Project Summary Report

The project involved exploring the process of classifying financial market regimes based on past stock market data and technical indicators, the idea behind this research being to classify financial markets regimes into uptrend, downtrend, and sideways markets. Systematic data preparation, exploratory data analysis, and the repeated modeling allow the study to mention opportunities and the inevitable disadvantages of the regime-based market prediction.

The dataset indicated quite good data quality and few missing values were limited to the roll-out of the indicators and no duplication or time violation was detected. Exploratory analysis also demonstrated that the trend behavior and volatility of the market regimes is mostly sustained but not isolated. Strong correlation coefficients between moving averages and volatility indicators implied redundancy in features whereas momentum-based indicators as well as RSI had secondary but complementary information. The features connected to volume did not have much standout predictive performance.

These findings were supported by modeling. A Logistic Regression baseline was inadequate to accommodate the dynamic in the regime that was found to be much biased on the sideways side and weak at separating the trends. Random Forest model enhanced predictive balance and explainability with non-linear feature interactions, although the total accuracy was weak. The feature importance analysis determined volatility and trend variables to be the most significant factors in classification decisions consistently with market behavior.

Practically speaking, the findings imply that regime classification is more applicable to strategic decision making rather than an exact timing of the trades. Individual period of high volatility and low volatility would be important in portfolio allocation, risk management and choice of strategy. The observed comparatively low consistency of the results of the different models is an aspect of ambiguity of market regimes and not bad models, especially when the market is in the transitional phase.

Subsequent development must focus less on expansion than on refinement of features by looking at redundancy and focusing on more regime sensitive changes like trend slopes or volatility percentiles. Performance indices must go beyond plain accuracy to indicate better class balance and regime-sensitive performance. More complex models such as gradient boosting techniques may be able to capture rather subtle interactions. On the whole, the project has developed a strong analytical base and offered distinct guidelines on how the same can be enhanced in the future in terms of regime-based financial modeling.