Investments in Cryptocurrencies

WANG Jiancheng 1155107822
Jhumka Abdoulaye 1155116361

Supervisor: Prof. HE Xuedong

April 2021

Abstract

This research project introduced Markowitz model, Fama French 3 Factor model, and Long Short Term Memory model separately to construct investment strategies for 68 cryptocurrencies. And results measured by a set of statistical and financial indicators such as Sharpe Ratio and Maximum Drawdown showed that all the investment strategies combined with different classical models outperform the market, though the LSTM strategy with long and short positions is the best one with an annualized return almost 10 times as much as the market return.

Contents

1 Objective

1.1 Background and Motivation

At the start of this year, on the 1st of January, the market capitalization of cryptocurrencies was approximately 191 billion U.S. dollars, as obtained from https://coinmarketcap.com. By the end of the year 2020, on the 31st of December, the market capitalization of cryptocurrencies was approximately 760 billion U.S. dollars. During one year, the cryptocurrency market increased by more than 570 billion U.S. dollars, an increase of nearly 400%.

The straightforward reason for such an exponential increase in the market size of cryptocurrencies is the presence of the pandemic caused by COVID-19. However, the pandemic had a negative impact on stock markets worldwide, when they reported, on 28 February 2020, their largest single-week declines since the 2008 financial crisis, the same crisis that led to the birth of Bitcoin. Satoshi Nakamoto, the creator of Bitcoin, must be proud now given that his creation and its predecessors bloomed during this current financial crisis, proving that decentralized digital currencies are more appealing to the public when the economy is undergoing a recession.

Moreover, there are many reasons why investing in cryptocurrencies is more beneficial than investing in stocks. Some reasons are listed as follows:

24/7 access to investment: to invest in the stock market, investors need to first find a broker. Only when they completed the administrative procedures can they start trading. Moreover, investors cannot trade all day long. Stock exchanges have trading sessions: timeframes when traders can place orders. As observed, access to the stock market is regulated. Therefore, it makes planning hard. Unlike trading stocks and commodities, the cryptocurrency market is not strictly regulated. Setting up an account to trade is relatively easy and there is no need for a broker. Moreover, it is accessible 24/7 across a growing number of exchanges, allowing for swift trade movements.

Low transaction fees: in addition to strict regulations, there are also expensive fees. A considerable portion of an investment is lost in the process of buying stocks. The costs accumulate as the investment traverse the stock exchange: brokers charge a fee or commission, banks will charge for the payments and the profits are taxed. In contrast, trading on cryptocurrency exchanges incurs relatively fewer costs. The costs associated with transacting on the blockchain are minuscule, consisting only of any mining fees. Exchanges themselves thus incur lower costs when buying and selling cryptocurrencies than do brokers for stock exchanges.

Global access and Transparency: stocks are usually traded in the country in which they are registered. Cryptocurrencies, led by Bitcoin, are accepted globally (more than 100 countries accept Bitcoin). Despite being globally available, there is no inter-border friction when transferring assets and transaction fees are still lower than stock exchanges. Moreover, the transaction would be clean, swift, transparent, and irreversible.

Not only assets: cryptocurrencies can be used as payment for goods and services gives them a great advantage over stocks.

Despite the potential of high returns, investors are still skeptical about investing in the cryptocurrency market due to high volatility. Early empirical results find Bitcoin to be partially ‘inefficient’ in Fama’s sense of efficient markets and predictability of returns, whereas a more recent work finds cryptocurrency asserts to be efficient[8]. In a recent paper, Brauneis and Mestel[2] concluded that efficiency of a single cryptocurrency is closely and positively related to its liquidity. Added to a portfolio, a sufficient amount of cryptocurrencies will have a market capitalization of billions of dollars, and paired with the fact that they usually provide lower transaction costs than the most efficient financial markets[6], it indicates ample liquidity.

Taken as a whole, these empirical findings, such as extreme volatility, bubble-behavior, and inefficiencies in the price discovery process of individual cryptocurrency, give rise to considering diversified cryptocurrency investments as a means of mitigating risk exposure in cryptocurrency markets.

1.2 Goal

There are plenty of cryptocurrencies being traded in the market and their influence on the financial world is enormous and unpredictable.

The objective of this project is to analyze the mechanism and market of cryptocurrencies, and the related regulation, so as to build up and apply portfolio selection models to cryptocurrency investment, and to test their performances. This research project applied 2 classical portfolio selection models and 1 modern portfolio selection model for cryptocurrency investment and compared their performances to determine the best portfolio model.

1.3 Summary of Contribution

This project was completed by Jhumka Abdoulaye (a third-year FTEC student) and WANG Jiancheng (a fourth-year FTEC student). Jhumka Abdoulaye completed constructing the Fama French 3-factor strategy; WANG Jiancheng completed constructing the LSTM strategy. Jhumka Abdoulaye and WANG Jiancheng both worked on the literature part and helped each other in the coding part.

1.4 Summarization of Final Year Project 1

The last research of Final Year Project 1 introduced CAPM and 1/N portfolio models to generate six optimal portfolios based on three criteria (whether short-selling is allowed; whether the chosen cryptocurrencies are market-capitalized or liquidized; and whether CAPM or 1/N model is adopted). The best portfolio model considering Return, Sharpe Ratio, and Maximum Utility is the CAPM portfolio model, which allows short-selling, with constituents based on their market capitalization. However, the overall performance of models applied in the last research project is relatively poor with a Sharpe Ratio less than 0.5 and there are some challenges and improvements (such as optimizing the cryptocurrency selection problem) that need to be overcome and achieved in the future research.

2 Mathematical and Financial Methods

2.1 Investment Models

2.1.1 Markowitz Model

The Markowitz model was the first that defined the concepts of risk and risk management related to financial and investment management[7]. The concept of risk in the model is defined as the volatility of asset prices which is usually represented by the variance of the asset's prices. The model manages the risk of a portfolio by choosing a combination of assets that yields the lowest overall portfolio's variance given an expected portfolio's return or the highest portfolio's return given a portfolio's variance.

The Markowitz model assumes that:

1. An investor is risk-averse.
2. An investor either maximizes their portfolio return for a given level of risk or minimizes their risk for a given return.
3. The market is perfect; it does not have taxes and transaction costs.
4. Analysis is based on a single period model of investment.
5. An investor is rational in nature.
6. There is no short sale.
7. Assets are infinitely divisible.

The Markowitz portfolio optimization can be stated mathematically as follows:

$$Min_{w_i}\sigma_p^2$$

Subject to:

$$\sigma_p^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}(w_i w_j \sigma_{ij})$$

$$r_p = r*$$

$$r_p = \sum_{i=1}^{N}(w_i r_i)$$

$$\sum_{i=1}^{N} w_i = 1$$

$$0 \leq w_i \leq 1 \quad \forall i$$

where:
σ_{ij} = covariance between asset i and j,
σ_p^2 = variance of the portfolio of assets,
r_i = expected return of asset i,
r_p = expected return of the portfolio,
$r*$ = a predefined level of return,
w_i = weight or proportion of asset i in the portfolio p.

This is the simplest form of Markowitz portfolio. The objective of the optimization is to minimize the portfolio variance, σ_p^2, at a certain return r_p. The no-short sell constraint it has on the weights of the assets, w_i, makes the model having no closed-form solution and NP-hard. It also has another unrealistic assumption of infinitesimal dividing of investment assets. However, for this research, a better version of the Markowitz portfolio optimization is implemented which has fewer restrictions and will be explained later in this paper.

2.1.2 Fama-French 3 Factor Model

In 1993, the professors Eugene Fama and Kenneth French developed the Fama-French 3 factor model[3]. This model is an extension of the Capital Asset Pricing Model (CAPM), which aims to explain the average stock returns through three factors: market risk, the outperformance of small-cap companies relative to large-cap companies, and the outperformance of high book-to-market value companies versus low book-to-market value companies. The logic behind the model is that high value and small-cap companies tend to regularly outperform the overall market. The model is explained mathematically as follows:

$$R_{i,t} - R_f = \alpha_i + \beta_{MKT,i}(MKT_t - R_f) + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \epsilon_{i,t}$$

where:
$R_{i,t} - R_f =$ the excess returns of the portfolio i over the risk-free rate,
$\alpha_i =$ the intercept of the equation,
$MKT_t - R_f =$ the excess of market returns over the risk-free rate,
$SMB_t =$ the return of the size factor,
$HML_t =$ the return on the value factor
$\beta_{XXX,i} =$ indicates how much the portfolio is loaded on each independent factor for each portfolio
$\epsilon_{i,t} =$ the error term.

Portfolios are constructed to mimic risk factors related to size and value. The size factor's portfolio is obtained by following these procedures:

1. Sorting the stocks on their market capitalization in descending order.
2. The stocks are parted by the median into two size groups, small and big.
3. The return of the portfolio is the difference between the returns of the portfolio with small stocks and the portfolio with large stocks.

The value factor's portfolio is obtained by following these procedures:

1. Sorting the stocks on their book-to-market ratio in descending order.
2. The stocks are split based on the breakpoints for the bottom 30% (Low), middle 40% (Medium).and top 30% (High) of the ranked values of book-to-market ratio for stocks.
3. The returns of the portfolio is the difference between the returns of the portfolio with high-value stocks and the portfolio with low-value stocks

Then they used cross-sections of average returns of the US stock market using the size and value factors to create portfolios. The excess return on the portfolios is used as dependent variables in the time-series regressions. The betas for the 3 factors for each portfolio are calculated by running a time series regression. The conclusion of Fama and French's research is that both the market value of a firm (size) and the book-to-market ratio (value) have high predictive power in the stock returns and can explain the differences in average returns across stocks that are not captured by the market premium in the CAPM.

2.1.3 Long Short Term Memory Model

LSTM was firstly proposed by Sepp Hochreiter and Jürgen Schmidhuber[5]. By introducing Constant Error Carousel (CEC) units, LSTM deals with the vanishing gradient problem. The initial version of the LSTM block included cells, input and output gates. In comparisons with real-time recurrent learning, backpropagation through time, and some other classical time-series algorithms, LSTM creates much more successful and effective runs and learns much faster. Moreover, LSTM is able to solve complicated, artificial long-time-lag issues and tasks that have never been solved by previous recurrent network algorithms.

RNNs with an LSTM architecture have more complex units that maintain an internal state and contain gates to keep track of dependencies between elements of the input sequence and regulate the cell's state accordingly. A common LSTM unit is composed of a cell, an input gate, an output gate and a forget gate. "The cell state is kind of like a conveyor belt. It runs straight down the entire chain, with only some minor linear interactions"[9], which remembers values over different arbitrary time intervals. The three gates recurrently connect to each other, aiming to address the problem of vanishing and exploding gradients by letting gradients pass through unchanged[5]. The picture below briefly shows the information flow for an unrolled LSTM unit and outlines its typical gating mechanism.

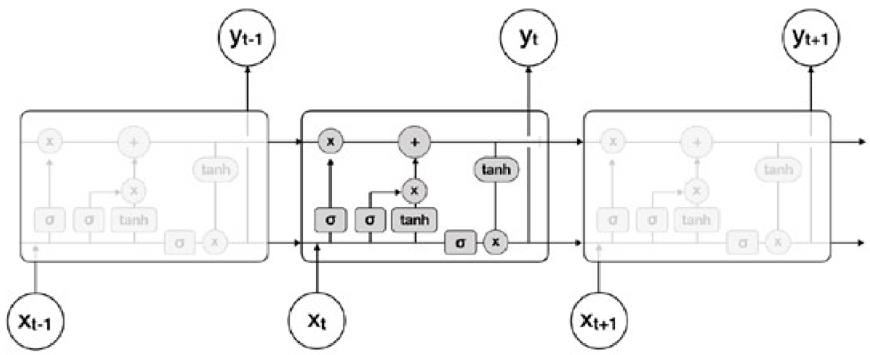

Figure 1: Information flow through an unrolled LSTM cell

A typical LSTM unit combines four parameterized layers which interact with each other and cell states via transforming and passing along vectors. These layers usually involve the three gates as above mentioned, but there are variations that may have additional gates or lack some of these three gates. The white nodes in Figure 2 identify element-wise operations, and the gray elements represent layers with weight and bias parameters learned during training:

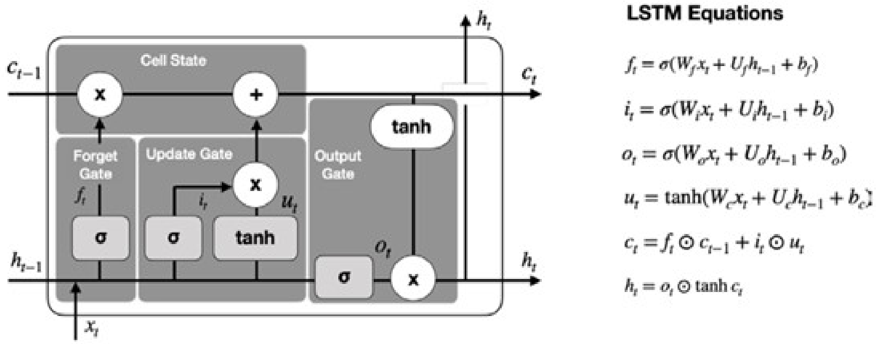

Figure 2: The logic of, math behind, an LSTM cell

The cell state, c, passes along the horizontal connection at the top of the cell. The cell state's interaction with the different gates results in a series of recurrent decisions:

1. The forget gate controls how much of that the cell's state is supposed to be voided to regulate the network's memory. It also receives the prior hidden state h_{t-1}, and the current input x_t, as inputs, computes a sigmoid activation, and multiplies the resulting value f_t, which has been normalized to the [0, 1] range.

2. The input gate computes a sigmoid activation from h_{t-1} and x_t which produces update candidates. A tan_h activation in the range from [-1, 1] multiplies the update candidates u_t, and adds or subtracts the result from the cell state depending on the resulting sign.

3. The output gate filters the updated cell state by using a sigmoid activation o_t, and multiplies it by the cell state normalized to the range [-1, 1] by using a tan_h activation

2.2 Financial and Statistical Indicators

In order to test the performances of portfolio strategies, this project adopts analytical criteria including Sharpe Ratio, Calmar Ratio, and Tail Ratio, most of which have become the most widely used methods for risk-adjusted return calculation and portfolio analysis.

2.2.1 Sharpe Ratio

The Sharpe ratio is the average return earned in excess of the risk-free rate per unit of volatility or total risk. Generally, the greater the value of the Sharpe ratio is, the more attractive the portfolio is.

$$Sharpe\ Ratio\ = \frac{R_p - r_f}{\sigma_p}$$

where R_p is the return of the portfolio, r_f is the risk-free rate, and σ_p is the standard deviation of the portfolio's excess return.

Usually, any Sharpe ratio greater than 1.0 is considered acceptable to good by investors. A ratio higher than 2.0 is rated as very good. A ratio of 3.0 or higher is considered excellent.

2.2.2 Calmar Ratio

The Calmar Ratio is the average compounded annual rate of return in excess of the risk-free rate versus its maximum drawdown. The higher the Calmar ratio, the better it performed on a risk-adjusted basis during the given time frame.

$$Calmar\ Ratio\ = \frac{R_p - r_f}{Maximum\ Drawdown}$$

A Calmar ratio greater than 1.0 is considered good, greater than 3 is very good, and greater than 5 is excellent.

2.2.3 Tail Ratio

The Tail Ratio is a ratio between the 95th and (absolute) 5th percentile of the distribution of the daily returns. For example, a ratio of 0.25 means that losses are four times as bad as profits.

$$Tail\ Ratio\ = \frac{|95_{th}\ of\ returns|}{|5_{th}\ of\ returns|}$$

3 Investment Models' Outline

3.1 Solution Approach

This subsection explains in detail how each model is used for investment.

3.2 Data & Portfolio Selection

This research project chose 68 cryptocurrencies having over 80% of the total market capitalization of the cryptocurrency market at all time. And the cryptocurrency data from January 2017 to December 2020 including daily trading data and market capitalization is fetched from finance.yahoo.com, coinmarketcap.com, and coingecko.com.

However, each model is fundamentally different and uses the data available differently. For example, the Markowitz model and Fama-French 3 Factor model use weekly data for investment but the Long Short Term Memory model uses daily data. The logic behind is that the latter model is more effective and requires more input data for prediction whereas the two former models are optimization method and sorting method respectively, therefore the daily volatility of a cryptocurrency is smoothed when using weekly data, leaving out noises and making the models more reliable.

The period of investment is from 1st January 2020 to 31st December 2020. Any data before the investment period may be used as training data, depending on the model. Thus, this subsection explains how each data is used and on which criteria the portfolios were created for each model.

3.3 Results

The results of each model will be analyzed and explained using their return and industrial financial and statistical tools such as Sharpe Ratio, Calmar Ratio, and Tail Ratio.

4 Markowitz Model

4.1 Solution Approach

In the Markowitz model, given a specific level of rate of return, there will be various portfolios with different variances, and only the portfolio with minimal variance is the optimal one. Then the model will adjust the rate of return into different values ranging from the minimal return of those selected cryptocurrencies to the maximal one of them. Consequently, an efficient frontier is constructed based on a set of optimal portfolios that offer the lowest risk for a given level of expected return. Thus, the optimal portfolio of the CAPM model consists of a risk-free asset and an optimal risky asset portfolio, and the optimal risky asset portfolio is at the point where the capital allocation line (CAL) is tangent to the efficient frontier. This portfolio is optimal because the slope of CAL is the highest, which means the model can achieve the highest returns per additional unit of risk. Therefore, the optimal weights of cryptocurrencies can be obtained from the optimal portfolio, which is used to construct future investment strategies.

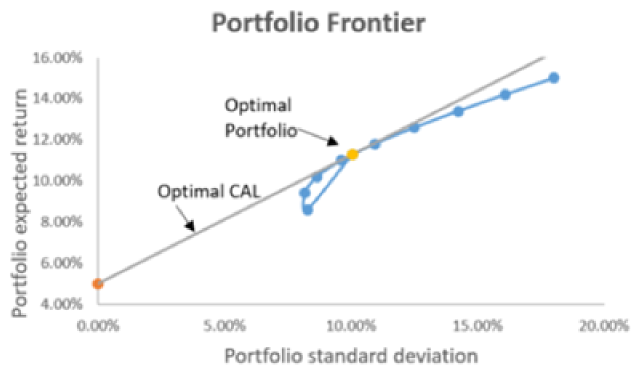

Figure 3: Optimal Portfolio on Efficient Frontier

This project mainly focuses on two optimization methods to solve the optimization problems in the Markowitz optimization model. In specific, the portfolio risk of return is quantified by σ_p^2 and in mean-variance analysis, only the first two moments (σ_p^2 and μ_p) are considered in the portfolio model. To minimize the portfolio risk no matter whether the portfolio model allows short-selling or not, the first method is to apply the closed form combined with quadratic programming; the second one is to use a function named "minimize" from the package "scipy.optimize" in Python.

4.1.1 Closed form and quadratic programming

The first approach is to apply the closed form to get the optimal portfolio with short-selling and quadratic programming to get the optimal one without short-selling since there is no closed-form solution in the inequality constrained optimization problem.

Firstly, for closed from, the solution can be calculated in the equality constrained optimization problem. To solve the problem, Markowitz's mean-variance analysis should be used and its mathematical formulation of given by:

$$\min \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(w_i w_j \sigma_{ij})$$

subject to $\sum_{i=1}^{N} w_i R_i = \mu_p$ and $\sum_{i=1}^{N} w_i = 1$. Given the target expected rate of return of portfolio μ_p, find the portfolio strategy that minimizes σ_p^2. To form the Lagrangian:

$$L = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}(w_i w_j \sigma_{ij}) - \lambda_1(\sum_{i=1}^{N} w_i - 1) - \lambda_2(\sum_{i=1}^{N} w_i R_i - \mu_p)$$

where λ_1 and λ_2 are Lagrangian multipliers. Then differentiate L with respect to w_i and Lagrangian multipliers, and set the derivative to zero.

$$\frac{\partial L}{\partial w_i} = \sum_{j=1}^{N} \sigma_{ij} w_j - \lambda_1 - \lambda_2 R_i = 0, i = 1, 2, ..., N$$

$$\frac{\partial L}{\partial \lambda_1} = \sum_{i=1}^{N} w_i - 1 = 0$$

$$\frac{\partial L}{\partial \lambda_2} = \sum_{i=1}^{N} w_i R_i - \mu_p = 0$$

Let Ω denote the covariance matrix of cryptocurrencies of the portfolio so that $a = 1^T \Omega^{-1} 1$, $b = 1^T \Omega^{-1} \mu$, $c = \mu^T \Omega^{-1} \mu$. Solving for λ_1 and λ_2: $\lambda_1 = \frac{c - b\mu_p}{\Delta}$ and $\lambda_2 = \frac{a\mu_p - b}{\Delta}$, where $\Delta = ac - b^2$. Thus, the optimal weight is given by:

$$w^* = \Omega^{-1}(\lambda_1 1 + \lambda_2 \mu)$$

Secondly, for quadratic programming, Markowitz's algorithm with no short selling requires the weights of all cryptocurrencies to be positive or zero. With inequality constraints, the Lagrange multiplier method no longer works because it imposes equality in the constraint. This optimization problem must be solved numerically, e.g. using the cvxopt.solver.qp in Python or the function solve.QP() in the R package quadprog. Quadratic programming problems are of the form:

$$\min \frac{1}{2} x^T D x - d^T x \ s.t.$$

$$A_{neq}^T x \geq b_{neq}$$

$$A_{eq}^T x = b_{eq}$$

where D is an $n \times n$ matrix, x and d are $n \times 1$ vectors, A_{neq}^T is an $m \times n$ matrix, b_{neq} is an $m \times 1$ vector, A_{eq}^T is an $l \times n$ matrix, b_{eq} is an $l \times 1$ vector. Now, consider the portfolio optimization problem:

$$\min_w \sigma^2 = w^T \Omega w \ s.t.$$

$$w^T R = \mu_p$$

$$w^T 1 = 1$$

$$w_i \geq 0 \ \ \forall i$$

Then combine the inequality constraints and equality constraints into matrices separately and use Python or R to solve the inequality constrained optimization problem. Thus, the optimal weight will be computed.

4.1.2 "scipy.optimize.minimize"

The second approach is to use Python package "scipy.optimize" to solve the minimization problem regardless of whether the optimal portfolio enables short-selling or not. The default method of "scipy.optimize.minimize" is BFGS to solve both unconstrained minimization problems and constrained minimization problems. The package provides L-BFGS-B algorithm (L-BFGS-B), a modification of Powell's method (Powell), and a truncated Newton algorithm (TNC) for bound-constrained minimization, and constrained optimization BY linear approximation method (COBYLA), sequential least squares programming (SLSQP), and trust-region algorithm (trust-constr) for constrained minimization.

4.1.3 1/N model

The 1/N model is used to assign each cryptocurrency with equal weight, which may be seen as a benchmark for data-driven portfolios. The reason why the naively diversified 1/N equal-weighted portfolio is adopted is that price changes in each cryptocurrency are unpredictable, which results in the inaccuracy of calculated portfolios. Therefore, the 1/N model can also be used to compare with other portfolio models.

4.2 Data & Portfolio selection

A first portfolio is created based on the market capitalization of cryptocurrencies. At each week t, the cryptocurrencies are sorted by their market capitalization as at week t-1 and the 10 largest cryptocurrencies are chosen as constituents of the portfolio. The model is run on the 10 assets and then the investment is performed according to the weights assigned to each asset.

A second portfolio is created based on their liquidity, based on the mean dollar trading volume. At each week t, the cryptocurrencies are sorted by their liquidity as at week t-1 and the 10 most liquid cryptocurrencies are chosen as constituents of the portfolio. The model is run on the 10 assets and then the investment is performed according to the weights assigned to each asset.

The risk-free rate is equal to 0. No risk-free asset was considered, implying that at any time all funds are entirely invested in the 10 cryptocurrencies. The constraint of short-selling is considered in this research project. The data used for the Markowitz optimization model is the historical data of the assets in the portfolio for the last 52 weeks (1 year) prior to the week of investment.

4.3 Results

Portfolio	Return	Sharpe ratio	Calmar ratio	Tail ratio
Liquid_Equal	214%	1.35	2.01	1.51
MarketCap_Equal	261%	1.64	2.92	1.67
Liquid_LongAndShort	411%	1.70	5.34	1.90
MarketCap_LongAndShort	223%	1.49	1.58	1.15
Liquid_LongOnly	275%	1.49	3.13	1.36
MarketCap_LongOnly	229%	1.39	2.47	1.46
Market Return	334%	2.19	4.84	1.54

Table 1: Result of the Markowitz model porfolios after investment.

Each portfolio's name represents a type of investment. For example, MarketCap_LongAndShort represents a portfolio model which allows short-selling, with constituents based on their market capitalization whereas Liquid_Equal is a portfolio model which uses the 1/N model, with constituents based on their liquidity.

The table above shows the result of each portfolio at the end of the year 2020 to facilitate comparison. A trend observed is that when investing with the Markowitz optimization model, the Liquid portfolio always has a better return than its MarketCap counterpart. However, when investing with the 1/N model, the MarketCap portfolio is better in terms of return and all the 3 financial ratios. It has a Tail ratio of 1.67, which means that its profits are 1.67 times better than its losses. It is the one of the 2 portfolios with a better Tail ratio than the market. The MarketCap_Equal portfolio is the simplest and safest portfolio to invest in as it requires no complex mathematical knowledge, it has the second-best Sharpe ratio and Tail ratio and the third-best Calmar ratio among the 6 portfolios created.

Another trend observed is that for portfolios with constituents based on their liquidity, the portfolio allowing short-selling has the highest profit while the 1/N portfolio has the lowest return. However, for portfolios with constituents based on their market capitalization, the trend is reversed with the 1/N model having the highest return.. A probable reason may be that choosing cryptocurrencies based on their previous week's liquidity is a better measure of whether the cryptocurrencies will yield a profit.

The most important observation is that the Liquid_LongAndShort portfolio is the most profitable portfolio with a return of 411% and the only portfolio exceeding the market return of 334%. It has a good portfolio with a Sharpe ratio of 1.70 but a lower than the market portfolio's Sharpe ratio of 2.19, which is the best Sharpe ratio. The Liquid_LongAndShort portfolio also has the highest

Calmar ratio and Tail ratio, indicating that it is the best portfolio to invest in compared to the other portfolios created. Nonetheless, the fact that it has a higher return but a lower Sharpe ratio indicates that investing in this portfolio bears too much risk; the variance of the portfolio is much higher than that of the market. Furthermore, it may be of pure luck that its return was better than that of the market. If the returns of most portfolios are lower than the return of the market when it is a profit, it can be safely assumed that if the market undergoes a recession, the Markowitz model cannot be used to obtain profits. In conclusion, it is wiser to not use the Markowitz model as an investment model for the cryptocurrency market.

5 Fama-French 3 Factor Model

5.1 Solution Approach

In this research, the Fama French 3 Factor model was modified so that it can be used as an out-of-sample investment model. The Fama French 3 Factor model is usually used to explain a market by using cross-sections of that market. Each cross-section represents a portfolio, and its returns are the average returns of all the constituents of that portfolio. This research tries to find out how much of the returns can be explained, and how much of it is statistically significant.

Firstly, the 3 factors will be created. Secondly, portfolios representing the cross-section of the cryptocurrency market will be formed. Thirdly, linear regression will be performed on the portfolios using the factors to determine whether these factors have explanatory power for the returns of the portfolios. The regression will be performed on the data from 1st January 2018 to 31st December 2019. The period mentioned is used as training data to assess whether it is worthwhile to invest. If the result of the linear regression is promising, then investment will be performed on the data from 1st January 2020 to 31st December 2020, which is used as testing data to assess whether the conclusion reached in the training data can be used as a good investment strategy.

5.2 Data & Portfolio Selection

The factors of the Fama-French 3 model are created differently from the procedures of Fama and French paper (1993) to suit the data used. The procedures are explained below.

Market Factor: The market return is obtained from bitwiseinvestments.com. It is the market capitalized weighted returns of the 100 largest cryptocurrencies. [1]

Size Factor: The size factor, SMB, is the difference between the return on a portfolio of small stocks and a portfolio of large stocks. The factor was calculated manually following the same steps as Fama and French's (2012) work.[4] Among the 68 cryptocurrencies, the cryptocurrency in which the market capitalization is above the 90th percentile is considered to have large capitalization while the bottom 10th percentile is classified as small capitalization.

Value Factor: Since cryptocurrencies do not have book-to-market ratio as they are different from stocks, a new factor is needed to evaluate the value of a cryptocurrency. A good measure is Network Value to Transaction ratio, NVT.

$$NVT = \frac{market\ capitalization}{transaction\ volume}$$

The NVT ratio was conceptualized by Woo and Chris Burniske. Their main idea was that in a traditional company, its utility is given by its earnings and therefore for a cryptocurrency, since it is in essence a payment and store of the value network, its utility would be given by its ability to move money.[11] A high NVT ratio indicates that the currency is expensive in relation to its actual transaction value which can indicate market optimism or overvaluation and similarly, a low ratio indicates either a pessimistic view from the market or undervaluation. Due to its similar interpretation to the PE ratio, the NVT metric is often considered a PE for the cryptocurrency market.[10]

Among the 68 cryptocurrencies, the cryptocurrencies in which the NVT is above the 90th percentile are considered to have high value while the bottom 10th percentile is classified as low value. The value factor, NVT, is the difference between the return on a portfolio of high-value cryptocurrencies and a portfolio of low-value cryptocurrencies.

Now that the factors are already obtained, the portfolios are created. To create portfolios out-of-sample, the 68 cryptocurrencies are sorted into 5 quantiles on week t-1 and the weekly return of each portfolio is the average weekly return of all its constituents on week t. The portfolios are sorted and rebalanced weekly. The sorting is performed in descending order on their size and their value separately, as shown below.

Quantile	Size	Value
From 0th percentile to 20th percentile	SMB1	NVT1
From 20th percentile to 40th percentile	SMB2	NVT2
From 40th percentile to 60th percentile	SMB3	NVT3
From 60th percentile to 80th percentile	SMB4	NVT4
From 80th percentile to 100th percentile	SMB5	NVT5

Table 2: Portfolio creating using the size and value factors.

5.3 Results

5.3.1 Regression Results

The results of the regression of the 10 portfolios are displayed below. It is to be recalled that the regression was performed on the data from 1st January 2018 to 31st December 2019. The period mentioned is used as training data to assess whether this model is worthwhile. The regression will not be performed on the investment period so as to simulate a real investment where future data is not known.

	SMB1	SMB2	SMB3	SMB4	SMB5
Intercept	-0.001	-0.005	-0.006	-0.010**	-0.002
	(0.003)	(0.005)	(0.004)	(0.005)	(0.004)
Market Return	0.042	0.197**	0.297***	0.498***	0.628***
	(0.048)	(0.086)	(0.071)	(0.076)	(0.065)
NVT	0.241***	0.541***	0.488***	0.314***	0.328***
	(0.048)	(0.086)	(0.072)	(0.076)	(0.065)
SMB	0.685***	0.267***	0.233***	0.174***	0.121**
	(0.042)	(0.075)	(0.062)	(0.066)	(0.057)
R^2	0.946	0.838	0.882	0.853	0.902
Adjusted R^2	0.944	0.833	0.878	0.849	0.899
Residual Std. Error	0.030	0.054	0.045	0.048	0.041
F Statistic	574.591***	171.096***	246.464***	191.670***	303.072***
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table 3: Regression result on the SMB portfolios.

	NVT1	NVT2	NVT3	NVT4	NVT5
Intercept	-0.009*	-0.003	-0.004	-0.003	-0.003
	(0.005)	(0.005)	(0.004)	(0.005)	(0.002)
Market Return	0.521***	0.393***	0.417***	0.208***	0.117***
	(0.084)	(0.080)	(0.069)	(0.073)	(0.039)
NVT	0.260***	0.234***	0.288***	0.426***	0.689***
	(0.084)	(0.080)	(0.069)	(0.073)	(0.039)
SMB	0.271***	0.409***	0.306***	0.346***	0.168***
	(0.073)	(0.070)	(0.060)	(0.064)	(0.034)
R^2	0.843	0.858	0.884	0.873	0.961
Adjusted R^2	0.838	0.854	0.881	0.869	0.960
Residual Std. Error	0.053	0.050	0.043	0.046	0.025
F Statistic	177.460***	200.106***	252.093***	225.926***	811.983***
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 4: Regression result on the NVT portfolios.

Since the R^2 for all the portfolios is over 0.80, it indicates that more than 80% of the returns can be explained by the 3 factors. The highest R^2 is 0.96 for portfolio NVT5. 9 out of 10 of the portfolios have statistically significant market return coefficients while all the NVT and SMB coefficients are statistically significant. All the intercepts are close to 0. In summary, it can be concluded that the 3 factors have an influence on the whole cryptocurrency market.

5.3.2 Investment Results

Since it was observed that the 3 factors have an influence over the cryptocurrency market, the trend in the portfolios sorted on the 2 factors must be observed. The portfolios are invested in the year 2020 using data from 1st January 2020 to 31st December 2020.

Portfolio	Return	Sharpe ratio	Calmar ratio	Tail ratio
SMB1	827%	2.75	13.01	1.91
SMB2	354%	1.96	4.72	1.03
SMB3	257%	1.53	2.72	1.27
SMB4	251%	1.51	2.75	1.51
SMB5	316%	1.87	3.75	1.51

Table 5: Result of SMB portfolios after investment.

Portfolio	Return	Sharpe ratio	Calmar ratio	Tail ratio
NVT1	253%	1.51	2.56	1.57
NVT2	263%	1.60	3.18	1.18
NVT3	456%	2.28	6.10	1.18
NVT4	529%	2.27	7.61	1.50
NVT5	408%	2.24	5.69	1.44

Table 6: Result of NVT portfolios after investment.

From Table 1, it is stated that the market return for the year 2020 was 334%. The assumptions of Fama and French (1993) apply also to the cryptocurrency market as it can be observed that the small-cap portfolio, SMB1, and the high-value portfolio, NVT5, outperformed the market while the large-cap portfolio, SMB5, and the low-value portfolio, NVT1, underperformed relative to the market in terms of return.

Firstly, the trend in the portfolios sorted using market capitalization will be observed by looking at the Table 5. The trend observed in the returns of the portfolios is that it decreases from SMB1 to SMB4 and it increases at SMB5. The SMB5 seems to not follow the trend observed in the other 4 portfolios. The portfolio SMB1 return far outperforms the other portfolios. Its Calmar ratio of 13.01 indicates that it is a much safer portfolio to invest in. Moreover, it has the best Sharpe ratio and Tail ratio, indicating that it is the best portfolio among the portfolios sorted using the size factor.

Secondly, the trend in the portfolios sorted using NVT will be observed. The trend observed in the returns of the portfolios is that it increases from NVT1 to NVT4 and it decreases at NVT5. The NVT5 seems to not follow the trend observed in the other 4 portfolios. The NVT4 portfolio outperforms the other NVT portfolios with a return of 529%, a Calmar ratio of 7.61 and a Tail ratio of 1.50. However, it does not have the best Sharpe ratio since the portfolio NVT4 has the best Sharpe ratio of 2.28.

A highly unusual observation is that NVT5 and SMB5 do not follow the trend observed in the other 4 portfolios in their respective tables. The return of NVT5 has a value in between the returns of NVT2 and NVT3, and the return of SMB5 has a value in between the returns of SMB2 and SMB3. A likely conclusion is that the cryptocurrency market is different from the stock market, therefore there are some fundamental differences not yet researched since it is a relatively newer market.

Among the 10 portfolios, the portfolio SMB1 outperforms the rest by far with a return more than twice of the market's return and an extraordinary Calmar ratio far superior to the other 9 portfolios.

6 Long Short Term Memory

6.1 Solution Approach

The Keras interface of TensorFlow 2 makes it very straightforward to build an RNN with two hidden layers with the following specifications:

- Layer 1: An LSTM module with 10 hidden units with input_shape = (window_size,1)
- Layer 2: A fully connected module with a single unit and linear activation

The loss function is set as mean squared error(MSE) to match the regression objective and optimizer is RMSProp which is recommended for RNN with default settings.

6.2 Data & Portfolio Selection

Firstly, the collected raw daily data is preprocessed via cleaning (processing missing data and noisy data), transformation (scaling and normalization), and splitting (splitting the data into two sets: training set (January 2017 – December 2019) and testing set(January 2020 – December 2020).

Secondly, the model training part comprises model selection and parameter configuration. The LSTM model set the window size equal to 90, which means only the previous 90 days' data will be treated as one sequence and that the return prediction of the day $T + 1$ should be based on the data of day $T - 90$ to T.

Thirdly, the well-trained model is going to predict the return in the test set. Furthermore, the predictions are supposed to be postprocessed by rescaling to the original numerical form.

Fourthly, this project introduced two strategies which are long-only strategy and long-short strategy. To construct them, for each day T after January 2020, the top 10 cryptocurrencies and their return predictions are used to calculate the optimal portfolio on that day by applying Markowitz's mean-variance model. However, there is a special condition for the strategy with only long positions that if the number of positive predicted returns of the top 10 cryptocurrencies on the day $T + 1$ is less than 5, implying a bad market on the day $T + 1$, the strategy will enter an empty position (without holding any cryptocurrency) that day.

6.3 Results

During the strategy backtesting of 2020, both of these two strategies (One with long-only position and the other with long-short position) achieved high returns with relatively low risks and outperformed the cryptocurrency market largely. The overall performance of the two strategies is showed and outlined in the following pictures including historical capital changes, returns, as well as some financial and statistical indicators of the two strategies.

As shown below, the capital of the strategy with only long position steadily increased from January to August, and then went into a vibration phase from the beginning of September up to the end of the test set, while the capital of the strategy with long and short positions had the same trend with that of the before mentioned strategy from the Jan to November and then rocketed up suddenly in December. Reasons why the two strategies did not perform well from August to November may include that there is no much training data (historical patterns) that resembles trends of this period and that the cryptocurrency market vibrated too much without an obvious trend during this period, resulting in the LSTM model's poor predictions. Furthermore, as the market itself rise hugely at the end of 2020, the strategy with long and short positions could easily take advantage of this market to buy low and sell high.

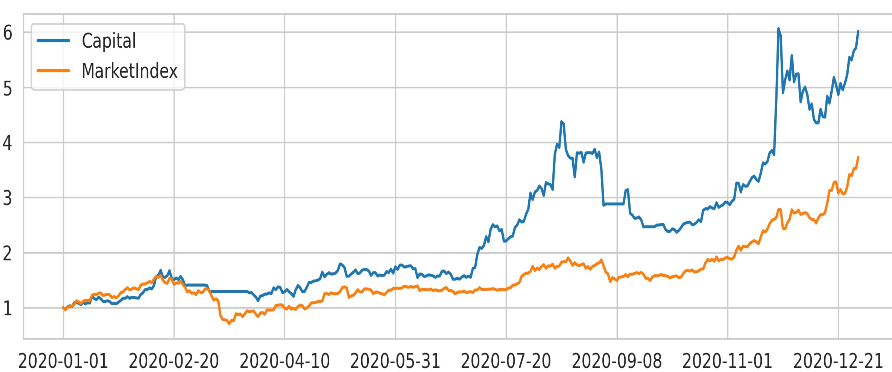

Figure 4: Capital line of long-only strategy

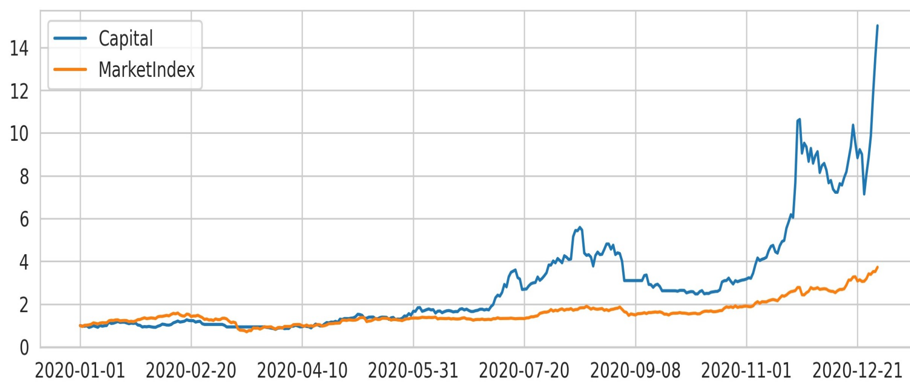

Figure 5: Capital line of long-short strategy

Compared to the long-only strategy, the historical returns of the long-short strategy are more volatile and there are several anomalies due to some wrong predictions of returns. Although the market suffered a huge decrease with a daily return of almost -30% in March 2020, both strategies were able to prevent such loss effectively since they successfully predicted a bad market and then entered into empty position with no cryptocurrency held.

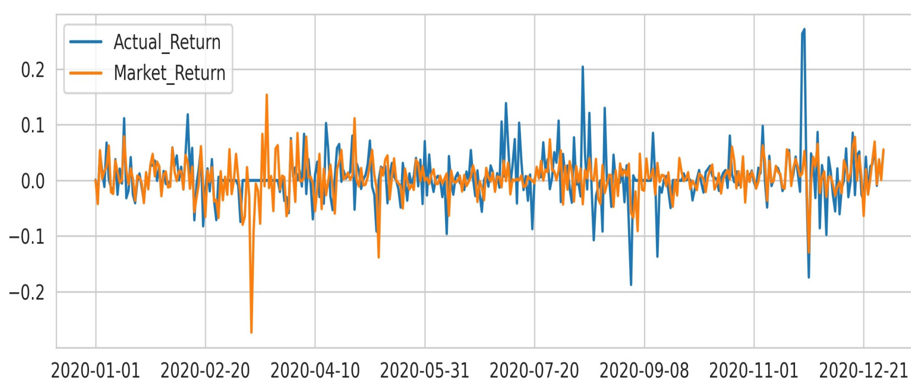

Figure 6: Daily returns of long-only strategy

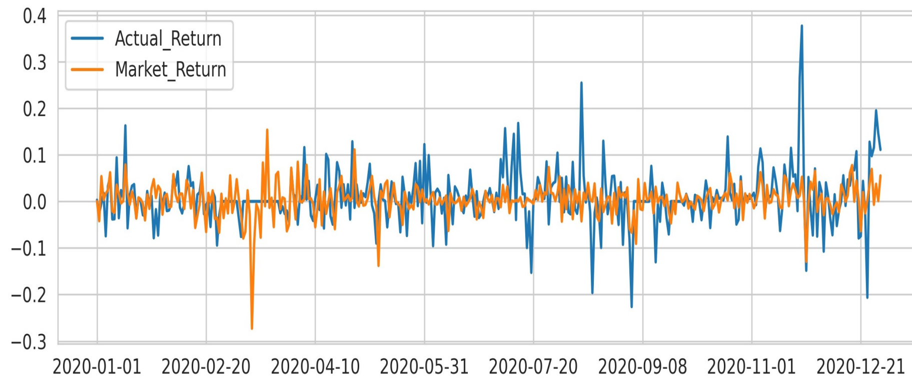

Figure 7: Daily returns of long-short strategy

The following two pictures display the overall performance of each strategy. The strategy with only long positions realized a return almost 2 times as much as the market return, while the strategy with long and short positions realized a return more than 5 times as much as the market return. Both strategies experienced a smaller maximal drawdown compared to the market index, however, each strategy's annual volatility is larger.

As for the financial and statistical indicators, both strategies achieved similar Sharpe Ratio of slightly larger than 2 and much higher Calmar Ratio compared to the market. As both of the ratios are financial metrics used to measure risk-adjusted returns, it is well informed that these two strategies are amazingly performed and pretty attractive to investors. And in terms of alpha, the strategy with long and short positions achieves quite a large alpha which is over 6.5. However, from the tail ratios of both strategies, there are fat-tailed distributions with some anomalies just as the paper mentioned before.

	Long-Only Portfolio	Long-Short Portfolio	Market
Return	502%	1401%	272%
Max Drawdown	-46%	-56%	-58%
Annual Volatility	0.73	0.95	0.57
Sharpe Ratio	2.08	2.46	1.50
Calmar Ratio	5.48	10.2	1.69
Alpha	2.60	6.85	0.00
Beta	0.28	0.32	1.00
Tail Ratio	1.43	1.48	1.04

Table 7: Results of the strategies and the market.

In summary, the long-only strategy and the long-short strategy performed well in the backtesting of the year 2020, though there are still some drawbacks that need to be changed and improved.

7 Discussion and Future Plan

Currently, Markowitz model and Fama French 3 factor model do not perform well as shown from the backtesting results. One possible explanation is the cryptocurrency market is highly volatile and unpredictable, and the classical models are not able to capture those risks and potential changes.

In the Markowitz model and LSTM model, when the portfolio models do not allow short-selling, i.e. when the portfolio can only long the cryptocurrencies, the weights allocated to each cryptocurrency are not distributed appropriately; the portfolios constructed by such models usually only focus on three to four cryptocurrencies, and do not consider the rest. One possible explanation for this could be that some cryptocurrencies dominantly outperform the others, providing higher returns while having a lower risk level. For future research, more cryptocurrencies could be included, and inefficient ones should be eliminated.

Since this project reallocates the portfolio on a regular basis, transaction fees are supposed to be taken into account. And for the strategy with both long and short positions, the margin part has not been considered. In future research, information and experiments about high frequency transaction fees and margin account need to be studied and carried out.

8 Conclusion

This research project introduced Markowitz model, Fama French 3 factor model and Long Short Term Memory model separately to construct investment strategies using 68 cryptocurrencies. And results measured by a set of statistical and financial indicators such as Sharpe Ratio and Calmar Ratio showed that all the strategies outperform the market. Moreover, LSTM model has the best performance and the LSTM strategy with long and short positions is the best one with an annualized return almost 10 times as much as the market return. In addition, there are still some challenges and improvements to be overcome and made in future research as mentioned in the above section.

References

[1] Bitwise. *Cryptocurrency Index and Beta Fund Provider.* URL: https://www.bitwiseinvestments.com/indexes/Bitwise-100.

[2] Alexander Brauneis and Roland Mestel. "Price discovery of cryptocurrencies: Bitcoin and beyond". In: *Economics Letters* 165 (2018), pp. 58–61. ISSN: 0165-1765. DOI: https://doi.org/10.1016/j.econlet.2018.02.001. URL: https://www.sciencedirect.com/science/article/pii/S0165176518300417.

[3] Eugene F. Fama and Kenneth R. French. "Common risk factors in the returns on stocks and bonds". In: *Journal of Financial Economics* 33.1 (1993), pp. 3–56. ISSN: 0304-405X. DOI: https://doi.org/10.1016/0304-405X(93)90023-5. URL: https://www.sciencedirect.com/science/article/pii/0304405X93900235.

[4] Eugene F. Fama and Kenneth R. French. "Size, value, and momentum in international stock returns". In: *Journal of Financial Economics* 105.3 (2012), pp. 457–472. ISSN: 0304-405X. DOI: https://doi.org/10.1016/j.jfineco.2012.05.011. URL: https://www.sciencedirect.com/science/article/pii/S0304405X12000931.

[5] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735. eprint: https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf. URL: https://doi.org/10.1162/neco.1997.9.8.1735.

[6] Thomas Kim. "On the transaction cost of Bitcoin". In: *Finance Research Letters* 23 (2017), pp. 300–305. ISSN: 1544-6123. DOI: https://doi.org/10.1016/j.frl.2017.07.014. URL: https://www.sciencedirect.com/science/article/pii/S1544612317301897.

[7] Harry Markowitz. "Portfolio Selection". In: *The Journal of Finance* 7.1 (1952), pp. 77–91. ISSN: 00221082, 15406261. URL: http://www.jstor.org/stable/2975974.

[8] Saralees Nadarajah and Jeffrey Chu. "On the inefficiency of Bitcoin". In: *Economics Letters* 150 (2017), pp. 6–9. ISSN: 0165-1765. DOI: https://doi.org/10.1016/j.econlet.2016.10.033. URL: https://www.sciencedirect.com/science/article/pii/S0165176516304426.

[9] Christopher Olah. *Understanding LSTM Networks.* Aug. 2015. URL: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.

[10] Ryan Vlastelica. *Is bitcoin in a bubble? This metric suggests there's more room to grow.* URL: https://www.marketwatch.com/story/is-bitcoin-in-a-bubble-this-metric-suggests-theres-more-room-to-grow-2017-06-08.

[11] Woo Willy and Chris Burniske. *Introducing NVT Ratio (Bitcoin's P/E Ratio), use it to detect bubbles.* URL: `https://woobull.com/introducing-nvt-ratio-bitcoins-pe-ratio-use-it-to-detect-bubbles/#fn1`.