Fuzzy Logic and Genetic Algorithm in Crude Palm Oil Futures Trading

Aswathy Leelakumari (A0178440A), Bharat Nagaraju (A0178258N), Koo Chee Kit (A0178200N),
Nandana Murthy (A0178563R), and Siew Yao Hoong (A0178544U)

I. Abstract

The Moving Average strategy is one of the most popular trading indicators that help traders in validating market movement and thus increasing profitable trades. Our system proposes an improved version of Fuzzy Moving Average strategy[4], by integrating Genetic Algorithm and Fuzzy Logic based on a combination of technical indicators. We are utilizing the Crude Palm Oil (FCPO) Futures trade stock and volume data from Bursa Malaysia to train and test the system such that an optimal trading rule set obtained from Fuzzy system is then applied by Genetic algorithm to implement trading decisions.

II. Introduction/Background/Objectives

Trading has three components: quantity, entry cost and exit price. Keeping in mind that trading is a probability game, it is important to maximize gains on the winners so that the inevitable losers can be overcome. This is possible when right volume of futures are bought at the right time. Though popular technical indicators such as MACD, RSI and Moving average Strategy help traders understand the prevailing trend and foresee trend reversals, they provide little or no information on the volume of futures to be bought. In this study, we use fuzzy logic to determine the strength of the trading signals and thus calculate trading volume and use GA to generate and select a best set of trading rules so as to maximize gains. We have used Crude Palm Oil (FCPO) Futures trade stock and volume data from Bursa Malaysia for our analysis.

III. System Design, Modeling & Simulation

This System consists of three components.
i) Trading
ii) Genetic Algorithm
iii) Fuzzy Logic
Trading system is integrated with GA and Fuzzy modules to read the incoming rates and make Buy / Sell / Hold decision and if it is Buy / Sell – Also recommends the number of trades.

We have been given FCPO Futures Trade stock price and trade volume for 6 years from 2011 – 2016 for 6 years. We split this into training and test data.
Training Steps:
1. Stock prices and volumes from 2011-2013 is taken for training. It is again split into 6 equal parts and each part holds data for approximately 6 months.

2. 6 months of data (As explained below in the data section) containing open price, close price, date, Cash balance is passed to Genetic Algorithm module.
3. Genetic Algorithm generates a set of trading rules and selects the best set of rules based on a fitness function containing various statistical indicators such as MACD – SMA, TMA, etc, RSI and passes a set of rules to Fuzzy module to choose the best fit. The best set of rules is returned to GA and then further to the trading module
4. After finishing up with the first set of data, the system is looping in to read the next set of data.
5. Meanwhile, A system date is maintained and after reading up of current day's trading prices that are received (to simplify a soft update in the pandas data frame, the "isAvailable" parameter is marked as 1 after the training / trading is performed) and the system date is incremented to next date for which prices of stock is available
6. Once the prices are processed for the next day, Trading module passes the prices to Genetic Algorithm to make Buy / Sell decisions.(Internally, GA module calculates the asset value if the trade is executed in combination with fuzzy). Once the buy / sell / hold decision is made – the trade is considered executed and a separate dataframe will be updated with the latest holdings, cash balance , etc.
7. Step 6 (Previous Step) is repeated till the end of cycle , approximately for the next 6 months of data.
8. This back testing provides an idea of how the strategy is performing and parameters can be fine-tuned
9. Then the system retrains using GA Evolution and prepares the next best set of rules. The whole process is repeated till all the data from 2011-2013 is consumed and the system is considered to have completed the training cycle
10. From 2014 Jan 1, The system receives prices every day and Based on GA / Fuzzy Module – Trades till 2016 Dec 31
11. Since the system is an online system, the system is retrained at the end of every 5 weeks to ensure the technical analysis trend is caught up.

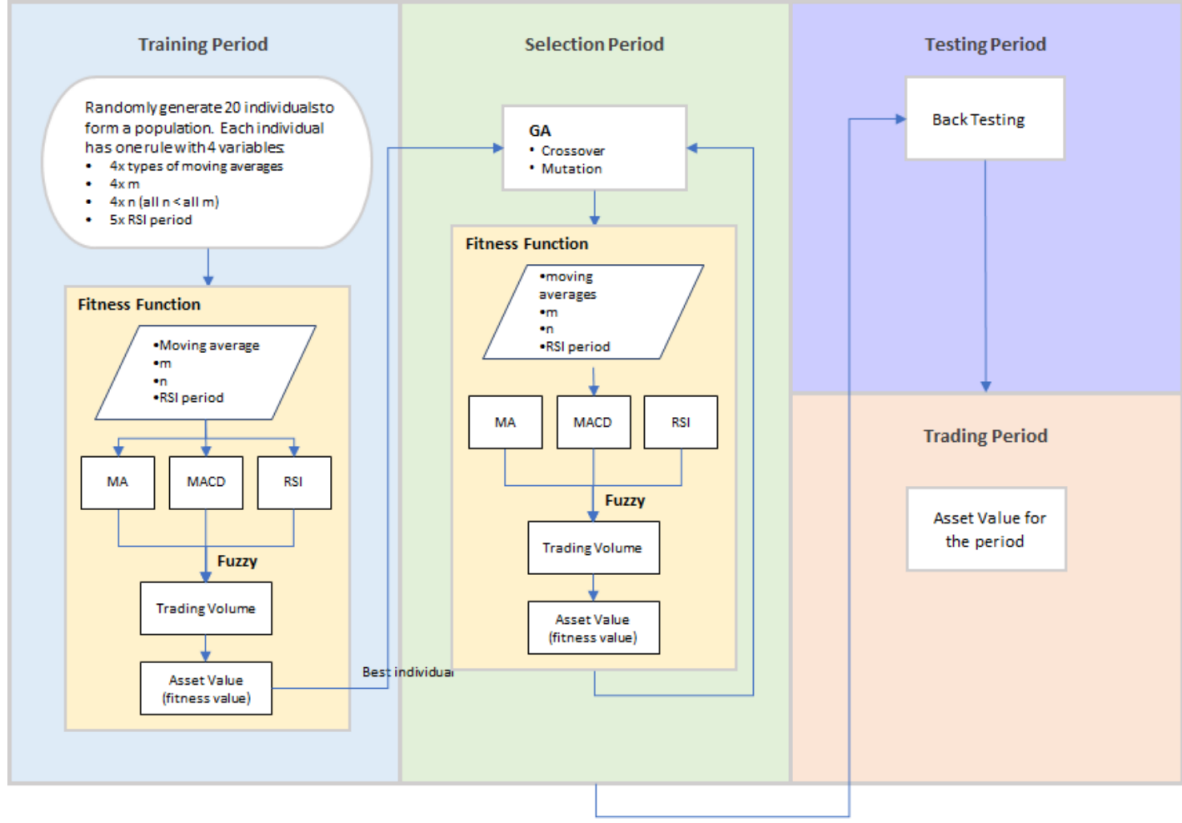|	Use 2011 data for	Use 2012 data for	Use 2013 data for	Use 2014 data for	Use 2015 data for	Use 2016 data for
Period 1 timenow = 2014	Training	GA Selection	Back Testing	Live Trading		
Period 2 timenow = 2015		Training	GA Selection	Back Testing	Live Trading	
Period 3 timenow = 2016			Training	GA Selection	Back Testing	Live Trading

Fig. 1. Process Flow Diagram

IV. DATA AND METHODS

A. Data

The original data given contains 6 years of trading information in one-minute interval, and there are over 350,000 rows of data in 9 columns. The visualization of the data can be seen in Figure (2).

In our project we made some modifications to the data and assumptions to the methodology:

- We reduced the per minute data to per day to make it more manageable computationally
- FCPO futures are usually in series and have a 60-day expiry date. For the purpose of this exercise we assumed the futures to be a single counter with no expiry dates
- Actual future trading requires deposit payment upfront. For this project we assumed full and immediate settlement during transaction
- The trading prices given in the data is in Malaysian ringgit, and the unit is one metric ton of palm oil. Since there are 25 metric tonnes of palm oil in each FCPO contract, each minimum trade of one contract will involve trading prices x 25.

B. Trading Rules and Indicators

We have used a combination of Technical Indicators in our system which Investors generally use while trading. These indicators are interpreted in the same way as any trader would do to decide whether to buy futures or sell them.

1) RSI: RSI stands for Relative Strength Indicator. In general the RSI is interpreted as follows;

1) If the indicator is below 30, then the price action is considered weak and possibly oversold.
2) If it is reading above 70, then the asset is after a strong uptrend and could be overbought.

We came up with the following rules with RSI indicator.

- If RSI≤30 then BUY
- If RSI≥70 then SELL

2) MACD: Moving average convergence divergence (MACD) is an extremely popular indicator used in technical analysis. It is a trend-following momentum indicator that shows the relationship between two moving averages of prices. Most notably these aspects are momentum, as well as trend direction and duration. [1]

First, MACD employs two Moving Averages of varying lengths (which are lagging indicators) to identify trend direction and duration. Then, it takes the difference in values between those two Moving Averages (MACD Line) and an EMA of those Moving Averages (Signal Line) and plots that difference between the two lines as a histogram which oscillates above and below a centre Zero Line. [2]

For example:

MACD Line: (12-day EMA - 26-day EMA)
Signal Line: 9-day EMA of MACD Line
MACD Histogram: MACD Line - Signal Line
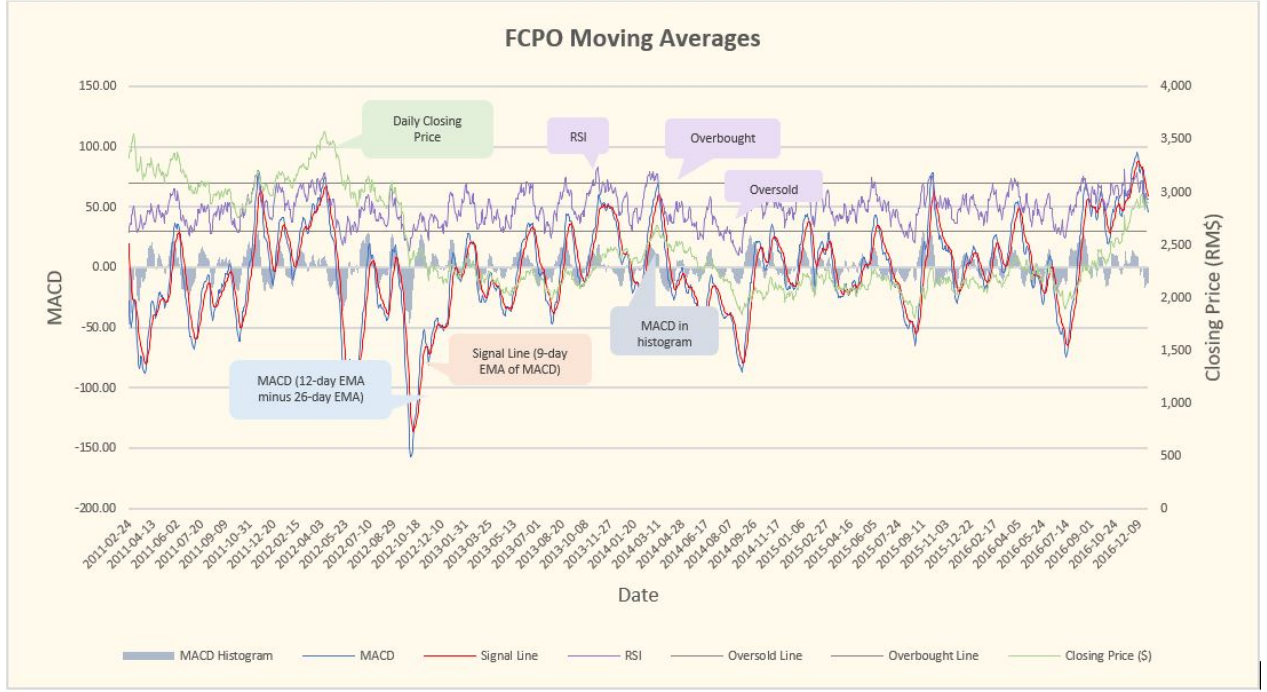Trading rules for MACD are as below:

Fig. 2. FCPO price range from 2011-2016 with MACD and RSI overlays

- if MACD Line-Signal Line > 0 then BUY
- if MACD Line-Signal Line < 0 then SELL

3) MA: Moving Averages is another popular technical indicator. It filters out the noise / random fluctuations from the time series. The moving average strategy uses the relative positions of a long term moving average line and a short term moving average line to estimate market trend. If the short-term average line is above the long term average line, then it suggests that the market trend is increasing. If otherwise, then the market trend is decreasing.

With this, we came up with two trading rules as below.

1) If short term MA – long term MA > 0 then SELL
2) If short term MA – long term MA < 0 then BUY

We have used four types of moving averages in our system.

a) SMA: Simple Moving Average calculates the mean prices of a period.

$$SMA(k) = \frac{1}{n}\sum_{i=1}^{n} c_{k-i}.$$

Where n indicates the period length,
c daily closing prices of futures
k day for which SMA is calculated

b) TMA: Triangular Moving Average is the average of the averages prices of a given period. In short, TMA is the average of SMA's of a given period. Its calculated as below.

$$TMA(k) = \frac{1}{n}\sum_{i}^{n} SMA(k-i).$$

Where n indicates the period length,
c daily closing prices of futures
k day for which SMA is calculated

c) TPMA: TPMA stands for Typical Price Moving Average. It calculates the average of the maximum closing price recording in a given period, minimum closing price recorded in the same period and the closing price on the previous trading day.

$$TPMA(k) = (high + low + close)/3$$
$$high = \max(c_{k-1}, c_{k-2}, \dots, c_{k-n}), \qquad low = \min(c_{k-1}, c_{k-2}, \dots, c_{k-n}), \qquad close = c_{k-1}.$$

Where n indicates the period length,
c daily closing prices of futures
k day for which SMA is calculated

d) EMA: Exponential Moving Average: it reduces the lag by applying additional weights on recent closing prices. EMA's are calculated in three steps as below:

1) Calculate the SMA for a time period.
2) Calculate the multiplier for weighting the EMA. [Symbol] [2 ÷ (time period + 1)]
3) Calculate the current EMA. [Symbol] [Closing price-EMA (previous day)] x multiplier + EMA (previous day)
   a) On Day 1, the initial EMA is considered as same as initial SMA

C. Fuzzy Logic

The fuzzy logic in the system is embedded in the fitness calculation function. When an individual with 4 rules (Types of Moving Averages, Long Period m, Short Period n, RSI period) is passed to the fitness function, the fitness function will calculate trading indices such as Moving Averages (SMA, EMA, TPMA and TMA), MACD, and RSI. As these 3 indices

may recommend different trading rules, we run them through a set of fuzzy rules to arrive at a single crisp trading volume.

The membership functions are given in Figure 5 , and the trading rules Figure 6 .

The sub goal range is given in Figure (4) . The range for MA and MACD are determined from looking at their distribution (Fig 5a and 5b), while RSI range was determined by the original definition of RSI, where below 30 means the contract is oversold (thus possibly a good time to buy) and higher than 70 means overbought (good time to sell).

From these sub-goal range, we construct the membership functions of the 3 variables. For Moving Averages the negative limit was set to -1000 as tests show that the moving averages difference can be quite large, especially when the long period is big and short period is small. The trading volume is in integer in the range of -10 to 10. Negative means a sell call, positive a buy call, and the integer represent the number of contract to trade.

We then constructed the trading rules from these 3 indices. In general, big MA and MACD, and small RSI mean buying opportunities, and vice versa.
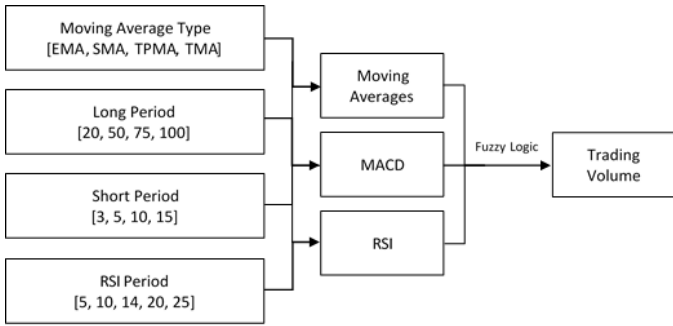

Fig. 3. Fuzzy Logic flow

Input	Low	Medium	High
MA	< -40	-60 - 60	> 40
MACD	< -8	-15 - 15	> 8
RSI	< 40	30 – 70	> 60

Fig. 4. Sub-Goal Logic and Rules

D. Genetic Algorithm

Our Genetic Algorithm (GA) implementation followed the paper ***Quantified moving average strategy of crude oil futures market based on fuzzy logic rules and genetic algorithms, section 2.4 GeneticAlgorithm [3]*** as closely as possible.

The chromosome is made up of [MA Method, M value, N value, RSI period]. One individual is encoded by one chromosome, each individual consists of rule sets. Genetic Algorithm generates the best set of rules from each individual population based on a fitness function. The fitness value is calculated via Fuzzy and the asset value is returned which proceeds to trading module.

Chromosome = [MA Method, M value, N value, RSI period]

*Population = 10 * Individual*

Our program population by default consists of 10 Individuals. An example of the initial random population generated is provided below. To evolve population, the control parameters are tuned by setting the new generation to 50 with a crossover probability of 0.7 and mutation probability of 0.3. All control parameters can be changed in the program.

['SMA', 100, 3, 20],
['SMA', 20, 10, 14],
['EMA', 50, 15, 25],
['EMA', 20, 10, 20],
['TMA', 75, 10, 5],
['SMA', 75, 15, 10],
['TPMA', 20, 10, 25],
['TPMA', 20, 10, 5],
['EMA', 75, 5, 10],
['TMA', 20, 5, 20]

We are using a computational framework named Distributed Evolutionary Algorithms in Python [4] for aiding the GA evolution strategies. For each generation, selection crossover and mutation is performed and a hall of fame of best individuals is taken to the next generation. Below is the pseudocode for the evolution procedure. Detailed program can be referred in https ://github.com/IssKeCi2Ca/GaFuzzySystem/blob/master/ga.py.

```
1  START
2  Generate the initial population
3  Compute fitness
4  Add previous best population
5  REPEAT
6    Selection
7    Crossover
8    Mutation
9    Compute fitness
10 UNTIL population has converged
11 STOP
```

Code Snippet 1. GA Strategy Pseudocode

E. Sentiment Analysis

Clearly this method does not consider the market situations and solely relies on technical and statistical analyses to predict the financial market. So, We consider using the information from news sources such as Yahoo Finance to obtain the latest news. Based on the latest news, Content is processed – tokenized, lemmatized, Filtered with Stop words and the polarity of the news is obtained. Based on the polarity of the Content – We either reduce or increase the intensity of buy signal – so that we don't solely rely on technical analysis parameters such as MACD , RSI, etc.

V. FITNESS CALCULATIONS

The fitness calculation function is called whenever we need to evaluate the fitness value of an individual. As the fitness value is the asset value at the end of the trading period, the fitness function is also used in executing the trade.
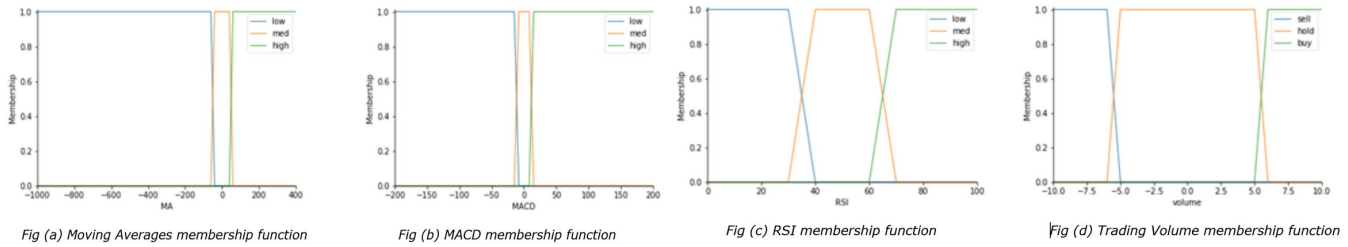
Fig (a) Moving Averages membership function
Fig (b) MACD membership function
Fig (c) RSI membership function
Fig (d) Trading Volume membership function

Fig. 5. Membership Function Graphs

MACD				
High	Buy	Hold	Sell	
Med	Hold	Sell	Sell	
Low	Sell	Sell	Sell	
	Low	Med	High	RSI

MA Low

MACD				
High	Buy	Buy	Hold	
Med	Buy	Hold	Sell	
Low	Hold	Sell	Sell	
	Low	Med	High	RSI

MA Medium

MACD				
High	Buy	Buy	Buy	
Med	Buy	Buy	Hold	
Low	Buy	Hold	Sell	
	Low	Med	High	RSI

MA High

Fig. 6. Trading Volume Rules

Relative Strength Indicator (RSI)	
Parameters	Values
RSI Period	14
RSI Buy Trigger	30
RSI Sell Trigger	70
Profit Earned	RM 855,690

Moving Average Convergence Divergence (MACD)	
EMA Weights	Values
9-day	0.2
12-day	0.15384615
26-day	0.07407407
Profit Earned	RM 11,660

Simple Moving Average (SMA)	
Parameters	Values
SMA Buy Trigger	0
SMA Sell Trigger	0
Profit Earned	RM 280,880

Fig. 7. Benchmarking Results

Once the trading volume is generated, it's executed via the following logic:

If trading volume > buy trigger

If there's enough cash, we buy the entire recommended volume

If there isn't enough cash, we buy what our cash can afford (no borrowing)

If trading volume < sell trigger

If there's enough contract at hand, we sell the entire recommended volume

If there isn't enough contract at hand, we sell the contract that we have

Else we hold

A sample outcome of the trade is attached at file 'dfFit_2011-2012_SMA_50_10_25.csv'

VI. BENCHMARKING

In order to have a measurement of the quality of our system, we performed live trading using each of the technical indicators used in our system individually without Fuzzy logic. We performed live trading for the period 03 Jan 2014 to 30 Dec 2016 using technical indicators such as RSI, MACD and SMA.

The parameters used and the results achieved are in Figure (7). For details, please refer to "Benchmarking" sheet in FCPO_Pivot.excel.

VII. RESULTS AND DISCUSSION

The results from the code for the period 2014-2016 run from the folder **/Fitness Calculation and Results Code** yields profit.

The total asset value is 1201924.37. Please note that GA / Fuzzy results may vary with execution since the approach does not yield the global maxima. The audit file containing the holdings and other details has been included under the same folder with the name - **dfFit_2014-2016_TMA_20_10_10.xlsx**. Also these results do not include the logic of sentiment mining and the codes for the GA/ Fuzzy / Sentiment Mining has been included under a separate directory (/**Main GA Fuzzy pycodes**)

The screenshot of the same has been included in Figure (8).

VIII. LIST OF ABBREVIATIONS

- RSI — Relative Strength Indicator
- MACD — Moving Average Convergence Divergence
- SMA — Simple Moving Average
- EMA — Exponential Moving Average

	MAMethod	MValue	NValue	RSIperiod	Fitness
0	EMA	50	5	25	9.337921e+06
1	TMA	20	5	5	9.988931e+06
2	EMA	50	5	5	9.559802e+06
3	TPMA	100	3	14	8.747421e+06
4	SMA	75	15	14	9.514080e+06
5	EMA	20	10	5	1.124262e+07
6	TMA	20	10	10	1.201929e+07
7	EMA	75	10	10	9.492351e+06
8	EMA	75	5	25	9.130922e+06
9	TMA	100	15	25	1.010834e+07
10	EMA	100	15	10	9.981567e+06
11	TPMA	75	3	5	9.484772e+06
12	TPMA	100	10	10	9.361637e+06
13	TPMA	50	3	10	9.143982e+06
14	TMA	100	10	14	1.033788e+07
15	EMA	20	5	20	9.244818e+06
16	EMA	50	15	5	1.005935e+07
17	SMA	100	15	5	9.729693e+06
18	EMA	75	10	25	9.440402e+06
19	TMA	100	15	5	1.007954e+07

Random individual with best fitness in year 2014 - 2016

	MAMethod	MValue	NValue	RSIperiod	Fitness
6	TMA	20	10	10	12019294.37

Fig. 8. Results

- TPMA — Typical Price Moving Average
- TMA — Triangular Moving Average
- GA — Genetic Algorithm

REFERENCES

[1] [Online]. Available: https://www.eatsleeptrade.net/technical-indicators-macd

[2] [Online]. Available: https://www.tradingview.com/wiki/MACD_(Moving_Average_Convergence/Divergence)

[3] X. Liu, H. An, L. Wang, and Q. Guan, "Quantified moving average strategy of crude oil futures market based on fuzzy logic rules and genetic algorithms," *Elsevier*, 2017. [Online]. Available: www.elsevier.com/locate/physa

[4] "Distributed Evolutionary Algorithms in Python." [Online]. Available: https://github.com/DEAP/deap