Introduction

Digital asset markets operate 24/7, and HFT algorithms require real-time price updates to make accurate decisions. For a given asset (e.g., BTC, ETH), multiple markets/pairs may exist. The most actively traded ones are spot and perpetual markets. In spot markets, the actual asset is traded directly, whereas in perpetual markets, derivatives of the asset are traded. Although the prices in these markets are correlated, they are often not identical.

Example:

- Spot: https://www.binance.com/en/trade/BTC_USDT
- Perpetual: https://www.binance.com/en/futures/BTCUSDT

There are two types of prices:

- The **bid** price represents the price that a buyer is willing to pay for an asset.
- The **ask** price is the price that a seller is willing to accept.

The **best bid** is the highest among all bid offers in the market, while the **best ask** is the lowest among all ask offers.

On Binance, the reaction time between these two markets is typically around 3–5 milliseconds (ms). A sudden price movement in one market is usually reflected in the other within 3–5 ms. However, not every price movement is reliable—some revert to their original level before affecting the other market.

Problem Definition

You are assigned to the Binance Spot-Perpetual Trading Team. You are provided with 24 hours of spot and perpetual **book ticker data** for the TRB/USDT pair, including timestamps, best bid, and best ask prices. You must keep in mind that the dataset **does not contain** data for **every single ms**. The timestamps between data points are **irregular**—For instance, there may be multiple entries within the same ms, or gaps of up to **100+ ms** between two consecutive entries.

The goal is to quantify and use the predictive power of **sudden price change** in one market on the other—specifically, to determine if and how a price movement in the spot (or perpetual) market causes a subsequent movement in the perpetual (or spot) market within a very short time window (5 ms).

If you're unsure how to define sudden price changes, you may use a ±0.07% (7 bps) price move within the last 3 ms as a working definition.

In analyzing price changes, you may choose to work with either best bid and ask prices separately, or use the mid-price defined as *(bid + ask) / 2*. This choice is left to your discretion based on what works best for your approach.

Additionally, a useful microstructure insight you may consider is that when prices are moving upward, the ask side tends to update first; conversely, when prices are declining, the bid side often reacts earlier. Of course, in some cases, both sides may move simultaneously.

Note that not all sudden price changes lead the other market. Some may revert quickly to their previous levels without affecting the second market. These are referred to as noise.

You are also provided **spot trade** data, including timestamps, price and quantity. Quantity could be beneficial to generate more insightful models.

Objectives

Primary Objective: Your task is to generate signals for i) spot trading by considering price movements in perpetual prices and ii) perpetual trading by considering price movements in spot prices. You need to predict the price that will be reached in one market within 5 ms, based on sudden price changes observed in the other one.

Secondary Objectives:

- Identify whether the spot or perpetual market more often leads price discovery. Which one drives the other one most?
- Detect and characterize noise events—sudden price changes that do not affect the other market.
- Estimate the momentum quality of the sudden price changes. You are encouraged to explore how long a price movement sustains its momentum—either in terms of **time duration** or **price delta**. For example, in the chart below, the movement in the red region shows high initial momentum but fades quickly, whereas the movement in the blue region, despite having lower intensity, persists over a longer period.