Пишу программу по генерации рекламного SMS с помощью LLM.

Есть описание предложения: «Необходимо предложить клиенту оформить дебетовую премиальную бизнес-карту Mastercard Preffered. Обслуживание карты стоит 700 рублей в месяц, но клиент может пользоваться ей бесплатно. Что необходимо сделать, чтобы воспользоваться предложением – например:

1. Оформить премиальную бизнес-карту в офисе банка или онлайн в интернет-банке СберБизнес.

2. Забрать карту.

3. В течение календарного месяца совершить по ней покупки на сумму от 100 000 рублей.

4. В течение следующего месяца пользоваться ей бесплатно.»

Есть преимущества – например: «Предложение по бесплатному обслуживанию – бессрочное.

Оплата покупок без отчётов и платёжных поручений.

Платёжные документы без комиссии.

Лимиты на расходы сотрудников.

Мгновенные переводы на карты любых банков.»

Есть ключевое сообщение (т.е. текст, который обязан остаться в финальной версии сообщения) – например: «Бесплатное обслуживание при покупках от 100 000 рублей в месяц.»

Есть подход – например: «Призыв к действию».

Есть промпт для создания неперсонализированного сообщения (частично на основе данных выше):

«Сгенерируй смс-сообщение для клиента. Начни сообщение с призыва к действию с продуктом.

Описание предложения: Необходимо предложить клиенту оформить дебетовую премиальную бизнес-карту Mastercard Preffered. Обслуживание карты стоит 700 рублей в месяц, но клиент может пользоваться ей бесплатно. Что необходимо сделать, чтобы воспользоваться предложением:

1. Оформить премиальную бизнес-карту в офисе банка или онлайн в интернет-банке СберБизнес.

2. Забрать карту.

3. В течение календарного месяца совершить по ней покупки на сумму от 100 000 рублей.

4. В течение следующего месяца пользоваться ей бесплатно.

Преимущества: Предложение по бесплатному обслуживанию — бессрочное.

Оплата покупок без отчётов и платёжных поручений.

Платёжные документы без комиссии.

Лимиты на расходы сотрудников.

Мгновенные переводы на карты любых банков.

В тексте смс запрещено использование:

- Запрещенные слова: № один, номер один, № 1, вкусный, дешёвый, продукт, спам, доступный, банкротство, долги, займ, срочно, сейчас, лучший, главный, номер 1, гарантия, успех, лидер;

- Обращение к клиенту;

- Приветствие клиента;

- Обещания и гарантии;

- Использовать составные конструкции из двух глаголов;

- Причастия и причастные обороты;

- Деепричастия и деепричастные обороты;

- Превосходная степень прилагательных;

- Страдательный залог;

- Порядковые числительные от 10 прописью;

- Цепочки с придаточными предложениями;

- Разделительные повторяющиеся союзы;

- Вводные конструкции;

- Усилители;

- Паразиты времени;

- Несколько существительных подряд, в том числе отглагольных;

- Производные предлоги;

- Сложные предложения, в которых нет связи между частями;

- Сложноподчинённые предложения;

- Даты прописью;

- Близкие по смыслу однородные члены предложения;

- Шокирующие, экстравагантные, кликбейтные фразы;

- Абстрактные заявления без поддержки фактами и отсутствие доказательства пользы для клиента;

- Гарантирующие фразы;

- Узкоспециализированные термины;

- Фразы, способные создать двойственное ощущение, обидеть;

- Речевые клише, рекламные штампы, канцеляризмы;

Убедись, что в готовом тексте до 250 знаков с пробелами.

Убедись, что готовый текст начинается с призыва к действию с продуктом.

Убедись, что в готовом тексте есть следующая ключевая информация: Бесплатное обслуживание при покупках от 100 000 рублей в месяц.»

По данному промпту получаем неперсонализированное SMS-сообщение – например: «Хотите получить бесплатное обслуживание вашей новой бизнес-карты Mastercard Preffered? Просто сделайте покупки на сумму от 100 000 рублей в месяц! Никаких отчетов и платежных поручений, только удобство и выгода.»

Есть параметры для персонализации сообщения, которые вводит пользователь:

Пол – например, мужской.

Поколение – например, X.

Психотип – например, Конструктор.

Стадия бизнеса – например, Новичок.

Отрасль – например, Фармацевтика.

Организационно-правовая форма (ОПФ) – например, ООО.

По данным параметрам персонализации создается промпт для персонализации (в который также добавляется неперсонализированное сообщение):

«Адаптируй, не превышая длину сообщения в 250 знаков с пробелами, текст с учетом следующих особенностей:

Рекомендуется использовать короткие описания, объяснения и уточнения, избегать нелогичных фраз, подкреплять информацию статистикой и фактами. Примеры использования продукта другими людьми могут помочь убедить клиентов. Обращение должно быть почтительным и выражать уважение к жизненному опыту клиента. Также важно подчеркнуть экономию и выгоду, продемонстрировать проверенность продукта временем.

Используйте простой и понятный язык, избегайте сложной терминологии и акцентируйте внимание на основных преимуществах продукта. Добавьте подробное описание ключевых функций и возможностей, чтобы клиент мог лучше понять, как продукт может помочь его бизнесу. Используйте мягкий и дружелюбный тон, минимум прямых призывов к действию, таких как «оформите», «подключите», чтобы снизить давление на клиента. Включите информацию о дополнительных продуктах и объясните их пользу, где это уместно.

Текст должен подчеркивать важность инвестиций в медицинское оборудование, закупку сырья и развитие частных медицинских услуг. Используйте профессиональный и заботливый тон, акцентируя внимание на таких аспектах, как повышение качества медицинских услуг, расширение ассортимента препаратов и улучшение условий работы. Включите термины, связанные с медицинскими технологиями, здравоохранением и фармацевтическим производством.

Текст должен учитывать, что у общества с ограниченной ответственностью, как правило, есть сотрудники. Подчеркните преимущества продукта в контексте управления операционными процессами. Отразите выгоды для бизнеса, связанные с масштабом команды, и подчеркните эффективность продукта для управления и контроля над работой сотрудников.

Убедись, что в готовом тексте до 250 знаков с пробелами.

Убедись, что готовый текст начинается с призыва к действию с продуктом.

Убедись, что в готовом тексте есть следующая ключевая информация: Бесплатное обслуживание при покупках от 100 000 рублей в месяц.

Текст для адаптации: Хотите получить бесплатное обслуживание вашей новой бизнес-карты Mastercard Preffered? Просто сделайте покупки на сумму от 100 000 рублей в месяц! Никаких отчетов и платежных поручений, только удобство и выгода.»

По данному промпту получаем персонализированное сообщение – например:

«Уважаемые клиенты, мы рады предложить вам новую бизнес-карту Mastercard Preffered с возможностью бесплатного обслуживания при ежемесячных покупках от 100 000 рублей. Никаких отчетов и платежных поручений – только удобство и экономия времени.».

Почему разбили генерацию промпта на последовательное создание неперсонализированного SMS-сообщения, а затем его персонализацию – поскольку если подавать сразу очень длинный промпт на создание финального SMS-сообщения, то LLM теряет часть инструкций.

Дальше у меня происходят проверки сгенерированных сообщений:

**Правила проверки сгенерированных SMS-сообщений**

В значительном количестве проверок используется морфологический анализатор, т.е. инструмент, позволяющий определять части речи (существительное, глагол, причастие и др.) и приводить слова к начальному виду (например, «вкусного» → «вкусный», «долгов» → «долг»). В качестве морфологического анализатора был использован наиболее мощный на сегодняшний день – pymorphy2 – однако, и он иногда допускает ошибки.

1. **Запрещенные слова**

Проверяется отсутствие запрещенных слов: № один, номер один, № 1, вкусный, дешёвый, продукт, спам, доступный, банкротство, долги, займ, срочно, сейчас, лучший, главный, номер 1, гарантия, успех, лидер.

При этом по позициям № один, номер один, № 1 проверяется отсутствие точного совпадения, а по остальным словам с помощью морфологического анализатора проверяется отсутствие всех возможных вариантов слов (различные падежи, ед. и мн. ч. и т.д.).

1. **Запрещено обращение к клиенту**

Проверяется, есть ли в сообщении слова «привет» «здравствуй», «добрый», «дорогой», «уважаемый», «дорогая», «уважаемая», «господин», «госпожа», «друг», «коллега», «товарищ», «приятель», «друг», «подруга» (возможные изменения падежей, числа не проверяются).

1. **Запрещено приветствие клиента**

Совпадает с проверкой №2 («запрещено обращение к клиенту»).

1. **Запрещены обещания и гарантии**

Проверяется с помощью морфологического анализатора отсутствие в сообщении слов «обещать», «гарантировать», «обязаться» в различных вариантах (род, вид, число, лицо и т.д. глагола могут варьироваться).

1. **Запрещены составные конструкции из двух глаголов**

Морфологический анализатор определяет части речи слов и проверяет, не идут ли подряд два глагола в раках одного предложения (разделение предложений по точке, восклицательному или вопросительному знаку).

1. **Запрещены причастия и причастные обороты**

Морфологический анализатор определяет части речи слов и проверяет, нет ли причастий.

1. **Запрещены деепричастия и деепричастные обороты**

Морфологический анализатор определяет части речи слов и проверяет, нет ли деепричастий.

1. **Запрещена превосходная степень прилагательных**

Морфологический анализатор определяет части речи слов и проверяет среди прилагательных, нет ли из них каких-либо в превосходной степени.

1. **Запрещен страдательный залог**

Морфологический анализатор определяет части речи слов и проверяет среди глаголов, нет ли из них каких-либо в страдательном залоге.

1. **Запрещены порядковые числительные от 10 прописью**

Проверяется с помощью морфологического анализатора отсутствие в сообщении слов «десятый», «одиннадцатый», «двенадцатый», «тринадцатый», «четырнадцатый», «пятнадцатый», «шестнадцатый», «семнадцатый», «восемнадцатый», «девятнадцатый», «двадцатый» в различных вариантах (род, число, падеж и т.д. прилагательного может варьироваться).

1. **Запрещены цепочки с придаточными предложениями**

Выполняется подсчет, сколько раз в сообщении встретились слова или словосочетания из набора «который, которая, которое, которые, если, потому что, так как, что, когда, хотя, несмотря на то что» – и если больше 2, то условие считается нарушенным.

1. **Запрещены разделительные повторяющиеся союзы**

Проверяется отсутствие в сообщении двух союзов «и…и», «ни…ни», «то…то», «не то…не то», «или…или», «либо…либо» с запятой перед вторым союзом. *Прим.: возможны редкие контрпримеры вроде «я пошел в магазин, и он тоже, и она тоже».*

1. **Запрещены вводные конструкции**

Проверяется отсутствие в сообщении какого-либо из следующих слов или словосочетаний: во-первых, во-вторых, с одной стороны, по сути, по правде говоря, может быть, кстати, конечно, естественно, безусловно, возможно.

1. **Запрещены усилители**

Проверяется отсутствие в сообщении какого-либо из следующих слов или словосочетаний: очень, крайне, чрезвычайно, совсем, абсолютно, полностью, чисто.

1. **Запрещены паразиты времени**

Проверяется отсутствие в сообщении какого-либо из следующих слов или словосочетаний: сейчас, немедленно, срочно, в данный момент, теперь.

1. **Запрещены несколько существительных подряд, в том числе отглагольных**

Морфологический анализатор определяет части речи слов и проверяет, не идут ли подряд три существительных в раках одного предложения (разделение предложений по точке, восклицательному или вопросительному знаку).

1. **Запрещены производные предлоги**

Проверяется отсутствие в сообщении какого-либо из следующих слов или словосочетаний: в течение, в ходе, вследствие, в связи с, по мере, при помощи, согласно, вопреки, на основании, на случай, в продолжение, по причине, вблизи, вдалеке, вокруг, внутри, вдоль, посередине, вне, снаружи, благодаря, невзирая на, исходя из.

1. **Запрещены сложные предложения, в которых нет связи между частями**

Реализовано с помощью обращения к LLM.

1. **Запрещены сложноподчинённые предложения**

Проверяется отсутствие в сообщении какого-либо из следующих слов или словосочетаний: когда, как только, пока, после того как, потому что, так как, оттого что, благодаря тому что, чтобы, для того чтобы, если, когда бы, если бы, хотя, несмотря на то что, как, будто, словно, как будто, что. Но при этом дополнительно проверяется с помощью морфологического анализатора, чтобы после найденного слова или словосочетания до конца предложения не было глагола (чтобы разделять не сложноподчиненное предложение «ты бежишь как ветер» и сложноподчиненное «ты бежишь, куда я тебе скажу»), а также проверяется, что предложение не вопросительное («что вы купите?»). *Прим.: возможны редкие контрпримеры вроде «ты бежишь, куда я тебе скажу?».*

1. **Запрещены даты прописью**

Проверяется отсутствие словосочетания из двух слов:

Первое слово – любое из набора: первого, второго, третьего, четвертого, пятого, шестого, седьмого, восьмого, девятого, десятого, одиннадцатого, двенадцатого, тринадцатого, четырнадцатого, пятнадцатого, шестнадцатого, семнадцатого, восемнадцатого, девятнадцатого, двадцатого, двадцать первого, двадцать второго, двадцать третьего, двадцать четвертого, двадцать пятого, двадцать шестого, двадцать седьмого, двадцать восьмого, двадцать девятого, тридцатого, тридцать первого.

Второе слово – любое из набора: января, февраля, марта, апреля, мая, июня, июля, августа, сентября, октября, ноября, декабря.

1. **Запрещены близкие по смыслу однородные члены предложения**

Реализовано с помощью обращения к LLM.

1. **Запрещены шокирующие, экстравагантные, кликбейтные фразы**

Реализовано с помощью обращения к LLM.

1. **Запрещены абстрактные заявления без поддержки фактами и отсутствие доказательства пользы для клиента**

Реализовано с помощью обращения к LLM.

1. **Запрещены гарантирующие фразы**

Совпадает с проверкой №4 («запрещены гарантирующие фразы»).

1. **Запрещены узкоспециализированные термины**

Реализовано с помощью обращения к LLM.

1. **Запрещены фразы, способные создать двойственное ощущение, обидеть**

Реализовано с помощью обращения к LLM.

1. **Запрещены речевые клише, рекламные штампы, канцеляризмы**

Реализовано с помощью обращения к LLM.

Дальше я отправляю эти результаты проверок с промптом – постарайся не допускать следующих ошибок (добавляю найденные ошибки), и модель генерирует сообщения заново. Так происходит несколько раз в цикле.