1.Background & Problem Statement

2.EDA & Key insights

3. Simulation of 2025 data

4.Semi-Markov Decision Processes

5.Results & Performance Comparison

6.Suggestions

7.Futurework

**1.Background & Problem Statement**

Google Ads Experts require a sophisticated staffing plan that adapts to market changes, as well as a dynamic demand. With this in mind, we created a model that optimizes revenue while prioritizing efficiency and customer needs.

We were given historical data from advertisers across 20 countries that provided us with insights regarding their Sign Up date and Projected Annual Budget. Each country has a threshold that makes their advertisers eligible to receive support from Google Ads, which is the main constraint for our project. We need to know how to assign up to 10 clients to each agent, taking into consideration staffing constraints such as hiring and firing personnel, which is exclusive to the country they work in.

Each advertiser can receive support from an agent for 60 days, after which that agent is available to help other customers. Our main goal is to meet advertiser demand, reducing the time it takes for eligible advertisers to get assigned to an agent, maximizing revenue split and adjusting for fluctuations in advertiser sign ups and hiring/firing constraints.

**2. EDA**

2.1 Monthly Trend by Country

By Visualizing the sign up quantity each month, we can find seasonal trends in a year. For example, there is a drop in June 2024, and for both 2023 and 2024, there was a rise in the last few months of the year.

2.2.1 Monthly Growth Rate by Country

Here we can have a better view, in 2024 there is a drop in June and rise in July. We did some research and we found out that this is because there is a change in the Google Ad Policy

2.2 Monthly Average Budget by Country

We plotted the average budget for each country. We can see the top 5 countries with the highest quantities, which are USA, India, China, Brazil, and the UK.

2.3 Growth Rate per Country (Mover over time):

Using the map, we have a clearer view of the amount of sign up in each month. The size of the circle stands for the amount of sign up each month, and the depth of color stands for average budget. From the change you can see the amount of sign-ups will go up during the last few months.

2.4 Growth rate of sign-up for each country from 2023 to 2024:

From the heat map, we can see the sign-up growth rate from 2023 to 2024. Argentina has a high growth rate, and Turkey has the lowest growth rate among all countries. We can reasonably invest more on those countries with higher growth rate

2.5 Distribution of budget of each country in 2023 and 2024:

From this visualisation we can see the distribution of budget of the advertiser in each country. The shape in each country matches the distribution of lognormal. We will mention that later. And from 2023 to 2024 you can see there is not much difference in shape and amount. We will reasonably assume that the budget in 2025 also follows such distribution.

2.6 Cluster of country by salary and average budget:

We can also cluster the country by the salary of the advisor and average budget of the advertiser. We classify the countries as 1.2.3.4.(refer to the dashboard). We might focus and invest more on the ‘investment focus market’ in future.

2.7 change of average budget over month each country:

From this plot we can see that the average budget doesn’t change much over the year, hence we will treat the distribution to be the same throughout the whole year.

**3. Simulation of 2025 data**

3.1.1Time series

We used a time series model and found that the prediction is not very accurate as long as 1. one/two year is to short for predicting. 2. no feature without and external data given

3.1.2 Simulate sign up dates and corresponding predicted budget

Since sign up dates follow the poisson distribution, we use Non-homogeneous Poisson Process to simulate 2025 sign up dates.

( we compare with 1. Time series （ARIMA/Prophet/LSTM）

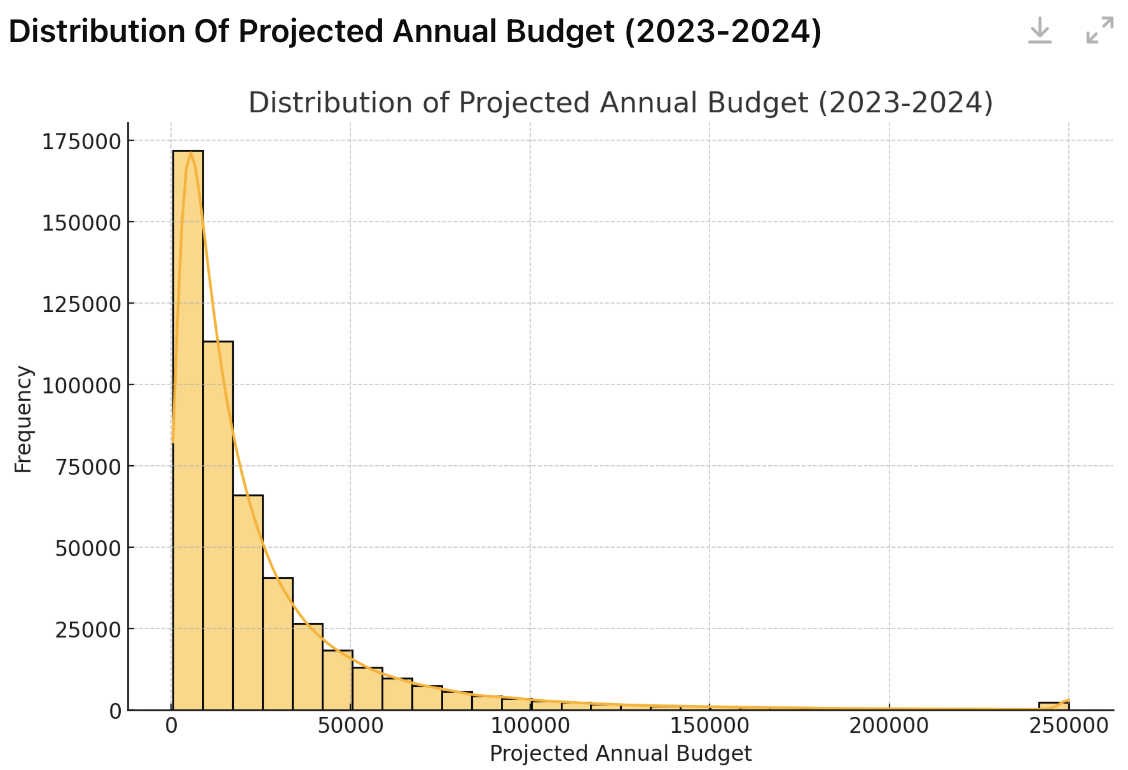
2. Brownian motion

3. Gaussian Process

4. Non-homogeneous Poisson Process

We then found out that the non-homogeneous Poisson Process most fit our situation)

For every new advertiser, we will assign a budget for them explicitly just like the data in 2023-2024. When we were doing EDA，we found that all countries have a lognormal distribution for predicted budgets. So we derived parameters of every country‘s lognormal distribution’s parameters. And we use the derived parameters to simulate the budget for every single new advertiser.



**4.Semi-Markov Decision Processes Model**

4.1Choice of Model ：Semi-MDP combined with Reinforce Learning：DQN

4.1.1 Dynamic Process Analysis

We will introduce this in the perspective of data processing:data input, data processing and data output.

(1)Data input:

1)Simulated 2025 data which contains every day’s sign up number and corresponding budget.

(2)Data processing：

1）Day1（in the mid month which is a more regular case to help you understand）: new sign ups

At the end of the day, check 2 pools, that is the waiting pool and the served pool.

2）At the first day of next month，

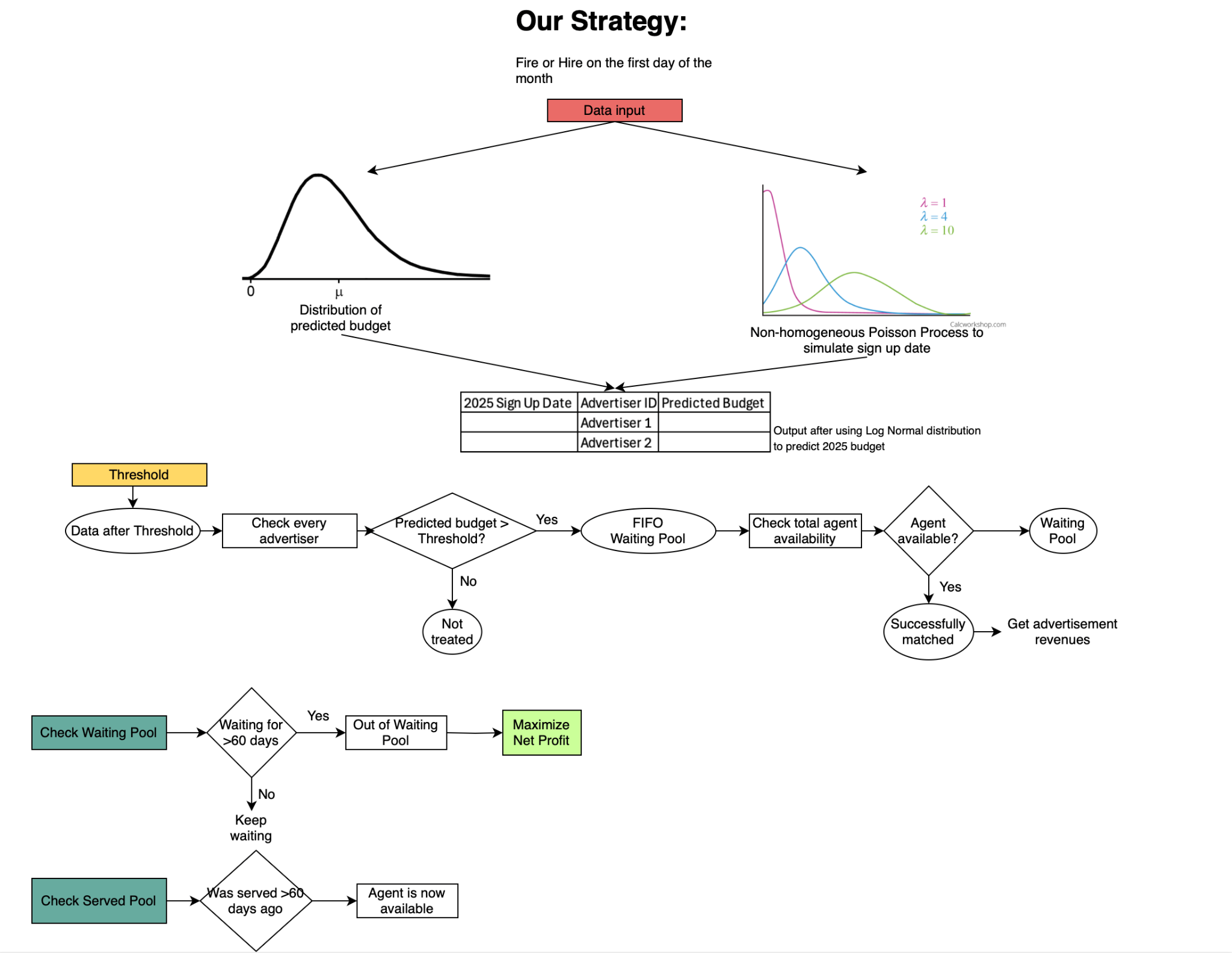
Then as time goes by，we should implement hiring/firing strategy for this month.

(3)Data output:

What we want to achieve：maximization Net Profit = Revenue（2 parts）-Cost（Firing Cost + Salary）

During the process, we cumulatively calculate the revenue and cost，so that we will get a net profit for 2025. Then we will optimize our strategy based on that.

Flow Chart of the whole process：



（连接到threshold中少了一个箭头，上下两个中间加一个layer，体现他们的时间先后顺序，然后data input 和data output搞得更明显一点）

Accordion to the deep analysis of our dynamic problem，the problem we are tackling presents two major challenges:

（1）One is dynamic and irregular decision timing：The decision-making process involves events occurring at irregular time intervals.Additionally, real-world systems evolve continuously rather than in discrete time steps, requiring a framework that can adapt dynamically.

Semi-MDP provides a structured way to handle dynamic state transitions and action timing.

Solution: We adopt a Semi-Markov Decision Process (semi-MDP) to explicitly handle variable-duration state transitions and better capture the real-time evolution of the system. Unlike traditional MDPs, semi-MDPs allow flexible timing between decisions, making them more suitable for problems where actions do not happen at fixed intervals. Furthermore, our model tracks the system state dynamically and enables strategic interventions at key decision points, allowing adjustments on a monthly basis.

（2）Another challenge is this problem requires us to continuously optimize the strategy

Given the complexity and uncertainty of real-world environments, a predefined strategy alone may not yield optimal results.Instead, the model should be able to learn from past experiences, adapt to changing trends, and refine its decision-making process over time.

Reinforcement Learning enables continuous policy improvement, ensuring adaptive and optimized decision-making.

Solution: We integrate Reinforcement Learning (RL) to continuously optimize decision policies. RL allows the model to learn from interactions with the environment, improving its strategy by maximizing long-term rewards. By iteratively updating policies based on observed outcomes, the model enhances decision efficiency and adaptability, leading to progressively better performance over time.

4.1.2 Choice of Model based on problem analysis

Based on what wue

We propose a semi-Markov Decision Process (semi-MDP) combined with Reinforcement Learning (RL) to tackle the problem efficiently. Our approach consists of two key components:

Process Tracking and Dynamic Adaptation (Semi-MDP)  
 The semi-MDP framework is designed to track the decision-making process while accommodating dynamic data changes over time. It allows us to model state transitions with variable time steps, which better captures real-world dynamics. Additionally, this structure enables us to input and adjust our strategy on a monthly basis, ensuring flexibility in decision-making.

Optimization and Learning (Reinforcement Learning)  
 The second component leverages Reinforcement Learning to continuously optimize decision policies based on observed outcomes. By learning from historical data and adapting to new trends, the RL model enhances decision-making efficiency and improves long-term performance.

This hybrid approach ensures that our model not only adapts to evolving conditions but also learns and improves decision strategies over time.

4.1.2 Model I:Semi-MDP （Enable Action Space）How to express these 2？

Compared with traditional LP, MILP,MDP and POMDP…..

Choice of Model：

Comparison：（just ask ChatGPT about their pros and cons and choose the most important and straightforward one）eg：

Why choose SMDP?

Fire/Hire is only executed once at the beginning of each month, but the state of the market (advertiser sign-up, Waiting Pool) changes every day, so standard MDP can't handle this, while SMDP allows for decisions with unequal time intervals.Fire takes effect 1 month in advance, Hire takes 1 month of training, and SMDP can model the impact of delayed decisions.

Why choose DQN?

The state space is large (daily changing advertiser registration, Waiting Pool, number of agents, etc).Actions are discrete (you can only Fire/Hire a whole number of agents, not 3.5).

（1）Notations：

（2）Assumptions：

1）The distribution of the predicted budget for the same country in 2025 is similar to its distribution in 2023 and 2024.

2）An employee who learns that he or she will be fired next month will not experience a reduction in the quality of his or her work that month, and thus will not result in a reduction in the advertiser uplift for the corresponding service.

（3）Model Structure

1）State Space St

（4）Semi-Markov Decision Processes Model

Initial Action（Fire/Hire）choice for Semi-MDP：Semi-MDP requires a reasonable fire/hire range input action space, so a Monte Carlo simulation is used to find a reasonable range

（Monte Carlo Simulation 只是“试探”不同策略，每次模拟都是独立的，所以可以并行计算，加快运行速度。Semi-MDP + DQN 需要“学习”最优策略，每次训练都要更新 Q-values，涉及大量状态转移和神经网络训练，计算时间更长（可能需要几小时））

4.1.3 Model II:Reinforce Learning：DQN （Enable Learning of Effect of delay ）

Compared with rule optimization,Q-Learning,PPO,.....(待finalize)

Comparison：（just ask ChatGPT about their pros and cons and choose the most important and straightforward one）eg：

DQN is suitable for SMDP with large state Spaces to learn long-term optimal Fire/Hire policies.

4.3

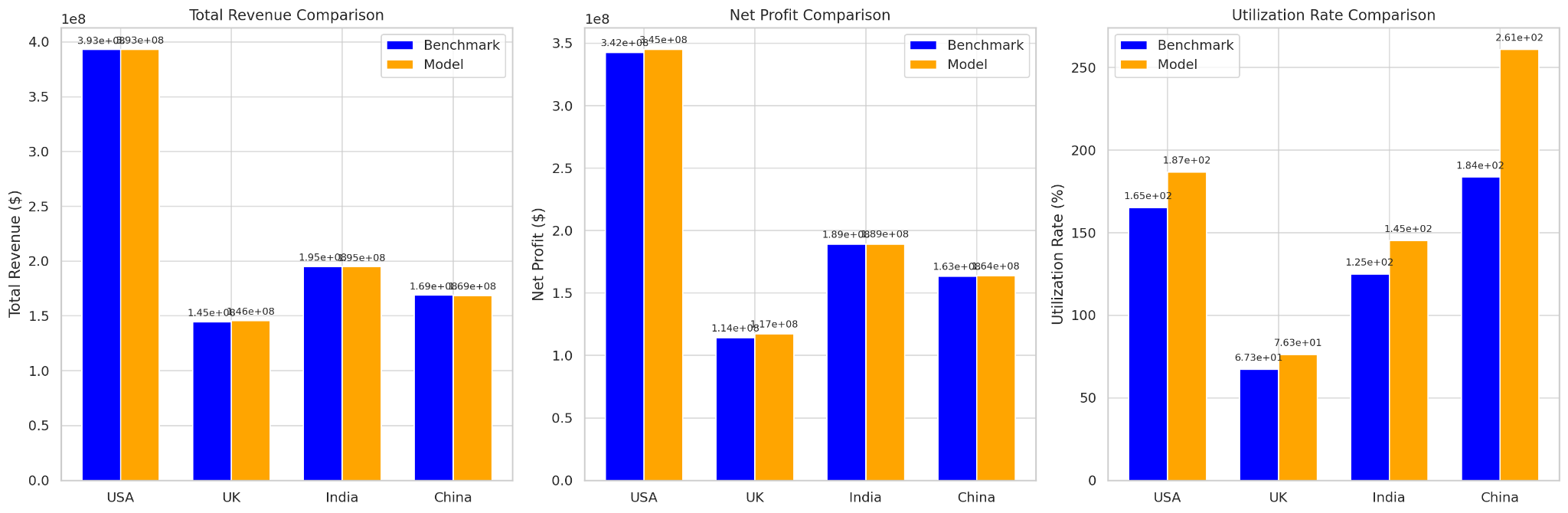
（1）Benchmark

（2）Semi-MDP+rule optimization（**using utilization rate or using short-term overload probability** ）

**Birth Death Process**

**Queueing Theory**

**Results:**

****

Increase in Net profit:

**USA**: $2,718,676.86

**UK**: $3,109,699.53

**India**: $150,586.55

**China**: $287,660.28

（3）Semi-MDP + **RL**

**Need more time to run:**

**The only output: USA: Net Profit: 393377328.99**

**5.Performance Comparison & Sensitivity test**

Metric：

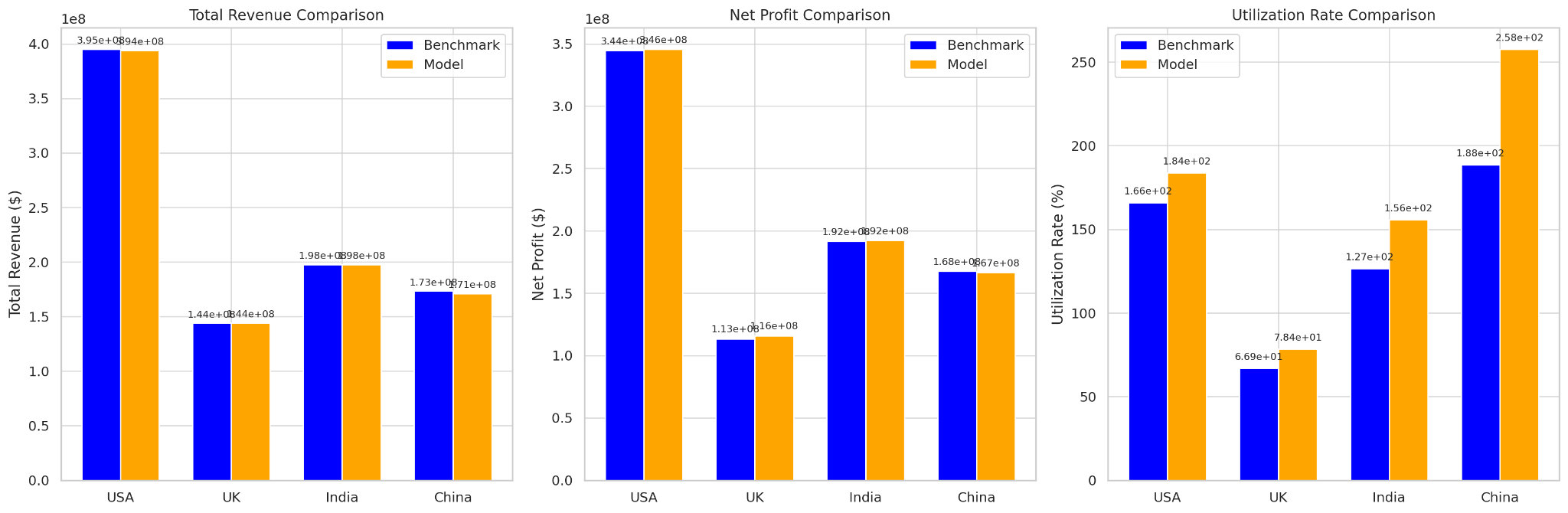
Total Revenue, Total Cost, Net Profit, Utilization Rate, Advertiser churn rate

Sensitivity Analysis：

Use simulation（use given distribution）to simulate different conditions.

（1）（联系之前的insight）由于google新推出的产品用户可能不适应，导致某1-2个月 sign up 数量骤降，后面回升到正常水平。

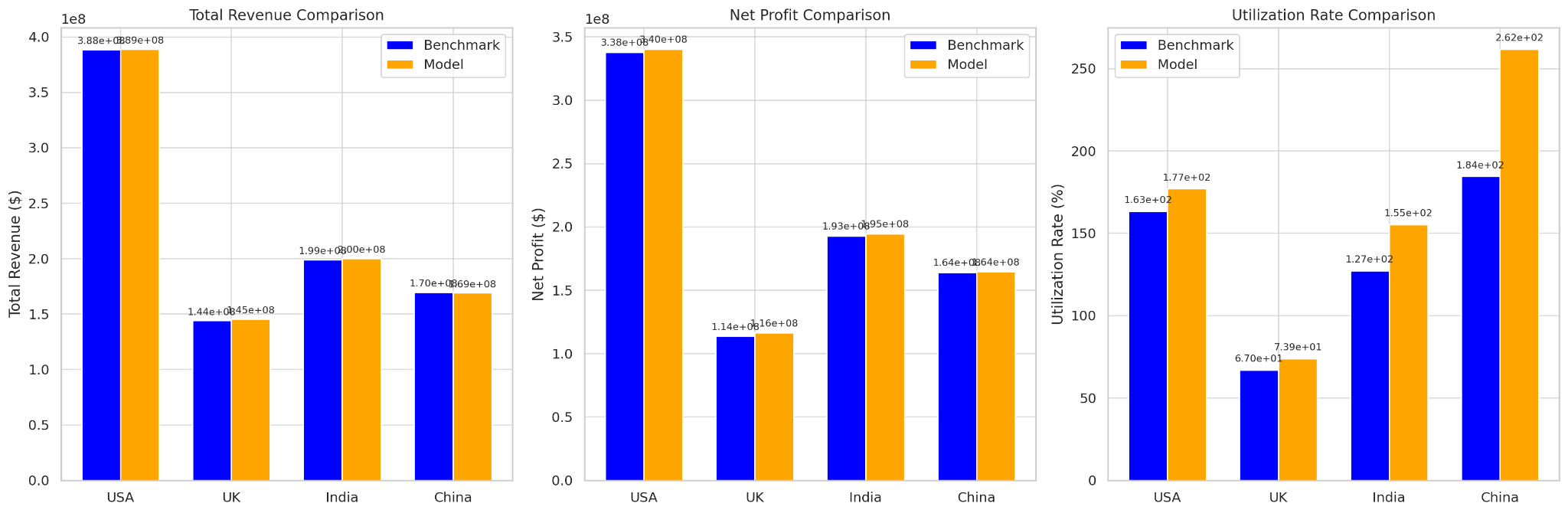
reference：搜索引擎的 AI 化转型：​Google 计划在 2025 年对搜索引擎进行全面的 AI 化改造，传统的搜索模式将被颠覆

Result: 

Increase in Net Profit:

* **USA**: $1,269,794.78
* **UK**: $2,481,857.65
* **India**: $533,471.82
* **China**: -$1,111,863.44 (decrease)

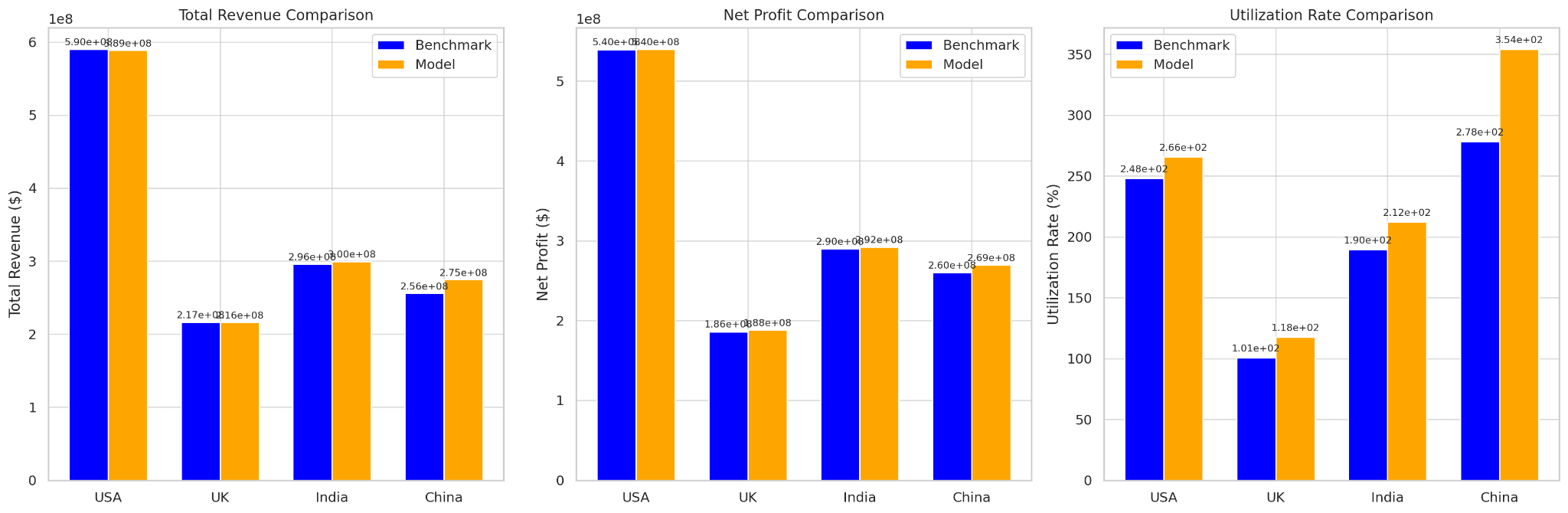
（2）google推出营销策略，突然导致用户大幅增长。Q4sign up 数量是原来的2倍。



Increase in Net Profit: **USA**: $2,385,481.15

* **UK**: $2,632,364.60
* **India**: $1,680,741.35
* **China**: $399,483.86

（3）2025年市场回暖，全年每个月sign up 数量都增长 50%。



Increase in net profit:

**USA**: $443,152.10

**UK**: $1,877,738.83

**India**: $1,965,008.37

**China**: $9,286,974.22

benchmark，semi-MDP + utility rate

**6.Suggestions**

（1）Help employees re-employment after layoff to reduce the company's reputation loss and employee negative psychology

…….

**7.Future Work：**

（1）Continuous work on RL：Takes more time and more computing power

（2）Adding Penalty function & reward function：

If advertisers receive service after 30 days，which may lead to lower service satisfaction, we could discount their expected uplifted budget accordingly.

（3）Optimize firing strategy：Firing agents with low uplift contributed first