Model Comparison Report

Economic Data

✅ Model 1: XGBoost & LightGBM

- **Model Type:** Tree-based Ensemble Models
- **Dataset:** US Unemployment Rate (UNRATE)
- **Feature Engineering:**
  - Lag features (1, 2, 3 months, etc.)
  - Rolling mean & std (window sizes: 3, 6, 12)
  - Train/test split (no future forecasting)
- **Interpretability:** SHAP values used for feature importance
- **Performance Metrics:**

Metric	XGBoost	LightGBM
MAE	0.303	**0.302**
RMSE	0.403	**0.401**
R^2	0.940	**0.941**

- **Strengths:**
  - High accuracy
  - Strong interpretability with SHAP

- Good for short-term forecasting and feature analysis

✅ Model 2: Explainable Boosting Machine (EBM)

- **Model Type:** Generalized Additive Model (interpretable ML)
- **Dataset:** US Retail Sales (RSAFS)
- **Feature Engineering:**
  - Extracted time components (year, month, weekday)
  - Lag features
- **Interpretability:**
  - Built-in EBM feature plots
  - SHAP for additional insights
- **Performance Metrics:**

Metric	EBM
MAE	**5.59**
RMSE	**6.55**
R^2	**0.973**

- **Strengths:**
  - Best overall accuracy
  - Extremely interpretable
  - Clear view of how features affect predictions
- **Limitations:**

- Not suitable for multi-step forecasting or seasonality
- No external indicators used

✅ Model 3: Prophet

- **Model Type:** Additive Time Series Model by Meta (Facebook)
- **Dataset:** US Retail Sales (RSAFS) with external indicators (CPI, UNRATE, FEDFUNDS)
- **Feature Engineering:**
  - Lag features, rolling stats
  - Built-in seasonality modeling (yearly, monthly)
- **Interpretability:**
  - Prophet components (trend, seasonality, holidays)
- **Performance Metrics:**

Metric	Prophet
MAE	5.85
RMSE	7.12
R^2	0.968
MAPE	4.7%

- **Strengths:**
  - Strong at modeling seasonality and trend
  - Supports long-term forecasting
  - Easy to visualize components

- **Limitations:**
    - Slightly lower accuracy than EBM
    - Less interpretable at the feature level

✅ Model 4: Temporal Fusion Transformer (TFT)

- **Model Type:** Deep Learning (Sequence-to-Sequence Transformer)
- **Dataset:** Bank Marketing Data (Binary Classification - subscription)
- **Feature Engineering:**
    - Categorical encoding
    - Time index creation
    - Grouped by campaign ID
    - Rolling time windows with encoder length = 30
- **Interpretability:**
    - TFT's built-in feature importance tools
    - Heatmaps of attention weights and variable relevance
- **Performance Metrics:**
    - ⚠️ **Not explicitly reported** (no MAE/RMSE/R^2 shown)
- **Strengths:**
    - Powerful deep learning for complex time series
    - Good for handling multiple variables and non-linear dynamics
- **Limitations:**

- High complexity
- No performance numbers shown (hard to compare directly)
- Requires more resources to train

Overall Model Comparison

Criteria	XGBoost/LightGBM	EBM	Prophet	Temporal Fusion Transformer (TFT)
Model Type	Tree Ensemble	Additive GAM	Additive Time Series	Deep Learning Transformer
Forecasting Capability	❌	❌	✅	✅
Interpretability	SHAP	Very High	Medium (components)	Medium-High (attention-based)
External Indicators	No	No	Yes	No
R^2 Score (Best)	0.941 (LightGBM)	✅ **0.973**	0.968	❌ Not available
MAE (Best)	**0.302**	5.59 (different scale)	5.85	❌ Not available
Use Case Suitability	Short-term prediction	Feature analysis	Forecasting with trends	Long-term, complex time series

---

🏆 Conclusion: Best Model by Use Case

- **Most Accurate Model:**
  ➢ **EBM** – Best R^2 and strong MAE performance. Great for interpretable business insights.

- **Best for Long-Term Forecasting:**
  ➢ **Prophet** – Handles trend/seasonality and supports future predictions.
- **Best for Complex Temporal Modeling:**
  ➢ **TFT** – Excellent architecture for deep time series tasks, but needs metrics to confirm.
- **Balanced Performance + Interpretability:**
  ➢ **LightGBM** – Very good accuracy and SHAP-based feature importance.