# **# 1 Успешность прототипа**

Данный вопрос касается оценки вероятности успеха следующего прототипа, исходя из предыдущих результатов вашей компании и компании-конкурента. Для решения этой задачи можно использовать следующие подходы:

**Байесовский подход**

Этот метод включает в себя обновление вероятности на основе новых данных. В данном случае, можно использовать информацию о успехе прототипов конкурента как априорную вероятность успеха, а затем обновить эту вероятность, учитывая данные вашей компании. Требуется предположение о том, что обе компании работают в схожих условиях и что опыт конкурента может быть полезен для оценки вероятности успеха вашего прототипа.

**Частотный подход**

Этот метод основан на прямом анализе частот успеха и неудач. Но в вашем случае, поскольку ваша компания не имела успешных прототипов, применение этого метода может быть ограничено (не будем рассматривать).

**Метод максимального правдоподобия**

Можно попытаться оценить параметр успеха для вашей компании, используя данные о прототипах конкурента. Это потребует предположения о том, что вероятность успеха прототипа одинакова для обеих компаний.

**Симуляционный метод (например, Монте-Карло симуляция)**

Можно создать модель, симулирующую процесс создания и тестирования прототипов, используя известные данные, и оценить вероятность успеха на основе этой модели.

**Классификация на основе признаков**

Если у вас есть данные не только о количестве успешных и неуспешных прототипов, но и дополнительная информация о каждом прототипе (например, жанр игры, команда разработчиков, бюджет, отзывы пользователей на ранних стадиях разработки и т.д.), можно использовать алгоритмы классификации для предсказания вероятности успеха нового прототипа. Модели машинного обучения, такие как логистическая регрессия, случайный лес или градиентный бустинг, могут быть обучены на этих данных для предсказания успеха.

# **Рассматриваемые решения**

Для начала рассмотрим Байесовский подход и метод максимального правдоподобия.

**1)**Для Байесовского подхода нам нужно определить априорную вероятность успеха (исходя из данных конкурента) и обновить её, учитывая данные вашей компании. В случае метода максимального правдоподобия, мы оценим вероятность успеха, исходя из общего количества прототипов и количества успешных среди них.

* Согласно байесовскому подходу, оценочная вероятность того, что следующий прототип (201-й) вашей компании будет успешным, составляет примерно 0.499%. Это означает, что исходя из прошлого опыта вашей компании и данных конкурента, вероятность успеха следующего прототипа довольно низкая.
* Теперь рассмотрим метод максимального правдоподобия. Для этого нам нужно будет оценить параметр вероятности успеха прототипа (p), исходя из общего количества прототипов и количества успешных среди них. В данном случае, мы можем использовать данные конкурента, поскольку у нас нет успешных прототипов. Для оценки этого параметра можно воспользоваться формулой максимального правдоподобия для биномиального распределения, где оценка p равна отношению количества успехов к общему количеству прототипов. Согласно методу максимального правдоподобия, оценка вероятности успеха прототипа, исходя из данных конкурента, составляет 0.5%. Это означает, что если мы предполагаем, что вероятность успеха прототипов нашей компании аналогична вероятности успеха конкурента, то вероятность того, что следующий прототип будет успешным, также составляет около 0.5%.

Оба метода предполагают, что вероятность успеха прототипа вашей компании аналогична вероятности успеха конкурента.

**2)** В контексте поставленного вопроса мы можем использовать симуляцию Монте-Карло для моделирования процесса создания прототипов и оценки вероятности их успеха. Для этого мы будем случайным образом генерировать исходы для большого числа прототипов, используя оцененную ранее вероятность успеха.

Используем оценку вероятности успеха прототипа, полученную методом максимального правдоподобия (0.005 или 0.5%), и проведем большое количество испытаний, чтобы увидеть, сколько из них окажутся успешными.

Согласно симуляции Монте-Карло, вероятность того, что следующий прототип вашей компании будет успешным, составляет примерно 0.495%. Этот результат получен путем моделирования 1000000 прототипов с вероятностью успеха 0.5% (оцененной ранее методом максимального правдоподобия) и подсчета количества успешных среди них.

Важно отметить, что результаты симуляции Монте-Карло могут немного варьироваться при каждом запуске из-за случайной природы генерации исходов. Однако при большом количестве симуляций результаты должны быть достаточно стабильными и надежными.

**3)** Рассмотрим пример, в котором мы будем использовать фиктивные данные для обучения простой модели машинного обучения для предсказания успеха игрового прототипа. В качестве метода мы можем использовать логистическую регрессию, так как она хорошо подходит для задач бинарной классификации (успех/неуспех).

Предположим, что у нас есть данные с несколькими признаками для каждого прототипа:

**genre** - Жанр игры (например, экшен, стратегия, головоломка и т.д.).

**team\_size** - Размер команды разработчиков.

**development\_time** - Время разработки в месяцах.

**budget** - Бюджет на разработку (в миллионах долларов).

**early\_feedback** - Ранний отзыв пользователей (оценка от 1 до 10).

**success** - Успешен ли прототип (1 для успеха, 0 для неудачи).

Создадим фиктивный набор данных и обучим на нем модель логистической регрессии.

**Шаг 1: Генерация фиктивных данных**

Сначала сгенерируем некоторые фиктивные данные для демонстрации:

**Шаг 2: Подготовка данных и обучение модели**

Теперь мы подготовим данные для обучения модели. Для этого

* Преобразуем категориальные данные (жанр игры) в числовой формат с помощью one-hot encoding.
* Разделим данные на обучающую и тестовую выборки.
* Масштабируем числовые признаки для улучшения обучения модели.
* Обучим модель логистической регрессии и оценим ее точность.

*Произведем эти шаги*

Точность модели логистической регрессии на тестовых данных составляет примерно 51.7%. Это означает, что модель правильно предсказывает исход (успех или неуспех) прототипа в 51.7% случаев.

Важно отметить, что результаты сильно зависят от качества и реалистичности фиктивных данных. В реальных условиях точность модели может быть выше или ниже в зависимости от:

* Качества и объема реальных данных.
* Сбалансированности классов (соотношение успешных и неуспешных прототипов).
* Релевантности и полноты выбранных признаков.

Также важно провести более тщательный анализ и оптимизацию модели, включая выбор наиболее подходящих признаков, настройку гиперпараметров и применение методов для борьбы с несбалансированными данными (если это необходимо).