



ОД А POSTERIORI КОН ИНФОРМИРАНИ А PRIORI РАСПРЕДЕЛБИ

БАЕСОВО ПРЕНОСНО УЧЕЊЕ ЗА ВРЕМЕНСКИ СЕРИИ

Јован Крајевски

2 јули 2025

Содржина

1 Дефиниција на проблемот

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet

- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

Прогноза на кратки временски серии

1 Дефиниција на проблемот

- Тренирачко множество: 500 серии \times 3 месеци
- Тестирачко множество: 365 дена
- Проблеми
 - over-fitting
 - сезонални ефекти со голема периода
- Како изгледа прогноза од класичен модел?

Прогноза од класичен модел

1 Дефиниција на проблемот



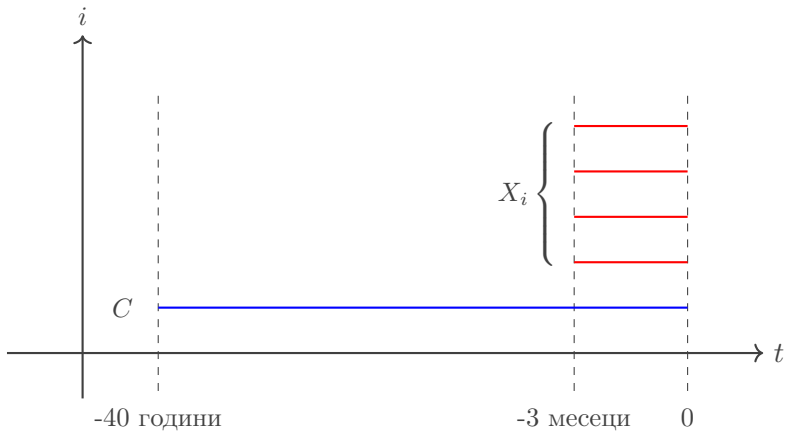
Решение со преносно учење

1 Дефиниција на проблемот

- Контекстна временска серија: 40 години
- Моделот е истрениран на контекстната временска серија
 - Заклучуваме дека временската серија расте во март
- Предиктор временски серии фино го подесуваат моделот
 - Го подесуваме растот во март за секоја од 500-те предиктор временски серии

Визуелизација на временските серии

1 Дефиниција на проблемот



Содржина

2 Теоретска рамка

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet
- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

Класична наспроти Баесова статистика

2 Теоретска рамка

- Објективисти наспроти Субјективисти
- Честота наспроти Верување
- Точкесто оценување наспроти *a posteriori* распределба
 - параметрите се случајни променливи!
 - параметарот секогаш има распределба - и пред, и после податоците
 - од *a priori* до *a posteriori* преку Баесовата теорема
- Интервали на доверба наспроти Интервали на кредибилитет

Зошто Баесова статистика?

2 Теоретска рамка

- Присуство на *a priori* верување
- Целосна распределба на параметрите
- Мали примероци

Зошто навистина Баесова статистика?

2 Теоретска рамка

- Хиерархиско моделирање
- Детално истражување на *a posteriori* распределбата
- Принципиелно преносно учење
 - $a\ priori \rightarrow a\ posteriori \rightarrow a\ priori \rightarrow a\ posteriori \rightarrow \dots$

Содржина

3 *A priori* распределби

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet
- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

Видови *a priori* распределби

3 *A priori* распределби

- Информирани *a priori* распределби
 - Потекнуваат од знаењето на експертите
 - Нема формален метод да се дојде до нив
- Слабо-информирани *a priori* распределби
 - Принципиелна регуларизација
- Генерички слабо-информирани *a priori* распределби
- Нејасни *a priori* распределби
- Рамномерни *a priori* распределби
 - Класична статистика + истражување на *a posteriori* распределба

Нормални *a priori* распределби

3 *A priori* распределби

$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$

$$\text{posterior} \propto \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \times \text{likelihood} \quad / \log$$

$$\text{logposterior} \propto -\frac{1}{2\sigma^2}(x - \mu)^2 + \text{loglikelihood}$$

- Ако $\mu = 0$, ова е L2 регуларизација со $\lambda = \frac{1}{2\sigma^2}$

$$\mathcal{L}_{L2} = \mathcal{L} + \lambda \|x\|_2^2$$

Лапласови *a priori* распределби

3 *A priori* распределби

$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$

$$\text{posterior} \propto \frac{1}{2b} e^{-\frac{|x-\mu|}{b}} \times \text{likelihood} \quad / \log$$

$$\text{logposterior} \propto -\frac{1}{b}|x - \mu| + \text{loglikelihood}$$

- Ако $\mu = 0$, ова е L1 регуларизација со $\lambda = \frac{1}{b}$

$$\mathcal{L}_{L1} = \mathcal{L} + \lambda \|x\|_1$$

Содржина

4 Facebook Prophet

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet

- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

Општа дефиниција на моделот

4 Facebook Prophet

- Обоштен адитивен модел:

$$y_i(t; \gamma_i, \theta_i) = g_i(t; \gamma_i) + \sum_{j=1}^m s_i(p^{(j)}, t; \theta_i^{(j)}) + h_i(t) + \epsilon_i^{(t)}$$

- g_i - линеарен тренд по делови
- s_i - сезоналност моделирана со Фуриеров ред

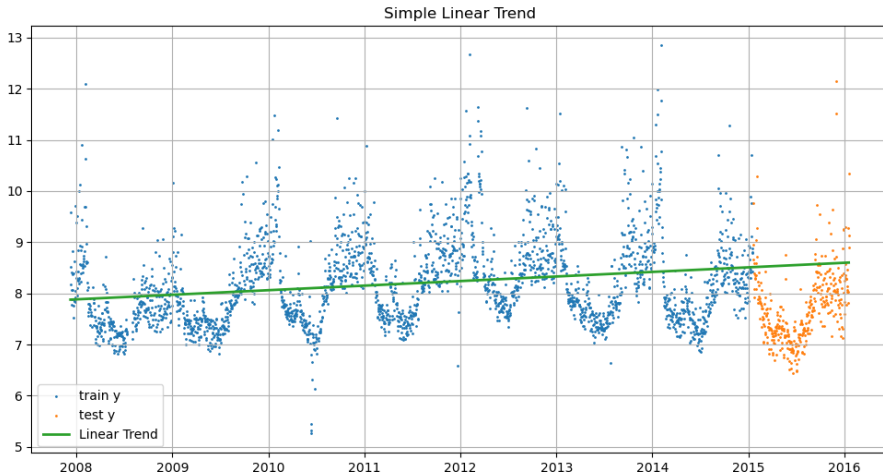
$$P = \{p^{(j)} : j \in \mathbb{N}, 1 \leq j \leq M\} \subseteq \mathbb{R}$$

$$\forall k, j \in \mathbb{N}, j \leq m$$

$$s(p^{(j)}, t^{(k)}; \theta_i^{(j)}) = s(p^{(j)}, t^{(k)} - p^{(j)}; \theta_i^{(j)})$$

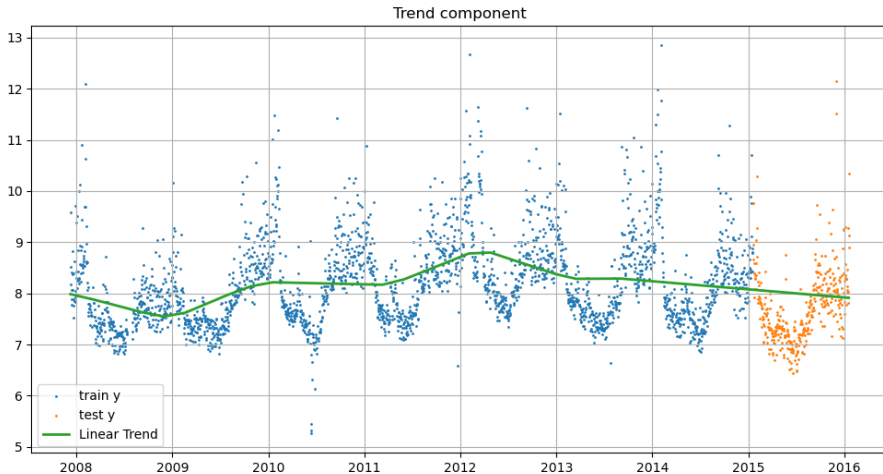
Визуелизација на линеарниот тренд

4 Facebook Prophet



Визуелизација на линеарниот тренд

4 Facebook Prophet



Параметри на линеарниот тренд

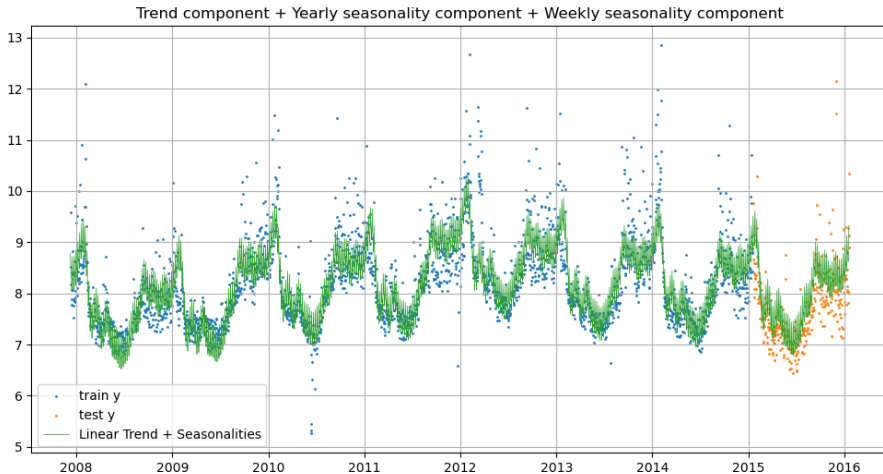
4 Facebook Prophet

- $w_i \sim \mathcal{N}(0, 5)$ - почетен раст - **параметар**
- $b_i \sim \mathcal{N}(0, 5)$ - почетен пресек со y -оската - **параметар**
- \mathbf{s}_i - точки на промена на растот w_i
- $\boldsymbol{\delta}_i \sim \text{Laplace}(0, 0.05)$ - промена на растот во точките \mathbf{s}_i - **параметар**
- \mathbf{A}_i - индикатор матрица, $t^{(k)} > s_{i,l} \implies a_{i,k}^{(l)} = 1$
 - корисно за векторизација

$$\mathbf{g}_i = (w_i + \mathbf{A}_i \boldsymbol{\delta}_i) \odot \mathbf{t}^T + b_i - \mathbf{A}_i (\boldsymbol{\delta}_i \odot \mathbf{s}_i)$$

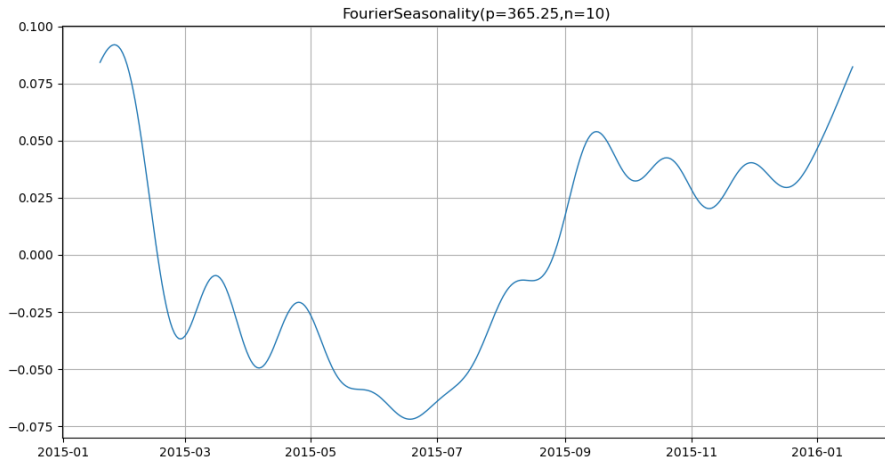
Визуелизација на Фуриеровата сезоналност

4 Facebook Prophet



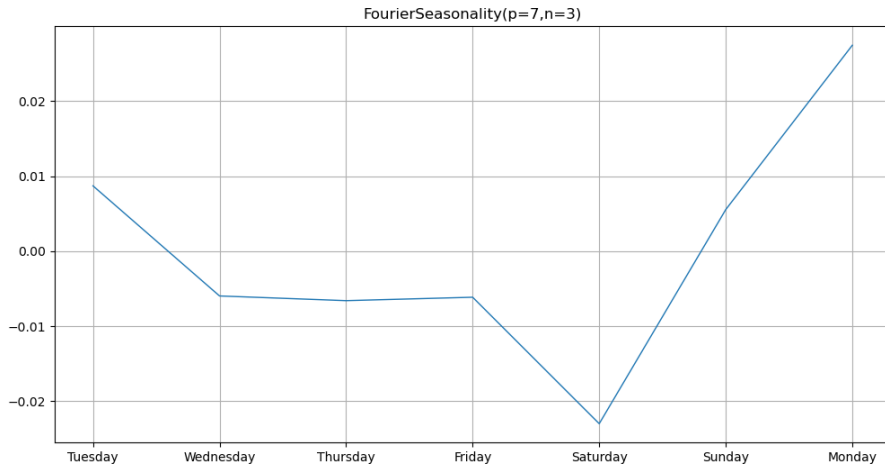
Визуелизација на Фуриеровата сезоналност

4 Facebook Prophet



Визуелизација на Фуриеровата сезоналност

4 Facebook Prophet



Параметри на Фуриеровата сезоналност

4 Facebook Prophet

- C_j - број на синусни и косинусни собироци за периодата $p^{(j)}$
- $\theta_i^{(j)} \sim \mathcal{N}(0, 10)$ - коефициенти на Фуриеровиот ред - **параметар**
- $\mathbf{F}^{(j)}$ - матрица од синусни и косинусни собироци
 - корисно за векторизација

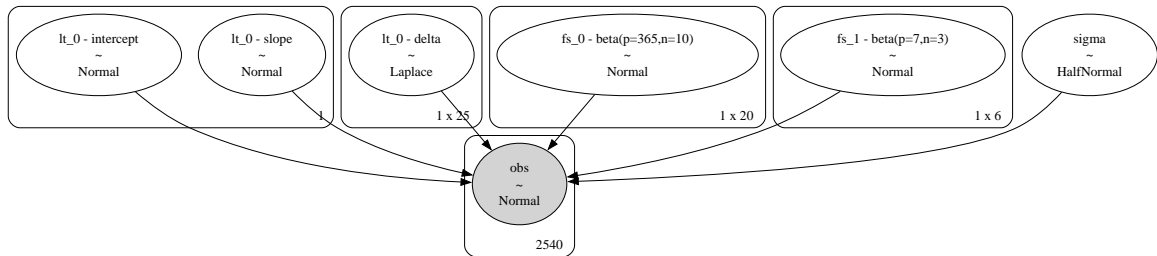
$$\mathbf{fs}_i^{(j)} = \mathbf{F}^{(j)} \theta_i^{(j)}$$

Целосен модел

4 Facebook Prophet

$\sigma_i \sim \mathcal{HN}(0, 0.5)$ - параметар

$$y_i \sim \mathcal{N}(\mathbf{g}_i + \sum_{j=1}^M \mathbf{fs}_i^{(j)}, \sigma_i)$$



Содржина

5 Vangja

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet

- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

Опис на Vangja

5 Vangja

- Надградба на Facebook Prophet
 - модуларен интерфејс
 - хиерархиско моделирање
 - преносно учење
 - разни методи за Баесово заклучување (MCMC, VI, MAP)
- Python, PyMC, JAX, GPU поддршка



Апликативен интерфејс

5 Vangja

„Течен“ интерфејс

```
1 model = Prophet(  
2     seasonality_mode="multiplicative",  
3     yearly_seasonality=10,  
4     weekly_seasonality=3,  
5 ).add_seasonality(  
6     name="monthly",  
7     period=30.5,  
8     fourier_order=5,  
9 )
```

Модуларен интерфејс

```
1 model = LinearTrend() * (  
2     1  
3     + FourierSeasonality(365.25, 10)  
4     + FourierSeasonality(30.5, 5)  
5     + FourierSeasonality(7, 3)  
6 )
```

Додатоци на Facebook Prophet

5 Vangja

- Промени на растот „од десно“
 - \mathbf{A}_i - индикатор матрица, $t^{(k)} < s_{i,l} \implies a_{i,k}^{(l)} = 1$
- Векторизирана анализа на повеќе временски серии
 - $\mathbf{s} = \mathbf{s}_1 = \dots = \mathbf{s}_N \implies \mathbf{A} = \mathbf{A}_1 = \dots = \mathbf{A}_N$

$$\mathbf{G} = (\mathbf{w} + \mathbf{A}\Delta) \odot \mathbf{t}^T + \mathbf{b} - \mathbf{A}(\Delta \odot \mathbf{s})$$

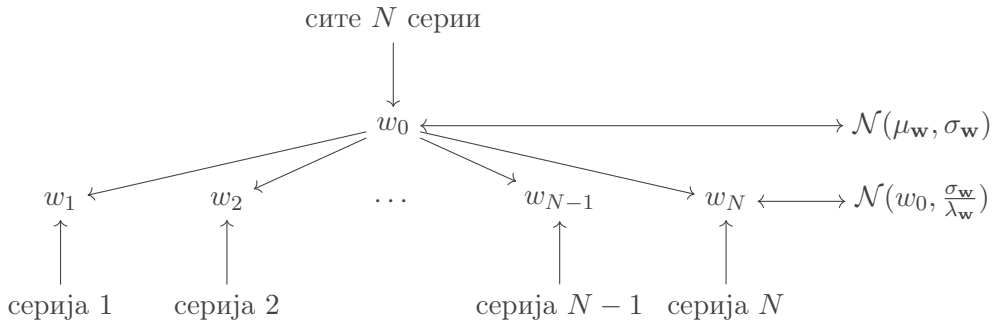
$$\mathbf{F}\mathbf{S}^{(j)} = \mathbf{F}^{(j)}\boldsymbol{\theta}^{(j)}$$

$$\mathbf{y} \sim \mathcal{N}(\mathbf{G} + \sum_{j=1}^M \mathbf{F}\mathbf{S}^{(j)}, \boldsymbol{\sigma}_\epsilon)$$

Хиерархиско моделирање

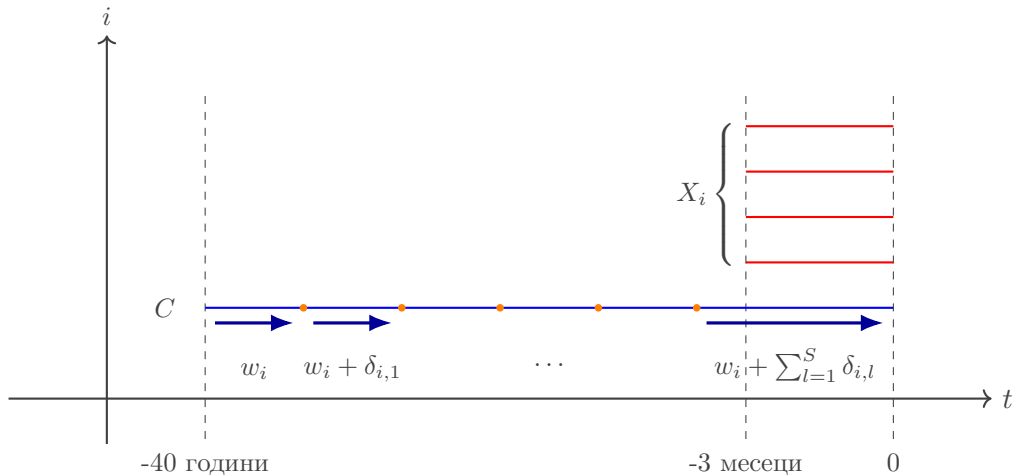
5 Vangja

локално $\sim \mathcal{D}(\mu = \text{глобално}, \sigma = \frac{\text{глобална несигурност}}{\text{стеснување}})$



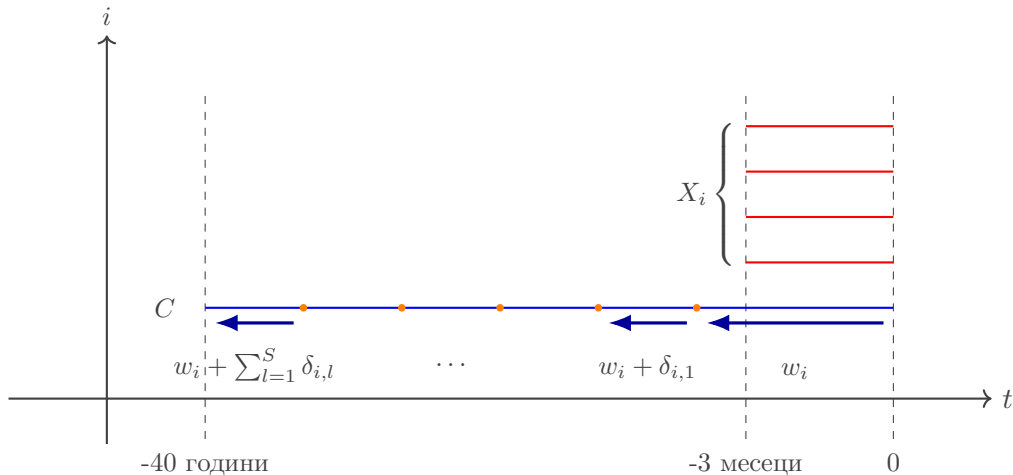
Зошто промени на растот од „десно“?

5 Vangja



Зошто промени на растот од „десно“?

5 Vangja



Репараметризација - проблем на инка

5 Vangja

локално = глобално + отстапување \times мерна единица

$$w_0 \sim \mathcal{N}(\mu_{\mathbf{w}}, \sigma_{\mathbf{w}})$$

$$\sigma_{\mathbf{w}} \sim \mathcal{HN}\left(\frac{\sigma_{\mathbf{w}}}{\lambda_{\mathbf{w}}}\right)$$

$$\mathbf{c}_{\mathbf{w}} \sim \mathcal{N}(0, 1)$$

$$\mathbf{w} = w_0 + \sigma_{\mathbf{w}} \odot \mathbf{c}_{\mathbf{w}}$$

$$\delta_0 \sim \mathcal{N}(\mu_{\Delta}, \sigma_{\Delta})$$

$$\sigma_{\Delta} \sim \mathcal{HN}\left(\frac{\sigma_{\Delta}}{\lambda_{\Delta}}\right)$$

$$\mathbf{C}_{\Delta} \sim \text{Laplace}(0, 1)$$

$$\Delta = \delta_0 + \sigma_{\Delta} \odot \mathbf{C}_{\Delta}$$

$$\theta_0^{(j)} \sim \mathcal{N}(\mu_{\theta^{(j)}}, \sigma_{\theta^{(j)}})$$

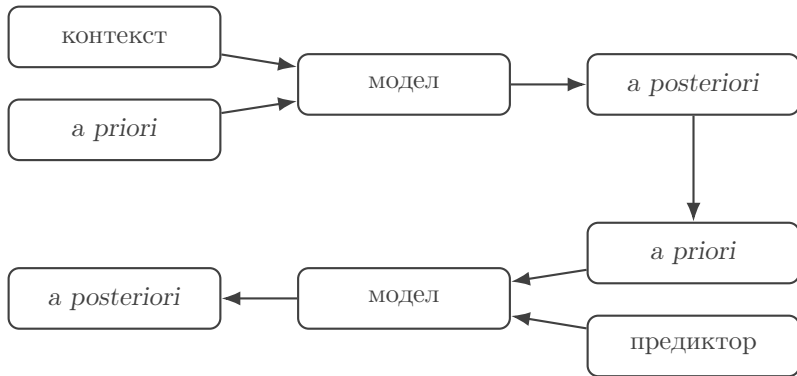
$$\sigma_{\theta^{(j)}} \sim \mathcal{HN}\left(\frac{\sigma_{\theta^{(j)}}}{\lambda_{\theta^{(j)}}}\right)$$

$$\mathbf{C}_{\theta^{(j)}} \sim \mathcal{N}(0, 1)$$

$$\theta^{(j)} = \theta_0^{(j)} + \sigma_{\theta^{(j)}} \odot \mathbf{C}_{\theta^{(j)}}$$

a posteriori \rightarrow *a priori*

5 Vangja



Иста форма на *a priori* распределбата

5 Vangja

контекстен *a posteriori* → предиктор *a priori*

контекстен *a posteriori* → глобално *a priori*

$$w_1^{MAP} = \operatorname{argmax}_{w_1} P(w_1|C)$$

$$\sigma_{w_1|C} = \sqrt{\mathbb{V}(w_1|C)}$$

$$w_i \sim \mathcal{N}(w_1^{MAP}, \sigma_{w_1|C})$$

$$\delta_1^{MAP} = \operatorname{argmax}_{\delta_1} P(\delta_1|C)$$

$$\sigma_{\delta_1|C} = \sqrt{\mathbb{V}(\delta_1|C)}$$

$$\delta_i \sim \text{Laplace}(\delta_1^{MAP}, \frac{\sigma_{\delta_1|C}}{\sqrt{2}})$$

$$\theta_1^{(j)MAP} = \operatorname{argmax}_{\theta_1^{(j)}} P(\theta_1^{(j)}|C)$$

$$\sigma_{\theta_1^{(j)}|C} = \sqrt{\mathbb{V}(\theta_1^{(j)})}$$

$$\theta_i^{(j)} \sim \mathcal{N}(\theta_1^{(j)MAP}, \sigma_{\theta_1^{(j)}|C})$$

MvN апроксимација на *a posteriori*

5 Vangja

- Ги моделира и корелациите помеѓу параметрите
- `prior_from_idata`

$$\boldsymbol{\mu}^C = \left(\mathbb{E}(w_1|C), \mathbb{E}(b_1|C), \mathbb{E}(\delta_{1,1}|C), \dots, \mathbb{E}(\beta_{1,1}^{(1)}|C), \dots, \mathbb{E}(\sigma_\epsilon|C) \right)^T$$

$$\boldsymbol{\Sigma}^C = \begin{pmatrix} \mathbb{V}(w_1|C) & Cov(w_1, b_1|C) & Cov(w_1, \delta_{1,1}|C) & \dots & Cov(w_1, \sigma_\epsilon|C) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Cov(\sigma_\epsilon, w_1|C) & Cov(\sigma_\epsilon, b_1|C) & Cov(\sigma_\epsilon, \delta_{1,1}|C) & \dots & \mathbb{V}(\sigma_\epsilon|C) \end{pmatrix}$$

$$(w_i, b_i, \delta_{i,1}, \dots, \beta_{i,1}^{(1)}, \dots, \sigma_\epsilon) \sim \mathcal{N}(\boldsymbol{\mu}^C, \boldsymbol{\Sigma}^C)$$

Регуларизација

5 Vangja

- Слични иницијални растови
- Слични норми на сезоналностите

$$\begin{aligned} \log(P(\gamma_i, \boldsymbol{\theta}_i^{(1)}, \dots, \boldsymbol{\theta}_i^{(M)}, \sigma_\epsilon | X_i)) &\propto \log(P(\gamma_i, \boldsymbol{\theta}_i^{(1)}, \dots, \boldsymbol{\theta}_i^{(M)}, \sigma_\epsilon)) \\ &\quad + \log(P(X_i | \gamma_i, \boldsymbol{\theta}_i^{(1)}, \dots, \boldsymbol{\theta}_i^{(M)}, \sigma_\epsilon)) \\ &\quad - \phi_{\mathbf{w}}(w_i - w_1^{MAP})^2 \\ &\quad + \sum_{j=1}^M \phi_{\boldsymbol{\theta}^{(j)}} \cdot \min(0, \|\mathbf{fs}_1^{(j)MAP}\|_2^2 - \|\mathbf{fs}_i^{(j)}\|_2^2) \end{aligned}$$

Содржина

6 Методологија

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet

- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

Податочно множество

6 Методологија

- Дневни берзански податоци - типична цена
- Контекстна временска серија: S&P 500 индекс, 40 години
- Предиктор временски серии: 500-те композитни акции, 3 месеци
- 730 последователни прозорци: од 01.01.2015 до 01.01.2017

Модели

6 Методологија

- Holt-Winters Exponential Smoothing
- Unobserved Components
- Facebook Prophet
- Timeseers
- Vangja со хиерархиско моделирање
- Vangja контекстен *a posteriori* → предиктор *a priori*
- Vangja контекстен *a posteriori* → глобално *a priori*

Содржина

7 Резултати

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet

- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

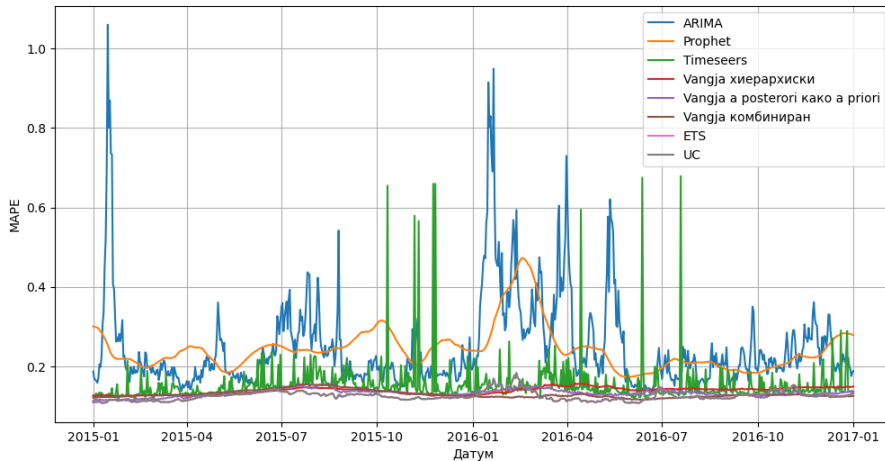
Метрики

7 Резултати

	MSE	RMSE	MAE	MAPE	↓
<i>ARIMA</i>	0.5400	0.2934	0.2439	0.2486	/
<i>Prophet</i>	0.1147	0.2696	0.2335	0.2421	2.61%
<i>Timeseers</i>	0.0505	0.1801	0.1535	0.1541	36.35%
<i>Vangja</i> хиерархиски	0.0413	0.1640	0.1424	0.1384	10.19%
<i>Vangja</i> предиктор <i>a priori</i>	0.0380	0.1583	0.1342	0.1324	4.34%
<i>Vangja</i> глобално <i>a priori</i>	0.0348	0.1493	0.1284	0.1297	2.04%
<i>ETS</i>	0.0357	0.1518	0.1285	0.1253	3.39%
<i>UC</i>	0.0357	0.1518	0.1285	0.1253	0%

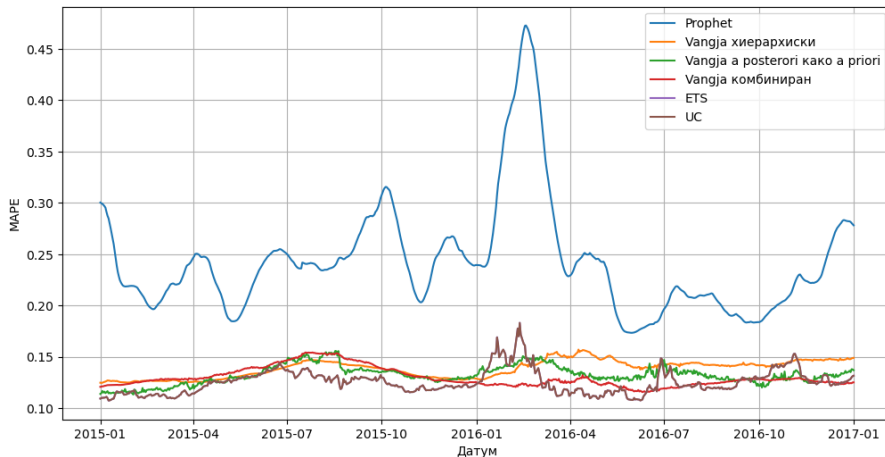
Визуелизација на МАРЕ

7 Резултати



Визуелизација на МАРЕ

7 Резултати



Хиперпараметри кај Vangja хиерархиски

7 Резултати

Стеснување	MSE	RMSE	MAE	MAPE
1000000	0.0413	0.1640	0.1424	0.1384
100000	0.0413	0.1640	0.1424	0.1384
10000000	0.0413	0.1640	0.1425	0.1384
10000	0.0441	0.1762	0.1496	0.1484
1000	0.0755	0.2404	0.1967	0.1995
100	0.4180	0.4625	0.3681	0.3748
1	3.1461	1.4366	1.0810	1.1251
10	9.6955	2.1440	1.5652	1.6445

Хиперпараметри кај Timeseers

7 Резултати

Стеснување	MSE	RMSE	MAE	MAPE
10000	0.0505	0.1801	0.1535	0.1541
1000000	0.0707	0.2008	0.1734	0.1749
10000000	0.0712	0.2018	0.1741	0.1758
100000	0.0725	0.2037	0.1758	0.1777
1000	0.0730	0.2305	0.1910	0.1971
100	0.2792	0.4236	0.3344	0.3497
1	1.2220	1.0768	0.8894	0.9073
10	3.7595	1.5247	1.1208	1.1746

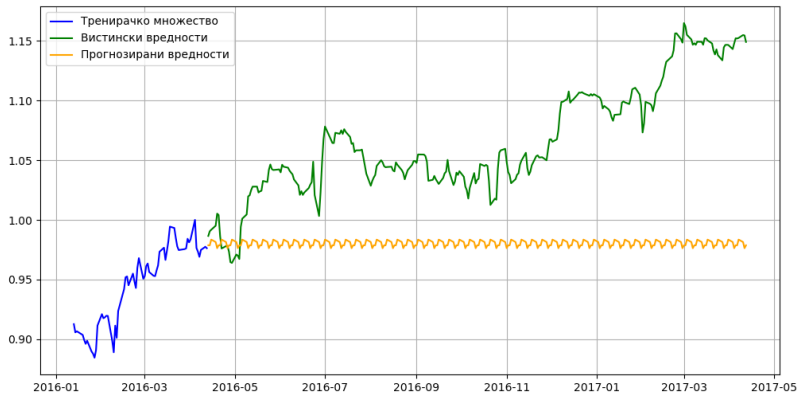
Хиперпараметри кај Vangja предиктор *a priori*

7 Резултати

Тренд	Сезоналност	λ_w	MSE	RMSE	MAE	MAPE
форма	prior_from_idata	1	0.0380	0.1583	0.1342	0.1324
форма	prior_from_idata	0	0.0386	0.1594	0.1353	0.1330
форма	prior_from_idata	-1	0.0392	0.1607	0.1365	0.1339
prior_from_idata	prior_from_idata	1	0.0374	0.1571	0.1337	0.1353
prior_from_idata	prior_from_idata	0	0.0374	0.1571	0.1337	0.1353
prior_from_idata	prior_from_idata	-1	0.0374	0.1571	0.1337	0.1353
форма	форма	1	0.0394	0.1609	0.1373	0.1360
форма	форма	0	0.0399	0.1618	0.1381	0.1364
форма	форма	-1	0.0401	0.1623	0.1386	0.1368
prior_from_idata	форма	1	0.0380	0.1582	0.1352	0.1368
prior_from_idata	форма	0	0.0380	0.1582	0.1352	0.1368
prior_from_idata	форма	-1	0.0380	0.1582	0.1352	0.1368

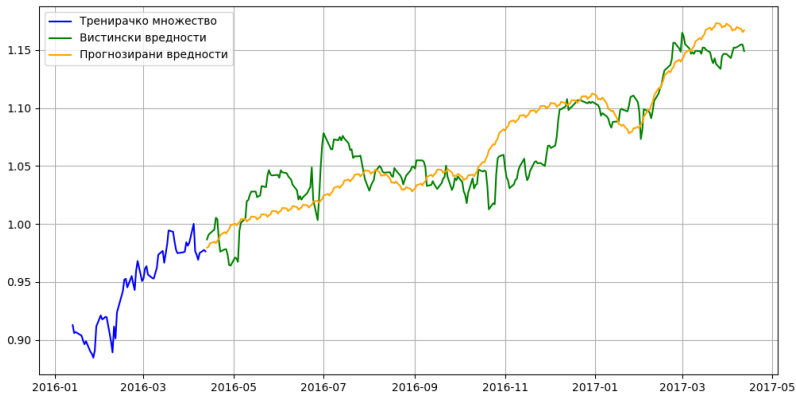
Визуелизација на прогнози

7 Резултати



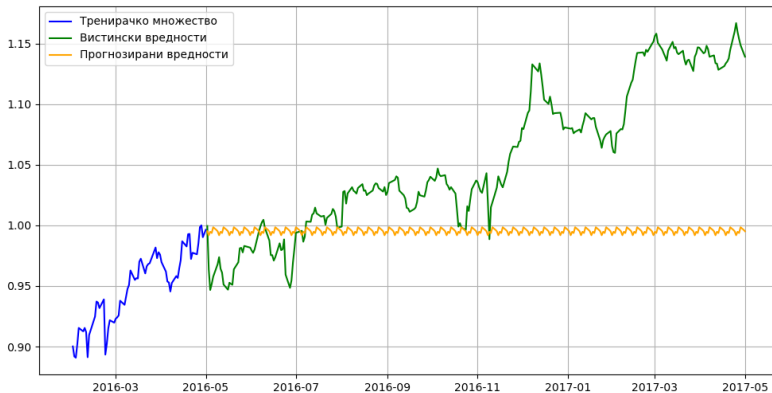
Визуелизација на прогнози

7 Резултати



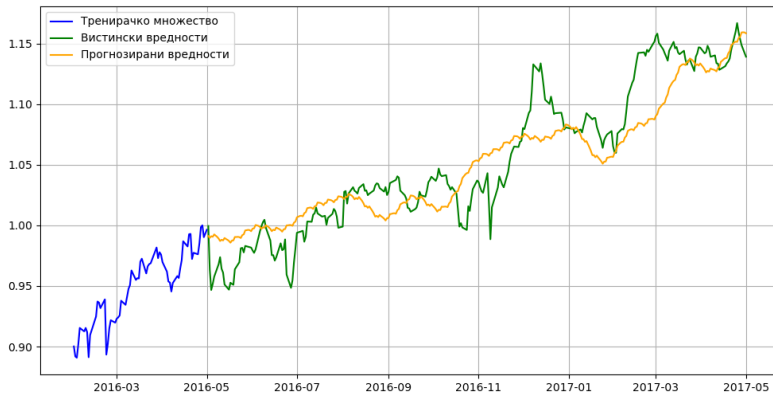
Визуелизација на прогнози

7 Резултати



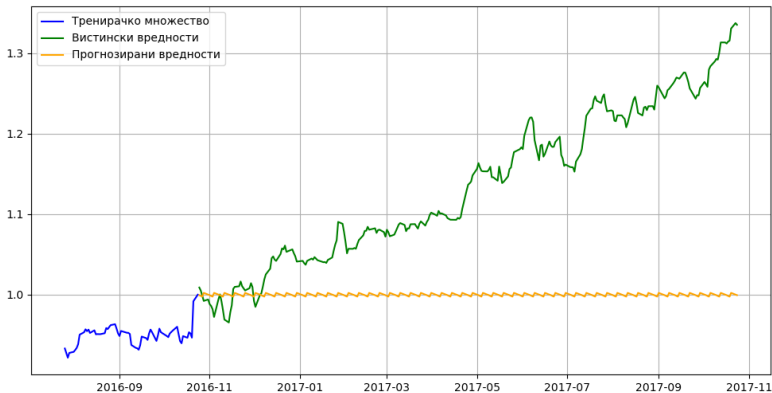
Визуелизација на прогнози

7 Резултати



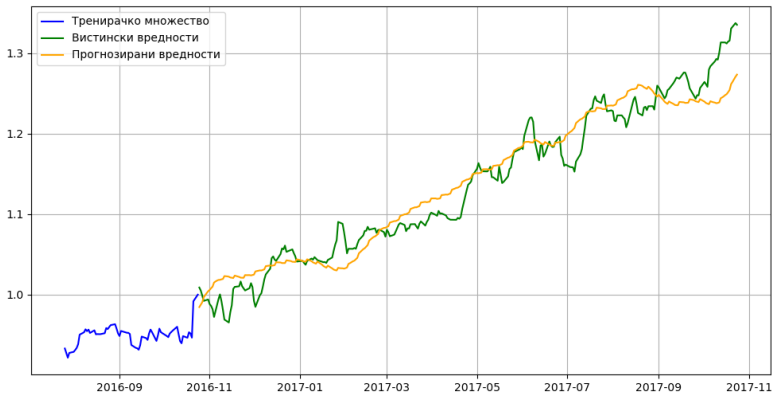
Визуелизација на прогнози

7 Резултати



Визуелизација на прогнози

7 Резултати



Содржина

8 Заклучок

- ▶ Дефиниција на проблемот
- ▶ Теоретска рамка
- ▶ *A priori* распределби
- ▶ Facebook Prophet

- ▶ Vangja
- ▶ Методологија
- ▶ Резултати
- ▶ Заклучок

Заклучок

8 Заклучок

- *a posteriori* \rightarrow *a priori* - соодветна замена за хиерархиско моделирање
 - пресметковна ефикасност
- Информирање на слабо-информираните *a priori* распределби
 - μ - преносно учење
 - σ, b - дискриминативна L2 и L1 регуларизација по параметар!
- Регуларизацијата конзистентно подобрува, но е занемарлива


Понатамошна работа

8 Заклучок

- Надоградба на Vangja со повеќе компоненти
 - логистички тренд (Facebook Prophet)
 - константи (Timeseers)
 - Radial basis functions за сезоналност (Timeseers)
 - авто-регресивност (Neural Prophet)
- Обоштена парадигма за Баесово преносно учење
- Емпириско информирање на *a priori* верувања

Репродукција на резултатите

8 Заклучок

 github.com/jovan-krajevski/magisterska

 github.com/jovan-krajevski/vangja

Q&A

*Благодариме на вниманието!
Секоја забелешка е добредојдена!*