PART 3: Market Regime Detection

Sunday, 20 April 2025 12:14 PM

Comprehensive Project Details: Market Regime Detection Using Multidimensional Time Series and Option Market Features

1. Project Overview

Objective:
Develop an unsupervised machine learning system to automatically detect, characterize, and monitor persistent **volatility regimes** in financial markets. The system will use a rich set of features from asset prices, options data, technical indicators, and sentiment indices. It should also highlight regime transitions (structural breaks), model regime persistence, and provide interpretable outputs for practitioners.

2. Motivation

Financial markets exhibit distinct periods ("regimes") with stable statistical and behavioral properties, including calm, trending, normal, and panicked/reactive phases. Detecting these regimes—especially their **transitions**—is crucial for risk management, portfolio construction, and trading. Integrating **option market behavior** (e.g., realized skew, vol of vol) and **sentiment indicators** (e.g., fear & greed index) into regime modeling provides deeper, more actionable insight than price-based approaches alone.

3. Data & Feature Engineering

3.1 Data Sources

- **Asset price and return data** (daily or higher frequency)
- **Option market data:**
  - ATM implied volatility
  - Option skew
  - Implied volatility surface
- **Sentiment data:**
  - Fear & greed index (or similar sentiment gauges)

3.2 Feature Set

A. Statistical Features

- **Realized volatility** (e.g., rolling 20-day std. dev. of returns)
- **Vol of vol:**
  - Rolling standard deviation of realized volatility (e.g., 10-day std. dev. of 20-day vol)
- **Rolling cross-asset correlations**
- **Skewness and kurtosis** of returns

B. Technical Indicators

- **Average True Range (ATR)**
- **ADX (Average Directional Index):** Trend strength
- **Moving average slopes** (e.g., 50-day MA slope)
- **MACD (Moving Average Convergence Divergence)**
- **RSI (Relative Strength Index)**
- **Bollinger Band width**
- **Stochastic Oscillator**
- **Sharpe ratio (rolling)**

C. Option Market Features

- **ATM implied volatility**
- **Option skew and realized skew**
  - Realized skew calculated as:
    ATM_vol_t1 = ATM_vol_t0 + skew_t0 * (1 + realized_skew)
    - Sticky delta: realized skew ≈ -1
    - Sticky strike: realized skew ≈ 0
    - Sticky local vol: realized skew ≈ 1
    - Jumpy vol: realized skew > 1

D. Event-Based Features

- **Gap count:** Number of large open-close gaps in a rolling window
- **Downside beta**

E. Sentiment Feature

- **Fear & Greed Index:** Quantitative sentiment measure

4. Volatility Regimes to Detect

Regime	Sentiment	Duration	Spot-Vol Correlation	Realized Skew	Description
Sticky Delta	Calm, trending	Long-term	Positive	≈ -1	Stable, trending markets, constant % strike volatility
Sticky Strike	Normal	Medium-term	Zero	≈ 0	Medium-term, constant fixed strike volatility
Sticky Local Vol	Normal	Medium-term	Negative	≈ 1	Medium-term, negative spot-vol correlation
Jumpy Vol	Panicked, reactive	Short-term	Very negative	> 1	Volatility jumps, rapid market moves, sentiment-driven spikes

5. Modeling Approach

5.1 Feature Preparation

- **Calculate all features** at each time step.
- **Standardize/normalize features** to ensure comparability and model stability.
- **Feature selection:** Use domain knowledge and correlation analysis to avoid redundancy.

5.2 Model Selection

A. Hidden Markov Models (HMM) / Hidden Semi-Markov Models (HSMM)

- **Rationale:**
  - HMMs/HSMMs naturally model regime persistence (duration), regime transitions, and can handle multivariate input.
  - HSMMs are preferred if you want explicit modeling of regime durations.
- **Setup:**
  - Number of hidden states: 4 (matching the four regimes)
  - Input: Standardized feature vectors at each time step
  - Output: Regime assignment, regime probability, and transition points

B. Clustering (Supplemental)

- **Gaussian Mixture Models (GMM):** For exploratory clustering and initial regime mapping.

C. Change-Point Detection (Supplemental)

- Use change-point detection algorithms (e.g., Bayesian Change Point, Kernel Change Point) on feature trajectories to independently flag structural breaks.

5.3 Training and Validation

- **Train model** on historical data, using the EM algorithm or similar for HMM/HSMM.
- **Validate detected regimes** by comparing average/median feature values in each regime to theoretical expectations (see regime table).
- **Check regime durations** and transitions against historical market events.

6. Output & Visualization

- **Time series plot** marked by regime segments (distinct colors for each regime)
- **Feature distribution plots** by regime (e.g., boxplots of vol of vol, realized skew, sentiment)
- **Transition point markers** (arrows or vertical lines) for regime changes
- **Regime duration statistics** (average length of time in each regime)
- **Real-time monitoring dashboard** (optional for deployment)

7. Evaluation Criteria

- **Internal validity:**
  - Regimes are distinct and stable in feature space
  - Feature averages by regime match theoretical expectations (e.g., high vol of vol in “jumpy vol”)
- **External validity:**
  - Detected regime transitions align with known market events
  - Regime persistence statistics are plausible
- **Practical usability:**
  - Clarity and interpretability of outputs for domain experts
  - Timeliness and reliability of regime transition signals

8. Project Workflow

plaintext
Copy

```
Raw Data (Prices, Options, Sentiment)
        ↓
Feature Engineering (statistical, technical, option, sentiment features)
        ↓
Feature Scaling & Selection
        ↓
Modeling (HMM/HSMM, optional GMM or change-point detection)
        ↓
Regime Assignment & Transition Detection
        ↓
Validation (theoretical alignment, event overlap, persistence analysis)
        ↓
Visualization & Reporting (plots, dashboards, regime stats)
        ↓
Deployment (real-time regime monitoring, alerts on transitions)
```

9. Implementation Notes

- Periodically retrain the model to adapt to market evolution.
- Regularly update feature definitions and windows as needed.
- Document feature importance and model decisions for transparency.
- Consider ensemble or hybrid approaches for robustness.

10. Expected Outcomes

- **Robust, interpretable regime detection** system for multidimensional financial data.
- **Clear identification of turning points** and regime durations.
- **Feature-driven regime characterization** (including vol of vol, realized skew, and sentiment).
- **Actionable outputs** for portfolio management, risk, and trading strategies.