

AI in Grading: A Practical Prompt-Based Framework for Research Paper and Project Presentation Assessments

Shibbir Ahmed

Assistant Professor,

Department of Computer Science

AI IN TEACHING & LEARNING SYMPOSIUM 2025

March 26, 2025



Outline

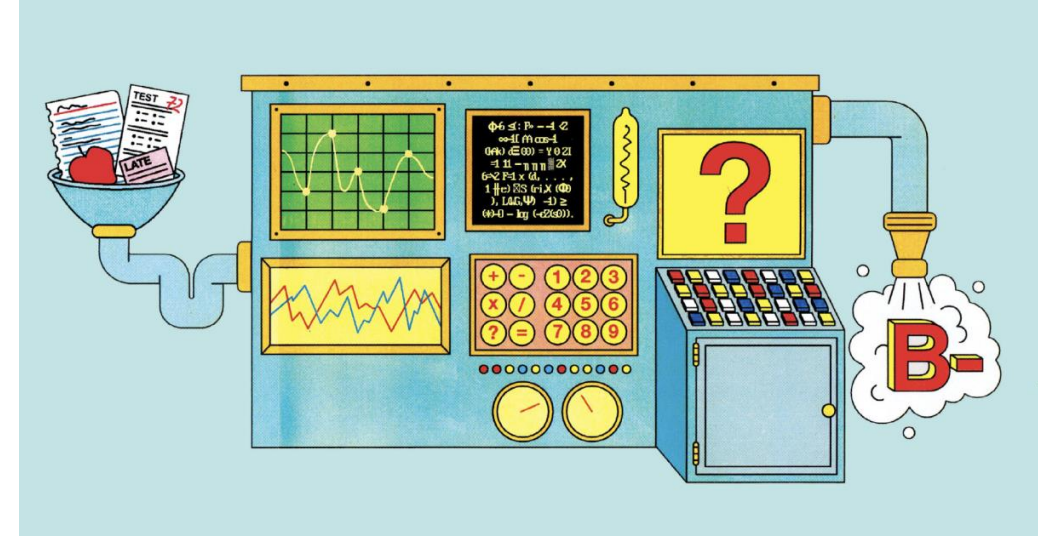
- Why AI in Grading?
- The Role of AI in Grading in Research Paper and Project Assessments
- The Prompt-Based Technique to Develop Rubrics
- How to practically use AI generated/refined Rubrics in Canvas for Grading
- Sample AI Prompts for Grading
- Benefits & Challenges Ahead
- Key Takeaways

Grading...



Challenges in Grading

- **Time-consuming** process
- **Inconsistent** evaluations
- **Subjectivity** in delivery & clarity
- **Detailed feedback** is difficult
- **Scalability** issues with large classes



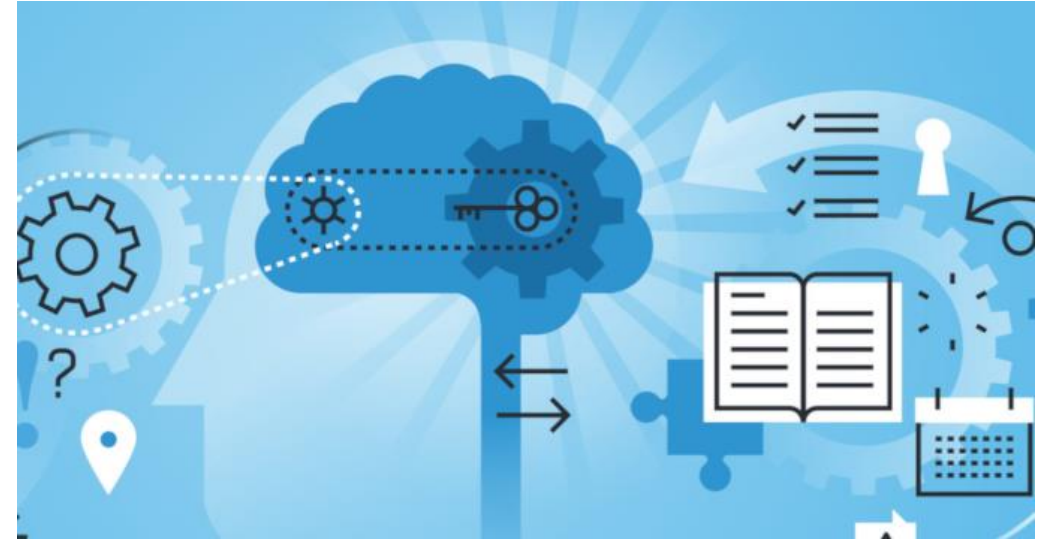
AI-Assisted Grading : Research Papers & Project Presentations

- **Use AI as a grading assistant**
- **Cross-check assessments** with AI feedback
- **Improve consistency** across evaluations
- **Enhance efficiency** in grading
- **Ensure fairness** in student evaluations



Benefits of AI-Assisted Grading

- **Saves time & effort**
- **Ensures fairness & consistency**
- **Provides structured feedback**
- **Improves student learning**
- **Enhances transparency**



AI-Supported Grading Rubrics

Software Quality Course: Research Paper Presentation Guidelines

Paper Selection:

Choose a research paper from **ICSE, FSE, or ASE** (top-tier Software Engineering conferences).

Presentation Format :

1. **Problem** – Define the research problem and its significance.
2. **Solution Approach** – Explain the proposed method or framework.
3. **Technical Contributions** – Highlight key contributions with examples.
4. **Demo/Reproducibility** – Show screenshots or evidence of reproducibility.
5. **Research Questions** – Outline the research questions addressed.
6. **Results Summary** – Present key findings and insights.
7. **Limitations** – Discuss the study's constraints.
8. **Future Direction** – Suggest potential improvements or extensions.

 **Note:** Each group member may be asked questions related to the paper and its reproducibility.



AI-Supported Grading Rubrics for Research Presentations



Research Paper Presentation Rubric						
Criteria	Ratings					Pts
Content (35 Points): i. Problem Statement (5 points)	<div>5 to >4 pts Excellent</div> <div>5 points: Clearly articulates the problem addressed in the paper, its significance, and its relevance.</div>	<div>4 to >2 pts Good</div> <div>3-4 points: States the problem but lacks clarity or depth in explaining its significance.</div>	<div>2 to >0 pts Needs Improvement</div> <div>1-2 points: Vague or incomplete description of the problem</div>	<div>0 pts Poor</div> <div>0 points (Poor): Fails to describe the problem.</div>	<div></div> <div>/ 5 pts</div>	
ii. Solution Approach (8 points)	<div>6 to >5 pts Excellent</div> <div>6 points (Excellent): Thorough explanation of the solution, including key methodologies or algorithms, with clarity and depth.</div>	<div>5 to >3 pts Good</div> <div>4-5 points (Good): Good explanation but missing minor details or lacks depth.</div>	<div>3 to >2 pts Needs Improvement</div> <div>3 points (Needs Improvement): Basic explanation with significant details missing.</div>	<div>2 to >0 pts Poor</div> <div>1-2 points (Poor): Minimal effort in explaining the solution.</div>	<div>0 pts Not Addressed</div> <div>0 points (Not Addressed): No explanation provided.</div>	<div></div> <div>/ 6 pts</div>
iii. Technical Contributions with Examples (6 points)	<div>6 to >5 pts Excellent</div> <div>6 points (Excellent): Effectively highlights the paper's contributions, supported by clear examples or evidence.</div>	<div>5 to >3 pts Good</div> <div>4-5 points (Good): Mentions contributions but with limited examples or evidence.</div>	<div>3 to >1 pts Needs Improvement</div> <div>2-3 points (Needs Improvement): Contributions are mentioned but lack explanation or examples.</div>	<div>1 to >0 pts Poor</div> <div>1 point (Poor): Minimal acknowledgment of contributions.</div>	<div>0 pts Not Addressed</div> <div>0 points (Not Addressed): No mention of contributions.</div>	<div></div> <div>/ 6 pts</div>
iv. Demo/Screenshots of Reproducibility (6 points)	<div>6 to >5 pts Excellent</div> <div>6 points (Excellent): Provides clear and accurate demos/screenshots with explanations, demonstrating reproducibility.</div>	<div>5 to >3 pts Good</div> <div>4-5 points (Good): Includes demos/screenshots but with minor gaps or unclear explanations.</div>	<div>3 to >1 pts Needs Improvement</div> <div>2-3 points (Needs Improvement): Screenshots/demos are vague or partially relevant.</div>	<div>1 to >0 pts Poor</div> <div>1 point (Poor): Screenshots/demos are minimal or lack relevance.</div>	<div>0 pts Poor</div> <div>1 point (Poor): Screenshots/demos are minimal or lack relevance.</div>	<div></div> <div>/ 6 pts</div>
v. Research Questions (4 points)	<div>4 to >3 pts Excellent</div> <div>4 points (Excellent): Clearly identifies and explains the key research questions.</div>	<div>3 to >2 pts Good</div> <div>3 points (Good): Mentions research questions, but explanations are incomplete or unclear.</div>	<div>2 to >0 pts Needs Improvement</div> <div>1-2 points (Needs Improvement): Minimal or vague mention of research questions.</div>	<div>0 pts Not Addressed</div> <div>0 points (Not Addressed): No research questions identified</div>	<div></div> <div>/ 4 pts</div>	
vi. Results Summary (4 points)	<div>4 to >3 pts Excellent</div> <div>4 points (Excellent): Concisely and accurately summarizes key results and implications.</div>	<div>3 to >2 pts Good</div> <div>3 points (Good): Provides a summary but lacks clarity or misses key implications.</div>	<div>2 to >0 pts Needs Improvement</div> <div>1-2 points (Needs Improvement): Minimal or unclear summary of results.</div>	<div>0 pts Not Addressed</div> <div>0 points (Not Addressed): No results summary provided.</div>	<div></div> <div>/ 4 pts</div>	
vii. Limitations (2 points)	<div>2 to >1 pts Excellent</div> <div>2 points (Excellent): Identifies and explains limitations comprehensively.</div>	<div>1 to >0 pts Good</div> <div>1 point (Good): Mentions limitations but explanation is minimal or unclear.</div>	<div>0 pts Not Addressed</div> <div>0 points (Not Addressed): Does not address limitations.</div>	<div></div> <div>/ 2 pts</div>		
viii. Future Direction (2 points)	<div>2 to >1 pts Excellent</div> <div>2 points (Excellent): Proposes thoughtful, well-reasoned future directions based on findings.</div>	<div>1 to >0 pts Good</div> <div>1 point (Good): Suggests future directions but lacks depth or alignment with findings.</div>	<div>0 pts Not Addressed</div> <div>0 points (Not Addressed): No future directions provided.</div>	<div></div> <div>/ 2 pts</div>		8

AI-Supported Grading Rubrics for Research Presentations



Delivery (15 Points): i. Presentation Based on Understanding and Answering Questions (7 points)	5 to >4 pts Excellent 5 points (Excellent): demonstrates an exceptional understanding of the material, confidently explains concepts in their own words with minimal reliance on slides, effectively engages with the audience, and answers questions perfectly.	4 to >3 pts Good 4 points (Good): shows a good grasp of the material, uses slides as a reference but explains most concepts independently, and engages moderately with the audience, answers questions moderately well.	3 to >2 pts Satisfactory 3 points (Satisfactory): has a moderate understanding, relies on slides for explanations, and shows limited engagement with the audience, providing somewhat satisfactory responses to questions.	2 to >1 pts Needs Improvement 2 points (Needs Improvement): heavily relies on reading slides, demonstrates limited understanding of the material, and lacks audience engagement, answers to the questions are not that much satisfactory.	1 to >0 pts Poor 0-1 points (Poor): reads directly from slides with minimal or no evidence of understanding or independent thought, answers to the questions are not at all satisfactory..	<input type="text"/> / 5 pts
ii. Presentation Quality (5 points)	5 to >4 pts Excellent 5 points (Excellent): Engaging and professional delivery, clear voice, good pacing, and appropriate body language.	4 to >3 pts Good 4 points (Good): Good delivery but with minor issues such as pacing, clarity, or engagement.	3 to >1 pts Needs Improvement 2-3 points (Needs Improvement): Adequate delivery but lacks engagement, clarity, or professionalism.	1 to >0 pts Poor 1 point (Poor): Minimal effort in delivery; difficult to follow.	0 pts Not Addressed 0 points (Not Addressed): Presentation quality significantly hampers understanding.	<input type="text"/> / 5 pts
iii. Slide Design (3 points)	5 to >4 pts Excellent 5 points (Excellent): Slides are visually appealing, well-organized, and complement the presentation. They effectively use visuals (e.g., charts, images) to enhance understanding, with minimal text and clear structure.	4 to >3 pts Good 4 points (Good): Slides are clear and organized with relevant visuals and text, but there is some room for improvement in design or structure.	3 to >2 pts Satisfactory 3 points (Satisfactory): Slides convey the key points but may be text-heavy, lack visuals, or have minor issues with organization or formatting.	2 to >1 pts Needs Improvement 2 points (Needs Improvement): Slides are cluttered, lack coherence, or fail to support the presentation effectively.	1 to >0 pts Poor 0-1 points (Poor): Slides are missing, disorganized, or significantly detract from the presentation quality.	<input type="text"/> / 5 pts

Grading using Canvas-Integrated Rubrics



Research Paper Presentation

Due: Dec 2, 2024 at 11:59pm - CS 5393.001

Criteria	Ratings					Pts
Content (35 Points): i. Problem Statement (5 points)	5 to >4 pts Excellent 5 points: Clearly articulates the problem addressed in the paper, its significance, and its relevance.	4 to >2 pts Good 3-4 points: States the problem but lacks clarity or depth in explaining its significance.	2 to >0 pts Needs Improvement 1-2 points: Vague or incomplete description of the problem	0 pts Poor 0 points (Poor): Fails to describe the problem.		<div>5 / 5 pts</div> <div></div>
ii. Solution Approach (8 points)	6 to >5 pts Excellent 6 points (Excellent): Thorough explanation of the solution, including key methodologies or algorithms, with clarity and depth.	5 to >3 pts Good 4-5 points (Good): Good explanation but missing minor details or lacks depth.	3 to >2 pts Needs Improvement 3 points (Needs Improvement): Basic explanation with significant details missing.	2 to >0 pts Poor 1-2 points (Poor): Minimal effort in explaining the solution.	0 pts Not Addressed 0 points (Not Addressed): No explanation provided.	<div>6 / 6 pts</div> <div></div>
iii. Technical Contributions with Examples (6 points)	6 to >5 pts Excellent 6 points (Excellent): Effectively highlights the paper's contributions, supported by clear examples or evidence.	5 to >3 pts Good 4-5 points (Good): Mentions contributions but with limited examples or evidence.	3 to >1 pts Needs Improvement 2-3 points (Needs Improvement): Contributions are mentioned but lack explanation or examples.	1 to >0 pts Poor 1 point (Poor): Minimal acknowledgment of contributions.	0 pts Not Addressed 0 points (Not Addressed): No mention of contributions.	<div>5 / 6 pts</div> <div></div>
iv. Demo/Screenshots of Reproducibility (6 points)	6 to >5 pts Excellent 6 points (Excellent): Provides clear and accurate demos/screenshots with explanations, demonstrating reproducibility.	5 to >3 pts Good 4-5 points (Good): Includes demos/screenshots but with minor gaps or unclear explanations.	3 to >1 pts Needs Improvement 2-3 points (Needs Improvement): Screenshots/demos are vague or partially relevant.	1 to >0 pts Poor 1 point (Poor): Screenshots/demos are minimal or lack relevance.	0 pts Poor 1 point (Poor): Screenshots/demos are minimal or lack relevance.	<div>6 / 6 pts</div> <div></div>
v. Research Questions (4 points)	4 to >3 pts Excellent 4 points (Excellent): Clearly identifies and explains the key research questions.	3 to >2 pts Good 3 points (Good): Mentions research questions, but explanations are incomplete or unclear.	2 to >0 pts Needs Improvement 1-2 points (Needs Improvement): Minimal or vague mention of research questions.	0 pts Not Addressed 0 points (Not Addressed): No research questions identified		<div>4 / 4 pts</div> <div></div>
vi. Results Summary (4 points)	4 to >3 pts Excellent 4 points (Excellent): Concisely and accurately summarizes key results and implications.	3 to >2 pts Good 3 points (Good): Provides a summary but lacks clarity or misses key implications.	2 to >0 pts Needs Improvement 1-2 points (Needs Improvement): Minimal or unclear summary of results.	0 pts Not Addressed 0 points (Not Addressed): No results summary provided.		<div>4 / 4 pts</div> <div></div>
vii. Limitations (2 points)	2 to >1 pts Excellent 2 points (Excellent): Identifies and explains limitations comprehensively.	1 to >0 pts Good 1 point (Good): Mentions limitations but explanation is minimal or unclear.	0 pts Not Addressed 0 points (Not Addressed): Does not address limitations.			<div>2 / 2 pts</div> <div></div>
viii. Future Direction (2 points)	2 to >1 pts Excellent 2 points (Excellent): Proposes thoughtful, well-reasoned future directions based on findings.	1 to >0 pts Good 1 point (Good): Suggests future directions but lacks depth or alignment with findings.	0 pts Not Addressed 0 points (Not Addressed): No future directions provided.			<div>2 / 2 pts</div> <div></div>

Grading using Canvas-Integrated Rubrics





Research Paper Presentation

Due: Dec 2, 2024 at 11:59pm - CS 5393.001

Delivery (15 Points): i. Presentation Based on Understanding and Answering Questions (7 points)	<div>5 to >4 pts Excellent</div> <div>5 points (Excellent): demonstrates an exceptional understanding of the material, confidently explains concepts in their own words with minimal reliance on slides, effectively engages with the audience, and answers questions perfectly.</div>	<div>4 to >3 pts Good</div> <div>4 points (Good): shows a good grasp of the material, uses slides as a reference but explains most concepts independently, and engages moderately with the audience, answers questions moderately well.</div>	<div>3 to >2 pts Satisfactory</div> <div>3 points (Satisfactory): has a moderate understanding, relies on slides for explanations, and shows limited engagement with the audience, providing somewhat satisfactory responses to questions.</div>	<div>2 to >1 pts Needs Improvement</div> <div>2 points (Needs Improvement): heavily relies on reading slides, demonstrates limited understanding of the material, and lacks audience engagement, answers to the questions are not that much satisfactory.</div>	<div>1 to >0 pts Poor</div> <div>0-1 points (Poor): reads directly from slides with minimal or no evidence of understanding or independent thought, answers to the questions are not at all satisfactory..</div>	<div>5 / 5 pts</div> <div></div>
ii. Presentation Quality (5 points)	<div>5 to >4 pts Excellent</div> <div>5 points (Excellent): Engaging and professional delivery, clear voice, good pacing, and appropriate body language.</div>	<div>4 to >3 pts Good</div> <div>4 points (Good): Good delivery but with minor issues such as pacing, clarity, or engagement.</div>	<div>3 to >1 pts Needs Improvement</div> <div>2-3 points (Needs Improvement): Adequate delivery but lacks engagement, clarity, or professionalism.</div>	<div>1 to >0 pts Poor</div> <div>1 point (Poor): Minimal effort in delivery; difficult to follow.</div>	<div>0 pts Not Addressed</div> <div>0 points (Not Addressed): Presentation quality significantly hampers understanding.</div>	<div>4 / 5 pts</div> <div></div>
iii. Slide Design (3 points)	<div>5 to >4 pts Excellent</div> <div>5 points (Excellent): Slides are visually appealing, well-organized, and complement the presentation. They effectively use visuals (e.g., charts, images) to enhance understanding, with minimal text and clear structure.</div>	<div>4 to >3 pts Good</div> <div>4 points (Good): Slides are clear and organized with relevant visuals and text, but there is some room for improvement in design or structure.</div>	<div>3 to >2 pts Satisfactory</div> <div>3 points (Satisfactory): Slides convey the key points but may be text-heavy, lack visuals, or have minor issues with organization or formatting.</div>	<div>2 to >1 pts Needs Improvement</div> <div>2 points (Needs Improvement): Slides are cluttered, lack coherence, or fail to support the presentation effectively.</div>	<div>1 to >0 pts Poor</div> <div>0-1 points (Poor): Slides are missing, disorganized, or significantly detract from the presentation quality.</div>	<div>3 / 5 pts</div> <div></div>
<div>The slides are good, but approaches could be enhanced using some examples. The Slide with Fig 1 was not properly scaled and was hard to read, and the evaluation setup slides contained too much text. research questions slides could be better using some visualization, Experimental evaluation, and experiment specification slides were good but the procedure slides have too many texts.</div>						
Total Points: 46 out of 50						



Effective AI Prompts for Grading & Feedback



Create a **detailed grading rubric for evaluating research paper presentations**. The rubric should assess both content and delivery, with a total of **50 points**. It should include specific criteria such as **problem statement, solution approach, technical contributions, research reproducibility, research questions, results summary, limitations, future direction**, presentation quality, and slide design. Each criterion should have a clear point distribution and descriptions for different **performance levels (Excellent, Good, Needs Improvement, and Poor)**. Ensure the **rubric is structured, easy to apply, and provides meaningful feedback**.



Effective AI Prompts for Grading & Feedback



◆ Problem Statement Evaluation Prompt

... "Define grading criteria for evaluating how well a research presentation explains the problem statement. Include different performance levels such as Excellent, Good, Needs Improvement, and Poor."

◆ Solution Approach Assessment Prompt

... "Describe a grading rubric for assessing the explanation of the solution approach in a research paper presentation. The rubric should include point distribution and descriptions of different levels of clarity and detail."

◆ Technical Contributions Prompt

... "Provide a structured evaluation rubric for grading how well the presenter highlights the technical contributions of the research. The rubric should differentiate between thorough explanations with examples and vague descriptions."

◆ Reproducibility (Demo/Screenshots) Prompt

... "Create grading criteria for assessing whether the presenter provides evidence of research reproducibility, such as demo results or screenshots. The rubric should reward clear demonstrations and penalize missing or vague explanations."

◆ Research Questions Prompt

... "Generate grading criteria for evaluating how well the presenter identifies and explains the research questions addressed in the paper. The rubric should differentiate between clear articulation and vague or missing questions."

◆ Results Summary Prompt

... "Create a grading rubric for evaluating how well the research results are summarized in a presentation. The rubric should distinguish between well-explained findings and unclear or missing results."



Effective AI Prompts for Grading & Feedback



◆ Limitations Prompt

💬 *"Provide grading criteria for assessing how well the presenter discusses the limitations of the research. The rubric should differentiate between a comprehensive explanation and missing or vague discussions."*

◆ Future Direction Prompt

💬 *"Develop a rubric for grading how well the presenter outlines potential future directions based on the research findings. The rubric should reward well-reasoned suggestions and deduct points for vague or missing directions."*

◆ Presentation Understanding & Delivery Prompt

💬 *"Create a grading rubric for evaluating the presenter's understanding of the research paper and delivery quality. Include criteria for confidence, ability to explain concepts, audience engagement, and reliance on slides."*

◆ Slide Design Prompt

💬 *"Generate a grading rubric for evaluating the design of presentation slides. Consider factors such as clarity, visual appeal, organization, and the effective use of visuals versus excessive text."*


💬 *"Combine all the above grading criteria into a cohesive research paper presentation rubric. Ensure that content and delivery are fairly weighted (e.g., 35 points for content, 15 points for delivery). Provide clear descriptions of performance levels (Excellent, Good, Needs Improvement, Poor) and ensure the total score sums to 50 points."*



Bonus Prompts for AI-Assisted Grading




If we want AI to **evaluate a presentation based on this rubric**, we can use:



Evaluate a research paper presentation using the following rubric: [Paste rubric here]. Provide a breakdown of points for each section and give constructive feedback.



OR



Based on this rubric, analyze the content, delivery, and slide design of a given presentation. Provide a detailed score and feedback for each criterion.

LLM as Judge while Grading



 SQ PPT.pptx
Presentation

You are a really good AI grading assistant. I took notes during presentation, now Evaluate a research paper presentation using the following rubric: [Paste rubric here]. Provide a breakdown of points for each section and give constructive feedback. "Data-Oriented Differential Testing of Object-Relational Mapping Systems"

Student 1

Presentation: The slides are good.

The problem is somewhat clear and the challenges were explained clearly.
The solution approach was presented in a diagram.
The methodology overview and results overview are presented well.

Question: Can explain the step in the approach based on the understanding, can explain the experimental setup as well.

Student 2

Slide: slides are fine

Presentation: read out from slides and notes, showed results using screenshots,
Reads from notes to explain results.

Question: Can answer the question related to approach vaguely, but can not answer basic questions like what is ORM "



LLM as Judge while Grading



Evaluation of Research Paper Presentation: "Data-Oriented Differential Testing of Object-Relational Mapping Systems"

Based on the provided rubric, here is a detailed breakdown of the evaluation for both students.

Final Comments:

- Student 1 had a **strong understanding** and effectively delivered the presentation with clear explanations.
- Student 2 should focus on **better preparation**, ensuring they understand the **fundamentals** and avoid excessive reliance on slides.



Student 1 Evaluation

Content (35 Points)

1. **Problem Statement (4/5)**
 - The problem was somewhat clear.
 - Challenges were explained well, which added clarity.
 - Could improve by explicitly stating the **significance** of the problem.
2. **Solution Approach (7/8)**
 - The approach was well presented with a **diagram**, which helped in understanding.
 - Minor improvements could include a **more detailed breakdown** of how the approach addresses the challenges.
3. **Technical Contributions with Examples (6/6)**
 - The **methodology overview and results** were well presented.
 - Contributions were clear and supported by examples.
4. **Demo/Screenshots of Reproducibility (6/6)**
 - The results were well demonstrated.
 - Screenshots helped in explaining the process effectively.
5. **Research Questions (4/4)**
 - Research questions were clearly stated.
6. **Results Summary (4/4)**
 - Summarized key findings accurately.
7. **Limitations (2/2)**
 - Limitations were addressed.
8. **Future Direction (4/4)**
 - Provided a well-thought-out future direction.

Delivery (15 Points)

9. **Presentation Based on Understanding (5/5)**
 - Answered questions confidently and explained the **experimental setup and steps** well.
 - Demonstrated **good understanding** of the approach.
10. **Presentation Quality (4/5)**
 - **Good delivery**, but could have been more engaging.
 - Voice and pacing were effective.
11. **Slide Design (4/5)**
 - Slides were **visually clear and well-structured**.
 - Some minor improvements could be made in organizing content more concisely.

Total Score for Student 1: 46/50

✔ **Strengths:** Clear technical explanation, well-structured presentation, confident delivery.

🔧 **Areas for Improvement:** More emphasis on the **significance** of the problem and slightly more engaging delivery.

Student 2 Evaluation

Content (35 Points)

1. **Problem Statement (3/5)**
 - The problem was introduced but lacked **depth** in explanation.
2. **Solution Approach (5/8)**
 - The solution was explained but mostly read from slides and notes.
 - Could have **provided more detail** beyond what was written.
3. **Technical Contributions with Examples (4/6)**
 - Contributions were mentioned but lacked **depth and strong examples**.
 - Screenshots were shown, which was helpful.
4. **Demo/Screenshots of Reproducibility (5/6)**
 - Results were presented using screenshots, which improved clarity.
 - More explanation would have made it stronger.
5. **Research Questions (3/4)**
 - Research questions were mentioned but **not fully explained**.
6. **Results Summary (3/4)**
 - Summarized key results but with **minimal elaboration**.
7. **Limitations (1/2)**
 - Limitations were briefly mentioned but **not well explained**.
8. **Future Direction (2/4)**
 - Future work was discussed but **lacked depth**.

Delivery (15 Points)

9. **Presentation Based on Understanding (2/5)**
 - Relied heavily on reading from slides and notes.
 - Answered **approach-related questions vaguely** but struggled with basic questions, such as "What is ORM?"
10. **Presentation Quality (2/5)**
 - **Minimal engagement**; mostly read aloud.
 - The flow of speech could be improved.
11. **Slide Design (4/5)**
 - Slides were fine and well-organized.

Total Score for Student 2: 32/50

✔ **Strengths:** Used screenshots effectively, slides were clear.

🔧 **Areas for Improvement:** Avoid reading directly from slides, improve understanding of **core concepts**, and engage more with the audience.



AI-Assisted Grading with GPT: Key Takeaways

- **Standardized rubrics for clarity**
- **Consistent & objective evaluations**
- **Detailed, structured feedback**
- **AI-assisted, human-driven grading**
- **Fair & efficient assessment with instructor-in-the-loop**



Challenges & Limitations of AI-Assisted Grading

Challenges:

- May miss deeper **conceptual gaps**
- Struggles with **engagement detection**
- Over-reliance on **slide reading**



Limitations:

- Lacks **context sensitivity**
- Some grading aspects need **human judgment**
- Requires **instructor oversight** for fairness

GUIDING

Using ChatGPT as a Grading Assistant

ChatGPT does not know your students.

- 1 It is unaware of IEPs or extenuating circumstances that would give reason to adapt a standard, expectation, or grade.

ChatGPT isn't perfect.

- 2 It's true regardless of the task that you are asking it to help with. Using ChatGPT to help grade is no different. You will need to review what it generates, edit, revise, and tweak.

You can't rely solely on ChatGPT for grading.

- 3 Can it be a tremendous help? Yes. But you should not take its assessments as final.

You are ultimately responsible.

- 4 You are responsible for the accuracy of the grading and the feedback. You will need to be able to defend grades and feedback you give students.

THOUGHTS







Thank you everyone for this opportunity to share how I use generative AI for grading research presentations. I'd like to hear your experiences as well, especially since my research focuses on understanding the reliability of LLMs: feel free to reach out!



Dr. Shibbir Ahmed

Assistant Professor — Computer Science
CMAL 311 F

✉ shibbir@txstate.edu

