

## Hyperparameter Search Results

Table 1: Optimal Hyperparameters for Dataset **QLD** (Dim: Hidden Dimensions; F. Size: Filter Size; Filters: Number of Filters, LR: Learning Rate; N. Layers: Number of Layers.)

Forecast Horizon	Model	Dim.	F. Size	Filters	LR	N. Layers
24	CNN-LSTM	128	7	64	0.001	-
	DLinear	-	-	-	0.001	-
	LSTM	256	-	-	0.001	1
	Mamba	64	-	-	0.005	-
	TimeMixer	32	-	-	0.001	2
	TimeXer	32	-	-	0.005	2
	TimesNet	256	-	-	0.001	2
	Transformer	32	-	-	0.001	2
	iTransformer	32	-	-	0.005	2
	CNN-LSTM	128	7	32	0.005	-
48	DLinear	-	-	-	0.001	-
	LSTM	64	-	-	0.005	2
	Mamba	32	-	-	0.01	-
	TimeMixer	256	-	-	0.001	2
	TimeXer	64	-	-	0.001	2
	TimesNet	32	-	-	0.005	2
	Transformer	256	-	-	0.001	2
	iTransformer	32	-	-	0.001	1

Table 2: Optimal Hyperparameters for Dataset **NSW** (Dim: Hidden Dimensions; F. Size: Filter Size; Filters: Number of Filters, LR: Learning Rate; N. Layers: Number of Layers.)

<b>Forecast Horizon</b>	<b>Model</b>	<b>Dim.</b>	<b>F. Size</b>	<b>Filters</b>	<b>LR</b>	<b>N. Layers</b>
24	CNN-LSTM	128	3	128	0.01	-
	DLinear	-	-	-	0.001	-
	LSTM	32	-	-	0.1	1
	Mamba	32	-	-	0.005	-
	TimeMixer	32	-	-	0.001	2
	TimeXer	64	-	-	0.001	2
	TimesNet	64	-	-	0.001	2
	Transformer	64	-	-	0.001	2
	iTransformer	32	-	-	0.001	1
	CNN-LSTM	512	3	64	0.001	-
48	DLinear	-	-	-	0.001	-
	LSTM	32	-	-	0.001	2
	Mamba	64	-	-	0.001	-
	TimeMixer	128	-	-	0.001	2
	TimeXer	128	-	-	0.001	2
	TimesNet	128	-	-	0.001	1
	Transformer	32	-	-	0.001	2
	iTransformer	32	-	-	0.001	2

Table 3: Optimal Hyperparameters for Dataset **SA** (Dim: Hidden Dimensions; F. Size: Filter Size; Filters: Number of Filters, LR: Learning Rate; N. Layers: Number of Layers.)

<b>Forecast Horizon</b>	<b>Model</b>	<b>Dim.</b>	<b>F. Size</b>	<b>Filters</b>	<b>LR</b>	<b>N. Layers</b>
24	CNN-LSTM	128	5	32	0.001	-
	DLinear	-	-	-	0.001	-
	LSTM	256	-	-	0.005	2
	Mamba	32	-	-	0.01	-
	TimeMixer	32	-	-	0.001	2
	TimeXer	32	-	-	0.001	2
	TimesNet	32	-	-	0.005	1
	Transformer	128	-	-	0.001	2
	iTransformer	64	-	-	0.001	1
	CNN-LSTM	512	3	512	0.01	-
48	DLinear	-	-	-	0.001	-
	LSTM	64	-	-	0.1	1
	Mamba	32	-	-	0.001	-
	TimeMixer	512	-	-	0.001	2
	TimeXer	32	-	-	0.005	2
	TimesNet	512	-	-	0.001	1
	Transformer	64	-	-	0.005	1
	iTransformer	256	-	-	0.001	2

Table 4: Optimal Hyperparameters for Dataset **TAS** (Dim: Hidden Dimensions; F. Size: Filter Size; Filters: Number of Filters, LR: Learning Rate; N. Layers: Number of Layers.)

<b>Forecast Horizon</b>	<b>Model</b>	<b>Dim.</b>	<b>F. Size</b>	<b>Filters</b>	<b>LR</b>	<b>N. Layers</b>
24	CNN-LSTM	128	3	128	0.005	-
	DLinear	-	-	-	0.001	-
	LSTM	128	-	-	0.001	1
	Mamba	32	-	-	0.001	-
	TimeMixer	32	-	-	0.001	2
	TimeXer	64	-	-	0.001	1
	TimesNet	32	-	-	0.001	2
	Transformer	32	-	-	0.005	2
	iTransformer	64	-	-	0.001	2
	CNN-LSTM	128	3	64	0.005	-
48	DLinear	-	-	-	0.001	-
	LSTM	128	-	-	0.01	1
	Mamba	32	-	-	0.001	-
	TimeMixer	32	-	-	0.001	2
	TimeXer	32	-	-	0.001	1
	TimesNet	512	-	-	0.001	1
	Transformer	32	-	-	0.01	2
	iTransformer	64	-	-	0.01	2

Table 5: Optimal Hyperparameters for Dataset **VIC** (Dim: Hidden Dimensions; F. Size: Filter Size; Filters: Number of Filters, LR: Learning Rate; N. Layers: Number of Layers.)

<b>Forecast Horizon</b>	<b>Model</b>	<b>Dim.</b>	<b>F. Size</b>	<b>Filters</b>	<b>LR</b>	<b>N. Layers</b>
24	CNN-LSTM	64	3	128	0.005	-
	DLinear	-	-	-	0.001	-
	LSTM	32	-	-	0.005	2
	Mamba	32	-	-	0.005	-
	TimeMixer	64	-	-	0.001	2
	TimeXer	64	-	-	0.001	1
	TimesNet	32	-	-	0.001	2
	Transformer	512	-	-	0.001	2
	iTransformer	128	-	-	0.001	1
	CNN-LSTM	32	7	128	0.05	-
48	DLinear	-	-	-	0.001	-
	LSTM	128	-	-	0.01	2
	Mamba	128	-	-	0.001	-
	TimeMixer	64	-	-	0.001	2
	TimeXer	32	-	-	0.005	2
	TimesNet	128	-	-	0.001	1
	Transformer	256	-	-	0.001	2
	iTransformer	32	-	-	0.001	2