Reporting of GlucoX Data Science Challenge at Castor

Ever Augusto Torres Silva

Executive Summary: Analysis of Stimulation Effectiveness and Notification Fatigue.

1. Objective:

To evaluate the effectiveness of "gentle reminders" and "urgent alerts" via push notifications to determine if the patient presents "nudge fatigue," meaning a drop in the number of measurements after several repeated notifications.

2. Proposed Strategy:

The CRISP-DM framework was partially proposed to solve Data Science problems. The following actions were taken:

1. **Problem Understanding:** Review of the project manager's request and understanding the problem to be solved.
2. **Data Understanding:** Exploratory data analysis through notebooks, describing and understanding the structures of the original data.
3. **Data Preparation:** Necessary actions were performed to enable raw data into actionable data to work without issues.
4. **Modeling:** The analysis involved identifying appropriate statistical tests to address the core question. Specifically, the Wilson test was selected to determine the error margin, and the chi-squared test was used to assess significant differences between rates.
5. **Evaluation:** Partial results were obtained, and iteration was performed to achieve more compelling visualizations to explain the problem.
6. **Deployment:** A .py script was created that deploys the entire preprocessing pipeline and the construction of evidence for the analysis.

3. Key Findings:

1. **Notification fatigue is evident:** The data confirms the product manager's hypothesis. The probability that a patient responds rapidly decreases once more notifications are sent.
2. **The global response rate is estimated at 60.8% for the first notification,** however,

significant differences in registration begin to appear, even by notification 3. Nevertheless, from notification 16 onward, it is evident that there are significant differences and rates drop to as low as 24%. (See chart 1)

3. **Urgency messages do not prevent notification fatigue:** Both types of notifications, "Friendly" and "Urgent," suffer from similar fatigue trajectories. While alerts contain a direct clinical message, patients stop reporting measurements similarly to "friendly" notifications when there are several notifications. (See chart 1, 2, 3)
4. **The risk segment is the most notable differentiating factor:** High-risk patients show the highest initial and subsequent response rates for both types of stimuli. For example, with the first urgent alert, high-risk patients can reach up to almost a 90% response rate. If age segments are added to this, younger patients have even higher response rates. (See chart 3)

4. Recommended Actions:

1. **Implement a frequency limit:** Establish a strict limit on the number of notifications sent within a specific period. While a drastic difference is seen at notification 16, there are also significant differences at 3, 7, and 9. I would recommend setting a limit around notification 5.
2. **Adopt a risk-stratified reminder strategy:**
    - **High Risk:** Maintain constant and structured notifications, with a wider notification limit.
    - **Low and Medium Risk:** Reduce the frequency of automatic reminders and consider other alternative strategies different from push notifications.
3. **Diversify interventions:** Given that urgent and friendly push notifications behave similarly, merely varying the text does not seem sufficient. Consider other types of strategies using other available data to measure their effects.

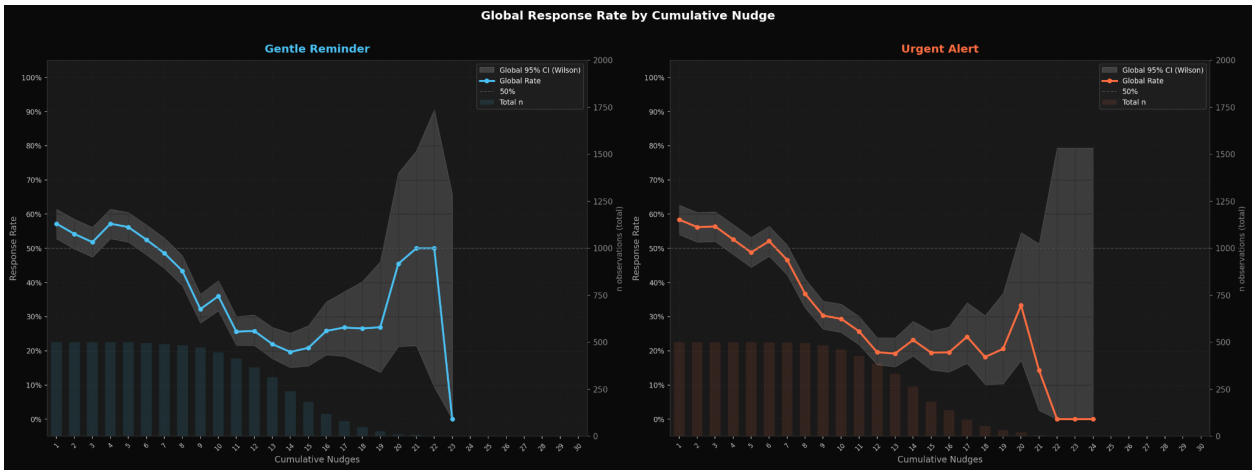

Chart 1: Global response rate by cumulative nudges

This visualization illustrates the impact of repeated push notifications on patient

measurement reporting behavior, confirming a significant trend of "notification fatigue." The global response rate, initially high, decays notably after the third notification, with a dramatic decrease observed from notification 16 onwards. This pattern is consistent across both "Friendly Reminders" and "Urgent Alerts," suggesting that message urgency alone does not mitigate fatigue.

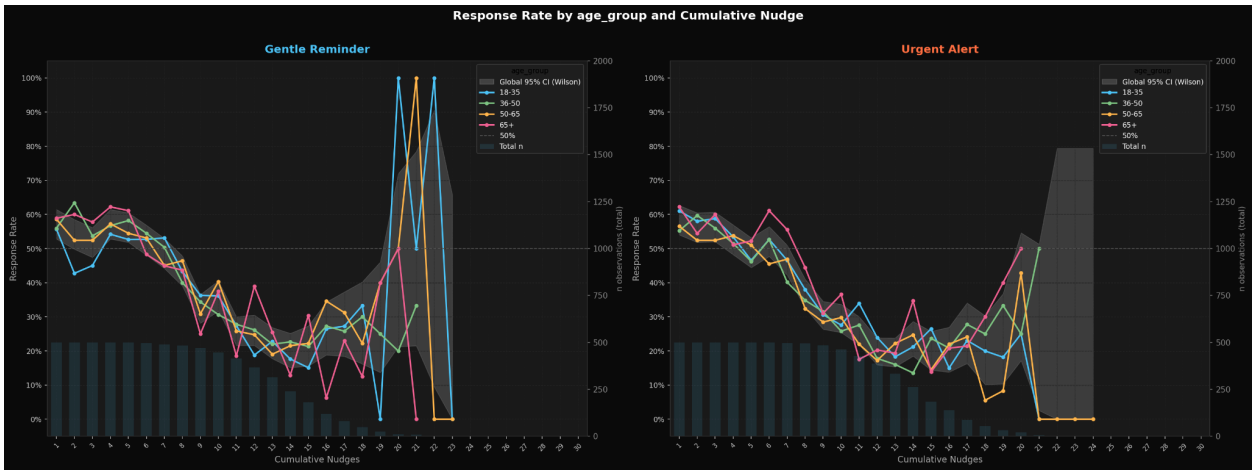

Chart 2: Response Rate Decay by Cumulative Notification Count (Gentle vs. Urgent, Stratified by Age Group)

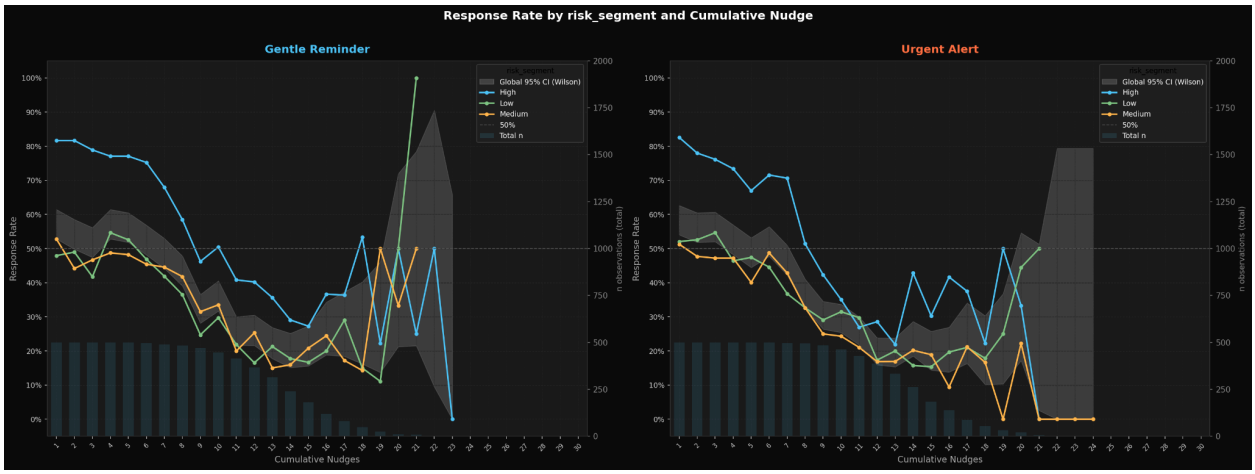

Chart 3: Response Rate Decay by Cumulative Notification Count (Gentle vs. Urgent, Stratified by Risk Segment)

4. AI Usage:

I utilized Artificial Intelligence throughout the code construction and reporting process, specifically Gemini.

I used AI to establish a plan for approaching the problem. For this, I performed a basic data preprocessing to visualize the structures of how the data was formed.

With the defined plan, I started developing it independently. Once I encountered coding issues, I relied on Code Assist with Gemini to resolve them.

Similarly, the generation of the graphs was an iterative process to achieve a single, informative graph that clearly displayed the results. This specific code originated from the AI, and I validated it.

Everything was executed in a notebook, and I performed the implementation in the `.py` file. I generated Docstrings and variable typing with the help of the AI.

Finally, within the code, I requested a grammatical correction and translation of the written texts (my native language is Spanish) to make it more presentable.

Lastly, for the presentation of this report, I wrote it in Spanish and requested a translation into English for better understanding, necessary formatting, and corrections.