Time series analysis using Prophet

Based on the Generalized Additive Models (GAMs)

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

- $g(t)$: Trend function (non-periodic changes)
- $s(t)$: Seasonality function (periodic changes)
- $h(t)$: Holiday effect function
- ϵ_t: Error term

Trend $g(t)$

- Logistic Trend Model
- Linear Trend Model

Logistic Trend Model

Based on the classic logistic growth model

$$g(t) = \frac{C}{1 + \exp(-k(t - m))}$$

- C: Carrying capacity
- k: Growth rate
- m: Offset parameter

Logistic Trend Model

Introduction of changepoints s_j, where $j \in S$: timestamps at which growth rate can change

- $k + \sum_{j|t \geq s_j} \delta_j$: growth rate at time t

Prophet utilizes several vectorial representations

- $\mathbf{a}(t) \in \{0, 1\}^S$
- $a_j(t) = \begin{cases} 1 & \text{if } t \geq s_j \\ 0 & \text{otherwise} \end{cases}$
- $k + \mathbf{a}(t)^T \boldsymbol{\delta}$: growth rate at time t

Logistic Trend Model

Phophet changes to the classical equation

$$g(t) = \frac{C(t)}{1 + \exp(-(k + \mathbf{a}(t)^T \boldsymbol{\delta})(t - (m + \mathbf{a}(t)^T \gamma)))}$$

- $C \to C(t)$: Carrying capacity changes over time
- $k \to k + \mathbf{a}(t)^T \boldsymbol{\delta}$: Growth rate can change over time at specific changepoints s_j
- $m \to m + \mathbf{a}(t)^T \gamma$: Adjustment to the offset parameter to account for changepoints

Linear Trend Model

$$g(t) = \underbrace{(k + \mathbf{a}(t)^T \boldsymbol{\delta})}_{\text{slope}} t + \underbrace{(m + \mathbf{a}(t)^T \gamma)}_{\text{intercept}}$$

- $k + \mathbf{a}(t)^T \boldsymbol{\delta}$: Growth rate
- $m + \mathbf{a}(t)^T \gamma$: Offset parameter

Seasonality $s(t)$

- Additive seasonality: $g(t) + s(t)$
- Multiplicative seasonality: $g(t)s(t)$

Seasonality $s(t)$

Based on Fourier series cosine-sine form

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

- $\mathbf{X}(t) = \left[\cos\left(\frac{2\pi t}{P}\right), \sin\left(\frac{2\pi t}{P}\right), \dots, \cos\left(\frac{2\pi Nt}{P}\right), \sin\left(\frac{2\pi Nt}{P}\right)\right]$
- $\boldsymbol{\beta} = [a_1, b_1, \dots, a_N, b_N]^T$
- N: Order of Fourier series
- P: Period

Holiday $h(t)$

Introducing holidays D_i, where $i \in L$: days at which an holiday occurs

$$h(t) = \mathbf{Z}(t)\kappa$$

- $\mathbf{Z}(t) = \begin{cases} z_i = 1 & \text{if } i \in L \\ z_i = 0 & \text{otherwise} \end{cases}$
- $\mathbf{Z}(t)$: vector of holiday dummies
- κ: Effect of the holiday

Final Model

$$y|m, \delta, \beta, \kappa, \epsilon \sim \mathcal{N}(g(t) + s(t) + h(t), \epsilon)$$

- The parameters are obtained as Maximum A Posterior (MAP) estimates

Additional Features

- Seasonalities: Custom seasonalities can be added with `add_seasonality()`
- Regressors: Additional regressors can be added using `add_regressor()`, although their future values have to be known
- Uncertainty: Confidence intervals are given by Markov chain Monte Carlo sampling

Practical Implementation

1. Import Prophet
2. Fit the model
3. Tune the parameters
4. Evaluate the results
5. Repeat step 3. and 4. as needed