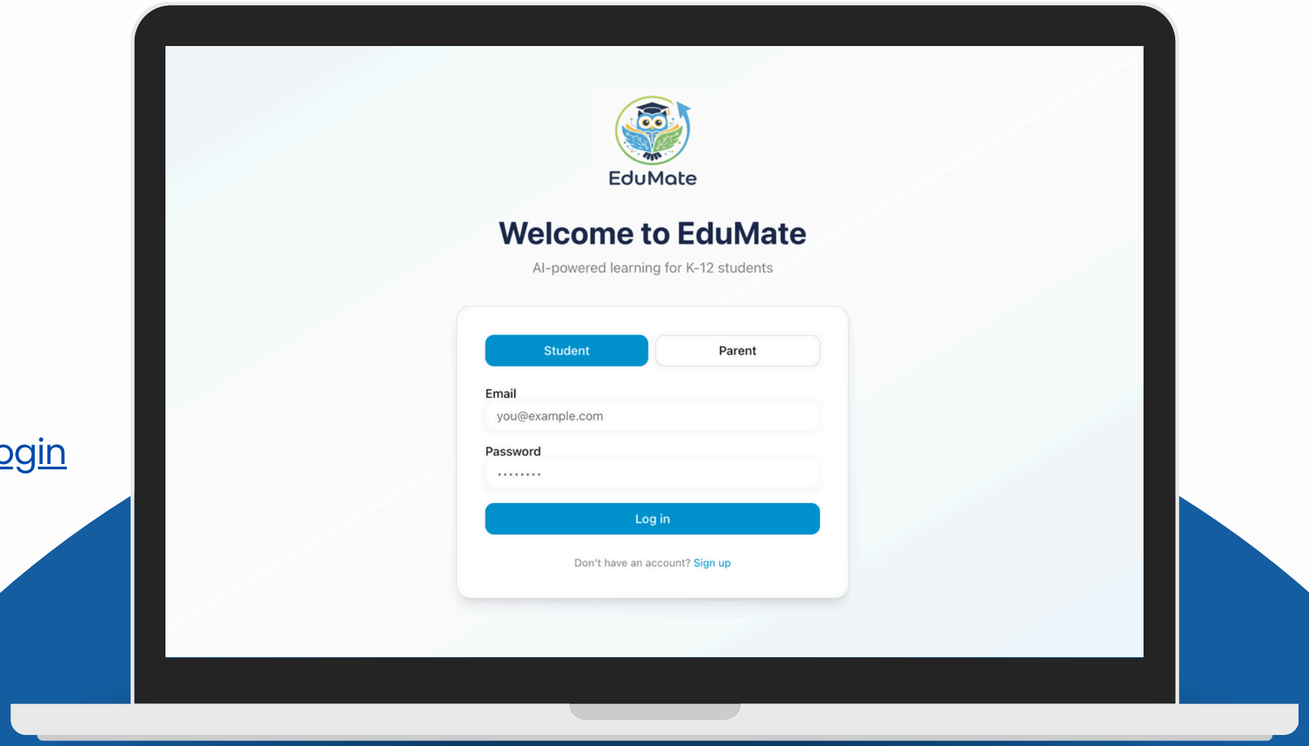
EduMate

EduMate

By: Aditya Kansara

Developed URL: https://project-check-point-1-edumate.vercel.app/login

Github URL: https://github.com/SALT-Lab-Human-AI/project-check-point-1-edumate.git



Recap & Final Goals

Recap from Checkpoint #3:

- Built RAG-powered tutoring system with three learning modules (S1, S2, S3)
- Implemented parent dashboard with progress tracking
- Deployed on Vercel (frontend) and Render (backend)

Target Users:

- K-12 students (grades 1-12) seeking math help
- Parents/guardians monitoring progress and controlling learning environment
- After-school staff/tutors running structured practice

Final Problem Statement:

Students need instant, reliable, grade-level help outside class. Generic chatbots hallucinate and aren't aligned to standards; parents rarely see which skills improved and where their child struggles.

Most Rewarding User Feedback:

- **100%** found system "Useful" or "Very useful" for overall ease of use (N=14)
- **100%** rated Student Dashboard positively
- **100%** found S1 guidance "Useful" or "Very useful"
- **86%** found S2 feedback "Effective" or "Very effective"

Improvements Based on Feedback:

- **Optimized website load speed**
  (addressed primary user concern)
- **Fixed S3 quiz generation bugs**
  (improved quiz answer visibility and functionality)
- **Memory optimization:**
  System now operates efficiently within 500MB RAM constraints

Evaluation Design Overview

Participants: 14 users (students and parents) for usability survey

Study Tasks:

- User Testing: Complete tasks in S1, S2, S3 modules
- User Survey: System usability across 9 dimensions
- System Performance Testing: API response times, memory usage, error rates

Metrics Collected

Quantitative:

- **Usability ratings:**
- "Useful" to "Very useful" scale across 9 dimensions
- **Task completion success** (comparative rubric scoring: 0-5 on 5 dimensions)
- **Response times** (tutoring: 1-2s, quiz generation: 2-3s, grading: <100ms)
- **Error rates** (<2% overall)
- **Vector search latency** (15-50ms)
- **Memory usage** (optimized to 500MB)

Qualitative:

- **User satisfaction themes** (positive feedback, areas for improvement)
- **Response quality analysis** (mathematical accuracy, grade appropriateness, educational quality)
- Safety filtering effectiveness

Study Materials & Protocol

Task:

- Create an Account and Login
- Complete practice problems in S1 (Structured Problem-Solving)
- Submit solutions for feedback in S2 (AI-Powered Solution Feedback)
- Generate and complete quizzes in S3 (Mathematical Quiz Generation)
- Navigate parent dashboard and review progress

Study Procedure:

- **Intro:** System overview and feature walkthrough
- **Task Completion:**
  - **Comparative evaluation:** Test prompts to baseline tools
  - **User testing:** Complete tasks in S1, S2, S3 modules
- **Survey:** 9-question usability survey (ease of use, module effectiveness, dashboard clarity)
- **Interview:** qualitative feedback collection

Interface participants interacted with:

- Student Dashboard with S1, S2, S3 modules
- Parent Dashboard with progress analytics and controls
- Grade selector (1-12), topic dropdowns, LaTeX math rendering

Quantitative Results

Key Highlights:

- **Overall Ease of Use: 100% positive**
- **Student Dashboard: 100% positive**
- **S1 & S2 Modules:**

  100% positive for **S1**; **86% positive** for **S2**
- **S3 Quiz Matching: 64% positive**
- **Math Expression Clarity: 93% positive**

Metric	Very Useful/Effective	Useful/Effective	Neutral	Not Very Well
Overall Ease of Use	9 (64%)	5 (36%)	0	0
Student Dashboard Clarity	11 (79%)	3 (21%)	0	0
S1 Guidance Usefulness	7 (50%)	7 (50%)	0	0
S2 Feedback Effectiveness	3 (21%)	9 (64%)	2 (14%)	0
S3 Quiz Grade Matching	4 (29%)	5 (36%)	3 (21%)	2 (14%)
Progress Tracking Helpfulness	6 (43%)	5 (36%)	3 (21%)	0
Parent Dashboard Usefulness	5 (36%)	9 (64%)	0	0
Parent Controls Ease	6 (43%)	8 (57%)	0	0
Math Expression Clarity	6 (43%)	7 (50%)	1 (7%)	0

System Performance:

- **API Response Times:** - Tutoring **1-2s,**
  - Quiz Generation **2-3s**,
  - Grading **<100ms**
- **Vector Search: 15-50ms** average
- **Error Rates: <2%** (RAG failures <1%, JSON parsing <2%)
- Memory Optimization: System optimized to operate efficiently within **500MB RAM** constraints

Post-Survey Improvements:

- Load speed optimization implemented
- S3 quiz generation bugs fixed
- Memory optimization completed

Qualitative Insights

Top 3 Positive Themes:

1) Clear Interface & Visual Design

- Color scheme and visual design appreciated
- Navigation menu layout praised
- Font size and readability appreciated

2) Useful Structured Guidance

- S1 step-by-step approach found effective **(100% positive)**
- S2 feedback provides precise error identification **(86% positive)**
- LaTeX math rendering clear and professional **(93% positive)**

3) Reduced Cognitive Load

- Module organization makes learning pathways clear **(100% positive ratings for Student Dashboard)**
- Parent controls easy to use **(100% "Easy")**
- Progress tracking **(79% positive)**

"Parent dashboard gives great visibility into my child's progress"

Top 2 Frustrations:

1) Load Speed Issues

- "Load speed of modules" mentioned as primary improvement area
- **Action Taken:** Implemented aggressive memory management and frontend/backend optimizations
- **Result:** Significantly improved module load speeds

2) Quiz Functionality Gaps

- Quiz answers not always visible after submission
- Some quiz functionality bugs affecting reliability
- **Action Taken:** Fixed S3 quiz generation bugs, improved answer visibility, resolved functionality issues
- **Result:** Quizzes are generated seamlessly without any issues.

"Would like faster loading times for the modules"

Interpretation & Discussion

Participants helped us find critical bugs and performance issues

1)**Load Speed Issues:** Multiple users reported slow module loading as the primary concern. This led to:

- Aggressive memory management optimizations (lazy loading, model unloading)
- Frontend/backend performance improvements
- Memory optimization to operate within 500MB constraints
- **Result:** Noticeable improvement in load times post-fix

2) **S3 Quiz Generation Bugs:** Users identified specific functionality problems:

- Quiz answers not displaying after submission (visibility bug)
- Quiz generation failures in certain scenarios
- Answer selection and grading inconsistencies
- **Result:** Fixed answer visibility, improved quiz generation reliability, resolved grading issues

Tie to Success Criteria:

- **Usability:** 100% found system "Useful" or "Very useful"
- **Module Effectiveness:** 100% positive for S1, 86% positive for S2
- **Parent Trust:** All parent controls rated "Easy" or "Very easy"
- **Responsive to Feedback:** Load speed and quiz bugs addressed post-survey

Limitations, Risks & Ethics

Limitations:

1) Small Sample Size

- User survey: N=14 (improved from N=5, but still limited generalizability)
- Need larger-scale studies for statistical significance

2) LaTeX Formatting Edge Cases

- The format_latex() post-processing function handles most LaTeX normalization cases
- May miss edge cases in deeply nested expressions or non-standard LaTeX syntax

Risks

1) Security and Privacy Considerations

- **Password Security:** Current implementation uses SHA-256 hashing, which is less secure than bcrypt. Future work should upgrade to bcrypt for improved password security.
- **Rate Limiting:** No rate limiting currently implemented. Future work should add rate limiting to prevent abuse and ensure fair resource usage.
- **Content Filtering:** While system prompts include safety instructions, more sophisticated content filtering would enhance safety for younger users.

2) Model Unpredictability

- LLM responses can vary despite RAG grounding
- LaTeX formatting edge cases may require manual review

Ethics

- Parent-controlled academic integrity features
- Safety filtering in system prompts (refuses cheating requests)
- Transparent progress reporting for parents

Conclusion & Future Work

Main Takeaways:

1) Structured Learning Pathways Are Highly Effective

- 100% user satisfaction with S1 structured guidance (N=14)
- Parent dashboard enables human-AI collaboration

2) User-Centered Design Achieves Strong Usability

- 100% positive ratings for overall ease of use and Student Dashboard
- User feedback directly informed improvements

3) Memory-Optimized RAG Enables Cost-Effective Deployment

- System operates efficiently within 500MB constraints

Future Work:

1) Technical:

- Interactive Socratic dialogue (multi-turn conversations)
- Adaptive difficulty based on performance history
- Multimodal input (image/handwriting recognition for S2)
- Security: Upgrade to bcrypt, add rate limiting

2) Research:

- Long-term learning outcomes impact of RAG grounding
- Optimal retrieval strategies for educational content
- Grade-adaptive prompting effectiveness across LLM architectures

3) UX:

- Enhanced progress analytics
- Gamification features

Contribution to Human-AI Collaboration

- RAG enhances AI tutoring reliability while maintaining pedagogical effectiveness
- Memory-optimized strategies enable cost-effective deployment
- Structured modules + parent controls balance learning support with academic integrity

Acknowledgments & Contributions

AI Development Tools:

- **Cursor:** Code completion and suggestions
- **Vercel:** Frontend deployment and hosting
- **ChatGPT (GPT-4):** Architecture discussions, debugging, documentation review

Deployment & Infrastructure:

- **Render:** Backend deployment and hosting
- **Supabase:** PostgreSQL database with pgvector extension

Acknowledgments:

- Survey participants for valuable feedback
- Open-source community for libraries and frameworks
- Groq for API access
- Vercel, Render, and Supabase for deployment and infrastructure services

AI Services:

- **Groq API:** LLM inference for tutoring and quiz generation
- **SentenceTransformers:** Embedding generation for RAG retrieval

Disclosures:

- All AI-generated code was reviewed and validated
- Prompts were iteratively refined based on evaluation results
- Core algorithms, architecture, and implementation decisions made independently
- AI tools used as development aids, not replacements for understanding

Individual project — all work completed independently

THANK YOU!