FIN-221: Machine Learning in Finance

HW1

Due on September 22, 2025

1. Read up on futures from any resource you like. I recommend the relevant chapter in Hulls book, Options, Futures, and Other Derivatives.

I have the 5th edition of Hull, it replaced the 1st edition I had in my office sometime in the early 2000s. I traded Stock Index Futures and Options for Timber Hill in the 1980s on the CME, CBOT and CBOE.

Most recently I worked on the US$ Swaps Pre-trade Pricing system for BofA including the transition from Libor (ED) to Sofr (SR3). Trade pricing for US$ Swaps, SOFR Futures, Bond Futures, and US Treasuries.

1. Exercise 2.1 of the textbook. You can use the programs and modules developed by Hud son and Thames. For futures roll data see https://raw.githubusercontent.com/hudson and-thames/example-data/main/futures stitched.csv I have also uploaded a zip le of sample ES data (courtesy of Hudson and Thames).
2. Read the paper The Volume Clock: Insights into the High Frequency Paradigm, David Easley, Marcos M. Lopez de Prado, Maureen OHara and answer the following questions:

* What is the volume-clock metric and why is it important?
  + Humans look at security price series in time buckets. Perhaps viewing returns from the last trade in a one-minute (hour, day, week, month, etc.,) interval to the last trade in the next one-minute (hour, day, week, month, etc.,) interval. Trade count, trade volume, VWAP etc., do not enter consideration or are of secondary interest. Traders may look at candle-stick charts (open, high, low, close) to extract trends with moving averages still looking at close to close prices.
  + The problem with close-to-close price returns is that they are very much non-normal – exhibiting leptokurtosis (fat tails) and non-stationarity in mean and variance
  + The volume-clock metric moves the analysis from time-based metrics to event-based metrics. Specifically, the volume-clock measures time in terms of trading volume or transaction counts.
  + Clark [1973] expands on Mandelbrot [1967] idea as a solution to “recover normality” in return data.
  + The ideas (volume clock) gain traction with the advent of HFT in the 1990s
* What is the main thesis of the paper regarding HFT vs LFT?
  + HFTs are characterized in the paper as predators using their informational advantage in market microstructure to identify LFT and exploit their predictable behavior.
  + LFTs leak information about their action and the HFT react to this information leakage to anticipate the LFT’s actions to exploit market microstructural opportunities.
* Explain how the authors have formed Figure 4 and how they interpret the results.
  + The authors start with a data set containing one year of E-mini futures trade between 11/7/2010 and 11/7/2011. I believe that they have quantity and date/time to the second of all trades. They cross tabulate the data bucketing volume data for each second (0 to 59) of a minute (for all minutes in the hour) vs hour (0 to 23) in the day (ignoring the date of observation) and calculating for each bucket the percentage of second buckets’ minute volume.