

Challenge 1: AI-Powered Design Evaluations

Generative AI ▾ Vision Model ▾ Vector DB ▾

To evaluate an AI model for Challenge 1: AI-Powered Design Evaluator, the criteria should encompass the effectiveness, accuracy, user-friendliness, and scalability of the solution. **Submissions will be evaluated based on the AI model and output recommendations.** Below are proposed evaluation criteria for such a model:

1. Accuracy & Relevance of Feedback:

- Principle Adherence: Can the model effectively identify if the design adheres to standard UX/UI principles such as alignment, contrast, spacing, hierarchy, etc.?
- Specificity: Does the model provide specific feedback points rather than generic recommendations?
- False Positives/Negatives: How frequently does the model misinterpret or overlook design elements, leading to inaccurate feedback?

2. Feedback Comprehensiveness:

- Coverage: Does the model evaluate all key aspects of UX/UI design, such as layout, typography, color schemes, usability, and interactivity?
- Depth: Does the feedback delve into nuanced details, providing both macro (overall design structure) and micro (specific element design) perspectives?

3. User Experience & Usability:

- Feedback Presentation: Is the feedback presented in a user-friendly, clear, and digestible manner? Consider factors like visual aids, annotations on the design, and a clear feedback structure.
- Upload & Processing Speed: How quickly can users upload their PDFs, and how fast does the model evaluate and return results?
- Interactivity: Can users interact with the feedback, perhaps clicking on certain elements for more detailed explanations or resources?

4. Scalability & Versatility:

- File Handling: Can the model effectively handle different file sizes, complexities, and types (e.g., single-page vs. multi-page PDFs)?



AI Innovation Challenge '23

- Design Style Adaptability: Does the model recognize and adapt feedback based on different design styles (e.g., minimalist vs. detailed, corporate vs. casual)?
- Continuous Learning: Can the model learn from feedback and user interactions to improve its recommendations over time?

5. Innovation & Creativity:

- Unique Features: Does the solution introduce innovative features or mechanisms not commonly found in traditional design evaluation tools?
- Integration Possibilities: Can the model easily be integrated into existing design workflows, tools, or platforms?

6. Ethical & Bias Considerations:

- Cultural Sensitivity: Does the model consider diverse and inclusive design principles? Does it recognize potential cultural biases in designs and provide feedback accordingly?
- Transparency: Does the model explain its evaluation criteria to users, ensuring designers understand the basis for the feedback?

7. General Robustness:

- Error Handling: How well does the model handle unexpected or corrupted inputs? Does it provide meaningful error messages?
- Security: How does the model ensure the privacy and security of the uploaded designs?

Resource: Available Web Data

<https://unblast.com/figma/>

<https://www.uistore.design/types/figma/>

<https://www.figma.com/community/tag/free%20download/files>

<https://www.lapa.ninja/freebies/figma/>

Resource: UX Design Principles

<https://www.uxdesigninstitute.com/blog/ux-design-principles/>

<https://www.springboard.com/blog/design/ux-design-principles/>

Resource: AI

<https://cloud.google.com/vertex-ai?hl=en>

<https://streamlit.io/>



AI Innovation Challenge '23

Challenge 2: Paired-Programming with TaskMaster and Junior Bot

Generative AI ▾ AutoGen ▾ Fullstack ▾

Submission will be evaluated upon the following criteria:

Task Completion:

Description: The extent to which the task provided by the TaskMaster Agent is successfully completed.

Guidance: Participants should demonstrate their ability to accomplish the specified task, which may involve data exploration, preprocessing, and predictive modeling.

Code Quality and Functionality:

Description: The quality of the code written, including adherence to best practices and its functionality in solving the task.

Guidance: Code should be well-structured, readable, and free from syntax and logic errors. It should effectively implement the predictive model and produce accurate results.

Interaction with Junior Bot:

Description: The extent to which participants interact and collaborate with the Junior Bot during the coding process.

Guidance: Participants should engage in effective communication with the Junior Bot, discussing coding decisions, debugging efforts, and problem-solving strategies as if simulating a real pair programming experience.

Simulated Pair Programming Experience:

Description: How well the participant simulates a collaborative pair programming experience with the Junior Bot.

Guidance: Participants should engage in collaborative decision-making, sharing ideas, and leveraging the expertise of the Junior Bot while working together to achieve the task.

Utilization of Synthetic Data:

Description: The effective use of synthetic data generated by the TaskMaster Agent in the data exploration and modeling process.



AI Innovation Challenge '23

Guidance: Participants should demonstrate an understanding of how to work with synthetic data, including data preprocessing, feature engineering, and model training.

TaskMaster Agent Interaction:

Description: The extent to which participants engage with and respond to guidance and feedback provided by the TaskMaster Agent.

Guidance: Participants should interact with the TaskMaster Agent to seek clarification, discuss the task, and address any questions or challenges.

Effective Code Execution:

Description: The code's ability to run successfully and produce desired outputs.

Guidance: The code editor should be capable of executing the code, allowing participants to reach the goals defined in the task.

Innovation and Problem-Solving:

Description: The creativity and innovation displayed in addressing challenges within the coding task.

Guidance: Participants should explore innovative solutions, tackle complex issues, and demonstrate problem-solving skills beyond the basics.

Resource: AutoGen

<https://www.microsoft.com/en-us/research/blog/autogen-enabling-next-generation-large-language-model-applications/>
<https://huggingface.co/spaces/MultiTransformer/autogen-tutorials>

Resource: AI

<https://cloud.google.com/vertex-ai?hl=en>
<https://streamlit.io/>

Resource: Synthetic Data

<https://towardsdatascience.com/gans-vs-autoencoders-comparison-of-deep-generative-models-985cf15936ea>
<https://ieeexplore.ieee.org/document/8834544>

Resource: Chat UI & Code IDE Boilerplate

https://github.com/Vy-X-S/radical_x_boilerplate_ui