MORPHEUS : A Foundation Model for Multivariate Time Series Forecasting

Harsh Deshpande, Manoj Cherukumalli, Amit Varshney, Prathamesh Patil, Leonard Eun, Dushyant Sahoo, Naren Chittar
JPMorgan Chase

Abstract

Multivariate time series are vital for capturing complex interactions among variables in a number of domains including finance, e-commerce, and climate science. Research in this space is predominantly focused on developing bespoke models tailored to specific tasks, with relatively little exploration into foundation models that offer universal forecasting capabilities. We address this gap with two contributions: a novel framework that adapts traditional tokenization techniques to multivariate time series, integrating multiple target and feature series into a unified model allowing us to leverage existing foundation models for language; and an innovative synthetic data generation process to overcome data scarcity, enabling robust model training. Our approach handles diverse covariates and is validated through extensive experiments, demonstrating superior performance over current state-of-the-art methods.

Introduction

Multivariate time series forecasting is a fundamental task in various domains, including finance, economics, meteorology, and supply-chain management. Accurate predictions can drive critical decision-making processes, yet the literature predominantly focuses on developing bespoke models tailored to specific tasks. These models often require extensive training and are prone to overfitting, especially in domains with small and noisy datasets like finance. This paper introduces MORPHEUS, a foundation model designed to offer universal forecasting capabilities. By leveraging existing language model architectures and addressing data scarcity through innovative synthetic data generation, MORPHEUS aims to provide a robust solution for multivariate forecasting across diverse applications.

Methodology

The MORPHEUS model employs a novel interleaving method to effectively manage high-dimensional multivariate data. This method integrates endogenous and exogenous covariates into a single sequence using special separator tokens, enabling the use of sequence-to-sequence architectures like the T5 transformer. The interleaving process ensures that recent observations are positioned closer to the end of the context, which is crucial for transformers known to focus on the end of the context. Additionally, MORPHEUS addresses data scarcity through a synthetic data generation process called MySTiC. This process generates synthetic multivariate time series data by sampling from an underlying Hidden Markov Model, augmenting real-world data and facilitating effective pretraining of the foundation model.

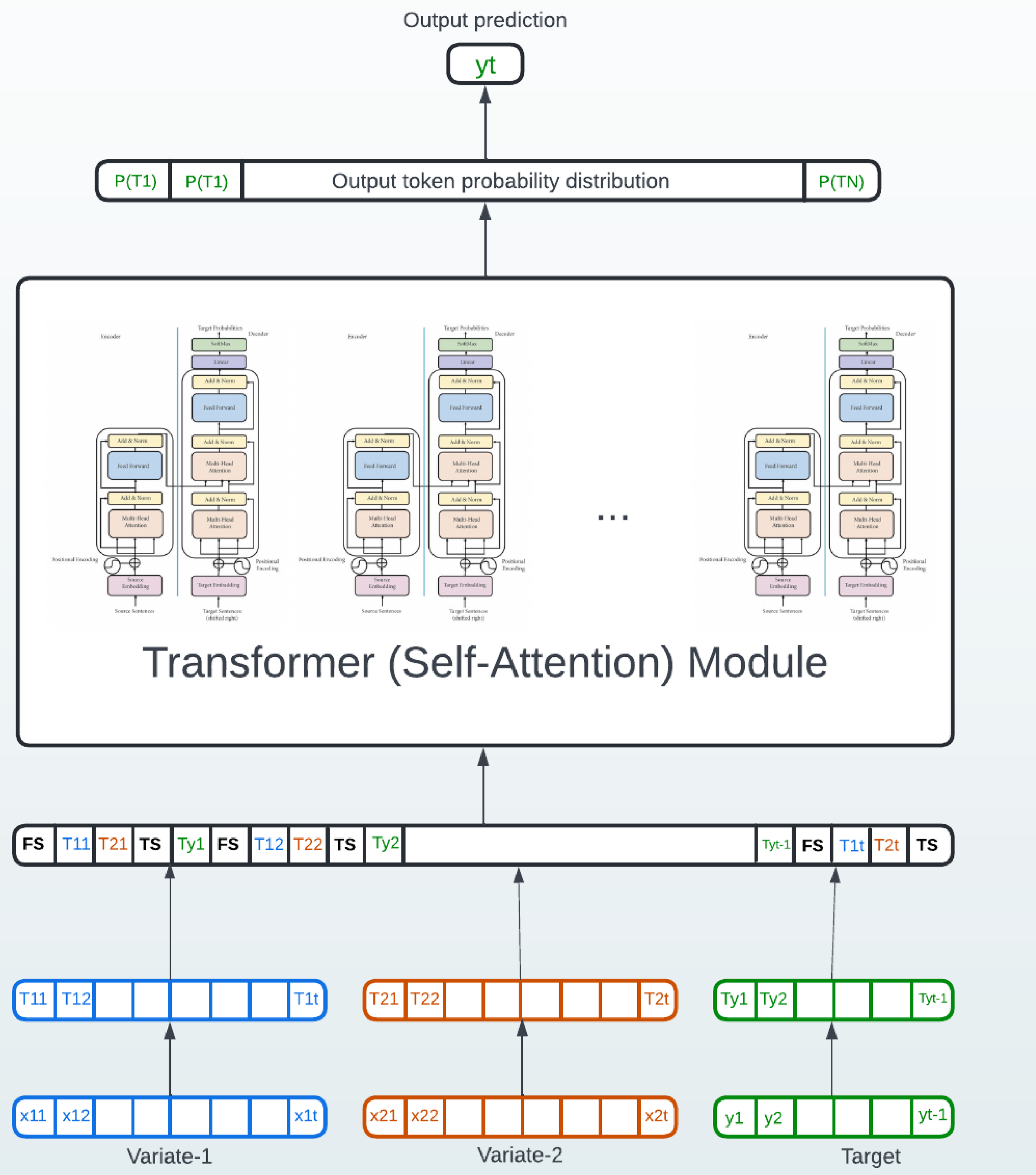

Figure 1. Architecture for MORPHEUS illustrating a case with 2 exogenous features and 1 endogenous target. The feature and target series are discretized to map real-valued observations into model tokens, and then interleaved using Feature and Target Separators (FS, TS). The interleaved series is passed through a stack of T5 transformer blocks to produce the next token, which is then mapped back to a real value which is our forecast.

Results

MORPHEUS was evaluated using the Mean Absolute Scaled Error (MASE) across synthetic and real-world datasets, demonstrating strong multivariate forecasting capabilities. On synthetic datasets, MORPHEUS achieved the lowest MASE scores, benefiting from pretraining on similar data. In real-world domains like finance, electricity, and traffic, MORPHEUS consistently outperformed or matched benchmark models, achieving MASE scores below 1. Even in challenging scenarios with complex dependencies, MORPHEUS maintained robust performance, highlighting its adaptability and competitive predictive accuracy. The model's innovative interleaving strategy effectively integrates multiple target and feature series, enhancing its ability to capture complex interactions.

		MySTiC-B	MySTiC-A	US Treasury	GB Treasury	Etth1	Etth2	Ettm1	Ettm2	Electricity	Traffic	Weather	Illness	All
2 Features	Chronos	0.43	0.30	1.09	1.21	1.01	0.60	1.08	0.73	**0.74**	0.51	1.04	1.06	0.82
	MOIRAI	0.83	0.77	1.23	1.67	**0.96**	0.95	1.14	1.13	0.80	0.57	1.19	1.22	1.04
	MORPHEUS	**0.25**	**0.27**	0.58	**0.68**	1.02	0.54	1.06	0.68	0.77	0.51	**1.00**	1.19	**0.71**
	DLinear	2.33	5.46	1.08	14.01	1.00	0.69	**1.00**	0.64	0.85	0.75	1.19	1.93	2.58
	LightTS	6.08	0.97	1.67	4.48	1.10	0.84	1.06	0.90	0.97	0.77	1.41	1.87	1.84
	Non-stationary Transformer	0.63	1.24	0.66	1.25	0.98	0.54	1.03	0.64	1.03	0.56	1.06	1.31	0.91
	TFT	1.42	1.57	0.77	1.46	0.97	**0.45**	1.02	0.57	0.90	0.50	1.02	2.00	1.05
	TiDE	1.54	1.57	1.10	3.62	1.03	0.81	**1.00**	0.67	0.91	0.89	1.26	1.86	1.35
	TimeMixer	0.82	0.74	1.02	1.48	0.99	0.52	1.01	0.60	0.91	0.61	1.09	1.70	0.96
	TimeXer	0.34	0.49	0.59	0.51	**0.96**	0.48	**1.00**	0.63	0.89	0.52	1.05	**1.01**	**0.71**
	Vanilla NN	0.39	0.57	**0.55**	0.96	**0.96**	0.51	1.01	0.61	0.92	0.52	1.05	1.06	0.76
	iTransformer	0.32	0.49	0.56	0.84	0.97	**0.45**	1.01	**0.59**	0.88	**0.49**	1.09	**1.01**	0.72
5 Features	Chronos	0.60	**0.24**	1.09	-	1.01	0.60	1.07	0.73	**0.74**	0.51	1.04	1.07	0.79
	MOIRAI	0.75	0.75	1.22	-	**0.92**	0.98	1.11	1.10	0.76	0.59	1.19	1.24	0.96
	MORPHEUS	**0.22**	0.30	0.50	-	1.04	0.57	1.09	0.70	0.77	0.45	**0.99**	1.39	0.73
	DLinear	2.07	2.43	1.07	-	1.00	0.69	**1.00**	0.64	0.85	0.74	1.19	1.93	1.24
	LightTS	1.04	1.28	1.70	-	1.12	0.87	1.07	0.79	0.98	0.76	1.29	1.88	1.16
	Non-stationary Transformer	0.70	1.32	0.59	-	0.99	0.57	1.03	0.66	1.07	0.59	1.07	-	0.86
	TFT	0.81	1.35	0.63	-	0.97	0.46	1.01	**0.58**	0.91	0.50	1.02	1.68	0.90
	TiDE	1.40	1.26	1.09	-	1.03	0.80	**1.00**	0.66	0.91	0.88	1.25	1.98	1.11
	TimeMixer	0.76	0.72	1.03	-	0.99	**0.52**	1.01	0.60	0.91	0.62	1.10	1.27	0.87
	TimeXer	0.39	0.40	0.53	-	0.96	0.48	1.01	0.63	0.88	0.52	1.05	1.01	0.71
	Vanilla NN	0.39	0.47	**0.45**	-	0.95	0.52	1.01	0.61	0.90	0.50	1.09	**0.99**	0.72
	iTransformer	0.34	0.41	0.49	-	0.97	0.45	1.01	0.59	0.88	**0.48**	1.07	1.00	**0.70**

Table 2. MASE of different models across feature sets (lower is better; bold for best, underlined for second-best). MASE is ratio of the model MAE (Mean Absolute Error) to a naive model MAE. Naive model is a model which repeats the last available target value as a forecast. Note: GB Treasury dataset only has 2 features, hence the 5 feature metrics cannot be computed

Conclusion

This paper presents significant advancements in multivariate time series forecasting by introducing MORPHEUS, a foundation model that integrates multiple target and feature series into a unified framework.

The interleaving strategy and synthetic data generation process enhance model training, offering predictive performance that is competitive with state-of-the-art bespoke models. The techniques are adaptable to any sequence-to-sequence architecture, paving the way for a universal model for multivariate forecasting.

Future research directions include enhancing synthetic data generation to better mimic real-world patterns and applying the interleaving scheme to base models other than T5 used in this work.