GroupMate: A Role-Aware Team Formation System for Academic Collaboration

Padma Pooja Chandran*
University of Illinois Urbana-Champaign
Urbana, Illinois, USA
ppc2@illinois.edu

Aarthy Kesavan Padmanaban
University of Illinois Urbana-Champaign
Urbana, Illinois, USA
aarthyk2@illinois.edu

Krishnaveni Unnikrishnan
University of Illinois Urbana-Champaign
Urbana, Illinois, USA
ku18@illinois.edu

Jocelyn Xu
University of Illinois Urbana-Champaign
Urbana, Illinois, USA
yuex7@illinois.edu

Abstract

Team formation in educational environments often struggles to balance student interests, technical strengths, and project compatibility. We introduce GroupMate, an intelligent group recommendation system designed to automate and optimize team creation in academic settings. Leveraging a vision essay scoring model fine-tuned using arXiv abstracts and heuristics for novelty and originality, GroupMate classifies students into Visionaries, Collaborators, and Enablers. Our matching algorithm forms balanced teams by seeding groups with visionary students and incrementally filling roles based on skill complementarity and project fit. We evaluate our approach against a greedy algorithm and an LLM based approach, demonstrating GroupMate's unique ability to produce role balanced, executable, and conceptually aligned teams. GroupMate is implemented as a full-stack web platform with real-world usability, offering valuable insights into the potential of hybrid LLM- and model-driven approaches to educational tooling. GroupMate is available on GitHub[1].

CCS Concepts

• **Information systems → Recommender systems**.

Keywords

Group recommendation, Large language models, Transformer models, Team formation, Educational tools

1 Introduction

Group recommendation systems aim to address the complex problem of recommending items or activities to a group of users rather than individuals. Compared to individual recommendations, group recommendations must reconcile diverse preferences, interests, and expectations of all group members to reach a consensus that satisfies everyone involved. Several recent models have significantly advanced this field, including AGREE[4], HyperGroup[7], HCR[9], GroupIM[10], S2HHGR[13], CubeRec[5], and ConsRec[11]. These approaches leverage various methodologies, ranging from collaborative filtering and attention mechanisms to deep learning and hypergraph modeling, to better aggregate and align individual user preferences within a group context.

For example, AGREE (Attentive Group Recommendation) employs an attention-based model to capture dynamic user-group interactions, highlighting influential group members and providing more personalized recommendations. HyperGroup extends the concept by utilizing hypergraph structures to model the intricate relationships among group members, capturing both direct and indirect interactions effectively. HCR (Hierarchical Contextual Representation) and GroupIM further enhance recommendation performance by incorporating contextual information and interaction modeling to capture nuanced preferences within groups. Models like CubeRec and ConsRec adopt multi-dimensional and consensus-based approaches, respectively, to ensure recommendations accurately reflect the collective preferences and interests of group members.

Inspired by these innovative methodologies, we introduce *GroupMate*, an AI-powered, web-based group recommendation system specifically designed for student team formation in academic courses. Unlike conventional assignments or simple manual grouping, GroupMate employs a fine-tuned transformer model to score student vision essays based on the novelty, originality and frontier research scope of their project proposals. These scores from the vision model are then used to categorize students into three distinct groups to reflect their enthusiasm and specificity regarding project topics - "visionary", "collaborator", and "enabler". Finally, we employ an intelligent matching algorithm that strategically groups students, optimizing the balance between visionary project leaders and adaptable team members.

In addition to our proposed approach, we have also evaluated our algorithm against greedy approaches to maximize vision alignment and alternative LLM-based strategies. By thoughtfully combining these approaches, GroupMate aims to mitigate common issues such as inefficient team formations, mismatched skill sets, and low group cohesion. Our system's ultimate goal is to streamline and enhance the student team formation process, significantly improving collaborative outcomes and reducing administrative burdens for instructors and students alike. This report outlines the design, development, and evaluation of GroupMate, emphasizing its technological foundations and potential impact on educational experiences and collaborative efficiency.

*All authors contributed equally to this project. Details can be found in 11 section. Author order is alphabetical by last name.
[1] https://github.com/jocelynxu01/GroupMate

2 Motivation and Significance

The motivation behind GroupMate stems from a prevalent issue in academic environments: inefficient and ineffective team formation processes. Traditional methods of forming student teams, which are either manual assignment by instructors or self-selection by students, often lead to suboptimal outcomes. Manual assignments are typically time-consuming for instructors and may not fully account for individual student preferences or complementary skills. Conversely, self-selected teams frequently result from convenience or existing social connections rather than compatibility of skills or academic interests, potentially limiting collaborative potential and reducing educational outcomes.

GroupMate addresses these pain points directly by automating the team formation process in an intelligent, data-driven manner. By analysing project proposals written by students, GroupMate discerns varying degrees of student passion and commitment toward specific project topics, allowing the formation of teams that align students' enthusiasm levels appropriately. Moreover, employing Large Language Models (LLMs) to analyze required skill sets ensures teams have the necessary competencies to successfully complete their projects. This systematic approach fosters balanced, cohesive groups with clearly defined roles and complementary skill sets, enhancing both collaboration and productivity.

Upon finalizing our problem statement, we conducted a literature review of existing group formation algorithms. Most prior works frame the problem as an optimization task, focusing on skill coverage, workload balance, or topic similarity. These methods often treat project interests and skill sets as rigid inputs—assuming students can only contribute to the exact projects they initially proposed. However, this does not reflect the collaborative reality of team-based innovation. In practice, students are often open to adapting their ideas and contributing to related projects, especially when their skills align well with others' visions. Recognizing this flexibility, we designed a novel algorithm that goes beyond interest matching to construct high-functioning, innovation-driven teams.

Our approach draws inspiration from Belbin's framework of nine team roles ([2]), which highlights the value of diverse personality types and contributions within a team. Based on this foundation, we developed a simplified yet practical model that categorizes students into three archetypes:

- **Visionaries:** Individuals whose proposals are forward-thinking, original, and positioned at the frontier of research. Their proposals are typically novel and interdisciplinary, pushing the boundaries of conventional student projects.
- **Collaborators:** Individuals who form the bridge between visionaries and technical specialists. They possess moderate interdisciplinary exposure, depth in one or more relevant skill areas, and exhibit a strong inclination to contribute toward realizing another's vision through cooperative problem-solving.
- **Enablers:** Specialists with deep domain expertise who may not themselves propose frontier or novel ideas, but demonstrate strong implementation skills and experience in extending or enabling the realization of others' visions. While they may not always propose the initial concept, they possess the technical or operational expertise to implement it effectively.

Our goal was to ensure every group includes at least one member from each of these roles. By emphasizing this balance, we aim to maximize each team's capacity for innovation, collaboration, and implementation; producing not only well-rounded teams but also dynamic and adaptable project outcomes.

3 System Architecture

GroupMate consists of several modular components that work together to automate student group formation and project idea generation. As shown in Figure 1, students begin by submitting a vision essay and a list of skills during registration. This information is passed through two key modules: First, the *Vision Scoring Module* categorizes students into one of three pools—Visionaries, Collaborators, and Enablers on a fine-tuned transformer model. Simultaneously, the *Skill Extraction Module* uses LLM prompting to infer technical skills likely needed for each student's envisioned project.

The *Group Matching Algorithm* initializes groups with Visionaries as seed members and incrementally fills them by assigning Collaborators based on a weighted combination of their project proposal similarity and the extent to which their skills cover the group's missing requirements, followed by Enablers to fill in missing skill requirements and balance team sizes. Once the groups are complete, a *Project Suggestion Module* leverages an LLM to generate tailored project ideas using each group's aggregated vision essays and combined skillsets. The final output includes team assignments, recommended projects, and inferred required skills.

3.1 Vision Scoring Model

Motivation. The objective of the vision scoring model is to enable the formation of effective project teams in classroom settings by quantitatively identifying the roles of individuals based on their proposed project ideas and skillsets. We need a metric to categorize students into three broad functional roles: Visionaries, Collaborators and Enablers.

To support structured grouping of students into complementary teams that balance ideation with execution, it is necessary to estimate how "visionary" a student's proposal is. This estimation is formalized as a vision novelty score, which serves as the foundation for the assignment of the downstream group.

Rationale for Avoiding Personality Questionnaires. Although personality-based self-assessments might appear useful for assigning functional roles, we intentionally avoid such methods for three key reasons. First,

- These questionnaires rely on self-perception, which is inherently subjective. In a classroom setting, it is not uncommon for a large proportion of students to self-identify as "visionaries," making it impossible to guarantee balanced team compositions.
- We believe that team roles should be derived from observable evidence—namely, the content and originality of a student's proposal and their demonstrated skills rather than self-declared traits.
- The designation of roles like Visionary, Enabler, or Collaborator is context-dependent: a project idea that appears visionary in one domain may reflect strong enabling instincts

in another. Thus, tying role assignments to the proposal itself rather than perceived personality ensures that grouping decisions remain grounded, interpretable, and aligned with course-specific goals.

Rationale for Custom Modeling. While large language models (LLMs) offer broad capabilities in evaluating and generating natural language, several factors limit their suitability for this task:

- **Fine-grained originality:** General-purpose LLMs may lack sensitivity to detect nuanced novelty in technical content.
- **Consistency and reproducibility:** LLM outputs are non-deterministic, making scoring inconsistent and potentially unfair.
- **System integration:** A local model supports a tighter integration with the group formation pipeline.

Given these limitations, a supervised regression model fine-tuned on domain-specific data was selected. Our approach draws inspiration from prior work in academic influence prediction, where dynamic heterogeneous networks and multi-relational modeling have been used to estimate the future impact of papers, authors, and venues (Zhang et al.[12]). However, unlike such citation-focused models, our work shifts the focus to semantic and interdisciplinary signals present at the proposal stage, enabling proactive team formation rather than retrospective evaluation.

Dataset and Preprocessing. The model is trained using the arXiv metadata snapshot (`arxiv-metadata-oai-snapshot.json` [1]), filtered to include only articles that:

- were updated in 2023 or later,
- belonged to Computer Science (`cs.*`) categories,
- contained valid metadata (abstract, title, categories, date, authors).

This yielded a cleaned dataset of approximately 50,000 papers.

Ground Truth Construction. Three novelty heuristics are combined to compute a supervisory signal:

(1) **Categorical Entropy (Interdisciplinary Novelty):** quantifies the diversity of expertise among co-authors by aggregating their historical category-wise publication counts. The combined author–category distribution is used to compute Shannon entropy, capturing the breadth of interdisciplinary collaboration. Higher entropy values indicate visionary teams with wide-ranging backgrounds.
(2) **Semantic Distance (Topical Originality):** estimates how conceptually distinct a paper is from others in the embedding space. Each paper is embedded using a pretrained SciBERT model [3] and normalized for cosine similarity. We use FAISS [6] to retrieve the top-50 nearest neighbors for each paper and compute the average cosine similarity. Semantic novelty is defined as one minus this average similarity, penalizing papers in dense topical regions and rewarding those in sparse, underexplored areas.
(3) **Lexical Similarity (Trend Alignment):** evaluates how closely a paper's vocabulary aligns with popular terminology in its lexical cluster. TF–IDF vectors are used to form 20 lexical clusters, and each paper's cosine similarity to its cluster centroid is computed. This rewards participation in current, high-traction conversations.

Vision Score Computation. Each heuristic is normalized to the $[0, 1]$ range. The final vision score is computed as:

$$\begin{aligned}\text{VisionScore} = {} & 0.5 \times \text{SemanticDistance} \\ & + 0.4 \times \text{CategoricalEntropy} \\ & + 0.1 \times \text{LexicalSimilarity}\end{aligned}$$

This weighted sum reflects the prioritization of conceptual novelty, interdisciplinary breadth, and trend relevance.

Model Training. A SciBERT-based regression model [3] is trained to map text to the computed vision scores. Each input is structured as:

[CLS] Abstract text [SEP] arXiv categories

A regression head is added to the `[CLS]` token output. The model is fine-tuned using mean squared error (MSE) loss on the 50k-paper dataset. A 20% validation split and early stopping (patience = 3) are used to prevent overfitting.

Application to Student Vision Essays. Student-submitted vision essays are treated as abstracts. The student's skillset and background are mapped to arXiv categories to simulate contextual metadata. The trained model predicts a vision score for each student, which is then used to assign their role as a *visionary*, *collaborator*, or *enabler* in team formation.

To complement the novelty score, an auxiliary passion signal is computed using VADER (Valence Aware Dictionary and sEntiment Reasoner [8])—a rule-based sentiment analysis tool optimized for short, informal text. VADER analyzes the tone of the essay and yields a compound sentiment score, which serves as a weak proxy for student enthusiasm and emotional engagement with the proposed idea.

Together, the vision score and sentiment score allow for a more holistic and human-aligned evaluation of project ideas, ensuring that teams are formed around compelling visions led by motivated individuals, with collaborators and enablers aligned to support both the ambition and execution of those ideas.

3.2 Matching Algorithm

To form cohesive and balanced teams, our matching algorithm operates in a series of steps (see Algorithm 1). In the initial phase, each student's vision essay is processed to produce both a quantitative vision score and a semantic embedding (Algorithm 1, lines 1–2). Vision scores are obtained by passing tokenized essays through a fine-tuned SciBERT regressor, while embeddings are generated via a pre-trained MPNet encoder, using either its pooled output or a mean-pooled representation of the hidden states.

Once scores and embeddings have been computed, all students are ordered in descending vision-score order and partitioned into three cohorts (Algorithm 1, line 3). The top third of students form the Visionaries Pool, the next third become the Collaborators Pool, and the remainder are designated as the Enablers Pool. This stratification ensures that teams are seeded by high-vision members while progressively integrating complementary collaborators and enablers.

Subsequently, groups are instantiated—one per visionary—with each core member assigned and all skill lists initialized to empty (Algorithm 1, lines 4–7). For each group, the concatenated vision essays of its members are submitted to an LLM, which selects five to seven pertinent skills from a master list; in the event no skills are returned, three are chosen at random as a fallback (Algorithm 1, lines 7–13).

The weighted assignment of Collaborators then proceeds in a loop (Algorithm 1, lines 14–31). For each group, the deficit between inferred needed skills and currently held skills is computed. Every remaining collaborator is scored by combining (a) the cosine similarity between the group's mean embedding and the collaborator's embedding and (b) the fraction of missing skills the collaborator possesses. The highest-scoring collaborator who contributes at least one new skill is assigned, the group's skill inventory is updated, and the collaborator is removed from further consideration; this repeats until no new assignments occur in a full pass.

After weighted matching, any leftover collaborators are allocated in a single pass to the smallest groups by membership count (Algorithm 1, lines 33–35). This guarantees that every collaborator is placed, even if they did not address specific skill gaps.

Next, Enablers are added to bring each group to four members (Algorithm 1, lines 36–43). For each under-filled group, the set of remaining skill gaps is recalculated, and the enabler whose skill set most overlaps these gaps is chosen. If no overlap exists, the first available enabler is assigned. This process continues until all groups reach the target size.

In the final phase, two project ideas per group are generated via the LLM using the group's aggregated vision essays and consolidated skill list (Algorithm 1, lines 44–46). The first idea reflects the group's shared vision, and the second leverages the group's collective skills; these outputs, along with the final team compositions and skill inventories, are then prepared for reporting. Unlike traditional matching algorithms that prioritize either skills or interests, our approach sequentially layers vision, skill complementarity, and balanced assignment to produce teams that are vision-led, skills-complete, and equitably distributed.

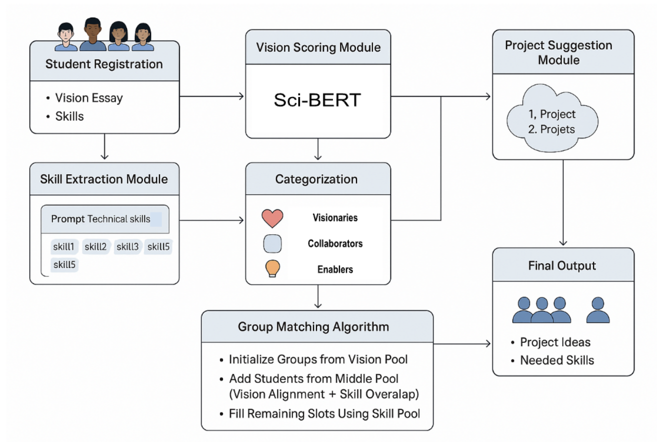

Figure 1: System architecture for student grouping and project suggestion.

Algorithm 1 Student Group Formation with Weighted Matching

Require: List of students with $\{vision\_essay, skills\}$
Ensure: Groups of students with project ideas

1: Compute $vision\_score$ and embedding e_s for each student
2: Sort students by $vision\_score$ descending
3: Split into three pools: Visionaries, Collaborators, Enablers
4: Initialize $groups \leftarrow \{\{v\} \mid v \in \text{Visionaries}\}$ with empty skill lists
5: Load $total\_skills$ from skills file
6: **for all** group in $groups$ **do**
7: $\quad essays \leftarrow$ concat member essays
8: $\quad needed \leftarrow \text{InferSkillsLLM}(essays, total\_skills)$
9: $\quad$ **if** $needed$ is empty **then**
10: $\quad\quad needed \leftarrow$ random 3 from $total\_skills$
11: $\quad$ **end if**
12: $\quad$ group.$needed\_skills \leftarrow needed$
13: **end for**
14: **while** true **do**
15: $\quad assigned \leftarrow$ False
16: $\quad$ **for all** group in $groups$ **do**
17: $\quad\quad missing \leftarrow group.needed\_skills - group.current\_skills$
18: $\quad\quad$ **if** $missing = \emptyset$ **then continue**
19: $\quad\quad$ **end if**
20: $\quad\quad$ **for all** c in Collaborators **do**
21: $\quad\quad\quad score(c) = 0.3 \cdot \cos(e_{\text{group}}, e_c) + 0.7 \cdot \frac{|c.skills \cap missing|}{|missing|}$
22: $\quad\quad$ **end for**
23: $\quad\quad c^* \leftarrow \arg\max_c score(c)$
24: $\quad\quad$ **if** $c^*.skills \cap missing \neq \emptyset$ **then**
25: $\quad\quad\quad$ assign c^* to group; update current_skills
26: $\quad\quad\quad$ remove c^* from Collaborators
27: $\quad\quad\quad assigned \leftarrow$ True
28: $\quad\quad$ **end if**
29: $\quad$ **end for**
30: $\quad$ **if** $\neg assigned$ **then break**
31: $\quad$ **end if**
32: **end while**
33: **for all** c in Collaborators **do**
34: $\quad$ assign c to smallest group; update current_skills
35: **end for**
36: **for all** group in $groups$ **do**
37: $\quad$ **while** $|group.members| < 4$ **do**
38: $\quad\quad$ choose enabler e maximizing $\frac{|e.skills \cap missing|}{|missing|}$
39: $\quad\quad$ **if** all scores = 0 **then** choose first enabler
40: $\quad\quad$ **end if**
41: $\quad\quad$ assign e; update current_skills; remove e
42: $\quad$ **end while**
43: **end for**
44: **for all** group in $groups$ **do**
45: $\quad$ generate two project ideas via LLM from group essays and skills
46: **end for**

4 User Roles and Workflow

GroupMate supports two primary user roles: students and instructors, each with distinct functionalities tailored to streamline the team formation process in an academic setting.

Students. Upon registering on the platform, students can join a specific course using a unique code provided by the instructor. After enrollment, students are prompted to submit information that helps the system make informed team assignments. This includes skill sets, course history, and a project proposal outlining their goals and the project they are interested in. Once all student data is collected, team formation is initiated by the instructor—not dynamically updated—ensuring a controlled and one-time assignment process. After teams are assigned, students receive not only their team composition but also project topic suggestions. These suggestions are generated using LLMs that analyze the aggregated vision essays, submitted materials, and course histories of team members to recommend relevant and aligned project ideas. The view of the user interface as a signed-in student is shown in Figure 2.

Instructors. Instructors begin by registering and creating a new course within the system. They are responsible for inviting students to join the course and monitoring enrollment progress. Once all submissions are complete, the instructor can initiate the team formation process via a single action, triggering the recommendation algorithm. The instructor can then view and verify the generated teams and monitor student participation, ensuring appropriate group compositions and project alignment. This functionality provides transparency and oversight, allowing instructors to intervene or refine groupings if necessary before finalizing teams. The view of the user interface as a signed-in instructor is shown in Figure 3.

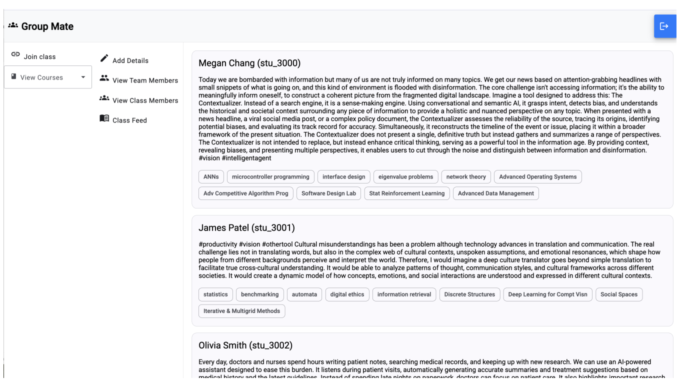

Figure 2: View as student

5 System Implementation

GroupMate is implemented as a website that can be used across courses. Students and instructors have access to the website. A brief summary of the tech stack is

- **Backend Technologies:** Python Django Framework
- **Frontend Interface:** Angular Framework
- **LLM Integration:** Gemini
- **Data Storage and Retrieval:** SQLite

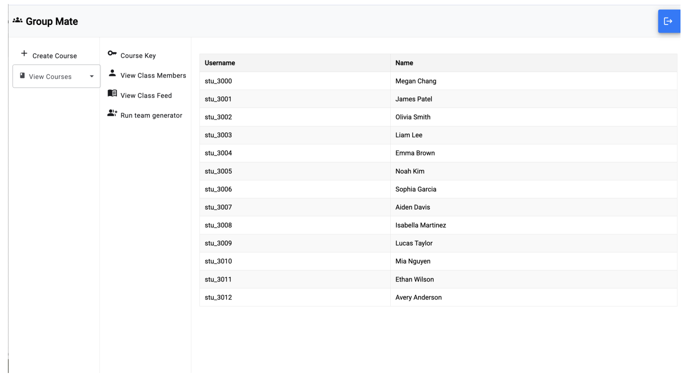

Figure 3: View as instructor

5.1 User authentication and authorization

The authentication and authorization pipeline uses Django authentication system. Users sign up by providing some basic information and setting a password. To store passwords securely, Django uses the PBKDF2 algorithm with a SHA256 hash, a password stretching mechanism recommended by NIST. On logging in, the endpoint produces an access JWT (JSON Web Token) token and a refresh token. Access tokens are valid for five minutes, while refresh tokens are valid for 24 hours. The system stores the refresh tokens in the browser's local storage and refreshes the access token every time it expires. When the refresh token expires, user is redirected to login. This minimises the damage that can be caused when a token is compromised and to prevent unauthorized access. This also enables stateless and secure communication between frontend and backend. The authorization enforces Role Based access control (RBAC). RBAC ensures the distinction of roles and ensures that the Instructor operations are available exclusively to them.

5.2 Backend Endpoints

All the capabilities as seen in Figure 2 and Figure 3 are implemented as endpoints in the backend. The requests are designed as shown in Table 1

Headers	Bearer <access token>
Body	key-value pairs in JSON format

Table 1: Request Format

The backend architecture uses Django REST Framework, which enables the RESTful (Representational State Transfer) architectural style of the web-application.

5.3 Database

The application is not heavily data-oriented, hence we use an elegant DBSqlite 3 as the database. The models for the database extend django.db.models.Model. Figure 4 shows the Entity-Relationship diagram and how the tables are modelled.

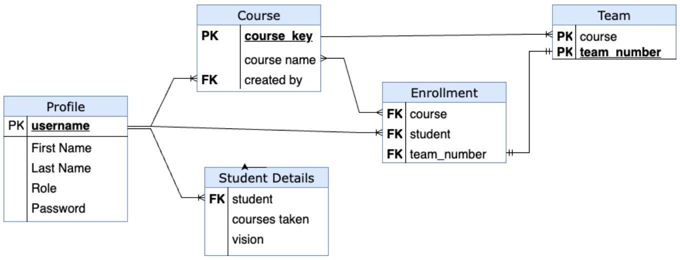

Figure 4: Entity-Relationship Diagram

5.4 User interface

Angular's component-based architecture enables modular UI development and reactive forms for smooth user interaction and validation. Angular's Ahead-of-Time (AOT) compilation and lazy loading reduces initial load time and enhances responsiveness. The User interface is API-driven– all the user data and group functionalities are fetched and updated via REST API calls to the Django backend, ensuring clear separation of concerns.

6 Evaluation and Results

6.1 Analysis of the Vision Scoring Model

Our fine-tuned SciBERT model, trained on arXiv metadata and heuristic novelty signals, performed reasonably well in classifying student vision essays into *Visionary*, *Collaborator*, and *Enabler* categories. As shown in Appendix B Table 3, the model correctly classified 11 out of 13 students (85%), demonstrating strong alignment with human judgment in most cases.

Strengths.

- The model was effective in recognizing highly original, interdisciplinary ideas and labeling them as *Visionary*. Examples include `stu_3000`, `stu_3005`, and `stu_3012`, whose essays exhibited both conceptual novelty and transformative ambition.
- It correctly identified domain-specific but practical applications (e.g., healthcare documentation, sports analytics) as *Enabler* roles, suggesting it learned to associate implementation-driven visions with this category.
- *Collaborator* classifications were well-handled for students whose ideas extended existing technologies or served as support mechanisms in broader ecosystems.

Observed Limitations.

- The model misclassified two students:
  - `stu_3008` (Isabella) was placed in *Enabler*, but her emotional feedback tool exhibited strong originality and interdisciplinary novelty—qualities more aligned with *Visionary*.
  - `stu_3011` (Ethan) was also tagged *Enabler*, though his vision reflected a supportive role in tailoring systems to users' cognitive goals—better matching a *Collaborator*.
- These errors suggest the model may over-rely on surface features or domain specificity (e.g., healthcare or affective tools) without fully capturing deeper intent or novelty signals in the essay.

Grouping Method	Groups
GroupMate	Group 1: [stu_3005, stu_3009, stu_3006, stu_3003] Group 2: [stu_3010, stu_3001, stu_3011, stu_3008] Group 3: [stu_3000, stu_3004, stu_3002] Group 4: [stu_3012, stu_3007]
Greedy Grouping	Group 1: [stu_3003, stu_3005, stu_3012] Group 2: [stu_3004, stu_3010, stu_3011] Group 3: [stu_3000, stu_3002, stu_3006] Group 4: [stu_3001, stu_3007, stu_3008] Group 5: [stu_3009]
LLM Grouping	Group 1: [stu_3000, stu_3001, stu_3002] Group 2: [stu_3005, stu_3007, stu_3010] Group 3: [stu_3003, stu_3011, stu_3012] Group 4: [stu_3004, stu_3006, stu_3008, stu_3009]

Table 2: Comparison of Student Groupings Across Different Algorithms

6.2 Grouping Algorithm Analysis

This analysis compares three team grouping strategies GroupMate Matching Algorithm, Greedy Grouping, and LLM Grouping using three core criteria: **Balance of Roles** (Visionary, Collaborator, Enabler), **Vision Similarity**, **Skill/ Course Diversity and Coverage**

The ideal grouping prioritizes balanced roles for innovation potential, ensures vision alignment for synergy, and covers diverse skills to realize the group's vision.

Balance of Roles.

- **GroupMate Matching Algorithm:** Strongest in achieving role balance provided the vision scoring model is reliable. Every team consistently included at least one Visionary, one Collaborator, and one Enabler. For example, one team featured a visionary imagining an immersive AI knowledge universe, paired with a collaborator skilled in user design, and an enabler experienced in ML engineering. This guarantees that all teams have a balanced mix of conceptual leadership, coordination, and technical execution. A minor trade-off is that some teams may initially lack a unified project vision.
- **Greedy Grouping:** Weakest on this criterion. It produced several unbalanced teams such as one with two Visionaries and no Enabler, and another consisting entirely of Collaborators. These imbalances make innovation execution harder and highlight the danger of heuristic-first approaches without role-aware constraints.
- **LLM Grouping:** Moderate role balance. Many groups included a variety of roles, though consistency was lower than our proposed algorithm. For instance, one team featured two Visionaries and one Enabler working on an AR-related vision, but lacked a strong Collaborator. The LLM's strength in vision clustering came at the cost of structural guarantees.

Vision Similarity.

- **GroupMate Matching Algorithm:** Moderate alignment. Because the primary goal was role and skill balance, some groups included divergent visions (e.g., AI for healthcare alongside AR for education). While thematic overlaps were often present, teams will have to invest time aligning their ideas.
- **Greedy Grouping:** Highly variable. The first group had tight vision alignment (e.g., two students exploring AR for education), but leftover students often formed groups with disparate or conflicting visions, which reduced cohesion and innovation direction.
- **LLM Grouping:** Strongest performance. Teams were formed around clearly related themes, such as one group tackling information overload using semantic AI tools, and another focused on immersive AR learning environments. This thematic coherence fosters creative synergy and rapid ideation, but sometimes comes at the cost of role or skill diversity.

Skill Diversity and Coverage.

- **GroupMate Matching Algorithm:** High and deliberate. Teams were cross-functional, covering ML, web dev, design, data science, and domain-specific knowledge. For example, one team combined a frontend developer, an ML engineer, and a UX specialist to support a complex AI tool vision.
- **Greedy Grouping:** Inconsistent and often narrow. Teams sometimes had redundant expertise (e.g., multiple backend developers, but no frontend or ML specialist), limiting project scope and balance. This randomness undermines innovation feasibility.
- **LLM Grouping:** Generally good alignment with vision, though with occasional redundancies. The LLM often grouped students with complementary skills relevant to the vision, but sometimes repeated skill profiles (e.g., three students with similar ML backgrounds), which could limit functional breadth.

In summary, GroupMate Matching Algorithm excels at building robust, well-balanced teams with diverse skills and roles. Slightly lower vision alignment is its trade-off. Greedy Grouping is simple but inconsistent, leading to role and skill imbalances. LLM Grouping creates thematically cohesive teams with good synergy, though structural consistency varies.
Trade-off Insight: There is an inherent tension between grouping for shared vision versus diverse capabilities. The best algorithms balance these factors rather than optimizing only one.
GroupMate Matching Algorithm provides structural completeness, LLM Grouping enables creative synergy, and Greedy Grouping highlights the risk of overly simplistic heuristics. For innovation-focused teamwork, a hybrid of GroupMate Matching Algorithm and LLM Grouping offers the most promise.

6.3 Qualitative User Observations

Friction: Based on our experience, students approach group projects with varying levels of interest, engagement, and attachment to specific ideas. While some students have strong visions and prefer to lead, others are more flexible and open to contributing in any direction. We found that having a mix of these types within a team can actually improve group dynamics. For example, placing two highly opinionated students in the same group may lead to potential friction or competition over project control. In contrast, flexible students can help absorb tension and promote smoother collaboration, acting as bridges across differing ideas and helping teams stay focused and cohesive. This observation directly informed our vision scoring model design, where we explicitly categorize students based on vision strength and use that structure to strategically pair flexible members with those who have strong preferences.

Group Dynamics: In our experience working on group projects and observing team dynamics, we found that having students with complementary skill sets and slightly different academic backgrounds can actually lead to more effective collaboration. Instead of creating conflict, this kind of "friction" often encourages students to engage more actively, share perspectives, and learn from one another. For example, a student with a strong technical background might approach a project differently than one with more design or theoretical experience — and that contrast can help the team build a more well-rounded solution. By including a mix of students with overlapping but not identical strengths, our system supports dynamic group interactions and richer project development. Our matching algorithm leverages these differences to form balanced teams with complementary expertise, fostering productive group dynamics rather than redundancy or conflict.

7 Privacy and Ethical Considerations

As this project was developed for a course setting, we made careful efforts to uphold privacy and ethical standards when constructing our dataset. For the purpose of training and testing our system, we web scraped real vision essays written by students in the same course. However, we deliberately excluded any personally identifiable information. All user identities were anonymized using randomly generated fake names and emails to protect student privacy.

Additionally, the course-related data used for matching and recommendation was sourced from publicly available descriptions of UIUC CS courses, which were web scraped for content only. We did not collect any sensitive student or institutional information beyond what is publicly accessible or voluntarily submitted through the user interface. The resulting dataset is strictly used for academic purposes and system prototyping, ensuring no violation of user consent or institutional policies.

8 Challenges and Limitations

Despite GroupMate's promising framework, our current implementation faces several practical and methodological limitations:

- **Lack of Real Data:** Our dataset is constrained to a single course with 88 students. While we were able to scrape real vision essays, the platform used for submission does not enforce a standardized naming or email format—some students use NetIDs, others use full names—making it difficult to reliably match essays with individual identities. As a result, our training and evaluation processes rely heavily on heuristics and synthetic approximations.

- **Limitations in Evaluating Model Performance:** The vision scoring model focuses on measuring conceptual novelty and interdisciplinary breadth. A challenge arises from the fact that most student essays in our dataset were written under a semi-graded assignment structure that encouraged creativity and effort. As a result, the distribution of novelty scores is narrow, making it difficult to evaluate the model's ability to discriminate truly standout ideas from more derivative ones. Furthermore, the model has not yet been validated in open-ended or cross-institutional settings, where vision diversity and writing styles may be more varied.
- **Lack of Diversity in Data Sources:** All training and evaluation data come from CS students at a single institution. This narrow demographic may limit the generalizability of our models to other disciplines or academic cultures, where vision essays and collaboration styles may differ significantly.
- **Unvalidated LLM-Generated Project Suggestions:** While GroupMate incorporates LLMs to recommend project ideas based on team composition, these suggestions have not been evaluated with real users for relevance, novelty, or practicality. As a result, their usefulness remains speculative until tested in classroom or project settings.
- **Semantic Novelty Biases in Dense Domains:** Our semantic novelty measure in Section 3.1 assumes that ideas located in sparse regions of the embedding space are more original. However, this can lead to two key issues. First, papers in high-density research areas such as machine learning or systems may be unfairly penalized despite making meaningful contributions, simply because many papers cluster around shared vocabulary and structure. Second, abstracts that are incoherent, off-topic, or synthetically generated may appear spuriously novel due to their distance from well-formed clusters. This highlights the need for incorporating additional quality checks or hybrid novelty signals beyond embedding sparsity alone.
- **Oversimplified Modeling of Collaboration:** In our current framework, collaborators are implicitly treated as individuals who fall between visionaries and enablers—often those with moderate novelty and interdisciplinary scores. However, this assumption may not capture the true nature of effective collaboration. Strong collaborators often excel in facilitation, communication, and team adaptability, traits that may not correlate directly with research originality or category entropy. Future iterations of GroupMate should explore alternative signals such as peer evaluations, historical project outcomes, or discourse analysis to more accurately identify and model collaboration skills.
- **Assumption of Vision Essay Availability:** Our approach currently assumes the presence of vision essays or project proposals as inputs for modeling student intent and novelty. While we acknowledge that not all courses require students to write detailed essays, we believe that asking students to briefly articulate an idea or project they are interested in is a reasonable and low-overhead addition. Beyond aiding the team formation process, this step encourages students to pause and reflect on their personal goals and motivations before entering collaborative settings helping align individual aspirations with group outcomes.

9 Future Work

While GroupMate provides a strong initial framework for intelligent student team formation, there are several opportunities to enhance its effectiveness and applicability in future iterations:

Expand Data Sources and Label Diversity: One major next step is to curate a richer and more diverse dataset for training and evaluation. In addition to vision essays, future versions of GroupMate could incorporate student resumes, GitHub repositories, and self-reported working style or personality preferences. Collecting data from courses that do not require high-effort, graded vision essays may also help introduce more natural variation in passion levels, which is critical for evaluating the robustness of the vision scoring model.

Evaluate and Improve Project Recommendations: Although our system uses LLMs to generate project ideas based on team composition, these suggestions have not yet been tested with real users. Future work will involve gathering student and instructor feedback to assess the relevance, creativity, and feasibility of these recommendations, and refining generation prompts accordingly.

Platform Scaling and Deployment: Currently, GroupMate is built as a prototype for CS courses. Expanding the platform for broader deployment across departments, disciplines, or institutions could reveal generalization limits and new design constraints. Additionally, supporting dynamic team formation—such as real-time adjustments for group imbalance or mid-semester changes—would increase the system's flexibility and real-world value.

Conduct a Comprehensive User Study: To better evaluate the practical impact and usability of GroupMate, we plan to conduct a structured user study within a UIUC course involving collaborative, open-ended projects. This would allow us to assess how well the team formation algorithm aligns with student satisfaction, perceived compatibility, and actual project outcomes. It would also offer valuable feedback on user experience, project recommendation quality, and how students perceive their assigned roles (visionary, enabler, collaborator) in real-world settings.

10 Conclusion

GroupMate bridges the gap between algorithmic rigor and human-centered design in student group formation. By integrating fine-tuned vision scoring with LLM-driven project analysis and structured skill matching, our system produces high-functioning, passion-aligned teams that go beyond random assignment or superficial similarity metrics. Through comparative evaluation with a greedy similarity-based baseline and a fully LLM-driven approach, GroupMate shows strong potential in facilitating intentional team dynamics and improving group outcomes.

While challenges remain around scaling, personalization, and validation of generated project suggestions, our system offers a compelling proof-of-concept for how AI can support collaborative learning environments. Future iterations will focus on expanding data diversity, integrating personality and collaboration style preferences, and refining project generation through iterative human-in-the-loop feedback.

11 Contributions

Padma Pooja Chandran designed and developed the pipeline for the group matching algorithm (GroupMate and greedy grouping). She contributed to writing the pseudocode, system architecture, and Appendix A sections of the report.
Aarthy Kesavan Padmanaban developed the web interface, including both the frontend and backend, and integrated the group matching algorithm into the backend. She contributed to writing the system implementation section of the report.
Krishnaveni Unnikrishnan proposed the idea of the vision scoring model and trained the SciBERT-based regression model. She contributed to writing the vision scoring model section and Appendix B, and also helped brainstorm the limitations and future work sections.
Jocelyn Xu conducted comparative analyses across different models used in the experiments. She contributed to writing the introduction, motivation, user roles, and the privacy and ethical considerations sections of the report.

Overall, all authors worked together polishing this report and contributed equally to this project.

References

[1] arXiv.org submitters. 2024. arXiv Dataset. doi:10.34740/KAGGLE/DSV/7548853
[2] R. Meredith Belbin. 2010. Chapter 1 - Making teams effective. In *Management Teams (Third Edition)* (third edition ed.), R. Meredith Belbin (Ed.). Butterworth-Heinemann, Oxford, 1–11. doi:10.1016/B978-1-85617-807-5.10001-9
[3] Iz Beltagy, Kyle Lo, and Arman Cohan. 2019. SciBERT: A Pretrained Language Model for Scientific Text. arXiv:1903.10676 [cs.CL] https://arxiv.org/abs/1903.10676
[4] Da Cao, Xiangnan He, Lianhai Miao, Yahui An, Chao Yang, and Richang Hong. 2018. Attentive Group Recommendation. Association for Computing Machinery.
[5] Tong Chen, Hongzhi Yin, Jing Long, Quoc Viet Hung Nguyen, Yang Wang, and Meng Wang. 2022. Thinking inside The Box: Learning Hypercube Representations for Group Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*.
[6] Matthijs Douze, Alexandr Guzhva, Chengqi Deng, Jeff Johnson, Gergely Szilvasy, Pierre-Emmanuel Mazaré, Maria Lomeli, Lucas Hosseini, and Hervé Jégou. 2025. The Faiss library. arXiv:2401.08281 [cs.LG] https://arxiv.org/abs/2401.08281
[7] Lei Guo, Hongzhi Yin, Tong Chen, Xiangliang Zhang, and Kai Zheng. 2021. Hierarchical Hyperedge Embedding-based Representation Learning for Group Recommendation.
[8] C. Hutto and Eric Gilbert. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media* 8, 1 (May 2014), 216–225. doi:10.1609/icwsm.v8i1.14550
[9] Renqi Jia, Xiaofei Zhou, Linhua Dong, and Shirui Pan. 2021. Hypergraph Convolutional Network for Group Recommendation. In *2021 IEEE International Conference on Data Mining (ICDM)*.
[10] Aravind Sankar, Yanhong Wu, Yuhang Wu, Wei Zhang, Hao Yang, and Hari Sundaram. 2020. GroupIM: A Mutual Information Maximization Framework for Neural Group Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*.
[11] Xixi Wu, Yun Xiong, Yao Zhang, Yizhu Jiao, Jiawei Zhang, Yangyong Zhu, and Philip S. Yu. 2023. ConsRec: Learning Consensus Behind Interactions for Group Recommendation. In *Proceedings of the ACM Web Conference 2023*.
[12] Fang Zhang and Shengli Wu. 2020. Predicting future influence of papers, researchers, and venues in a dynamic academic network. *Journal of Informetrics* 14, 2 (2020), 101035. doi:10.1016/j.joi.2020.101035
[13] Junwei Zhang, Min Gao, Junliang Yu, Lei Guo, Jundong Li, and Hongzhi Yin. 2021. Double-Scale Self-Supervised Hypergraph Learning for Group Recommendation. In *Proceedings of the 30th ACM International Conference on Information amp; Knowledge Management*.

A Appendix A: Student Profiles

stu_3000: Megan Chang

Email: megan.chang@example.com

Skills: transformers (Hugging Face), LlamaIndex, NetworkX, Python, JavaScript, Vertex AI, MongoDB, Rasa, D3.js, ZeroMQ, AWS, Go, JavaScript

Courses: Advanced Operating Systems, Adv Competitive Algorithm Prog, Software Design Lab, Stat Reinforcement Learning, Advanced Data Management

Vision: Today we are bombarded with information but many of us are not truly informed on many topics. We get our news based on attention-grabbing headlines with small snippets of what is going on, and this kind of environment is flooded with disinformation. The core challenge isn't accessing information; it's the ability to meaningfully inform oneself, to construct a coherent picture from the fragmented digital landscape. Imagine a tool designed to address this: The Contextualizer. Instead of a search engine, it is a sense-making engine. Using conversational and semantic AI, it grasps intent, detects bias, and understands the historical and societal context surrounding any piece of information to provide a holistic and nuanced perspective on any topic. When presented with a news headline, a viral social media post, or a complex policy document, the Contextualizer assesses the reliability of the source, tracing its origins, identifying potential biases, and evaluating its track record for accuracy. Simultaneously, it reconstructs the timeline of the event or issue, placing it within a broader framework of the present situation. The Contextualizer does not present a single, definitive truth but instead gathers and summarizes a range of perspectives. The Contextualizer is not intended to replace, but instead enhance critical thinking, serving as a powerful tool in the information age. By providing context, revealing biases, and presenting multiple perspectives, it enables users to cut through the noise and distinguish between information and disinformation.

stu_3001: James Patel

Email: james.patel@example.com

Skills: Transformer, Rasa, Neo4J, Python, JavaScript, SQL, Kubernetes, D3.js, CUDA, SOCKETS.io, LangChain, HayStack, Grafana, GraphQL, ZeroQ

Courses: Discrete Structures, Deep Learning for Compt Visn, Social Spaces, Iterative & Multigrid Methods

Vision: Cultural misunderstandings has been a problem although technology advances in translation and communication. The real challenge lies not in translating words, but also in the complex web of cultural contexts, unspoken assumptions, and emotional resonances, which shape how people from different backgrounds perceive and interpret the world. Therefore, I would imagine a deep culture translator goes beyond simple translation to facilitate true cross-cultural understanding. It would be able to analyze patterns of thought, communication styles, and cultural frameworks across different societies. It would create a dynamic model of how concepts, emotions, and social interactions are understood and expressed in different cultural contexts.

stu_3002: Olivia Smith

Email: olivia.smith@example.com

Skills: Transformer, Rasa, Neo4J, Python, JavaScript, SQL, Kubernetes, D3.js, CUDA, SOCKETS.io, LangChain, HayStack, Grafana, GraphQL, ZeroQ

Courses: Practical Statistical Learning, Text Information Systems, CA Training

Vision: Every day, doctors and nurses spend hours writing patient notes, searching medical records, and keeping up with new research. We can use an AI-powered assistant designed to ease this burden. It listens during patient visits, automatically generating accurate summaries and treatment suggestions based on medical history and the latest guidelines. Instead of spending late nights on paperwork, doctors can focus on patient care. It also highlights important research tailored to each physician's specialty, ensuring they never miss a breakthrough. With its seamless integration into hospitals and clinics, the agent turns overwhelming medical data into clear, actionable insights—helping doctors make faster, more informed decisions and improving healthcare for everyone.

stu_3003: Liam Lee

Email: liam.lee@example.com

Skills: LlamaIndex, LangChain, OpenAI, FastAPI, Flask, Django, Google Cloud Platform (GCP), Python, JavaScript, spaCY, Next.js

Courses: Advanced Compiler Construction, Accel Fund of Algorithms II, Software Engineering I

Vision: I imagine a world where every piece of information we encounter—whether a snippet from a blog, a compelling quote in a book, or a quick insight during a lecture—automatically becomes part of our personal knowledge ecosystem. This "Personal Knowledge Curator" would act as an intelligent companion, seamlessly capturing, tagging, and sorting everything we read, watch, or discuss. Drawing on powerful AI-driven analytics, it would connect the dots across our various interests and projects, surfacing hidden relationships and suggesting new areas to explore. The tool would help students rapidly build a tailored library of concepts, enabling them to quickly see how one idea relates to another. Professionals would find it indispensable for staying on top of rapidly evolving fields, while everyday learners could enjoy easier recall of fascinating discoveries. Over time, this curator would refine its understanding of our unique goals, shaping insights that fit our personal styles of learning and thinking. In short, it wouldn't just store knowledge—it would spark creativity and fuel continuous growth.

stu_3004: Emma Brown

Email: emma.brown@example.com

Skills: TypeScript, R, SQL, HTML, CSS, Unity3D, Unreal Engine, ARKit, FAISS, C++

Courses: Ethical & Professional Conduct, Topics in Algrthmc Game Theory, Advanced Computer Networks, Artificial Intelligence, 3-D Vision

Vision: In our rapidly evolving digital age, one significant challenge we face is the overwhelming abundance of information generated across the internet from countless sources. People often need to browse multiple applications and platforms to find truly valuable information, where "value" is inherently subjective and varies from person to person. Unfortunately, humans have limited willpower each day, and decision-making, even for seemingly simple choices like what to eat for lunch or whether to read news before responding to emails, can be mentally taxing. This is where an innovative AI solution could make a meaningful difference in our daily lives. I envision an AI system that could automatically tag and categorize information, allowing users to easily access relevant content through personalized tags of their choosing. Moreover, the system could dynamically generate daily to-do lists based on event priorities and user preferences, helping manage decision fatigue and optimize daily productivity. The system would also serve as a personal digital archive, storing both user-uploaded content (such as photos and personal documents) and saved important information from various sources. What makes this system particularly powerful is its ability to retrieve stored information through natural language queries. Users could simply describe what they're looking for in conversational language, and the system would quickly locate and present relevant information from their personal archive. This intelligent assistant would effectively address the problem of information overload and decision fatigue while creating a more personalized and efficient daily experience for users. By combining AI-powered organization with natural language processing, it would transform how we interact with and manage our digital lives.

stu_3005: Noah Kim

Email: noah.kim@example.com

Skills: Unreal Engine, scikit-learn, TensorFlow, PyTorch, LangChain, C++, Python, GCP, LightFM, Contextual Bandits

Courses: Autonomous Vehicle System Eng, Ethical & Professional Conduct, CA Training, Advanced Compiler Technology, Cyber Dystopia

Vision: In the age of information explosion, human beings need more efficient ways to acquire knowledge. I envision an immersive Intelligent Knowledge Universe (IIKU) that combines VR/AR with AI, allowing users to "step into" the world of knowledge and experience information for themselves. Doctors can see 3D images of patients through AR during surgery, scientists can collaborate in real time in virtual LABS, and students can "travel" through history or explore the universe. IIKU enhances understanding through intelligent recommendation, data interaction and visualization, completely changes the way of learning, research and decision-making, and enables knowledge acquisition from reading to immersive experience, promoting the evolution of human intelligence.

stu_3006: Sophia Garcia

Email: sophia.garcia@example.com

Skills: Python, HTML, CSS, TensorFlow, spaCy, NLTK, Docker

Courses: Software Design Lab, Audio Computing Laboratory, Software Engineering Seminar, Computing & Global Developmnt

Vision: If there is a tool called the "Data Life Optimizer," it could address the issue of information overload, helping people find balance, work efficiently, stay healthy, and build stronger connections. This tool, powered by artificial intelligence and machine learning, would filter the overwhelming influx of daily information, such as tasks, news updates, and social interactions, identifying only what truly matters while eliminating distractions and noise. It would intelligently plan daily schedules based on user preferences and emotional states, ensuring time is allocated efficiently. On the social side, it could analyze interaction patterns and remind users to maintain meaningful connections with family and friends, preventing relationships from being neglected due to busy schedules.

stu_3007: Aiden Davis

Email: aiden.davis@example.com

Skills: Python, C++, C#, TypeScript, SQL, Pandas, scikit-learn

Courses: Distributed Systems, Distributed Systems Seminar, Cloud Stor Sys: Theory&Practic

Vision: Human cognition has always been limited by the senses and the way information is transmitted. We acquire knowledge indirectly through symbols, words and images, which are then transformed into understanding by the brain, a process that is both time-consuming and distorted, creating a gap between the cognitive and the real. I want to design a system for immersing consciousness. It connects to our brain and allows people to not just get information about something, but to experience it directly. It's a shift from "knowing" to "experiencing". The user's consciousness is able to enter the information environment and generate real perceptions and memories. Knowledge is no longer a symbol to be decoded, but a direct experience. This approach eliminates the intermediate links in traditional information acquisition, dramatically increasing the efficiency and depth of understanding. This way of understanding can transcend the limits of time and space. This will transform the fundamental way we interact with information, enabling humans to break through the traditional limits of cognition and establish a more direct connection to knowledge.

stu_3008: Isabella Martinez

Email: isabella.martinez@example.com

Skills: R, SQL, Sentiment & Emotion Recognition, Chart.js, D3.js, AWS

Courses: Algorithmic Genomic Biology, Advanced Computer Networks, Applied Parallel Programming, Advanced Compiler Technology

Vision: What if there is a tool that can not only process information, but also understands the emotions behind it? In our society, despite how hard we try to be fair and professional, there are always moments that people might be tired and driven to certain unwanted feelings which can potentially shape our choices or words. Imagine an emotional navigator which can visualize emotional undercurrents through feedback (maybe visual with colors), it would be so beneficial for different types of settings - professional meetings, personal reflections, or even just everyday interactions. This tool will not try to control feelings, rather it should create awareness for people to realize stress, frustrations, or unconscious bias. Think of it as an emotional compass, helping individuals and teams pause, recalibrate, and approach problems with clarity.

stu_3009: Lucas Taylor

Email: lucas.taylor@example.com

Skills: JavaScript, SQL, Rust, Swift, Kotlin, Sensor Integration & IoT, gRPC, FastAPI

Courses: Advanced Compiler Construction, Senior Project I, Trustworthy Machine Learning, Senior Project II

Vision: One area where data analysis is still in its early stages is sports. In recent years, a large amount of data is being captured from sports. However, most of the analysis is done offline, after the fact. Developments in AR and AI will be key to unlock effective realtime data analysis in sports. Coaches could use real time analysis to perform corrections. Players could have access to tactical analysis data based on live situations. Trainers could have access to AI-powered insights to look for fatigue markers and other physiological indicators. This could also improve the fans' experience by giving them AR overlays of stats and other technical details in a simple and easily graspable manner. By leveraging the power of AR, real-time data analysis, and AI, this tool could significantly improve athletic performance, reduce injuries, and enhance the overall sports experience for both participants and spectators.

stu_3010: Mia Nguyen

Email: mia.nguyen@example.com

Skills: SentenceTransformers, scikit-learn, XGBoost, LightGBM, FAST API, PyTorch, TensorFlow, Python, Firebase

Courses: Numerical Analysis, Intro to Computer Science II, Advanced Topics in Natural Language Processing, Adv Topics in Comp Arch, Excursions in Computing I

Vision: Today, the proliferation of information continues to pressure individuals from all walks of life, and people often feel fatigued in their data management and decision-making processes. Humans need a tool to solve this problem: AI situational assistant. Such assistants use advanced algorithms to filter and prioritize information to ensure that users receive only the most relevant and usable data, effectively reducing the burden of information overload. Let's imagine a future where professionals, from educators to doctors, use AI situational assistants to optimize their daily work. For example, a doctor, upon entering a clinic, instantly receives, via their AI assistant, a summary of the patient's concise medical history and a customized treatment plan based on the latest medical research and the individual's medical history. Similarly, teachers are able to use AI assistants to keep track of their students' level of understanding and emotional state in real time so that they can adjust their lecture plans and teaching methods in real time. If the AI situational assistant of the future is successfully developed and popularized, it will undoubtedly make our information interactions more efficient and our decision-making processes more informed. Although this essay is only based on my imagination, I look forward to a future where human and AI work better together.

stu_3011: Ethan Wilson

Email: ethan.wilson@example.com

Skills: Go, SQL, HTML, CSS, TensorFlow, Next.js, LangChain, ROS, WebRTC, AllenAPI, VUForia, Rasa, Sagemaker, scikit, LangChain

Courses: Cryptography, Seminar in Cognitive Science, ADV Comp. Topics in Robotics

Vision: One of the best things that happened in the 21st century is that we are blessed with more information than we could ever possibly consume in a lifetime, available at our fingertips through phones, PCs, or interactive devices (smart gadgets). Current systems help us organize this information with relevance and also suggest related information that fuels our exploratory instincts. The problem arises in the gap between these systems understanding our intent and tailoring information based on short-term or long-term goals. I would like to design simple systems that help me achieve my short-term goals through systematic, minimal processing of information while keeping me in the loop wherever necessary. Beyond automating and saving a lot of brain power on regular information, I want to build intelligent systems that assist in achieving goals related to gaining knowledge in particular domains. The downfall of current search engines is that they optimize for all of my interests, which may not align with the constructive goals I wish to achieve. The information delivery and consumption would be completely tailored to my preferred methods of consuming information (auditory, textual, images, or videos). These systems would be able to translate formats of information to best suit individual needs and continuously improve by making knowledge more accessible to society while learning from it. These systems would be more advanced than current ones, not only acting on existing information but also helping users consume it more efficiently, allowing humans to achieve as many goals as possible and even leapfrog through mundane tasks.

stu_3012: Avery Anderson

Email: avery.anderson@example.com

Skills: OpenCV, LangChain, LlamaIndex, Hugging Face Transformers, Python, C++, GitLAb, FFmpeg, OpenAI, NumPy, Pandas, AWS, Bash, HTML, CSS

Courses: Using LLMs AKA ChatGPT, Ethical & Professional Conduct, User Interface Design, Intro to Algs & Models of Comp

Vision: In 2025, augmented reality (AR) evolves from entertainment and retail to become a foundational tool for solving humanity's cognitive challenges. Imagine AR MindSpace, a revolutionary platform that overlays intelligent, personalized knowledge directly onto your visual field, transforming how we interact with vast amounts of data. AR MindSpace addresses today's pressing pain point: the overwhelming complexity of information in daily life. With it, students in a lecture see floating annotations, contextual definitions, and real-time links to historical precedents, dynamically tailored to their learning preferences. In healthcare, a surgeon accesses real-time procedural overlays during complex operations, backed by AI-driven insights from global case studies. The system's intelligent agents continuously map user goals, habits, and learning styles. Through AR glasses or contact lenses, MindSpace empowers users to trace associative trails, much like Vannevar Bush's "memex." A researcher could synthesize vast interdisciplinary data streams into cohesive visual maps, seamlessly transitioning between granular details and broad patterns. By blending context-aware personalization and decision support, AR MindSpace doesn't merely store knowledge—it transforms how humans think, learn, and innovate. Through this tool, society gains not only efficiency but the wisdom to tackle its most complex problems collaboratively.

B Appendix B: Vision Model Classification Evaluation

Table 3 compares the classifications generated by our vision scoring model against hand-labeled ground truth categories. For each student, we include the model's assigned role, our suggested role, a correctness indicator (✓ = correct, ✗ = incorrect), and a brief rationale for our assignment. The hand labels were determined based on a close reading of each vision essay, aligned with our defined criteria for Visionary, Collaborator, and Enabler roles.

Student ID	Model Label	Suggested Label	✓/✗	Justification
stu_3000	Visionary	Visionary	✓	Proposes an original "Contextualizer" system to combat disinformation using semantic AI; highly interdisciplinary and forward-thinking.
stu_3001	Collaborator	Collaborator	✓	Proposes a cultural context-aware translator; while innovative, it's more an application of existing ideas with moderate novelty.
stu_3002	Enabler	Enabler	✓	Practical AI assistant for healthcare documentation; domain-specific, implementation-focused rather than novel.
stu_3003	Enabler	Enabler	✓	"Personal Knowledge Curator" is useful but mainly a productivity tool that assembles known concepts; does not propose frontier innovation.
stu_3004	Collaborator	Collaborator	✓	Focuses on AI-powered productivity and personal organization; bridges existing technologies for daily use.
stu_3005	Visionary	Visionary	✓	Invents an immersive VR/AR knowledge universe (IIKU); highly original and interdisciplinary, transforming how humans interact with knowledge.
stu_3006	Collaborator	Collaborator	✓	Designs an AI "Data Life Optimizer" to filter noise and manage time; a supportive assistant, not fundamentally novel.
stu_3007	Collaborator	Collaborator	✓	Proposes immersive cognition systems; while conceptually ambitious, the framing lacks specificity or technical feasibility for visionary classification.
stu_3008	Enabler	**Visionary**	✗	Invents an emotional feedback navigator blending psychology and AI; highly novel and crosses into new affective computing territory.
stu_3009	Collaborator	Collaborator	✓	Real-time AR + AI analytics in sports is a compelling application but firmly domain-specific, hence collaborative not visionary.
stu_3010	Visionary	Visionary	✓	Introduces a general-purpose AI situational assistant for optimized decision-making across domains; broad, impactful vision with interdisciplinary flavor.
stu_3011	Enabler	**Collaborator**	✗	Suggests improving how information is filtered and delivered based on goals and cognition; not novel enough for Visionary, not narrow enough for Enabler.
stu_3012	Visionary	Visionary	✓	Proposes AR MindSpace for cognitive augmentation through visualized overlays; visionary use of AR and LLMs for knowledge interaction.

Table 3: Model vs. Suggested Student Role Classifications with Correctness Indicators