Applied Data Science Department

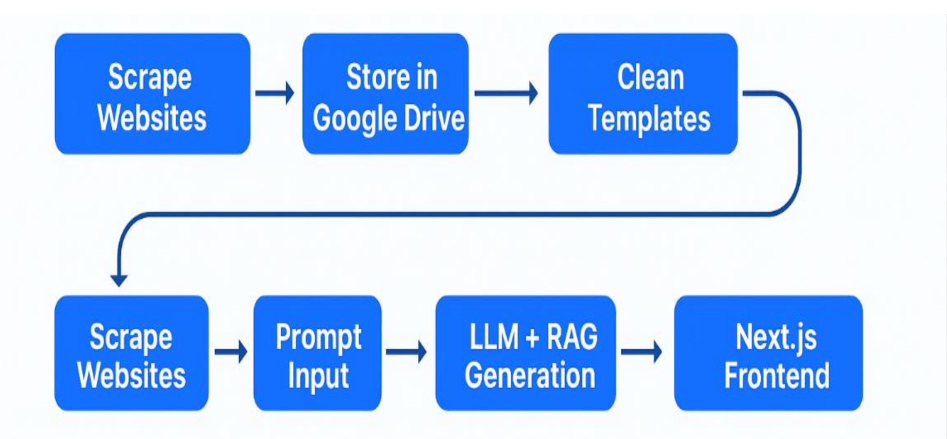
AI Powered Website Generator For B2B Saas Industries

Harsh, Shinde
Yogavarshni, Ramachandran
Shivram, Sriramulu
Pranavi, Avula
Ashwini Hiremath

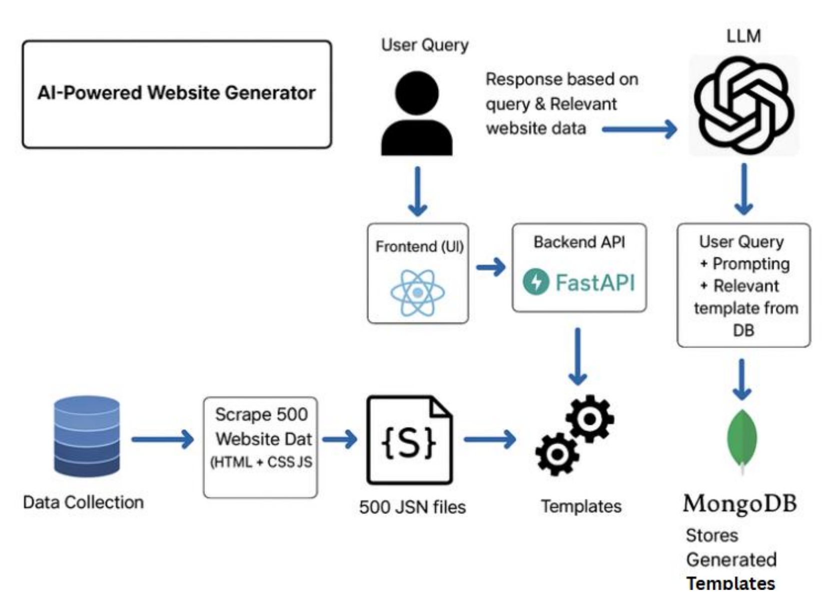
Project Advisor: Dr. Simon Shim

Introduction

Developing modern responsive websites is an ongoing struggle for very early stage B2B SaaS startups owing to metals for limited resources, materials or time. Classic WEB development implies the need for hands-on work done in design, content generating and coding – which usually increases the time taken to launch MVP. This project introduces an AI driven website generator that makes use of Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) to automate the process of website creation from a single user prompt. The solution provides made-to-order industry-specific SaaS landing pages within 30 seconds. As allowed in the process of the data pipeline, We did JavaScript-based automation to scrape through 500+ B2B SaaS websites. Google drive was used to store the raw HTML/CSS data. AWS was used to build an ETL pipeline extracting transforming and loading structured HTML/CSS components. Cleaned and re-usable templates were then passed back to the Google Drive organized by Industry and Layout type.



Methodology



This system features a React frontend that collects key inputs This There's a React frontend that allows us to collect the most important inputs from users, like for example, company name, industry domain, and website features wanted. These inputs are fed into a FastAPI back end that pulls and correlates pre-extracted HTML templates scraped from 500+ B2B SaaS websites. The backend then builds structured prompts from these inputs and sends these to fine-tuned LLMs – including Claude, Gemini and GPT, to produce tailored content and structure. The created websites are dynamically displayed through injection of this content into applicable templates. The finished templates are then saved in MongoDB for version control and user history to be useful in the future. This workflow makes it easy for users to get instantly deployable, industry-ready websites without even needing to code.

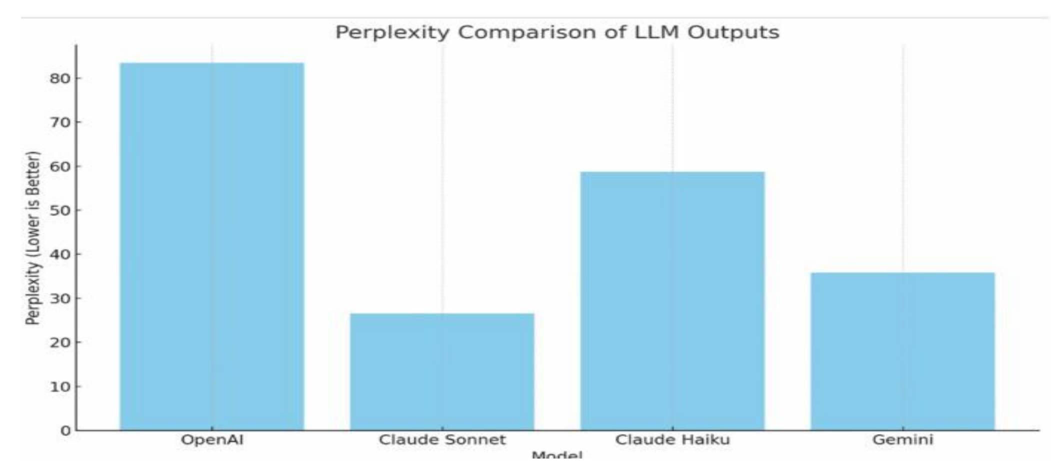
Analysis and Results

Perplexity Score

To evaluate the predictive confidence of various Large Language Models (LLMs), we used **perplexity**, a standard NLP metric. It measures how "surprised" a model is by the actual output. Lower perplexity indicates the model is better at predicting the next word in a sequence.

Comparative Results

Claude Sonnet achieved the lowest perplexity (26.59), indicating the most confident predictions and highest language fluency.
Gemini followed with a solid performance (35.87), outperforming Claude Haiku and OpenAI GPT.
Claude Haiku showed moderate perplexity (58.70), suggesting balanced performance.
OpenAI GPT (likely GPT-3.5) showed the highest perplexity (83.44), reflecting more uncertainty in predictions.



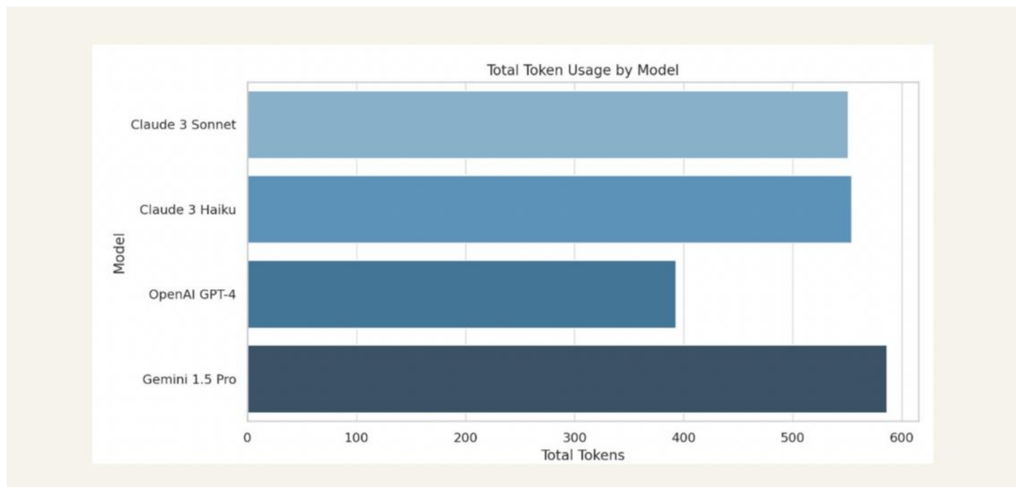
Rank	Model	Perplexity
1	Claude Sonnet	26.59
2	Gemini	35.87
3	Claude Haiku	58.70
4	OpenAI GPT	83.44

Token Usage Analysis: Conciseness vs. Verbosity

Gemini 1.5 Pro used the most tokens (~590), suggesting a more verbose or detailed style of response.
Claude 3 Haiku and Claude 3 Sonnet used a similar number of tokens (~550), indicating a balanced but slightly wordy output.
OpenAI GPT-4 used the fewest tokens (~390), making it the most concise, possibly due to optimized encoding or more compact phrasing.

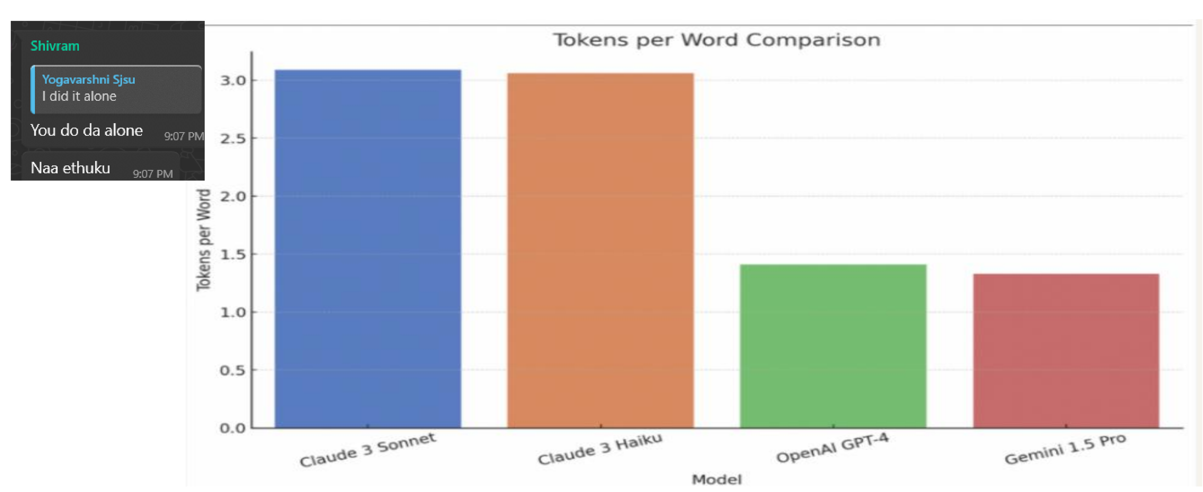
Interpretation

High perplexity may signal weak generalization, outdated pretraining data, or limitations in contextual understanding.
Low perplexity suggests more robust language modeling, better understanding of context, and less reliance on "guesswork."



Output Word Count Analysis

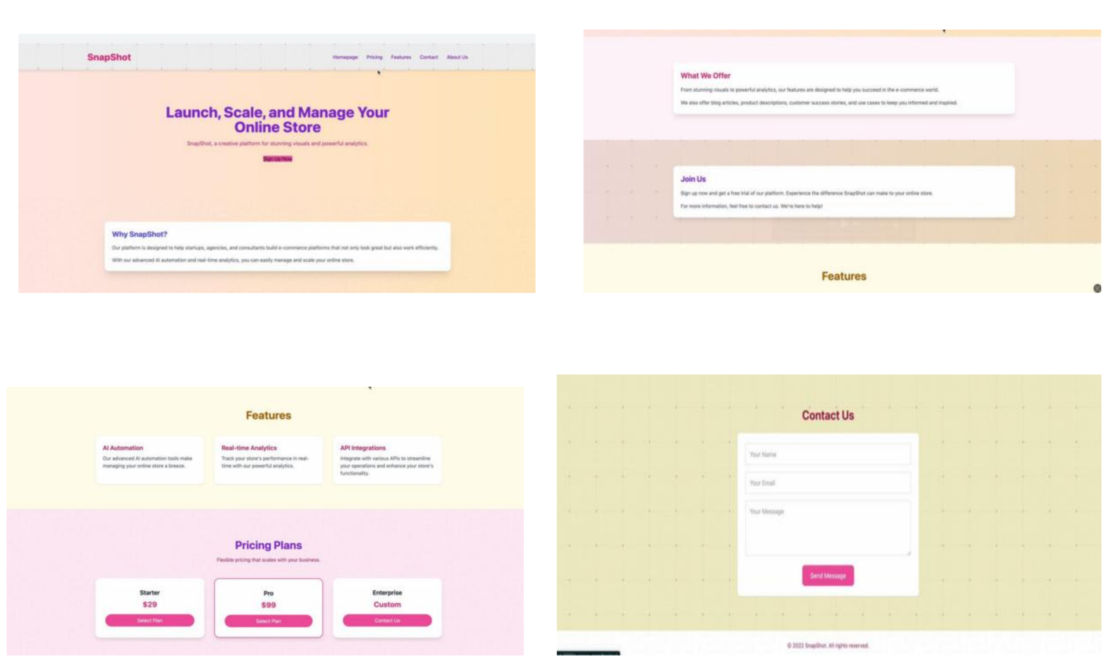
Gemini 1.5 Pro has the highest word count (~430), reaffirming its verbose style. Ideal for generating elaborate or narrative-rich content.
OpenAI GPT-4 ranks second (~270 words), indicating it balances detail with conciseness.
Claude 3 Sonnet and Claude 3 Haiku both produced fewer words (~180–190), leaning toward more compact responses.

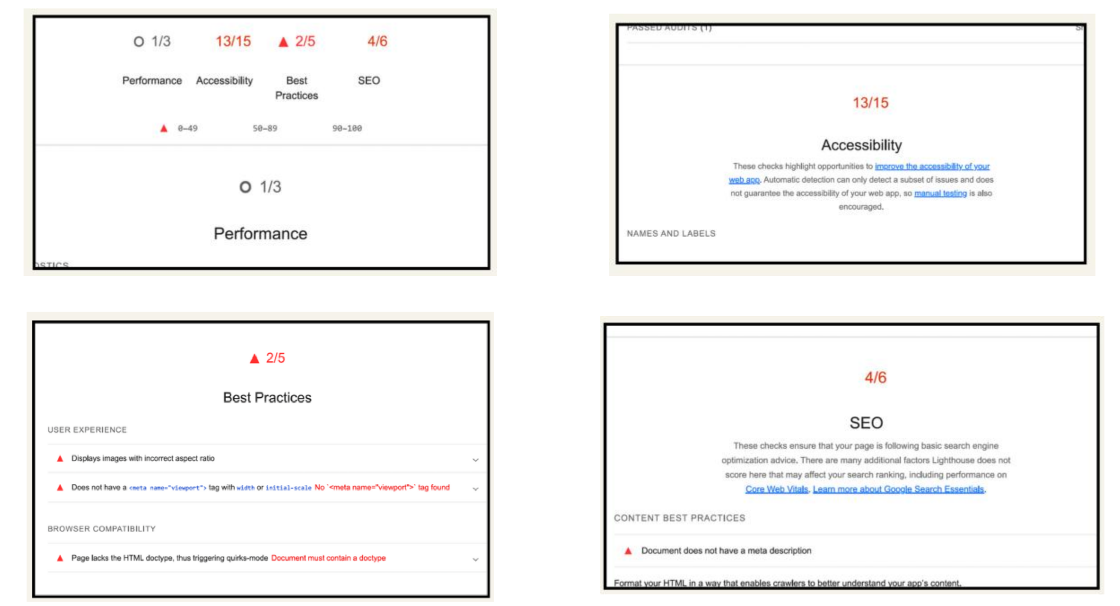


Output Word Count Analysis

Claude 3 Sonnet and Haiku use over 3 tokens per word, making them less efficient and potentially more expensive if pricing is based on token count.
OpenAI GPT-4 and Gemini 1.5 Pro both achieve leaner language, using just ~1.4 tokens per word on average..

Results: Sample result using GPT-4 model





Lighthouse Report

Performance: Low score (1/3) due to poor load optimization.
Accessibility: 13/15 – Good, but manual testing still recommended.
Best Practices: Missing viewport tag and doctype; image ratios incorrect.
SEO: Meta description missing; basic SEO needs improvement.

Summary/Conclusions

GPT-4 delivered the best content quality, ensuring accurate structure, natural tone, and relevant messaging. The generated websites maintain a polished look and feel with correct **placeholders** and effective **call-to-action (CTA)** sections. This tool simplifies website creation for businesses.
Future scope includes enhancing SEO, optimizing for mobile and performance, and **automatically generating and embedding relevant images** for each section.

Key References

[1] Roth, R. E. (2017, April). User interface and user experience (ui/ux) design. Geographic Information Science & Technology Body of Knowledge, 2017(Q2). doi: 10.22224/gistbok/2017.2.5
[2] Han, Y., Liu, C., & Wang, P. (2024). A comprehensive survey on vector database: Storage and retrieval technique, challenge. Preprint or Journal Name (update this).
[3] Anunphop, P., & Chongstitvatana, P. (2022). Web components template generation from web screenshot. In Proceedings of the [conference name]. Thailand.
[4] Ouyang, L., Wu, J., Jiang, X., et al. (2022). Aligning language models to follow instructions. arXiv preprint arXiv:2203.02155.

Acknowledgements

We would like to express our sincere gratitude to **Professor Dr. Simon Shim** for his invaluable guidance and insights throughout this project. His deep expertise in the B2B SaaS industry and advanced knowledge of Large Language Models (LLMs) provided critical direction, particularly in identifying emerging trends and evaluating the competitive landscape of AI-powered website generation. **Dr. Shim's** encouragement to explore innovative approaches and consider real-world applications greatly enriched the quality and relevance of this work.