Mathematical Formulation of ML-Driven Pricing and Dispatching Strategy

Strategy & Operations Analytics

2026 年 2 月 26 日

1 Time Series Forecasting: Generalized Additive Model (GAM)

The Prophet forecasting algorithm decomposes the time series of bike-sharing demand into a Generalized Additive Model (GAM). For any discrete time step t, the predicted demand $y(t)$ is formulated as:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \tag{1}$$

Nomenclature (符号释义):

- $y(t)$: The target observable variable (e.g., predicted actual rides at time t).
- $g(t)$: **Trend component (趋势项)**. Captures the non-periodic long-term growth or decline of the business baseline.
- $s(t)$: **Seasonality component (季节项)**. Captures periodic fluctuations (e.g., daily commutes, weekly cycles, yearly winter/summer patterns).
- $h(t)$: **Holiday/Event component (节假日项)**. Captures the deterministic shocks caused by anomalous events (e.g., snowstorms, Thanksgiving).
- $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$: **Error term (白噪声误差)**. Represents the unexplainable random variance in the real world.

1.1 Piecewise Linear Trend $g(t)$

To handle sudden business shifts (e.g., sudden temperature drops), the trend is modeled as piecewise linear:

$$g(t) = \left(k + \mathbf{a}(t)^T\boldsymbol{\delta}\right) t + \left(m + \mathbf{a}(t)^T\boldsymbol{\gamma}\right) \tag{2}$$

Nomenclature (符号释义):

- k: Base growth rate (初始基础增长斜率).
- m: Base offset parameter (初始截距).
- $\mathbf{a}(t) \in \{0,1\}^S$: Indicator vector for S changepoints (变点激活开关). $a_j(t) = 1$ if time t has passed the j-th changepoint.
- $\boldsymbol{\delta} \in \mathbb{R}^S$: Rate adjustments at each changepoint (斜率变化幅度). Mathematically constrained by a Laplace prior $\delta_j \sim \text{Laplace}(0, \tau)$ to prevent overfitting (L1 Regularization).
- $\boldsymbol{\gamma}$: Offset adjustments (截距补偿向量) to ensure the trend line remains continuous at the changepoint.

1.2 Fourier-based Seasonality $s(t)$

Periodic patterns are approximated using truncated Fourier series:

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right) \right) = \mathbf{X}(t)\boldsymbol{\beta} \tag{3}$$

Nomenclature (符号释义):

- P: The physical period of the seasonality (物理周期, e.g., $P = 365.25$ for yearly, $P = 7$ for weekly).
- N: The order of Fourier approximation (傅里叶截断阶数). Higher N fits more complex wave shapes but risks overfitting.
- a_n, b_n: Fourier coefficients (波浪的振幅参数), jointly represented as vector $\boldsymbol{\beta} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$.

---

2 Operations Research: Constrained NLP & Bayesian Optimization

Once the baseline demand is forecasted, an Operations Research (OR) model is constructed to find the optimal pricing and asset dispatching combination.

2.1 Objective Function: Profit Maximization

Let the decision variables be $\mathbf{x} = [P_e, P_c, Q_e, Q_c]^T \in \mathcal{X}$, representing the prices and physical supply quantities for e-bikes and classic bikes. The objective is to maximize the expected profit $\Pi(\mathbf{x})$:

$$\max_{\mathbf{x}\in\mathcal{X}} \Pi(\mathbf{x}) = (P_e - C_e)\min(\hat{D}_e(\mathbf{x}, \mathbf{w}), Q_e) + (P_c - C_c)\min(\hat{D}_c(\mathbf{x}, \mathbf{w}), Q_c) - F_e Q_e - F_c Q_c \quad (4)$$

Nomenclature (符号释义):

- $\mathbf{x}$: The strategy vector (决策向量), the ultimate answer we want to find.
- $\Pi(\mathbf{x})$: Total System Profit (系统总利润).
- P_e, P_c: Target ARPU for E-bikes and Classic bikes (目标定价/客单价).
- C_e, C_c: Marginal fulfillment & ops costs (边际履约/换电调度成本).
- Q_e, Q_c: Physical dispatched supply (实际物理投放车辆数).
- F_e, F_c: Fixed daily depreciation costs per bike (单车日均固定折旧费).
- $\hat{D}(\mathbf{x}, \mathbf{w})$: The ML Surrogate Demand Function (机器学习预估真实需求), outputted by XGBoost decision trees. $\mathbf{w}$ represents exogenous variables like weather.
- $\min(\hat{D}, Q)$: **The Truncation Function (供需短板截断算子)**. Enforces the physical reality that actual rides cannot exceed physical supply or user demand.

2.2 SLA Constraints & Big M Penalty

To enforce the local government's Service Level Agreement (SLA), the total fleet size must exceed $Q_{\min}$. We transform the constrained problem into an unconstrained one via a Penalty Function $\mathcal{L}(\mathbf{x})$:

$$\mathcal{L}(\mathbf{x}) = \Pi(\mathbf{x}) - M \cdot \mathbb{I}(Q_e + Q_c < Q_{\min}) \quad (5)$$

Nomenclature (符号释义):

- M: A sufficiently large positive scalar (大 M 惩罚项, e.g., 10^9).
- $\mathbb{I}(\cdot)$: Indicator function (指示函数). Returns 1 if the SLA constraint is violated, imposing an infinite penalty to the profit.

2.3 TPE Algorithm: Expected Improvement (EI)

Since $\mathcal{L}(\mathbf{x})$ is non-differentiable (due to the min function and Tree-based $\hat{D}$), the Tree-structured Parzen Estimator (TPE) is utilized to find the global optimum by maximizing the Expected Improvement $EI_{y^*}(\mathbf{x})$:

$$EI_{y^*}(\mathbf{x}) \propto \left(\gamma + \frac{g(\mathbf{x})}{l(\mathbf{x})}(1-\gamma)\right)^{-1} \tag{6}$$

Nomenclature (符号释义):

- y^*: The profit threshold (利润分位数阈值, e.g., Top 15% historical trials).
- $l(\mathbf{x}) \equiv p(\mathbf{x}|y \geq y^*)$: The Probability Density Function (PDF) of parameters that yield **High Profit** (产出好结果的参数核密度分布).
- $g(\mathbf{x}) \equiv p(\mathbf{x}|y < y^*)$: The PDF of parameters that yield **Low Profit or Violate Constraints** (产出坏结果或触碰红线的参数分布).
- γ: The prior probability $p(y \geq y^*)$ (全局好结果先验概率).

Theoretical Conclusion: To maximize EI, the TPE algorithm explicitly searches for parameters $\mathbf{x}$ that maximize the likelihood ratio $\frac{l(\mathbf{x})}{g(\mathbf{x})}$, circumventing the need for mathematical gradients entirely.