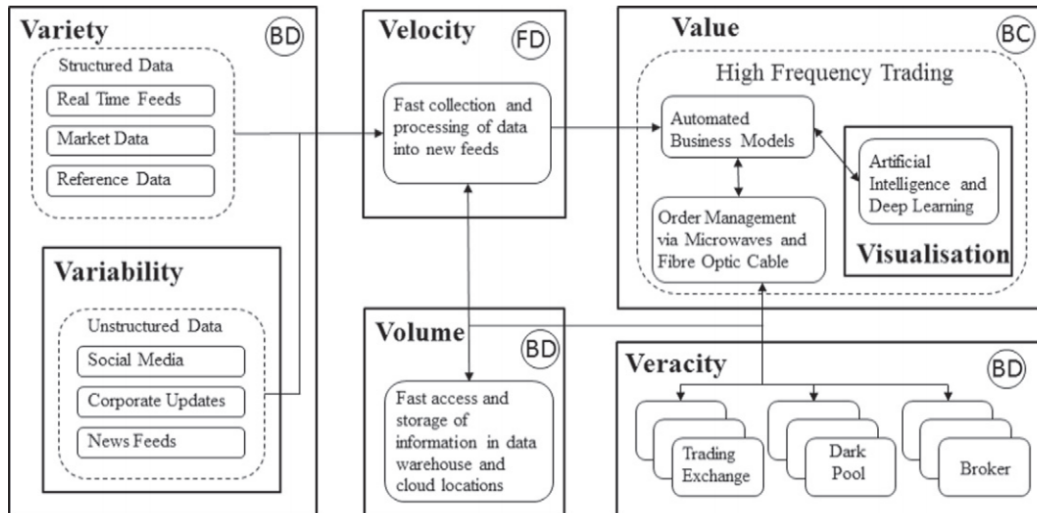


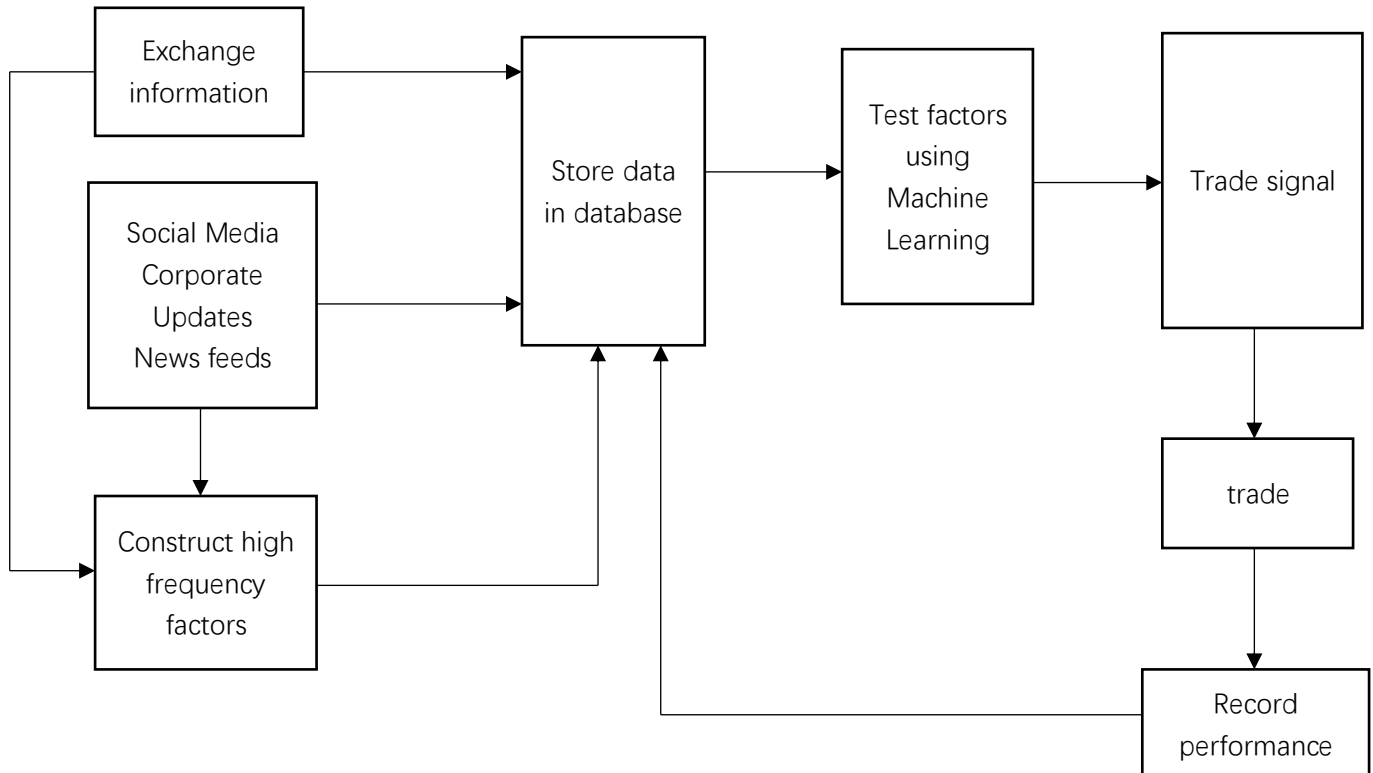
1. Big data application

Big data in quantitative finance is very popular recently, especially in high frequency trading, such as using high frequency factor to predict the price movement and algorithm trading. The intrinsic big data properties as follows:

(BD = Big Data; FD = Fast Data and BC = Big Compute)



2. Workflow



3. Database selection

For structured data like market data and real time feeds, I will choose time series database to store them, such as InfluxDB. Because in this kind of database, time is treated as first class with finest granularity and time-window based operations are highly optimized. Besides, most of exchange information are time series data

For social media data or news, I will use the MongoDB to store them. Because the volume of the social media data is very large, MongoDB performs better than SQL in bigdata especially in searching speed.