Optimizing Automotive Logistics: Enhancing Efficiency in HiFlow's Vehicle Transportation Services

Author Name

February 26, 2024

Contents

1 Introduction

HiFlow, a leader in vehicle transportation services, is known for its adept delivery of single vehicles across multiple locations. The recent investment by Crédit Agricole Consumer Finance marks a pivotal moment for HiFlow, promising to bolster its offerings and propel its international expansion.
This partnership is set to enhance HiFlow's capabilities and fast-track its growth, aiming to establish a presence in 10 countries by 2026. Leveraging advanced technology and a suite of applications available in 6 languages, alongside a vast network of over 7,000 independent drivers, HiFlow delivers innovative and flexible logistics solutions within the automotive sector.
Currently operational in France, Belgium, and Spain, HiFlow is on a trajectory to redefine automotive logistics through its expansion and technological innovation.

2 Problem Statement

As HiFlow strides forward in vehicle transportation services, it faces a pivotal challenge in its operational model, particularly in the manual and inefficient driver-request matching process.
Despite its advanced platform that connects clients with drivers for vehicle transfers, the system's reliance on drivers to manually select transfer requests based on personal preferences often leads to last-minute scrambles to find suitable matches.
This not only introduces operational delays and increases costs by necessitating price hikes to attract drivers but also diminishes client and driver satisfaction. Furthermore, as HiFlow expands into new markets, such inefficiencies threaten to escalate costs and complicate operations further.
The core challenge HiFlow faces is developing a more efficient system for matching drivers with transfer requests. The goal is to minimize manual intervention, reduce the need for price increases as a driver incentive, and optimize the overall vehicle transportation process.
By addressing this challenge, HiFlow aims to enhance operational efficiency, lower costs, and improve service satisfaction for both clients and drivers.

3 Approach

In response to the inefficiency in HiFlow's driver-request matching process, our solution is the implementation of a Two-Tower recommendation system.
This system is designed to automate and optimize the matching of drivers with vehicle transfer requests, reducing the need for manual intervention and mitigating the necessity of last-minute price adjustments. The system is composed of two key components:

- **Driver Tower:** Utilizes advanced data analysis techniques to interpret driver-specific information. By converting this data into a high-dimensional embedding, the system constructs a profile for each driver that encapsulates their preferences and availability.
- **Request Tower:** Processes detailed attributes of each transfer request, such as pickup and drop-off locations. Similar to the Driver Tower, it transforms these details into an embedding that represents the unique characteristics of the request.

The core innovation lies in the interaction between these two towers. Upon receiving a new transfer request, the system generates its embedding and compares it against the embeddings of available drivers using cosine similarity metrics. This comparison yields a ranked list of drivers, ordered by their compatibility with the request, thereby facilitating an efficient and precise match.

In this model, $\mathbf{W}$, and $\mathbf{W}'$ Matrices denote the neural network weights for the Transfer and Driver Towers, respectively, while $\mathbf{B}$, and $\mathbf{B}'$ represent the corresponding biases. $\mathbf{D}$ captures the driver features, and $\mathbf{T}$ encapsulates the trip features .
Transfer Tower and Driver Tower operations are given by:

$$\left.\mathbf{E} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1} & w_{m2} & \cdots & w_{mn} \end{pmatrix} \begin{pmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{pmatrix} \right\} \text{Transfer Tower}$$

$$\left.\mathbf{E}' = \begin{pmatrix} w'_{11} & w'_{12} & \cdots & w'_{1n} \\ w'_{21} & w'_{22} & \cdots & w'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w'_{m1} & w'_{m2} & \cdots & w'_{mn} \end{pmatrix} \begin{pmatrix} D_1 \\ D_2 \\ \vdots \\ D_n \end{pmatrix} + \begin{pmatrix} B'_1 \\ B'_2 \\ \vdots \\ B'_m \end{pmatrix} \right\} \text{Driver Tower}$$

The score is given by the cosine similarity of the resulting vectors $\mathbf{E}$ and $\mathbf{E}'$.

Given the vectors $\mathbf{E}$ and $\mathbf{E}'$ from the Transfer Tower and Driver Tower, respectively, the score can be represented as the cosine similarity between these two embeddings, denoted as $\text{cosine}(\mathbf{E}, \mathbf{E}')$. This is given by:

$$\text{score} = \text{cosine}(\mathbf{E}, \mathbf{E}') = \frac{\mathbf{E} \cdot \mathbf{E}'}{\|\mathbf{E}\| \|\mathbf{E}'\|}$$

4 Challenges

In our analysis, we focused on records from France, as data from other countries were comparatively limited. A critical challenge in developing HiFlow's Two-Tower recommendation system was the presence of "attendance bias," where the system tends to favor drivers with extensive historical data and higher popularity. This creates a cycle that increases the visibility and opportunities for these drivers, overshadowing others who are equally skilled but less exposed.

Within our dataset, we identified two distinct subsets of drivers: the first comprised 691 drivers who received the majority of requests. Within this group, a specific subset of 253 drivers displayed an unusually high acceptance rate, leading to overrepresentation in the model's recommendations. The remaining drivers, totaling 2,765, were associated with a smaller number of requests.
The accompanying pie chart illustrates the distribution of records across the three subsets: Subset 1 includes drivers with many requests and typical acceptance rates, Subset 2 consists of drivers with numerous requests and exceptionally high acceptance rates, and Subset 3 encompasses drivers with a fewer number of requests.

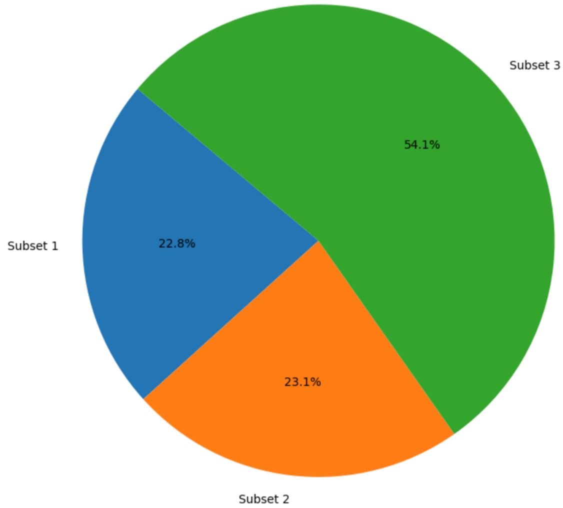

Figure 1: Pie chart showing the number of records for each driver subset.

A similar issue arose with drivers' regional distribution, with the system predisposing towards drivers from regions with higher request volumes, inadvertently sidelining drivers from less active regions.

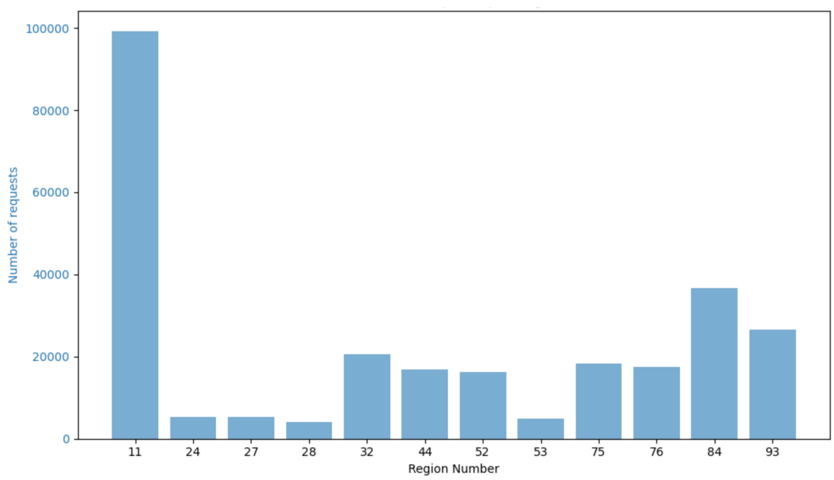

Figure 2: Bar chart showing the number of records for each region.

Such biases not only undermines the fairness of the system but also its ability to serve the varied needs of all transfer requests equitably. Therefore, addressing these biases was crucial to ensuring a balanced distribution of opportunities among drivers.

5 Solutions for Addressing Model Bias

To address the inherent biases in our model and enhance its fairness, we devised a strategic approach that includes data augmentation for underrepresented drivers and corrective adjustments for overrepresented ones. Additionally, to counteract regional bias, we implemented a weighted sampling method.

5.1 Weighted Sampling

A balanced training set was constructed by assigning weights to each driver record, compensating for imbalances across various driver regions. This ensures that each region contributes proportionately to the model's training process.

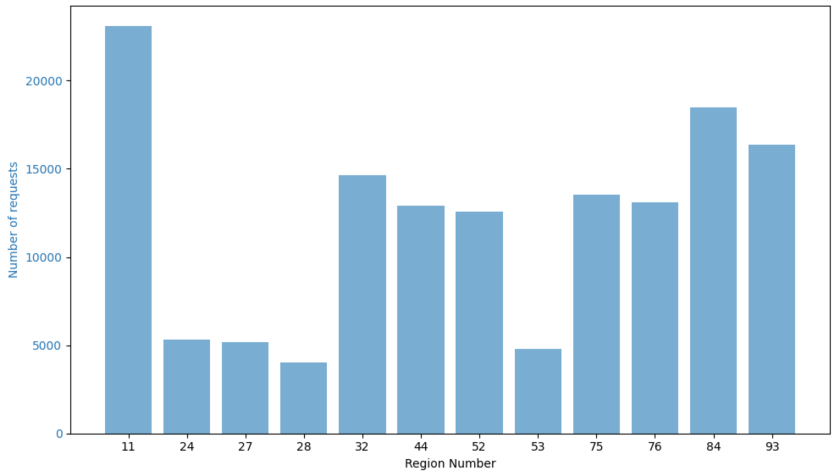

Figure 3: Bar chart showing the number of records for each region.

This chart displays a more uniform distribution of requests across various regions, with a majority of the regions receiving a significant number of requests.

5.2 Data Augmentation for Drivers with Fewer Requests

We focused on expanding the training dataset for the 2,765 drivers who had fewer requests by:
Quantile Binning: We used `pd.qcut` to divide distance data into 20 equal-frequency quantiles.
Interest Calculation: We computed the average match value within each quantile.
Match Scoring: We introduced a linear function, $y = ax + b$, for scoring matches. The coefficients a and b are calibrated based on the mean interest and the respective quantile's range.

The "distance" mentioned here includes the sum of the driver's location to the departure point and from the drop-off point back to the driver's location, providing a comprehensive measure of each trip's scope for analysis.

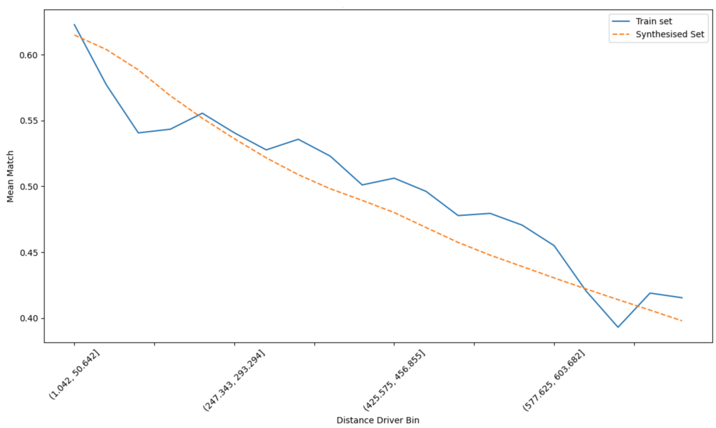

Figure 4: Mean Match Values Across Distance Bins for Original and Synthesised Driver Data.

The graph presents a comparative analysis of mean match values across binned driver distances, depicted by two distinct lines. The solid blue line signifies the initial distribution within the training dataset, while the dashed orange line corresponds to the synthesized dataset. The close alignment of the two lines across the distance bins indicates a negligible deviation in mean match values between the original and synthesized records, affirming the efficacy of the data augmentation method in preserving the inherent data characteristics.

5.3 Correction for Overrepresented Drivers

For the 267 drivers with abnormally high request and acceptance rates, we generated synthetic negative records to neutralize their disproportionate influence. This involved the same binning and scoring techniques, reflecting a more authentic request and acceptance distribution.

6 Analysis of Model Results

6.1 Metrics

To rigorously evaluate the effectiveness and fairness of our recommendation system, we established a suite of metrics tailored to capture both performance and bias:

i. **Rank Match Metric:** examines the rank assigned to drivers who were accepted for a transfer. This metric helps in understanding the system's accuracy in identifying suitable matches.

ii. **Rank No Match Metric:** Focuses on the system's ability to correctly identify unsuitable matches, by ranking drivers who were not accepted for a transfer. A higher rank for non-matching drivers indicates the system's effectiveness in discerning incompatible pairings, which is crucial for avoiding inefficient recommendations.

iii. **Diversification Metric:** This evaluates the variety in the top-n drivers recommended by the system across 100 transfers, assessing whether the model shows a bias toward certain drivers. It is essential for ensuring that the system fairly represents a broad spectrum of drivers, rather than favoring a select few.

Given the data's tendency to favor a specific group of drivers, we prioritize the Rank No Match and Diversification metrics in our evaluation strategy. These metrics are pivotal in addressing the system's inherent bias by fostering diversity and accurately identifying mismatches between drivers and transfer requests. This approach is essential for realigning the recommendation system with our goals of fair opportunity distribution and effectiveness in fulfilling diverse transfer needs.