



# Chronos-2

## Time Series Forecasting

# Vấn đề & Mục tiêu nghiên cứu

**Liệu các mô hình nền tảng (foundation models) dự báo chuỗi thời gian ở chế độ zero-shot có thể sánh ngang với các phương pháp dự báo truyền thống hay không?**

## Mục tiêu

- So sánh Chronos-2 (zero-shot) với 3 mô hình baseline mạnh
- Đánh giá mức độ hiệu chuẩn (calibration) của dự báo xác suất
- Đánh giá ý nghĩa thống kê của sự khác biệt
- Phân tích các dạng sai số và các trường hợp mô hình thất bại

## Thiết lập bài toán

- Horizon:  $H = 30$  ngày
- Tính mùa vụ:  $m = 7$  (chu kỳ tuần)
- Chỉ số đánh giá chính: MASE
- Bài toán: Dự báo chuỗi thời gian đơn biến

# Dữ liệu

## Nguồn

API: Wikimedia Pageviews API

**Mục tiêu:** Trang Wikipedia về Bitcoin

**Thời gian:** Tháng 01/2020 – Tháng 12/2024 (5 năm)

**Số bản ghi:** 1.827 quan sát theo ngày

## Tiền xử lý

- Giá trị thiếu: Không có
- Ngoại lệ (2,08%): Được xử lý bằng phương pháp winsorization

## Chia dữ liệu (theo thứ tự thời gian)

### Tập huấn luyện (60%)

- 2020–2022
- 1.096 ngày

### Tập xác thực (20%)

- 2023
- 365 ngày

### Tập kiểm tra (20%)

- 2024
- 366 ngày

# Phương pháp

## Seasonal Naive

- Baseline
- Giá trị từ cùng ngày trong tuần của tuần trước

## ETS

- Thống kê
- Holt–Winters (xu hướng cộng + mùa vụ cộng)

## Gradient Boosting

- Học máy
- LightGBM với các đặc trưng trễ [1, 7, 14, 28]

## Chronos-2

- Foundation model
- T5 encoder–decoder (200 triệu tham số), zero-shot

## Quy trình đánh giá

- Backtesting rolling-origin (k = 5 folds, cửa sổ mở rộng)

## Độ quan trọng đặc trưng của Gradient Boosting

- Các đặc trưng quan trọng nhất:
  - lag\_7 (35%)
  - lag\_1 (22%)
  - rolling\_mean\_7 (18%)

# Kết quả

Hiệu năng trên tập validation (5-Fold CV)

Model	MAE	MASE	Rank
Gradient Boosting	873 $\pm$ 293	0.344 $\pm$ 0.116	1st ★
Chronos-2	999 $\pm$ 200	0.394 $\pm$ 0.079	2nd
Seasonal Naive	1,007 $\pm$ 243	0.397 $\pm$ 0.096	3rd
ETS	1,288 $\pm$ 92	0.508 $\pm$ 0.036	4th
Test Set: Test set (2024): GB (1.080) > Naive (1.082) > Chronos (1.118) > ETS (2.754)			

Test Set: Test set (2024): GB (1.080) > Naive (1.082) > Chronos (1.118) > ETS (2.754)

Gradient Boosting achieves 14% lower MASE than Chronos-2

# Phân tích

## Ý nghĩa thống kê

- Kiểm định Wilcoxon Signed-Rank ( $\alpha = 0,05$ )
- Chronos-2 so với GB:  $p = 0,047 \checkmark$
- Gradient Boosting vượt trội có ý nghĩa thống kê so với tất cả các mô hình

## Hiệu chuẩn dự báo xác suất

- Độ bao phủ khoảng 80%: chỉ đạt 14,2%
- Các khoảng dự báo của Chronos-2 quá hẹp
- Mô hình thể hiện sự quá tự tin (overconfident)

## Sai số theo horizon

- MAE tăng gấp đôi từ  $h = 1$  (~600) lên  $h = 30$  (~1.200)
- Tất cả các mô hình đều suy giảm hiệu năng theo cách tương tự
- Các thời kỳ có lưu lượng truy cập cao là khó dự báo nhất

# Thảo luận

## Vì sao Gradient Boosting vượt trội

- lag\_7 nắm bắt tốt tính mùa vụ theo tuần (35%)
- Kỹ thuật xây dựng đặc trưng theo miền bài toán
- Mã hoá tường minh các quy luật đã biết
- Được tối ưu hoá cho chuỗi dữ liệu cụ thể này

## Giá trị của Chronos-2

- Không cần xây dựng đặc trưng
- Zero-shot: không cần huấn luyện
- Dự báo xác suất được tích hợp sẵn
- Phù hợp hơn cho các danh mục chuỗi đa dạng

**Với một chuỗi đơn có quy luật rõ ràng → kỹ thuật xây dựng đặc trưng chiếm ưu thế.**

**Với danh mục chuỗi đa dạng → các mô hình zero-shot mang lại giá trị thực tiễn.**

# Kết luận & Hướng nghiên cứu tiếp theo

## Phát hiện chính

- Gradient Boosting vượt trội hơn Chronos-2 một cách có ý nghĩa thống kê (MASE 0,344 so với 0,394;  $p = 0,047$ )
- Kỹ thuật xây dựng đặc trưng nắm bắt mùa vụ theo tuần là lợi thế then chốt
- Các khoảng dự báo của Chronos-2 cần được hiệu chỉnh hậu nghiệm (post-hoc calibration)

## Hạn chế

- Chỉ xét một miền dữ liệu (chỉ trang Bitcoin)
- Horizon cố định ( $H = 30$ )
- Không sử dụng biến ngoại sinh
- Sử dụng thiết lập mặc định của Chronos-2

## Hướng nghiên cứu tiếp theo

- Mở rộng sang nhiều trang Wikipedia
- Fine-tune Chronos-2 trên dữ liệu Wikipedia
- Khám phá mô hình tổ hợp (ensemble) giữa Gradient Boosting và Chronos-2



The background is a light blue gradient with dark blue wavy borders at the top and bottom. Scattered throughout are numerous rounded squares of various sizes, some with dark blue outlines and others with light blue outlines. The text "Thank You" is centered in a bold, dark blue font.

**Thank  
You**