Semester : VIII Subject : AIFB Academic Year: 2024-2025

Model stability in backtesting refers to how consistently a trading or investment model performs across different market conditions, timeframes, and datasets. A model that's only profitable on one dataset but fails elsewhere may be overfitted or unstable.

Key Techniques to Ensure Model Stability

1. Split Testing (Train-Test Separation)

- **Divide your historical data** into two parts:
  - **Training set**: Used to build or optimize the model.
  - **Testing set**: Used to evaluate the model's performance on unseen data.
- **Purpose**: Confirms that the model generalizes well and is not overfit to the training data.

2. Walk-Forward Testing

- Simulate **real-time evaluation** by moving through data step-by-step:
  1. Train on an initial window.
  2. Test on the next window.
  3. Slide the window forward and repeat.
- **Purpose**: Measures performance consistency over time and prevents data snooping bias.

3. Cross-Validation on Time Series

- Instead of random splits (which violate time sequence), use **time-based folds** to test performance across different historical periods.
- Helps expose performance fluctuations during:
  - Bull markets
  - Bear markets
  - Sideways/volatile markets

4. Out-of-Sample Testing

- Test your model on a **completely new dataset** not used at all during model building.
- This is often from a **different year**, **different market**, or even **different instrument** (e.g., use NIFTY after training on SENSEX).

5. Performance Metrics Tracking

- Monitor stability across:
  - **Sharpe ratio**
  - **Win rate**
  - **Max drawdown**
  - **Average return**
  - **Profit factor**
- If these vary too much between periods, your model may lack stability.

6. Stress Testing / Scenario Analysis

- Test the model during known extreme periods like:

- 2008 financial crisis
- COVID-19 crash (March 2020)
- High-interest rate environments

- Does the model still behave logically? If not, it may need redesigning.

7. Robustness to Parameter Changes

- Slightly change model parameters (e.g., moving average length, threshold values) and observe how the performance changes.
- A **stable model performs similarly across small parameter variations**.

8. Avoiding Overfitting

- Overfitting = A model that performs extremely well in backtest but fails in real-world data.
- Warning signs:
  - Too many indicators
  - Perfect equity curves
  - Unrealistically high win rates
- Keep models **simple and interpretable** for better stability.

9. Rolling Performance Analysis

- Divide your testing period into **rolling windows** (e.g., 6-month or 1-year chunks).
- Evaluate if performance is **steady or degrading** over time.

10. Compare Against Benchmark

- Always compare model performance to a **benchmark** (like a Buy-and-Hold strategy).
- If your model consistently beats the benchmark over time, it's a sign of stability.