

## ІННОВАЦІЙНИЙ ПІДХІД У ПРОГНОЗУВАННІ ЧАСОВИХ РЯДІВ: ВІД ТРАДИЦІЙНИХ МЕТОДІВ ДО НОВАТОРСЬКОЇ МОДЕЛІ TIMESFM

Перцев Ю.О.<sup>1</sup>, Коротка Л.І.<sup>1</sup>

<sup>1</sup>ДВНЗ «Український державний хіміко-технологічний університет», Україна

**Анотація:** *Прогнозування часових рядів є важливим інструментом для компаній у сучасному світі. Воно дозволяє приймати обґрунтовані рішення на основі аналізу історичних даних та передбачень майбутніх тенденцій. У цьому контексті використовуються різні моделі, такі як ARIMA, LSTM, Prophet та XGBoost. Проте їх застосування часто вимагає значного часу для доналаштування. З метою вирішення цієї проблеми Google розробив модель TimesFM, яка забезпечує прогнозування часових рядів без потреби в навчанні. TimesFM показує високу продуктивність та точність на різних типах даних, що робить його потужним інструментом для бізнесу.*

**Ключові слова:** *Прогнозування часових рядів, ARIMA, LSTM, TimesFM*

В сучасному світі прогнозування часових рядів має важливе значення для компаній, оскільки воно дозволяє їм приймати рішення, використовуючи історичні дані про результати діяльності. Вони допомагають передбачити майбутні тенденції. Компанії та підприємці використовують фінансове прогнозування, щоб визначити, як розподілити свої ресурси, чи будуть очікувані витрати на певний період.

Інвестори використовують фінансове прогнозування, щоб визначити, чи вплинуть певні події на акції компанії. Інші аналітики використовують прогнози для екстраполяції того, як зміняться такі тенденції, як внутрішній валовий продукт чи безробіття наступного року. Що далі за часом, то менш точним буде прогноз.

### **Традиційний спосіб прогнозування часових рядів**

На даний момент для прогнозування часових рядів існують різні підходи, які використовуються в промисловості та академічних дослідженнях. Обраний метод зазвичай залежить від характеристик даних і конкретного завдання прогнозування. На теперішній час найвідомішими моделями які широко поширені є:

1. ARIMA (Auto Regressive Integrated Moving Average). Як відомо, ARIMA – це класична статистична модель, яка поєднує авторегресію (AR) та ковзне середнє (MA). Модель забезпечує хорошу продуктивність для даних часових рядів із сильною сезонністю або трендом.

2. Мережі LSTM (Long Short-Term Memory). Типом рекурентної нейронної мережі (RNN) є LSTM – це модель глибокого навчання, призначена для фіксації довгострокових залежностей у послідовних даних.

3. Бібліотека Prophet. FBProphet була розроблена Facebook. Інструмент прогнозування Prophet розроблений для даних часових рядів, які демонструють закономірності в різних часових масштабах; добре обробляє відсутні дані та викиди.

4. Високоєфективна бібліотека для машинного навчання XGBoost (Extreme Gradient Boosting). XGBoost є алгоритмом ансамблевого навчання, який використовує градієнтний бустинг для побудови моделей. Може використовувати прогнози з кількох слабких моделей та об'єднати результати для створення остаточного прогнозу. Для пошуку складних зв'язків у даних часових рядів це дуже добре [1].

Які моделі використовувати залежить від типу даних [2 - 4]. Однак варто зауважити, що під час використання цих моделей значний час витрачається на їх доопрацювання. Щоб вирішити цю проблему та використати останні досягнення в LLMs (generative AI), Google запустив нову модель TimesFM, яка є базовою моделлю декодера, яка може виконувати прогнозування часових рядів із нульовим навчанням.

### **Google TimesFM: базова модель декодера для прогнозування часових рядів**

TimesFM розроблено як попередньо навчену модель прогнозування з 200 мільйонами параметрів, що використовує великий корпус часових рядів із 100 мільярдів реальних часових точок.

На відміну від великих мовних моделей (LLM), які зазвичай навчаються за допомогою декодера для обробки природної мови, TimesFM адаптує архітектуру декодера до прогнозування часових рядів. У моделі

використовуються складені трансформаторні шари, які розглядають суміжні моменти часу як маркери. Унікальною особливістю є використання багат шарового блоку перцептрона із залишковими зв'язками для перетворення фрагментів часових рядів у токени. Під час навчання модель вчиться прогнозувати наступні часові точки, причому довжина вихідного фрагмента потенційно перевищує довжину вхідного фрагмента. Ця архітектура забезпечує ефективне прогнозування з меншою кількістю кроків генерації під час висновку.

Щоб підвищити продуктивність моделі, для попереднього навчання використовується комбінація синтетичних і реальних даних. Синтетичні дані допомагають досягнути фундаментальні часові закономірності, тоді як підібраний корпус із 100 мільярдів часових точок, включаючи дані з Google Trends і переглядів сторінок у Вікіпедії, надає відтінок реального світу, допомагаючи моделі узагальнювати різні домени.

### **Архітектура моделі**

Трансформери показали багатообіцяючі результати в адаптації до іншої тривалості контексту. Під час навчання модель розбиває часовий ряд на кілька патчів. Патч часового ряду є природним аналогом токена в мовних моделях і покращує продуктивність. Крім того, це покращує швидкість висновку, оскільки кількість токенів, що подаються в трансформатор, зменшується на коефіцієнт довжини патча.

TimesFM – це модель лише для декодера, що означає, що для заданої послідовності вхідних патчів модель передбачає наступний патч як функцію всіх попередніх патчів. Має блок попередньої обробки у вигляді вхідних шарів, який перетворює часові ряди на вхідні токени.

Спочатку розбиваються дані на суміжні фрагменти, які не перекриваються. Потім кожен патч обробляється залишковим блоком у вектор розміру моделі. Бінарна маска заповнення постачається разом із вхідним шаром. Залишковий блок – це, по суті, блок багаторівневого перцептрона (MLP) з одним прихованим шаром із з'єднанням пропуску.

Ця модель має ряд трансформаторних шарів, які розміщені один на одному. Кожен із цих рівнів має стандартну багатоголову самоувагу (SA), за

якою слідує мережа прямого зв'язку (FFN). Модель використовує причинну увагу (рис. 1), тобто кожен вихідний токен може відповідати лише вхідним токенам, які передують йому в послідовності (включаючи відповідний вхідний токен). Це можна описати рівнянням:

$$o_j = \text{StackedTransformer}((t_1, m_1), \dots, (t_j, m_j)), j \in [N] \quad (1)$$

### Продуктивність

Для вимірювання прогнозу моделі використовується метрика MSE (середня квадратична похибка), яка кількісно визначає близькість між фактичними та прогнозованими значеннями.

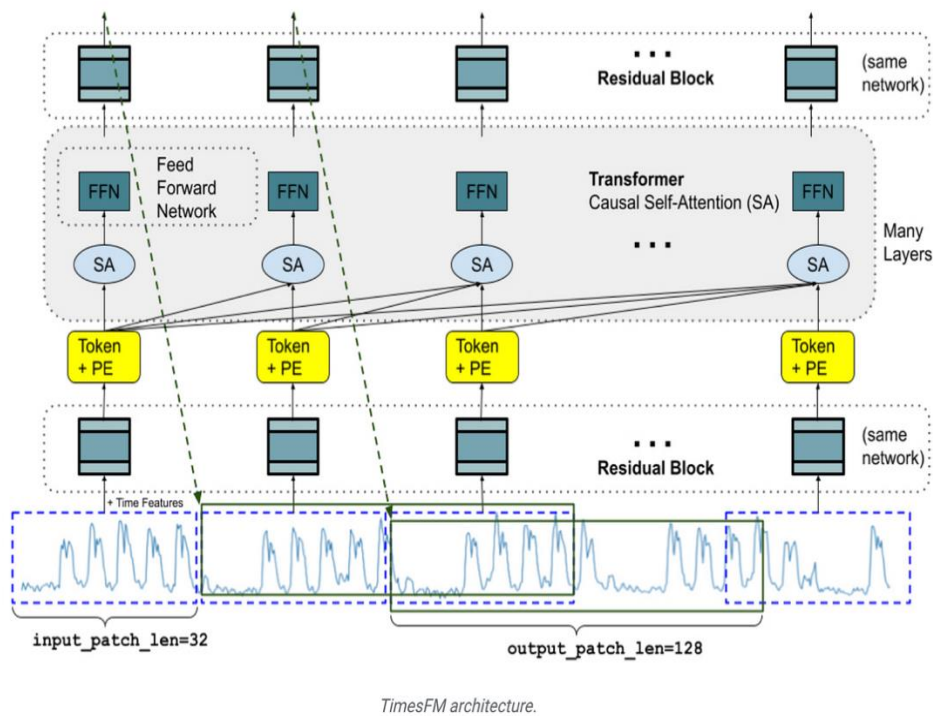


Рисунок 1 Архітектура мережі

TimesFM оцінюється шляхом нульового тестування на різних еталонних показниках часових рядів, демонструючи продуктивність, близьку до найсучаснішого рівня, порівняно зі статистичними методами, такими як ARIMA, і такими моделями DL, як DeepAR і PatchTST. Ефективність моделі продемонстрована в архіві прогнозів Монаша та популярних еталонних тестах для довгострокового прогнозування, перевершуючи як контрольовані, так і нульові підходи.

Запропонована архітектура TimesFM, незважаючи на свій менший розмір, виявляється потужним інструментом для точного та ефективного прогнозування в різноманітних програмах.

### **ЛІТЕРАТУРА**

- 1.Перцев Ю. О., Коротка Л. І. Порівняння нейронних мереж RNN та LSTM типу при прогнозуванні цін на фондовому ринку. Матеріали VIII Міжнародної науково-технічної конференції комп'ютерне моделювання та оптимізація складних систем (1-3 листопада 2023 року м. Дніпро, Україна). С. 124-127.
- 2.Abhimanyu Das, Weihao Kong, Rajat Sen, Yichen Zhou A decoder-only foundation model for time-series forecasting arXiv:2310.10688
3. [Electronic resource] - <https://blog.research.google/2024/02/a-decoder-only-foundation-model-for.html>  
(Доступ 31.03.2024)
4. Зеленцов Д.Г. Технології обчислювального інтелекту в задачах моделювання динамічних систем: монографія / Д.Г. Зеленцов, Л.І. Коротка // - Дніпро: Баланс-Клуб, 2018. - 178 с.

### **Innovative Approach to Time Series Forecasting: From Traditional Methods to the Cutting-Edge Model TimesFM**

Pertsev Y., Korotka L.

**Abstract.** *Forecasting time series is crucial for companies as it enables them to make informed decisions using historical data and future predictions. Various approaches exist for time series forecasting, each tailored to different data characteristics and forecasting tasks. Currently, widely known models include ARIMA, LSTM networks, Prophet, and XGBoost. However, these models often require significant preprocessing time. To address this issue and leverage recent advancements in generative AI, Google introduced TimesFM, a decoder-only model specifically designed for time series forecasting. TimesFM utilizes transformer layers and a multi-layer perceptron block to transform time series fragments into tokens, enabling efficient forecasting with minimal generation steps. Synthetic and real-world data are combined for pretraining to capture fundamental temporal patterns and enhance model generalization. Evaluation demonstrates TimesFM's competitive performance across various benchmark time series datasets compared to traditional statistical methods and DL models.*

**Keywords:** *Time series forecasting, ARIMA, LSTM, TimesFM*

### **REFERENCE**

- 1.Pertsev Yu. O., Korotka L. I. Porivniannia neironnykh merezh RNN ta LSTM typu pry prohnouzuvanni tsin na fondovomu rynku. Materialy VIII Mizhnarodnoi naukovo-

tekhnichnoi konferentsii kompiuterne modeliuvannia ta optymizatsiia skladnykh system (1-3 lystopada 2023 roku m. Dnipro, Ukraina). S. 124-127.

Abhimanyu Das, Weihao Kong, Rajat Sen, Yichen Zhou A decoder-only foundation model for time-series forecasting arXiv:2310.10688

3. [Electronic resource] - <https://blog.research.google/2024/02/a-decoder-only-foundation-model-for.html>

(Dostup 31.03.2024)

4. Zelentsov, D.G. and Korotkaya, L.I. (2018) Tehnologii vyichislitel'nogo intellekta v zadachah modelirovaniya dinamicheskikh sistem: monografiya [Technologies of Computational Intelligence in Tasks of Dynamic Systems Modeling: Monograph], Balans-Klub, Dnepr, 178pp.