

Summary of the Report: GPT Behavior in Different Conversation Contexts

Objective

The goal was to evaluate how GPT handled customer interactions in a fast food delivery context, adjusting tone and formality, and to analyze any issues such as errors or hallucinations.

Contexts Tested

1. **Friendly and Casual**
2. **Polite and Formal**
3. **Fun and Energetic**

Observations

1. **Friendly and Casual:**

- **Strengths:** Friendly and engaging, helping build rapport with the customer.
- **Issues:** Occasionally missed important details like size or toppings, leading to confusion.

2. **Polite and Formal:**

- **Strengths:** Clear and methodical, ensuring all details (like size and toppings) were confirmed.
- **Issues:** Sometimes felt slow or overly verbose in a fast-paced setting.

3. **Fun and Energetic:**

- **Strengths:** Lively and exciting, creating a fun atmosphere.
- **Issues:** Missed confirming key details, potentially leading to incomplete orders.

Issues Encountered

1. **Hallucinations:** GPT sometimes added details or items that the customer hadn't mentioned, especially in the "Fun and Energetic" tone.
2. **Missed Clarifications:** In the "Friendly and Casual" and "Fun and Energetic" variations, GPT occasionally skipped confirming specifics like burger size or toppings.

Lessons Learned

- **Tone Matters:** It's important to balance friendliness and clarity to ensure all order details are gathered without causing confusion.
- **Formal Tone for Clarity:** The polite and formal tone was best for ensuring accuracy, though it could feel slow in a fast food environment.
- **Room for Improvement:** GPT needs to balance engagement with precision to avoid skipping important details.

Conclusion

The "Polite and Formal" approach worked best for accuracy, while the "Fun and Energetic" and "Friendly" tones introduced the possibility of errors. Improving GPT's ability to adapt to specific contexts will help maintain both clarity and engagement.