بهینه‌سازی پیش‌بینی میزان پخت غذا در سلف‌های دانشگاه تهران

ارائه به استاد راهنما

مهدی وجهی

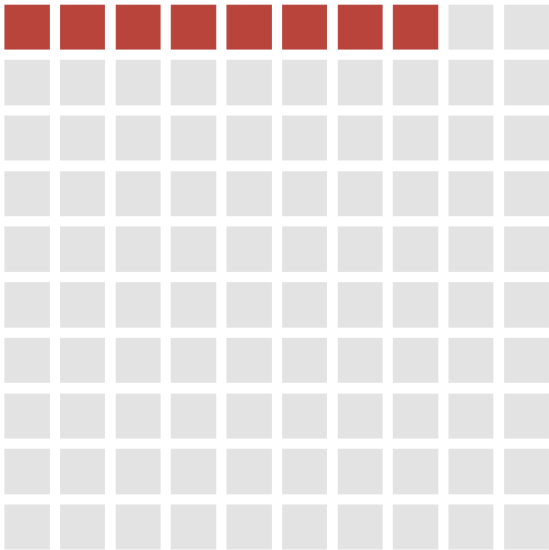
شهریور ۱۴۰۵

فهرست مطالب

- چرا این مسئله مهم است
- بازتعریف مسئله و مبانی نظری
- آنچه داده گفت
- مدل
- چقدر صرفه‌جویی
- محدودیت و گام بعد

از هر ۱۰۰ پرس، ۸ تا دور ریخته می‌شود

۱ چرا این مسئله مهم است



- 8.05% رزروها تحویل گرفته نمی‌شوند
- سالانه حدود ۲۳۲ هزار پرس، حدود ۸۷ میلیارد تومان به قیمت شهریور ۱۴۰۵

چرا تا حالا حل نشده

۱ چرا این مسئله مهم است

- پخت روی «تعداد رزرو» بسته می‌شود؛ عددی مطمئن ولی همیشه بیش‌برآورد
- کمتر پختن هدررفت را کم می‌کند ولی ریسک کمبود دارد
- هزینه‌ی این دو خطا برابر نیست، و تا وقتی این نامتقارنی کمّی نشود هیچ مدیری جرئت کم‌کردن ندارد

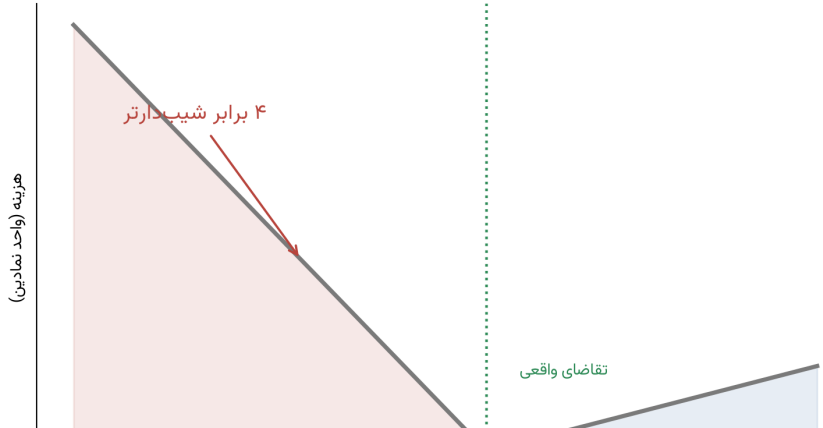
فهرست مطالب

۲ بازتعریف مسئله و مبانی نظری

- چرا این مسئله مهم است
- بازتعریف مسئله و مبانی نظری
- آنچه داده گفت
- مدل
- چقدر صرفه‌جویی
- محدودیت و گام بعد

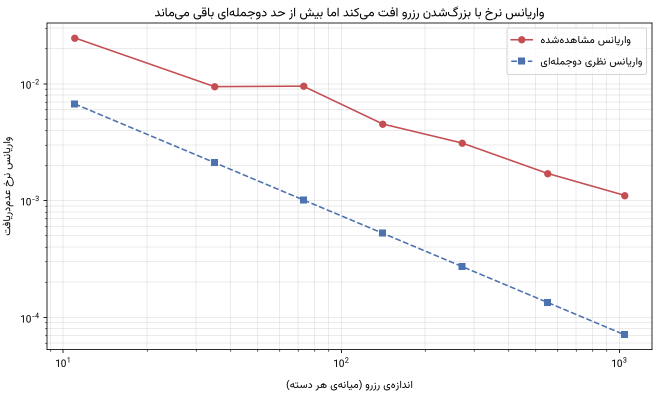
مسئله‌ی روزنامه‌فروش

- ۲۰۰ نفر رزرو کرده‌اند. اگر ۲۰۰ پرس بپزی و ۱۵ نفر نیایند، ۱۵ پرس دور ریخته می‌شود، حدود 5.6 میلیون تومان
- اگر ۱۸۵ پرس بپزی و ۱۹۰ نفر بیایند، ۵ دانشجو غذا ندارند — کدام بدتر است؟
- جواب کلاسیک (Newsvendor): $\tau^* = \frac{C_u}{C_u + C_o}$ — مقدار بهینه میانگین تقاضا نیست، چندک بحرانی توزیع تقاضاست



شکست فرمول‌های کلاسیک: نوسان واقعی تا 15.6 برابر فرضیات نظری است

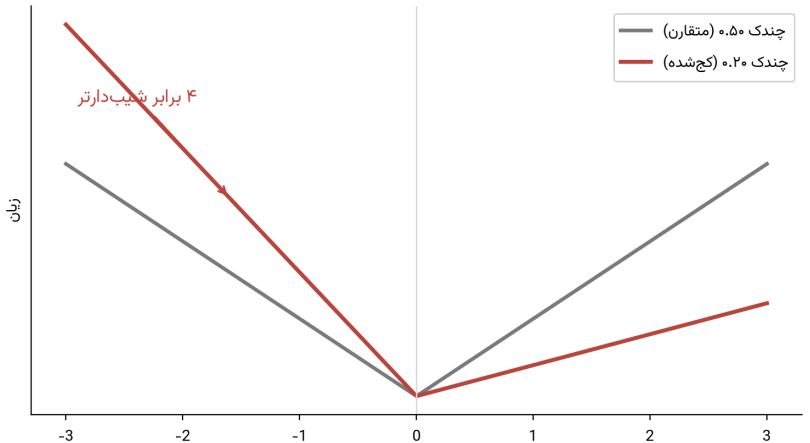
- **رد فرض استقلال افراد:** فرمول‌های تحلیلی غیبت را مستقل می‌دانند (توزیع دوجمله‌ای)، اما رفتار جمعی نوسان را تا 15.6 برابر فراتر از تئوری می‌برد
- **ماندگاری نوسان پس از کنترل ساختاری:** حتی با کنترل سلف، وعده، تقویم و روز هفته، همچنان 7.6 برابر نوسان اضافه باقی می‌ماند
- **پیامد راهبردی:** در میان ۷,۵۷۹ سلول ناهمگن، چندک بهینه فرمول بسته ندارد و باید **مستقیماً از داده آموخته شود**



نیوزوندور داده‌محور و رگرسیون کوانتایل

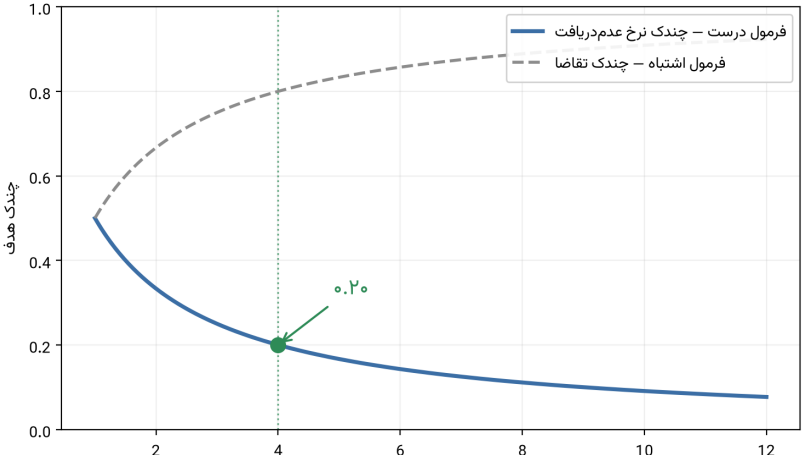
۲ بازتعریف مسئله و مبانی نظری

- به‌جای «اول توزیع را تخمین بزن، بعد بهینه کن»، مستقیماً تصمیم از داده یاد گرفته می‌شود (Ban & Rudin ۲۰۱۹؛ Huber و همکاران ۲۰۱۹)
- ابزار: رگرسیون کوانتایل (Koenker & Bassett ۱۹۷۸) — تابع زیانی با وزن τ برای کم‌زدن و $1-\tau$ برای زیادزدن
- کمینه‌کردن این زیان و حل مسئله‌ی Newsvendor یک کار واحدند، نه دو کار پشت سر هم



یک تصحیح که مال خود پروژه است

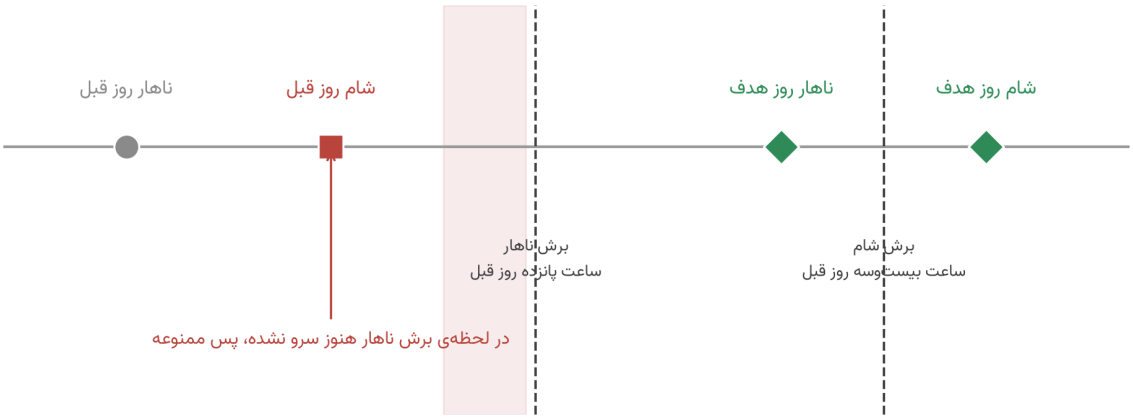
- فرمول کلاسیک در فضای **تقاضا** تعریف شده؛ خروجی این پروژه چندکی از **نرخ عدم‌دریافت** است که متمم تقاضاست: $D = Res(1-\rho)$
- پس فرمول برمی‌گردد: $\tau_\rho^* = \dfrac{C_o}{C_u + C_o}$
- تأیید عددی با شبیه‌سازی مستقیم نیوزوندور روی ۴۰۰ هزار نمونه



داده و قاعده‌ی برش

۲ بازتعریف مسئله و مبانی نظری

- ۲ میلیون رزرو فردی، ۱۴۲ روز سرو، ۳۰ سلف، ۶ شهر
- تصمیم پخت ساعت ۱۵ روز قبل برای ناهار و ۲۳ برای شام گرفته می‌شود — مدل فقط چیزی را می‌بیند که آن لحظه واقعاً در دسترس بوده



فهرست مطالب

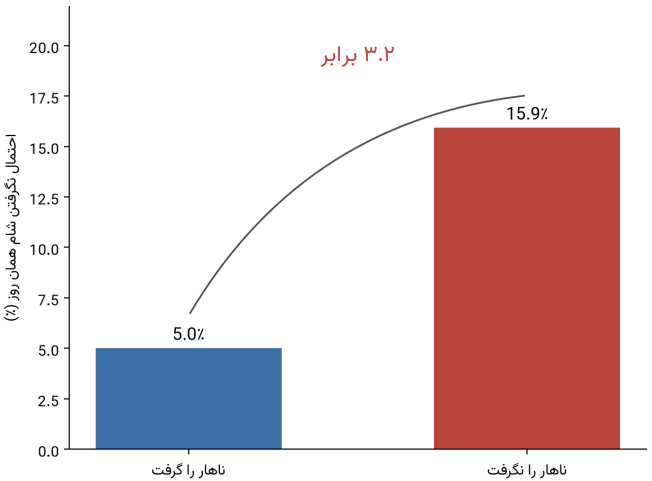
۳ آنچه داده گفت

- چرا این مسئله مهم است
- بازتعریف مسئله و مبانی نظری
- آنچه داده گفت
- مدل
- چقدر صرفه‌جویی
- محدودیت و گام بعد

عدم‌دریافت یک تصمیم غذایی نیست، غیبت از دانشگاه است

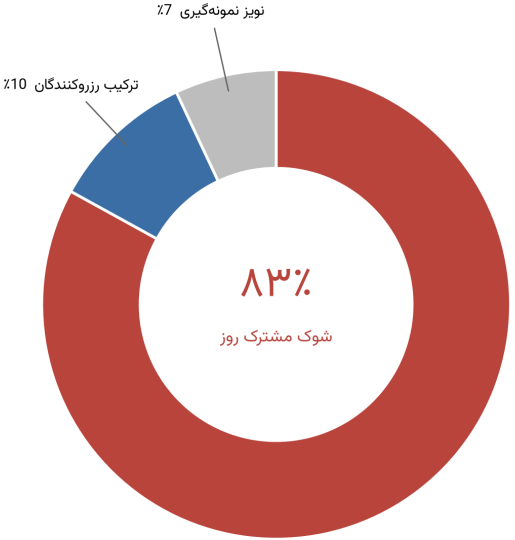
۳ آنچه داده گفت

- اگر کسی ناهارش را نگیرد، احتمال نگرفتن شامش 3.2 برابر می‌شود
- فرضیه‌ی «ترجیح غذایی شخصی» رد شد: پایداری جفت (فرد، غذا) 0.24 در برابر پایداری خود فرد 0.43
- **پیامد:** تغییر منو ابزار کاهش هدررفت نیست



83% نوسان از یک شوک مشترک روزانه می‌آید

۳ آنچه داده گفت

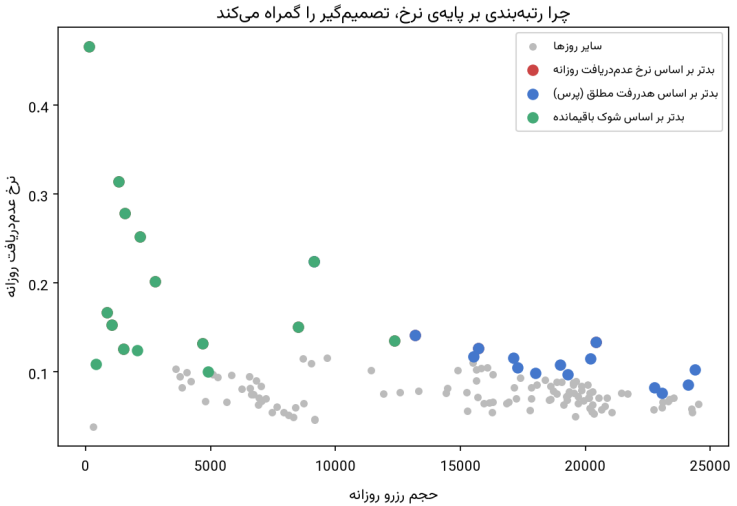


- تجزیه‌ی واریانس نرخ سلول: 7% نویز نمونه‌گیری، 10% ترکیب رزروکنندگان، 83% شوک مشترک روز
- در 99.5% جفت‌سلف‌ها این شوک هم‌جهت است — سراسری دانشگاه است نه محلی
- خبر خوب: با اطلاعات مجاز در لحظه‌ی برش تا $R^2 = 0.60$ پیش‌بینی‌پذیر است

روزهایی که بد به‌نظر می‌رسند، روزهایی نیستند که پول می‌سوزانند

۳ آنچه داده گفت

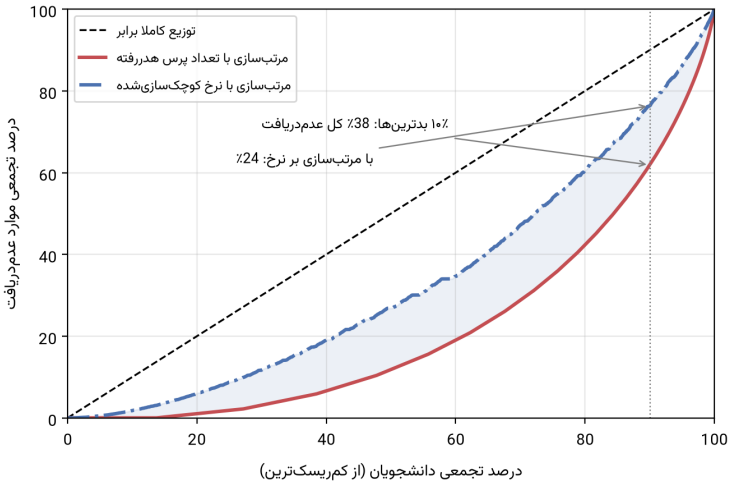
- همبستگی حجم رزرو با نرخ منفی است (-0.35) ولی با هدررفت مطلق شدیداً مثبت ($+0.75$)
- فهرست «بدترین روزها» با معیار نرخ و با معیار پرس فقط ۴ عضو از ۱۵ عضو مشترک دارند



هدف‌گیری دانشجویان پرغیبت جواب نمی‌دهد

۳ آنچه داده گفت

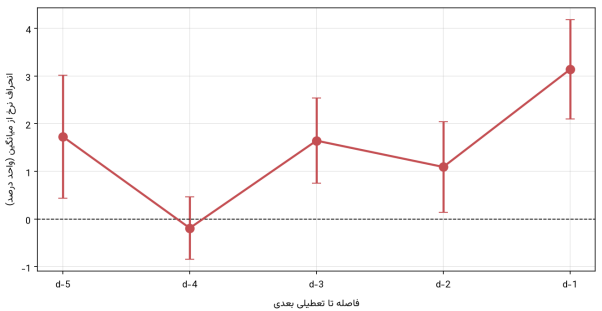
- سقف نظری هر سیاستی که 10% افراد را هدف بگیرد، 38% هدررفت است
- هدف‌گیری بر پایه‌ی مشخصات جمعیتی به 8.7% می‌رسد — بدتر از انتخاب تصادفی (10%)، چون افراد پرنرخ عمدتاً کم‌رزرواند

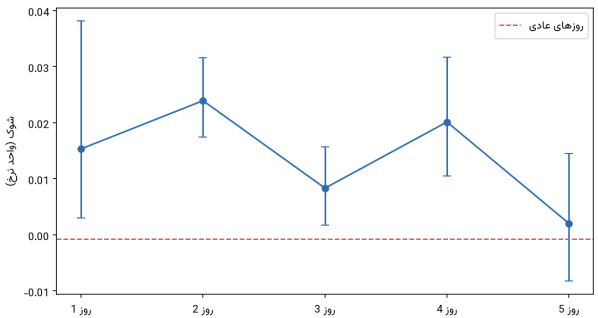


تقویم سیگنال رایگان می‌دهد

۳ آنچه داده گفت

- روز قبل از تعطیلی: 3.1 واحد درصد بالاتر، حدود 39% نسبی
- چهار روز اول بازگشت پس از وقفه‌ی چهارروزه یا بیشتر: 22% نسبی بالاتر — بدترین روز، روز **دوم** است نه اول
- شام رمضان تقریباً دو برابر

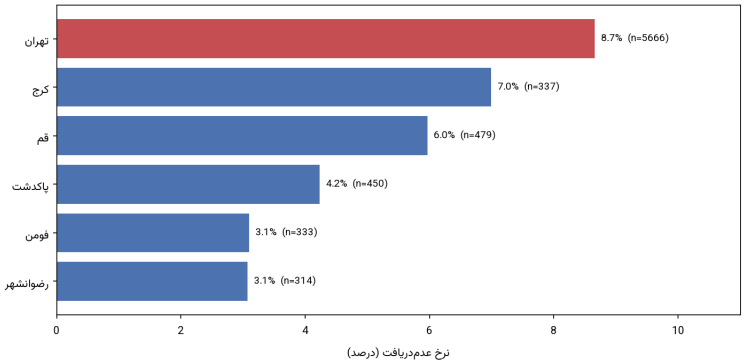




جغرافیا

۳ آنچه داده گفت

- تهران 8.7% در برابر رضوانشهر 3.1%، با تفکیک تقریباً کامل بین پردیس‌ها
- سازوکارش رفتار افراد است نه ترکیب سلف‌ها — در پردیس دورافتاده سلف تنها گزینه‌ی غذایی است
- **پیامد:** تمرکز منابع روی تهران



فهرست مطالب

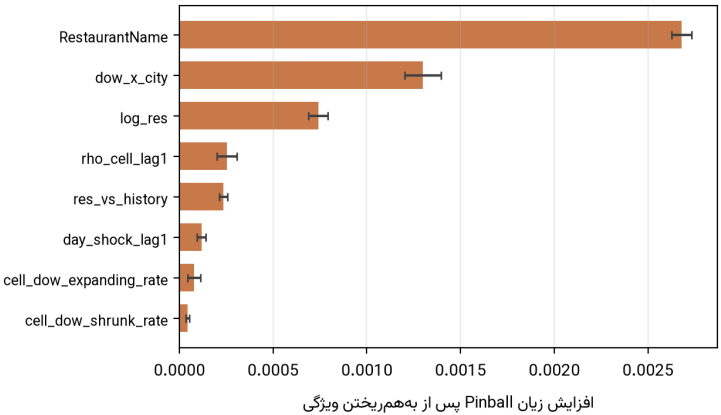
۴ مدل

- چرا این مسئله مهم است
- بازتعریف مسئله و مبانی نظری
- آنچه داده گفت
- مدل
- چقدر صرفه‌جویی
- محدودیت و گام بعد

چه ساخته شد و مدل به چه چیزی نگاه می‌کند

۴ مدل

- ۱۰۰ ویژگی، ۱۳ خانواده‌ی مدل، غربال چهارمرحله‌ای، انتخاب نهایی یک مدل درختی کوانتایلی
- پروتکل: تقسیم زمانی با مبدأ غلتان؛ پنجره‌ی آزمون فقط یک‌بار باز شد
- مهم‌ترین ورودی‌ها: هویت سلف، برهم‌کنش روزهفته با شهر، حجم رزرو، نرخ دیروزِ همان سلول، شوک دیروز



۱۳ خانواده‌ی مدل امتحان شدند

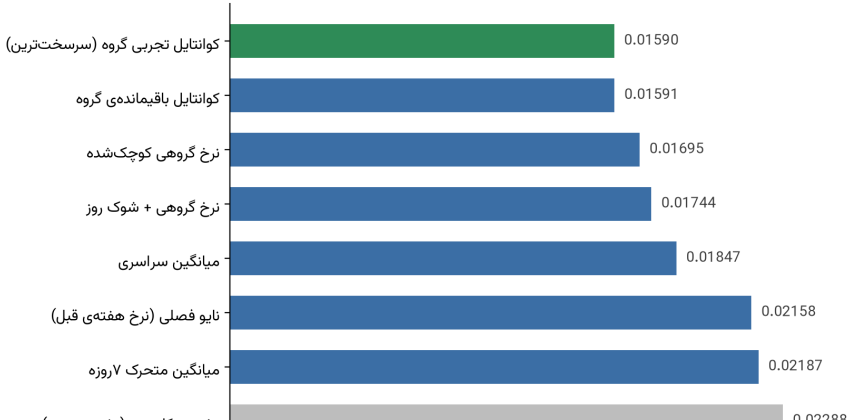
۴ مدل

- از رگرسیون خطی ساده تا شبکه‌ی عصبی و مدل‌های بیزی — طیف کامل، نه فقط چند حدس
- هر خانواده حداقل یک نماینده تا انتهای مسیر آموزش و ارزیابی رفت
- خانواده‌ی درختی/بوستینگ (خ۲) در نهایت مدل برتر را داد؛ ترکیب و کالیبراسیون (خ۱۲، خ۱۳) لایه‌ای روی همه‌اند، نه یک رقیب مستقل

خط پایه: معیاری که هر مدلی باید ازش بهتر باشد

۴ مدل

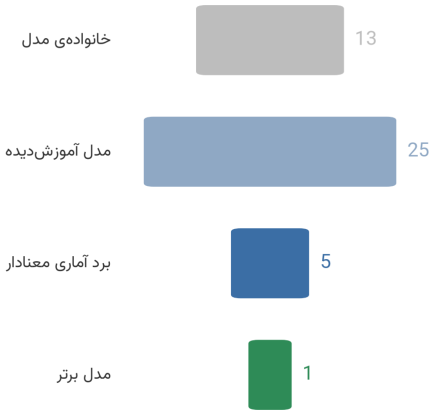
- خط پایه یک قاعده‌ی ساده و بدون یادگیری ماشین است — همان کاری که یک مدیر سلف بدون هیچ مدلی می‌تواند انجام دهد
- هشت خط پایه از ساده‌ترین (پخت = کل رزرو، وضع موجود) تا پیچیده‌ترین (کوانتایل گروهی) طیف را پوشش می‌دهند
- B3 (کوانتایل تجربی همان سلف×وعده×روزهفته) سرسخت‌ترین است و در بقیه‌ی این ارائه همان «قوی‌ترین خط پایه»ای است که مدل‌های آموزش‌دیده باید از آن جلو بزنند



نتایج منفی

۴ مدل

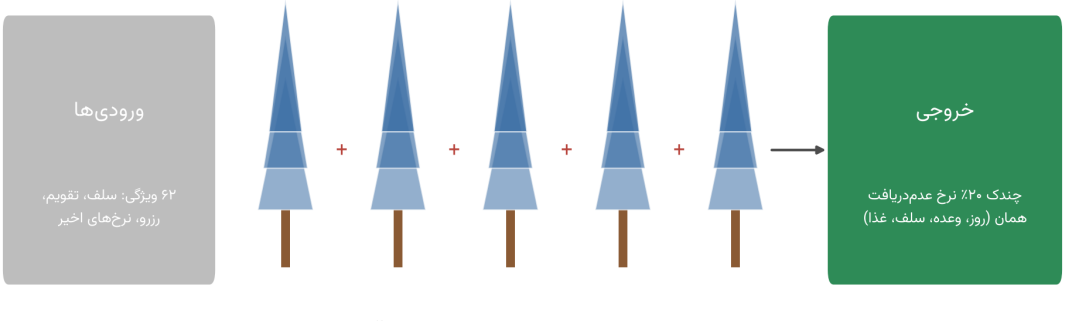
- از ۲۵ مدل فقط ۵ تا برد آماری معنادار داشتند
- شبکه‌های عصبی و مدل سطح فرد کمکی نکردند
- مدل برتر در اعتبارسنجی با مدل برتر در پنجره‌ی آزمون یکی نبود، و طبق پروتکل تغییر نکرد



مدل برتر چطور کار می‌کند

- یک LightGBM کوانتایلی: صدها درخت تصمیم پشت‌سرهم، هرکدام فقط خطای باقی‌مانده‌ی درخت‌های قبلی را اصلاح می‌کند
- خروجی مستقیماً همان چندک $\tau = 0.20$ است که برای مسئله انتخاب شده بود، نه یک میانگین که بعداً تبدیل شود
- روی foldهای رسمی از قوی‌ترین خط پایه (B3) به‌طور معنادار برد

هر درخت فقط خطای باقی‌مانده‌ی درخت‌های قبلی را اصلاح می‌کند



فهرست مطالب

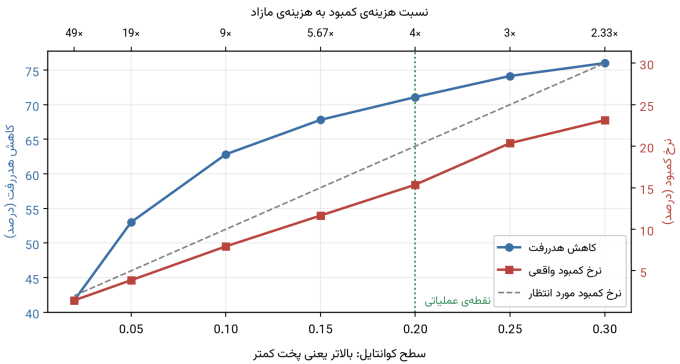
۵ چقدر صرفه‌جویی

- چرا این مسئله مهم است
- بازتعریف مسئله و مبانی نظری
- آنچه داده گفت
- مدل
- چقدر صرفه‌جویی
- محدودیت و گام بعد

τ به‌عنوان اهرم سیاستی: سه سناریو

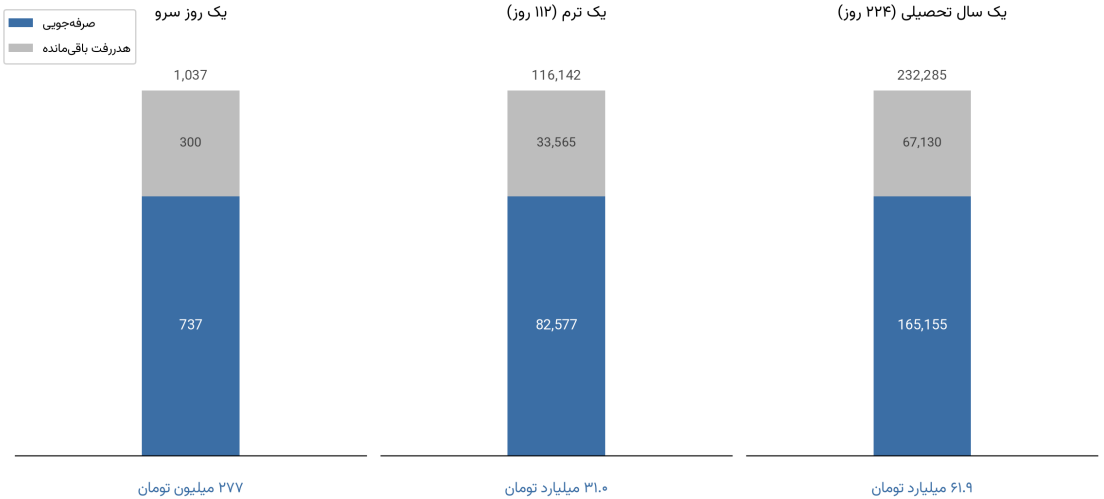
۵ چقدر صرفه‌جویی

سناریو	τ	نرخ کمبود سلول	تقاضای برآورده‌نشده	صرفه‌جویی سالانه
محافظه‌کارانه	0.02	1.4%	0.02%	36.3 میلیارد
متعادل	0.10	7.9%	0.14%	54.7 میلیارد
تهاجمی	0.20	15.4%	0.28%	61.9 میلیارد



نردبان: از یک روز تا یک سال تحصیلی

۵ چقدر صرفه‌جویی

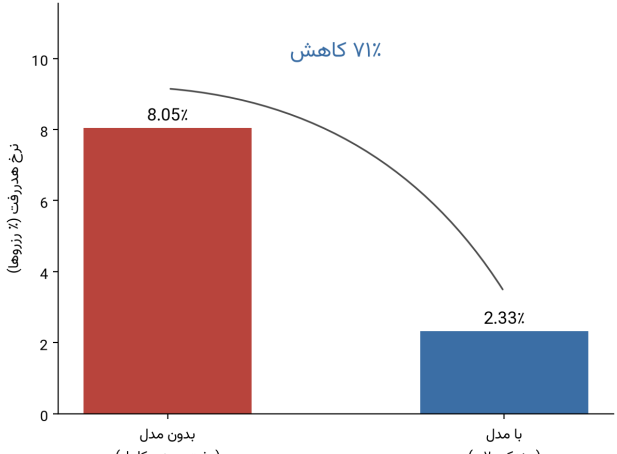


مبنا: 48.5 ردیف در هر روز سرو؛ 15.202 پرس صرفه‌جویی به‌ازای هر ردیف؛ ۳۲ هفته‌ی آموزشی؛ ۳۷۵ هزار تومان هر پرس. تابستان و نوروز بیرون گذاشته شده‌اند.

چند درصد از هدررفت جلوگیری می‌شود

۵ چقدر صرفه‌جویی

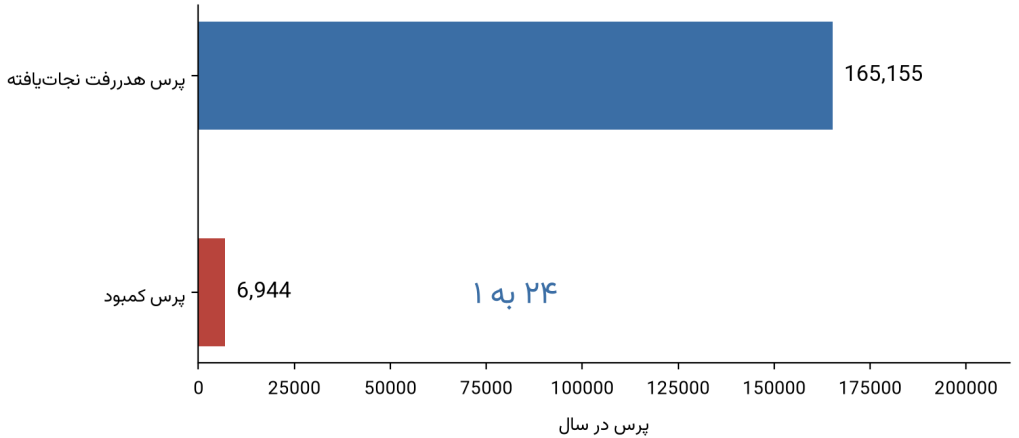
- 71% **هدررفت حذف می‌شود،** 29% **می‌ماند** — باقی‌مانده حدود ۶۷ هزار پرس در سال، یعنی 25.2 میلیارد تومان
- برحسب نرخ عدم‌دریافت: از 8.05% رزروها به 2.33% می‌رسد
- سقف تئوریک (پیش‌بینی کامل و بی‌خطا) 100% است؛ فاصله تا آن سقف عمداً برای اجتناب از کمبود گذاشته شده، نه ضعف مدل



طرف دیگر معامله: کمبود

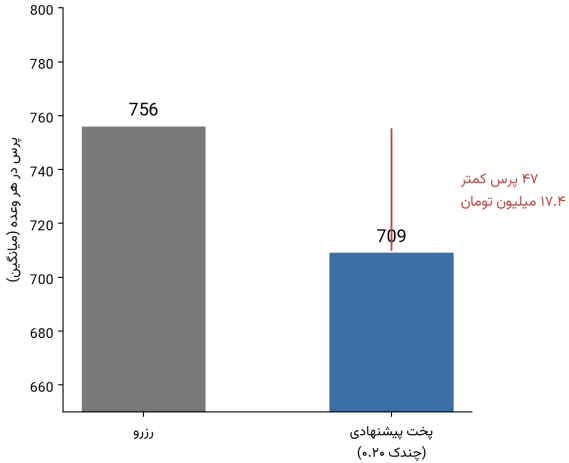
۵ چقدر صرفه‌جویی

- 15.4% سلول‌ها کمبود می‌خورند، ولی **عمق میانگین هر کمبود** 4.2 **پرس** است
- در یک وعده در کل دانشگاه: میانگین ۱۸ پرس، میانه ۸ پرس؛ سالانه حدود 0.28% تقاضا (سطح خدمت 99.72%)
- به ازای هر ۱ پرس کمبود، حدود **۲۴ پرس** هدررفت جلوگیری می‌شود



یک سلف مشخص: فنی امیرآباد

۵ چقدر صرفه‌جویی



- بزرگ‌ترین سلف دانشگاه: میانگین **۷۵۶ رزرو در هر وعده** ($\times 39$ کوچک‌ترین سلف)
- میانگین پیش‌بینی مدل در $\tau = 0.20$: نرخ عدم‌دریافت 6.2%
- یعنی دستور پخت **۷۰۹ پرس به‌جای ۷۵۶**، حدود ۴۷ پرس کمتر — حدود 17.4 میلیون تومان در هر وعده، حدود 3.9 میلیارد تومان در سال

سلف	وعده	رزرو	دریافت واقعی	پخت پیشنهادی	نتیجه
کشاورزی	شام	۱۰۷	۹۶	۱۰۱	۵ پرس صرفه‌جویی، بدون کمبود
علوم اجتماعی	ناهار	۲۶۶	۲۴۰	۲۴۷	۷ پرس صرفه‌جویی، بدون کمبود
تربیت بدنی	ناهار	۱۶۰	۱۵۲	۱۴۹	کمبود ۳ پرس

فهرست مطالب

۶ محدودیت و گام بعد

- چرا این مسئله مهم است
- بازتعریف مسئله و مبانی نظری
- آنچه داده گفت
- مدل
- چقدر صرفه‌جویی
- محدودیت و گام بعد

محدودیت‌ها، با همان دقت نتایج

۶ محدودیت و گام بعد

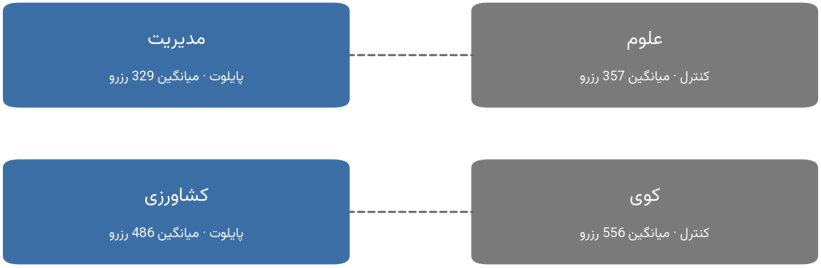
- پنج ماه داده، بدون تابستان و بدون سیگنال چندترمی
- قیمت ۳۷۵ هزار تومان نرخ امروز است، نه نرخ بازه‌ی داده — و هیچ آزمون میدانی انجام نشده
- هزینه‌ی کمبود تثبیت نشده، یعنی τ نهایی یک تصمیم مدیریتی است نه خروجی مدل؛ نردبان بخش پیش روی فرض تکرار الگوی هفتگی در ۳۲ هفته بسته شده

۷ از ۱۲ ماه سال در داده دیده شده — تابستان و نوروز نه

پایلوت پیشنهادی

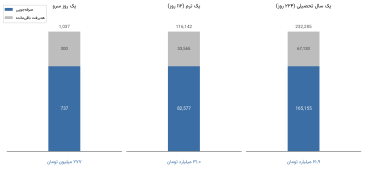
۶ محدودیت و گام بعد

- دو جفت همتاشده از دو خوشه‌ی پایدار: مدیریت/علوم و کشاورزی/کوی
- چهار هفته، ۷۲ مشاهده در هر بازو
- این آزمونِ «شدنی‌بودن عملیاتی» است، نه آزمونی با توان آماری کافی



جمع‌بندی در سه جمله

۶ محدودیت و گام بعد



۱. هر روز سرو، حدود هزار پرس دور ریخته می‌شود؛ سالانه ۲۳۲ هزار پرس و ۸۷ میلیارد تومان
۲. با یک تغییر تصمیم — پخت روی چندک به‌جای پخت روی رزرو — 71% آن برمی‌گردد (61.9 میلیارد تومان در سال) به قیمت 0.28% تقاضای برآورده‌نشده، یعنی ۲۴ پرس نجات‌یافته به ازای هر ۱ پرس کمبود
۳. τ تنها دکمه‌ای است که باید دست مدیریت باشد

سؤال و پاسخ

با تشکر از توجه شما

اسلایدهای پشتیبان

شماره‌گذاری‌شده تا سریع بپری رویشان

پ۱ — تابع زیان Pinball

$$\rho_\tau(u) = u\,(\tau - \mathbb{1}[u < 0])\,, \qquad u = y - \hat{y}$$

- برای $u \geq 0$ (کم‌زدن): شیب τ؛ برای $u < 0$ (زیادزدن): شیب $1 - \tau$
- کمینه‌کننده‌ی امید ریاضی این زیان دقیقاً چندک τاُم توزیع شرطی y است — اثبات با مشتق‌گیری از انتگرال تابع توزیع تجمعی

شاهد: Literature_Review.md بند ۲-۲.

پ۲ – چرا زیان با وزن حجم رزرو وزن‌دهی شد

- واریانس ρ به‌شدت به اندازه‌ی رزرو وابسته است: std چارک کوچک ÷ بزرگ $\times 3.03 =$
- بدون وزن‌دهی، سلف‌های کوچکِ پرنوسان بر تابع زیان اثر نامتناسب می‌گذارند
- راه‌حل: $w = Res$ در تابع زیان

شاهد: F06؛ ردیف ۲۵ ؛decision_log یافته‌ی ۱۶.

پ۳ – جدول کامل تبدیل نسبت هزینه به چندک

C_u/C_o	چندک تقاضا	τ_ρ (چندک نرخ عدم‌دریافت)
$2\times$	0.67	0.33
$4\times$	0.80	0.20
$9\times$	0.90	0.10

$$\tau_\rho = \frac{C_o}{C_u + C_o} = 1 - \frac{C_u}{C_u + C_o}$$

شاهد: reports/tau_sensitivity.md، reports/phase10/10.2_newsvendor.md.

پ۴ — قاعده‌ی برش اطلاعاتی، کامل

۷ اسلایدهای پشتیبان

مجاز	ممنوع
`Reservation` تا $d+2$	`ReceiveWithCard/Code`، `DontReceive` برای روز d
`Recv/NoRecv`، نرخ ρ تا لحظه‌ی برش همان وعده‌ی هدف	هر ستون پس از پخت، و هر وعده که هنوز سرو نشده
منو و ویژگی‌های تقویمی هر روز	—
`ReserveStatus` تاریخچه‌ی فرد p تا لحظه‌ی برش (مدل B)	`ReserveStatus` برای روز هدف d

تله‌ی رایج: شام روز $d-1$ در لحظه‌ی برش ناهار روز d (ساعت ۱۵) هنوز سرو نشده — با اینکه تاریخش $d-1$ است، استفاده از آن نشتی است. شاهد: AGENTS.md؛ بند ۴-۲ project-definition.md؛ ردیف ۹ decision_log.

پ۵ — هشت خط پایه روی $\tau = 0.20$

۷ اسلایدهای پشتیبان

خط پایه	pinball	نرخ کمبود	% کاهش هدررفت
B3_empirical_quantile	0.01590	9.6%	60.1%
B7_group_residual_quantile	0.01591	8.9%	59.8%
B2_group_shrunk	0.01695	9.8%	54.7%
B6_day_factor	0.01744	13.6%	58.1%
B1_global_mean	0.01847	9.9%	43.4%
B4_seasonal_naive	0.02158	11.7%	43.5%
B5_rolling_mean	0.02187	18.3%	57.8%
B0_cook_all	0.02288	0.0%	0.0%

شاهد: reports/baselines_tau020.md.

پ۶ — مدل برتر در برابر خط پایه، در نقطه‌ی سربه‌سر

- در $\tau = 0.20$ روی پنجره‌ی آزمون، مدل برتر از قوی‌ترین خط پایه (B3) عملاً چیزی جلو نزد: -0.011 میلیارد تومان
- مزیت مدل در τهای کوچک‌تر ظاهر می‌شود: در $\tau = 0.02$ برابر $+0.94$ میلیارد تومان
- بیشتر ارزش از خودِ سوئیچ به تصمیم چندکی می‌آید، نه از پیچیدگی مدل — مزیت واقعی مدل در کالیبراسیون است

شاهد: reports/phase10/10.2_newsvendor.md.

پ۷ — اندازه‌ی نمونه‌ی مؤثر و بوت‌استرپ بلوکی دوبعدی

۷ اسلایدهای پشتیبان

- اندازه‌ی نمونه: خام 7,579 سلول، مؤثر تنها 593 ICC) روز 0.225 =) — ضریب تورم واریانس 12.8×
- آزمون‌های آماری فاز ۷ باید روی این اندازه‌ی مؤثر، نه خام، تفسیر شوند
- بوت‌استرپ روی هر دو بُعد (روز و سلف) بلوک می‌بندد چون ICC روز و سلف تقریباً برابرند

شاهد: reports/baselines_tau020.md؛ F10.

پ۸ — بیش‌پراکندگی و رد فرض استقلال افراد

- در سطح فرد: واریانس گروهی $14.4\times$ دوجمله‌ای؛ $\chi^2/df = 18.2$
- ICC واحد سکونت $+0.064$ =، دانشکده $+0.033$ =
- **پیامد حیاتی:** جمع ساده‌ی احتمالات فردی در مدل B عدم‌قطعیت را حدود $3.8\times$ کم‌برآورد می‌کند — کوانتایل خوش‌بینانه، ریسک کمبود بیشتر از ادعای مدل

شاهد: F08.

پ۹ — پنج سطح داده

۷ اسلایدهای پشتیبان

سطح	واحد رکورد	حجم
L1 سلول	(d, m, r, f)	7,579
L2 سلف×وعده	(d, m, r)	حدود 4,100
L3 عامل روز	(d, m)	256
L4 سری زمانی	سری (m, r)	۴۱ سری، میانه ۱۹۷ نقطه
L5 رزرو فردی	(p, d, m, r, f)	2,049,322 رزرو

«مدل A/B» فازهای ۱ تا ۶: A ≡ L1/L2، B ≡ L5. شاهد: AGENTS.md.

پ۱۰ — کالیبراسیون و ACI

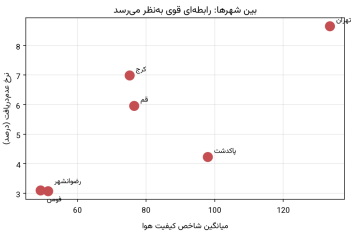
- بدون کالیبراسیون: پوشش تجربی سیستماتیک بیشتر از اسمی است — بیشترین شکاف در $\tau = 0.30$ (+18.0%)، یعنی مدل محافظه‌کارتر از حد اسمی
- در نقطه‌ی عملیاتی $\tau = 0.20$: پوشش از 33.6% (شکاف +13.6%) با ACI به 20.1% (شکاف +0.1%) می‌رسد
- pinball هم کمی بهتر می‌شود: از 0.01008 به 0.00975

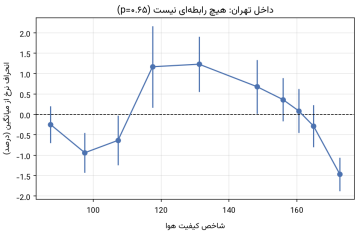
شاهد: reports/phase8/8.8_calibration.md.

پ۱۱ – ماجرای آلودگی هوا

۷ اسلایدهای پشتیبان

- اثر خام قوی: $r = +0.28$ $(p \approx 6 \times 10^{-134})$
- پس از کنترل سلف×وعده×روزهفته: -0.013 $(p = 0.25)$ – کاملاً صفر
- سازوکار: تهران هم‌زمان بالاترین AQI و بالاترین نرخ را دارد

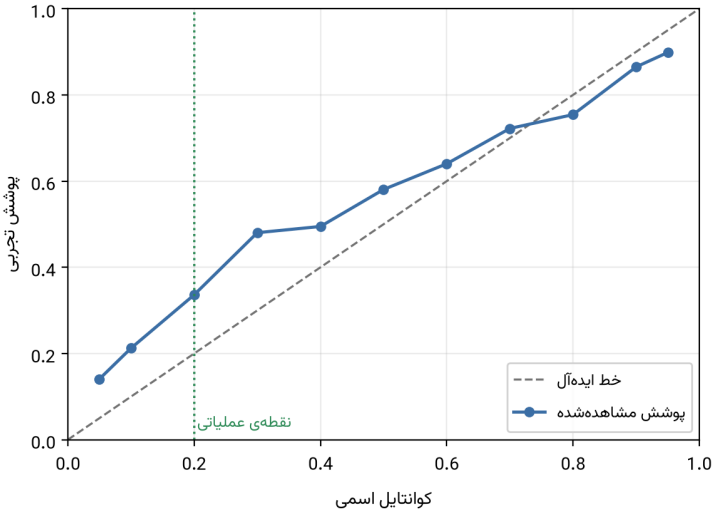




شاهد: F25.

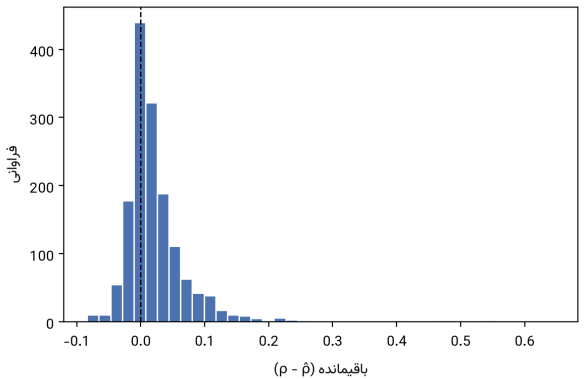
پ۱۲ — نمودار کالیبراسیون

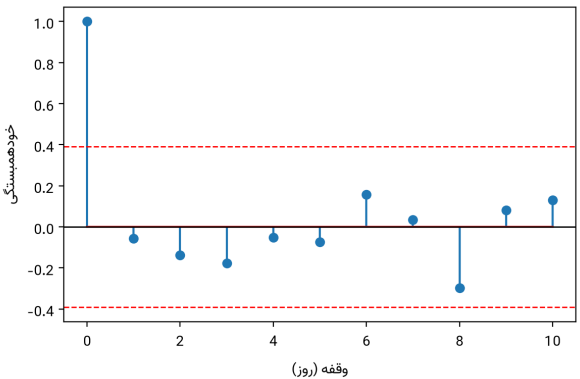
۷ اسلایدهای پشتیبان



پ۱۳ — تحلیل باقیمانده

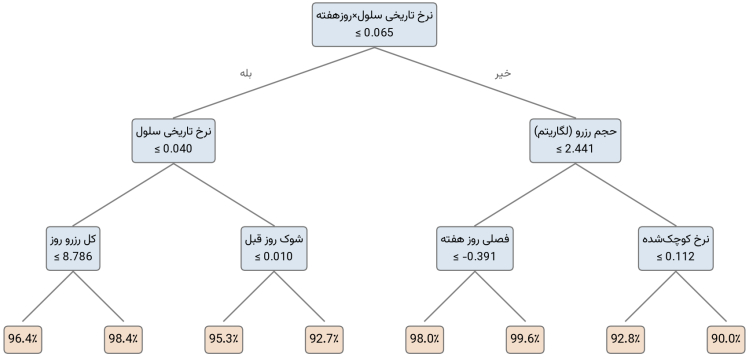
۷ اسلایدهای پشتیبان





پ۱۴ — درخت جانشین قابل‌فهم

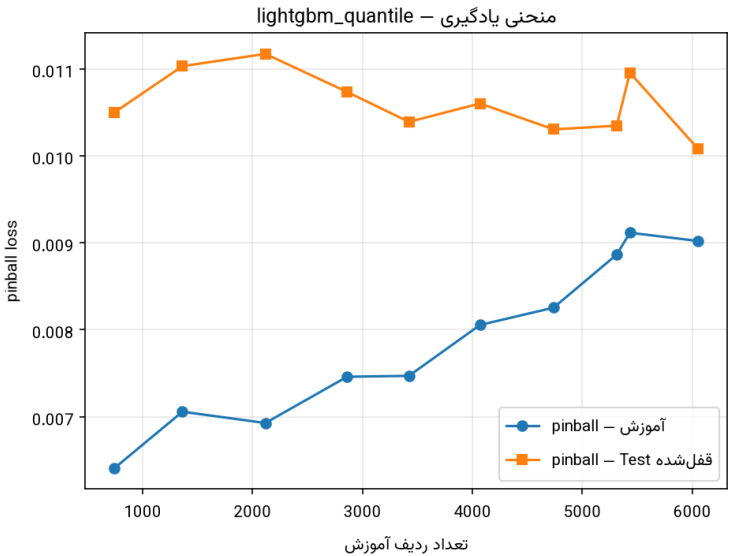
۷ اسلایدهای پشتیبان



عدد هر برگ: درصدی از رزرو که باید پخته شود

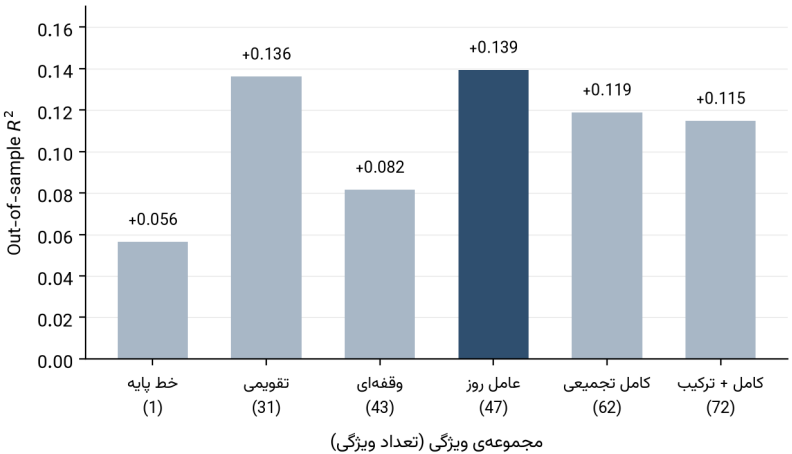
پ۱۵ — منحنی یادگیری

۷ اسلایدهای پشتیبان



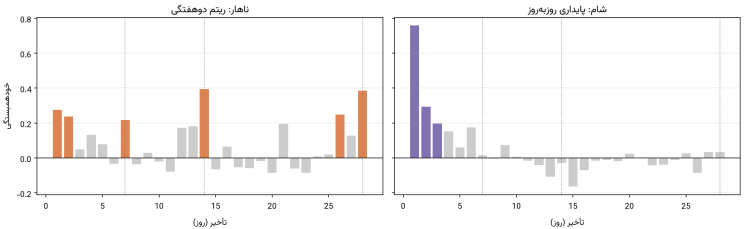
پ۱۶ — مقایسه‌ی شش مجموعه‌ی ویژگی

۷ اسلایدهای پشتیبان



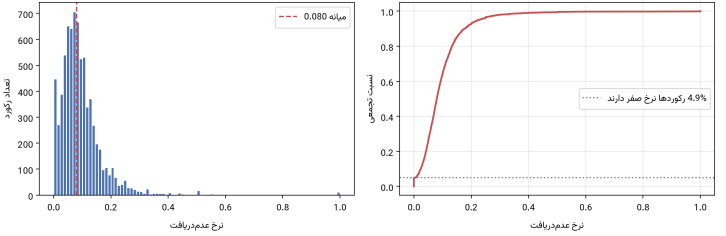
پ۱۷ – خودهمبستگی به تفکیک وعده

۷ اسلایدهای پشتیبان



پ۱۸ — توزیع متغیر هدف

۷ اسلایدهای پشتیبان



پرسش‌های محتمل

جواب کوتاه

«مدل پیچیده چه چیزی بیشتر از یک قاعده‌ی ساده داد؟»

خطرناک‌ترین سؤال — خودت اول مطرحش کن (اسلاید ۱۵).

- در $\tau = 0.20$ روی پنجره‌ی آزمون، مدل برتر از خط پایه‌ی B3 عملاً چیزی جلو نزد (-0.011 میلیارد)
- بیشتر ارزش از خودِ سوئیچ به تصمیم چندکی می‌آید، نه پیچیدگی مدل
- مزیت مدل در τهای کوچک‌تر ظاهر می‌شود ($\tau = 0.02$: $+0.94$ میلیارد) و در کالیبراسیون

«۶۲ میلیارد چقدر قابل اتکاست؟» / «چرا رگرسیون کوانتایل؟»

۷ اسلایدهای پشتیبان

- بازه‌ی 44 تا 81 میلیارد با بوت‌استرپ بلوکی، روی ۳۲ هفته‌ی آموزشی، بدون تابستان، بدون تورم، بدون رشد ثبت‌نام — فرض‌ها در 10.4_annual_savings.md
- تصمیم فقط به یک نقطه از توزیع نیاز دارد؛ تخمین کل توزیع خطای جاهای بی‌ربط را هم وارد تصمیم می‌کند

«معادل‌سازی Pinball با نیوزوندور از کجاست؟» / «چرا تک‌مرحله‌ای اجرا نشد؟»

- Rudin & Ban (۲۰۱۹)، Huber و همکاران (۲۰۱۹)، صورت‌بندی رسمی در Risks (۲۰۲۶) — همان مقاله هشدار می‌دهد کالیبراسیون باید جدا از دقت نقطه‌ای بررسی شود
- اجرا شد: خانواده‌ی ۱۱ (`f11_decision.py`) همین آزمون هم‌ارزی را دارد؛ نتیجه: هدف واقعی نسخه‌ی وزن‌دار است، نه pinball خام

«هزینه‌ی کمبود را چطور تعیین کردی؟» / «نشتی داده چطور کنترل شد؟»

- تعیین نشد — این یک تصمیم سیاستی است، نه کمیت قابل‌اندازه‌گیری از داده؛ کل بازه‌ی نسبت هزینه از ۲ تا ۴۹ آزموده شد. $\tau = 0.20$ توصیه است نه ادعا
- نشتی: قاعده‌ی برش به تفکیک وعده با ممیزی خودکار، فولدهای منجمد با دروازه‌ی هش — شام روز قبل در لحظه‌ی تصمیمِ ناهار فردا هنوز سرو نشده، پس ممنوع است

«چرا داده‌ی سطح فرد کمک نکرد؟» / «آلودگی هوا واقعاً بی‌اثر است؟»

۷ اسلایدهای پشتیبان

- ترکیب افراد فقط 10% واریانس را توضیح می‌دهد، شوک روز 83% — سیگنال فردی افزونه بود نه تازه
- اثر AQI خام قوی بود، پس از کنترل سلف و وعده صفر شد؛ آزمون آستانه‌ای هم چیزی پیدا نکرد — سازوکارش این‌که تهران هم‌زمان بالاترین آلودگی و بالاترین نرخ را دارد

«چرا نرخ کمبود 15.4% ولی مشکلی نیست؟» / «چرا فقط ۵ ماه داده؟»

۷ اسلایدهای پشتیبان

- 15.4% فقط رخداد را می‌شمارد نه اندازه را؛ عمق میانگین هر سلول کمبوددار 4.2 پرس است
- ۵ ماه همان چیزی است که از معاونت دانشجویی تحویل گرفته شد — در محدودیت‌ها آمده و اولین قلم کار آینده است

«۲۲۴ روز سرو از کجا آمد؟» / «۲۳۲ هزار پرس با ۱۶۲ هزار پرسِ ۵ ماه نمی‌خواند»

- ۳۲ هفته‌ی آموزشی = دو نیم‌سال ۱۶هفته‌ای؛ تابستان و نوروز عمداً بیرون گذاشته شدند چون الگوی سرو متفاوتی دارند که در این داده مشاهده نشده — هر ترم ۱۱۲ روز سرو
- داده‌ی خام 53.4 ردیف در روز سرو دارد، ولی نردبان روی مبنای فولدهای رسمی با 48.5 ردیف در روز بسته شده — یعنی برآورد محافظه‌کارانه است، نه خوش‌بینانه

«عدد 3.9 میلیارد فنی امیرآباد چقدر دقیق است؟»

۷ اسلایدهای پشتیبان

- برآورد مرتبه‌ی بزرگی است، نه عدد گزارش‌شده در مخزن
- از میانگین رزرو این سلف و میانگین پیش‌بینی مدل برای آن، ضرب در ۲۲۴ روز سرو ساخته شده — هر دو ورودی از مخزن می‌آیند، ولی خود ضرب در گزارش نیامده