



HAL
open science

A Temporal Linear Network for Time Series Forecasting

Rémi Genet, Hugo Inzirillo

► **To cite this version:**

Rémi Genet, Hugo Inzirillo. A Temporal Linear Network for Time Series Forecasting. 2025. hal-04923919

HAL Id: hal-04923919

<https://hal.science/hal-04923919v1>

Preprint submitted on 1 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Temporal Linear Network for Time Series Forecasting

Rémi Genet^{*§} and Hugo Inzirillo^{†§}

^{*} DRM, Université Paris Dauphine - PSL

[†] CREST-ENSAE, Institut Polytechnique de Paris

Abstract—Recent research has challenged the necessity of complex deep learning architectures for time series forecasting, demonstrating that simple linear models can often outperform sophisticated approaches. Building upon this insight, we introduce a novel architecture the Temporal Linear Net (TLN), that extends the capabilities of linear models while maintaining interpretability and computational efficiency. TLN is designed to effectively capture both temporal and feature-wise dependencies in multivariate time series data. Our approach is a variant of TSMixer that maintains strict linearity throughout its architecture. TSMixer removes activation functions, introduces specialized kernel initializations, and incorporates dilated convolutions to handle various time scales, while preserving the linear nature of the model. Unlike transformer-based models that may lose temporal information due to their permutation-invariant nature, TLN explicitly preserves and leverages the temporal structure of the input data. A key innovation of TLN is its ability to compute an equivalent linear model, offering a level of interpretability not found in more complex architectures such as TSMixer. This feature allows for seamless conversion between the full TLN model and its linear equivalent, facilitating both training flexibility and inference optimization. Importantly, we demonstrate that TLN outperforms standard linear regression models. This superior performance is attributed to the unique training structure of the TLN and the inherent relationship of weights established during model construction, which can not be found in conventional regression approaches. Our findings suggest that TLN strikes a balance between the simplicity of linear models and the expressiveness needed for complex time series forecasting tasks. It offers improved interpretability compared to TSMixer while demonstrating greater resilience in multivariate cases than basic linear models and standard linear regression. This work contributes to the ongoing discussion about the trade-offs between model complexity and performance in time series analysis, and opens new possibilities for developing efficient and interpretable forecasting models that leverage the strengths of neural network architectures and linear models.

I. INTRODUCTION

Time series forecasting is a critical task with wide-ranging applications in various domains such; traffic flow estimation, energy management, and financial investment [1]. Multivariate time series have attracted a great interest in the scientific community [2], [3], becoming an essential field of study for many industrial applications. The constant increase in available data justifies the development of increasingly sophisticated predictive frameworks. Unlike univariate time series that focus on the evolution of a single variable, multivariate series framework integrate several interdependent variables which significantly

complicates the learning task. The analysis of these data requires models capable of managing temporal dependencies and dynamic interactions across different dimensions and time horizons. To address these challenges, researchers have explored a variety of approaches, from neural networks [4] to state-space models [5], [6], vector autoregressive models [7], [8] and advanced neural architectures [9], [10]. Despite the increase in the availability of dataset for time series forecasting, it remains a very complex task. Powerful models have been proposed through the years [11], [12]. Some research proposed dynamic system to model different existing states within a time series [13], [14], [15]. Deep learning techniques, particularly Long Short-Term Memory (LSTM) [16] networks and attention [17] mechanisms, have shown promising results due to their ability to handle nonlinearities and long-term dependencies in data. In this paper, we introduce a novel architecture that combines the interpretability of linear models with the structural advantages of deep learning approaches. We demonstrate TLN’s effectiveness and consistency through extensive experiments on two different multistep forecasting tasks, showing robust performance across varying sequence lengths and prediction horizons. Our results show that TLN achieves superior or comparable performance to more complex architectures while maintaining complete interpretability through its linear equivalent form and demonstrating remarkable stability across different input configurations. The implementation of TLN is available as an open-source Python package `temporal_linear_network` on PyPI, and the complete source code can be found at TLN. This repository includes all the code necessary to reproduce our experiments and implement TLN in various time series forecasting applications.

II. RELATED WORK

Econometrics methods for predicting time series, such as ARIMA [18] and its variants, have been the predominant choices for stationary and linear time series. These models rely on assumptions of stationarity and linearity which limit their ability to capture complex dynamics. These complex dynamics are often observed in real world data. Approaches such as Prophet [19] have been proposed to improve the use of statistical models by incorporating seasonal and holiday effects, but remain limited in contexts where non-linear interactions are predominant. Over the past decades, time series forecasting solutions have evolved from traditional statistical methods to sophisticated deep learning approaches. RNNs [20] has

[§]These authors contributed equally.

emerged as a fundamental architecture with the capacity of maintaining internal memory which allowed sequential data management. Limitation of standard RNNs lies in vanishing gradient [21] which motivates the development of more sophisticated architecture, LSTMs [16]. LSTMs use gating mechanism to increase the long term memory management. [22] proposed the GRU (Gated Recurrent Units) offering an alternative to LSTMs. They are less complex while maintaining comparable performances. Recent research has turned to Mixture of Experts (MoE) [23], architectures that improve flexibility and performance by dynamically selecting specialized models or "experts" based on the input data. Many expert choice methods have been proposed [24], [25]. In our previous work we proposed a method using gated mechanism relying on kolmogorov-arnold networks [26] to weight the contribution of each expert to the global output [27]. We refer the reader to [24] for further information. Recently, there has been a surge in Transformer-based architectures [28] with extensions for time series analysis and forecasting [29], [30], particularly for the challenging task of long-term time series forecasting (LTSF) [1]. Transformer have demonstrated remarkable success in natural language processing [?], machine translation [31] and computer vision tasks. We mentioned it before, they have been adapted for time series forecasting, Lim et al. [29] proposed a Temporal fusion Transformer (TFT), built using LSTMs and attention mechanism. However, the effectiveness of Transformer-based models for time series forecasting has been called into question. Zeng et al. [1] argue that while Transformers excel at extracting semantic correlations among elements in a long sequence, time series modeling primarily involves extracting temporal relations. The permutation-invariant nature of the self-attention mechanism in Transformers may result in temporal information loss, which is crucial for time series analysis. In contrast, Chen et al. [32] propose TSMixer, a novel architecture designed by stacking multi-layer perceptrons (MLPs) [33]. TSMixer is based on mixing operations along both the time and feature dimensions to extract information efficiently. This approach aims to leverage the strengths of MLPs while addressing the limitations of Transformer-based models in capturing temporal dependencies. The debate surrounding the efficacy of Transformer-based models versus simpler architectures like linear models or MLPs for time series forecasting highlights the need for further research and comparison of different approaches. This paper aims to explore and compare various models for time series forecasting, focusing on their ability to capture temporal relations and their performance on benchmark datasets.

III. ARCHITECTURE

We introduce the Temporal Linear Network (TLN), a novel architecture designed to capture complex temporal relationships while maintaining interpretability through its ability to be transformed into an equivalent linear model. The TLN model consists of multiple layers, each comprising a SequentialDense

operation followed by a 1D convolution. For a model with L layers, the forward pass can be expressed as:

$$\tilde{X}^{(l)} = f_{\omega_l}(X^{(l)}), \quad (1)$$

For each feature we compute the 1D convolution along the sequence:

$$\text{Conv1D}_{\omega_{l,q}}(\tilde{X}_q^{(l)}) = \sum_{k=0}^{K^{(l)}-1} W_{k,q}^{(l)} \cdot \tilde{X}_{t+k,q}^{(l)} + b_q, \quad (2)$$

where $\tilde{X}_q^{(l)}$ is the input sequence, $W \in \mathbb{R}^{K \times Q}$ is the associated weights and $b \in \mathbb{R}^Q$ the bias. Writing the equation for each layer of network, we obtain:

$$\tilde{X}^{(l)} = \text{Conv1D}_{\omega_l}(f_{\omega_l}(\tilde{X}^{(l-1)})), \quad (3)$$

where ω_l denotes all the weights shared for each layer. The final layer form will be:

$$\tilde{X}^{(L)} = \text{Conv1D}_{\omega_L}(f_{\omega_L}(\tilde{X}^{(L-1)})). \quad (4)$$

A. SequentialDense Layer

The SequentialDense layer applies separate transformations along the feature and time dimensions. For an input tensor $X \in \mathbb{R}^{S \times F}$, where B denotes the batch size, S the input sequence length, and F is the number of input features. The SequentialDense layer performs the following operations. The first step consist in an element wise scaling:

$$\tilde{x} = K \odot X \quad (5)$$

where the base kernel is represented by $K \in \mathbb{R}^{S \times 1}$. Followed by the feature transformation

$$x_{out} = F \odot (W_x \tilde{x} + b_x). \quad (6)$$

The feature kernel is given by $F \in \mathbb{R}^{1 \times F'}$, where F' denotes the number of output features. The final step of the sequential dense layer consist in applying a time kernel along each sequence.

$$Y = \tau \odot (W_t x_{out} + B_t)', \quad (7)$$

where the time kernel is $\tau \in \mathbb{R}^{S' \times 1}$, where S' is the output sequence length.

B. Comparison with Standard Linear Regression

In a standard linear regression on flattened inputs, the model would have a single weight matrix $W \in \mathbb{R}^{(S \times F) \times (S' \times F')}$ and a bias vector $b \in \mathbb{R}^{S' \times F'}$. The prediction would be computed as:

$$Y_{ijk} = \sum_{l=1}^{SF} W_{ij,l} X_l + b_{ij} \quad (8)$$

It's important to note that in many cases, the number of parameters in this standard linear regression model ($S \times F \times S' \times F' + S' \times F'$) can be significantly larger than the number of parameters in our TLN model. This is because TLN uses parameter sharing and structured transformations to reduce the total number of parameters while still capturing complex temporal relationships.

C. Equivalent Linear Model

Despite its complex structure, TLN can be transformed into an equivalent linear model. This is possible because all operations in TLN (SequentialDense and Conv1D) are linear transformations. The composition of linear transformations is itself a linear transformation, which guarantees the existence of an equivalent linear model. Let f_1, f_2, \dots, f_n be linear transformations. Then, for any vectors x and y , and any scalar c :

$$\begin{aligned} f_n(\dots f_2(f_1(x + cy))) &= f_n(\dots f_2(f_1(x) + cf_1(y))) \\ &= f_n(\dots (f_2(f_1(x)) + cf_2(f_1(y)))) \\ &= f_n(\dots f_2(f_1(x))) + cf_n(\dots f_2(f_1(y))) \end{aligned} \quad (9)$$

This proves that the composition $f_n \circ \dots \circ f_2 \circ f_1$ is also a linear transformation. To find the equivalent linear model, we can compute the equivalent weights $W_{eq} \in \mathbb{R}^{S \times F \times S' \times F'}$ and bias $b_{eq} \in \mathbb{R}^{S' \times F'}$ as follows: First, compute the bias by passing a zero input through the model:

$$\begin{aligned} b_{eq} &= \text{TLN}(0) \\ &= \text{TLN}(0_{S' \times F'}) \end{aligned} \quad (10)$$

Then, for each input position (i, j) , compute:

$$W_{eq}^{(i,j)} = \text{TLN}(e_{ij}) - b_{eq} \quad (11)$$

where e_{ij} is a unit vector with a 1 at position (i, j) and 0 elsewhere. The equivalent linear model can then make predictions using:

$$Y = \sum_{i=1}^S \sum_{j=1}^F W_{eq}^{(i,j)} \cdot X_{i,j} + b_{eq} \quad (12)$$

This transformation allows us to interpret the complex TLN model as a single linear operation, providing insights into how the model uses different parts of the input sequence to make predictions, while maintaining a more parameter-efficient structure during training and standard inference.

IV. LEARNING

To evaluate the effectiveness of the Temporal Linear Network (TLN), we conducted experiments on two different tasks, designed to study multiple aspects of the model's performance.

A. Task 1: Market Traded Notional Predictions

Our first task focuses on predicting market traded notional values, inspired by recent works in financial time series forecasting [34], [35], [30], [27]. Financial markets are known for their high volatility and noise, making forecasting challenging. However, traded volume series often exhibit stronger patterns, including seasonality, autocorrelation, and correlations between different assets. For this task, we applied a min-max scaler to preprocess the data, ensuring all values fall within a consistent range.

B. Task 2: ETTh1 Dataset

For our second task, we utilized the ETTh1 dataset, a well-known benchmark in time series forecasting literature. This dataset consists of hourly measurements related to electricity transformer temperatures. To align with previous studies and enable fair comparisons, we preprocessed this data using a standard scaler.

C. Experimental Setup

For both tasks, we designed our experiments to study the impact of various factors on model performance; we varied the input sequence length by testing values of 3, 10, 20, 50, 100, 200, and 400 time steps to understand how sequence length impacts model accuracy. We also compared the performance of models using univariate inputs (single variable) versus those utilizing multivariate inputs (multiple variables) to determine the benefits of incorporating additional features. We also explored the effect of adding temporal information, such as the hour and day of the week. We stacked this information to the input data to capture any underlying patterns related to "time". Finally, we evaluated the models across different forecasting horizons. We made predictions for 1, 6, 15, 30, 90, 180, and 360 hours ahead. The objective was to assess the capability of the different models in making both short-term and long-term forecasts. Both tasks use hourly data. This configuration allows us to perform a consistent comparison across different forecasting horizons and input sequence lengths.

D. Comparison

We compared the TLN against a diverse set of baseline models.

Linear Regression: This model flattens the sequential dimension, transforming the input from shape (sequence_length, features) to (sequence_length * features). It then applies a standard linear regression to predict the target values.

Linear Models from "Are Transformers Effective for Time Series Forecasting?": These models were a key inspiration for our work, as they demonstrate the effectiveness of simple linear approaches. While similar to TLN in their linear nature, they lack some of the mechanisms present in our model. In our experiment, we implemented three types of linear layers to process the time series data. The first is "CLinear", which is a classic linear layer applied directly across the time dimension. This setup provides a straightforward transformation of the input data. The second, "NLinear", introduces a normalization step by subtracting the last value in the sequence before applying the linear transformation. The aim of this transformation is to reduce bias from recent trends. Finally, we implemented the "DLinear" which decompose the time series into its trend and seasonal components. After the decomposition processed, linear layer is applied over each component. This decomposition allows the model to better capture and process different patterns within the time series. It's worth noting that these models are not inherently designed for multivariate to univariate prediction.

To address this limitation, we added a final dense layer to aggregate the results.

Recurrent Neural Networks (RNNs) are often the default choice for time series forecasting and are a key component in more complex models like the Temporal Fusion Transformer. However, given recent criticisms of transformer models, we aimed to reassess the efficacy of RNNs for standard time series forecasting. While RNNs can capture complex non-linear patterns, they come with a computational cost due to their non-parallelizable sequential processing.

LSTM Encoder-Decoder uses separate LSTM layers for encoding and decoding. The encoder (50 units) processes the input sequence and passes its final state to the decoder (50 units), which generates the output sequence. A final dense layer produces the predictions.

LSTM Shared Encoder-Decoder Uses a single LSTM layer (units equal to the number of input features) shared between the encoding and decoding phases. We also make a comparison against two variations of GRU (Gated Recurrent Unit) networks

GRU Last Sequence, consists of a two-layer GRU network with 10 units in each layer, where only the last output of the second layer is used for prediction. This approach focuses on capturing the final state of the sequence, which can represent the most relevant information for forecasting. The second variant, *GRU Full Sequence*, also features a two-layer GRU network with 10 units per layer but differs by using the full output sequence of the second layer. The outputs are flattened and then used for prediction, allowing the model to leverage information from all time steps within the sequence, potentially capturing more intricate temporal patterns.

Multilayer Perceptrons (MLPs): While not typically used for time series forecasting, we included MLPs as a fundamental deep learning baseline. As with linear regression, we flattened the inputs to accommodate the sequential data. We used linear activations to maintain comparability with our model. In previous tests, we observed that ReLU activations in small architectures like these often hindered performance in this scenario. *Single-layer MLP*: Flattens the input and applies a single dense layer to produce the output. *Two-layer MLP*: Flattens the input, applies a dense layer with $(n_ahead * 10 * 2)$ units, followed by another dense layer for the final prediction.

TSMixer: The TSMixer, which inspired our work, offers an alternative to RNNs and builds upon the ideas presented in "Are Transformers Effective for Time Series Forecasting?" It consists of the following key components: First, *Reversible Instance Normalization* is employed to normalize the data while preserving the ability to revert to the original representation, which helps in maintaining stability during training. Next, a *Temporal Linear layer* is utilized; it permutes the input dimensions and applies a dense layer, allowing for effective handling of temporal dependencies within the data. Finally, a *Feature Linear layer* consisting of two dense layers with an activation function in between is integrated, enabling complex feature

transformations and non-linear interactions. We tested four different variants; We tested four variants: *TSMixer-1* consists of a single block utilizing ReLU activation, while *TSMixer-2* extends this to two blocks, also with ReLU activation, allowing for deeper feature extraction. In contrast, *TSMixer-no-relu-1* features a single block with linear activation, and *TSMixer-no-relu-2* uses two blocks with linear activation, offering an alternative approach without non-linear transformations. These configurations enable flexibility in handling different types of data and learning patterns. All TSMixer variants use layer normalization and a feed-forward dimension of 5. We included variants with linear activation to align more closely with TLN approach. As with the linear models mentioned earlier, we added a dense layer for result aggregation in multivariate to univariate prediction tasks.

TLN (LinearNet): Our proposed Temporal Linear Network, implemented in two variants: *LinearNet* which uses convolution operations and *LinearNet-no-convolution* which does not embed convolution filter. Both variants use two hidden layers with default parameters for internal hidden size. This comprehensive set of models allows us to compare our proposed TLN against a wide range of approaches, from simple linear models to complex deep learning architectures, providing a thorough evaluation of its performance in various time series forecasting scenarios.

E. Evaluation Metrics

We used different evaluation metrics for our two tasks, chosen to facilitate comparison with existing literature and to provide the most meaningful and interpretable results for each specific dataset.

1) *Task 1: Market Traded Notional Predictions*: For the first task, we used the R-squared (R^2) score as our primary evaluation metric. The R^2 score is defined as:

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \quad (13)$$

where y_i are the true values, \hat{y}_i are the predicted values, and \bar{y} is the mean of the true values.

2) *Task 2: ETTh1 Dataset*: For the second task, we used Mean Squared Error (MSE) as our primary evaluation metric. MSE is defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2, \quad (14)$$

where n is the number of predictions.

3) *Rationale for Metric Selection*: The choice of different metrics for the two tasks serves multiple purposes:

Consistency with existing literature: By using R^2 for the financial task and MSE for the ETTh1 dataset, we maintain consistency with the metrics commonly used in these domains, facilitating direct comparisons with other papers.

Readability of results: In the BTC tasks, the R^2 scores typically fall in the range of 0.1 to 0.4. In this range of values, differences between models are more easily discernible with this metric. Conversely, for the ETTh1 tasks, R^2 scores would often exceed 0.9, compressing the visible differences between models. Using MSE for ETTh1 provides a clearer distinction between model performances.

Scale-independent interpretation: R^2 has the advantage of being scale-independent and directly interpretable as the proportion of variance explained by the model. This makes it particularly useful for the financial task where the absolute scale of errors may be less informative than the proportion of variance explained. It's important to note that R^2 and MSE are intrinsically related. Both metrics measure the same fundamental quantity - the discrepancy between predicted and actual values. The choice between them is primarily a matter of presentation and interpretation in the context of the specific task and data characteristics. The R^2 has an additional advantage in that it inherently accounts for the variance of the target variable. This makes it more immediately interpretable than MSE, as it provides a clear benchmark (with 1 being a perfect score and 0 indicating performance no better than always predicting the mean). By using these two complementary metrics across our tasks, we aim to provide a comprehensive and nuanced evaluation of model performance, balancing the needs for comparability with existing literature, readability of results, and meaningful interpretation in the context of each specific forecasting task.

F. Results Analysis

Our comprehensive experimental results, presented in the appendix for clarity, reveal key insights about model behaviors across different scenarios and demonstrate how TLN (referred to as LinearNet in the tables) provides an effective alternative to existing approaches. For Bitcoin volume prediction in univariate settings without time features, short-term predictions (1-step ahead) show similar performance across simpler models including linear regression, CLinear, NLinear, DLinear, MLP, TSMixer, and TLN. Surprisingly, RNNs perform poorly in these scenarios despite their prevalence in time-series forecasting applications. More significantly, deep learning linear models (CLinear, NLinear, and DLinear) and TSMixer show degraded performance with longer sequence lengths, while both linear regression and TLN maintain consistent performance regardless of sequence length.

Regarding medium-term predictions (6-30 steps ahead), the performance hierarchy shifts substantially, with recurrent architectures demonstrating superior capabilities. However, our proposed TLN approach maintains stable performance while offering a significant advantage in computational efficiency. While RNN training times increase dramatically from 8 to 150 seconds as sequence length grows from 3 to 400, TLN training time remains nearly constant, increasing only marginally from 6.5 to 7.1 seconds. Long-term prediction scenarios particularly highlight TLN's strengths. While TSMixer occasionally

matches its performance depending on the specific run, TLN consistently delivers stable results across all runs, contrasting sharply with the highly variable performance of other models. These results emphasize TLN's two key advantages: consistent handling of long sequences regardless of input characteristics and minimal training time compared to deep learning alternatives. Adding temporal features improves performance across all models, demonstrating the importance of this basic information. RNNs benefit significantly, showing stronger performance in both short and medium-term predictions without their previous short-term limitations. Even basic linear regression and ridge regression models perform well in these scenarios. TLN maintains consistent strong performance across configurations, with one notable exception: a simple 3-layer MLP achieves surprisingly good results for 360-step predictions with very short input sequences. However, TLN matches this performance when provided longer sequences, suggesting that while it may not extract optimal predictions from limited data, it effectively leverages longer sequences for reliable predictions.

The transition to multivariate tasks, both with and without temporal features, reveals particularly interesting dynamics. As noted in previous research, increased input dimensionality often degrades model performance and complicates calibration, evident in our results tables. While some models become practically unusable (particularly TSMixer), TLN emerges as consistently superior across nearly all prediction horizons. This stability extends from its earlier demonstrated robustness to sequence length - as feature dimensionality increases, TLN maintains consistent performance while other models degrade. This advantage becomes clear when examining parameter counts: most models show dramatic parameter increases with added features, while TLN maintains a more modest growth rate, particularly compared to ridge regression or linear regression. Results on the ETTh1 dataset, which has previously highlighted the limitations of complex deep learning approaches, confirm earlier research showing simpler models outperforming RNNs. While individual models may excel in specific scenarios, TLN demonstrates unmatched consistency across all configurations. Frequently, other models achieve strong performance with certain sequence lengths but fail dramatically when input length changes. This consistency across varying conditions, combined with efficient training times and robust performance across input dimensions, establishes TLN as a practical advancement in time series prediction. This analysis underscores TLN's primary contributions: consistent performance across varying input conditions, efficient computational requirements, and reliable results across different prediction horizons. These characteristics make it particularly valuable for practical applications requiring robust performance across diverse scenarios.

V. COMPARING WITH MACHINE LEARNING MODELS

As demonstrated in the learning tasks, linear models often prove not only sufficient but also more efficient for time series forecasting. These models should be considered as primary

candidates before exploring non-linear models, which are frequently the default choice in deep learning literature. Having proposed a new architecture that essentially offers an innovative method for calibrating a linear model, we aim to highlight its distinctions from standard machine learning methods in terms of final model weights. To accomplish this, we reemployed the same two tasks from above, calibrating a model with a sequence length of 1000 for ETTh1 and 200 for the BTC task. The difference in sequence lengths is attributed to the higher number of input features in the BTC task, which significantly increases the complexity of standard model calibration. For both tasks, we trained the model to predict 128 steps ahead, utilizing multivariate inputs that include time information. We then leveraged the TLN’s capacity to be rewritten as a linear equation and visualized the weights in a heatmap. The heatmap axes represent the features and the sequence length. We compare the weights using TLN, linear regression, ridge [36], lasso [37], and ElasticNet [38] .

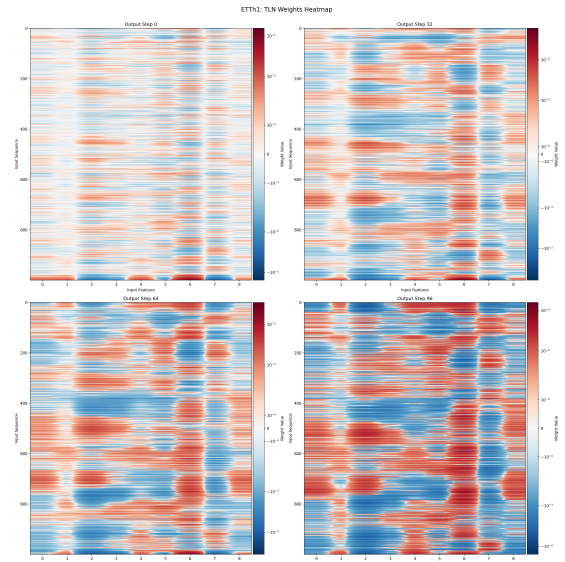


Fig. 2: TLN weights for ETTh1 task

In contrast, the TLN demonstrates a more balanced distribution of weights across all features, albeit with varying magnitudes. An interesting and expected observation is that weights for 1-step-ahead predictions are heavily biased towards recent events, while longer-term predictions incorporate more information from earlier elements in the sequences.

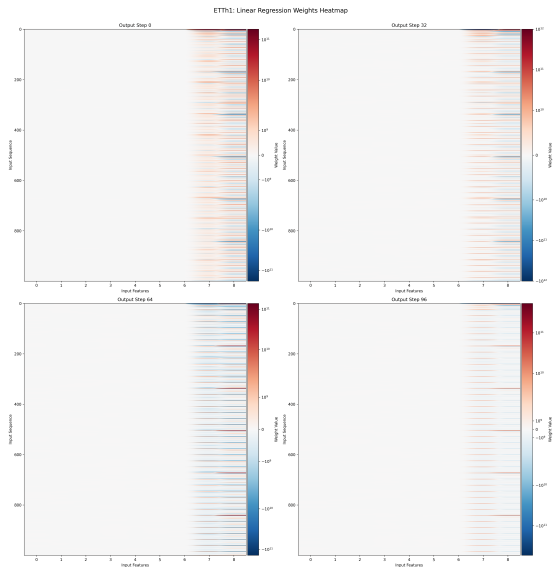


Fig. 1: Linear Regression weights for ETTh1 task

In the Linear Regression model, the weights appear to be disproportionately concentrated on the last two features, which represent temporal information. Notably, the magnitude of these values is exceptionally high, considering our targets are scaled to the $[0,1]$ range.

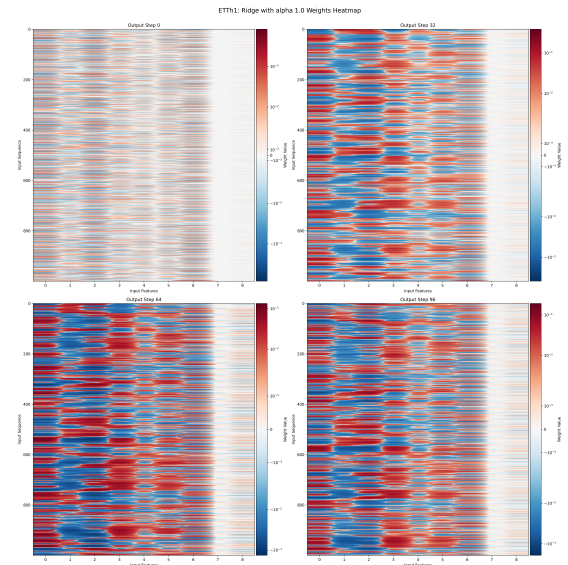


Fig. 3: Ridge Regression weights for ETTh1 task (alpha = 1.0)

The Ridge Regression model shows less dominance of temporal information compared to Linear Regression. However, it lacks the TLN’s characteristic of emphasizing recent values for short-term predictions, which is a logical consequence of the applied weight regularization.

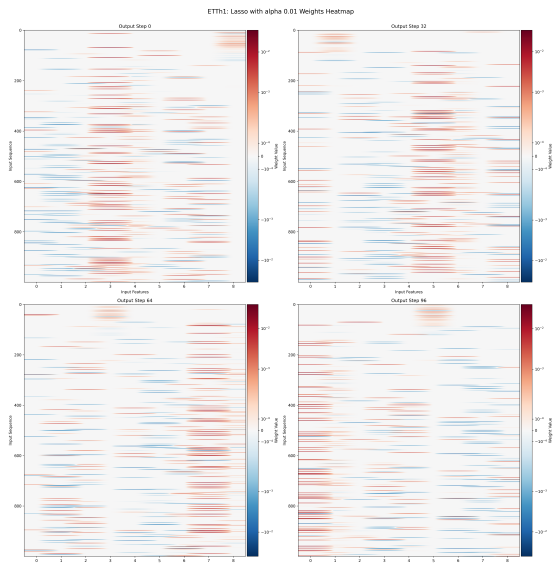


Fig. 4: Lasso Regression weights for ETTh1 task (alpha = 0.01)

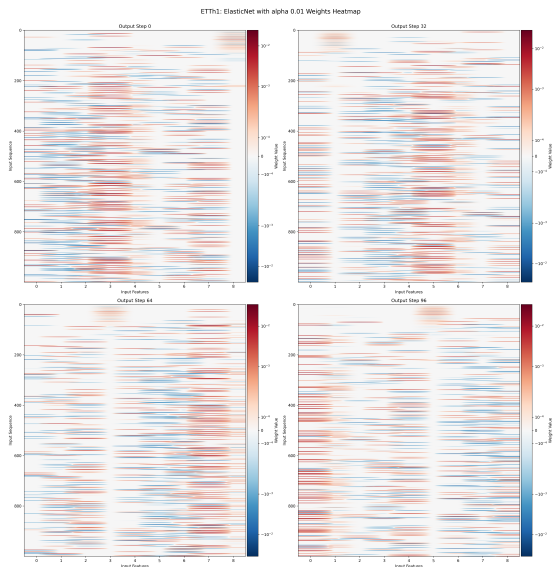


Fig. 5: ElasticNet Regression weights for ETTh1 task (alpha = 0.01)

Lasso and ElasticNet models exhibit distinct characteristics. While both utilize a broader range of parameters than Linear Regression, they demonstrate their strength in feature selection by setting many weights to zero. Interestingly, temporal information is only used at specific points in the sequence, which is logical given the continuous nature of the series. In practice, only one time feature would be necessary, although we included multiple for matrix formatting purposes.

For the ETTh1 tasks, Lasso and ElasticNet outperform the TLN, while Linear Regression and Ridge fall short. This aligns with our observations across all models: Linear Regression excelled in ETTh1 tasks but became less efficient

with increased inputs, while Lasso and ElasticNet’s corrective effects enhanced performance. Conversely, for the BTC tasks, the TLN outperformed other models, underscoring its value as a linear alternative that operates distinctly from traditional approaches. A noteworthy observation from our tests is the TLN’s robustness in handling larger inputs. While standard ML models can become computationally intensive to calibrate with low alpha values or numerous features, the TLN consistently performs well. Although it may require more time for smaller datasets, it demonstrates superior scalability when calibrating on substantially larger inputs.

VI. CONCLUSION

In this paper, we introduced the Temporal Linear Network (TLN), a novel architecture that bridges the gap between simple linear models and complex deep learning approaches for time series forecasting. Our extensive experimental results across multiple tasks and configurations demonstrate several key findings.

First, TLN achieves robust performance across varying input conditions while maintaining the interpretability advantages of linear models. The architecture’s ability to be transformed into an equivalent linear model provides transparency through explicit weight relationships. The successive transformations applied in the model (alternating between feature and temporal dimensions) create structured weight dependencies that act as an implicit regularization mechanism. This internal structure results in a more compact representation compared to standard linear models, requiring fewer parameters while maintaining or improving performance.

The reduced parameter space, combined with the structured transformations, explains both the model’s stability across different sequence lengths and its computational efficiency. Where a standard linear model would require a full $S \times F \times S' \times F'$ weight matrix for a transformation from an input of shape (S, F) to an output of shape (S', F') , our approach decomposes this into a series of smaller transformations. This decomposition not only reduces the memory footprint but also results in more efficient matrix multiplications during both training and inference.

Our comparative analysis with traditional machine learning methods reveals that TLN’s weight distributions show more balanced patterns than standard linear regression approaches. The model’s performance on both the ETTh1 dataset and Bitcoin volume prediction tasks demonstrates consistent accuracy across different prediction horizons, from short-term to long-term forecasting. This consistency is particularly noteworthy in multivariate settings, where many competing approaches show significant degradation in performance.

The key advantage of TLN lies in its ability to maintain the interpretability of linear models while achieving superior performance through its structured parameter space. This is accomplished without the computational overhead typically associated with deep learning approaches, making it particularly

valuable for real-world applications where both performance and interpretability are crucial.

Our results suggest that TLN represents a significant advancement in time series forecasting, offering an optimal balance between model complexity and interpretability. Its consistent performance across various scenarios, combined with its computational efficiency and theoretical guarantees of linearity, makes it a valuable addition to the existing toolkit of time series forecasting methods.

REFERENCES

- [1] A. Zeng, M. Chen, L. Zhang, and Q. Xu, "Are transformers effective for time series forecasting?" *arXiv preprint arXiv:2205.13504*, 2023.
- [2] W. W. Wei, *Multivariate time series analysis and applications*. John Wiley & Sons, 2018.
- [3] M. Marcellino, J. H. Stock, and M. W. Watson, "A comparison of direct and iterated multistep ar methods for forecasting macroeconomic time series," *Journal of econometrics*, vol. 135, no. 1-2, pp. 499–526, 2006.
- [4] Z. Tang, C. De Almeida, and P. A. Fishwick, "Time series forecasting using neural networks vs. box-jenkins methodology," *Simulation*, vol. 57, no. 5, pp. 303–310, 1991.
- [5] J. D. Hamilton, "State-space models," *Handbook of econometrics*, vol. 4, pp. 3039–3080, 1994.
- [6] C.-J. Kim, "Dynamic linear models with markov-switching," *Journal of econometrics*, vol. 60, no. 1-2, pp. 1–22, 1994.
- [7] E. Zivot and J. Wang, "Vector autoregressive models for multivariate time series," *Modeling financial time series with S-PLUS®*, pp. 385–429, 2006.
- [8] H. Lütkepohl, "Vector autoregressive models," in *Handbook of research methods and applications in empirical macroeconomics*. Edward Elgar Publishing, 2013, pp. 139–164.
- [9] M. Binkowski, G. Marti, and P. Donnat, "Autoregressive convolutional neural networks for asynchronous time series," in *International Conference on Machine Learning*. PMLR, 2018, pp. 580–589.
- [10] K. Ma and H. Leung, "A novel lstm approach for asynchronous multivariate time series prediction," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–7.
- [11] D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski, "Deepar: Probabilistic forecasting with autoregressive recurrent networks," *International journal of forecasting*, vol. 36, no. 3, pp. 1181–1191, 2020.
- [12] B. N. Oreshkin, D. Carпов, N. Chapados, and Y. Bengio, "N-beats: Neural basis expansion analysis for interpretable time series forecasting," *arXiv preprint arXiv:1905.10437*, 2019.
- [13] S. S. Rangapuram, M. W. Seeger, J. Gasthaus, L. Stella, Y. Wang, and T. Januschowski, "Deep state space models for time series forecasting," *Advances in neural information processing systems*, vol. 31, 2018.
- [14] L. Li, J. Yan, X. Yang, and Y. Jin, "Learning interpretable deep state space model for probabilistic time series forecasting," *arXiv preprint arXiv:2102.00397*, 2021.
- [15] H. Inzirillo, "Deep state space recurrent neural networks for time series forecasting," *arXiv preprint arXiv:2407.15236*, 2024.
- [16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, E. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.
- [18] G. Box and G. Jenkins, *Time Series Analysis: Forecasting and Control*, ser. Holden-Day series in time series analysis and digital processing. Holden-Day, 1976. [Online]. Available: <https://books.google.fr/books?id=1WVHAAAAMAAJ>
- [19] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.
- [20] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [21] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE transactions on neural networks*, vol. 5, no. 2, pp. 157–166, 1994.
- [22] K. Cho, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [23] M. I. Jordan and R. A. Jacobs, "Hierarchical mixtures of experts and the em algorithm," *Neural computation*, vol. 6, no. 2, pp. 181–214, 1994.
- [24] S. Masoudnia and R. Ebrahimpour, "Mixture of experts: a literature survey," *Artificial Intelligence Review*, vol. 42, pp. 275–293, 2014.
- [25] S. E. Yuksel, J. N. Wilson, and P. D. Gader, "Twenty years of mixture of experts," *IEEE transactions on neural networks and learning systems*, vol. 23, no. 8, pp. 1177–1193, 2012.
- [26] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, "Kan: Kolmogorov-arnold networks," *arXiv preprint arXiv:2404.19756*, 2024.
- [27] H. Inzirillo and R. Genet, "A gated residual kolmogorov-arnold networks for mixtures of experts," *arXiv preprint arXiv:2409.15161*, 2024.
- [28] T. Lin, Y. Wang, X. Liu, and X. Qiu, "A survey of transformers," *AI open*, vol. 3, pp. 111–132, 2022.
- [29] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," *International Journal of Forecasting*, vol. 37, no. 4, pp. 1748–1764, 2021.
- [30] R. Genet and H. Inzirillo, "A temporal kolmogorov-arnold transformer for time series forecasting," *arXiv preprint arXiv:2406.02486*, 2024.
- [31] Q. Wang, B. Li, T. Xiao, J. Zhu, C. Li, D. F. Wong, and L. S. Chao, "Learning deep transformer models for machine translation," *arXiv preprint arXiv:1906.01787*, 2019.
- [32] S.-A. Chen, C.-L. Li, N. C. Yoder, S. O. Arık, and T. Pfister, "Tsmixer: An all-mlp architecture for time series forecasting," *Transactions on Machine Learning Research*, 2023.
- [33] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [34] R. Genet and H. Inzirillo, "Tkan: Temporal kolmogorov-arnold networks," *arXiv preprint arXiv:2405.07344*, 2024.
- [35] H. Inzirillo and R. Genet, "Sigkan: Signature-weighted kolmogorov-arnold networks for time series," *arXiv preprint arXiv:2406.17890*, 2024.
- [36] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.
- [37] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 58, no. 1, pp. 267–288, 1996.
- [38] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 67, no. 2, pp. 301–320, 2005.

APPENDIX A
BTC NOTIONNALS PREDICTION TASKS RESULTS

A. R2 scores

[H]

TABLE I: Univariate without Time: R2 Score (x1e-2)

n	sequence ahead	length	Linear		LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet			TSMixer		TSMixer	TSMixer	
			Regression	Ridge	Encoder Decoder	Shared Encoder Decoder	Last Sequence	Full Sequence	1 Layer linear	2 Layer linear	3 Layer ReLU	LinearNet	no conv	CLinear	NLinear	DLinear	1	2	no relu
1	3	41.15	41.02	<u>39.68</u>	<u>37.71</u>	35.88	40.24	40.32	40.74	<u>39.92</u>	40.92	40.69	40.52	40.89	40.72	<u>41.26</u>	<u>42.22</u>	<u>40.84</u>	32.43
	10	<u>41.08</u>	40.94	<u>37.56</u>	32.44	<u>40.46</u>	39.55	40.57	40.73	33.75	40.32	40.70	40.65	41.02	40.79	32.19	39.95	33.08	39.69
	20	<u>41.51</u>	41.37	33.50	31.87	37.38	38.41	41.13	<u>41.37</u>	33.90	41.17	41.04	40.98	<u>41.19</u>	40.97	40.27	29.82	40.67	<u>40.97</u>
	50	<u>41.64</u>	<u>41.51</u>	37.01	28.10	38.37	40.15	<u>41.41</u>	41.01	32.73	<u>41.52</u>	<u>41.36</u>	40.99	26.60	40.75	39.16	35.00	30.87	29.07
	100	<u>41.57</u>	41.44	31.96	33.51	38.69	<u>40.50</u>	41.05	40.88	30.02	41.36	40.90	40.79	31.78	<u>40.99</u>	33.15	20.71	21.85	30.14
	200	nan	nan	36.15	31.68	38.42	38.68	40.99	40.69	21.19	<u>41.29</u>	40.98	<u>41.03</u>	40.23	40.86	29.74	22.44	-4.80	37.20
400	nan	nan	32.21	30.77	36.34	36.62	40.12	40.31	29.74	40.39	<u>40.52</u>	40.02	38.68	39.48	13.69	14.23	9.87	33.47	
6	3	13.47	13.46	13.66	<u>12.99</u>	14.53	14.21	13.30	13.26	<u>14.58</u>	13.37	12.87	13.11	11.79	12.78	14.08	7.97	8.30	<u>13.88</u>
	10	14.80	14.77	9.31	7.70	<u>15.41</u>	15.30	14.42	14.47	<u>13.48</u>	13.95	14.17	14.52	14.10	14.41	7.66	15.10	7.54	8.09
	20	16.97	16.95	15.57	6.15	16.51	<u>17.06</u>	16.76	16.66	12.30	16.68	16.66	16.60	16.53	16.65	<u>16.98</u>	<u>16.21</u>	15.96	7.91
	50	<u>17.30</u>	<u>17.29</u>	-7.38	12.10	15.78	13.17	17.02	17.03	12.74	17.23	16.92	16.90	7.72	16.89	16.41	13.69	16.00	13.42
	100	17.02	17.02	-10.49	11.61	16.53	<u>18.01</u>	16.79	16.80	11.41	16.90	16.89	16.40	<u>16.57</u>	16.57	12.68	5.63	-26.01	6.41
	200	nan	nan	0.73	12.08	16.53	<u>18.77</u>	<u>17.63</u>	<u>17.85</u>	12.16	<u>18.19</u>	<u>17.82</u>	<u>17.14</u>	13.18	<u>17.65</u>	5.77	10.08	<u>16.53</u>	-31.98
400	nan	nan	<u>16.99</u>	<u>10.95</u>	<u>16.54</u>	<u>18.48</u>	16.66	16.73	9.67	17.39	17.66	16.85	16.10	16.73	0.28	-55.01	15.01	8.64	
15	3	7.05	7.06	<u>8.27</u>	0.03	7.72	8.10	6.65	6.88	7.58	7.02	6.74	6.84	5.50	6.84	7.29	3.29	7.58	6.95
	10	9.38	9.38	8.20	6.64	9.76	<u>10.04</u>	9.16	9.33	<u>8.35</u>	9.26	9.25	9.21	9.00	9.24	3.61	2.87	3.38	9.46
	20	<u>11.74</u>	11.74	0.83	0.93	11.67	<u>12.39</u>	11.50	11.56	4.79	11.53	11.37	11.45	11.28	11.46	4.72	<u>11.36</u>	<u>10.53</u>	<u>10.81</u>
	50	11.73	<u>11.75</u>	-14.12	6.10	<u>12.76</u>	11.30	11.40	11.56	3.50	11.17	11.15	11.45	11.18	11.26	<u>12.26</u>	2.68	5.54	10.71
	100	11.43	<u>11.46</u>	-13.85	7.17	11.05	7.63	11.22	11.07	6.05	11.31	11.31	10.78	10.98	11.23	-1.63	7.12	-9.99	-42.14
	200	nan	nan	2.47	0.19	<u>14.50</u>	<u>14.31</u>	<u>12.18</u>	<u>12.25</u>	3.39	<u>12.84</u>	<u>12.64</u>	<u>12.14</u>	10.70	<u>12.29</u>	9.16	6.77	9.17	-7.62
400	nan	nan	<u>12.47</u>	<u>7.53</u>	<u>13.23</u>	<u>13.35</u>	11.53	11.68	-1.69	12.68	<u>12.79</u>	11.58	<u>11.39</u>	11.50	-12.58	-13.82	10.11	6.23	
30	3	5.54	5.54	6.37	6.03	6.23	<u>6.52</u>	5.26	5.32	6.43	5.27	5.40	5.35	4.76	5.38	5.52	5.99	5.54	5.95
	10	7.60	7.60	8.04	-0.19	<u>8.43</u>	8.43	7.30	7.33	<u>7.95</u>	7.37	7.46	7.41	7.02	7.41	2.59	2.42	<u>7.44</u>	2.75
	20	9.01	9.01	7.55	<u>6.50</u>	<u>9.82</u>	7.96	8.92	8.75	2.80	8.64	8.80	8.82	8.39	8.69	3.30	<u>8.88</u>	2.48	3.34
	50	<u>9.17</u>	<u>9.19</u>	8.73	-0.01	9.15	8.94	8.86	8.91	3.95	8.77	8.93	8.91	8.70	8.92	<u>8.19</u>	5.25	5.73	<u>7.59</u>
	100	8.88	8.91	1.89	-0.14	9.05	<u>10.10</u>	8.68	8.70	2.89	8.90	8.88	8.61	8.61	8.70	-0.10	4.60	0.50	7.17
	200	nan	nan	9.22	0.02	9.13	8.91	<u>9.81</u>	<u>9.82</u>	1.25	<u>10.26</u>	10.22	<u>9.81</u>	<u>9.74</u>	<u>9.80</u>	-32.94	7.60	-27.45	7.06
400	nan	nan	<u>11.21</u>	-0.19	9.22	-2.56	9.25	9.42	-6.61	<u>10.50</u>	<u>10.28</u>	9.61	8.87	9.39	4.13	-15.09	-0.34	-30.88	
90	3	3.19	3.20	<u>3.65</u>	-0.24	3.63	3.62	3.07	3.11	3.63	3.02	-0.01	2.82	2.26	3.11	2.98	2.72	2.85	2.87
	10	4.65	4.65	4.43	2.97	4.40	<u>5.23</u>	4.43	4.48	<u>4.75</u>	4.43	4.46	4.57	3.76	4.41	4.16	1.18	1.28	4.27
	20	5.58	5.58	4.97	3.96	<u>5.83</u>	<u>5.91</u>	5.45	5.38	2.94	5.41	5.40	5.35	4.82	5.28	1.69	4.74	1.46	4.27
	50	5.85	<u>5.87</u>	4.83	3.65	5.19	5.56	5.45	5.65	2.19	5.80	5.73	5.50	0.22	5.44	<u>4.25</u>	<u>5.34</u>	4.18	4.41
	100	<u>6.49</u>	<u>6.51</u>	5.14	-0.54	5.49	5.09	6.27	6.27	3.76	<u>6.61</u>	6.53	6.34	6.01	6.31	3.95	2.53	2.29	-17.78
	200	nan	nan	<u>7.46</u>	2.31	5.22	<u>8.20</u>	7.58	7.79	1.24	7.98	8.18	7.87	4.98	7.85	3.42	-7.09	<u>6.81</u>	-35.22
400	nan	nan	5.33	<u>4.25</u>	5.50	0.91	<u>7.79</u>	<u>8.06</u>	0.14	<u>9.07</u>	<u>8.92</u>	<u>8.17</u>	<u>8.04</u>	<u>8.42</u>	-1.92	-12.81	-1.08	<u>7.33</u>	
180	3	3.33	3.33	3.52	3.39	3.63	3.41	2.93	3.22	<u>3.89</u>	3.03	-0.04	3.04	2.23	3.14	0.93	0.11	3.10	0.90
	10	4.98	4.98	4.93	-0.03	5.01	<u>5.36</u>	4.84	4.83	5.16	4.79	4.70	4.90	3.80	4.65	4.42	1.75	4.49	4.52
	20	6.22	<u>6.23</u>	5.73	<u>4.62</u>	<u>6.00</u>	6.07	5.96	6.03	5.46	6.04	5.89	5.79	5.69	5.96	2.14	5.76	5.22	1.97
	50	7.23	<u>7.25</u>	5.63	-0.35	5.73	<u>7.38</u>	6.76	6.83	5.11	6.97	7.11	6.95	-0.74	6.99	6.12	0.21	5.08	4.16
	100	<u>8.37</u>	<u>8.40</u>	<u>5.74</u>	3.39	5.88	5.59	8.12	7.94	<u>6.25</u>	8.12	8.14	8.09	7.29	8.10	<u>6.79</u>	-24.32	4.03	7.18
	200	nan	nan	5.30	3.88	5.76	7.02	10.08	10.02	5.53	<u>10.52</u>	<u>10.36</u>	10.37	9.15	10.11	6.13	<u>8.65</u>	-29.69	-7.09
400	nan	nan	5.69	-0.01	5.78	<u>10.46</u>	<u>10.28</u>	<u>10.47</u>	-0.56	<u>11.34</u>	-0.02	<u>11.04</u>	<u>10.56</u>	<u>10.29</u>	6.00	5.07	<u>10.96</u>	<u>11.48</u>	
360	3	2.80	2.80	2.35	-0.35	<u>2.85</u>	2.70	2.40	2.25	2.82	2.37	-0.17	2.48	1.06	2.53	0.18	0.17	0.95	0.81
	10	4.04	4.04	2.49	1.39	3.34	3.73	3.52	3.59	3.76	3.47	3.45	3.52	2.30	3.49	1.49	0.09	1.01	1.42
	20	4.99	<u>5.00</u>	3.77	1.12	<u>4.48</u>	4.48	4.45	4.04	4.37	4.43	4.37	4.28	3.13	4.32	1.53	1.03	2.74	2.80
	50	5.83	<u>5.85</u>	3.38	-0.95	3.82	5.30	5.05	5.29	5.37	5.39	5.08	5.25	2.86	5.01	3.86	-1.72	3.93	1.71
	100	<u>7.09</u>	<u>7.12</u>	2.33	<u>1.63</u>	3.85	6.77	6.34	6.44	<u>7.56</u>	6.55	6.38	6.43	3.56	6.18	<u>5.43</u>	<u>1.31</u>	-3.14	-22.15
	200	nan	nan	<u>4.19</u>	1.58	3.85	<u>8.88</u>	7.49	7.99	<u>7.74</u>	8.30	8.21	<u>7.98</u>	5.58	7.84	-17.30	-8.22	1.88	<u>4.47</u>
400	nan	nan	2.42	-0.31	3.85	6.12	<u>7.91</u>	<u>8.27</u>	-4.19	<u>8.79</u>	<u>9.33</u>	7.97	<u>7.45</u>	<u>7.90</u>	-26.43	-6.83	<u>8.26</u>	-1.89	

TABLE II: Univariate with Time: R2 Score (x1e-2)

n	sequence ahead length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer		
		Regression	Ridge	Encoder Decoder	Shared Encoder Decoder	Last Sequence	Full Sequence	Layer linear	Layer linear	Layer ReLU	LinearNet	no conv	CLinear	NLinear	DLinear	1	2	no relu 1	no relu 2
1	3	41.37	41.24	39.07	32.98	40.38	38.88	40.80	41.12	40.05	41.17	40.57	3.31	41.04	41.18	42.00	42.62	34.95	35.02
	10	41.63	41.47	41.67	34.71	38.31	39.75	41.00	41.23	38.53	41.34	40.47	0.82	41.25	8.04	39.64	40.05	40.52	41.62
	20	42.80	42.66	37.33	35.81	40.38	40.78	41.61	41.84	40.73	42.09	41.00	41.44	41.48	41.54	42.72	41.35	40.33	30.41
	50	42.84	42.70	36.42	34.63	42.07	40.47	41.81	42.10	39.44	41.88	40.79	28.73	32.99	41.25	27.92	29.52	28.17	37.15
	100	42.64	42.51	40.83	37.46	37.90	40.48	40.29	40.65	38.69	41.80	38.76	40.86	33.08	8.81	25.40	30.25	39.55	27.18
	200	nan	nan	39.15	35.29	40.38	41.16	40.35	40.92	37.73	41.55	38.96	4.53	33.80	39.12	24.39	-0.17	35.41	19.76
	400	nan	nan	38.67	31.50	40.41	40.78	39.51	40.13	36.84	41.21	38.45	6.41	26.25	8.93	25.09	-6.93	36.99	37.61
6	3	14.62	14.61	15.24	13.85	15.71	15.87	14.35	14.52	17.70	14.30	13.31	13.05	12.12	4.47	14.14	9.89	14.21	13.79
	10	17.95	17.92	17.63	14.86	17.93	18.33	17.30	17.27	17.40	16.84	14.57	14.31	14.56	8.62	10.38	17.59	8.80	10.84
	20	20.44	20.42	17.63	14.84	19.30	20.69	19.14	19.31	14.75	18.32	15.89	15.75	15.74	15.81	5.70	19.82	20.44	9.93
	50	20.49	20.48	17.53	14.22	18.20	21.04	18.52	19.44	17.36	18.84	17.41	12.47	17.49	3.32	15.81	17.35	17.19	10.08
	100	20.05	20.05	14.90	14.00	18.94	20.47	17.53	17.93	16.49	18.87	17.70	7.12	9.36	3.60	7.49	15.19	15.20	15.62
	200	nan	nan	13.69	14.43	18.47	20.75	18.21	19.79	17.22	20.50	18.26	14.76	11.63	13.64	11.01	13.29	5.66	17.11
	400	nan	nan	17.58	15.47	18.58	20.64	18.05	16.96	17.85	20.27	17.10	11.21	13.78	15.53	14.75	8.72	14.43	10.76
15	3	9.05	9.06	11.54	8.82	11.01	11.32	8.48	8.55	13.33	8.72	7.40	0.87	6.28	0.05	5.87	4.68	5.07	8.40
	10	13.21	13.21	10.39	8.43	14.18	14.32	12.33	12.66	12.95	12.02	9.62	4.97	9.19	1.87	12.32	12.09	11.68	11.56
	20	15.11	15.11	11.83	10.97	15.02	15.72	13.62	13.81	7.73	13.11	11.28	4.86	10.98	3.31	12.66	4.66	13.45	12.76
	50	15.07	15.09	5.64	10.28	14.98	16.25	12.99	14.07	7.79	13.26	12.73	12.68	12.91	5.07	7.42	11.50	5.60	12.14
	100	15.24	15.26	-19.03	10.37	15.05	15.58	13.03	14.02	6.89	13.82	13.12	3.06	11.11	12.71	7.72	11.04	12.41	11.51
	200	nan	nan	1.42	10.06	14.84	16.22	13.88	15.19	12.10	15.72	13.46	12.43	10.43	11.43	8.79	8.06	2.71	-5.42
	400	nan	nan	9.56	9.30	15.32	16.18	12.78	13.74	13.35	15.42	14.49	10.26	12.66	6.49	13.13	-7.67	11.36	15.32
30	3	7.15	7.16	11.00	6.03	9.57	9.50	6.51	6.80	11.83	6.65	5.46	1.13	4.97	0.88	6.31	2.62	3.11	2.60
	10	10.61	10.60	13.14	7.23	11.69	10.94	9.85	10.22	11.99	9.94	7.61	7.05	6.96	0.06	4.34	4.64	8.56	9.00
	20	12.46	12.46	12.31	7.83	12.44	13.02	11.58	11.77	9.93	11.37	10.13	2.16	8.58	0.79	5.20	6.11	5.54	10.87
	50	12.62	12.64	9.69	8.96	12.34	13.67	11.43	11.50	6.59	11.47	10.67	10.69	9.93	3.41	7.60	7.36	7.51	6.88
	100	13.68	13.70	9.49	8.38	12.54	13.34	8.61	12.26	11.12	12.58	10.68	4.26	5.63	2.61	8.31	9.44	8.05	10.20
	200	nan	nan	9.45	7.71	12.19	14.07	10.26	12.91	12.00	13.59	11.73	6.64	10.74	11.54	8.31	4.50	11.78	3.87
	400	nan	nan	9.28	7.37	11.86	13.84	10.41	11.23	11.67	13.76	12.90	5.42	9.64	6.70	11.98	-5.86	1.10	-5.52
90	3	4.91	4.92	8.01	4.32	7.61	7.64	4.57	4.66	10.71	4.54	2.85	0.22	2.25	1.07	2.92	2.96	0.92	0.63
	10	7.97	7.97	11.20	4.31	9.16	9.20	7.10	7.44	10.81	7.50	4.40	2.65	3.85	2.30	5.03	3.29	6.01	5.49
	20	9.65	9.65	12.41	5.76	9.56	10.52	8.53	8.85	10.61	8.72	7.06	2.66	6.59	3.78	7.52	7.46	7.31	7.92
	50	10.67	10.68	10.52	5.32	10.15	11.20	8.93	9.82	10.66	9.83	8.28	3.92	6.95	4.53	7.94	8.66	6.62	5.52
	100	12.31	12.34	11.25	5.61	9.84	11.56	10.41	11.00	11.02	11.43	8.63	4.91	7.43	5.66	10.33	9.82	9.08	8.28
	200	nan	nan	11.26	5.90	9.69	12.31	10.26	11.21	11.61	12.42	11.20	5.64	8.21	6.45	10.63	4.58	9.61	11.23
	400	nan	nan	10.40	6.19	9.35	12.48	10.43	9.92	12.21	12.37	12.46	8.70	9.22	6.05	9.68	9.46	11.10	5.66
180	3	5.67	5.67	4.51	1.30	8.38	8.57	5.37	5.40	11.17	5.07	1.26	2.15	2.13	0.69	2.65	0.25	0.53	-0.16
	10	9.27	9.27	6.09	1.35	8.62	10.41	8.68	8.65	12.18	8.69	4.91	2.23	3.92	2.73	2.88	2.10	6.29	2.36
	20	11.44	11.45	6.37	4.97	9.48	11.63	10.08	10.62	12.67	10.72	8.25	3.25	7.98	0.73	8.08	5.62	5.18	4.30
	50	12.60	12.62	6.47	3.79	10.08	12.51	11.20	11.41	12.29	11.15	9.45	4.98	6.23	5.45	5.49	6.93	5.31	9.35
	100	13.75	13.78	5.94	3.13	9.51	12.80	11.57	12.08	12.82	12.41	9.98	6.05	9.08	5.45	10.16	9.37	9.61	11.01
	200	nan	nan	6.74	4.98	10.21	14.61	10.21	12.94	14.81	14.20	13.55	10.97	9.70	7.73	13.11	9.32	12.32	13.54
	400	nan	nan	6.43	0.18	10.45	14.41	8.80	10.00	14.87	14.83	14.78	8.87	11.69	9.24	8.12	3.96	12.76	11.46
360	3	5.54	5.54	2.16	0.73	6.31	5.70	4.91	4.77	11.61	4.72	1.46	2.17	1.13	2.48	-0.06	0.73	0.36	0.25
	10	8.77	8.78	3.36	0.12	8.26	9.35	7.61	7.29	11.74	8.08	3.15	0.32	1.81	2.08	2.69	0.70	2.51	2.04
	20	10.72	10.73	3.93	-0.83	8.10	9.89	8.96	9.33	11.75	9.39	5.47	3.72	4.34	4.43	4.81	3.73	4.48	4.15
	50	11.80	11.82	2.80	-0.09	8.18	10.62	9.11	9.86	11.50	10.32	7.53	5.42	6.12	3.50	5.88	6.10	5.68	6.57
	100	13.18	13.21	3.97	-0.73	8.11	11.93	10.61	10.73	11.91	11.28	6.77	6.73	5.96	6.82	8.63	6.74	8.66	7.92
	200	nan	nan	2.32	0.79	8.05	12.76	10.96	11.47	12.50	12.44	12.47	9.12	7.26	9.67	11.15	10.29	10.56	10.63
	400	nan	nan	2.79	-0.56	8.26	11.85	8.96	9.53	11.60	12.59	11.63	10.06	10.06	9.09	11.58	8.77	-0.15	-13.34

[H]

TABLE III: Multivariate without Time: R2 Score (x1e-2)

n	sequence ahead length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer		
		Regression	Ridge	Encoder Decoder	Shared Encoder Decoder	Last Sequence	Full Sequence	Layer 1 linear	Layer 2 linear	Layer 3 ReLU	LinearNet	no conv	CLinear	NLinear	DLinear	1	2	no relu 1	no relu 2
1	3	40.80	40.51	41.14	39.82	38.69	40.66	38.80	40.59	34.61	39.37	39.13	40.06	40.15	39.91	41.18	42.47	32.76	33.81
	10	40.16	39.90	39.25	37.56	39.31	38.30	37.52	39.57	30.22	40.27	40.03	40.46	41.25	39.86	39.59	35.81	37.25	38.74
	20	39.80	39.69	39.98	39.33	39.10	38.08	37.21	38.36	33.45	41.22	40.58	40.51	38.06	41.01	19.10	34.98	36.36	38.90
	50	38.09	38.57	39.49	40.29	40.22	38.00	36.43	36.66	32.51	41.08	40.90	40.78	40.59	41.15	31.47	31.29	32.08	33.91
	100	35.61	37.06	39.54	39.91	40.46	39.54	33.15	34.42	28.88	41.12	40.86	40.49	40.42	40.97	31.11	26.25	16.79	33.36
	200	nan	nan	39.44	39.20	41.09	37.27	28.04	29.35	25.44	40.68	40.83	40.57	34.14	40.65	20.98	21.25	11.68	28.37
	400	nan	nan	39.45	39.38	40.08	34.23	17.96	17.77	23.11	40.27	40.39	39.61	40.29	39.22	-3.90	6.14	21.89	30.04
6	3	13.63	13.67	12.72	13.26	14.15	14.65	12.38	13.45	12.58	13.14	13.47	13.47	12.19	13.28	14.37	10.10	9.90	12.89
	10	14.05	14.16	-1.96	11.06	14.09	15.20	13.22	14.00	12.37	14.11	14.70	14.24	14.17	14.31	14.88	11.67	8.28	8.06
	20	15.50	15.75	11.58	10.53	16.38	16.02	14.17	14.65	12.08	16.91	16.86	16.82	16.62	16.45	1.37	11.91	8.31	5.07
	50	13.78	14.66	7.74	8.38	15.59	13.76	12.14	13.97	7.42	17.19	16.93	17.11	16.93	16.09	11.07	14.80	5.26	3.41
	100	10.89	12.65	13.26	12.78	16.10	15.30	10.08	13.02	5.73	17.13	17.14	16.84	16.38	16.80	14.36	9.95	-11.38	12.90
	200	nan	nan	12.24	9.13	16.55	13.78	4.82	5.78	4.07	17.10	18.21	17.82	17.38	17.63	-1.83	2.56	1.84	-2.88
	400	nan	nan	8.57	15.60	16.20	10.78	-5.13	-7.57	4.45	17.46	17.83	16.82	17.68	16.67	-19.56	-7.59	12.53	5.97
15	3	6.72	6.80	-12.46	5.13	7.26	7.45	6.18	6.78	4.20	6.86	7.01	6.60	5.81	6.20	3.38	6.28	3.47	6.37
	10	8.54	8.68	-9.55	5.31	8.01	8.52	7.26	8.35	1.32	9.09	8.86	8.44	8.56	7.74	8.92	1.86	3.62	9.02
	20	10.13	10.46	4.71	7.10	10.56	10.83	8.45	9.11	1.32	11.39	11.18	11.39	11.23	11.17	-0.70	-1.55	8.04	7.75
	50	8.67	9.62	-21.25	8.83	9.86	9.77	6.73	8.90	-6.50	10.99	11.32	11.15	11.12	10.63	8.97	9.20	7.91	7.93
	100	5.15	6.98	-7.06	-13.81	9.79	8.98	4.09	6.99	-9.21	11.08	11.60	11.14	5.00	11.08	3.93	3.28	8.44	7.34
	200	nan	nan	-0.06	7.55	10.17	9.95	0.11	3.28	-14.00	12.71	12.77	12.26	12.40	12.16	-0.41	-4.62	7.57	1.05
	400	nan	nan	-10.76	-0.88	-3.59	7.41	-8.39	-12.46	2.12	12.61	13.07	11.83	11.12	11.93	-2.45	-13.03	-8.07	-0.69
30	3	5.28	5.36	4.99	4.64	6.11	5.99	4.71	4.80	4.30	5.30	5.47	4.63	4.31	5.53	1.43	1.46	4.96	5.07
	10	6.75	6.93	3.84	3.20	7.43	7.52	5.93	6.29	2.90	7.24	7.10	7.33	6.96	7.22	2.32	1.65	7.19	6.73
	20	7.48	7.86	0.68	6.25	8.24	7.33	6.21	6.88	1.32	8.58	8.59	8.53	7.72	7.70	0.83	-0.72	7.66	4.12
	50	6.36	7.30	-7.36	6.38	8.30	6.37	4.92	6.70	-6.13	8.46	8.71	8.79	8.47	8.84	2.89	5.38	7.51	6.98
	100	2.91	4.71	1.69	4.12	8.17	8.15	2.28	5.14	-6.50	8.62	8.78	8.42	8.11	8.29	4.51	4.90	3.40	-2.71
	200	nan	nan	4.14	6.52	7.95	7.83	-2.71	0.90	-16.43	10.03	10.06	9.90	7.25	9.37	-3.31	2.78	-9.69	3.38
	400	nan	nan	-4.83	6.72	8.11	5.22	-13.41	-16.94	-9.67	10.07	10.64	9.43	9.59	9.85	-5.63	2.17	-0.69	0.63
90	3	2.58	2.65	2.61	2.10	2.95	2.58	1.79	2.26	-0.85	2.48	2.50	2.03	1.94	3.12	2.22	0.53	2.03	0.19
	10	3.78	3.93	1.93	2.91	3.19	3.92	2.78	3.12	-2.10	3.69	3.83	3.68	2.57	3.73	2.09	1.35	4.17	2.55
	20	4.43	4.74	2.69	3.35	4.07	4.68	3.05	3.48	-3.57	4.55	4.62	4.11	2.66	3.14	2.03	3.36	1.44	1.72
	50	3.17	4.02	1.74	3.93	3.88	5.12	1.55	2.43	-7.19	4.96	4.91	2.52	4.11	4.66	2.53	-12.62	5.05	-2.46
	100	-0.63	1.31	2.40	3.33	4.20	4.99	-2.70	-0.31	-9.16	5.72	5.46	3.56	3.94	4.29	-7.68	5.00	-2.95	3.76
	200	nan	nan	1.59	2.92	4.40	4.20	-7.05	-5.60	-11.80	6.84	7.31	7.02	6.01	5.06	1.72	-8.59	1.84	2.10
	400	nan	nan	2.59	3.60	3.57	-1.96	-29.53	-31.37	-14.51	3.35	7.72	5.63	6.88	4.32	-0.98	-7.08	-5.75	-2.01
180	3	2.19	2.29	2.22	2.51	2.20	2.36	0.27	1.08	-2.53	2.36	1.93	1.85	0.88	0.59	-3.36	-0.07	-0.42	-0.83
	10	3.36	3.57	1.81	2.05	3.25	3.60	1.09	1.75	-5.87	3.38	3.64	1.48	2.94	1.19	0.20	-2.77	3.12	0.86
	20	3.95	4.36	2.69	2.82	3.94	4.33	1.02	0.93	-7.34	4.45	4.31	1.81	5.05	5.88	3.08	1.61	3.26	-1.02
	50	2.50	3.54	3.00	3.26	3.89	2.92	-2.58	-1.15	-13.86	5.51	4.51	1.81	4.74	0.34	-1.36	3.04	1.98	-6.84
	100	-2.09	0.23	-0.12	2.05	3.29	-0.10	-9.33	-6.15	-11.75	4.94	5.54	1.00	-0.86	5.37	-20.33	4.51	-2.22	-3.13
	200	nan	nan	-0.81	2.58	1.67	2.40	-21.21	-18.84	-6.77	2.42	8.14	7.93	-0.17	2.99	1.92	-5.74	-7.18	-1.94
	400	nan	nan	1.88	1.72	2.87	-27.58	-82.84	-94.04	-9.05	-19.66	5.89	4.50	3.57	4.99	-10.34	-10.17	-19.69	-18.58
360	3	-0.11	0.25	-0.96	-0.43	0.66	-0.17	-0.77	0.20	-9.54	1.51	1.43	0.36	0.93	-0.62	0.37	-17.90	-1.80	-12.36
	10	-0.34	0.12	-2.68	-1.18	0.71	0.42	-1.65	-0.88	-13.59	-0.07	1.52	0.32	0.03	0.01	-6.16	-3.35	-5.22	-2.07
	20	-0.84	-0.13	-10.98	-3.39	1.18	0.23	-2.91	-2.30	-19.42	0.90	2.09	1.43	-0.13	3.69	-5.23	0.02	-1.51	-3.53
	50	-4.45	-2.80	-16.71	-4.85	-4.15	-3.02	-8.28	-7.23	-27.94	-1.09	1.65	-1.74	-8.32	1.60	-7.01	-12.56	-6.91	-8.74
	100	-11.72	-8.15	-13.90	-6.32	-2.69	-3.98	-20.69	-18.50	-27.88	-5.11	0.52	0.56	-1.57	5.52	0.02	-23.24	-9.90	-3.11
	200	nan	nan	-23.79	-20.17	-12.20	-10.60	-36.56	-32.77	-16.33	-10.39	3.18	4.73	-2.30	3.84	-10.34	-1.39	-1.27	-4.26
	400	nan	nan	-17.93	-23.18	-8.22	-36.51	-100.34	-104.08	-11.77	-42.35	-1.47	2.41	-1.63	5.24	-7.79	-5.99	-28.90	-55.32

[H]

TABLE IV: Multivariate with Time: R2 Score ($\times 1e-2$)

n	sequence ahead length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer		
		Regression	Ridge	Encoder Decoder	Shared Encoder Decoder	Last Sequence	Full Sequence	Layer 1 linear	Layer 2 linear	Layer 3 ReLU	LinearNet no conv	CLinear	NLinear	DLinear	1	2	no relu 1	no relu 2	
1	3	40.83	40.52	<u>40.95</u>	<u>39.93</u>	37.19	37.11	38.90	<u>40.30</u>	<u>37.36</u>	39.17	39.13	39.55	<u>39.80</u>	39.58	<u>37.00</u>	<u>35.86</u>	38.48	40.64
	10	40.42	40.13	39.45	37.35	<u>41.62</u>	36.15	38.23	40.02	30.64	39.53	40.45	40.32	38.97	<u>40.55</u>	36.08	34.94	<u>39.75</u>	38.05
	20	<u>41.13</u>	<u>40.96</u>	40.25	37.29	40.07	38.04	<u>39.00</u>	40.27	33.35	<u>41.34</u>	<u>41.20</u>	<u>41.21</u>	38.93	38.49	22.56	23.91	36.31	<u>41.01</u>
	50	39.33	39.76	39.87	39.22	38.96	<u>38.53</u>	35.99	38.28	31.67	<u>42.09</u>	40.70	41.07	29.37	32.99	35.43	21.35	35.71	21.31
	100	36.38	37.74	39.91	38.64	40.16	37.30	33.86	34.44	28.05	40.57	39.51	<u>40.88</u>	38.15	35.04	22.35	26.50	34.10	35.83
	200	nan	nan	38.37	37.88	40.65	37.81	28.68	29.25	28.10	<u>41.63</u>	39.49	40.90	27.81	18.87	24.60	14.93	29.58	17.19
	400	nan	nan	40.40	38.01	40.12	36.84	19.23	19.69	23.33	<u>40.50</u>	39.61	38.77	38.26	7.60	15.36	3.02	25.98	23.99
6	3	14.45	14.50	11.10	11.99	15.37	<u>16.00</u>	13.77	13.83	<u>14.58</u>	14.33	13.54	12.60	12.20	13.51	10.41	<u>14.91</u>	12.65	14.13
	10	16.61	16.73	9.23	15.11	16.29	<u>17.43</u>	11.15	15.99	13.04	16.97	14.15	14.71	14.35	14.23	11.09	5.56	15.46	10.52
	20	<u>18.86</u>	<u>19.06</u>	<u>16.22</u>	<u>17.28</u>	<u>17.90</u>	<u>19.57</u>	<u>17.25</u>	<u>16.94</u>	14.45	18.90	15.83	15.78	16.35	16.60	<u>17.60</u>	<u>13.32</u>	<u>15.88</u>	14.44
	50	16.68	17.48	8.50	16.83	17.75	<u>19.72</u>	14.96	14.96	10.99	19.09	17.22	17.26	<u>16.57</u>	17.24	5.73	11.93	9.38	<u>16.79</u>
	100	12.85	14.57	11.44	15.99	17.40	<u>19.25</u>	10.99	11.68	14.46	19.12	<u>17.74</u>	<u>17.81</u>	16.05	<u>17.91</u>	15.45	6.09	9.44	9.69
	200	nan	nan	15.41	15.50	17.87	18.64	7.95	5.79	13.48	<u>20.09</u>	15.23	13.90	16.20	14.28	8.61	12.48	6.33	1.92
	400	nan	nan	6.67	14.24	17.38	16.50	-1.06	-3.17	10.57	<u>19.79</u>	16.76	13.55	16.36	4.34	1.21	-6.11	-0.67	5.66
15	3	8.55	8.63	1.15	5.90	9.77	<u>10.34</u>	7.96	8.13	10.08	8.68	7.23	6.47	5.46	7.26	8.21	8.14	8.05	6.68
	10	12.13	12.28	-5.76	10.17	12.60	<u>13.37</u>	11.15	11.06	7.57	11.91	9.11	9.25	8.98	9.37	<u>11.17</u>	4.01	<u>10.08</u>	<u>10.31</u>
	20	<u>13.47</u>	<u>13.76</u>	-2.39	<u>11.94</u>	12.74	<u>13.96</u>	<u>11.75</u>	<u>11.83</u>	6.69	13.86	11.96	12.05	11.46	12.12	5.83	<u>11.52</u>	6.34	4.79
	50	11.37	12.26	3.19	10.19	13.27	<u>14.27</u>	9.43	10.29	6.93	13.87	12.50	12.64	11.62	12.39	7.81	9.24	9.84	7.26
	100	7.91	9.67	-35.14	8.73	12.33	12.90	5.82	7.36	10.35	<u>13.47</u>	12.94	12.39	12.09	<u>12.67</u>	10.93	9.23	9.59	8.39
	200	nan	nan	6.00	10.65	<u>13.56</u>	13.93	3.22	2.69	<u>10.66</u>	<u>15.09</u>	12.29	12.10	11.00	11.52	10.27	1.53	-1.18	6.10
	400	nan	nan	-14.51	9.67	12.61	11.12	-7.46	-9.77	6.69	<u>14.86</u>	<u>13.76</u>	<u>13.30</u>	<u>12.36</u>	12.44	2.85	3.09	6.77	7.77
30	3	6.76	6.84	-0.73	6.66	7.30	8.61	6.28	6.40	<u>9.60</u>	6.75	5.64	4.76	4.71	5.26	5.67	3.34	5.88	5.80
	10	9.45	9.63	1.84	6.30	10.09	<u>10.73</u>	8.20	8.58	7.28	9.72	7.11	7.40	6.84	6.69	6.86	6.36	5.13	8.06
	20	<u>10.66</u>	<u>11.00</u>	0.14	9.78	<u>11.35</u>	<u>12.13</u>	<u>9.05</u>	<u>9.26</u>	5.17	11.20	9.68	8.60	8.82	9.46	<u>9.43</u>	7.13	6.64	6.85
	50	8.91	9.81	8.33	9.65	11.12	<u>11.63</u>	7.25	8.45	3.65	11.03	10.32	10.17	9.87	10.40	1.84	6.53	7.79	2.01
	100	6.82	8.62	6.69	<u>10.80</u>	10.47	<u>10.77</u>	4.04	5.95	<u>9.88</u>	<u>11.55</u>	9.85	4.09	8.85	10.09	1.54	7.98	4.26	8.38
	200	nan	nan	9.99	9.10	11.23	10.38	1.21	0.42	4.16	<u>12.48</u>	10.41	4.94	10.11	<u>11.30</u>	4.04	<u>8.37</u>	<u>8.90</u>	3.07
	400	nan	nan	6.65	9.69	11.18	9.86	-10.06	-14.77	3.98	<u>12.25</u>	<u>12.02</u>	<u>11.71</u>	<u>11.67</u>	10.35	3.59	3.35	6.28	<u>8.76</u>
90	3	4.24	4.30	3.62	3.67	6.20	<u>6.45</u>	3.44	3.59	<u>6.20</u>	4.22	2.25	1.14	1.40	1.22	2.28	1.54	2.35	2.05
	10	6.80	6.95	4.76	4.80	<u>7.90</u>	7.63	5.23	5.59	4.82	6.45	3.85	3.39	1.94	2.11	4.10	0.46	3.32	1.01
	20	<u>7.98</u>	<u>8.28</u>	<u>5.38</u>	5.42	<u>8.08</u>	<u>8.88</u>	<u>5.89</u>	<u>6.77</u>	4.95	7.72	6.53	5.82	3.81	3.15	6.85	4.98	3.72	3.01
	50	7.43	8.24	4.14	5.14	7.48	8.73	4.69	6.30	2.53	<u>8.94</u>	6.95	4.60	5.01	4.32	4.10	5.98	<u>7.69</u>	6.15
	100	5.35	7.28	5.09	4.89	7.44	8.80	1.75	4.08	3.26	<u>10.12</u>	6.81	2.22	5.97	6.42	6.70	<u>9.18</u>	7.62	5.05
	200	nan	nan	2.04	<u>5.48</u>	7.71	<u>8.96</u>	-3.55	-2.61	3.04	<u>9.78</u>	9.68	<u>9.24</u>	7.81	<u>9.52</u>	<u>7.58</u>	0.01	-3.28	<u>7.76</u>
	400	nan	nan	4.22	5.11	7.10	8.00	-18.64	-27.44	-1.17	<u>10.06</u>	<u>9.99</u>	7.57	<u>9.98</u>	4.44	4.05	5.67	4.93	3.31
180	3	4.52	4.62	2.89	2.08	4.46	<u>6.35</u>	2.81	3.20	<u>4.28</u>	4.33	2.39	0.87	1.11	0.80	-1.13	2.43	2.26	-0.50
	10	7.47	<u>7.67</u>	3.09	3.40	7.04	7.01	<u>4.68</u>	5.14	-0.37	7.60	3.55	2.27	3.76	2.51	0.47	4.34	<u>4.16</u>	4.73
	20	<u>8.97</u>	<u>9.34</u>	3.66	4.46	<u>7.82</u>	7.92	4.52	<u>5.42</u>	0.41	8.84	6.93	<u>4.93</u>	4.01	3.88	4.48	4.82	2.88	<u>5.14</u>
	50	7.88	8.85	3.24	3.77	7.47	<u>7.96</u>	1.87	3.13	-0.14	<u>9.84</u>	7.50	3.67	5.55	5.08	-0.11	<u>7.73</u>	-11.01	-1.04
	100	3.74	6.04	2.06	3.64	6.46	7.15	-5.36	-3.42	2.48	<u>9.49</u>	6.50	4.24	6.59	5.74	<u>5.52</u>	4.80	1.34	2.02
	200	nan	nan	1.15	<u>5.01</u>	7.02	5.22	-16.28	-15.62	-0.60	5.38	<u>9.78</u>	-1.36	<u>9.94</u>	<u>13.65</u>	3.95	3.58	-3.72	-11.91
	400	nan	nan	<u>3.72</u>	4.93	7.38	-15.72	-73.55	-82.04	-9.67	-12.31	<u>8.45</u>	3.19	4.00	6.42	-6.81	-5.93	-7.53	-15.29
360	3	2.61	2.96	<u>0.36</u>	-0.17	3.46	<u>3.60</u>	<u>1.60</u>	1.74	<u>-3.29</u>	3.14	1.17	1.05	-3.25	0.74	-2.23	-1.10	-2.30	-0.62
	10	4.22	4.67	-1.39	0.18	3.20	4.32	1.52	<u>2.53</u>	-5.56	<u>5.92</u>	2.18	3.05	-0.22	1.65	-0.90	-0.85	-2.07	-1.87
	20	<u>4.67</u>	<u>5.35</u>	-3.97	0.21	<u>5.70</u>	<u>6.06</u>	0.81	1.40	-11.88	<u>6.91</u>	4.72	3.83	2.16	4.15	-3.05	0.27	-2.19	1.50
	50	1.71	3.26	-6.69	-0.53	4.73	3.75	-4.96	-2.89	-21.13	4.85	4.65	2.82	-1.87	<u>5.38</u>	-0.45	-2.80	-4.10	3.11
	100	-5.68	-1.37	-7.89	0.75	1.84	1.21	-15.02	-13.04	-18.44	3.89	2.16	<u>5.48</u>	-2.15	3.98	-1.15	-2.04	2.08	-11.48
	200	nan	nan	-17.98	<u>1.36</u>	-5.26	-1.09	-33.13	-28.96	-13.34	0.73	<u>6.24</u>	<u>7.47</u>	<u>6.13</u>	<u>8.99</u>	<u>3.34</u>	<u>0.49</u>	<u>2.24</u>	<u>3.13</u>
	400	nan	nan	-5.08	0.27	1.60	-30.51	-91.38	-100.45	-16.73	-22.82	-3.50	1.59	-8.06	<u>2.01</u>	-16.88	-14.82	-21.30	-10.91

B. Training Times

[H]

TABLE V: Univariate without Time: Training time

n ahead	sequence length	Linear		LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer	
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	1 Layer	2 Layer	3 Layer	no conv	CLinear	NLinear	DLinear	1	2	no relu 1	no relu 2	
1	3	0.0	0.0	6.6	9.8	7.8	7.1	5.4	3.3	4.7	7.2	6.1	3.7	4.7	4.2	10.1	10.9	11.3	15.1
	10	0.0	0.0	7.6	13.6	8.8	9.2	4.4	3.2	5.3	7.6	6.4	4.4	4.9	4.8	13.4	10.8	15.0	9.1
	20	0.0	0.0	9.0	12.5	12.8	16.6	4.8	4.1	4.9	7.4	7.3	5.0	5.4	4.4	10.4	14.6	9.7	8.9
	50	0.0	0.0	13.1	25.0	25.4	26.3	4.3	3.4	5.7	7.0	7.3	4.9	6.9	4.0	8.9	11.3	10.8	10.6
	100	0.1	0.0	23.7	60.1	38.4	30.0	4.4	3.8	5.7	8.2	8.4	7.0	8.8	4.1	9.1	11.9	12.1	12.1
	200	nan	nan	32.5	81.0	50.7	72.9	4.0	3.6	5.9	8.0	7.6	5.3	5.1	5.6	8.1	10.6	12.0	11.3
	400	nan	nan	59.2	128.1	144.0	195.2	4.3	4.9	5.3	7.2	7.2	4.3	4.8	5.7	8.1	10.1	9.8	9.2
6	3	0.0	0.0	8.9	9.9	10.1	9.4	4.5	5.7	8.5	7.3	6.5	6.6	5.0	4.9	8.9	16.4	13.5	12.7
	10	0.0	0.0	10.0	9.4	9.8	10.1	4.8	5.3	7.8	7.0	7.5	6.2	4.5	4.9	13.1	11.9	15.6	16.0
	20	0.0	0.0	11.2	8.8	18.4	14.5	6.0	6.3	8.4	7.0	6.3	5.3	4.5	4.6	9.0	9.4	10.6	15.8
	50	0.0	0.0	19.7	10.1	33.0	40.8	4.4	5.3	7.1	7.4	7.4	5.7	5.1	5.1	10.2	11.0	11.1	12.6
	100	0.1	0.1	30.4	14.9	50.9	49.7	5.8	5.9	8.4	7.8	9.4	5.5	6.4	6.6	9.7	10.5	13.7	10.0
	200	nan	nan	47.1	20.6	73.0	86.1	5.3	6.7	7.7	8.3	6.9	5.7	7.3	6.2	7.9	9.7	9.5	9.2
	400	nan	nan	85.5	163.9	144.3	134.4	5.5	6.4	5.8	7.2	7.0	5.9	4.6	7.9	7.5	9.5	9.4	9.5
15	3	0.0	0.0	10.8	12.0	8.9	11.0	5.5	8.6	10.8	7.2	8.9	5.7	4.5	5.0	7.9	15.2	9.0	8.5
	10	0.0	0.0	12.9	15.0	12.0	11.3	4.7	7.8	10.7	7.2	6.0	5.7	6.7	5.1	13.0	14.5	11.4	9.7
	20	0.0	0.0	13.6	12.2	17.2	15.1	5.5	6.6	8.8	8.3	6.5	4.9	4.9	4.6	12.5	11.7	9.3	9.6
	50	0.0	0.0	18.8	39.1	41.0	35.6	5.4	9.9	10.6	7.6	7.7	6.0	6.3	5.5	10.0	17.3	11.9	13.3
	100	0.1	0.1	25.9	34.2	60.7	59.3	5.2	7.9	10.2	7.5	7.8	5.9	5.0	5.7	9.6	10.9	13.5	11.6
	200	nan	nan	39.0	31.1	147.6	67.9	5.4	9.3	10.8	7.9	6.9	5.1	4.8	6.1	7.4	9.5	9.4	10.4
	400	nan	nan	60.4	39.6	246.2	124.4	5.4	9.2	9.2	7.9	6.5	5.2	5.1	7.3	7.6	9.3	9.1	8.5
30	3	0.0	0.0	9.8	21.2	7.9	8.4	4.0	6.0	8.5	6.5	6.1	4.0	4.9	4.5	8.1	12.4	9.9	11.5
	10	0.0	0.0	15.0	13.4	10.7	11.3	4.8	5.5	7.8	6.7	6.2	4.5	4.2	4.3	9.1	13.3	9.4	11.2
	20	0.0	0.0	14.8	13.5	18.6	17.3	5.2	5.6	8.1	6.8	6.7	4.6	4.4	4.4	14.1	9.4	15.3	11.9
	50	0.0	0.1	19.2	15.6	28.2	32.2	5.9	7.3	9.0	7.1	7.3	6.2	6.7	5.6	9.2	11.9	15.6	16.8
	100	0.1	0.1	31.5	31.0	52.8	47.5	6.3	7.7	8.5	7.5	8.5	6.2	7.6	7.0	10.4	12.1	12.7	14.0
	200	nan	nan	37.9	25.7	72.8	94.6	6.4	8.5	9.1	7.7	7.9	5.6	5.7	6.6	9.1	10.5	12.1	12.0
	400	nan	nan	60.2	53.4	123.3	155.3	6.2	7.9	6.7	7.1	6.8	4.8	7.1	7.5	7.5	9.3	9.1	9.1
90	3	0.0	0.0	17.1	33.8	8.5	8.7	4.3	9.3	12.0	6.3	5.1	3.1	3.7	4.3	7.2	12.0	8.8	8.1
	10	0.0	0.0	18.4	29.8	11.3	13.5	4.7	9.3	13.0	6.4	6.6	5.2	3.9	4.8	8.7	14.6	12.8	8.6
	20	0.0	0.1	19.6	33.3	17.3	14.8	4.7	8.2	11.6	6.4	6.0	4.4	5.9	4.6	8.7	12.6	13.0	8.8
	50	0.1	0.0	21.4	35.5	21.1	25.7	5.3	9.4	12.1	7.2	6.1	4.5	5.9	5.0	8.6	11.0	16.7	11.9
	100	0.2	0.1	26.1	69.8	42.3	56.9	5.5	9.4	12.5	7.5	7.0	5.2	6.8	6.2	10.1	10.6	10.7	10.5
	200	nan	nan	51.7	43.3	51.9	86.8	6.7	9.6	10.5	8.0	7.5	5.4	6.3	7.8	9.0	11.8	11.7	10.1
	400	nan	nan	54.7	52.3	132.3	118.2	5.4	9.0	9.4	8.3	6.8	5.4	5.0	7.7	7.2	9.4	10.5	9.4
180	3	0.0	0.1	28.6	46.2	8.1	8.3	3.2	6.6	8.6	6.4	5.0	3.1	4.9	3.8	7.8	12.2	8.3	9.5
	10	0.1	0.2	27.5	44.6	10.6	13.2	4.4	8.2	10.0	6.1	5.9	4.6	5.0	4.3	8.3	14.5	9.5	9.2
	20	0.1	0.1	29.5	48.1	13.8	13.1	4.0	7.9	9.4	6.8	5.3	3.4	6.5	4.5	11.5	8.6	11.2	11.4
	50	0.1	0.1	30.2	54.8	20.2	26.6	5.5	8.5	11.6	6.5	7.1	4.3	3.6	5.6	7.8	14.9	11.0	15.0
	100	0.2	0.0	34.7	61.7	30.7	45.5	5.8	9.3	11.1	9.2	6.5	4.6	4.7	6.1	7.7	10.7	9.9	9.5
	200	nan	nan	53.3	57.9	51.1	67.1	4.8	8.1	9.4	7.3	6.3	4.7	4.0	6.2	8.0	10.1	12.7	9.7
	400	nan	nan	57.7	69.0	101.6	118.6	5.9	9.3	8.7	8.3	5.6	5.0	7.8	8.6	8.5	10.9	10.9	9.9
360	3	0.1	0.0	52.3	84.1	8.3	9.1	3.4	5.9	17.4	6.6	5.6	3.1	5.2	4.2	8.8	11.1	8.9	9.5
	10	0.2	0.1	46.8	87.7	10.0	10.6	3.6	6.1	16.3	6.4	5.9	3.3	3.7	4.4	11.1	11.4	9.3	8.5
	20	0.1	0.1	47.4	83.8	13.7	13.6	3.8	6.1	18.6	6.5	5.1	3.2	3.3	4.6	8.3	11.2	9.2	8.7
	50	0.7	0.2	48.1	98.6	23.1	21.9	4.4	6.9	16.4	6.6	6.3	4.4	6.0	5.5	8.6	14.8	11.9	14.5
	100	0.8	0.2	52.1	90.8	30.7	34.9	5.0	7.1	19.6	7.1	5.7	4.8	4.1	5.7	7.4	9.7	10.6	12.2
	200	nan	nan	78.7	101.3	51.3	61.8	5.3	8.0	17.0	7.1	6.2	4.6	6.2	6.1	10.3	11.5	10.3	10.2
	400	nan	nan	79.7	142.5	95.9	121.9	4.8	7.0	14.7	7.6	8.3	5.3	4.3	8.4	10.1	11.0	9.9	9.2

TABLE VI: Univariate with Time: Training time

n	sequence ahead length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet			TSMixer		TSMixer	TSMixer			
		Regression	Ridge	Encoder	Shared	Last	Full	1	2	3	no	CLinear	NLinear	DLinear	1	2	no	no	
				Decoder	Encoder	Sequence	Sequence	Layer	Layer	Layer	conv				1	2	relu	relu	
1	3	0.0	0.0	7.5	9.3	7.5	8.5	8.6	4.7	4.9	7.8	7.0	7.6	8.1	11.8	15.7	12.6	17.4	13.6
	10	0.0	0.0	7.5	16.5	12.1	9.9	5.6	4.3	6.0	8.9	7.2	4.3	8.4	6.0	15.9	11.9	17.3	17.4
	20	0.0	0.0	11.1	12.4	16.0	14.9	8.5	5.5	6.0	8.6	6.1	8.7	7.0	12.6	16.3	18.6	21.1	15.8
	50	0.1	0.1	15.4	23.6	22.1	35.6	7.4	6.6	6.3	9.2	10.3	9.4	9.2	8.5	19.3	16.2	19.1	12.2
	100	0.6	0.2	24.1	51.8	62.4	53.6	7.7	5.5	7.4	11.2	9.9	9.3	6.9	5.0	22.7	21.5	16.1	15.0
	200	nan	nan	41.0	74.9	81.2	113.3	8.2	5.6	7.2	11.7	9.7	5.0	8.4	8.2	13.9	14.5	18.4	12.1
	400	nan	nan	65.4	101.6	208.6	224.9	5.1	6.7	8.4	13.0	7.6	6.6	5.8	4.8	17.8	12.7	19.4	11.7
6	3	0.0	0.0	11.1	9.6	9.6	10.1	8.0	5.3	6.8	9.4	8.4	8.3	5.4	4.8	15.4	17.7	13.9	18.8
	10	0.0	0.0	10.6	14.7	13.4	13.4	9.1	5.6	6.2	8.5	7.2	11.8	9.2	7.0	17.1	17.9	23.7	19.1
	20	0.0	0.0	11.3	17.0	21.9	20.9	8.7	5.6	8.4	8.7	7.9	13.2	10.6	8.7	18.0	25.9	17.8	19.3
	50	0.2	0.1	12.7	22.9	30.4	40.5	7.5	6.4	8.0	11.4	9.2	10.8	12.7	7.7	16.3	13.1	20.7	16.6
	100	0.8	0.2	21.9	34.4	56.9	70.5	8.4	6.6	7.6	12.2	10.0	11.5	8.5	8.9	21.5	21.0	17.3	18.4
	200	nan	nan	31.2	66.4	107.6	90.1	7.5	7.8	7.3	11.9	11.3	9.8	9.7	9.6	18.2	16.9	18.7	16.5
	400	nan	nan	42.8	80.5	187.0	231.0	8.4	7.3	8.1	14.1	14.3	13.3	9.6	16.6	17.9	20.6	23.0	15.1
15	3	0.0	0.0	11.1	13.3	9.4	9.7	7.0	6.6	8.1	8.9	6.3	6.0	8.1	4.8	12.3	18.9	19.6	15.6
	10	0.0	0.0	12.3	12.7	14.1	15.9	8.5	8.7	9.2	9.4	9.6	15.1	7.4	6.4	15.0	22.1	19.3	18.3
	20	0.0	0.0	16.8	17.5	16.5	21.4	8.2	6.8	9.0	9.2	8.8	7.7	8.3	7.9	17.4	15.4	18.9	20.3
	50	0.2	0.1	16.4	23.2	40.1	36.9	8.4	8.8	9.8	12.2	9.8	11.0	8.5	7.6	17.9	19.1	18.4	16.0
	100	0.7	0.2	26.3	32.3	52.5	46.8	9.0	9.7	9.7	12.2	11.1	8.1	7.4	11.6	21.2	22.2	23.2	20.9
	200	nan	nan	34.7	25.3	121.3	98.8	7.4	9.6	8.5	11.7	10.1	9.5	10.2	12.8	17.7	18.6	17.1	17.0
	400	nan	nan	53.2	123.7	181.3	181.7	7.2	10.1	9.8	11.7	15.5	14.5	11.8	19.4	21.0	17.0	19.0	22.9
30	3	0.0	0.0	16.0	20.2	12.2	10.7	6.2	5.0	6.9	9.3	6.7	7.4	5.3	6.3	14.9	16.5	17.8	14.5
	10	0.0	0.0	18.3	25.0	16.2	11.2	8.1	5.8	6.3	9.2	7.8	8.3	6.7	6.1	18.1	19.7	18.6	15.9
	20	0.0	0.0	17.3	22.0	19.6	21.9	9.2	7.5	6.7	10.5	8.6	11.7	8.0	7.0	20.2	25.6	24.2	24.3
	50	0.2	0.1	20.4	37.9	39.9	44.6	10.1	6.8	7.0	12.7	10.3	12.9	6.9	7.9	20.2	19.0	18.8	16.0
	100	1.4	0.5	28.6	24.9	81.4	56.0	7.7	6.7	5.8	14.3	9.7	11.0	8.1	9.0	21.4	21.1	22.5	15.6
	200	nan	nan	54.9	52.4	115.3	90.2	9.1	7.9	6.3	13.1	11.1	12.5	8.3	13.9	18.1	18.0	17.8	14.7
	400	nan	nan	60.1	65.8	214.4	165.3	9.1	6.8	7.8	11.1	11.5	10.2	6.5	9.7	19.3	14.1	16.5	14.1
90	3	0.0	0.0	33.8	53.9	10.9	11.7	5.7	6.3	9.0	8.0	8.3	6.0	4.1	5.2	13.0	14.1	16.5	18.6
	10	0.0	0.2	32.0	42.3	14.9	13.7	7.1	7.1	10.0	9.4	9.3	5.7	9.3	7.5	22.0	19.9	22.4	17.4
	20	0.1	0.1	41.8	48.8	18.8	19.4	9.9	7.5	8.9	10.9	9.2	7.9	8.0	7.6	18.0	21.0	15.9	16.4
	50	0.2	0.1	36.3	49.5	47.6	34.8	8.6	10.8	9.0	12.6	10.8	10.4	8.7	8.7	19.2	13.8	16.2	18.7
	100	1.0	0.2	41.1	128.5	78.0	48.7	10.3	8.7	9.0	12.8	11.4	11.0	12.4	12.8	19.6	18.3	21.9	17.2
	200	nan	nan	61.6	102.0	115.2	84.6	9.4	7.8	7.9	14.4	15.1	12.6	8.6	12.2	18.6	19.5	17.4	16.9
	400	nan	nan	132.6	133.4	240.5	180.9	9.4	7.7	9.5	9.1	13.1	14.4	9.2	9.8	19.9	17.7	18.0	14.1
180	3	0.1	0.1	29.7	65.7	10.1	10.9	5.7	6.6	8.4	7.8	7.9	5.8	7.3	5.0	10.8	15.1	12.5	12.9
	10	0.1	0.1	35.8	78.3	10.8	15.2	7.8	7.5	8.6	8.1	9.2	5.6	5.6	5.6	14.5	16.1	12.7	11.8
	20	0.1	0.2	33.8	101.8	15.4	16.7	7.5	7.3	7.1	9.7	8.4	6.9	10.8	6.1	16.9	12.1	15.3	14.2
	50	0.5	0.1	44.5	67.9	28.3	30.5	9.1	8.0	6.9	9.9	9.9	9.3	7.5	9.3	18.6	17.1	16.9	13.9
	100	1.2	0.2	33.6	81.2	40.7	45.4	8.8	7.4	7.9	10.3	9.4	8.4	7.3	7.6	20.4	18.2	20.1	12.4
	200	nan	nan	57.6	106.9	112.0	75.4	9.5	8.2	8.9	9.5	9.8	10.4	8.5	8.4	16.4	15.9	17.2	17.2
	400	nan	nan	75.9	86.8	197.0	159.0	7.1	8.0	9.5	8.8	12.8	11.6	9.2	11.4	17.5	16.0	18.6	16.4
360	3	0.1	0.0	45.2	142.4	10.7	9.3	5.0	4.8	14.7	7.9	8.4	4.3	4.9	5.3	15.0	15.8	12.1	12.8
	10	0.3	0.2	57.1	123.8	12.6	13.3	6.4	4.8	14.5	8.4	8.4	4.4	10.5	5.2	13.2	14.5	13.8	16.2
	20	0.3	0.2	69.7	116.8	12.4	15.5	6.3	4.4	13.2	8.6	8.6	8.2	8.3	5.9	15.3	17.9	15.7	15.2
	50	0.6	0.2	49.4	118.4	25.5	24.0	7.8	5.1	15.6	10.0	9.3	8.2	8.2	7.7	18.1	18.1	18.9	13.1
	100	1.4	0.3	70.0	125.7	42.0	47.3	8.9	5.0	13.8	9.2	9.7	10.7	7.3	6.1	17.4	18.3	17.7	17.8
	200	nan	nan	59.7	141.7	55.1	71.7	9.5	6.6	15.5	9.6	11.6	7.8	5.9	8.8	18.4	16.6	15.4	15.4
	400	nan	nan	89.3	161.3	166.7	101.9	7.9	8.3	14.0	8.4	11.7	10.2	7.8	9.0	18.6	15.4	17.6	17.0

[H]

TABLE VII: Multivariate without Time: Training time

n ahead	sequence length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet			TSMixer		TSMixer	TSMixer			
		Regression	Ridge	Encoder	Shared	Last	Full	1 Layer	2 Layer	3 Layer	no conv	CLinear	NLinear	DLinear	1	2	no relu	no relu	
1	3	0.0	0.0	6.9	7.2	7.4	7.6	3.9	3.9	4.6	7.9	6.7	6.3	4.7	6.0	12.9	13.2	15.0	12.7
	10	0.2	0.2	7.0	7.6	10.5	9.2	4.0	4.3	4.4	8.4	7.7	6.0	5.2	5.9	16.4	13.2	15.2	10.6
	20	0.9	0.2	10.0	9.7	13.6	12.9	4.2	3.8	4.8	9.7	7.6	5.4	5.8	6.4	17.6	11.9	15.1	12.4
	50	4.3	0.6	14.8	14.7	19.0	29.5	5.0	4.3	5.2	8.3	8.3	8.0	6.3	7.6	18.1	12.3	17.7	12.5
	100	11.0	1.0	18.0	22.3	41.9	43.6	5.1	4.7	4.4	9.9	8.8	5.4	7.3	6.4	19.1	12.8	16.2	12.8
	200	nan	nan	37.3	40.1	76.3	73.4	5.0	6.1	4.9	11.5	8.9	6.3	6.7	8.3	15.4	12.1	14.6	12.1
	400	nan	nan	51.8	73.5	121.3	111.3	6.1	6.4	6.2	15.2	15.7	8.3	8.7	7.7	18.3	14.5	16.1	11.6
6	3	0.0	0.0	7.6	8.7	8.1	9.2	5.3	4.7	5.7	7.8	6.5	7.6	5.2	5.5	13.5	17.5	15.4	14.6
	10	0.2	0.2	9.7	11.0	10.3	10.8	5.6	5.5	6.0	7.6	7.4	5.5	6.1	6.3	13.8	13.6	15.0	14.1
	20	0.8	0.2	11.2	11.7	20.5	17.7	5.5	4.7	5.8	8.3	8.6	6.4	6.7	8.2	16.6	13.6	13.5	11.7
	50	4.6	0.6	12.1	15.4	29.8	30.4	5.6	5.5	5.9	9.7	9.2	8.3	6.5	9.3	15.6	12.8	15.2	14.6
	100	11.7	1.0	16.8	25.0	60.6	50.2	5.8	6.5	6.1	10.8	11.2	8.9	6.9	11.0	16.4	14.2	16.3	13.8
	200	nan	nan	27.0	29.1	116.2	73.6	5.5	7.1	6.3	12.0	12.2	9.8	8.9	13.3	15.6	14.1	15.5	13.0
	400	nan	nan	44.5	57.6	183.2	133.1	5.1	7.4	7.3	14.2	15.0	16.7	10.9	21.4	19.1	18.6	19.6	15.9
15	3	0.0	0.0	10.5	10.3	7.8	8.2	5.9	6.5	7.7	8.1	7.1	5.2	5.4	5.9	13.0	15.2	15.6	13.2
	10	0.3	0.2	9.9	12.0	12.5	11.7	5.2	7.0	8.2	9.2	7.3	5.4	6.2	6.5	14.3	14.6	14.2	12.3
	20	1.1	0.2	10.6	14.1	18.0	15.1	5.1	6.7	7.8	10.6	8.5	6.9	8.0	8.8	16.1	12.8	13.3	12.0
	50	4.0	0.5	13.2	16.3	33.9	28.7	5.0	7.4	7.8	8.9	9.4	8.3	6.3	10.4	17.3	14.7	14.4	13.1
	100	12.2	0.9	17.7	24.0	54.4	42.7	5.4	6.8	8.0	10.0	10.4	9.0	8.9	9.1	15.6	13.9	16.6	12.6
	200	nan	nan	26.7	29.2	85.4	53.8	5.8	9.3	8.2	12.9	12.7	10.7	10.1	14.7	15.9	14.4	15.8	12.9
	400	nan	nan	57.1	54.4	294.8	105.8	6.3	9.3	9.7	12.4	17.6	16.8	15.4	22.2	22.6	21.0	22.7	18.8
30	3	0.0	0.2	9.9	15.1	8.3	7.7	5.4	4.4	5.5	8.5	7.0	5.4	4.9	6.1	14.7	15.9	14.8	12.6
	10	0.3	0.1	10.8	15.3	11.3	11.5	5.6	4.4	5.3	8.4	7.2	5.4	5.3	6.5	13.5	12.1	13.8	12.5
	20	1.1	0.2	16.3	16.0	17.8	18.6	5.1	5.1	5.4	9.9	8.5	5.9	5.2	7.2	15.6	13.9	13.2	11.5
	50	4.5	0.6	18.0	18.6	27.6	27.9	5.7	6.6	5.6	9.1	9.2	8.3	6.9	7.6	16.9	14.6	15.1	12.4
	100	12.7	1.0	21.8	27.8	53.3	38.4	5.4	6.0	5.8	9.9	11.2	7.5	7.9	12.0	15.6	14.8	17.1	14.9
	200	nan	nan	35.2	42.2	71.0	53.7	6.6	7.2	6.2	12.6	12.4	11.2	10.9	14.3	15.8	15.3	16.1	14.3
	400	nan	nan	60.0	47.8	210.0	105.3	6.7	6.9	7.2	11.4	12.7	9.4	9.8	13.5	16.3	15.0	16.0	12.5
90	3	0.2	0.0	16.1	37.5	7.5	7.9	5.7	6.9	8.2	7.8	7.7	5.2	4.8	6.4	12.2	15.3	14.1	11.5
	10	0.4	0.1	19.5	37.9	9.3	9.5	6.4	6.0	7.8	7.9	9.1	5.5	5.7	6.4	11.3	12.4	12.8	13.6
	20	1.2	0.1	19.3	38.7	13.7	11.9	6.6	6.2	7.8	8.0	9.1	7.4	4.2	7.9	13.6	11.6	13.4	10.5
	50	5.2	0.7	20.5	42.0	19.7	28.0	6.8	7.4	8.5	8.5	10.8	6.8	6.1	7.9	17.2	14.6	15.1	12.3
	100	13.9	1.0	26.2	46.8	30.6	33.7	6.1	7.5	8.4	10.0	11.7	7.9	6.5	10.3	19.9	15.7	16.8	13.9
	200	nan	nan	35.8	54.6	56.3	54.2	7.0	8.8	9.2	11.8	14.5	12.0	6.8	12.6	17.0	14.6	15.6	12.6
	400	nan	nan	55.0	74.8	99.7	99.8	6.0	7.6	10.4	11.1	16.5	11.4	10.8	14.4	16.7	16.7	17.6	14.7
180	3	0.2	0.2	28.7	67.3	8.0	8.0	4.4	5.7	7.6	8.1	8.5	4.1	6.3	5.4	10.6	15.6	14.7	14.3
	10	0.7	0.3	26.4	65.6	9.2	9.3	5.7	6.5	6.7	7.5	8.4	4.6	4.6	4.9	10.6	10.6	13.1	10.3
	20	1.6	0.2	26.5	70.6	12.5	12.7	5.8	6.4	7.3	8.1	8.9	5.3	5.1	8.3	12.2	11.7	13.3	14.9
	50	5.5	0.8	29.3	73.1	18.9	19.1	6.2	7.2	7.5	8.3	9.7	5.9	4.8	7.5	17.0	13.2	14.7	13.8
	100	14.2	1.3	33.8	77.8	30.8	32.5	6.3	7.3	7.9	10.5	10.2	7.5	5.4	9.3	17.1	14.6	16.5	12.9
	200	nan	nan	42.9	82.9	53.6	53.5	6.4	12.3	8.9	11.3	11.3	7.4	5.3	7.2	13.6	15.0	13.7	11.5
	400	nan	nan	61.0	99.7	99.8	99.7	5.5	10.9	11.8	10.3	14.3	9.1	7.3	10.8	17.3	15.9	17.0	14.9
360	3	0.4	0.3	42.8	126.0	7.7	8.3	4.5	4.0	13.8	8.6	8.9	4.7	4.9	5.4	11.5	12.8	13.7	11.8
	10	0.9	0.2	43.7	133.2	9.4	9.7	5.6	4.5	13.0	8.5	8.7	4.7	5.7	5.5	11.9	12.7	11.9	11.2
	20	1.8	0.3	44.4	128.5	11.9	12.9	5.6	5.2	12.5	8.6	9.3	5.2	4.3	6.1	12.7	12.7	13.4	13.6
	50	6.8	0.9	54.3	131.3	18.9	19.8	6.0	6.2	13.7	9.1	11.1	5.4	4.0	6.9	15.2	12.3	13.4	12.1
	100	16.0	1.4	51.5	135.7	30.8	30.2	5.3	7.1	13.4	10.8	11.7	5.3	5.4	6.0	14.4	13.2	15.7	12.1
	200	nan	nan	66.4	154.8	52.9	72.8	5.8	7.5	15.3	12.4	13.1	7.6	5.8	7.0	16.3	12.8	14.0	11.5
	400	nan	nan	78.0	169.3	103.3	97.5	6.1	10.6	18.4	9.8	12.5	9.0	6.5	10.6	15.9	14.1	14.9	12.8

[H]

TABLE VIII: Multivariate with Time: Training time

n ahead	sequence length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet	LinearNet	CLinear	NLinear	DLinear	TSMixer	TSMixer	TSMixer	TSMixer	
		Regression	Ridge	Encoder	Shared	Last	Full	1	2								3	no	no
1	3	0.0	0.0	7.3	7.9	8.2	8.0	4.8	4.2	5.4	7.8	7.5	4.5	6.8	6.1	16.1	14.4	16.0	16.5
	10	0.3	0.2	9.6	8.7	11.2	11.8	4.2	6.0	5.2	7.8	8.4	7.6	5.7	9.2	15.1	14.0	15.3	14.0
	20	1.1	0.2	10.0	10.4	18.2	15.8	5.5	5.9	6.0	8.2	8.9	12.3	5.7	11.8	18.4	16.3	18.4	11.5
	50	4.5	0.4	13.0	16.6	29.8	29.5	5.1	6.0	6.1	11.3	10.1	12.1	7.9	8.1	20.3	15.5	21.8	17.0
	100	14.9	1.1	21.4	25.5	51.4	48.1	6.6	6.9	5.6	10.0	11.1	5.7	10.2	9.3	23.1	16.1	17.1	15.8
	200	nan	nan	31.3	41.5	85.2	111.0	8.1	8.4	6.0	12.5	12.7	11.4	9.4	9.3	18.5	12.7	17.6	14.3
	400	nan	nan	59.1	58.9	139.2	166.8	7.3	10.7	7.9	15.4	19.4	23.8	15.3	25.7	24.6	20.9	25.5	22.2
6	3	0.0	0.0	8.4	9.2	9.3	9.7	6.8	5.4	5.8	7.8	7.1	8.5	6.6	8.7	19.0	16.2	16.6	15.5
	10	0.3	0.2	10.5	10.1	11.8	14.5	6.4	5.8	6.4	9.4	7.5	8.1	5.6	7.0	19.3	13.6	14.7	13.8
	20	1.6	0.2	10.1	14.2	16.9	19.3	7.1	6.0	6.2	8.7	9.0	11.0	8.7	6.9	16.3	13.2	15.2	13.8
	50	4.7	0.3	14.7	15.5	35.6	38.7	7.5	6.3	6.4	12.1	11.8	9.8	10.6	11.6	21.1	16.5	20.3	16.6
	100	14.6	1.2	18.7	25.3	56.3	53.9	6.6	7.4	7.3	12.0	13.5	14.4	10.8	10.2	18.8	18.3	18.0	15.1
	200	nan	nan	33.4	39.9	114.3	100.1	8.1	7.7	6.9	11.8	16.2	12.3	10.3	15.4	18.8	16.0	19.7	15.5
	400	nan	nan	59.1	67.9	230.1	173.6	8.2	11.5	8.3	17.5	20.7	21.1	17.0	23.1	23.6	21.9	23.1	19.3
15	3	0.0	0.0	11.8	13.7	11.9	10.0	6.9	7.4	8.4	9.4	7.1	7.6	6.8	7.1	13.9	14.5	18.3	15.3
	10	0.3	0.2	12.4	13.3	13.2	13.1	7.7	7.2	8.6	9.0	8.0	8.8	6.8	10.7	17.2	16.8	15.2	14.7
	20	1.6	0.2	13.7	17.1	21.4	17.7	8.0	7.8	8.5	9.3	9.5	7.5	8.2	7.2	17.8	14.4	15.1	13.6
	50	5.2	0.6	14.5	21.9	35.5	37.8	8.2	10.9	9.3	10.9	11.3	10.7	9.1	8.9	23.3	22.8	18.4	19.5
	100	13.8	1.1	20.8	28.0	76.0	58.1	8.6	8.9	9.2	12.6	13.8	9.3	9.7	12.7	20.3	17.1	18.9	15.0
	200	nan	nan	28.5	40.5	131.9	63.0	8.6	10.5	9.6	14.0	16.0	13.4	9.2	15.6	17.6	15.4	17.4	15.9
	400	nan	nan	57.8	61.7	246.2	141.7	10.4	12.0	10.3	13.6	21.6	24.8	20.8	32.6	24.6	21.0	22.3	18.6
30	3	0.0	0.1	15.6	22.1	8.7	11.6	6.4	5.6	6.5	9.4	7.5	6.3	6.4	7.7	13.2	14.3	17.3	14.7
	10	0.4	0.1	13.6	26.5	13.2	16.4	6.7	5.5	5.7	9.6	7.8	7.0	6.9	6.7	18.1	19.7	16.1	15.1
	20	1.5	0.3	13.5	25.5	20.4	18.4	7.5	5.8	5.8	11.1	9.3	9.5	7.4	12.3	17.5	17.8	19.3	13.5
	50	6.1	0.4	18.3	29.4	44.6	43.1	8.1	8.0	5.9	10.6	11.3	12.0	9.9	12.9	21.6	16.9	17.8	14.1
	100	15.7	1.2	21.6	35.2	49.5	35.8	6.3	8.5	6.3	11.4	12.9	9.4	9.5	11.6	21.1	16.3	16.9	16.9
	200	nan	nan	38.4	63.2	114.3	77.5	7.3	7.1	6.6	11.2	15.1	13.7	10.1	17.3	17.2	17.8	17.8	15.4
	400	nan	nan	69.1	85.4	214.2	136.6	13.0	13.0	8.6	14.7	16.2	16.1	9.5	18.3	22.1	19.3	17.7	18.3
90	3	0.1	0.1	16.2	43.1	12.9	11.2	7.1	6.7	8.6	7.8	8.5	5.4	8.0	6.4	12.7	12.1	16.8	15.9
	10	0.5	0.1	17.1	43.5	14.2	13.6	7.4	7.1	8.2	8.0	9.7	5.5	8.3	6.2	14.0	15.0	15.0	16.6
	20	1.6	0.2	18.1	43.6	17.8	18.3	7.3	9.1	8.2	8.3	10.3	9.2	5.6	9.1	14.8	15.8	16.2	12.1
	50	5.4	0.6	22.6	46.8	26.5	42.8	8.6	11.0	8.7	9.2	11.6	11.2	6.8	10.1	20.1	18.1	20.3	14.4
	100	17.3	1.3	25.7	49.7	37.4	35.7	8.6	9.4	9.6	10.2	13.5	9.0	9.8	10.9	25.4	18.5	19.9	17.7
	200	nan	nan	35.6	60.3	95.0	56.0	9.1	9.1	9.7	11.0	16.7	16.8	8.6	16.6	17.4	17.4	19.2	14.5
	400	nan	nan	54.9	85.5	197.6	125.1	8.1	9.8	12.4	11.6	20.2	17.8	15.1	15.6	18.6	19.4	19.5	17.0
180	3	0.2	0.2	27.2	75.4	8.3	10.9	5.2	7.2	7.2	8.5	9.4	4.9	5.9	5.5	14.1	13.6	15.6	15.3
	10	0.7	0.2	25.5	76.2	10.1	11.0	6.8	7.4	7.1	7.7	9.0	5.9	6.8	6.0	13.3	15.9	13.9	12.0
	20	1.9	0.2	28.2	79.4	15.2	13.8	6.4	7.8	7.4	8.4	10.4	5.6	6.3	6.5	14.1	14.7	19.2	11.6
	50	6.3	0.6	31.5	83.3	21.7	21.2	8.2	8.6	7.7	9.0	12.0	7.1	6.8	8.7	19.1	17.7	17.4	18.6
	100	18.2	1.3	34.4	88.5	31.6	33.2	7.3	9.1	8.7	9.8	13.7	8.6	10.1	10.2	20.2	19.6	21.1	16.8
	200	nan	nan	54.3	99.4	54.1	54.0	7.2	10.7	9.3	9.8	13.2	6.9	9.0	10.4	16.8	13.5	17.6	15.7
	400	nan	nan	62.9	122.4	118.9	102.0	7.6	13.2	12.4	11.9	16.7	12.6	8.0	17.4	17.9	17.6	18.7	15.1
360	3	0.2	0.2	43.2	160.0	8.2	8.8	5.2	5.0	12.6	8.9	11.4	5.6	6.3	5.4	15.3	12.0	16.6	13.4
	10	0.9	0.3	48.3	150.4	12.6	11.7	6.4	5.7	13.2	8.7	10.7	5.7	6.6	5.7	12.4	15.0	17.2	13.7
	20	2.5	0.3	45.0	153.7	14.3	16.9	5.9	5.2	13.5	8.7	10.4	5.3	6.2	7.0	13.4	14.9	21.4	15.6
	50	6.5	0.6	54.6	157.0	19.4	22.3	7.0	5.9	14.3	9.7	11.8	7.0	6.4	8.4	18.9	17.1	18.0	13.1
	100	19.9	1.6	55.3	165.4	31.6	32.7	9.5	8.1	14.4	10.4	15.7	10.1	10.2	11.3	23.0	17.4	19.9	20.8
	200	nan	nan	61.5	188.1	62.1	65.0	6.9	10.4	16.3	11.8	13.5	8.4	8.2	7.6	21.3	17.4	16.1	13.9
	400	nan	nan	79.2	195.9	100.6	99.4	7.7	20.8	21.2	9.8	14.6	9.7	7.4	12.4	17.8	17.4	17.5	14.4

C. Number of parameters

[H]

TABLE IX: Univariate without Time: Number of Parameters

n	sequence ahead	length	Linear		LSTM		GRU Last	GRU Full	MLP 1 Layer	MLP 2 Layer	MLP 3 Layer ReLU	LinearNet			TSMixer 1	TSMixer 2	TSMixer no relu 1	TSMixer no relu 2	
			Regression	Ridge	Encoder Decoder	Shared Encoder Decoder						LinearNet no conv	CLinear	NLinear					DLinear
1	3	4	4	30651	14	1061	1081	4	101	521	62	18	6	6	10	42	76	76	76
	10	11	11	30651	14	1061	1151	11	241	661	104	32	13	13	24	161	307	307	307
	20	21	21	30651	14	1061	1251	21	441	861	164	52	23	23	44	501	977	977	977
	50	51	51	30651	14	1061	1551	51	1041	1461	344	112	53	53	104	2721	5387	5387	5387
	100	101	101	30651	14	1061	2051	101	2041	2461	644	212	103	103	204	10421	20737	20737	20737
	200	nan	nan	30651	14	1061	3051	201	4041	4461	1244	412	203	203	404	40821	81437	81437	81437
	400	nan	nan	30651	14	1061	5051	401	8041	8461	2444	812	403	403	804	161621	322837	322837	322837
6	3	24	24	30651	14	1116	1236	24	1206	15726	157	93	26	26	50	62	96	96	96
	10	66	66	30651	14	1116	1656	66	2046	16566	234	142	68	68	134	216	362	362	362
	20	126	126	30651	14	1116	2256	126	3246	17766	344	212	128	128	254	606	1082	1082	1082
	50	306	306	30651	14	1116	4056	306	6846	21366	674	422	308	308	614	2976	5642	5642	5642
	100	606	606	30651	14	1116	7056	606	12846	27366	1224	772	608	608	1214	10926	21242	21242	21242
	200	nan	nan	30651	14	1116	13056	1206	24846	39366	2324	1472	1208	1208	2414	41826	82442	82442	82442
	400	nan	nan	30651	14	1116	25056	2406	48846	63366	4524	2872	2408	2408	4814	163626	324842	324842	324842
15	3	60	60	30651	14	1215	1515	60	5715	96015	454	354	62	62	122	98	132	132	132
	10	165	165	30651	14	1215	2565	165	7815	98115	594	466	167	167	332	315	461	461	461
	20	315	315	30651	14	1215	4065	315	10815	101115	794	626	317	317	632	795	1271	1271	1271
	50	765	765	30651	14	1215	8565	765	19815	110115	1394	1106	767	767	1532	3435	6101	6101	6101
	100	1515	1515	30651	14	1215	16065	1515	34815	125115	2394	1906	1517	1517	3032	11835	22151	22151	22151
	200	nan	nan	30651	14	1215	31065	3015	64815	155115	4394	3506	3017	3017	6032	43635	84251	84251	84251
	400	nan	nan	30651	14	1215	61065	6015	124815	215115	8394	6706	6017	6017	12032	167235	328451	328451	328451
30	3	120	120	30651	14	1380	1980	120	20430	381030	1309	1149	122	122	242	158	192	192	192
	10	330	330	30651	14	1380	4080	330	24630	385230	1554	1366	332	332	662	480	626	626	626
	20	630	630	30651	14	1380	7080	630	30630	391230	1904	1676	632	632	1262	1110	1586	1586	1586
	50	1530	1530	30651	14	1380	16080	1530	48630	409230	2954	2606	1532	1532	3062	4200	6866	6866	6866
	100	3030	3030	30651	14	1380	31080	3030	78630	439230	4704	4156	3032	3032	6062	13350	23666	23666	23666
	200	nan	nan	30651	14	1380	61080	6030	138630	499230	8204	7256	6032	6032	12062	46650	87266	87266	87266
	400	nan	nan	30651	14	1380	121080	12030	258630	619230	15204	13456	12032	12032	24062	173250	334466	334466	334466
90	3	360	360	30651	14	2040	3840	360	169290	3411090	9229	8829	362	362	722	398	432	432	432
	10	990	990	30651	14	2040	10140	990	181890	3423690	9894	9466	992	992	1982	1140	1286	1286	1286
	20	1890	1890	30651	14	2040	19140	1890	199890	3441690	10844	10376	1892	1892	3782	2370	2846	2846	2846
	50	4590	4590	30651	14	2040	46140	4590	253890	3495690	13694	13106	4592	4592	9182	7260	9926	9926	9926
	100	9090	9090	30651	14	2040	91140	9090	343890	3585690	18444	17656	9092	9092	18182	19410	29726	29726	29726
	200	nan	nan	30651	14	2040	181140	18090	523890	3765690	27944	26756	18092	18092	36182	58710	99326	99326	99326
	400	nan	nan	30651	14	2040	361140	36090	883890	4125690	46944	44956	36092	36092	72182	197310	358526	358526	358526
180	3	720	720	30651	14	3030	6630	720	662580	13626180	34609	33849	722	722	1442	758	792	792	792
	10	1980	1980	30651	14	3030	19230	1980	687780	13651380	35904	35116	1982	1982	3962	2130	2276	2276	2276
	20	3780	3780	30651	14	3030	37230	3780	723780	13687380	37754	36926	3782	3782	7562	4260	4736	4736	4736
	50	9180	9180	30651	14	3030	91230	9180	831780	13795380	43304	42356	9182	9182	18362	11850	14516	14516	14516
	100	18180	18180	30651	14	3030	181230	18180	1011780	13975380	52554	51406	18182	18182	36362	28500	38816	38816	38816
	200	nan	nan	30651	14	3030	361230	36180	1371780	14335380	71054	69506	36182	36182	72362	76800	117416	117416	117416
	400	nan	nan	30651	14	3030	721230	72180	2091780	15055380	108054	105706	72182	72182	144362	233400	394616	394616	394616
360	3	1440	1440	30651	14	5010	12210	1440	2621160	54468360	133969	132489	1442	1442	2882	1478	1512	1512	1512
	10	3960	3960	30651	14	5010	37410	3960	2671560	54518760	136524	135016	3962	3962	7922	4110	4256	4256	4256
	20	7560	7560	30651	14	5010	73410	7560	2743560	54590760	140174	138626	7562	7562	15122	8040	8516	8516	8516
	50	18360	18360	30651	14	5010	181410	18360	2959560	54806760	151124	149456	18362	18362	36722	21030	23696	23696	23696
	100	36360	36360	30651	14	5010	361410	36360	3319560	55166760	169374	167506	36362	36362	72722	46680	56996	56996	56996
	200	nan	nan	30651	14	5010	721410	72360	4039560	55886760	205874	203606	72362	72362	144722	112980	153596	153596	153596
	400	nan	nan	30651	14	5010	1441410	144360	5479560	57326760	278874	275806	144362	144362	288722	305580	466796	466796	466796

TABLE X: Univariate with Time: Number of Parameters

n	sequence ahead length	Linear Regression	Ridge	LSTM		GRU Last Sequence	GRU Full Sequence	MLP	MLP	MLP	LinearNet				TSMixer 1	TSMixer 2	TSMixer no relu 1	TSMixer no relu 2	
				Encoder	Shared Encoder			Layer 1 linear	Layer 2 linear	Layer 3 ReLU	LinearNet no conv	CLinear	NLinear	DLinear					
1	3	10	10	31051	88	1121	1141	10	221	641	160	38	8	8	12	82	150	150	150
	10	31	31	31051	88	1121	1211	31	641	1061	202	52	15	15	26	229	437	437	437
	20	61	61	31051	88	1121	1311	61	1241	1661	262	72	25	25	46	609	1187	1187	1187
	50	151	151	31051	88	1121	1611	151	3041	3461	442	132	55	55	106	2949	5837	5837	5837
	100	301	301	31051	88	1121	2111	301	6041	6461	742	232	105	105	206	10849	21587	21587	21587
	200	nan	nan	31051	88	1121	3111	601	12041	12461	1342	432	205	205	406	41649	83087	83087	83087
	400	nan	nan	31051	88	1121	5111	1201	24041	24461	2542	832	405	405	806	163249	326087	326087	326087
6	3	60	60	31051	88	1176	1296	60	1926	16446	255	113	28	28	52	102	170	170	170
	10	186	186	31051	88	1176	1716	186	4446	18966	332	162	70	70	136	284	492	492	492
	20	366	366	31051	88	1176	2316	366	8046	22566	442	232	130	130	256	714	1292	1292	1292
	50	906	906	31051	88	1176	4116	906	18846	33366	772	442	310	310	616	3204	6092	6092	6092
	100	1806	1806	31051	88	1176	7116	1806	36846	51366	1322	792	610	610	1216	11354	22092	22092	22092
	200	nan	nan	31051	88	1176	13116	3606	72846	87366	2422	1492	1210	1210	2416	42654	84092	84092	84092
	400	nan	nan	31051	88	1176	25116	7206	144846	159366	4622	2892	2410	2410	4816	165254	328092	328092	328092
15	3	150	150	31051	88	1275	1575	150	7515	97815	552	374	64	64	124	138	206	206	206
	10	465	465	31051	88	1275	2625	465	13815	104115	692	486	169	169	334	383	591	591	591
	20	915	915	31051	88	1275	4125	915	22815	113115	892	646	319	319	634	903	1481	1481	1481
	50	2265	2265	31051	88	1275	8625	2265	49815	140115	1492	1126	769	769	1534	3663	6551	6551	6551
	100	4515	4515	31051	88	1275	16125	4515	94815	185115	2492	1926	1519	1519	3034	12263	23001	23001	23001
	200	nan	nan	31051	88	1275	31125	9015	184815	275115	4492	3526	3019	3019	6034	44463	85901	85901	85901
	400	nan	nan	31051	88	1275	61125	18015	364815	455115	8492	6726	6019	6019	12034	168863	331701	331701	331701
30	3	300	300	31051	88	1440	2040	300	24030	384630	1407	1169	124	124	244	198	266	266	266
	10	930	930	31051	88	1440	4140	930	36630	397230	1652	1386	334	334	664	548	756	756	756
	20	1830	1830	31051	88	1440	7140	1830	54630	415230	2002	1696	634	634	1264	1218	1796	1796	1796
	50	4530	4530	31051	88	1440	16140	4530	108630	469230	3052	2626	1534	1534	3064	4428	7316	7316	7316
	100	9030	9030	31051	88	1440	31140	9030	198630	559230	4802	4176	3034	3034	6064	13778	24516	24516	24516
	200	nan	nan	31051	88	1440	61140	18030	378630	739230	8302	7276	6034	6034	12064	47478	88916	88916	88916
	400	nan	nan	31051	88	1440	121140	36030	738630	1099230	15302	13476	12034	12034	24064	174878	337716	337716	337716
90	3	900	900	31051	88	2100	3900	900	180090	3421890	9327	8849	364	364	724	438	506	506	506
	10	2790	2790	31051	88	2100	10200	2790	217890	3459690	9992	9486	994	994	1984	1208	1416	1416	1416
	20	5490	5490	31051	88	2100	19200	5490	271890	3513690	10942	10396	1894	1894	3784	2478	3056	3056	3056
	50	13590	13590	31051	88	2100	46200	13590	433890	3675690	13792	13126	4594	4594	9184	7488	10376	10376	10376
	100	27090	27090	31051	88	2100	91200	27090	703890	3945690	18542	17676	9094	9094	18184	19838	30576	30576	30576
	200	nan	nan	31051	88	2100	181200	54090	1243890	4485690	28042	26776	18094	18094	36184	59538	100976	100976	100976
	400	nan	nan	31051	88	2100	361200	108090	2323890	5565690	47042	44976	36094	36094	72184	198938	361776	361776	361776
180	3	1800	1800	31051	88	3090	6690	1800	684180	13647780	34707	33869	724	724	1444	798	866	866	866
	10	5580	5580	31051	88	3090	19290	5580	759780	13723380	36002	35136	1984	1984	3964	2198	2406	2406	2406
	20	10980	10980	31051	88	3090	37290	10980	867780	13831380	37852	36946	3784	3784	7564	4368	4946	4946	4946
	50	27180	27180	31051	88	3090	91290	27180	1191780	14155380	43402	42376	9184	9184	18364	12078	14966	14966	14966
	100	54180	54180	31051	88	3090	181290	54180	1731780	14695380	52652	51426	18184	18184	36364	28928	39666	39666	39666
	200	nan	nan	31051	88	3090	361290	108180	2811780	15775380	71152	69526	36184	36184	72364	77628	119066	119066	119066
	400	nan	nan	31051	88	3090	721290	216180	4971780	17935380	108152	105726	72184	72184	144364	235028	397866	397866	397866
360	3	3600	3600	31051	88	5070	12270	3600	2664360	54511560	134067	132509	1444	1444	2884	1518	1586	1586	1586
	10	11160	11160	31051	88	5070	37470	11160	2815560	54662760	136622	135036	3964	3964	7924	4178	4386	4386	4386
	20	21960	21960	31051	88	5070	73470	21960	3031560	54878760	140272	138646	7564	7564	15124	8148	8726	8726	8726
	50	54360	54360	31051	88	5070	181470	54360	3679560	55526760	151222	149476	18364	18364	36724	21258	24146	24146	24146
	100	108360	108360	31051	88	5070	361470	108360	4759560	56606760	169472	167526	36364	36364	72724	47108	57846	57846	57846
	200	nan	nan	31051	88	5070	721470	216360	6919560	58766760	205972	203626	72364	72364	144724	113808	155246	155246	155246
	400	nan	nan	31051	88	5070	1441470	432360	11239560	63086760	278972	275826	144364	144364	288724	307208	470046	470046	470046

TABLE XI: Multivariate without Time: Number of Parameters

n	sequence ahead length	Linear		LSTM		GRU		MLP	MLP	MLP	LinearNet				TSMixer		TSMixer	TSMixer	
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	Layer linear	Layer linear	Layer ReLU	LinearNet no conv	CLinear	NLinear	DLinear	1	2	1	2	
1	3	58	58	34251	2984	1601	1621	58	1181	1601	4616	630	24	24	28	402	742	742	742
	10	191	191	34251	2984	1601	1691	191	3841	4261	4658	644	31	31	42	773	1477	1477	1477
	20	381	381	34251	2984	1601	1791	381	7641	8061	4718	664	41	41	62	1473	2867	2867	2867
	50	951	951	34251	2984	1601	2091	951	19041	19461	4898	724	71	71	122	4773	9437	9437	9437
	100	1901	1901	34251	2984	1601	2591	1901	38041	38461	5198	824	121	121	222	14273	28387	28387	28387
	200	nan	nan	34251	2984	1601	3591	3801	76041	76461	5798	1024	221	221	422	48273	96287	96287	96287
	400	nan	nan	34251	2984	1601	5591	7601	152041	152461	6998	1424	421	421	822	176273	352087	352087	352087
6	3	348	348	34251	2984	1656	1776	348	7686	22206	4711	705	44	44	68	422	762	762	762
	10	1146	1146	34251	2984	1656	2196	1146	23646	38166	4788	754	86	86	152	828	1532	1532	1532
	20	2286	2286	34251	2984	1656	2796	2286	46446	60966	4898	824	146	146	272	1578	2972	2972	2972
	50	5706	5706	34251	2984	1656	4596	5706	114846	129366	5228	1034	326	326	632	5028	9692	9692	9692
	100	11406	11406	34251	2984	1656	7596	11406	228846	243366	5778	1384	626	626	1232	14778	28892	28892	28892
	200	nan	nan	34251	2984	1656	13596	22806	456846	471366	6878	2084	1226	1226	2432	49278	97292	97292	97292
	400	nan	nan	34251	2984	1656	25596	45606	912846	927366	9078	3484	2426	2426	4832	178278	354092	354092	354092
15	3	870	870	34251	2984	1755	2055	870	21915	112215	5008	966	80	80	140	458	798	798	798
	10	2865	2865	34251	2984	1755	3105	2865	61815	152115	5148	1078	185	185	350	927	1631	1631	1631
	20	5715	5715	34251	2984	1755	4605	5715	118815	209115	5348	1238	335	335	650	1767	3161	3161	3161
	50	14265	14265	34251	2984	1755	9105	14265	289815	380115	5948	1718	785	785	1550	5487	10151	10151	10151
	100	28515	28515	34251	2984	1755	16605	28515	574815	665115	6948	2518	1535	1535	3050	15687	29801	29801	29801
	200	nan	nan	34251	2984	1755	31605	57015	1144815	1235115	8948	4118	3035	3035	6050	51087	99101	99101	99101
	400	nan	nan	34251	2984	1755	61605	114015	2284815	2375115	12948	7318	6035	6035	12050	181887	357701	357701	357701
30	3	1740	1740	34251	2984	1920	2520	1740	52830	413430	5863	1761	140	140	260	518	858	858	858
	10	5730	5730	34251	2984	1920	4620	5730	132630	493230	6108	1978	350	350	680	1092	1796	1796	1796
	20	11430	11430	34251	2984	1920	7620	11430	246630	607230	6458	2288	650	650	1280	2082	3476	3476	3476
	50	28530	28530	34251	2984	1920	16620	28530	588630	949230	7508	3218	1550	1550	3080	6252	10916	10916	10916
	100	57030	57030	34251	2984	1920	31620	57030	1158630	1519230	9258	4768	3050	3050	6080	17202	31316	31316	31316
	200	nan	nan	34251	2984	1920	61620	114030	2298630	2659230	12758	7868	6050	6050	12080	54102	102116	102116	102116
	400	nan	nan	34251	2984	1920	121620	228030	4578630	4939230	19758	14068	12050	12050	24080	187902	363716	363716	363716
90	3	5220	5220	34251	2984	2580	4380	5220	266490	3508290	13783	9441	380	380	740	758	1098	1098	1098
	10	17190	17190	34251	2984	2580	10680	17190	505890	3747690	14448	10078	1010	1010	2000	1752	2456	2456	2456
	20	34290	34290	34251	2984	2580	19680	34290	847890	4089690	15398	10988	1910	1910	3800	3342	4736	4736	4736
	50	85590	85590	34251	2984	2580	46680	85590	1873890	5115690	18248	13718	4610	4610	9200	9312	13976	13976	13976
	100	171090	171090	34251	2984	2580	91680	171090	3583890	6825690	22998	18268	9110	9110	18200	23262	37376	37376	37376
	200	nan	nan	34251	2984	2580	181680	342090	7003890	10245690	32498	27368	18110	18110	36200	66162	114176	114176	114176
	400	nan	nan	34251	2984	2580	361680	684090	13843890	17085690	51498	45568	36110	36110	72200	211962	387776	387776	387776
180	3	10440	10440	34251	2984	3570	7170	10440	856980	13820580	39163	34461	740	740	1460	1118	1458	1458	1458
	10	34380	34380	34251	2984	3570	19770	34380	1335780	14299380	40458	35728	2000	2000	3980	2742	3446	3446	3446
	20	68580	68580	34251	2984	3570	37770	68580	2019780	14983380	42308	37538	3800	3800	7580	5232	6626	6626	6626
	50	171180	171180	34251	2984	3570	91770	171180	4071780	17035380	47858	42968	9200	9200	18380	13902	18566	18566	18566
	100	342180	342180	34251	2984	3570	181770	342180	7491780	20455380	57108	52018	18200	18200	36380	32352	46466	46466	46466
	200	nan	nan	34251	2984	3570	361770	684180	14331780	27295380	75608	70118	36200	36200	72380	84252	132266	132266	132266
	400	nan	nan	34251	2984	3570	721770	1368180	28011780	40975380	112608	106318	72200	72200	144380	248052	423866	423866	423866
360	3	20880	20880	34251	2984	5550	12750	20880	3009960	54857160	138523	133101	1460	1460	2900	1838	2178	2178	2178
	10	68760	68760	34251	2984	5550	37950	68760	3967560	55814760	141078	135628	3980	3980	7940	4722	5426	5426	5426
	20	137160	137160	34251	2984	5550	73950	137160	5335560	57182760	144728	139238	7580	7580	15140	9012	10406	10406	10406
	50	342360	342360	34251	2984	5550	181950	342360	9439560	61286760	155678	150068	18380	18380	36740	23082	27746	27746	27746
	100	684360	684360	34251	2984	5550	361950	684360	16279560	68126760	173928	168118	36380	36380	72740	50532	64646	64646	64646
	200	nan	nan	34251	2984	5550	721950	1368360	29959560	81806760	210428	204218	72380	72380	144740	120432	168446	168446	168446
	400	nan	nan	34251	2984	5550	1441950	2736360	57319560	109166760	283428	276418	144380	144380	288740	320232	496046	496046	496046

TABLE XII: Multivariate with Time: Number of Parameters

n	sequence ahead length	Linear Regression	Ridge	LSTM		GRU Last Sequence	GRU Full Sequence	MLP	MLP	MLP	LinearNet	LinearNet no conv	CLinear	NLinear	DLinear	TSMixer 1	TSMixer 2	TSMixer no relu 1	TSMixer no relu 2
				Encoder	Shared Encoder Decoder			Layer 1 linear	Layer 2 linear	Layer 3 ReLU									
1	3	64	64	34651	3634	1661	1681	64	1301	1721	5632	758	26	26	30	442	816	816	816
	10	211	211	34651	3634	1661	1751	211	4241	4661	5674	772	33	33	44	841	1607	1607	1607
	20	421	421	34651	3634	1661	1851	421	8441	8861	5734	792	43	43	64	1581	3077	3077	3077
	50	1051	1051	34651	3634	1661	2151	1051	21041	21461	5914	852	73	73	124	5001	9887	9887	9887
	100	2101	2101	34651	3634	1661	2651	2101	42041	42461	6214	952	123	123	224	14701	29237	29237	29237
	200	nan	nan	34651	3634	1661	3651	4201	84041	84461	6814	1152	223	223	424	49101	97937	97937	97937
	400	nan	nan	34651	3634	1661	5651	8401	168041	168461	8014	1552	423	423	824	177901	355337	355337	355337
6	3	384	384	34651	3634	1716	1836	384	8406	22926	5727	833	46	46	70	462	836	836	836
	10	1266	1266	34651	3634	1716	2256	1266	26046	40566	5804	882	88	88	154	896	1662	1662	1662
	20	2526	2526	34651	3634	1716	2856	2526	51246	65766	5914	952	148	148	274	1686	3182	3182	3182
	50	6306	6306	34651	3634	1716	4656	6306	126846	141366	6244	1162	328	328	634	5256	10142	10142	10142
	100	12606	12606	34651	3634	1716	7656	12606	252846	267366	6794	1512	628	628	1234	15206	29742	29742	29742
	200	nan	nan	34651	3634	1716	13656	25206	504846	519366	7894	2212	1228	1228	2434	50106	98942	98942	98942
	400	nan	nan	34651	3634	1716	25656	50406	1008846	1023366	10094	3612	2428	2428	4834	179906	357342	357342	357342
15	3	960	960	34651	3634	1815	2115	960	23715	114015	6024	1094	82	82	142	498	872	872	872
	10	3165	3165	34651	3634	1815	3165	3165	67815	158115	6164	1206	187	187	352	995	1761	1761	1761
	20	6315	6315	34651	3634	1815	4665	6315	130815	221115	6364	1366	337	337	652	1875	3371	3371	3371
	50	15765	15765	34651	3634	1815	9165	15765	319815	410115	6964	1846	787	787	1552	5715	10601	10601	10601
	100	31515	31515	34651	3634	1815	16665	31515	634815	725115	7964	2646	1537	1537	3052	16115	30651	30651	30651
	200	nan	nan	34651	3634	1815	31665	63015	1264815	1355115	9964	4246	3037	3037	6052	51915	100751	100751	100751
	400	nan	nan	34651	3634	1815	61665	126015	2524815	2615115	13964	7446	6037	6037	12052	183515	360951	360951	360951
30	3	1920	1920	34651	3634	1980	2580	1920	56430	417030	6879	1889	142	142	262	558	932	932	932
	10	6330	6330	34651	3634	1980	4680	6330	144630	505230	7124	2106	352	352	682	1160	1926	1926	1926
	20	12630	12630	34651	3634	1980	7680	12630	270630	631230	7474	2416	652	652	1282	2190	3686	3686	3686
	50	31530	31530	34651	3634	1980	16680	31530	648630	1009230	8524	3346	1552	1552	3082	6480	11366	11366	11366
	100	63030	63030	34651	3634	1980	31680	63030	1278630	1639230	10274	4896	3052	3052	6082	17630	32166	32166	32166
	200	nan	nan	34651	3634	1980	61680	126030	2538630	2899230	13774	7996	6052	6052	12082	54930	103766	103766	103766
	400	nan	nan	34651	3634	1980	121680	252030	5058630	5419230	20774	14196	12052	12052	24082	189530	366966	366966	366966
90	3	5760	5760	34651	3634	2640	4440	5760	277290	3519090	14799	9569	382	382	742	798	1172	1172	1172
	10	18990	18990	34651	3634	2640	10740	18990	541890	3783690	15464	10206	1012	1012	2002	1820	2586	2586	2586
	20	37890	37890	34651	3634	2640	19740	37890	919890	4161690	16414	11116	1912	1912	3802	3450	4946	4946	4946
	50	94590	94590	34651	3634	2640	46740	94590	2053890	5295690	19264	13846	4612	4612	9202	9540	14426	14426	14426
	100	189090	189090	34651	3634	2640	91740	189090	3943890	7185690	24014	18396	9112	9112	18202	23690	38226	38226	38226
	200	nan	nan	34651	3634	2640	181740	378090	7723890	10965690	33514	27496	18112	18112	36202	66990	115826	115826	115826
	400	nan	nan	34651	3634	2640	361740	756090	15283890	18525690	52514	45696	36112	36112	72202	213590	391026	391026	391026
180	3	11520	11520	34651	3634	3630	7230	11520	878580	13842180	40179	34589	742	742	1462	1158	1532	1532	1532
	10	37980	37980	34651	3634	3630	19830	37980	1407780	14371380	41474	35856	2002	2002	3982	2810	3576	3576	3576
	20	75780	75780	34651	3634	3630	37830	75780	2163780	15127380	43324	37666	3802	3802	7582	5340	6836	6836	6836
	50	189180	189180	34651	3634	3630	91830	189180	4431780	17395380	48874	43096	9202	9202	18382	14130	19016	19016	19016
	100	378180	378180	34651	3634	3630	181830	378180	8211780	21175380	58124	52146	18202	18202	36382	32780	47316	47316	47316
	200	nan	nan	34651	3634	3630	361830	756180	15771780	28735380	76624	70246	36202	36202	72382	85080	133916	133916	133916
	400	nan	nan	34651	3634	3630	721830	1512180	30891780	43855380	113624	106446	72202	72202	144382	249680	427116	427116	427116
360	3	23040	23040	34651	3634	5610	12810	23040	3053160	54900360	139539	133229	1462	1462	2902	1878	2252	2252	2252
	10	75960	75960	34651	3634	5610	38010	75960	4111560	55958760	142094	135756	3982	3982	7942	4790	5556	5556	5556
	20	151560	151560	34651	3634	5610	74010	151560	5623560	57470760	145744	139366	7582	7582	15142	9120	10616	10616	10616
	50	378360	378360	34651	3634	5610	182010	378360	10159560	62006760	156694	150196	18382	18382	36742	23310	28196	28196	28196
	100	756360	756360	34651	3634	5610	362010	756360	17719560	69566760	174944	168246	36382	36382	72742	50960	65496	65496	65496
	200	nan	nan	34651	3634	5610	722010	1512360	32839560	84686760	211444	204346	72382	72382	144742	121260	170096	170096	170096
	400	nan	nan	34651	3634	5610	1442010	3024360	63079560	114926760	284444	276546	144382	144382	288742	321860	499296	499296	499296

APPENDIX B
ETTH1 PREDICTION TASKS RESULTS

A. MSE

TABLE XIII: Univariate without Time: Mean Squared Error (x1e2)

n ahead	sequence length	Linear		LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet no conv	CLinear	NLinear	DLinear	TSMixer 1	TSMixer 2	TSMixer	TSMixer
		Regression	Ridge	Encoder Decoder	Shared Encoder Decoder	Last Sequence	Full Sequence	1 Layer linear	2 Layer linear	3 Layer ReLU							no relu 1	no relu 2
1	3	0.87	0.87	0.91	3.40	0.93	0.90	0.88	0.88	0.88	0.88	0.89	0.87	0.87	0.87	0.86	0.93	0.87
	10	0.86	0.86	0.88	1.73	0.94	0.89	0.89	0.87	0.89	0.87	0.89	0.87	0.88	0.85	0.86	0.86	0.86
	20	0.84	0.84	0.86	1.56	0.88	0.86	0.85	0.85	0.87	0.85	0.86	0.84	0.86	93.89	0.84	131.75	0.85
	50	0.81	0.81	0.87	1.40	0.91	0.85	0.82	0.82	0.88	0.84	0.83	0.84	0.81	0.83	92.80	106.00	0.80
	100	0.79	0.79	0.85	3.96	0.88	0.82	0.83	0.81	0.88	0.81	0.82	0.82	135.52	0.79	121.46	128.87	0.80
	200	nan	nan	0.86	1.78	0.93	0.79	0.82	0.82	1.00	0.80	0.82	0.81	130.67	0.79	175.63	1.16	0.84
	400	nan	nan	0.83	1.48	0.92	0.80	0.87	0.85	1.19	0.81	0.80	0.84	123.58	0.79	12.43	19.54	0.79
6	3	2.63	2.63	2.75	5.86	3.02	2.76	2.67	2.68	2.85	2.74	2.71	2.69	2.63	2.62	2.73	2.62	2.86
	10	2.53	2.53	2.77	7.39	2.82	2.69	2.60	2.58	2.78	2.59	2.58	2.60	2.53	2.54	2.60	2.54	2.55
	20	2.32	2.32	2.58	7.53	2.45	2.46	2.37	2.35	2.44	2.35	2.36	2.37	2.34	2.34	2.28	111.79	128.14
	50	2.15	2.15	2.66	8.43	2.66	2.47	2.22	2.22	2.49	2.21	2.23	2.23	2.17	2.18	2.15	2.20	2.12
	100	2.04	2.04	2.29	4.67	2.45	2.28	2.15	2.09	2.65	2.12	2.11	2.14	2.07	2.08	93.80	2.24	101.00
	200	nan	nan	2.54	9.33	2.62	2.13	2.09	2.05	3.03	2.05	2.05	2.06	2.02	2.00	165.85	129.13	282.62
	400	nan	nan	2.50	4.74	2.70	2.09	2.14	2.09	3.55	2.04	2.05	2.11	125.47	2.00	9.63	6.68	3.12
15	3	4.92	4.91	6.15	9.22	5.79	5.53	5.05	5.15	5.47	5.05	5.09	5.05	5.29	4.90	5.57	4.99	5.00
	10	4.35	4.35	5.73	10.09	5.32	4.93	4.47	4.48	5.00	4.43	4.43	4.44	4.46	4.46	4.43	4.70	4.62
	20	4.04	4.04	5.20	9.92	4.90	4.84	4.18	4.17	4.29	4.11	4.11	4.20	4.07	4.08	4.08	4.33	4.04
	50	3.75	3.75	5.93	9.04	4.91	4.96	3.91	3.84	4.44	3.78	3.85	3.91	3.81	3.82	3.71	111.73	104.85
	100	3.51	3.51	5.42	72.91	4.77	4.29	3.67	3.61	3.97	3.60	3.68	3.72	3.60	3.56	3.87	124.74	3.75
	200	nan	nan	5.10	30.28	4.67	3.75	3.58	3.48	4.88	3.52	3.55	3.55	3.54	3.45	4.22	230.36	3.80
	400	nan	nan	4.51	69.70	4.51	3.57	3.55	3.54	4.49	3.53	3.52	3.57	134.25	3.44	365.10	5.85	505.74
30	3	6.60	6.60	10.37	20.55	8.50	8.35	6.75	6.89	8.66	6.78	6.72	6.80	7.30	6.57	7.38	7.50	7.81
	10	6.08	6.08	8.96	45.86	8.31	7.85	6.28	6.43	7.89	6.38	6.33	6.33	6.14	6.16	7.30	6.70	6.38
	20	5.81	5.81	9.41	138.11	7.80	8.35	6.10	5.99	8.52	5.91	5.98	6.11	5.82	5.85	5.62	5.66	5.73
	50	5.39	5.39	12.84	12.39	7.64	8.08	5.58	6.10	5.98	5.70	5.54	5.84	5.53	5.62	5.26	5.42	6.36
	100	5.08	5.08	7.84	37.29	7.19	7.45	5.41	5.24	5.89	5.23	5.26	5.30	5.16	5.14	5.48	5.82	5.37
	200	nan	nan	7.98	28.77	7.06	5.81	5.20	5.02	6.15	5.11	5.05	5.20	5.08	5.03	5.82	5.93	5.23
	400	nan	nan	11.45	14.79	6.28	5.50	5.17	5.04	6.46	5.34	5.11	5.34	5.14	4.99	19.72	10.76	6.97
90	3	11.89	11.89	25.79	25.24	17.56	18.57	12.58	12.29	20.56	12.01	12.17	12.05	12.06	11.95	13.60	13.30	12.12
	10	10.90	10.90	21.15	29.11	17.95	16.87	11.29	14.69	17.55	11.32	11.17	11.36	13.50	10.96	10.87	12.13	11.06
	20	10.34	10.34	20.47	29.27	16.62	16.73	10.82	10.67	24.61	10.54	11.03	10.85	12.40	11.11	9.71	9.80	9.98
	50	9.38	9.38	21.11	29.20	16.45	14.26	10.01	10.00	11.46	10.30	10.16	9.84	10.22	9.84	91.12	8.77	9.10
	100	8.69	8.69	20.18	137.06	15.63	14.45	9.10	9.34	15.78	8.85	8.91	9.12	8.95	8.88	8.86	11.53	85.41
	200	nan	nan	22.18	49.61	13.90	11.67	8.87	8.58	11.62	8.70	8.43	8.70	8.61	8.62	8.52	9.06	129.28
	400	nan	nan	21.20	126.35	13.28	8.34	8.88	9.02	10.96	8.47	8.78	8.92	8.86	8.64	579.99	13.07	169.39
180	3	14.99	14.99	29.39	31.26	22.97	24.27	16.93	14.93	24.84	15.22	15.30	15.48	19.09	15.11	17.35	14.88	15.05
	10	13.85	13.85	27.44	33.66	23.26	23.09	14.57	13.71	22.13	14.07	14.08	14.74	15.70	14.12	14.77	13.85	13.69
	20	13.10	13.10	28.22	84.44	22.41	22.64	13.91	14.10	25.00	13.99	13.39	13.81	14.92	13.44	13.89	103.46	12.36
	50	11.82	11.82	23.23	96.41	22.56	19.49	12.81	12.70	17.51	12.05	12.89	12.69	12.32	12.21	84.12	98.34	11.76
	100	10.90	10.90	27.08	32.86	20.22	19.20	11.47	11.38	13.80	11.18	11.79	11.35	11.39	11.67	330.02	13.45	97.18
	200	nan	nan	21.33	111.73	18.78	12.22	10.59	10.77	13.70	10.46	10.67	10.66	10.67	10.52	127.48	14.98	12.63
	400	nan	nan	21.83	25.80	16.20	11.30	11.05	11.66	12.32	10.83	10.88	10.96	11.23	10.66	15.03	494.41	12.22
360	3	18.20	18.20	31.21	34.31	26.11	27.42	21.35	18.46	34.21	19.99	18.46	18.94	18.36	18.58	20.75	21.48	20.35
	10	16.93	16.93	32.16	133.02	25.56	27.61	17.92	17.47	28.98	17.79	18.07	17.95	19.58	17.69	16.49	16.31	17.70
	20	16.09	16.09	30.18	61.65	25.96	25.48	17.05	17.55	28.61	17.08	16.19	17.06	19.12	16.88	94.27	14.58	121.18
	50	14.57	14.57	26.64	131.28	22.15	21.89	15.93	15.32	22.83	15.04	15.17	15.70	15.19	15.09	129.07	107.08	13.49
	100	13.42	13.42	27.03	34.54	23.28	19.40	14.42	15.05	15.85	14.14	14.15	14.33	14.11	14.32	12.38	12.42	12.14
	200	nan	nan	26.98	126.63	18.22	14.63	13.65	14.23	15.45	13.54	13.48	13.31	13.41	13.41	22.01	22.34	162.30
	400	nan	nan	27.53	81.75	19.69	13.74	14.21	13.75	16.23	13.32	13.83	14.67	14.22	14.18	19.59	919.92	16.13

TABLE XIV: Univariate with Time: Mean Squared Error

n	sequence ahead length	Linear		LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer		
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	1 Layer	2 Layer	3 Layer	no conv	CLinear	NLinear	DLinear	1	2	no relu	no relu		
1	3	0.87	0.87	0.86	2.13	0.93	0.90	0.87	0.87	0.88	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.87		
	10	0.86	0.86	0.89	5.47	0.97	0.94	0.88	0.86	0.87	0.90	0.87	0.88	0.86	0.86	0.86	0.86	0.87		
	20	0.84	0.84	0.87	1.83	0.88	0.85	0.85	0.85	0.85	0.85	0.86	0.85	0.84	0.84	0.85	111.00	131.07	133.02	
	50	0.81	0.81	0.87	1.40	0.89	0.85	0.83	0.83	0.88	0.83	0.84	0.84	0.82	0.81	75.17	98.08	109.19	89.21	
	100	0.79	0.79	0.83	1.93	0.87	0.87	0.83	0.80	0.87	0.82	0.82	0.82	0.82	135.43	0.80	112.55	0.92	0.79	124.02
	200	nan	nan	0.83	1.32	0.90	0.79	0.83	0.81	0.94	0.80	0.81	0.82	132.47	0.79	462.56	121.72	0.85	0.79	
400	nan	nan	0.83	3.24	0.92	0.80	0.84	0.84	1.11	0.80	0.81	0.82	121.44	0.79	973.50	5.26	2.65	228.86		
6	3	2.63	2.63	2.88	7.61	2.96	2.76	2.66	2.63	2.87	2.69	2.75	2.64	2.70	2.64	2.81	2.62	2.77	2.69	
	10	2.53	2.53	2.80	4.13	2.91	2.83	2.58	2.63	2.63	2.59	2.60	2.62	2.55	2.55	2.55	2.54	2.56	2.61	
	20	2.32	2.32	2.55	138.78	2.52	2.58	2.36	2.35	2.34	2.35	2.37	2.39	2.33	2.33	110.58	2.25	2.26	2.30	
	50	2.15	2.15	3.00	11.81	2.62	2.37	2.21	2.19	2.84	2.23	2.26	2.23	2.18	2.18	2.15	2.17	109.46	2.16	
	100	2.04	2.04	2.72	52.14	2.30	2.27	2.13	2.09	2.44	2.11	2.09	2.14	2.07	2.06	92.63	106.23	2.05	2.04	
	200	nan	nan	2.50	10.70	2.55	2.05	2.07	2.06	2.71	2.04	2.02	2.02	2.05	2.00	3.34	11.02	2.21	166.53	
400	nan	nan	2.42	6.83	2.57	2.03	2.14	2.04	3.07	2.04	2.05	2.13	2.05	2.02	10.59	303.03	231.66	2.55		
15	3	4.92	4.91	5.99	10.37	5.58	5.66	5.01	4.91	5.59	5.10	5.02	5.01	5.15	4.88	5.73	5.30	5.26	5.39	
	10	4.35	4.35	5.13	12.97	5.12	5.22	4.46	4.41	6.00	4.67	4.41	4.45	4.36	4.39	4.42	4.44	4.84	4.64	
	20	4.04	4.04	5.08	12.81	4.96	5.05	4.17	4.10	4.70	4.28	137.60	4.15	4.07	4.12	3.96	3.94	3.99	4.28	
	50	3.75	3.75	5.06	9.50	4.63	4.84	3.90	3.84	4.84	3.87	3.92	4.05	3.82	3.83	86.47	108.10	3.73	4.09	
	100	3.51	3.51	5.94	11.59	4.51	4.22	3.70	3.62	5.54	3.70	3.69	3.66	3.57	3.58	97.35	90.28	3.67	3.56	
	200	nan	nan	6.40	137.32	4.79	3.87	3.48	3.49	4.22	3.54	3.56	3.54	3.46	3.47	366.13	389.26	508.11	3.76	
400	nan	nan	6.04	17.51	4.40	3.70	3.63	3.50	4.96	3.49	3.54	3.61	3.53	3.44	772.86	635.55	3.90	180.85		
30	3	6.60	6.60	8.67	138.02	8.88	8.35	6.82	6.85	9.12	6.85	6.75	6.82	6.90	6.60	6.64	6.97	7.63	6.66	
	10	6.08	6.08	9.23	138.03	8.06	8.27	6.24	6.58	8.83	6.33	6.30	6.41	6.72	6.18	6.48	6.80	6.53	6.56	
	20	5.81	5.81	9.54	19.03	7.89	8.90	6.20	6.62	8.27	6.02	5.90	6.02	6.15	5.94	5.82	99.96	127.98	102.47	
	50	5.39	5.39	10.47	137.07	7.52	7.18	5.65	5.64	7.47	5.73	5.54	5.67	5.48	5.56	94.11	104.69	5.36	5.31	
	100	5.08	5.08	7.84	20.44	7.64	6.94	5.34	5.62	7.27	5.37	5.28	5.51	5.21	5.21	93.07	5.70	5.23	105.88	
	200	nan	nan	6.22	13.89	6.94	5.86	5.23	5.05	7.10	5.18	5.16	5.18	5.04	5.07	280.56	560.11	146.90	141.87	
400	nan	nan	8.89	13.53	6.50	5.38	5.26	5.28	6.17	5.01	5.06	5.19	5.07	4.97	276.90	34.54	14.31	743.94		
90	3	11.89	11.89	21.77	85.51	18.13	17.49	12.77	12.56	24.56	12.34	11.77	12.40	14.84	12.08	11.89	13.74	14.13	13.53	
	10	10.90	10.90	23.11	26.31	17.64	17.66	11.37	11.13	19.01	11.44	11.27	11.30	14.18	11.13	12.26	11.74	12.50	12.73	
	20	10.34	10.34	20.94	26.59	17.52	18.61	11.14	10.37	20.01	10.52	10.83	11.11	10.67	10.29	102.98	102.10	9.88	123.79	
	50	9.38	9.38	20.92	138.15	16.92	16.85	10.49	9.84	10.21	9.77	9.75	9.80	9.85	9.97	88.88	103.36	9.16	110.85	
	100	8.69	8.69	23.35	24.28	17.36	13.88	9.28	9.15	12.46	8.82	9.20	9.04	9.26	9.07	78.34	8.94	8.80	89.02	
	200	nan	nan	19.94	24.93	15.35	10.47	8.69	8.39	10.95	8.86	8.48	8.63	8.60	8.55	8.76	34.74	522.98	112.63	
400	nan	nan	17.99	99.05	12.21	9.93	9.19	8.68	10.67	8.47	8.70	8.70	8.72	8.58	24.81	418.33	11.02	15.35		
180	3	14.99	14.99	27.42	135.23	23.68	23.19	16.26	15.60	26.64	15.15	15.41	16.47	17.02	15.30	17.76	15.04	17.70	15.86	
	10	13.85	13.85	26.86	135.94	23.02	22.33	14.72	14.03	25.61	14.17	14.32	14.51	14.36	14.50	13.33	13.72	16.47	14.31	
	20	13.10	13.10	28.20	30.83	23.38	21.15	14.01	14.40	24.18	14.42	13.95	13.71	14.78	13.87	100.05	97.59	95.76	99.48	
	50	11.82	11.82	26.59	36.62	22.49	20.82	12.37	12.48	17.05	12.82	12.61	12.38	13.78	12.78	11.09	100.64	10.89	11.26	
	100	10.90	10.90	27.16	133.13	18.96	16.50	11.78	11.34	13.84	11.04	11.36	11.41	11.37	11.89	10.20	132.51	139.68	80.01	
	200	nan	nan	21.99	78.18	16.96	11.20	10.44	10.46	14.80	10.41	10.95	10.69	10.78	10.53	25.52	236.41	10.21	10.73	
400	nan	nan	23.12	78.83	15.51	11.23	11.25	10.81	13.33	10.95	10.91	11.09	11.29	10.82	72.45	71.79	16.08	38.84		
360	3	18.20	18.20	33.05	135.46	30.08	27.49	21.18	18.61	32.06	20.06	18.61	18.95	22.01	18.84	21.26	19.40	20.14	18.34	
	10	16.93	16.93	29.06	76.02	25.93	26.45	18.18	19.18	29.14	18.09	18.41	18.01	17.69	17.73	18.00	16.45	121.64	17.63	
	20	16.09	16.09	29.64	132.37	27.00	25.64	16.97	16.76	28.02	17.01	16.85	17.05	18.08	16.84	14.73	14.53	14.86	120.54	
	50	14.57	14.57	26.33	68.75	23.91	28.54	16.08	15.67	23.52	15.37	15.27	15.25	15.50	15.46	79.70	13.29	13.03	101.02	
	100	13.42	13.42	30.84	82.32	21.54	18.96	14.28	15.52	18.49	14.07	14.07	14.76	14.12	14.11	125.29	12.68	121.28	12.49	
	200	nan	nan	27.10	125.77	21.88	13.75	13.73	13.31	14.86	13.29	13.46	13.93	14.15	13.32	17.34	18.89	12.12	126.29	
400	nan	nan	30.96	43.55	20.64	13.38	14.21	14.34	15.46	14.03	13.19	14.18	14.16	14.01	589.04	100.91	16.83	16.57		

TABLE XV: Multivariate without Time: Mean Squared Error

n	sequence ahead length	Linear		LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer	
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	1 Layer	2 Layer	3 Layer	LinearNet	no conv	CLinear	NLinear	DLinear	1	2	no relu	relu
1	3	0.87	0.87	1.09	1.62	1.38	1.59	1.66	0.95	1.34	0.87	0.95	1.04	1.10	0.92	0.93	4.16	1.58	1.91
	10	0.84	0.84	1.24	1.09	1.03	1.15	1.22	1.13	1.38	0.89	0.98	0.95	1.26	0.96	0.99	3.59	2.34	0.95
	20	0.81	0.81	1.23	1.04	1.07	0.99	2.94	2.14	1.98	0.95	0.96	0.86	1.92	1.54	2.74	228.76	0.93	1.52
	50	0.80	0.79	1.43	2.08	1.39	1.20	3.86	0.88	5.24	0.99	0.84	0.86	0.91	0.91	1.15	4.62	0.83	322.19
	100	0.79	0.78	1.14	1.38	1.38	1.08	1.78	2.85	7.61	0.98	0.87	1.53	0.86	0.86	14.06	8.12	124.68	0.99
	200	nan	nan	1.05	12.48	1.06	1.26	1.52	2.96	17.34	0.82	0.81	272.42	1.91	4.06	27.81	246.04	331.66	121.42
400	nan	nan	1.58	1.14	1.04	2.48	2.28	4.35	45.70	0.96	1.31	3.11	2.50	1.35	42.97	129.12	105.51	141.03	
6	3	2.75	2.75	4.07	4.26	5.97	4.79	3.71	3.18	3.85	3.38	3.55	3.91	4.64	12.00	3.40	250.81	5.29	3.47
	10	2.45	2.45	4.28	8.75	5.89	5.93	3.57	3.00	4.97	2.98	3.09	5.92	3.37	4.29	189.52	3.27	3.23	4.40
	20	2.16	2.16	3.59	4.74	4.44	4.41	2.68	3.09	5.80	3.10	2.88	4.17	2.73	2.80	3.79	3.02	3.74	3.01
	50	2.07	2.07	5.28	6.33	6.00	4.32	2.61	3.31	8.58	2.51	2.60	2.41	2.88	4.28	2.44	2.80	65.80	4.59
	100	1.99	1.99	3.90	7.91	4.92	5.73	3.31	2.77	13.38	2.52	2.30	3.79	2.28	2.63	4.35	7.12	536.31	30.47
	200	nan	nan	4.09	4.61	6.02	8.17	3.39	2.98	22.64	2.06	2.08	2.10	2.26	2.42	26.11	18.45	4.40	3.52
400	nan	nan	4.09	9.70	4.35	9.52	5.07	3.85	41.20	2.31	2.23	4.84	2.16	7.73	91.44	191.59	213.75	90.30	
15	3	5.29	5.29	9.09	9.79	12.59	11.76	6.72	6.21	7.33	5.78	7.38	7.22	6.88	7.21	9.68	7.89	7.45	6.48
	10	4.24	4.24	9.30	16.19	8.39	7.77	6.00	5.07	9.59	5.02	5.91	6.25	5.56	5.79	5.19	5.11	5.02	5.55
	20	3.87	3.87	8.59	9.23	11.69	11.70	5.19	4.88	11.18	4.29	4.81	4.68	5.32	4.78	4.32	5.40	536.29	4.75
	50	3.68	3.68	9.83	18.83	9.15	8.92	5.19	4.98	15.33	4.21	4.29	4.82	4.26	4.27	7.21	16.00	6.01	4.17
	100	3.48	3.48	10.32	10.22	7.87	8.59	6.01	5.10	15.28	5.05	5.19	254.80	3.94	4.47	28.35	141.99	4.56	4.09
	200	nan	nan	12.43	8.42	10.40	14.22	6.66	5.30	38.12	3.95	3.90	5.38	3.81	3.73	31.06	41.91	218.27	284.20
400	nan	nan	9.52	30.14	9.78	14.99	7.92	6.74	47.23	3.85	4.23	6.73	171.07	4.10	102.91	196.28	150.91	84.00	
30	3	7.01	7.01	11.77	16.51	18.91	14.80	8.09	7.52	12.53	7.61	8.14	8.82	9.96	8.50	7.87	7.79	8.13	8.17
	10	6.09	6.09	14.67	17.02	13.31	14.56	7.55	6.75	15.14	7.91	8.22	10.36	7.38	7.53	7.50	408.61	8.20	6.53
	20	5.72	5.71	17.05	16.37	17.78	20.04	7.18	6.90	20.16	6.60	7.87	7.35	162.49	7.04	6.95	6.54	9.78	7.40
	50	5.30	5.30	20.03	22.94	16.85	20.95	7.72	7.33	27.28	6.03	6.33	6.27	426.30	6.46	8.06	7.86	289.12	14.71
	100	5.01	5.00	18.20	18.09	21.63	14.43	7.68	7.09	23.66	5.69	6.10	5.86	5.48	6.46	34.59	112.88	6.78	9.02
	200	nan	nan	18.35	25.15	13.88	30.23	9.80	9.15	42.38	5.44	6.51	6.58	6.17	5.19	205.53	283.72	124.93	358.75
400	nan	nan	18.61	23.51	14.53	19.49	11.41	10.35	47.16	6.17	5.37	7.50	7.39	5.32	131.13	95.38	447.65	298.91	
90	3	12.90	12.91	29.70	38.01	33.52	33.24	15.51	15.65	33.70	15.93	14.94	12.65	15.95	16.16	15.16	14.97	13.21	14.72
	10	11.10	11.10	37.24	39.70	32.73	32.05	14.76	14.59	32.82	14.76	14.55	12.16	14.71	13.13	367.88	13.79	15.45	315.88
	20	10.21	10.21	40.58	43.18	31.18	34.78	13.16	14.44	43.43	12.32	15.21	13.24	12.36	12.48	10.34	11.70	17.48	10.22
	50	9.05	9.05	45.48	44.17	30.06	46.04	14.42	12.26	41.71	10.54	11.89	17.18	10.64	134.55	139.70	322.35	381.22	492.01
	100	8.66	8.66	52.84	39.72	31.33	34.55	15.91	12.99	41.28	10.72	9.98	11.65	11.99	15.75	16.47	41.08	98.58	379.43
	200	nan	nan	41.03	43.49	32.45	61.13	20.43	21.64	95.29	10.12	8.68	8.99	267.08	9.83	531.07	246.84	163.73	84.05
400	nan	nan	34.99	36.33	34.43	30.50	21.19	19.08	83.16	10.35	9.97	15.83	10.34	114.68	245.02	33.27	299.67	168.89	
180	3	15.49	15.49	40.35	49.65	45.37	36.30	18.96	19.64	44.75	18.47	19.50	19.99	19.42	19.31	18.10	18.68	18.30	16.92
	10	13.39	13.39	49.11	51.50	43.57	39.11	18.41	17.66	39.25	17.17	18.54	18.10	17.56	18.09	14.66	19.53	18.44	352.11
	20	12.35	12.35	49.07	70.99	42.60	39.04	17.99	18.70	42.62	19.31	15.45	16.96	15.90	197.46	14.38	22.20	15.38	186.87
	50	11.42	11.41	47.01	63.70	39.42	49.35	18.28	18.27	57.26	13.73	16.86	13.80	11.61	15.68	615.11	40.94	13.05	403.75
	100	11.58	11.57	72.81	48.98	41.86	65.99	20.82	18.87	77.79	13.21	12.31	12.20	10.83	13.42	19.85	456.92	19.32	19.26
	200	nan	nan	80.79	52.46	37.80	25.08	25.20	23.21	84.99	14.17	11.11	11.06	247.20	12.57	124.35	53.66	108.01	246.80
400	nan	nan	54.95	53.63	42.19	48.65	27.70	26.32	87.67	12.24	12.65	13.57	18.77	148.21	176.55	508.44	107.24	72.38	
360	3	17.78	17.78	42.46	49.07	43.98	39.55	22.10	20.33	48.30	21.14	22.79	22.82	22.70	22.12	21.60	20.62	267.71	24.73
	10	15.69	15.69	53.77	48.96	45.69	37.68	21.12	20.19	53.21	20.38	22.50	17.87	70.22	21.67	19.41	22.48	19.15	16.57
	20	14.78	14.77	58.54	59.96	40.30	45.76	20.29	20.05	62.87	20.23	19.27	19.24	33.13	18.58	15.28	21.28	14.89	23.32
	50	13.97	13.97	72.86	55.62	34.56	54.77	21.36	19.47	83.15	17.21	16.92	23.27	50.90	18.76	98.77	20.50	15.22	185.22
	100	13.92	13.91	83.80	49.71	32.95	35.76	21.65	21.36	68.60	15.01	14.03	19.88	257.18	15.16	271.71	45.54	19.61	21.68
	200	nan	nan	84.15	60.97	35.52	46.58	24.86	23.32	87.26	15.64	13.88	14.46	14.33	15.21	144.87	380.25	25.82	64.30
400	nan	nan	75.32	52.22	42.29	50.71	32.99	27.54	65.54	16.11	16.19	105.63	14.72	19.22	199.03	779.54	85.02	93.26	

TABLE XVI: Multivariate with Time: Mean Squared Error

n	sequence ahead length	Linear		LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet			TSMixer	TSMixer	TSMixer	TSMixer		
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	1 Layer	2 Layer	3 Layer	no conv	CLinear	NLinear	DLinear	1	2	no relu	no relu	
1	3	0.86	0.86	1.05	1.27	1.66	1.13	0.92	1.00	1.17	0.87	0.94	0.88	0.88	1.22	0.91	1.03	2.35	1.84
	10	0.82	0.82	1.28	1.83	1.11	1.35	0.98	1.20	1.54	0.89	0.96	1.13	1.09	1.19	0.97	0.92	1.62	0.98
	20	0.80	0.80	1.11	1.29	1.27	1.16	1.28	0.89	2.39	0.86	0.94	0.87	0.88	2.01	234.07	0.91	0.91	1.08
	50	0.79	0.79	1.88	1.71	3.24	1.25	1.16	1.22	3.90	0.85	0.93	0.96	0.93	2.21	1.02	0.97	1.02	0.86
	100	0.79	0.79	1.62	2.30	1.31	2.04	1.36	1.04	9.77	0.85	1.45	1.39	0.80	2.96	147.32	3.10	1.62	185.25
	200	nan	nan	1.49	1.21	1.48	2.35	1.47	2.62	18.57	0.82	1.39	1.09	1.34	5.32	393.69	151.92	1.29	4.69
400	nan	nan	1.28	2.00	1.24	3.06	2.12	1.98	46.64	1.05	0.84	3.19	1.58	1.82	15.97	144.73	253.28	122.98	
6	3	2.57	2.57	3.52	5.30	4.52	5.22	3.32	3.09	4.12	2.97	3.27	3.32	3.21	3.26	2.94	2.99	3.81	3.17
	10	2.26	2.26	3.44	21.34	4.56	3.68	2.83	3.21	5.10	2.67	3.12	3.63	3.11	3.47	2.86	2.99	2.68	3.01
	20	2.08	2.08	4.84	6.15	4.24	3.77	3.22	2.64	5.31	2.50	2.82	3.08	2.74	2.76	2.50	2.56	2.90	2.98
	50	2.02	2.02	4.84	5.89	4.96	4.54	3.04	3.08	7.97	2.31	2.64	2.76	2.77	3.00	3.03	3.29	162.76	2.17
	100	1.98	1.97	4.11	6.47	4.52	4.26	3.13	2.94	15.54	2.30	2.25	5.57	215.63	2.28	8.76	4.65	2.69	2.63
	200	nan	nan	3.69	7.05	4.62	6.62	3.41	3.42	18.48	2.11	2.27	216.08	2.20	2.67	337.89	245.81	200.58	6.25
400	nan	nan	5.41	12.91	3.74	7.33	4.82	4.27	26.19	2.22	2.58	3.32	3.14	7.01	153.95	56.12	55.62	72.38	
15	3	4.69	4.69	8.22	9.34	10.89	8.81	5.81	5.12	6.53	5.75	6.17	6.24	6.47	6.03	5.72	6.46	5.69	5.61
	10	3.98	3.98	8.17	10.56	7.87	8.59	4.94	4.81	8.21	4.46	5.21	5.54	5.28	5.52	6.69	428.52	5.09	5.58
	20	3.72	3.72	9.46	9.06	9.88	10.05	5.20	4.85	10.18	4.50	5.16	7.87	4.76	4.67	4.39	5.80	4.69	4.88
	50	3.55	3.55	9.62	14.34	11.35	10.82	4.91	4.97	13.03	4.97	5.50	4.72	4.31	4.83	4.93	5.08	4.70	4.01
	100	3.45	3.45	9.17	20.99	9.08	11.15	5.42	4.81	20.42	4.24	4.02	4.82	5.12	5.71	5.20	20.90	5.46	4.08
	200	nan	nan	10.02	14.10	9.55	11.89	6.45	5.42	37.55	3.89	3.84	5.66	4.13	4.67	216.23	111.26	241.16	4.61
400	nan	nan	9.33	9.85	14.08	10.57	7.93	6.57	26.30	3.92	3.68	3.87	4.29	6.09	241.69	162.49	13.80	42.47	
30	3	6.62	6.63	12.77	14.02	14.35	15.16	7.62	7.42	11.48	7.41	8.68	8.04	9.40	8.50	9.36	7.77	7.59	7.38
	10	5.88	5.88	15.48	23.43	16.73	16.34	7.36	7.44	14.50	7.04	8.38	8.38	7.21	7.25	7.35	68.11	7.40	6.57
	20	5.62	5.62	20.45	23.11	17.30	16.78	8.04	6.84	19.64	6.93	7.05	7.01	7.02	7.67	7.74	7.23	242.46	7.16
	50	5.25	5.25	23.41	33.70	16.17	16.06	8.00	7.97	25.92	6.32	6.38	6.25	6.24	263.83	8.84	215.71	10.65	348.21
	100	5.13	5.12	17.60	26.79	15.96	15.62	7.85	7.56	30.12	6.69	5.76	6.38	6.67	7.78	7.33	74.18	306.91	11.84
	200	nan	nan	16.95	21.36	15.70	20.80	9.82	9.28	49.29	6.40	5.56	5.23	5.85	8.13	16.37	196.66	367.85	7.71
400	nan	nan	17.83	27.29	14.35	12.39	11.51	9.66	31.07	5.96	5.96	7.14	6.14	8.48	222.01	185.17	352.89	437.86	
90	3	12.41	12.42	31.50	37.67	33.82	31.98	14.79	15.00	31.98	15.24	14.68	15.01	15.04	14.75	21.40	14.15	15.42	148.51
	10	11.07	11.07	37.68	37.63	31.12	27.11	14.59	14.87	34.04	14.20	14.83	14.15	15.03	14.53	13.34	13.08	17.15	12.29
	20	10.37	10.37	35.67	44.00	33.85	30.75	15.28	14.01	42.45	13.64	13.28	12.06	38.67	17.38	117.39	14.47	12.38	12.98
	50	9.24	9.24	47.81	46.38	30.79	31.32	15.27	14.86	61.98	13.55	11.95	13.90	12.08	12.21	462.62	22.74	10.76	115.39
	100	8.97	8.96	58.02	34.52	31.25	43.02	15.36	17.78	56.27	14.44	10.13	11.31	10.06	10.72	182.93	95.11	320.99	18.80
	200	nan	nan	46.83	33.35	31.20	24.05	20.85	19.72	75.53	15.51	9.36	208.19	9.91	8.57	92.94	68.51	21.53	212.53
400	nan	nan	40.97	37.55	33.14	43.22	21.68	20.72	81.35	11.23	9.86	10.02	9.13	12.79	60.18	46.20	115.60	79.14	
180	3	14.94	14.94	43.71	43.77	43.95	35.60	18.42	17.66	49.05	15.82	18.45	238.87	18.37	18.58	17.81	17.00	15.41	17.56
	10	13.36	13.36	55.17	56.24	39.95	36.24	18.33	17.68	51.81	17.99	16.94	20.29	13.36	17.06	61.89	27.39	14.03	322.67
	20	12.50	12.50	54.74	51.52	38.60	39.86	18.27	18.12	53.12	16.97	17.60	242.93	16.70	209.43	17.36	32.19	107.52	135.71
	50	11.55	11.54	68.51	41.78	44.40	58.33	19.69	18.77	67.61	16.91	15.18	13.91	14.79	15.16	14.80	14.50	16.68	16.40
	100	11.83	11.83	46.66	58.48	39.14	46.16	21.25	19.78	69.66	14.20	12.86	11.50	12.36	147.25	17.59	22.84	310.78	15.62
	200	nan	nan	58.66	57.23	33.58	28.46	27.24	23.46	85.32	12.02	11.37	11.04	356.74	10.68	167.82	33.78	756.57	67.43
400	nan	nan	76.19	48.52	44.65	50.48	28.90	29.27	72.34	13.32	13.41	13.75	222.25	93.58	49.34	542.91	91.09	434.84	
360	3	17.26	17.26	44.28	54.76	50.34	43.67	21.15	21.78	56.22	21.31	20.95	21.36	22.09	21.12	20.58	22.50	20.39	20.83
	10	15.72	15.72	51.88	57.81	43.68	40.35	21.07	23.13	59.94	19.26	19.29	20.29	17.07	20.23	29.64	20.07	23.27	20.96
	20	14.92	14.92	59.43	51.84	50.76	36.47	21.34	19.71	67.56	19.12	19.60	17.04	214.95	17.95	14.55	20.09	30.83	20.98
	50	14.09	14.09	68.81	50.27	42.97	51.25	22.97	19.00	76.30	19.14	17.49	17.87	359.46	198.78	27.40	19.78	14.40	44.04
	100	14.13	14.12	61.50	56.05	45.72	64.54	22.79	20.58	63.41	18.87	15.25	17.62	13.93	14.54	21.44	58.37	20.36	25.77
	200	nan	nan	74.01	50.55	37.52	37.41	25.15	23.82	65.03	15.48	14.49	240.85	67.08	16.76	135.37	49.14	25.16	357.21
400	nan	nan	68.65	71.22	40.35	46.68	31.09	27.46	80.68	15.24	15.92	15.38	19.92	16.13	80.83	203.56	127.84	339.34	

B. Training Times

TABLE XVII: Univariate without Time: Training time

n ahead	sequence length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet			TSMixer		TSMixer	TSMixer			
		Regression	Ridge	Encoder Decoder	Shared Encoder Decoder	Last Sequence	Full Sequence	1 Layer linear	2 Layer linear	3 Layer ReLU	LinearNet no conv	CLinear	NLinear	DLinear	1	2	no relu 1	no relu 2	
1	3	0.0	0.0	6.7	7.5	7.7	7.7	9.5	5.2	5.5	9.3	10.9	8.0	5.9	7.4	10.9	9.2	9.5	9.5
	10	0.0	<u>0.0</u>	7.9	16.3	9.5	12.0	8.5	3.5	7.2	10.5	10.1	6.9	7.8	4.4	12.2	12.7	12.3	15.5
	20	0.0	<u>0.0</u>	9.4	22.6	16.9	15.0	9.7	4.3	7.5	9.2	10.4	13.3	8.7	3.6	7.8	11.8	8.1	10.2
	50	0.0	<u>0.0</u>	14.6	20.7	25.0	32.0	7.6	5.3	7.9	10.2	9.8	8.8	8.3	4.4	10.8	16.3	15.1	15.6
	100	0.0	<u>0.0</u>	21.9	14.4	35.0	44.9	7.2	7.5	8.6	10.6	9.8	9.7	4.1	3.5	6.3	15.1	11.3	9.7
	200	nan	nan	37.0	60.2	82.5	85.4	9.1	6.8	7.6	10.6	11.7	9.1	7.2	4.6	6.2	8.5	9.7	10.1
	400	nan	nan	62.5	162.1	97.5	178.7	13.0	13.9	6.7	10.7	13.9	11.1	4.9	5.5	8.4	10.1	12.6	8.7
6	3	0.0	0.0	8.7	14.5	7.5	8.8	8.2	5.2	9.1	7.5	6.2	7.8	3.6	4.0	8.1	8.3	9.2	7.1
	10	0.0	<u>0.0</u>	10.2	12.8	10.3	11.6	7.4	5.2	7.3	10.7	8.3	9.0	4.8	5.2	12.3	15.3	10.2	11.7
	20	0.0	<u>0.0</u>	12.5	13.7	23.2	19.7	10.0	4.4	6.6	7.8	8.6	9.2	6.5	4.9	8.5	8.8	8.2	10.5
	50	0.0	<u>0.0</u>	13.4	29.5	28.5	26.4	7.9	6.0	8.0	8.6	7.4	8.2	5.4	5.5	9.6	11.4	12.0	9.4
	100	<u>0.0</u>	0.1	16.1	271.2	66.9	45.9	9.2	7.1	6.5	9.0	9.1	8.0	5.8	6.3	8.8	17.3	17.1	15.5
	200	nan	nan	27.3	50.9	92.0	85.8	15.3	14.6	10.3	10.4	12.9	11.6	8.6	11.0	8.8	11.3	12.0	9.9
	400	nan	nan	43.6	117.2	220.6	135.4	12.0	10.6	8.3	9.4	10.1	10.0	4.3	9.8	8.1	9.9	11.5	9.0
15	3	0.0	0.0	8.2	12.0	7.7	9.1	7.2	8.4	12.5	7.7	4.3	26.1	5.1	4.3	8.6	10.1	7.8	10.6
	10	0.0	<u>0.0</u>	7.9	14.8	12.3	11.3	5.7	7.4	10.5	7.1	7.1	6.6	5.6	5.3	8.5	9.7	11.3	7.4
	20	0.0	<u>0.0</u>	10.0	13.4	18.4	14.7	6.7	7.3	10.6	8.0	7.8	6.4	6.6	4.8	8.9	8.8	12.6	8.7
	50	0.0	<u>0.0</u>	10.5	19.7	29.8	31.6	7.6	7.7	9.6	8.2	7.1	6.1	4.1	5.3	8.5	9.8	10.9	11.6
	100	0.0	<u>0.0</u>	15.5	11.9	50.6	37.3	8.1	9.5	10.0	8.6	7.8	6.6	5.6	7.0	6.7	9.0	9.3	7.8
	200	nan	nan	23.6	26.9	196.2	71.0	14.7	15.9	15.3	9.3	11.3	10.9	8.9	9.8	10.4	11.5	11.2	12.5
	400	nan	nan	50.9	25.5	815.2	123.8	11.4	10.2	11.0	9.3	10.2	8.3	3.8	10.2	7.6	9.5	9.2	8.3
30	3	0.0	0.0	7.8	126.1	8.1	9.1	20.5	6.3	11.2	6.9	4.0	2.9	2.4	3.7	9.2	8.3	9.8	8.6
	10	0.0	<u>0.0</u>	12.9	14.5	10.5	11.3	5.4	5.1	7.7	8.0	5.5	4.9	3.9	4.5	7.5	10.1	11.5	8.4
	20	0.0	<u>0.0</u>	13.4	11.2	23.7	17.9	5.4	4.8	8.1	8.2	6.2	4.7	4.5	5.2	8.0	10.7	12.0	8.9
	50	<u>0.0</u>	0.0	18.1	26.8	30.2	25.9	6.8	4.8	6.9	7.4	6.5	5.8	6.4	5.7	8.6	9.8	9.7	8.9
	100	0.0	<u>0.0</u>	19.4	18.1	50.2	26.7	6.5	6.5	6.7	8.4	7.0	5.8	4.8	6.7	7.0	9.8	9.6	9.4
	200	nan	nan	25.0	32.8	76.3	62.8	8.8	7.7	9.3	7.0	8.9	7.9	6.7	7.8	7.8	10.0	10.3	8.5
	400	nan	nan	27.0	143.7	98.0	60.2	7.9	8.4	8.6	7.3	8.4	5.4	9.2	10.7	8.6	10.8	12.1	10.0
90	3	0.0	<u>0.0</u>	12.2	274.2	7.9	8.3	3.1	9.5	13.7	5.9	4.2	2.3	2.7	3.3	8.1	11.5	7.4	6.9
	10	0.0	0.0	20.3	37.7	10.3	12.5	8.7	7.0	10.4	5.9	5.3	4.8	2.7	4.0	7.4	11.3	8.0	8.7
	20	<u>0.0</u>	0.2	13.8	40.4	12.1	13.3	4.2	7.7	11.2	6.0	5.8	4.9	3.4	5.9	8.6	9.2	11.4	7.1
	50	<u>0.0</u>	0.1	20.8	145.4	18.6	19.4	4.9	7.2	9.7	5.5	5.8	4.4	4.4	5.7	8.2	10.0	9.2	8.8
	100	<u>0.1</u>	0.1	20.7	30.8	24.3	22.1	6.4	7.7	11.2	6.2	6.0	4.2	3.2	8.2	8.4	10.8	11.7	8.7
	200	nan	nan	30.1	35.3	56.0	51.6	9.8	11.1	12.1	5.9	7.4	6.1	4.7	9.1	8.5	10.5	10.2	8.9
	400	nan	nan	44.5	41.0	76.1	66.4	6.6	7.7	10.9	6.6	6.1	3.7	5.5	10.7	6.7	8.8	8.3	7.9
180	3	0.0	0.0	19.2	52.4	7.8	9.7	2.9	8.9	12.2	5.2	3.6	2.1	4.5	3.3	6.1	8.2	7.2	9.7
	10	<u>0.0</u>	0.2	23.1	479.5	10.8	11.5	4.0	7.1	9.5	6.2	4.2	2.8	4.5	3.9	10.5	10.8	9.1	9.1
	20	0.0	0.1	18.2	40.2	12.7	11.0	4.5	7.1	9.0	6.6	3.9	3.3	2.6	4.8	7.8	9.1	9.2	7.6
	50	<u>0.0</u>	0.1	25.2	42.3	16.8	18.5	4.7	8.0	9.7	5.6	5.1	3.9	2.6	5.5	8.9	11.4	9.6	11.2
	100	0.1	<u>0.0</u>	28.5	539.4	28.2	24.8	6.0	7.8	8.4	6.4	5.7	3.3	4.2	5.4	6.8	9.8	10.3	9.0
	200	nan	nan	30.5	53.1	68.6	39.6	6.9	8.4	10.6	6.4	5.6	4.7	6.9	8.2	7.6	10.1	10.9	8.9
	400	nan	nan	47.4	267.6	72.8	62.6	8.7	7.3	9.9	5.5	4.9	4.5	3.8	7.0	6.3	8.5	8.5	7.4
360	3	0.1	0.0	28.3	117.1	10.2	9.7	3.1	7.1	15.2	4.9	4.7	2.4	2.8	4.0	7.7	9.3	11.2	8.3
	10	<u>0.1</u>	0.1	27.0	77.1	8.6	11.4	4.7	5.2	12.5	5.0	4.3	3.1	3.0	3.9	7.9	10.4	9.8	8.7
	20	0.1	<u>0.0</u>	65.9	72.8	10.3	12.4	4.5	5.4	12.7	5.5	4.0	3.1	4.9	4.6	7.4	8.6	8.1	9.6
	50	0.2	<u>0.2</u>	29.8	88.6	13.4	14.7	4.3	4.6	11.2	5.8	5.2	3.9	2.5	5.3	6.6	11.0	9.6	9.7
	100	0.2	<u>0.2</u>	36.4	915.7	25.4	31.1	4.6	4.5	11.7	6.0	5.5	2.8	3.0	5.2	7.3	8.9	10.1	9.0
	200	nan	nan	50.2	90.8	33.0	38.5	7.9	6.8	13.9	5.3	5.2	3.9	4.3	7.1	7.5	9.8	9.4	8.3
	400	nan	nan	43.3	122.9	55.3	52.7	7.0	5.7	12.5	5.7	5.5	4.7	3.2	7.2	6.5	8.5	8.1	7.5

TABLE XVIII: Univariate with Time: Training time

n	sequence ahead length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet			TSMixer	TSMixer	TSMixer	TSMixer			
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	1 Layer	2 Layer	3 Layer	no conv	CLinear	NLinear	DLinear	1	2	no relu	no relu	
1	3	0.0	0.0	7.1	8.5	7.2	7.2	7.7	3.3	4.6	9.2	5.2	4.1	5.1	3.8	8.6	10.6	12.2	7.9
	10	0.0	0.0	9.0	6.4	8.5	9.9	9.5	3.9	5.3	9.4	9.3	11.7	7.8	4.6	9.6	15.2	13.8	9.5
	20	0.0	0.0	9.4	41.1	13.9	15.3	9.5	5.5	7.7	10.5	9.7	11.4	4.8	3.1	7.9	8.4	8.0	7.7
	50	0.0	0.0	19.1	35.0	30.6	33.0	9.2	5.5	6.7	11.0	9.8	8.8	4.6	4.6	9.5	14.6	14.5	15.7
	100	0.0	0.0	28.8	52.3	44.9	46.8	7.8	7.6	7.9	10.2	10.8	9.1	5.3	6.3	6.2	9.6	9.2	9.3
	200	nan	nan	44.7	60.0	70.9	89.5	9.7	9.7	8.0	11.1	13.6	10.0	5.0	4.3	6.6	8.9	9.3	10.5
	400	nan	nan	63.9	106.5	114.8	192.1	9.4	12.7	8.4	12.3	12.5	10.5	2.3	6.3	6.3	8.8	10.6	8.0
6	3	0.0	0.0	6.9	35.5	7.4	8.3	8.1	3.4	4.5	8.5	6.0	7.6	3.1	4.4	8.7	9.9	12.8	9.9
	10	0.0	0.0	9.7	17.1	11.3	10.0	7.7	4.2	5.4	8.2	7.6	8.6	4.5	7.0	9.0	12.7	9.6	9.3
	20	0.0	0.0	11.6	7.7	18.2	15.3	10.0	5.1	5.9	7.9	6.5	9.8	5.7	5.5	6.7	9.9	11.4	10.7
	50	0.0	0.0	13.7	125.7	32.4	31.3	8.2	5.2	4.8	8.3	9.1	8.5	6.3	5.8	8.4	9.6	10.1	9.7
	100	0.0	0.0	24.0	21.8	78.6	49.3	11.0	6.4	4.9	9.0	10.0	9.6	5.9	6.3	7.6	15.1	17.1	18.8
	200	nan	nan	55.9	72.1	109.7	73.1	8.2	6.0	5.3	8.6	10.4	10.7	5.6	6.0	7.0	8.4	7.5	8.8
	400	nan	nan	84.5	130.1	162.5	171.0	7.6	6.8	4.7	8.8	9.3	7.9	6.3	6.2	8.2	8.8	8.5	9.6
15	3	0.0	0.0	7.4	12.7	7.5	7.9	12.7	5.2	7.5	7.7	4.1	5.5	5.6	3.9	6.1	9.7	11.6	9.1
	10	0.0	0.0	10.1	22.9	11.4	10.2	5.9	5.7	7.3	7.3	6.5	5.7	3.9	4.8	12.2	10.3	8.7	7.4
	20	0.0	0.0	12.1	15.9	21.4	14.0	7.0	6.0	7.2	6.7	3.8	7.5	4.7	4.0	9.4	10.2	13.3	9.1
	50	0.0	0.0	10.7	23.3	38.2	33.0	7.3	6.7	6.5	9.1	7.1	6.5	6.4	6.3	7.9	9.9	10.5	8.5
	100	0.0	0.0	18.5	39.8	60.1	41.7	8.0	6.4	7.9	8.0	7.8	7.1	5.4	5.4	9.2	11.2	8.3	8.9
	200	nan	nan	21.3	13.8	103.5	66.7	8.2	8.7	6.7	7.5	8.8	8.3	5.9	5.9	5.8	7.8	7.6	9.6
	400	nan	nan	34.8	89.3	244.8	77.2	7.6	8.2	6.8	8.3	8.0	7.2	5.0	7.3	6.7	8.0	9.0	8.5
30	3	0.0	0.0	8.1	12.7	7.3	8.7	3.2	3.9	5.0	5.9	4.9	5.5	2.5	3.4	6.4	11.2	10.4	8.5
	10	0.0	0.0	9.8	11.3	12.2	11.0	5.2	3.4	5.2	7.1	6.2	4.7	4.2	5.6	10.5	8.3	9.8	10.2
	20	0.0	0.0	13.0	25.3	18.3	14.9	5.7	3.5	4.8	7.5	6.6	5.9	5.0	4.2	6.3	9.7	8.2	9.6
	50	0.0	0.0	13.8	17.8	99.1	29.8	6.7	3.4	4.2	8.1	7.3	7.7	4.1	5.2	7.4	9.4	9.5	10.2
	100	0.0	0.1	20.8	74.7	47.6	37.4	6.8	4.7	4.5	7.8	7.5	5.0	4.6	5.1	8.7	10.7	8.1	10.4
	200	nan	nan	29.1	41.6	59.6	66.9	6.9	4.8	4.4	7.2	5.8	7.5	5.3	5.0	6.0	8.0	9.5	7.6
	400	nan	nan	35.7	68.4	91.5	66.8	6.9	4.7	4.8	9.7	7.8	8.7	7.2	6.3	7.0	9.5	8.9	9.4
90	3	0.0	0.0	13.7	31.4	6.6	8.8	2.4	4.1	6.5	6.3	4.1	2.2	5.0	3.0	7.9	9.7	10.2	10.2
	10	0.0	0.0	12.2	59.7	9.8	8.4	3.7	4.6	6.9	6.4	5.3	3.8	3.3	3.7	6.8	10.5	11.7	11.0
	20	0.0	0.0	15.5	90.8	13.8	12.7	3.7	4.4	6.6	6.8	5.6	3.9	2.4	3.5	8.4	9.0	11.5	8.2
	50	0.0	0.0	19.3	26.9	26.2	18.0	4.5	4.8	6.3	6.6	5.1	5.0	4.3	4.1	7.1	11.5	8.6	9.9
	100	0.1	0.1	16.6	60.3	30.4	22.3	4.7	5.1	7.0	6.1	5.4	6.4	6.3	5.2	9.0	8.8	9.0	11.6
	200	nan	nan	55.6	642.2	70.9	45.4	6.8	5.4	6.3	6.3	5.3	7.0	6.5	6.0	7.0	8.3	8.1	8.9
	400	nan	nan	41.5	44.4	77.4	52.6	5.3	4.8	6.8	6.6	4.9	8.1	5.9	6.5	6.4	11.3	8.2	9.7
180	3	0.0	0.0	16.1	41.5	6.9	8.3	2.3	5.1	5.9	5.5	4.2	2.2	2.2	2.9	6.3	7.4	9.7	7.6
	10	0.0	0.1	58.1	38.6	8.6	9.9	3.5	4.9	5.5	5.6	4.3	7.6	2.5	4.1	8.1	13.0	11.6	10.4
	20	0.0	0.0	22.5	444.6	15.3	11.0	2.9	4.7	5.0	5.8	4.3	3.5	4.3	3.4	7.0	9.4	10.4	9.6
	50	0.0	0.1	19.6	591.2	16.6	21.8	4.2	5.4	5.0	5.9	5.3	3.7	4.7	3.5	6.6	10.8	9.6	9.2
	100	0.2	0.1	26.6	44.0	22.3	23.2	5.2	6.0	5.8	6.5	5.6	4.8	4.1	3.9	8.0	12.0	15.5	12.4
	200	nan	nan	34.2	48.3	43.8	37.9	5.9	5.3	5.2	6.0	5.2	3.7	5.7	4.2	6.2	7.9	7.7	8.0
	400	nan	nan	41.9	58.2	68.3	63.1	5.5	6.1	5.9	5.6	6.0	4.4	3.1	5.5	6.0	8.0	7.8	8.1
360	3	0.1	0.0	27.1	71.3	6.8	8.0	2.3	2.6	10.9	5.3	3.8	2.2	2.5	2.7	7.5	11.8	8.4	8.7
	10	0.1	0.1	27.6	72.6	7.4	8.8	3.3	3.0	16.2	5.4	3.9	3.6	3.2	3.5	8.7	9.4	8.7	10.6
	20	0.1	0.1	32.8	75.2	9.1	11.1	3.8	3.3	10.3	5.7	5.1	3.1	5.2	3.5	7.3	8.3	9.1	8.3
	50	0.1	0.1	31.8	73.9	13.3	17.2	3.0	3.2	9.4	5.8	5.4	2.9	3.9	3.3	7.4	8.7	9.5	10.2
	100	0.3	0.3	39.5	78.0	18.5	20.8	4.8	3.0	9.1	6.0	5.9	3.3	3.9	4.5	6.6	9.6	10.1	10.7
	200	nan	nan	51.5	87.2	29.3	30.7	4.4	5.0	10.5	6.2	7.7	3.6	8.7	5.1	6.7	8.7	9.2	8.5
	400	nan	nan	46.7	84.5	51.1	58.3	8.0	3.5	8.8	5.5	4.6	6.3	2.8	6.4	6.5	8.7	7.9	8.1

TABLE XIX: Multivariate without Time: Training time

n ahead	sequence length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet			TSMixer	TSMixer	TSMixer	TSMixer			
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	1 Layer	2 Layer	3 Layer	no conv	CLinear	NLinear	DLinear	1	2	no relu	no relu	
1	3	0.0	0.0	7.6	8.1	8.0	49.4	4.4	4.3	6.0	7.7	8.0	22.2	9.4	5.8	13.6	13.5	15.4	17.8
	10	0.0	0.0	7.4	12.9	14.1	13.5	4.6	4.6	6.4	9.1	7.9	11.0	40.1	14.4	14.8	11.5	12.4	13.7
	20	0.1	0.1	7.6	15.2	19.6	22.1	4.5	3.8	7.5	7.9	11.0	11.4	60.6	10.6	16.5	9.8	21.3	14.1
	50	0.6	0.1	9.8	11.7	43.6	31.4	4.3	6.4	5.5	12.1	12.1	9.2	7.7	6.0	20.8	19.8	22.7	9.4
	100	1.7	0.3	12.9	30.2	42.2	60.5	4.0	3.4	4.6	8.3	9.2	8.5	7.7	7.3	23.8	15.7	13.2	19.2
	200	nan	nan	28.1	14.9	118.9	107.9	5.4	3.4	4.8	11.3	12.4	2.8	7.0	3.7	14.1	10.1	12.0	11.3
	400	nan	nan	28.4	127.9	232.7	177.0	4.8	3.3	3.7	13.3	10.2	5.1	28.3	6.3	14.2	11.2	11.9	9.1
6	3	0.0	0.0	6.7	9.9	6.9	8.2	7.9	4.2	4.9	8.3	8.8	37.9	46.3	3.8	15.0	10.5	18.2	13.1
	10	0.0	0.0	7.1	7.2	8.4	8.5	4.6	4.1	5.2	9.3	7.1	58.8	9.5	64.3	9.9	14.7	12.3	13.2
	20	0.1	0.1	7.4	14.1	13.7	16.1	6.3	4.1	5.2	9.5	11.0	9.7	16.2	9.2	13.4	16.6	13.4	15.3
	50	0.4	0.2	10.0	11.9	21.6	36.0	7.0	4.5	5.0	9.7	10.1	13.5	6.0	5.4	16.3	14.0	11.4	16.3
	100	2.4	0.3	12.2	17.8	32.0	42.9	5.3	5.2	5.7	9.0	12.6	7.9	8.5	10.5	20.1	15.4	13.1	10.7
	200	nan	nan	17.7	44.9	53.5	37.3	5.9	5.4	5.4	10.2	11.3	11.0	11.7	9.5	16.0	14.1	14.8	17.5
	400	nan	nan	30.5	32.4	90.0	142.2	5.7	7.5	5.7	12.1	18.0	13.8	13.5	48.0	17.3	17.9	16.9	13.7
15	3	0.0	0.0	6.9	9.1	6.5	7.2	5.1	5.4	7.2	7.7	5.9	6.0	6.3	5.3	10.6	14.2	17.5	13.2
	10	0.0	0.0	7.3	11.4	10.2	10.4	5.4	6.8	8.0	8.0	7.2	6.1	4.5	41.2	12.1	14.1	13.4	10.2
	20	0.1	0.1	8.5	12.1	9.2	9.4	5.1	6.6	8.6	8.6	9.0	7.3	7.0	10.8	16.9	16.7	11.5	13.9
	50	0.5	0.1	10.9	72.6	16.5	14.9	4.8	6.1	7.0	8.5	9.0	5.6	7.7	6.5	17.7	85.8	17.9	12.8
	100	2.4	0.3	13.0	22.7	33.0	19.4	4.5	7.1	8.0	7.7	9.2	5.4	8.7	9.3	77.4	14.0	16.8	14.9
	200	nan	nan	17.6	33.2	44.8	53.5	4.3	8.8	7.3	11.2	11.0	8.7	7.9	9.9	14.4	15.4	14.3	11.4
	400	nan	nan	30.6	110.3	90.4	60.1	4.7	6.9	6.9	11.0	13.2	9.8	4.8	14.1	19.9	14.0	14.0	10.5
30	3	0.0	0.0	7.6	12.3	6.4	7.2	6.3	3.4	4.6	8.4	5.7	3.5	5.4	4.3	15.9	11.3	15.7	13.9
	10	0.0	0.0	8.6	14.2	7.7	6.9	4.8	3.6	4.4	7.8	7.0	3.9	4.0	5.9	14.8	9.3	14.8	10.5
	20	0.1	0.0	9.0	18.2	9.3	10.0	4.9	3.6	4.7	9.0	7.9	5.9	3.0	10.1	16.8	13.5	27.0	14.6
	50	0.7	0.1	10.4	17.4	15.4	12.5	4.3	3.5	4.4	7.9	8.5	7.9	3.0	8.3	16.0	13.5	11.6	14.5
	100	2.1	0.3	13.4	22.9	21.1	19.7	4.5	4.4	4.8	10.4	10.2	17.7	7.4	10.5	17.1	58.1	17.8	13.3
	200	nan	nan	18.5	33.8	49.2	39.5	4.4	5.0	5.3	10.3	10.0	7.1	9.1	12.0	12.7	28.2	13.0	10.3
	400	nan	nan	32.4	42.4	75.3	60.2	4.5	4.5	4.6	9.9	11.0	7.7	10.0	10.9	17.1	13.0	12.2	10.2
90	3	0.0	0.1	11.0	30.0	6.2	6.4	3.7	4.7	6.7	7.5	6.5	3.7	3.7	6.9	11.4	12.6	12.8	17.0
	10	0.1	0.1	11.4	32.5	7.1	7.2	5.0	4.7	6.6	7.0	7.2	3.8	4.5	4.7	9.5	12.4	14.9	8.6
	20	0.1	0.2	12.2	29.0	9.2	8.9	4.8	4.9	7.0	7.5	7.8	5.8	6.8	8.4	12.9	11.0	95.8	11.1
	50	0.9	0.1	13.5	32.5	12.5	12.8	3.7	4.4	6.6	8.0	9.7	5.1	3.8	6.4	11.5	11.0	11.5	9.6
	100	2.3	0.3	17.2	36.5	21.9	20.6	4.8	6.1	7.7	7.9	10.5	9.5	6.5	10.9	16.8	18.7	14.5	13.0
	200	nan	nan	21.5	38.4	35.7	32.7	4.3	4.7	6.8	9.8	13.3	7.6	4.9	21.0	14.2	12.5	15.7	12.3
	400	nan	nan	35.6	46.3	56.8	56.7	4.6	5.1	7.3	9.6	14.2	6.9	7.2	7.1	12.7	11.9	19.9	9.3
180	3	0.0	0.1	17.5	52.1	6.5	6.7	4.8	5.6	5.4	6.6	7.4	32.5	5.8	5.0	13.1	14.8	15.3	10.8
	10	0.1	0.0	18.1	50.3	7.3	7.6	4.1	5.0	5.2	8.0	7.3	3.8	4.1	4.3	11.8	10.2	14.5	8.2
	20	0.3	0.1	18.4	55.6	8.8	9.2	3.9	4.9	5.4	7.4	8.1	4.9	6.7	4.6	12.2	13.5	14.3	9.0
	50	1.0	0.4	20.7	52.6	12.6	12.9	3.7	5.2	5.4	7.0	9.4	12.6	4.3	7.5	12.0	64.1	15.0	9.7
	100	2.8	0.3	22.1	66.6	21.0	20.3	4.0	5.5	6.1	7.5	11.0	6.6	4.2	10.6	15.6	16.4	16.2	14.1
	200	nan	nan	26.4	65.4	32.4	32.3	4.9	5.6	5.8	7.0	11.2	6.6	3.5	7.6	11.5	11.0	12.1	9.5
	400	nan	nan	38.8	68.7	56.5	56.4	4.2	6.3	6.5	7.2	10.2	7.2	4.5	7.0	12.0	11.0	11.7	10.2
360	3	0.1	0.0	25.8	704.2	6.7	7.0	3.7	3.0	9.9	6.5	7.8	5.5	4.9	5.1	20.7	13.5	11.6	13.0
	10	0.3	0.1	26.1	255.4	7.6	8.1	4.3	3.3	9.5	7.6	8.3	3.6	3.1	5.0	12.5	12.7	13.1	11.2
	20	0.4	0.1	47.0	104.9	9.4	10.7	4.0	3.1	8.6	6.5	8.4	7.9	3.2	5.4	12.3	16.1	13.4	12.1
	50	1.2	0.2	64.6	107.9	12.6	13.2	4.0	3.3	9.0	7.4	9.6	6.0	4.3	6.9	12.2	11.5	13.0	9.6
	100	3.0	0.4	43.6	113.9	21.4	19.8	3.8	3.9	10.0	6.5	10.7	5.4	2.9	7.7	12.4	15.6	12.2	13.4
	200	nan	nan	35.4	110.6	33.6	38.6	4.0	4.0	10.2	6.7	9.3	7.0	5.5	9.0	12.2	15.2	12.6	11.5
	400	nan	nan	44.5	127.9	55.7	55.3	3.9	4.3	10.9	6.6	10.1	6.1	6.5	6.7	14.4	10.8	11.5	16.0

TABLE XX: Multivariate with Time: Training time

n	sequence ahead length	Linear	LSTM	LSTM	GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer		
		Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	1 Layer	2 Layer	3 Layer	no conv	CLinear	NLinear	DLinear	1	2	no relu	no relu	
1	3	0.0	0.0	6.9	10.1	9.1	6.9	6.7	3.7	5.5	8.4	7.5	11.6	6.7	3.7	14.5	16.3	56.2	12.8
	10	0.0	0.0	7.4	8.5	12.8	14.0	5.2	3.8	6.6	11.3	6.1	4.4	20.1	11.0	15.8	13.7	16.0	12.8
	20	0.1	0.1	10.3	9.2	13.1	16.0	4.6	4.4	6.0	10.2	9.0	11.5	9.9	9.2	11.6	17.6	17.7	17.4
	50	0.8	0.2	17.1	19.6	23.2	36.3	5.6	4.0	5.5	10.8	13.4	10.7	10.4	7.5	17.5	16.4	16.7	18.2
	100	3.0	0.4	18.5	18.5	68.7	130.4	5.5	5.3	4.9	8.2	8.5	8.3	10.9	4.9	17.8	18.1	17.3	11.9
	200	nan	nan	27.1	55.5	91.9	49.7	7.2	4.2	3.7	13.6	10.1	14.1	7.3	3.6	12.7	11.5	20.1	40.9
	400	nan	nan	49.7	43.3	253.5	216.9	5.1	5.6	5.2	13.6	12.7	4.2	7.9	7.2	25.3	11.9	13.1	10.5
6	3	0.0	0.0	7.5	8.7	8.2	7.2	6.8	3.7	4.6	7.3	6.6	7.8	5.6	7.4	12.0	14.6	13.0	14.5
	10	0.0	0.0	7.7	7.4	8.6	9.5	4.9	4.1	4.8	10.2	6.2	62.3	6.8	9.1	14.0	15.9	16.3	11.7
	20	0.1	0.2	7.4	13.1	12.9	14.8	5.5	5.5	5.4	9.7	9.8	9.1	8.2	7.6	16.4	18.8	16.7	16.4
	50	0.9	0.2	9.9	12.4	18.7	24.3	5.2	7.2	5.0	11.1	10.5	9.0	9.4	8.9	87.1	14.3	13.8	16.6
	100	2.7	0.4	12.7	18.3	36.6	36.6	6.2	6.7	5.5	10.0	13.2	37.2	3.5	9.1	22.3	16.8	17.2	15.3
	200	nan	nan	17.6	23.0	62.4	64.7	6.0	5.0	5.1	11.0	11.9	6.3	9.4	12.5	13.1	12.6	13.3	13.4
	400	nan	nan	29.8	33.8	128.3	72.7	4.9	6.5	7.8	11.8	16.7	13.8	14.9	14.5	39.0	18.6	14.7	12.6
15	3	0.0	0.0	6.9	9.2	6.5	7.2	4.3	5.6	6.9	8.3	5.8	7.3	5.0	4.7	13.6	16.7	14.6	14.3
	10	0.0	0.0	7.4	10.7	10.0	9.5	5.1	5.9	7.1	9.2	6.7	6.6	5.4	6.0	13.9	9.3	15.8	10.6
	20	0.1	0.1	8.6	11.8	11.4	9.6	4.9	6.4	7.9	8.2	7.8	6.2	11.5	19.5	14.6	80.1	16.2	16.2
	50	1.1	0.2	11.1	13.4	17.1	16.2	4.8	7.4	8.1	7.8	9.9	9.0	8.8	6.9	14.7	15.9	17.3	13.9
	100	4.0	0.4	14.2	18.0	29.4	22.4	4.9	6.0	7.0	8.9	10.3	11.1	8.2	7.8	17.0	16.7	17.3	15.5
	200	nan	nan	19.2	21.1	51.9	48.8	5.6	8.0	7.3	10.8	12.7	7.7	8.4	10.6	14.2	15.2	24.9	11.7
	400	nan	nan	27.6	42.9	69.5	57.3	6.1	7.9	9.2	10.4	18.8	12.2	8.6	14.3	13.7	16.2	13.9	12.7
30	3	0.0	0.0	7.7	13.8	7.2	7.0	4.2	3.6	4.9	8.1	6.1	4.6	5.8	4.1	14.7	13.2	17.0	13.4
	10	0.0	0.1	8.1	13.5	9.4	7.6	4.9	3.6	4.5	8.8	6.1	3.4	5.0	5.5	13.9	9.4	13.7	9.9
	20	0.2	0.2	9.4	14.8	9.5	9.3	4.9	4.1	5.1	7.8	8.1	7.5	6.5	32.5	16.5	14.5	11.1	12.6
	50	0.9	0.2	11.7	18.8	15.7	15.1	4.5	3.7	4.7	8.7	12.2	12.3	7.8	5.3	15.7	10.8	76.3	9.5
	100	3.9	0.4	13.8	22.2	26.8	22.5	4.0	4.3	4.6	9.0	10.5	7.4	6.8	8.4	15.3	15.9	13.4	19.1
	200	nan	nan	18.6	27.0	44.2	34.5	4.7	5.7	5.6	9.9	13.6	13.9	9.3	8.3	12.8	11.5	12.9	12.1
	400	nan	nan	30.1	50.0	94.2	57.3	5.4	6.0	6.9	8.3	12.1	6.8	7.8	24.2	14.3	15.1	12.1	10.1
90	3	0.0	0.0	11.6	31.2	6.4	7.0	5.1	4.8	7.2	6.7	7.4	6.9	5.6	35.3	11.4	19.3	17.8	11.5
	10	0.0	0.1	12.2	31.3	7.9	7.2	4.0	4.6	6.9	6.9	6.6	6.1	7.2	7.6	10.7	12.2	13.3	12.2
	20	0.2	0.1	12.9	30.1	9.1	9.6	4.4	4.9	6.8	7.0	8.3	4.4	4.5	6.3	10.1	14.7	11.1	12.1
	50	1.4	0.2	14.3	32.5	13.5	13.7	3.8	4.9	6.6	7.5	10.0	7.2	4.5	5.9	11.9	12.7	13.1	9.7
	100	3.7	0.5	16.8	35.9	19.8	20.2	4.9	5.0	7.2	7.7	10.6	7.4	9.6	60.7	14.6	12.7	13.5	12.3
	200	nan	nan	21.6	45.1	40.5	34.4	4.0	5.6	7.2	9.3	12.1	6.9	12.5	13.0	14.4	14.2	14.1	42.3
	400	nan	nan	31.2	51.4	57.1	57.3	4.9	6.8	7.6	9.0	12.7	7.1	5.8	11.4	14.3	10.7	15.1	10.4
180	3	0.0	0.2	19.8	61.8	7.2	6.6	3.9	5.5	5.7	7.2	7.2	3.5	6.1	31.2	13.3	12.9	12.8	10.9
	10	0.1	0.2	17.2	59.0	7.3	7.7	3.9	4.8	5.2	6.9	7.4	5.1	3.1	5.3	9.8	11.0	12.3	8.4
	20	0.2	0.1	20.6	232.9	9.8	9.6	4.3	5.0	5.4	6.7	8.1	3.8	5.2	4.6	14.1	16.2	10.6	9.7
	50	1.3	0.2	21.0	61.3	14.1	13.7	4.1	5.1	5.5	7.0	9.5	36.8	63.6	7.5	12.7	12.5	69.7	9.4
	100	3.1	0.4	21.8	86.1	19.8	20.2	3.6	5.2	6.2	7.5	10.8	5.8	7.6	5.8	15.0	13.3	13.0	11.8
	200	nan	nan	31.5	70.4	32.4	32.3	4.1	5.9	5.9	7.4	10.6	5.7	3.7	40.6	43.9	34.8	12.3	11.3
	400	nan	nan	37.4	79.9	367.0	57.0	5.2	7.3	6.8	7.4	10.9	8.5	3.9	7.0	12.2	11.3	11.8	10.4
360	3	0.1	0.1	30.8	110.1	7.4	7.7	4.0	3.0	9.2	6.4	8.0	34.8	6.6	7.6	11.9	19.5	16.0	13.0
	10	0.1	0.1	26.7	111.7	7.5	7.9	4.1	2.9	9.4	6.7	9.2	6.3	3.8	5.9	12.5	11.4	14.2	11.6
	20	0.4	0.1	29.3	114.0	9.6	10.4	3.8	3.0	9.1	6.6	9.2	4.2	3.0	5.2	11.2	10.1	101.4	12.6
	50	1.9	0.1	32.5	117.2	13.5	14.0	3.8	3.5	9.8	7.0	10.6	6.2	3.4	5.6	76.3	12.7	13.1	10.2
	100	3.9	0.3	211.0	115.1	20.7	20.0	3.6	3.3	10.5	7.3	10.6	4.5	5.8	6.7	14.6	12.8	14.0	13.9
	200	nan	nan	39.3	120.4	32.0	32.4	3.8	5.4	10.1	6.9	10.0	5.2	3.2	6.6	26.0	11.4	12.5	10.3
	400	nan	nan	44.9	145.5	56.1	55.6	4.5	5.2	12.1	6.6	10.6	11.7	3.7	9.4	14.3	10.9	11.5	9.3

C. Number of parameters

TABLE XXI: Univariate without Time: Number of Parameters

n	sequence ahead	length	Linear		LSTM		GRU		MLP			LinearNet				TSMixer		TSMixer	
			Regression	Ridge	Encoder	Shared	Last	Full	Layer	Layer	Layer	LinearNet	conv	CLinear	NLinear	DLinear	1	2	no relu
1	3	4	4	30651	14	1061	1081	4	101	521	62	18	6	6	10	42	76	76	76
	10	11	11	30651	14	1061	1151	11	241	661	104	32	13	13	24	161	307	307	307
	20	21	21	30651	14	1061	1251	21	441	861	164	52	23	23	44	501	977	977	977
	50	51	51	30651	14	1061	1551	51	1041	1461	344	112	53	53	104	2721	5387	5387	5387
	100	101	101	30651	14	1061	2051	101	2041	2461	644	212	103	103	204	10421	20737	20737	20737
	200	nan	nan	30651	14	1061	3051	201	4041	4461	1244	412	203	203	404	40821	81437	81437	81437
	400	nan	nan	30651	14	1061	5051	401	8041	8461	2444	812	403	403	804	161621	322837	322837	322837
6	3	24	24	30651	14	1116	1236	24	1206	15726	157	93	26	26	50	62	96	96	96
	10	66	66	30651	14	1116	1656	66	2046	16566	234	142	68	68	134	216	362	362	362
	20	126	126	30651	14	1116	2256	126	3246	17766	344	212	128	128	254	606	1082	1082	1082
	50	306	306	30651	14	1116	4056	306	6846	21366	674	422	308	308	614	2976	5642	5642	5642
	100	606	606	30651	14	1116	7056	606	12846	27366	1224	772	608	608	1214	10926	21242	21242	21242
	200	nan	nan	30651	14	1116	13056	1206	24846	39366	2324	1472	1208	1208	2414	41826	82442	82442	82442
	400	nan	nan	30651	14	1116	25056	2406	48846	63366	4524	2872	2408	2408	4814	163626	324842	324842	324842
15	3	60	60	30651	14	1215	1515	60	5715	96015	454	354	62	62	122	98	132	132	132
	10	165	165	30651	14	1215	2565	165	7815	98115	594	466	167	167	332	315	461	461	461
	20	315	315	30651	14	1215	4065	315	10815	101115	794	626	317	317	632	795	1271	1271	1271
	50	765	765	30651	14	1215	8565	765	19815	110115	1394	1106	767	767	1532	3435	6101	6101	6101
	100	1515	1515	30651	14	1215	16065	1515	34815	125115	2394	1906	1517	1517	3032	11835	22151	22151	22151
	200	nan	nan	30651	14	1215	31065	3015	64815	155115	4394	3506	3017	3017	6032	43635	84251	84251	84251
	400	nan	nan	30651	14	1215	61065	6015	124815	215115	8394	6706	6017	6017	12032	167235	328451	328451	328451
30	3	120	120	30651	14	1380	1980	120	20430	381030	1309	1149	122	122	242	158	192	192	192
	10	330	330	30651	14	1380	4080	330	24630	385230	1554	1366	332	332	662	480	626	626	626
	20	630	630	30651	14	1380	7080	630	30630	391230	1904	1676	632	632	1262	1110	1586	1586	1586
	50	1530	1530	30651	14	1380	16080	1530	48630	409230	2954	2606	1532	1532	3062	4200	6866	6866	6866
	100	3030	3030	30651	14	1380	31080	3030	78630	439230	4704	4156	3032	3032	6062	13350	23666	23666	23666
	200	nan	nan	30651	14	1380	61080	6030	138630	499230	8204	7256	6032	6032	12062	46650	87266	87266	87266
	400	nan	nan	30651	14	1380	121080	12030	258630	619230	15204	13456	12032	12032	24062	173250	334466	334466	334466
90	3	360	360	30651	14	2040	3840	360	169290	3411090	9229	8829	362	362	722	398	432	432	432
	10	990	990	30651	14	2040	10140	990	181890	3423690	9894	9466	992	992	1982	1140	1286	1286	1286
	20	1890	1890	30651	14	2040	19140	1890	199890	3441690	10844	10376	1892	1892	3782	2370	2846	2846	2846
	50	4590	4590	30651	14	2040	46140	4590	253890	3495690	13694	13106	4592	4592	9182	7260	9926	9926	9926
	100	9090	9090	30651	14	2040	91140	9090	343890	3585690	18444	17656	9092	9092	18182	19410	29726	29726	29726
	200	nan	nan	30651	14	2040	181140	18090	523890	3765690	27944	26756	18092	18092	36182	58710	99326	99326	99326
	400	nan	nan	30651	14	2040	361140	36090	883890	4125690	46944	44956	36092	36092	72182	197310	358526	358526	358526
180	3	720	720	30651	14	3030	6630	720	662580	13626180	34609	33849	722	722	1442	758	792	792	792
	10	1980	1980	30651	14	3030	19230	1980	687780	13651380	35904	35116	1982	1982	3962	2130	2276	2276	2276
	20	3780	3780	30651	14	3030	37230	3780	723780	13687380	37754	36926	3782	3782	7562	4260	4736	4736	4736
	50	9180	9180	30651	14	3030	91230	9180	831780	13795380	43304	42356	9182	9182	18362	11850	14516	14516	14516
	100	18180	18180	30651	14	3030	181230	18180	1011780	13975380	52554	51406	18182	18182	36362	28500	38816	38816	38816
	200	nan	nan	30651	14	3030	361230	36180	1371780	14335380	71054	69506	36182	36182	72362	76800	117416	117416	117416
	400	nan	nan	30651	14	3030	721230	72180	2091780	15055380	108054	105706	72182	72182	144362	233400	394616	394616	394616
360	3	1440	1440	30651	14	5010	12210	1440	2621160	54468360	133969	132489	1442	1442	2882	1478	1512	1512	1512
	10	3960	3960	30651	14	5010	37410	3960	2671560	54518760	136524	135016	3962	3962	7922	4110	4256	4256	4256
	20	7560	7560	30651	14	5010	73410	7560	2743560	54590760	140174	138626	7562	7562	15122	8040	8516	8516	8516
	50	18360	18360	30651	14	5010	181410	18360	2959560	54806760	151124	149456	18362	18362	36722	21030	23696	23696	23696
	100	36360	36360	30651	14	5010	361410	36360	3319560	55166760	169374	167506	36362	36362	72722	46680	56996	56996	56996
	200	nan	nan	30651	14	5010	721410	72360	4039560	55886760	205874	203606	72362	72362	144722	112980	153596	153596	153596
	400	nan	nan	30651	14	5010	1441410	144360	5479560	57326760	278874	275806	144362	144362	288722	305580	466796	466796	466796

TABLE XXII: Univariate with Time: Number of Parameters

n	sequence ahead	length	Linear		LSTM		GRU	GRU	MLP	MLP	MLP	LinearNet				TSMixer	TSMixer	TSMixer	TSMixer
			Regression	Ridge	Encoder	Shared Encoder	Last Sequence	Full Sequence	Layer linear	Layer linear	Layer ReLU	LinearNet no conv	CLinear	NLinear	DLinear	1	2	1	2
1	3	4	4	30651	14	1061	1081	4	101	521	62	18	6	6	10	42	76	76	76
	10	11	11	30651	14	1061	1151	11	241	661	104	32	13	13	24	161	307	307	307
	20	21	21	30651	14	1061	1251	21	441	861	164	52	23	23	44	501	977	977	977
	50	51	51	30651	14	1061	1551	51	1041	1461	344	112	53	53	104	2721	5387	5387	5387
	100	101	101	30651	14	1061	2051	101	2041	2461	644	212	103	103	204	10421	20737	20737	20737
	200	nan	nan	30651	14	1061	3051	201	4041	4461	1244	412	203	203	404	40821	81437	81437	81437
400	nan	nan	30651	14	1061	5051	401	8041	8461	2444	812	403	403	804	161621	322837	322837	322837	
6	3	24	24	30651	14	1116	1236	24	1206	15726	157	93	26	26	50	62	96	96	96
	10	66	66	30651	14	1116	1656	66	2046	16566	234	142	68	68	134	216	362	362	362
	20	126	126	30651	14	1116	2256	126	3246	17766	344	212	128	128	254	606	1082	1082	1082
	50	306	306	30651	14	1116	4056	306	6846	21366	674	422	308	308	614	2976	5642	5642	5642
	100	606	606	30651	14	1116	7056	606	12846	27366	1224	772	608	608	1214	10926	21242	21242	21242
	200	nan	nan	30651	14	1116	13056	1206	24846	39366	2324	1472	1208	1208	2414	41826	82442	82442	82442
400	nan	nan	30651	14	1116	25056	2406	48846	63366	4524	2872	2408	2408	4814	163626	324842	324842	324842	
15	3	60	60	30651	14	1215	1515	60	5715	96015	454	354	62	62	122	98	132	132	132
	10	165	165	30651	14	1215	2565	165	7815	98115	594	466	167	167	332	315	461	461	461
	20	315	315	30651	14	1215	4065	315	10815	101115	794	626	317	317	632	795	1271	1271	1271
	50	765	765	30651	14	1215	8565	765	19815	110115	1394	1106	767	767	1532	3435	6101	6101	6101
	100	1515	1515	30651	14	1215	16065	1515	34815	125115	2394	1906	1517	1517	3032	11835	22151	22151	22151
	200	nan	nan	30651	14	1215	31065	3015	64815	155115	4394	3506	3017	3017	6032	43635	84251	84251	84251
400	nan	nan	30651	14	1215	61065	6015	124815	215115	8394	6706	6017	6017	12032	167235	328451	328451	328451	
30	3	120	120	30651	14	1380	1980	120	20430	381030	1309	1149	122	122	242	158	192	192	192
	10	330	330	30651	14	1380	4080	330	24630	385230	1554	1366	332	332	662	480	626	626	626
	20	630	630	30651	14	1380	7080	630	30630	391230	1904	1676	632	632	1262	1110	1586	1586	1586
	50	1530	1530	30651	14	1380	16080	1530	48630	409230	2954	2606	1532	1532	3062	4200	6866	6866	6866
	100	3030	3030	30651	14	1380	31080	3030	78630	439230	4704	4156	3032	3032	6062	13350	23666	23666	23666
	200	nan	nan	30651	14	1380	61080	6030	138630	499230	8204	7256	6032	6032	12062	46650	87266	87266	87266
400	nan	nan	30651	14	1380	121080	12030	258630	619230	15204	13456	12032	12032	24062	173250	334466	334466	334466	
90	3	360	360	30651	14	2040	3840	360	169290	3411090	9229	8829	362	362	722	398	432	432	432
	10	990	990	30651	14	2040	10140	990	181890	3423690	9894	9466	992	992	1982	1140	1286	1286	1286
	20	1890	1890	30651	14	2040	19140	1890	199890	3441690	10844	10376	1892	1892	3782	2370	2846	2846	2846
	50	4590	4590	30651	14	2040	46140	4590	253890	3495690	13694	13106	4592	4592	9182	7260	9926	9926	9926
	100	9090	9090	30651	14	2040	91140	9090	343890	3585690	18444	17656	9092	9092	18182	19410	29726	29726	29726
	200	nan	nan	30651	14	2040	181140	18090	523890	3765690	27944	26756	18092	18092	36182	58710	99326	99326	99326
400	nan	nan	30651	14	2040	361140	36090	883890	4125690	46944	44956	36092	36092	72182	197310	358526	358526	358526	
180	3	720	720	30651	14	3030	6630	720	662580	13626180	34609	33849	722	722	1442	758	792	792	792
	10	1980	1980	30651	14	3030	19230	1980	687780	13651380	35904	35116	1982	1982	3962	2130	2276	2276	2276
	20	3780	3780	30651	14	3030	37230	3780	723780	13687380	37754	36926	3782	3782	7562	4260	4736	4736	4736
	50	9180	9180	30651	14	3030	91230	9180	831780	13795380	43304	42356	9182	9182	18362	11850	14516	14516	14516
	100	18180	18180	30651	14	3030	181230	18180	1011780	13975380	52554	51406	18182	18182	36362	28500	38816	38816	38816
	200	nan	nan	30651	14	3030	361230	36180	1371780	14335380	71054	69506	36182	36182	72362	76800	117416	117416	117416
400	nan	nan	30651	14	3030	721230	72180	2091780	15055380	108054	105706	72182	72182	144362	233400	394616	394616	394616	
360	3	1440	1440	30651	14	5010	12210	1440	2621160	54468360	133969	132489	1442	1442	2882	1478	1512	1512	1512
	10	3960	3960	30651	14	5010	37410	3960	2671560	54518760	136524	135016	3962	3962	7922	4110	4256	4256	4256
	20	7560	7560	30651	14	5010	73410	7560	2743560	54590760	140174	138626	7562	7562	15122	8040	8516	8516	8516
	50	18360	18360	30651	14	5010	181410	18360	2959560	54806760	151124	149456	18362	18362	36722	21030	23696	23696	23696
	100	36360	36360	30651	14	5010	361410	36360	3319560	55166760	169374	167506	36362	36362	72722	46680	56996	56996	56996
	200	nan	nan	30651	14	5010	721410	72360	4039560	55886760	205874	203606	72362	72362	144722	112980	153596	153596	153596
400	nan	nan	30651	14	5010	1441410	144360	5479560	57326760	278874	275806	144362	144362	288722	305580	466796	466796	466796	

TABLE XXIII: Multivariate without Time: Number of Parameters

n	sequence ahead length	Linear Regression	Ridge	LSTM		GRU Last	GRU Full	MLP 1 Layer linear	MLP 2 Layer linear	MLP 3 Layer ReLU	LinearNet no conv	CLinear	NLinear	DLinear	TSMixer 1	TSMixer 2	TSMixer no relu 1	TSMixer no relu 2
				LSTM Encoder Decoder	Shared Encoder Decoder													
1	3	22	22	31851	428	1241	1261	22	461	881	662	114	12	16	162	298	298	298
	10	71	71	31851	428	1241	1331	71	1441	1861	704	128	19	30	365	697	697	697
	20	141	141	31851	428	1241	1431	141	2841	3261	764	148	29	50	825	1607	1607	1607
	50	351	351	31851	428	1241	1731	351	7041	7461	944	208	59	110	3405	6737	6737	6737
	100	701	701	31851	428	1241	2231	701	14041	14461	1244	308	109	210	11705	23287	23287	23287
	200	nan	nan	31851	428	1241	3231	1401	28041	28461	1844	508	209	410	43305	86387	86387	86387
	400	nan	nan	31851	428	1241	5231	2801	56041	56461	3044	908	409	810	166505	332587	332587	332587
6	3	132	132	31851	428	1296	1416	132	3366	17886	757	189	32	56	182	318	318	318
	10	426	426	31851	428	1296	1836	426	9246	23766	834	238	74	140	420	752	752	752
	20	846	846	31851	428	1296	2436	846	17646	32166	944	308	134	260	930	1712	1712	1712
	50	2106	2106	31851	428	1296	4236	2106	42846	57366	1274	518	314	620	3660	6992	6992	6992
	100	4206	4206	31851	428	1296	7236	4206	84846	99366	1824	868	614	1220	12210	23792	23792	23792
	200	nan	nan	31851	428	1296	13236	8406	168846	183366	2924	1568	1214	2420	44310	87392	87392	87392
	400	nan	nan	31851	428	1296	25236	16806	336846	351366	5124	2968	2414	4820	168510	334592	334592	334592
15	3	330	330	31851	428	1395	1695	330	11115	101415	1054	450	68	128	218	354	354	354
	10	1065	1065	31851	428	1395	2745	1065	25815	116115	1194	562	173	338	519	851	851	851
	20	2115	2115	31851	428	1395	4245	2115	46815	137115	1394	722	323	638	1119	1901	1901	1901
	50	5265	5265	31851	428	1395	8745	5265	109815	200115	1994	1202	773	1538	4119	7451	7451	7451
	100	10515	10515	31851	428	1395	16245	10515	214815	305115	2994	2002	1523	3038	13119	24701	24701	24701
	200	nan	nan	31851	428	1395	31245	21015	424815	515115	4994	3602	3023	6038	46119	89201	89201	89201
	400	nan	nan	31851	428	1395	61245	42015	844815	935115	8994	6802	6023	12038	172119	338201	338201	338201
30	3	660	660	31851	428	1560	2160	660	31230	391830	1909	1245	128	248	278	414	414	414
	10	2130	2130	31851	428	1560	4260	2130	60630	421230	2154	1462	338	338	668	1016	1016	1016
	20	4230	4230	31851	428	1560	7260	4230	102630	463230	2504	1772	638	638	1268	1434	2216	2216
	50	10530	10530	31851	428	1560	16260	10530	228630	589230	3554	2702	1538	1538	3068	4884	8216	8216
	100	21030	21030	31851	428	1560	31260	21030	438630	799230	5304	4252	3038	3038	6068	14634	26216	26216
	200	nan	nan	31851	428	1560	61260	42030	858630	1219230	8804	7352	6038	6038	12068	49134	92216	92216
	400	nan	nan	31851	428	1560	121260	84030	1698630	2059230	15804	13552	12038	24068	178134	344216	344216	344216
90	3	1980	1980	31851	428	2220	4020	1980	201690	3443490	9829	8925	368	368	728	518	654	654
	10	6390	6390	31851	428	2220	10320	6390	289890	3531690	10494	9562	998	998	1988	1344	1676	1676
	20	12690	12690	31851	428	2220	19320	12690	415890	3657690	11444	10472	1898	1898	3788	2694	3476	3476
	50	31590	31590	31851	428	2220	46320	31590	793890	4035690	14294	13202	4598	4598	9188	7944	11276	11276
	100	63090	63090	31851	428	2220	91320	63090	1423890	4665690	19044	17752	9098	9098	18188	20694	32276	32276
	200	nan	nan	31851	428	2220	181320	126090	2683890	5925690	28544	26852	18098	18098	36188	61194	104276	104276
	400	nan	nan	31851	428	2220	361320	252090	5203890	8445690	47544	45052	36098	36098	72188	202194	368276	368276
180	3	3960	3960	31851	428	3210	6810	3960	727380	13690980	35209	33945	728	728	1448	878	1014	1014
	10	12780	12780	31851	428	3210	19410	12780	903780	13867380	36504	35212	1988	1988	3968	2334	2666	2666
	20	25380	25380	31851	428	3210	37410	25380	1155780	14119380	38354	37022	3788	3788	7568	4584	5366	5366
	50	63180	63180	31851	428	3210	91410	63180	1911780	14875380	43904	42452	9188	9188	18368	12534	15866	15866
	100	126180	126180	31851	428	3210	181410	126180	3171780	16135380	53154	51502	18188	18188	36368	29784	41366	41366
	200	nan	nan	31851	428	3210	361410	252180	5691780	18655380	71654	69602	36188	36188	72368	79284	122366	122366
	400	nan	nan	31851	428	3210	721410	504180	10731780	23695380	108654	105802	72188	72188	144368	238284	404366	404366
360	3	7920	7920	31851	428	5190	12390	7920	2750760	54597960	134569	132585	1448	1448	2888	1598	1734	1734
	10	25560	25560	31851	428	5190	37590	25560	3103560	54950760	137124	135112	3968	3968	7928	4314	4646	4646
	20	50760	50760	31851	428	5190	73590	50760	3607560	55454760	140774	138722	7568	7568	15128	8364	9146	9146
	50	126360	126360	31851	428	5190	181590	126360	5119560	56966760	151724	149552	18368	18368	36728	21714	25046	25046
	100	252360	252360	31851	428	5190	361590	252360	7639560	59486760	169974	167602	36368	36368	72728	47964	59546	59546
	200	nan	nan	31851	428	5190	721590	504360	12679560	64526760	206474	203702	72368	72368	144728	115464	158546	158546
	400	nan	nan	31851	428	5190	1441590	1008360	22759560	74606760	279474	275902	144368	144368	288728	310464	476546	476546

TABLE XXIV: Multivariate with Time: Number of Parameters

n	sequence ahead length	Linear Regression	Ridge	LSTM		GRU Last	GRU Full	MLP 1 Layer linear	MLP 2 Layer linear	MLP 3 Layer ReLU	LinearNet				TSMixer 1	TSMixer 2	TSMixer no relu 1	TSMixer no relu 2	
				LSTM Encoder Decoder	Shared Encoder Decoder						LinearNet no conv	CLinear	NLinear	DLinear					
1	3	28	28	32251	694	1301	1321	28	581	1001	1066	170	14	14	18	202	372	372	372
	10	91	91	32251	694	1301	1391	91	1841	2261	1108	184	21	21	32	433	827	827	827
	20	181	181	32251	694	1301	1491	181	3641	4061	1168	204	31	31	52	933	1817	1817	1817
	50	451	451	32251	694	1301	1791	451	9041	9461	1348	264	61	61	112	3633	7187	7187	7187
	100	901	901	32251	694	1301	2291	901	18041	18461	1648	364	111	111	212	12133	24137	24137	24137
	200	nan	nan	32251	694	1301	3291	1801	36041	36461	2248	564	211	211	412	44133	88037	88037	88037
	400	nan	nan	32251	694	1301	5291	3601	72041	72461	3448	964	411	411	812	168133	335837	335837	335837
6	3	168	168	32251	694	1356	1476	168	4086	18606	1161	245	34	34	58	222	392	392	392
	10	546	546	32251	694	1356	1896	546	11646	26166	1238	294	76	76	142	488	882	882	882
	20	1086	1086	32251	694	1356	2496	1086	22446	36966	1348	364	136	136	262	1038	1922	1922	1922
	50	2706	2706	32251	694	1356	4296	2706	54846	69366	1678	574	316	316	622	3888	7442	7442	7442
	100	5406	5406	32251	694	1356	7296	5406	108846	123366	2228	924	616	616	1222	12638	24642	24642	24642
	200	nan	nan	32251	694	1356	13296	10806	216846	231366	3328	1624	1216	1216	2422	45138	89042	89042	89042
	400	nan	nan	32251	694	1356	25296	21606	432846	447366	5528	3024	2416	2416	4822	170138	337842	337842	337842
15	3	420	420	32251	694	1455	1755	420	12915	103215	1458	506	70	70	130	258	428	428	428
	10	1365	1365	32251	694	1455	2805	1365	31815	122115	1598	618	175	175	340	587	981	981	981
	20	2715	2715	32251	694	1455	4305	2715	58815	149115	1798	778	325	325	640	1227	2111	2111	2111
	50	6765	6765	32251	694	1455	8805	6765	139815	230115	2398	1258	775	775	1540	4347	7901	7901	7901
	100	13515	13515	32251	694	1455	16305	13515	274815	365115	3398	2058	1525	1525	3040	13547	25551	25551	25551
	200	nan	nan	32251	694	1455	31305	27015	544815	635115	5398	3658	3025	3025	6040	46947	90851	90851	90851
	400	nan	nan	32251	694	1455	61305	54015	1084815	1175115	9398	6858	6025	6025	12040	173747	341451	341451	341451
30	3	840	840	32251	694	1620	2220	840	34830	395430	2313	1301	130	130	250	318	488	488	488
	10	2730	2730	32251	694	1620	4320	2730	72630	433230	2558	1518	340	340	670	752	1146	1146	1146
	20	5430	5430	32251	694	1620	7320	5430	126630	487230	2908	1828	640	640	1270	1542	2426	2426	2426
	50	13530	13530	32251	694	1620	16320	13530	288630	649230	3958	2758	1540	1540	3070	5112	8666	8666	8666
	100	27030	27030	32251	694	1620	31320	27030	558630	919230	5708	4308	3040	3040	6070	15062	27066	27066	27066
	200	nan	nan	32251	694	1620	61320	54030	1098630	1459230	9208	7408	6040	6040	12070	49962	93866	93866	93866
	400	nan	nan	32251	694	1620	121320	108030	2178630	2539230	16208	13608	12040	12040	24070	179762	347466	347466	347466
90	3	2520	2520	32251	694	2280	4080	2520	212490	3454290	10233	8981	370	370	730	558	728	728	728
	10	8190	8190	32251	694	2280	10380	8190	325890	3567690	10898	9618	1000	1000	1990	1412	1806	1806	1806
	20	16290	16290	32251	694	2280	19380	16290	487890	3729690	11848	10528	1900	1900	3790	2802	3686	3686	3686
	50	40590	40590	32251	694	2280	46380	40590	973890	4215690	14698	13258	4600	4600	9190	8172	11726	11726	11726
	100	81090	81090	32251	694	2280	91380	81090	1783890	5025690	19448	17808	9100	9100	18190	21122	33126	33126	33126
	200	nan	nan	32251	694	2280	181380	162090	3403890	6645690	28948	26908	18100	18100	36190	62022	105926	105926	105926
	400	nan	nan	32251	694	2280	361380	324090	6643890	9885690	47948	45108	36100	36100	72190	203822	371526	371526	371526
180	3	5040	5040	32251	694	3270	6870	5040	748980	13712580	35613	34001	730	730	1450	918	1088	1088	1088
	10	16380	16380	32251	694	3270	19470	16380	975780	13939380	36908	35268	1990	1990	3970	2402	2796	2796	2796
	20	32580	32580	32251	694	3270	37470	32580	1299780	14263380	38758	37078	3790	3790	7570	4692	5576	5576	5576
	50	81180	81180	32251	694	3270	91470	81180	2271780	15235380	44308	42508	9190	9190	18370	12762	16316	16316	16316
	100	162180	162180	32251	694	3270	181470	162180	3891780	16855380	53558	51558	18190	18190	36370	30212	42216	42216	42216
	200	nan	nan	32251	694	3270	361470	324180	7131780	20095380	72058	69658	36190	36190	72370	80112	124016	124016	124016
	400	nan	nan	32251	694	3270	721470	648180	13611780	26575380	109058	105858	72190	72190	144370	239912	407616	407616	407616
360	3	10080	10080	32251	694	5250	12450	10080	2793960	54641160	134973	132641	1450	1450	2890	1638	1808	1808	1808
	10	32760	32760	32251	694	5250	37650	32760	3247560	55094760	137528	135168	3970	3970	7930	4382	4776	4776	4776
	20	65160	65160	32251	694	5250	73650	65160	3895560	55742760	141178	138778	7570	7570	15130	8472	9356	9356	9356
	50	162360	162360	32251	694	5250	181650	162360	5839560	57686760	152128	149608	18370	18370	36730	21942	25496	25496	25496
	100	324360	324360	32251	694	5250	361650	324360	9079560	60926760	170378	167658	36370	36370	72730	48392	60396	60396	60396
	200	nan	nan	32251	694	5250	721650	648360	15559560	67406760	206878	203758	72370	72370	144730	116292	160196	160196	160196
	400	nan	nan	32251	694	5250	1441650	1296360	28519560	80366760	279878	275958	144370	144370	288730	312092	479796	479796	479796

APPENDIX C
BTC VOLUMES: WEIGHTS

BTC Volume: Linear Regression Weights Heatmap

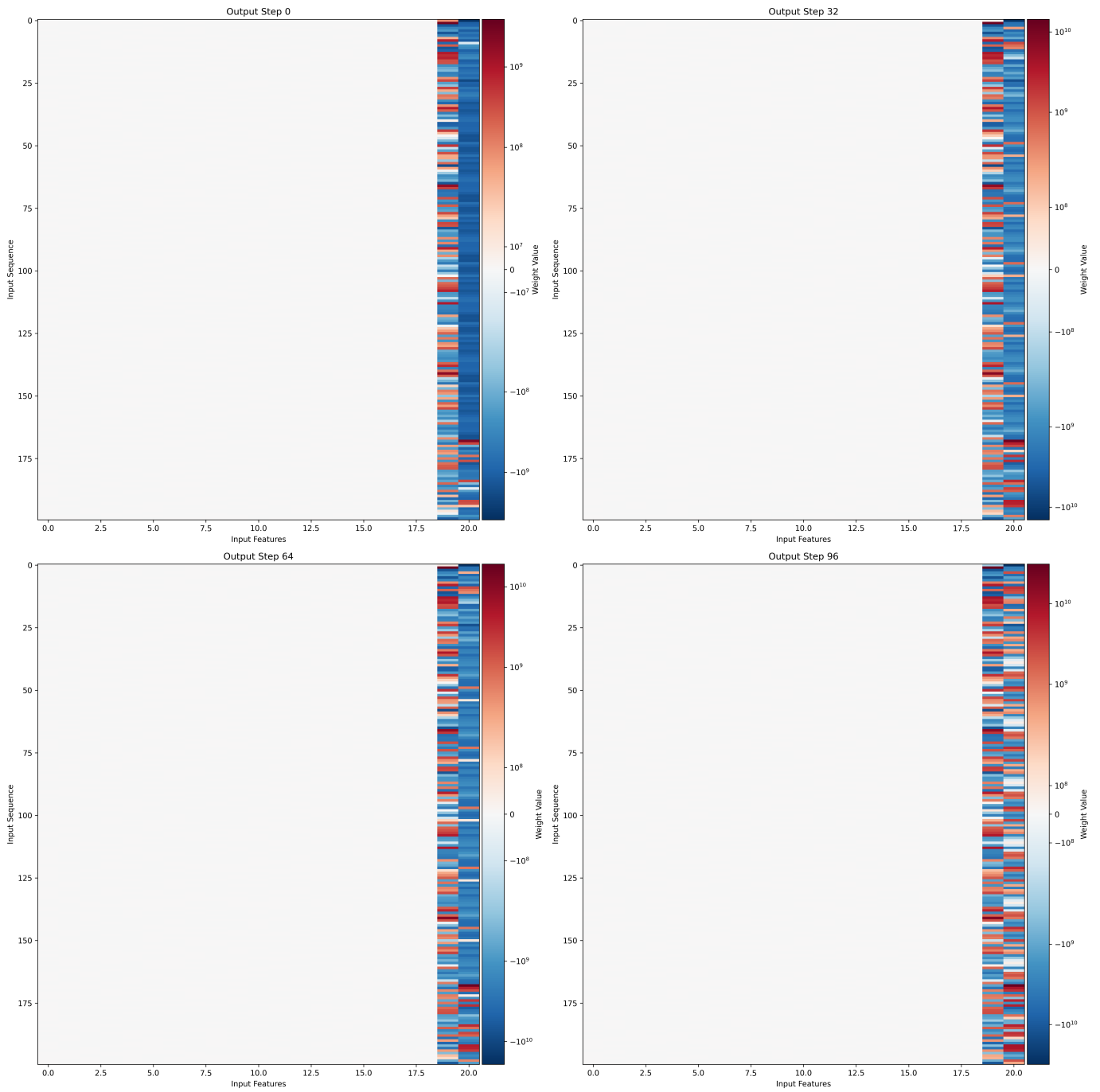


Fig. 6: Linear Regression weights for BTC Volume task

BTC Volume: TLN Weights Heatmap

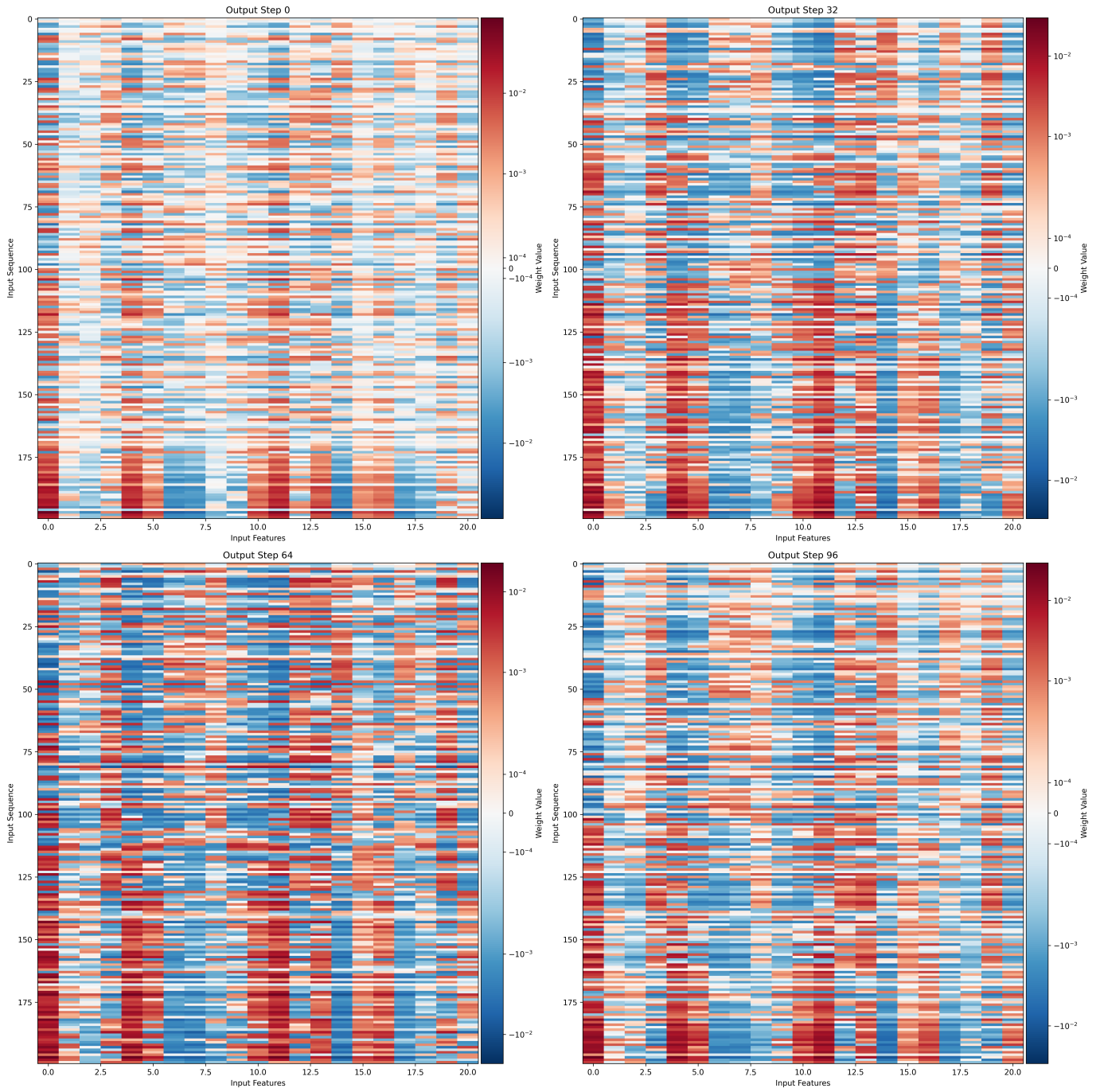


Fig. 7: TLN weights for BTC Volume task

BTC Volume: Ridge with alpha 1.0 Weights Heatmap

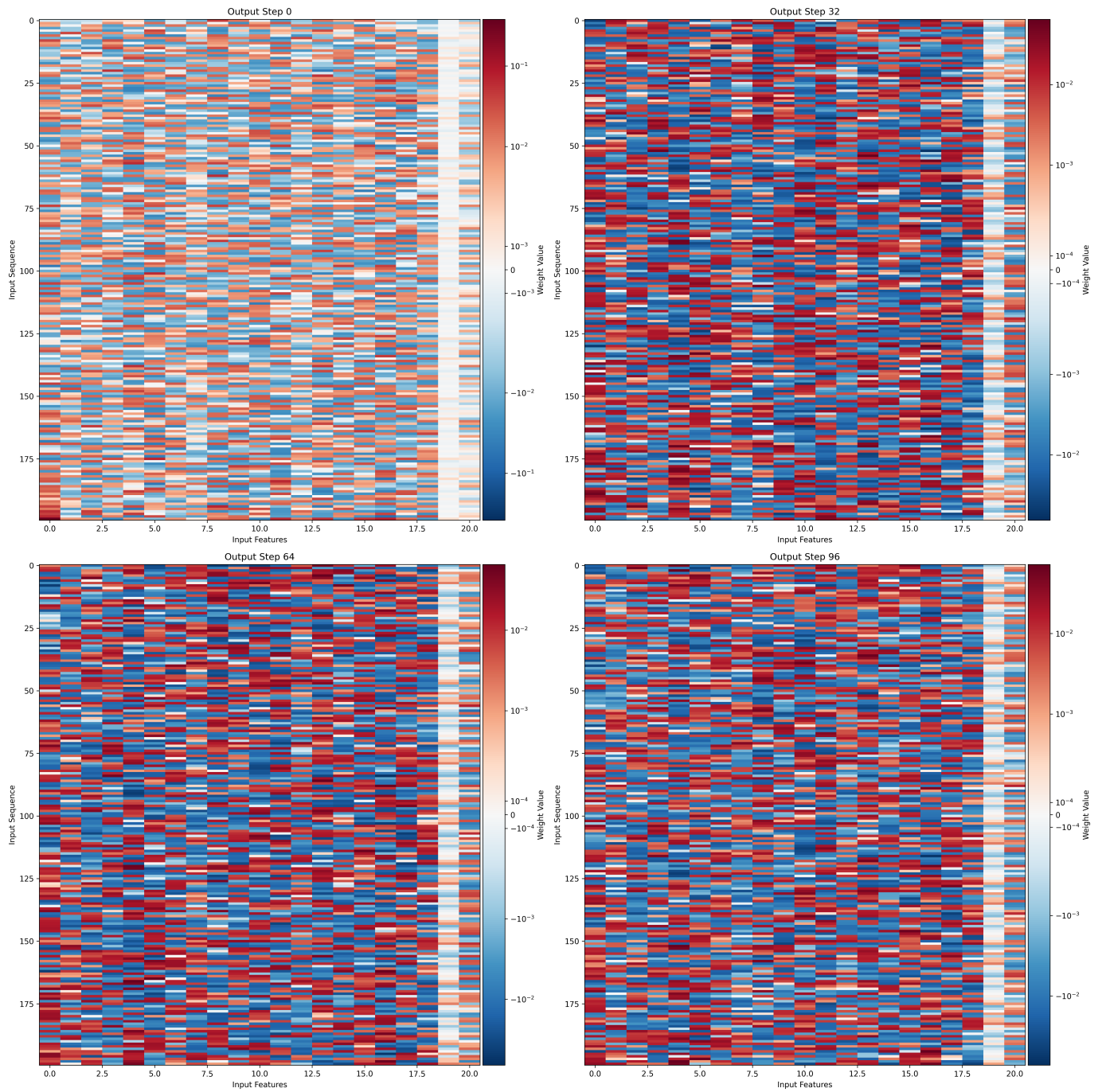


Fig. 8: Ridge Regression weights for BTC Volume task (alpha = 1.0)

BTC Volume: Lasso with alpha 0.001 Weights Heatmap

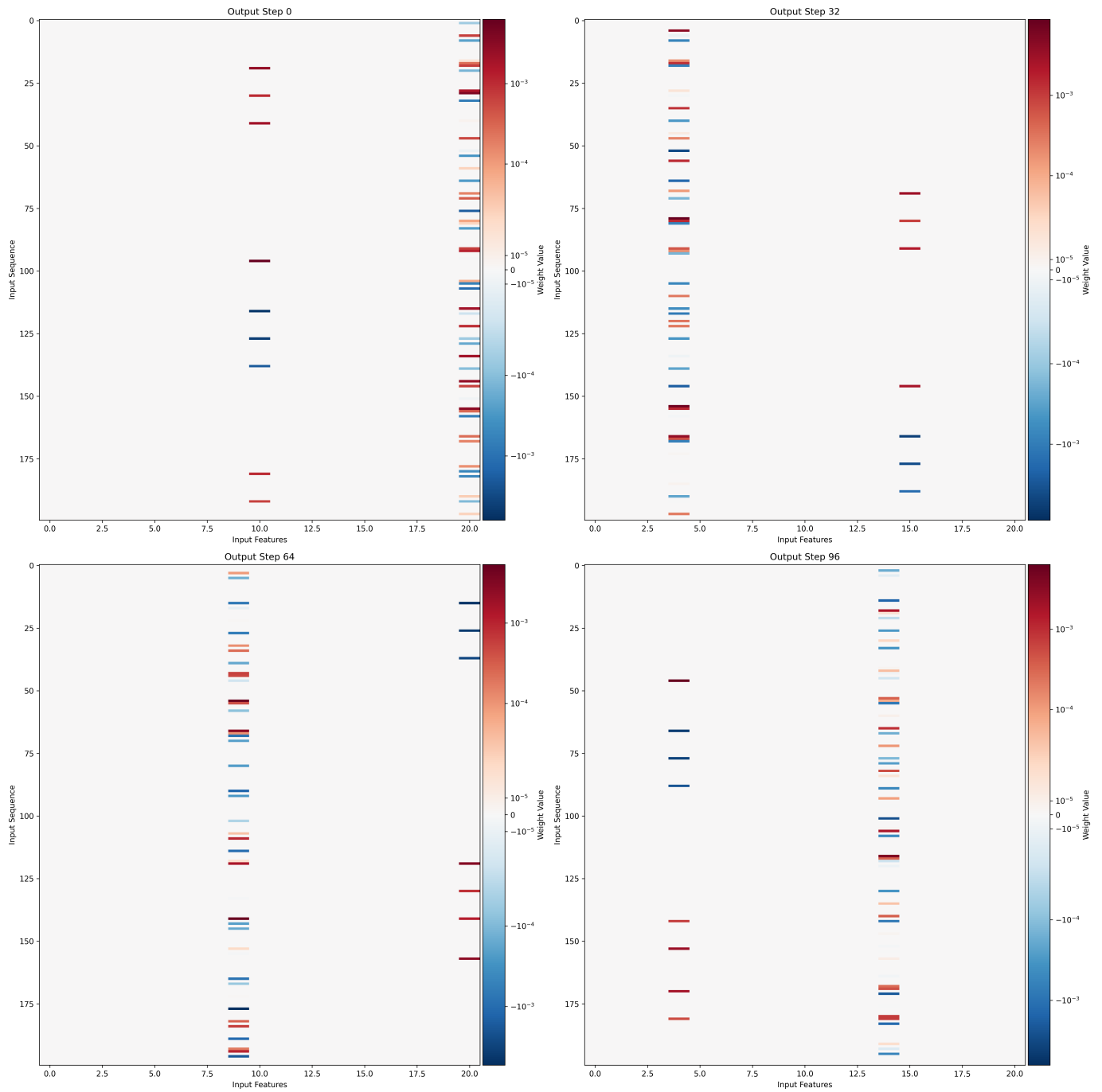


Fig. 9: Lasso Regression weights for BTC Volume task (alpha = 0.01)

BTC Volume: ElasticNet with alpha 0.001 Weights Heatmap

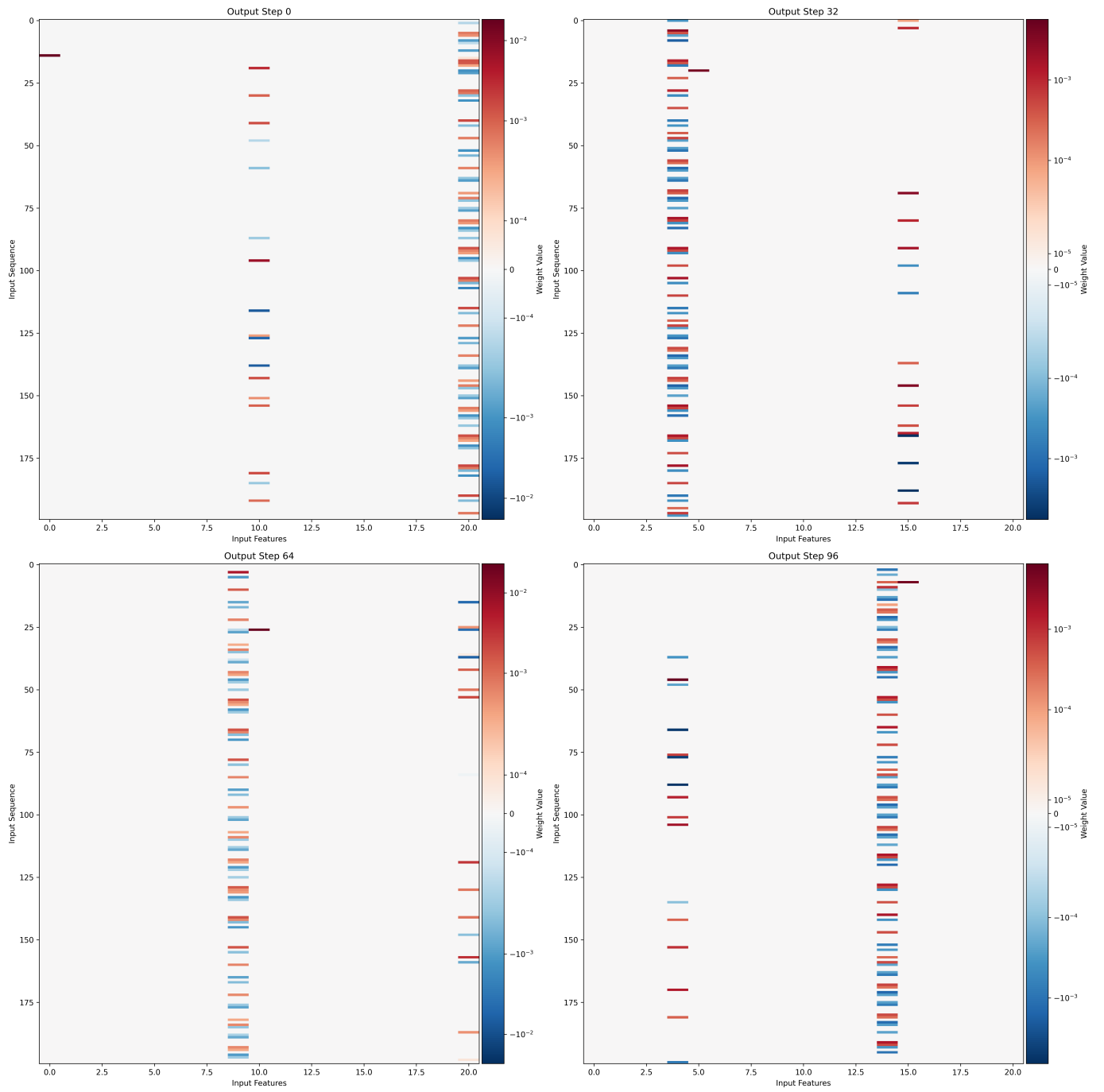


Fig. 10: ElasticNet Regression weights for BTC Volume task (alpha = 0.01)