

暗号資産市場における予測モデル構築のための 高度な特徴量エンジニアリング: OHLCV、 ファンディングレート、建玉(OI)に基づく定量的 アプローチ

1. エグゼクティブサマリー

暗号資産(仮想通貨)市場は、その24時間365日の稼働体制、複数の取引所への流動性の分断、そして無期限先物(Perpetual Swaps)という独自のデリバティブ商品の存在により、従来の金融市場とは一線を画す市場構造を有しています。この市場において、機械学習モデルやアルゴリズム取引戦略の収益性を決定づける最も重要な要素の一つが、**特徴量エンジニアリング(Feature Engineering)**です。ユーザーの要求に基づき、本レポートでは注文板(L2/L3)データを使用せず、**OHLCV**(始値・高値・安値・終値・出来高)、ファンディングレート(**FR**)、**建玉(OI: Open Interest)**の3つの主要データセットのみを用いて、予測力の高い「強い特徴量」を生成するための手法を網羅的に分析します。

本稿の分析は、最新の定量的金融工学の文献および実証研究に基づいています。分析の結果、単一の生データ(Raw Data)が予測力を持つことは稀であり、市場のマイクロ構造的な不均衡や参加者の行動心理を捉えるための高度な変換が必要であることが明らかになりました。具体的には、以下の4つのカテゴリーにおいて「強い特徴量」が特定されました。

1. 定常性を維持した時系列変換: 従来の整数階差分($d=1$)ではなく、**実数階差分(Fractional Differencing)**を用いることで、価格データの長期記憶(Memory)を保持しつつ、機械学習モデルに必要な定常性を確保する手法。
2. 高度なボラティリティ推定: 単純な終値ベースの分散ではなく、日中の高値・安値を考慮した**Garman-Klass**推定量や、トレンド成分を除去した**Yang-Zhang**推定量を用いることで、市場の不確実性をより正確に捉える手法。
3. デリバティブ市場の歪みの定量化: ファンディングレートと建玉の相互作用から、市場のポジショニングの偏り(Overcrowdedness)を検出する**OI加重ファンディングレート**や、累積ファンディングレートの乖離(**Divergence**)。
4. 強制決済(**Liquidation**)の予兆検知: 建玉と出来高の比率や、市場規模に対するレバレッジ比率を用いることで、価格変動に対する市場の脆弱性(Fragility)を測定し、ショートスクイズや**清算連鎖(Liquidation Cascade)**の発生確率を数値化する手法。

本レポートは、これらの特徴量の数学的定義、経済学的背景、および実装上の注意点を詳細に解説し、実務家が堅牢な取引モデルを構築するための包括的なガイドラインを提供することを目的としています。

2. 序論:暗号資産データの特性と課題

2.1 暗号資産市場の異質性

特徴量エンジニアリングの議論に入る前に、対象となるデータの特性を深く理解する必要があります。暗号資産市場は、株式市場や外国為替市場と比較して、極めて高いノイズ率と非定常性を持っています。

まず、**OHLCV**データの概念自体が伝統的市場とは異なります。株式市場には明確な「大引け(Close)」が存在し、その価格はクロージング・オークションによって決定されるため、極めて高い流動性と情報価を持ちます。一方、暗号資産市場は連続的であり、例えば「日足の終値」は協定世界時(UTC)00:00という恣意的な瞬間の価格に過ぎません。このため、単一の時点価格(終値)のみに依存した指標は、その瞬間のミクロな価格変動ノイズの影響を強く受けやすく、予測モデルの汎化性能を低下させる要因となります¹。

次に、**無期限先物(Perpetual Swaps)の存在です。これは暗号資産市場の取引高の大部分を占める商品であり、満期日がありません。現物価格との乖離を防ぐためにファンディングレート(資金調達率)**というメカニズムが存在し、これが市場のセンチメントとレバレッジコストを直接的に反映するユニークなデータポイントとなります²。

さらに、**建玉(Open Interest)**の透明性が挙げられます。多くの伝統的市場ではOIの開示は日次や週次ですが、暗号資産ではリアルタイム(あるいは分単位)でAPI経由で取得可能です。これにより、資金の流出入を極めて高い粒度で追跡することが可能となります⁴。

2.2 機械学習における「定常性」のジレンマ

「強い特徴量」を定義する上で避けて通れないのが、時系列データの**定常性(Stationarity)**の問題です。ランダムフォレストや勾配ブースティング(LightGBM, XGBoost)、ニューラルネットワークなどの多くの機械学習アルゴリズムは、入力データの統計的特性(平均、分散、自己相関)が時間を通じて一定であることを前提としています(あるいは、その方が学習効率が良い)。

しかし、暗号資産の価格データは明らかに非定常です(トレンドがあり、分散が時間とともに変動する)。これに対処するために、実務では対数収益率(Log Returns)や階差(Differencing)を取るのが一般的です。

$$r_t = \ln(P_t) - \ln(P_{t-1})$$

この変換はデータを定常化させますが、同時に**記憶(Memory)**を完全に消去してしまいます。価格が過去の最高値に近いのか、暴落後の底値圏にあるのかという「水準」の情報が失われるのです。長期的なトレンドやサイクルを捉えるためには、この記憶の保持が不可欠です⁵。

本レポートでは、この「定常性」と「記憶」のトレードオフを解消する**実数階差分(Fractional Differencing)**を、現代の定量的取引における最も強力な前処理手法の一つとして詳述します。

3. データ前処理と定常性の確保: 実数階差分

強い特徴量を作成するための第一歩は、情報を破壊せずにデータをクリーニングすることです。特に OHLCV データにおいて、実数階差分は必須のテクニックとなりつつあります。

3.1 整数階差分の限界と実数階差分の導入

従来の時系列分析では、1次の階差 ($d=1$) を取ることでトレンドを除去していました。これは、昨日の価格と今日の価格の差分のみを見ることであり、一昨日以前の情報は(数学的には)切り捨てられます。しかし、金融時系列、特に暗号資産のような投機性の高い市場では、過去の価格推移が現在の参加者の心理状態(含み益・含み損の状態)に影響を与えるため、長期記憶には予測力があります。

****実数階差分(Fractional Differencing)****は、差分の次数 d を整数(0や1)ではなく、実数(例えば 0.4 や 0.6)として扱います。これにより、データの定常性を確保しつつ、過去のデータの重みを緩やかに減衰させながら保持することが可能になります⁷。

3.2 数学的定式化

ラグ演算子(Backshift Operator) B を用いて、時点 t の価格 X_t に対する d 階の差分は以下のように定義されます。

$$(1 - B)^d X_t = \sum_{k=0}^{\infty} \omega_k X_{t-k}$$

ここで、重み係数 ω_k は以下の式で計算されます。

$$\omega_k = (-1)^k \prod_{i=0}^{k-1} \frac{d-i}{i+1}$$

具体的には、

- $\omega_0 = 1$
- $\omega_1 = -d$
- $\omega_2 = \frac{d(d-1)}{2}$
- ...と続きます。

$d=1$ の場合、 $\omega_0=1$, $\omega_1=-1$ となり、それ以降の重みはすべて0になります(これが通常の階差です)。しかし、 $0 < d < 1$ の場合、重みは無限に続きますが、徐々にゼロに収束します。これにより、遠い過去のデータも(わずかながら)現在の値に含まれることになります⁵。

3.3 暗号資産における最適な d の探索

では、具体的にどのような d の値を用いれば「強い特徴量」となるのでしょうか。研究によれば、

以下の手順で最適な d を探索することが推奨されています。

1. 拡張Dickey-Fuller検定(ADF検定)の実施: 対象となる暗号資産の対数価格系列に対し、 d を 0.0 から 1.0 まで 0.05 刻みで変化させながら差分系列を作成し、それぞれのADF検定のp値を計算します。
2. 閾値の決定: p値が有意水準(通常は0.05)を下回り、定常とみなせる最小の d を特定します。
3. 相関の確認: その d で作成した差分系列と、元の価格系列との相関係数を計算します。相関が高いほど、元の「記憶」が保たれていることを意味します。

実証研究によると、ビットコイン(BTC)やイーサリアム(ETH)などの主要通貨において、最適な d はしばしば $0.35 \sim 0.45$ の範囲に収束します⁶。これは、完全な階差($d=1$)を取る必要はなく、約0.4次の差分で十分に定常性が確保できることを示唆しています。

特徴量としての実装:

OHLCの各値(対数変換後)に対して、この最適化された d を適用して得られる系列(FracDiff_Close, FracDiff_High など)は、単なるリターン(Return)よりも情報量が多く、機械学習モデルの入力として極めて強力です。これにより、モデルは「現在の価格変動」だけでなく「現在の価格水準の歴史的な位置づけ」を同時に学習することが可能になります。

4. OHLCVベースの特徴量エンジニアリング: 価格行動の深化

OHLCVデータは最も基本的なデータですが、単純な移動平均線やRSIのような古典的な指標は、現代の効率的な市場ではエッジ(優位性)を失っています。ここでは、市場のミクロ構造や統計的性質に基づいた、より高度な特徴量を設計します。

4.1 高度なボラティリティ推定量

ボラティリティ(変動率)は、リスクの指標であると同時に、情報が市場に流入する速度の代理変数でもあります。一般的に用いられる「終値ベースの標準偏差(Close-to-Close Volatility)」は、日中の高値・安値の情報を捨ててしまっているため、情報の利用効率が非常に悪い(非効率な)推定量です。暗号資産のように日中の変動が激しい市場では、**レンジベース(Range-based)**の推定量が圧倒的に強い特徴量となります¹⁰。

4.1.1 Parkinson推定量

Parkinson推定量は、特定の期間(日足や時間足)の高値(High)と安値(Low)のみを用います。

$$\sigma_{\text{Park}}^2 = \frac{1}{4 \ln 2} (\ln H_t - \ln L_t)^2$$

理論的には、終値ベースの推定量と比較して約5倍の効率性を持つとされています。これは、価格が

始値からどれだけ離れたか(最大エクスカッション)を捉えるため、トレンドの強さや反転の可能性を示唆する重要なシグナルとなります。

4.1.2 Garman-Klass推定量

Parkinson推定量は、市場が開いていない時間帯(ギャップ)や、始値と終値の変動を無視するという欠点があります。Garman-Klass推定量はこれを拡張し、始値(Open)と終値(Close)の情報も統合します。

$$\sigma_{GK}^2 = 0.5 (\ln \frac{H_t}{L_t})^2 - (2 \ln 2 - 1) (\ln \frac{C_t}{O_t})^2$$

暗号資産市場は24時間取引であるため、従来の株式市場のような「オーバーナイト・ギャップ」は存在しませんが、システムメンテナンスや流動性の空白による突発的な価格乖離は発生します。また、短い時間足(1時間足など)で見た場合、前の足の終値と次の足の始値がずれることは頻繁にあります。Garman-Klassはこれらの情報を漏らさず捉えるため、短期的な価格変動リスクの予測において非常に強力です¹²。

4.1.3 Yang-Zhang推定量

さらに高度なのがYang-Zhang推定量です。これは「ドリフト(トレンド)独立」かつ「始値ギャップ対応」の最小分散不偏推定量として知られています。

$$\sigma_{YZ}^2 = \sigma_{\text{Overnight}}^2 + k \sigma_{\text{OpenClose}}^2 + (1-k) \sigma_{\text{RogersSatchell}}^2$$

ここで、 k はデータの数に依存する定数です。

暗号資産市場は強力なトレンド(ドリフト)が発生しやすい市場です。通常の標準偏差では、強力な一方通行のトレンドが発生しているときに「ボラティリティが高い」と判定されてしまいます(平均からの乖離が大きくなるため)。しかし、トレーダーにとって、滑らかな上昇トレンドは「リスク(不確実性)」ではありません。Yang-Zhang推定量は、このトレンド成分とボラティリティ成分を分離できるため、純粋な「市場の乱高下」のみを抽出できます¹³。

機械学習モデルにおいて、トレンド指標(移動平均など)とボラティリティ指標(YZ推定量)を同時に入力する際、両者の相関が低いと、多重共線性の問題を回避しつつ、よりリッチな情報をモデルに与えることができます。

4.2 出来高加重平均価格(VWAP)とその乖離

出来高(Volume)は価格変動の燃料ですが、単独ではノイズが多いデータです。価格と出来高の相互作用を捉える最も信頼性の高い指標が**VWAP(Volume Weighted Average Price)**です。機関投資家のアルゴリズム取引の多くがVWAPをベンチマークとして執行されるため、VWAP自体が強力なサポート・レジスタンスラインとして機能します¹⁴。

4.2.1 アンカー付きVWAP(Anchored VWAP)からの乖離

「強い特徴量」として提案したいのは、固定期間(20日など)の移動平均ではなく、特定のイベントを起点としたAnchored VWAPからの乖離率(Zスコア)です。

$$Z_{\text{VWAP}} = \frac{P_t - \text{VWAP}_t}{\sigma_{\text{VWAP}}}$$

ここでの σ_{VWAP} は、VWAPを基準とした価格の標準偏差です。

起点の例:

- 日次リセット (**Session VWAP**): デイトレードにおいて、その日の「フェアバリュー」からの乖離を測る。
- 週次リセット: スイングトレードにおける中期トレンドの判定。
- イベント起点: 米雇用統計発表時や、重要なプロトコルアップデート時を起点とする。

シグナルの解釈:

価格がVWAPから $+2\sigma$ 以上乖離した場合、統計的に「買われすぎ (Overextended)」と判断され、平均回帰 (Mean Reversion) の圧力が強まります。逆に、トレンドが強い場合、価格は $+1\sigma$ バンドに沿って上昇を続けます (バンドウォーク)。この「バンド幅の拡大・縮小」と「価格の位置」を組み合わせた特徴量は、ブレイクアウトの予兆検知に有効です。

4.3 相対出来高 (RVOL) と吸収 (Absorption)

単純な出来高の移動平均との比較に加え、**相対出来高 (RVOL: Relative Volume)** は、特定の時間帯における「異常な」出来高を検知するために重要です。暗号資産市場は24時間動いていますが、地域ごとの活動時間 (東京時間、ロンドン時間、NY時間) によってベースとなる出来高は異なります。

$$\text{RVOL}_t = \frac{V_t}{\text{Avg}(V_t, \text{same_time_of_day})}$$

過去数週間の「同時刻」の平均出来高と比較することで、真のサプライズを抽出できます。

吸収 (Absorption) パターンの検知:

RVOLが極めて高い (例: 通常時の3倍以上) にもかかわらず、価格変動幅 (High - Low) が極端に小さい場合、それは「吸収」を示唆しています。大口投資家が指値注文 (Limit Orders) で市場の売り (または買い) をすべて吸収している状態であり、強力な反転シグナルとなります。

特徴量としては、以下のように定義できます。

$$\text{Absorption_Score} = \frac{\text{RVOL}_t}{\text{Range}_t}$$

このスコアが高いほど、価格が動かないまま出来高が消化されていることを示し、次の価格変動のエネルギーが蓄積されていることを示唆します。

5. 無期限先物のダイナミクス: ファンディングレート (FR)

ここからは、暗号資産特有のデータセットであるデリバティブ指標に焦点を当てます。無期限先物に

における**ファンディングレート(資金調達率)**は、市場のセンチメントとポジションの偏りを直接的に数値化したものであり、最強の予測因子の一つです²。

5.1 ファンディングレートのメカニズムと裁定機会

ファンディングレートは、無期限先物価格を現物指数価格(Spot Index Price)に収束させるために設計されたメカニズムです。

- **FR > 0 (プラス)**: 先物価格 > 現物価格。ロングポジション保有者がショートポジション保有者に金利を支払う。強気相場(コンタンゴ)を示唆。
- **FR < 0 (マイナス)**: 先物価格 < 現物価格。ショートポジション保有者がロングポジション保有者に金利を支払う。弱気相場(バックワーデーション)を示唆。

このレートは通常8時間ごとに決済されます(取引所により異なる)。

5.2 FRを用いた強力な特徴量

5.2.1 アニュアルライズド・キャリー (Annualized Carry)

これは「市場がロング(またはショート)ポジションに対してどれだけのコストを要求しているか」を年率換算したものです。

$$\text{Carry} = \text{FR}_{\text{current}} \times 3 \text{ (1日3回)} \times 365$$

例えば、8時間ごとのFRが0.05%の場合、年率では約54%になります。

予測ロジック: キャリーが極端に高い(例: 年率50%~100%超)場合、現物買い・先物売りの「デルタニュートラル裁定(Cash and Carry Arbitrage)」を行うアービトラージ・ボットが市場に参入します。これにより、先物には強力な売り圧力がかかり、価格の上昇を抑制する要因となります。逆に、この高コストを支払ってでもロングが積み上がっている場合、後述する「ロングスクイーズ」のリスクが高まります¹⁷。

5.2.2 累積ファンディングレート (Cumulative Funding Rate: CFR) の乖離

ファンディングレートを価格のように累積してプロットしたものをCFRと呼びます。

$$\text{CFR}_t = \sum_{i=0}^t \text{FR}_i$$

このCFRと実際の価格推移の**ダイバージェンス(乖離)**は、非常に精度の高い先行指標となります¹⁸。

- 強気のダイバージェンス: 価格が安値を更新している(下落トレンド)にもかかわらず、CFRが横ばい、あるいは上昇し始めた場合。これは、価格は下がっているものの、ショート(売り)の勢いが弱まり、むしろアグレッシブなショートがいなくなっていることを示します。底打ちのサインとして機能します。
- 弱気のダイバージェンス: 価格が高値を更新している(上昇トレンド)にもかかわらず、CFRが上

昇していない(あるいは低下している)場合。これは、価格の上昇が現物主導 (Spot-driven) であり、過度なレバレッジ(ロング)によるものではないことを示す健全な上昇か、あるいはショート勢が燃料となっていないため上昇力が枯渇しつつあることを示唆します。文脈によりますが、CFRの急低下を伴う価格上昇は、ショートカバー (踏み上げ) による一時的な上昇であることが多いです。

5.2.3 予測レートと実現レートのサプライズ

取引所は、次回のFR決済に向けた「予測ファンディングレート (Predicted Funding Rate)」をリアルタイムで公開しています。これと、実際に確定した「実現ファンディングレート (Realized Funding Rate)」の差分は、直前の需給の急変を表します 19。

$$FR_Surprise_t = Realized_FR_t - Predicted_FR_{t-\Delta t}$$

決済直前に大口の注文が入り、板のバランスが一気に崩れた場合、このサプライズ値が大きくなります。これは短期間のボラティリティ急拡大の予兆として機能します。

5.2.4 FRのローリングZスコア

FRの絶対値は、相場環境 (強気相場か弱気相場か) によってベースラインが異なります。機械学習モデルに入力する際は、過去の期間 (例: 30日) の平均と標準偏差を用いて正規化 (Zスコア化) することが推奨されます 20。

$$Z_FR_t = \frac{FR_t - \mu_{FR}}{\sigma_{FR}}$$

$Z_FR > 3$ のような極端な値は、平均回帰の絶好の機会 (逆張りシグナル) となります。

6. ポジショニングとレバレッジ: 建玉 (OI) の分析

建玉 (Open Interest) は、市場に参加している「未決済の契約総数」を表します。出来高が「フロー (流れ)」であるのに対し、OIは「ストック (蓄積)」であり、市場に溜まったエネルギー (ポテンシャル) を示します。

6.1 建玉と価格の相互作用マトリクス

OIの変化と価格の変化の組み合わせは、トレンドの性質を分類する上で最も基本的なフレームワークです ²¹。これをカテゴリカル特徴量としてモデルに組み込むことは非常に有効です。

価格変動 (ΔP)	OI変動 (ΔOI)	市場の状態	解釈と予測
上昇	増加	Long Build-up	新規資金の流入に

			よる強い上昇トレンド。順張り(Follow Trend)。
上昇	減少	Short Covering	売り方の買い戻しによる上昇。新規買いではないため、買い戻し一巡後に下落しやすい。
下落	増加	Short Build-up	新規売りによる強い下落トレンド。順張り(Follow Trend)。
下落	減少	Long Liquidation	買い方の投げ売り(ロスカット)による下落。セリングクライマックス(底)に近い可能性。

特徴量としては、これらの状態を数値化(例:1~4のフラグ)するか、あるいはこれらを連続値として表現する「OI-Price Correlation」などが考えられます。

6.2 レバレッジ比率と市場の脆弱性

OIの絶対値だけでは、それが「危険な水準」なのかどうか判断できません。市場規模に対してOIが大きくなりすぎると、わずかな価格変動で大量のロスカットが発生する「脆弱な(Fragile)」状態になります。これを捉えるための比率指標を作成します。

6.2.1 OI / 時価総額レシオ (Leverage Ratio)

$$\text{Leverage_Ratio} = \frac{\text{Total OI (USD)}}{\text{Market Cap (USD)}}$$

ビットコインの場合、この比率が歴史的に高い水準(例えば3%以上)に達すると、大規模なデレバレッジ(暴落)が発生する確率が飛躍的に高まります²³。この指標は、機械学習モデルにおいて「高リスクレジーム」を識別するためのスイッチとして機能します。

6.2.2 OI / 出来高レシオ (Liquidity Efficiency)

$$\text{OI_Vol_Ratio} = \frac{\text{OI}_t}{\text{Volume}_t}$$

OIが積み上がっているにもかかわらず、出来高が低い状態は危険です。これはポジションが膠着し

ており、逃げ場（流動性）が少ないことを意味します。この状態で価格が逆行すると、薄い板を食い破りながらロスカットが連鎖するため、価格変動が加速します 25。

逆に、出来高が巨大でOIが増えない場合は、高頻度取引（HFT）や短期勢による回転売買が主であることを示唆します。

6.3 OI加重ファンディングレート

市場のセンチメントをより正確に測るために、ファンディングレートにOIで重み付けを行います 26。

$$\text{Weighted_Sentiment} = \text{FR}_t \times \text{OI}_t$$

単にFRが高い（0.1%など）だけでなく、その状態でOIが過去最大級に積み上がっている場合、ロングポジションの保有者は巨額の金利を支払い続けていることになります。これは「耐えている」状態であり、価格が上昇しなければ即座に決済売りに転じる潜在的な売り圧力となります。

7. 相互作用と複合特徴量: アルファの源泉

最も強力な「強い特徴量」は、これまで解説したOHLCV、FR、OIを組み合わせた複合指標（Interaction Features）にあります。これらは特定の市場イベント、特に「清算（Liquidation）」や「スクイズ（Squeeze）」をピンポイントで検出するために設計されます。

7.1 強制決済（Liquidation）カスケードの検知シグナル

「清算連鎖（Liquidation Cascade）」とは、価格変動がマージンコールを誘発し、強制決済注文（成行注文）が発生、それがさらに価格を動かし、次のマージンコールを呼ぶという負のフィードバックループです。これを予測することは、逆張り戦略において最大の利益機会となります²⁸。

特徴量設計のロジック:

- 脆弱性の確認: $\$Z_{OI} > 2.0\$$ （レバレッジが過剰に積み上がっている）。
- トリガーの検知: 短期間（例: 15分）での価格急変。
- 連鎖の確認: 価格変動と同時にOIが急減していること（ポジションが消滅している）。

数式化（Cascade Signal）:

$$\text{Cascade_Score} = -1 \times \text{sign}(\Delta P_t) \times \Delta \text{OI}_t \times \text{Volume}_t$$

（注: OIが減少するとき ΔOI は負の値になるため、符号を調整します）

このスコアがスパイクした瞬間は、まさに「セリングクライマックス」や「バイイングクライマックス」であり、平均回帰の可能性が極めて高いポイントです。

7.2 ショートスクイズ確率指数

ショートスクイズは、踏み上げ相場です。これが発生する条件は明確です。

- FRがマイナス: ショートが金利を支払っている。

2. OIが増加: それでもショートが増え続けている。
3. 価格が下がらない: ショートの売り圧力を現物の買いが吸収している。

特徴量 (Squeeze Probability):

$$\text{Squeeze_Prob} = \mathbb{I}(\text{FR} < 0) \times |\text{FR}| \times \Delta \text{OI} \times (\text{Price}_t - \text{Low}_n)$$

ここで \mathbb{I} は指示関数です。価格が直近の安値を割っていないにもかかわらず、FRがマイナスでOIが増えている場合、この指標は大きくなります。これは「罠にかかったショート」の量を表しており、価格が少しでも上昇すれば一気に買い戻し(スクイーズ)が発生します。

7.3 VWAPとOIのダイバージェンス

トレンドの終焉を見極めるための特徴量です。

- 価格は上昇トレンド(VWAPより上、かつ高値更新中)。
- しかし、OIは減少傾向にある。

これは、価格上昇の動力が「新規買い」ではなく「ショートの買い戻し(ロスカット)」であることを意味します。買い戻しが終われば(OIが減りきれば)、買う人がいなくなるため、トレンドは急停止します。

特徴量:

$$\text{Trend_Quality} = \text{Correlation}(\Delta P, \Delta \text{OI}, \text{window}=20)$$

上昇トレンド中にこの相関が負になった場合、トレンド転換の強いシグナルとなります。

8. 機械学習モデルへの統合と実装戦略

これまでに挙げた「強い特徴量」を実際に機械学習モデルに組み込む際のベストプラクティスをまとめます。

8.1 ターゲットラベルの設計:トリプルバリア法

特徴量がどれだけ優れていても、予測対象(ラベル)が不適切であればモデルは学習しません。単純な「次の足の終値」を予測するのはノイズが多すぎます。暗号資産には**トリプルバリア法(Triple Barrier Method)**が最適です⁷。

1. **Upper Barrier:** 利食いライン(例: ボラティリティ $\times 2$)。
2. **Lower Barrier:** 損切りライン(例: ボラティリティ $\times 2$)。
3. **Vertical Barrier:** 時間切れ(例: 24時間後)。

価格が最初にどのバリアに接触したかをラベル(+1, -1, 0)とします。これにより、ボラティリティに動的に適応した、より実戦的な予測が可能になります。先述の「Yang-Zhangボラティリティ」をバリア幅の決定に使用することで、一貫性のあるラベリングが可能になります。

8.2 モデル選択と特徴量重要度

これらの特徴量(特に相互作用特徴量)を扱うには、**勾配ブースティング決定木(LightGBM, XGBoost, CatBoost)**が最も適しています。これらは非線形な関係や、データの欠損、スケーリングの違いに対して堅牢であり、金融時系列データ(特にテーブルデータ形式)において深層学習モデル(LSTMなど)を凌駕するパフォーマンスを示すことが多いです³¹。

一方、LSTMなどのシーケンスモデルを使用する場合は、前述の実数階差分による定常化処理が必須となります。非定常なデータをそのままRNNに入力しても、モデルは過去の価格水準に過剰適合(Overfit)し、未来の未知の価格帯に対応できなくなるからです。

8.3 評価と検証: Look-ahead Biasの排除

特にファンディングレートを使用する際、バックテストで「未来の情報を参照(Look-ahead Bias)」してしまうミスが頻発します。FRは8時間ごとに決定されますが、その期間中の「予測レート」は変動します。モデルの学習には、その瞬間に実際に取得可能だった値(Predicted Funding Rateの過去のスナップショット)を使用するか、あるいは確定したRealized Funding Rateのみを、確定時刻以降に使用するように厳密に管理する必要があります。

また、検証には時系列を分割するWalk-Forward Validationや、学習データとテストデータの間にパージ期間(Purge Period)を設けるPurged K-Fold Cross Validationを用いることで、情報のリークを防ぎ、実運用に近い性能評価を行うべきです。

9. 結論

暗号資産価格予測のための「強い特徴量」は、単なる価格の加工データではありません。それは、市場のマイクロ構造(Microstructure)、資金調達コスト(Cost of Carry)、そして**投資家のポジショニング(Positioning)**の3要素を統合し、市場の歪みを定量化したものです。

本レポートの分析により、以下の特徴量セットが推奨されます:

1. **トレンド**: 実数階差分を適用した価格系列(定常性と記憶の両立)。
2. **ボラティリティ**: Yang-Zhang推定量(トレンドとノイズの分離)。
3. **需給バランス**: OI加重ファンディングレート、および累積FRのダイバージェンス。
4. **イベント検知**: Liquidation Cascade Score(OI、価格、出来高の相互作用によるパニック検知)。
5. **脆弱性**: OI / 時価総額レシオ(レバレッジ過多の警告)。

これらをLightGBM等の非線形モデルに投入し、トリプルバリア法でラベリングを行うアプローチは、現在の定量的暗号資産取引における「State-of-the-Art(最先端)」のスタンダードと言えます。これらの特徴量は、市場の非効率性を捉え、ランダムウォークに見える価格変動の中から有意なシグナルを抽出するための強力な武器となるでしょう。

引用文献

1. Historical Data for Perpetual Futures - CoinAPI.io Blog, 12月 4, 2025にアクセス、
<https://www.coinapi.io/blog/historical-data-for-perpetual-futures>
2. Bitcoin: Funding Rates - All Exchanges - CryptoQuant, 12月 4, 2025にアクセス、
<https://cryptoquant.com/asset/btc/chart/derivatives/funding-rates>
3. Understanding Funding Rates in Perpetual Futures and Their Impact - Coinbase, 12月 4, 2025にアクセス、
<https://www.coinbase.com/learn/perpetual-futures/understanding-funding-rates-in-perpetual-futures>
4. Open Interest & Trading Volume: What Every Crypto Options Trader Must Know - OSL, 12月 4, 2025にアクセス、
<https://www.osl.com/hk-en/academy/article/open-interest-and-trading-volume-what-every-crypto-options-trader-must-know>
5. Time-Series Forecasting: Unleashing Long-Term Dependencies with Fractionally Differenced Data - arXiv, 12月 4, 2025にアクセス、
<https://arxiv.org/pdf/2309.13409>
6. Machine Learning Trading Essentials (Part 2): Fractionally ..., 12月 4, 2025にアクセス、
<https://hudsonthames.org/machine-learning-trading-essentials-part-2-fractionally-differentiated-features-filtering-and-labelling/>
7. Supervised Autoencoders with Fractionally Differentiated Features and Triple Barrier Labelling Enhance Predictions on Noisy Data - arXiv, 12月 4, 2025にアクセス、
<https://arxiv.org/html/2411.12753v2>
8. Fractional differencing as viewed from the frequency domain | by NTT - Medium, 12月 4, 2025にアクセス、
<https://medium.com/@nttp/fractional-differencing-as-viewed-from-the-frequency-domain-8ff33741408f>
9. Fractional differentiation and its use in machine learning - ResearchGate, 12月 4, 2025にアクセス、
https://www.researchgate.net/publication/354130032_Fractional_differentiation_and_its_use_in_machine_learning
10. (PDF) Bitcoin Return Volatility Forecasting: A Comparative Study between GARCH and RNN, 12月 4, 2025にアクセス、
https://www.researchgate.net/publication/353390822_Bitcoin_Return_Volatility_Forecasting_A_Comparative_Study_between_GARCH_and_RNN
11. Beyond GARCH in cryptocurrency volatility modelling: superiority of range-based estimators, 12月 4, 2025にアクセス、
<https://www.tandfonline.com/doi/abs/10.1080/13504851.2024.2363295>
12. Powerful Features: Volatility Estimators! : r/quant - Reddit, 12月 4, 2025にアクセス、
https://www.reddit.com/r/quant/comments/s21y0k/powerful_features_volatility_estimators/
13. Demystifying Time-Series Momentum Strategies: Volatility Estimators, Trading Rules, and Pairwise Correlations - Return Stacked® Portfolio Solutions, 12月 4, 2025にアクセス、
<https://www.returnstacked.com/academic-review/demystifying-time-series-mo>

[mentum-strategies-volatility-estimators-trading-rules-and-pairwise-correlations/](#)

14. Volume Weighted Average Price (VWAP) — Indicators and Strategies - TradingView, 12月 4, 2025にアクセス、
<https://www.tradingview.com/scripts/vwap/page-2/>
15. Master VWAP Swing Trading: 7 Powerful Strategies for Success - Opofinance Blog, 12月 4, 2025にアクセス、
<https://blog.opofinance.com/en/vwap-swing-trading-strategy/>
16. Fundamentals of Perpetual Futures We are grateful to Lin William Cong, Urban Jermann, Shimon Kogan, Tim Roughgarden, Adrien Verdelhan, as well as conference participants at the 2024 Utah Winter Finance Conference and seminar participants at a16z Crypto, Hebrew University, Reichman University, and the Virtual Derivatives Workshop for their insightful feedback and helpful comments. Songrun He - arXiv, 12月 4, 2025にアクセス、
<https://arxiv.org/html/2212.06888v5>
17. Crypto Stat Arb I: Quantifying & Combining Alphas · Analytic Musings, 12月 4, 2025にアクセス、<https://analytic-musings.com/2024/03/10/crypto-stat-arb-i/>
18. Perpetual Futures Liquidity Assessment Report - Chaos Labs, 12月 4, 2025にアクセス、
https://chaoslabs.xyz/resources/chaos_Labs_ethena_perpetual_assessment_risk_report.pdf
19. Predicted Funding Rates | Product Docs - Coin Metrics, 12月 4, 2025にアクセス、
<https://gitbook-docs.coinmetrics.io/market-data/market-data-overview/funding-rates/predicted-funding-rates>
20. Understanding Perpetual Futures: A Guide for Cryptocurrency Traders - Investopedia, 12月 4, 2025にアクセス、
<https://www.investopedia.com/what-are-perpetual-futures-7494870>
21. How to Use Open Interest in Intraday Trading: Strategies & Analysis | Kotak Securities, 12月 4, 2025にアクセス、
<https://www.kotaksecurities.com/investing-guide/intraday-trading/how-to-use-open-interest-for-intraday-trading/>
22. Open Interest: Overview, Calculation, Analysis, Trading Guide, Tool - Strike Money, 12月 4, 2025にアクセス、<https://www.strike.money/options/open-interest>
23. Marketcap — Penunjuk dan Strategi - TradingView, 12月 4, 2025にアクセス、
<https://my.tradingview.com/scripts/marketcap/>
24. Coinbase Research - Crypto Market Positioning (November 2025), 12月 4, 2025にアクセス、
<https://www.coinbase.com/institutional/research-insights/research/trading-insights/crypto-market-positioning-november-2025>
25. OI Volume Oscillator Cross Dynamics — Crypto_Moses的指標 - TradingView, 12月 4, 2025にアクセス、
<https://tw.tradingview.com/script/58IGqRsC-OI-Volume-Oscillator-Cross-Dynamics/>
26. Bitfinex Analysis Shows Bitcoin Presents a Promising Buying Opportunity - Crypto News, 12月 4, 2025にアクセス、

- <https://cryptonews.com.au/news/bitfinex-analysis-shows-bitcoin-presents-a-promising-buying-opportunity-120285/>
27. Bitfinex Alpha | Bitcoin Buying Opportunities Emerge as Interest in Alts Spread, 12月 4, 2025にアクセス、
<https://blog.bitfinex.com/bitfinex-alpha/bitfinex-alpha-bitcoin-buying-opportunities-emerge-as-interest-in-alts-spread/>
 28. The Leverage Purge: How \$8.55B in Liquidations Reset the Market - Amberdata Blog, 12月 4, 2025にアクセス、
<https://blog.amberdata.io/the-leverage-purge-how-8.55b-in-liquidations-reset-the-market>
 29. Risk Management in Cryptocurrency Exchanges: Protecting Users from Liquidation, 12月 4, 2025にアクセス、
<https://medium.com/@gwrx2005/risk-management-in-cryptocurrency-exchanges-protecting-users-from-liquidation-430acfd1b304>
 30. Liquidations — Indicadores y estrategias - TradingView, 12月 4, 2025にアクセス、
https://es.tradingview.com/scripts/liquidations/?script_access=all
 31. Assets Forecasting with Feature Engineering and ... - arXiv, 12月 4, 2025にアクセス、
<https://arxiv.org/pdf/2501.07580>
 32. (PDF) Assets Forecasting with Feature Engineering and Transformation Methods for LightGBM - ResearchGate, 12月 4, 2025にアクセス、
https://www.researchgate.net/publication/388029138_Assets_Forecasting_with_Feature_Engineering_and_Transformation_Methods_for_LightGBM