Applied Data Science Department

StarryStudioAI
Project Advisor: Dr. Simon Shim

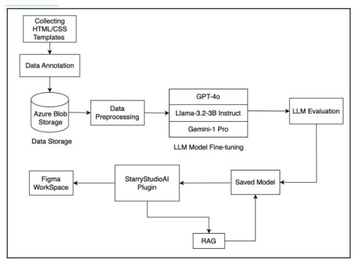
Gagrani Poojan,
Patel Bhavik,
Sheth Smeet,
Thakur Kashish,
Varahan Priya

Introduction

StarryStudioAI is a innovative Figma plugin that transforms natural language prompts into high-fidelity, interactive design templates. Leveraging fine-tuned Large Language Models (LLMs) like GPT-4o, Llama-3.2, and Gemini 1 Pro, enhanced with Retrieval-Augmented Generation (RAG), StarryStudioAI automates the generation of HTML/CSS code. These outputs are seamlessly integrated into Figma, enabling real-time visualization and refinement. Designed for users of all technical backgrounds, the tool accelerates workflows by reducing manual effort, ensuring design consistency, and bridging the gap between creative ideation and implementation. With StarryStudioAI, web development is more efficient, accessible, and collaborative, paving the way for the next generation of UI/UX innovation.

Methodology

System Architecture

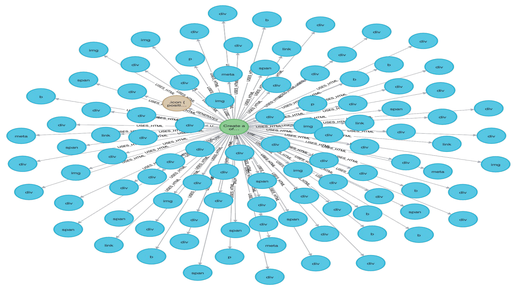


- **Data Collection:** Collection of HTML/CSS templates from publicly available Figma-compatible datasets and repositories.
- **Data Preprocessing:** Cleaning, formatting, and standardizing data to ensure compatibility with LLMs for training.
- **Data Storage:** Annotated templates are securely stored in Azure Blob Storage for easy retrieval and scalability.
- **LLM Fine-Tuning:** Fine-tuning GPT-4o, Llama-3.2-3B Instruct, and Gemini-1 Pro models on preprocessed HTML/CSS data.
- **RAG:** Dynamically retrieves relevant design patterns or code snippets to enhance context and accuracy.
- Model Deployment: Fine-tuned models are stored and integrated into the StarryStudioAI plugin for real-time inference.
- **StarryStudioAI Plugin:** Accepts natural language prompts and renders generated designs directly into the Figma Workspace.
- **LLM Evaluation:** Models are evaluated using metrics like BLEU, RUBY, and BERTScore to ensure usability and accuracy.

Analysis and Results

Knowledge Graph and RAG

The hierarchical knowledge graph from the project's dataset, mapping relationships between instructions, HTML components, and CSS styles. The central node represents the root instruction, branching into child nodes like div, img, span, and meta, illustrating their breakdown and interdependencies. The radial layout emphasizes the modular organization, reflecting principles of reusability and systematic web design.



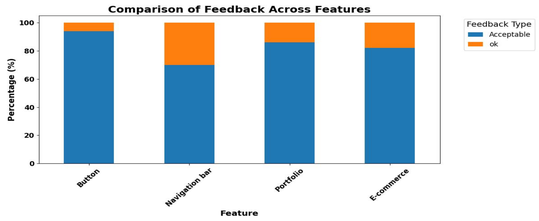
Models

- **GPT-4o:** A multimodal transformer optimized for generating semantically robust and reusable HTML/CSS code. Pre-trained on a vast corpus, it leverages instruction fine-tuning and excels in managing hierarchical structures, enhancing StarryStudioAI's design-to-code workflows.
- **Llama-3.2-3B Instruct:** An instruction-tuned Code Llama variant with 3 billion parameters, designed for structured code generation. It ensures semantic consistency and efficient resource utilization through mixed-precision training, making it ideal for lightweight integrations.
- **Gemini-1 Pro:** A multimodal transformer by Google DeepMind tailored for structured code generation. It employs a dual-stream architecture and robust attention mechanisms to align user prompts with design metadata, handling complex workflows efficiently.

Evaluation Results

Evaluation of Metrics/Models	Llama-3.2	GPT-4o	Gemini-1 Pro
BLEU	0.75	0.48	0.25
Ruby	0.40	0.59	0.30
ChrF	0.33	0.73	0.44
BertScoreF1	0.74	0.79	0.64
Pass@ 1	0.0	0.33	0.16
Pass@3	0.16	0.50	0.33

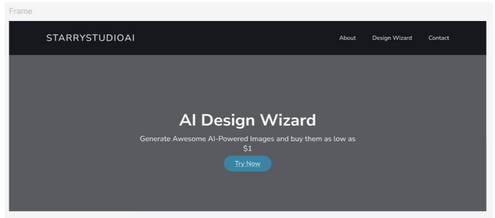
The experimental results demonstrate StarryStudioAI's credible performance in automating web design processes, with each model contributing unique strengths to address various aspects of UI/UX development. GPT-4o emerges as the frontrunner, showcasing remarkable proficiency in generating semantically accurate and structurally sound designs. Its character F-score of 0.73 and Pass@1 rate of 0.33 highlight its ability to produce high-fidelity, Interactive web designs on the first attempt. The model's BERTScore F1 of 0.79 further underscores its capacity to align generated code with user intent. LlaMa 3.2 excels in preserving structural fidelity, boasting a BLEU score of 0.75 and a BERTScore F1 of 0.74, although it faces challenges in generating fully functional designs on the first attempt. Gemini-1 Pro demonstrates balanced performance across metrics, showing promise in semantic understanding and efficiency.



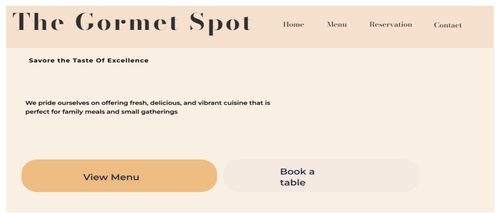
Human evaluation results validate StarryStudioAI's effectiveness in generating practical UI components, with Button designs and Portfolio templates achieving exceptional acceptance rates of 95% and 85% respectively. The survey of 50 participants, including both technical and non-technical users, demonstrates the system's ability to satisfy the user requirements by producing high-quality, production-ready designs that require minimal modifications.

StarryStudio AI Generated UI Designs

Create a responsive e-commerce header and hero component using HTML and CSS. The header should feature a vector brand logo for 'StarryStudioAI' and navigation links: About, Design Wizard, and Contact. The hero section should use an image background with a subtle gradient overlay, displaying the title 'AI Design Wizard,' the tagline 'Generate Awesome AI-Powered Images and buy them as low as $1,' and a 'Try Now' button. Focus on a clean, modern design with accessibility and responsiveness

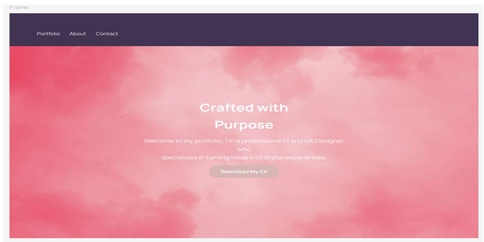


Create a responsive header and hero section for a restaurant website using HTML and CSS. The header should have a left-aligned logo ("The Gourmet Spot") and right-aligned navigation links ("Home," "Menu," "Reservations," "Contact") styled with hover effects.



Create a centered button labeled 'StarryStudioAI' with a gradient background that transitions on hover.

Create a responsive e-commerce header and hero component using HTML and CSS. The header should feature a vector brand logo for "StarryStudioAI" and navigation links: About, Design Wizard, and Contact. The hero section should use an image background with a subtle gradient overlay, displaying the title "AI Design Wizard," the tagline "Generate Awesome AI-Powered Images and buy them as low as $1," and a "Try Now" button. Focus on a clean, modern design with accessibility and responsiveness.



Summary/Conclusions

StarryStudioAI's performance metrics reveal a groundbreaking leap in AI-driven web design automation. With a character F-score of 0.73, a BERTScore F1 of 0.79, and a Pass@1 rate of 0.33 for the GPT-4o model, the system demonstrates remarkable precision in generating semantically robust and structurally accurate Figma designs. These metrics not only validate the system's technical prowess but also showcase its ability to transform natural language prompts into pixel-perfect design layouts with minimal manual intervention. This innovative approach streamlines the design-to-development process, empowering users to produce high-quality, interactive web designs with remarkable speed and precision.

Key References

[1] OpenAI, "GPT-4 Technical Report," arXiv preprint arXiv:2303.08774, Mar. 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2303.08774:contentReference[oaicite:0]{index=0}.

[2] B. Rozière, J. Gehring, F. Gloeckle, et al., "Code Llama: Open Foundation Models for Code," arXiv preprint arXiv:2308.12950, Aug. 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2308.12950:contentReference[oaicite:1]{index=1}.

[3] S. Feng et al., "Designing with Language: Wireframing UI Design Intent with Generative Large Language Models," arXiv preprint arXiv:2312.07755, Dec. 2023. [Online]. Available: https://arxiv.org/abs/2312.07755:contentReference[oaicite:2]{index=2}.

[4] Y. Gui et al., "VISION2UI: A Real-World Dataset with Layout for Code Generation from UI Designs," arXiv preprint arXiv:2404.06369, Apr. 2024. [Online]. Available: ahttps://arxiv.org/abs/2404.06369:contentReference[oaicite:3].

Acknowledgements

We thank the Applied Data Science Department at San Jose State University and Prof. Simon Shim for their guidance and mentorship throughout this project. We acknowledge the use of Figma-compatible datasets, Azure Blob Storage, OpenAI's GPT-4o, and Google's Gemini for their contributions. Special thanks to our collaborators and teammates for their dedication and efforts. We also appreciate the participants of our user evaluation studies for their valuable insights, which helped refine our approach.