PROJECT REPORT

Advanced Time Series Forecasting with Deep Learning and Explainability

1. Dataset Description

A synthetic multivariate time series dataset was generated consisting of three interacting signals

(signal1, signal2, signal3). These series contain noise, trends, and nonlinear patterns, simulating real world sensor or financial data.

A 30 step sliding window was used for supervised learning. Data was scaled using MinMaxScaler.

2. Preprocessing Steps

• Normalization (MinMax)

• Windowing (30-step sequential samples)

• Train/test split (80:20)

• Target: signal1

3. Deep Learning Model – Transformer

A Transformer model was implemented with:

• Multi-head Attention (4 heads)

• Layer Normalization

• Dense(64, ReLU)

• GlobalAveragePooling1D

• Dense output layer

Trained for 10 epochs using Adam optimizer and MSE loss.

4. Classical Baseline Models

A. ARIMA(5,1,2)

B. Prophet (Facebook)

5. Performance Comparison

Transformer clearly outperformed ARIMA and Prophet.

Metrics:

Transformer MAE: 0.0507

Transformer RMSE: 0.0613

Transformer MAPE: 21.10%

ARIMA MAE: 0.5207

Prophet MAE: 0.5542

The Transformer is ~10× more accurate.

6. SHAP Explainability

KernelSHAP was applied using flattened sequence windows (30×3 → 90 features).

A wrapper reshaped flattened inputs back to 3D before prediction.

SHAP showed:

- Recent timesteps have higher influence.
- signal1 contributes most strongly.
- Attention focuses on last 5–10 timesteps.

7. Findings

Transformer captures long range temporal dependencies and nonlinear multivariate patterns better than classical methods.

8. Final Model Justification

- Best accuracy
- Handles multivariate sequences
- Self-attention improves long range modeling

• SHAP interpretability validates model behavior

9. Conclusion

Transformers provide superior forecasting performance on complex multivariate time series data.

Explainability confirms meaningful temporal feature importance.