INDIUM

LLM testing of a leading social media engagement platform

Social media!

Our client is a social media platform that stands out as one of the early adopters of Generative AI. They used LLM models which can help customers streamline social media static post creation and enhance the quality and reach of user-generated content by providing a communication channel and featuring tools typical of social networking sites, such as posts, comments, and live streaming.

Refining the art of social media engagement

The client identified several key areas for improvement within their content generation tool:

1 **Boosting engagement:** Their primary goal was to validate the effectiveness of the LLM model in assisting users with crafting posts that resonate with their audience.

2 **Accuracy & consistency:** Maintaining accuracy and consistency across AI-powered features was crucial. This included ensuring flawless grammar correction, relevant hashtag suggestions, and text enhancements that didn't alter the user's intended meaning.

3 **Seamless integration & performance:** Integration with the existing platform and achieving optimal response times for the AI features were critical for a smooth user experience. timely product approvals.

The client aimed to solidify their position as a platform that fosters engaging and effective social media interactions by addressing these challenges.

Indium delivers: A multi-faceted approach to LLM validation

We at Indium understood the importance of crafting a comprehensive solution addressing each client's challenges. Here's how we tackled it:

Model output validation: We employed rigorous LLM testing methodologies to ensure the generated content was accurate, consistent, and reliable. This involved evaluating the AI's ability to handle:

▶ **Automated content acquisition:** We developed an AI automation solution ("AI box") specifically designed for content-specific automation to streamline the training data collection process. This eliminated manual effort and ensured a constant flow of high-quality training data for the LLM model.

▶ **Our in-built LLM model (custom AI tools) for enhanced content generation:** We went beyond basic validation and we leveraged our inbuilt LLM model that can accelerate the testing activities tailored to the client's specific needs. These custom tools facilitated:

Prompt engineering:	User persona development:	Iterative refinement:	Improved content quality:
We created a comprehensive set of questions categorized by user intent and audience demographics. This diverse question bank helped the AI learn to interpret prompts effectively, considering factors like tone, purpose, and target audience.	We collaborated with the client to develop user personas representing their key audience segments. This allowed the LLM model to tailor its content generation based on different user groups' specific communication styles and preferences.	Through a process of testing and refinement, we helped the AI learn from various prompts, including equivalent questions, tonal variations, and rephrased versions. This continuous learning loop ensured the AI's ability to adapt to diverse user styles and effectively translate their intent into engaging content.	Our in-built LLM models further refined the AI's ability to enhance text, resulting in content that resonated better with target audiences.

- **Multiple prompts & input types:** The tool was tested with diverse prompts and content formats to guarantee its adaptability across various use cases.

- **Word limitations:** We ensured the AI adhered to character limitations for different content types (e.g., post captions and hashtags) to avoid truncation issues.

- **Grammar corrections & error detection:** We meticulously evaluated the AI's grammar correction capabilities and its ability to identify typos and misspellings for flawless writing.

- **Offensive content filtering:** To maintain a positive and inclusive platform environment, we validated the AI's effectiveness in filtering out offensive language and inappropriate content.

Model integration validation: Seamless integration with the existing platform was paramount. Our testing ensured smooth interaction and compatibility of the AI feature with the client's system, guaranteeing a user-friendly experience.

▶ **Contextual Functionality:** We understood the importance of the user interface (UI) in facilitating a positive user experience. Our validation encompassed:

Functionality testing: We meticulously tested the integration (with the existing system) to ensure all features, such as threading functionality and feedback collection mechanisms, worked flawlessly. This ensured users could interact seamlessly with the AI chatbot and provide valuable feedback.

UI validation: The UI was rigorously tested for clarity, ease of use, and visual appeal. This ensured users could intuitively navigate the AI tool and understand its functionalities.

▶ **Error-free experience:** We prioritized minimizing errors to maintain user trust and satisfaction. This involved:

Error validation: We conducted comprehensive testing to identify and eliminate potential errors within the AI system. This included testing for unexpected responses, crashes, or malfunctions.

▶ **Multilingual support:** We validated the AI's ability to handle content creation and user interaction in multiple languages for a broader user base. This testing ensured the AI functioned effectively across diverse language groups.

Performance evaluation: We conducted thorough LLM performance testing to assess the AI feature's speed, responsiveness, and efficiency. This ensured optimal real-world usage by evaluating metrics such as:

Consistency: A crucial aspect of user trust is consistent behavior from the AI tool. We achieved this by measuring the AI's consistency in generating similar outputs for similar inputs. This involved iterating the validation process multiple times based on accuracy and consistency metrics. By ensuring predictable behavior, we guaranteed the AI would produce the same high-quality answer for similar user prompts, fostering user confidence and trust.

Accuracy: The accuracy of the AI's text enhancements, grammar corrections, and hashtag suggestions was rigorously evaluated.

Error Rate: We minimized the error rate by identifying and addressing any instances of incorrect outputs or malfunctions.

Response Rate: Rapid response times were critical for a seamless user experience. We optimized the AI for fast and efficient content generation.

We empowered the client to confidently launch their AI-powered content generation tool by implementing these comprehensive validation and LLM testing procedures.

The power of Generative AI:
A win-win for users and the platform

The client's vision of empowering users with an AI-powered content creation tool translated into a resounding success story. Here are the key results:

Client benefits:

▶ **Rapid adoption:** Indium's validation efforts fostered user confidence, leading to a remarkable 20%+ increase in the platform's user base.

▶ **Content explosion:** Users embraced the AI tool, resulting in a 50% surge in content generation. This enriched the platform with fresh and engaging content.

Enhanced user experience:

▶ **Lightning Speed:** Our performance optimization ensured a sub-30-second response time for content generation, keeping users engaged and productive.

▶ **Quality Boost:** The AI's accuracy and effectiveness in enhancing text resulted in a 10x increase in perceived content quality by users.

Indium's multi-faceted validation approach ensured a flawless AI feature launch and empowered the client to unlock significant business growth and enhance user experience.