

Problem Statement:

Build a multi-agentic AI system using image processing, computer vision & AI/LLM to auto-align and evaluate handwritten answer sheets

Challenges

- Slow, repetitive grading — teachers spend hours per exam.
- Inconsistent scoring — subjectivity across multiple graders.
- Poor feedback — students get marks, not actionable guidance.
- Rubric drift — teachers apply rubrics inconsistently.
- Scalability pain — exams across many classes/boards are hard to scale.
- Low trust in automation — teachers don't trust AI that can't explain marks.
- Connectivity & access — rural schools lack reliable internet for cloud tools.

What we deliver

An explainable, multi-agent AI pipeline that auto-aligns & semantically evaluates handwritten answer sheets.

Core pieces:

- Transformer OCR → reads handwriting + preserves layout.
- RAG (Qdrant + Sentence-Transformers) → grounds grading in teacher's model answers.
- Llama 2 (LLM) → assigns marks, confidence, and human-readable explanations.

- Whisper → transcribes oral answers for multimodal exams.
- Human-in-the-loop dashboard (React + FastAPI) → quick teacher review & learning loop.

Tech Stach

Component	Short use-case / role
Transformer OCR	High-accuracy OCR for printed + handwritten answers (extract text + layout + bounding boxes for highlighting).
Sentence-Transformers (all-MiniLM-L6-v2)	Create semantic embeddings for model-answer chunks and student answers (fast, lightweight).
Qdrant (Vector DB)	Store embeddings + metadata (question_id, chunk_id, source offsets). Fast semantic retrieval for RAG.
Llama 2 (LLM)	Core grader: consume retrieved contexts + student text, produce score, confidence, explanation, and matched/missing points.
Whisper	Transcribe audio submissions (student verbal answers) into text for semantic grading with the LLM.
LangChain	Orchestrate RAG flows, prompt templates, retrieval chaining, few-shot examples, and agent-like pipelines.
FastAPI	Backend API: upload endpoints, OCR jobs, retrieval, LLM invocation, verification, and logging.
React + Tailwind	Teacher UI: upload PDFs, review/override grades, highlight evidence, dashboards and analytics.
PostgreSQL	Store metadata, audit logs, grading history, user accounts, and labeled corrections (training data).

Our Solution

- **AutoAlign & DeepGrade++:** Multi-agent AI system that automatically maps handwritten/typed/audio answers to the correct questions and evaluates them using semantic understanding.
- **Pipeline:** Transformer OCR → chunking → embeddings (Sentence-Transformers) → RAG retrieval (Qdrant) → Llama 2 grading → Whisper for voice → human-in-loop verification → dashboard analytics.
- **Goal:** Fast, accurate, and explainable grading for teachers & institutions at scale.

Uniqueness

- **Semantic evaluation, not keyword matching:** Understands context, synonyms, and phrasing rather than checking isolated words.
- **Multi-modal support:** Handles handwritten text, typed answers, diagrams, and math equations.
- **Viva support:** Uses **Whisper** to transcribe spoken answers and evaluates them semantically.
- **Explainable scoring:** Each mark linked to evidence (source IDs + highlighted text), building trust with teachers.
- **Automatic alignment:** Maps student answers to exact question IDs, even in messy layouts.

Innovation

- **RAG-grounded LLM grading:** Reduces hallucination, ensures AI marks are anchored in teacher model answers.
- **Human-in-the-loop learning:** Teacher corrections feed back into the system for continuous improvement.
- **Edge + cloud hybrid:** Supports offline grading for low-connectivity schools.

- **Confidence scoring & alerts:** Automatically flags uncertain answers for human review, improving reliability.

Fuseability (Integrations)

- **Existing LMS / ERP:** Can plug into school management software to automatically fetch exams & submit results.
- **APIs:** FastAPI endpoints allow integration with other grading or analytics tools.
- **Export options:** CSV, PDF, dashboard reports for administrators & teachers.
- **Scalable architecture:** Vector DB + LLM modularity allows adding new subjects, languages, or exam types seamlessly.

Reliability

- **Audit-ready:** Logs prompts, retrieved chunks, scores, confidence, and teacher overrides.
- **Deterministic scoring:** Low LLM temperature ensures consistent grading across students.
- **Error handling:** OCR confidence thresholds, verification pipeline, and fallback to human review for ambiguous answers.
- **Batch & async processing:** Handles large classes & multiple subjects efficiently.

Business Model

- **B2B SaaS for schools & universities**
 - Subscription-based: per-student or per-exam.
 - Tiered plans: basic auto-grading → full multi-modal + analytics.

- **Enterprise licensing:** Custom deployments for boards or state-level exam authorities.
- **Add-on services:**
 - Analytics & insights dashboards (student performance, skill gaps).
 - AI-assisted feedback & remedial content generation.
 - Multi-lingual grading support.
- **Upsell potential:** Extend to corporate assessments, online coaching platforms, or competitive exams.