Combining SARIMA (Seasonal Autoregressive Integrated Moving Average) and LSTM (Long Short-Term Memory) models can be a powerful approach for stock price prediction. SARIMA can handle the seasonality aspects while LSTM can capture long-term dependencies and nonlinear patterns in the data. Here’s a step-by-step approach on how to integrate SARIMA statistical features into an LSTM model:

Step 1: Data Preparation

Collect Data: Gather the historical stock price data that you want to predict.

Preprocessing: Normalize the data to ensure the LSTM model can train effectively. You might also need to create lagged features depending on the structure of your model.

Step 2: SARIMA Model Implementation

Model Identification: Use ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots to identify the order of ARIMA model (p, d, q) and seasonal components (P, D, Q, S).

Parameter Estimation: Fit the SARIMA model to your historical data to find the best coefficients.

Residual Analysis: Analyze the residuals of the SARIMA model to ensure there are no patterns (like seasonality or trends) left in the residuals.

Step 3: Feature Engineering from SARIMA

Extract Features: Use the coefficients and residuals of the SARIMA model as features for the LSTM model. Common features include:

Lagged Values of Residuals: Incorporate one or more lagged residuals.

Seasonal Adjustments: Values adjusted for seasonality.

Predictions from SARIMA: Use SARIMA predictions as an input feature to LSTM.

Combine Features: Alongside traditional features (like raw prices, volume, etc.), add these new SARIMA-derived features.

Step 4: LSTM Model Implementation

Create LSTM Architecture: Design the LSTM network with appropriate input layers (considering the number of features including SARIMA outputs), hidden layers, and an output layer.

Train LSTM: Train the LSTM model using the combined dataset (original features plus SARIMA-derived features).

Model Tuning: Adjust hyperparameters such as learning rate, number of layers, dropout rate, etc., to optimize performance.

Step 5: Prediction and Evaluation

Predict: Use the trained LSTM model to make predictions on the test data.

Evaluate: Assess the model using metrics such as RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), etc.

Compare: Optionally, compare these predictions with standard LSTM and standalone SARIMA models to evaluate the improvement.

Step 6: Integration and Deployment

Backtesting: Apply the model to historical data to perform backtesting and see how it would have performed in real-world trading.

Real-time Testing: Integrate the model into a simulation or real-time trading environment to test its practical effectiveness.

By leveraging the statistical strengths of SARIMA in capturing seasonality and the powerful memory capabilities of LSTM for complex patterns, this combined approach can potentially improve the accuracy of stock price predictions. Always ensure to conduct thorough backtesting to validate the model's effectiveness before any practical implementation.