

Optimasi Portofolio Berbasis Prediksi *Return* Saham Menggunakan *Hybrid XGBoost dan Improved Firefly Algorithm* untuk Saham – Saham dalam Indeks LQ45

1st Muhammad Ridho Saputra

Fakultas Informatika

Universitas Telkom

Bandung, Indonesia

ridhosaputra@student.telkomuniversity.ac.id

2nd Deni Saepudin

Fakultas Informatika

Universitas Telkom

Bandung, Indonesia

denisaepudin@telkomuniversity.ac.id

Abstrak—Portofolio merupakan kumpulan aset investasi yang dimiliki oleh suatu perusahaan atau perorangan. *Return* yang maksimal dengan *risk* yang minimum adalah harapan bagi setiap investor. Namun dalam memprediksi pergerakan harga saham, kenaikan dan penurunan harga saham dari waktu ke waktu sulit ditebak. Oleh karena itu, *machine learning* digunakan untuk menjadi jalan keluar dalam mempelajari dan memprediksi data saham tersebut. Pada tugas akhir ini, model prediksi *return* saham menggunakan metode *XGBoost* dan *Improved Firefly Algorithm*. Hasil prediksi tersebut selanjutnya digunakan untuk pertimbangan membangun portofolio. Prediksi *return* saham yang lebih tinggi dari *threshold* akan masuk ke dalam portofolio. Dari saham yang sudah terseleksi, portofolio dibangun dengan menggunakan *Equal Weight (EW)*. Portofolio dengan dan tanpa optimasi kemudian dibandingkan untuk mengetahui *mean return*, standar deviasi, dan *Sharpe Ratio* tertinggi. Hasil eksperimen menunjukkan bahwa kinerja portofolio 7 saham dengan mempertimbangkan prediksi memiliki hasil terbaik apabila dibandingkan dengan portofolio dengan tanpa mempertimbangkan prediksi dengan *mean return* 0.0029, standar deviasi 0.0158, dan *Sharpe Ratio* 0.1837.

Kata kunci—portofolio, prediksi *return* saham, *xgboost*, *firefly algorithm*, LQ45

Abstract—A portfolio is a collection of investment assets owned by a company or individual. Maximum return with minimum risk is the hope for every investor. However, in predicting stock price movements, stock price increases and decreases from time to time are difficult to predict. Therefore, machine learning is used to be a way to study and predict stock data. In this final project, the stock return prediction model uses the *XGBoost* method and the *Improved Firefly Algorithm*. The prediction results are then used for consideration in building a portfolio. Predicted stock returns that are higher than the threshold will enter the portfolio. From the stocks that have been selected, the portfolio is constructed using *Equal Weight (EW)*. Portfolios with and without optimization are then compared to find the highest mean return, standard deviation, and *Sharpe Ratio*. The experimental results show that the performance of the 7 stock portfolio by considering the predictions has the best results when compared to the portfolio without considering the predictions with a mean return of 0.0029, standard deviation of 0.0158, and *Sharpe Ratio* of 0.1837.

Keywords—portfolio, stock return prediction, *xgboost*, *firefly algorithm*, LQ45

A. PENDAHULUAN

A. Latar Belakang

Portofolio merupakan kumpulan aset investasi yang dimiliki oleh suatu perusahaan atau perorangan. Dalam melakukan investasi tentunya investor memiliki tujuan untuk memaksimalkan *return* dan meminimalkan *risk* yang didapat. Untuk meraih tujuan tersebut, dapat dilakukan dengan beberapa cara, salah satunya dengan melakukan optimasi portofolio.

Dalam pasar saham, optimasi portofolio merupakan salah satu topik yang menarik untuk diperbincangkan. Hal itu terjadi karena kenaikan dan penurunan saham dari waktu ke waktu sangat sulit diprediksi. Dengan melakukan optimasi portofolio, investor dapat memprediksi *return* saham sehingga mempermudah dalam mengambil keputusan apakah saham perlu dijual atau lebih baik dipertahankan. Banyak model portofolio yang digunakan untuk melakukan optimasi portofolio dan salah satu yang mudah dipahami adalah *Equal Weight (EW)*.

Pada tugas akhir ini, akan dilakukan prediksi *return* saham dalam indeks LQ45 dengan menggunakan metode *XGBoost* dan *Improved Firefly Algorithm*. Hasil prediksi tersebut selanjutnya akan digunakan sebagai bahan pertimbangan membangun portofolio yang dioptimasi dan selanjutnya portofolio dibangun dengan menggunakan metode *Equal Weight (EW)*.

B. Topik dan Batasannya

Topik yang dibahas pada tugas akhir ini yaitu bagaimana implementasi metode *XGBoost* dan *Improved Firefly Algorithm* dalam memprediksi *return* saham dan bagaimana kerja portofolio yang diperoleh dari hasil pertimbangan prediksi *return* saham apabila dibandingkan dengan tanpa mempertimbangkan prediksi.

Adapun batasan masalah pada tugas akhir ini yaitu data yang digunakan adalah data saham harian dalam indeks LQ45 yang diperoleh dari Yahoo Finance dengan rentang waktu 12 tahun, dimulai dari 14 Februari 2010 hingga 14 Februari 2022.

C. Tujuan

Tujuan dari tugas akhir ini yaitu untuk mengimplementasi metode *XGBoost* dan *Improved Firefly Algorithm* dalam memprediksi *return* saham dan

menentukan apakah kinerja portofolio akan lebih baik jika diperoleh dari pertimbangan prediksi *return* saham apabila dibandingkan dengan tanpa mempertimbangkan prediksi.

D. Organisasi Tulisan

Organisasi tulisan dari tugas akhir ini yaitu setelah pendahuluan berupa studi terkait mengenai prediksi harga dan *return* saham, metode XGBoost, dan Improved Firefly Algorithm. Kemudian, pada bagian sistem yang dibangun berisikan bentuk sistem, alur kerja sistem dan bentuk pengujian sistem. Bagian selanjutnya yaitu evaluasi yang berisikan pemaparan hasil pengujian serta analisis seperti kinerja model machine learning, dan kinerja portofolio dengan dan tanpa mempertimbangkan prediksi. Bagian terakhir berupa kesimpulan mengenai hasil dan analisis pengujian.

B. KAJIAN TEORI

A. Prediksi Harga dan Return Saham

Prediksi harga dan *return* saham sudah banyak dilakukan sebelumnya, beberapa diantaranya menggunakan metode *Support Vector Regression* (SVR), *Artificial Neural Network* (ANN), *Recurrent Neural Network* (RNN), dan *Long Short-Term Memory Units*. Seperti pada tahun 2013, Messad dkk[1] melakukan penelitian untuk memprediksi harga saham dengan metode *Support Vector Regression* (SVR). Bentuk data yang digunakan merupakan time series yang diambil dari *Dhaka Stock Exchange* (DSE) selama kurang lebih empat tahun, mulai dari tahun 2009 sampai dengan tahun 2012. Dalam mengevaluasi performa, digunakan *Mean Absolute Percentage Error* (MAPE) untuk menghitung error rate antara harga saham aktual dengan harga saham prediksi yang berasal dari hasil percobaan. Dan didapatkan hasil perhitungan dengan error-rate sebesar 0.22.

Selanjutnya, Selvamuthu dkk [2] juga pernah melakukan penelitian untuk memprediksi harga saham pada tahun 2019. Metode yang digunakan adalah *Artificial Neural Network* (ANN) dengan data yang digunakan berupa tick data dari Reliance Private Limited selama satu tahun, terhitung mulai dari tahun 2017 hingga 2018. Harga saham yang diambil dari tick data merupakan harga saham di awal setiap 15 menit. Untuk algoritma pembelajaran yang digunakan pada penelitian ini, yaitu *Levenberg-Marquardt* (LM), *Scaled Conjugate Gradient* (SCG), dan *Bayesian Regularization* (BR). Pada persiapan data, dilakukan split data sebesar 75% untuk data training dan 25% untuk data testing. Kemudian, di akhir penelitian dilakukan evaluasi performa untuk ketiga algoritma pembelajaran dengan menggunakan *Mean Squared Error* (MSE) dan MAPE, dan didapatkan hasil akurasi sebesar 96.2%, 97.0%, dan 98.9% untuk LM, SCG, dan BR.

Masih di tahun 2019, Peng dkk [3] juga melakukan penelitian terkait memprediksi harga saham dengan menggunakan metode *Recurrent Neural Network* (RNN) dan *Long Short-Term Memory Units*. Data yang digunakan pada penelitian ini merupakan data historis dari American Airlines selama 20 tahun, terhitung mulai tahun 1970. Terdapat dua skema prediksi optimasi yang digunakan, yaitu metode hybrid dan recursive. Dalam evaluasi

performanya didapatkan hasil MSE sebesar $5.631e-4$ dan $1.042e-2$ untuk metode hybrid dan recursive.

B. eXtreme Gradient Boosting

XGBoost atau *eXtreme Gradient Boosting* diusulkan oleh Chen pada tahun 2016 [6]. Pada dasarnya XGBoost merupakan salah satu metode *machine learning* dengan model *ensemble learning* berbasis *decision tree*, yaitu suatu algoritma untuk mengoptimalkan klasifier (pohon) lemah secara efektif sehingga meningkatkan kinerja pembelajaran sistem. Dalam hal ini, klasifier – klasifier lemah yang terbentuk bertujuan agar setiap klasifier (pohon) baru berfokus untuk mempelajari kelemahan (data yang salah diklasifikasi) dari klasifier sebelumnya. Berdasarkan *ensemble learning*, maka klasifier atau pohon yang baru mengalami peningkatan gradien berdasarkan persamaan berikut:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F, \quad (1)$$

dimana $\hat{y}_i$, x_i , dan K adalah respon prediksi, input dan jumlah fungsi dalam ruang fungsi F. Untuk memastikan fungsi yang paling sesuai, f_k , pada fungsi tersebut diperkenalkan sebagai parameter yang akan menyesuaikan data selama pelatihan dan secara otomatis menemukan kelompok terkait. Dalam algoritma XGBoost, faktor regularisasi $\Omega(f_k)$ ditambahkan untuk mengekspresikan kompleksitas pohon yang berdasarkan GBDT. Untuk f_k dipelajari dengan meminimumkan fungsi objektif dari model pelatihan sebagaimana persamaan berikut:

$$L(j) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (2)$$

dimana $l(\hat{y}_i, y_i)$ adalah *loss function* dan Ω adalah *regularisation*. *Loss function* dapat berupa *logistic loss* atau *square loss*, yang mewakili tingkat kesamaan antara *training set* dan model. *Regularisation* merupakan istilah untuk mengontrol kompleksitas model dan biasanya mengontrol *overfitting*. Kompleksitas setiap pohon dapat didefinisikan sebagai rumus berikut:

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \quad (3)$$

di mana:

γ , λ = Koefisien *regularization* L1, dan Koefisien *regularization* L2

T , ω = Jumlah *node*, dan bobot *node*

C. Improved Firefly Algorithm

Pengaturan *hyperparameter* adalah sebuah masalah optimasi dimana pada prinsipnya bertujuan untuk mencari nilai yang meminimalkan *loss function*. Semua pendekatan

untuk otomatisasi pencarian *hyperparameter* terbagi menjadi dua model yaitu *model-free* dan *model-based*. Salah satu algoritma yang digunakan untuk pencarian *hyperparameter*, yaitu *Firefly Algorithm (FA)*. *Firefly Algorithm* merupakan algoritma optimasi bersifat *nature-inspired metaheuristic* yang terinspirasi oleh kunang – kunang dengan melihat dari segi perilaku sosial dan cara komunikasi sekelompok kunang – kunang dengan cahaya yang dipancarkan dari ekornya [7]. Karakteristik kunang – kunang dipandang lebih ideal sehingga dapat dijadikan sebuah algoritma.

Dr. Xin-She Yang merumuskan 3 persamaan yang dibuat dalam perilaku pergerakan kunang – kunang yaitu:

1. Intensitas Cahaya

Ada dua hal yang berkaitan dan sangat penting dalam *Firefly Algorithm* yaitu intensitas cahaya dan fungsi keatraktifan. Untuk kasus yang paling sederhana, contohnya optimasi maksimum, tingkat intensitas cahaya pada sebuah kunang – kunang x dapat dilihat sebagai,

$$I(x) = f(x) \quad (4)$$

dengan nilai I merupakan tingkat intensitas cahaya pada x kunang – kunang yang sebanding terhadap solusi fungsi tujuan permasalahan yang akan dicari $f(x)$. Keatraktifan β bernilai relatif karena intensitas cahaya yang harus dilihat dan dinilai oleh kunang – kunang akan berbeda. Dengan demikian, hasil penilaian juga akan berbeda tergantung dari jarak antara kunang – kunang satu dengan yang lainnya. Selain itu intensitas cahaya akan menurun dilihat dari sumbernya dikarenakan terserap oleh media lain seperti udara γ . Fungsi keatraktifan ialah sebagai berikut:

$$\beta = \beta_0 e^{-\gamma r_{ij}^2} \quad (5)$$

2. Jarak

Jarak antara dua kunang – kunang setiap i dan j , pada posisi x_i dan x_j , masing – masing, dapat didefinisikan sebagai jarak *Cartesian* atau *Euclidean* sebagai berikut:

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (6)$$

3. Pergerakan Kunang – Kunang

Pergerakan kunang – kunang i yang tertarik oleh lebih terangnya kunang – kunang j diberikan oleh persamaan berikut:

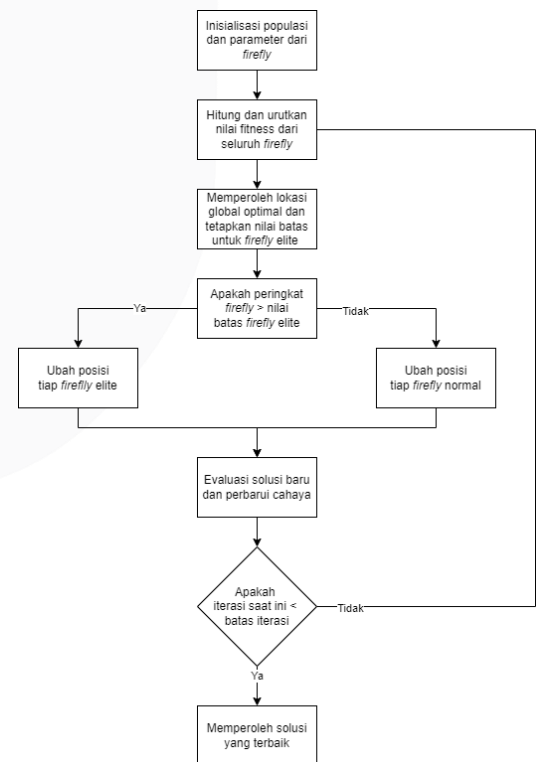
$$x_{id}(t+1) = x_{id}(t) + \beta_0 e^{-\gamma r_{ij}^2} (x_{jd}(t) - x_{id}(t)) + \alpha \left(\delta - \frac{1}{2} \right) \quad (7)$$

dimana x_{jd} dan x_{id} adalah nilai dari dimensi d dari kunang – kunang x_j dan x_i , masing – masing.

Prosedur standar untuk menerapkan *Firefly Algorithm* adalah sebagai berikut:

- inisialisasi populasi *firefly*, jumlah iterasi dan parameter *firefly algorithm*.
- Evaluasi fungsi *fitness* pada setiap *firefly*.
- inisialisasi fungsi *fitness* awal sebagai penentuan tingkat intensitas cahaya awal.
- Perbarui pergerakan setiap *firefly* menggunakan persamaan pergerakan kunang – kunang.
- Bandingkan tiap calon *firefly* terbaik dari nilai fungsi *fitness* agar mendapatkan nilai *firefly* terbaik.
- Lakukan sampai batas iterasi atau sampai mendapatkan *firefly* dengan fungsi *fitness* yang cukup baik.

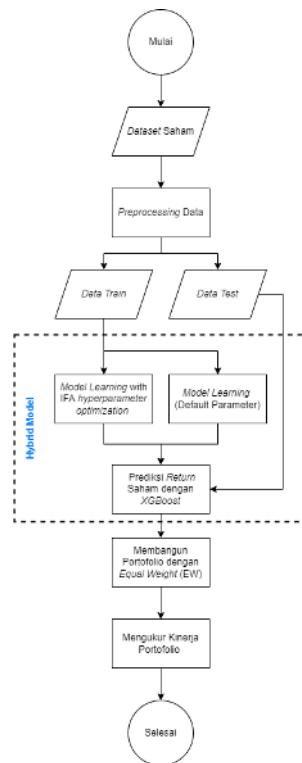
Namun, *Firefly Algorithm* biasa masih mempunyai masalah terkait optimasi dengan skala besar. Untuk mengatasi masalah ini, pertama – tama bagi seluruh kelompok kunang – kunang menjadi sub kelompok elite dan sub kelompok normal. Strategi *chaotic search* digunakan untuk sub kelompok elite untuk meningkatkan pencarian lokal. Pada saat yang sama, strategi *Particle Swarm Optimization (PSO)* digunakan untuk sub kelompok normal untuk meningkatkan pencarian global dan mempercepat konvergensi global. *Flowchart* dari *Improved Firefly Algorithm* dijelaskan pada Gambar 1.



GAMBAR 1
FLOWCHART IMPROVED FIREFLY ALGORITHM

C. METODE

Pada tugas akhir ini, pembentukan portofolio terdiri dari saham – saham yang sudah dipilih. Pemilihan tersebut didasarkan pada prediksi *return* saham menggunakan metode *XGBoost* dan *Improved Firefly Algorithm*. Jika hasil prediksi *return* saham melebihi *threshold*, saham akan dipilih untuk dimasukkan ke dalam portofolio. Dari portofolio yang sudah dioptimasi akan dibangun dengan menggunakan *Equal Weight (EW)*. Kinerja portofolio dengan dan tanpa optimasi akan dibandingkan untuk mengetahui *mean return*, standar deviasi, dan *Sharpe Ratio* tertinggi. Metodologi dari tugas akhir ini ditunjukkan pada Gambar 2.



GAMBAR 2
METODOLOGI PENELITIAN

A. Dataset

Dataset yang digunakan pada tugas akhir ini adalah data harga penutupan (*close price*) harian saham indeks LQ45 yang didapatkan dari Yahoo Finance dalam jangka waktu 12 tahun, mulai dari Februari 2010 sampai dengan Februari 2022. Dari harga harian saham tersebut kemudian digunakan untuk menghitung *Simple Moving Average (SMA)*, *Exponential Moving Average (EMA)*, *Moving Average Convergence Divergence (MACD)*, *Relative Strength Index (RSI)*, dan juga *Stochastic*. Hasil tersebut yang digunakan sebagai fitur – fitur dalam membangun model prediksi *return* saham.

B. Preprocessing Data

Data yang sudah dikumpulkan merupakan data mentah yang memungkinkan adanya noise, nilai kosong dan bahkan tidak konsisten, oleh karena itu perlu dilakukannya teknik preprocessing data. Teknik ini dilakukan guna menghasilkan data yang berkualitas.

Adapun tahapan yang akan dilakukan dalam *preprocessing* data sebagai berikut.

1. Filling Missing Value

Filling Missing Value merupakan suatu teknik yang digunakan untuk menangani data yang bernilai kosong atau *null*. Sebab *dataset* merupakan data *time series*, maka *imputation* dilakukan menggunakan metode *Last Observation Carried Forward (LOCF)* atau mengisi nilai kosong dengan nilai di hari sebelumnya.

2. Data Splitting

Data Splitting atau pembagian data dilakukan untuk membagi *dataset* menjadi tiga bagian, yaitu data *training*, data *validation*, dan data *testing*. Data *training* digunakan untuk pembentukan model dan melatih algoritma, data *validation* digunakan untuk proses validasi model dan mencegah *overfitting*, sedangkan data *testing* digunakan untuk menguji model yang sudah dilatih. Pembagian data dilakukan sebesar 70% untuk data *training*, 15% untuk data *validation* dan 15% untuk data *testing*.

C. Optimasi hyperparameter

Implementasi *Improved Firefly Algorithm (IFA)* bertujuan untuk mencari parameter yang optimal sehingga saat membangun model prediksi menggunakan *XGBoost* menghasilkan keluaran yang terbaik. Implementasi IFA ini terdiri dari dua proses utama, yaitu penentuan parameter yang digunakan dan proses iterasi berulang. Parameter yang digunakan pada IFA dijelaskan pada Tabel 1.

TABEL 1
PENGATURAN PARAMETER *IMPROVED FIREFLY ALGORITHM*

Symbol	Keterangan	Nilai
n	Jumlah populasi	100
t	Jumlah iterasi	100
α	Konstanta kecepatan perpindahan <i>firefly</i>	0.9
β_0	Konstanta ketertarikan	1
γ	Konstanta penyerapan cahaya	2

D. Pembentukan Portofolio

Dengan mempertimbangkan hasil prediksi *return* saham, saham yang melebihi *threshold* akan dimasukkan ke dalam portofolio. Dari saham – saham yang sudah terpilih, dibangun portofolio dengan menggunakan *Equal Weight (EW)*. *Equal Weight (EW)* merupakan metode paling sederhana untuk membangun sebuah portofolio dengan menentukan bobot yang sama untuk semua saham [8]. Bobot portofolio *EW* dihitung menggunakan persamaan (8).

$$w_i = \frac{1}{N}, i = 1, 2, \dots, N \quad (8)$$

di mana w_i adalah bobot saham ke- i , dan N adalah banyaknya saham dalam portofolio.

E. Evaluasi Kinerja

1. Evaluasi hyperparameter

Untuk mengukur kinerja *hyperparameter* yang sudah dioptimalkan maka akan dilakukan pembandingan dengan kinerja *hyperparameter* tanpa dioptimasi dengan Improved Firefly Algorithm Perbandingan tersebut dinilai dari perhitungan *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE) dengan persamaan (9) dan (10).

$$RMSE = \sqrt{\frac{\sum_{i=1}^T (y_i - y'_i)^2}{T}} \quad (9)$$

$$MSE = \frac{\sum_{i=1}^T (y_i - y'_i)^2}{T} \quad (10)$$

di mana T adalah jumlah data yang diprediksi, kemudian y adalah nilai aktual dan y' adalah nilai hasil prediksi.

2. Evaluasi Data Splitting

Untuk mengetahui apakah model benar – benar belajar menggunakan data yang ada, dilakukanlah evaluasi *data splitting* sehingga baik dari data *training*, data *validation*, dan data *testing* akan dilihat tingkat kesalahan dalam prediksi dan kinerja model dengan pertimbangan dari nilai perhitungan RMSE dan MAE.

3. Evaluasi Prediksi Return Saham

Untuk mengukur kinerja prediksi *return* saham dilakukan dengan mempertimbangkan tingkat kesalahan dalam prediksi dan kinerja model *XGBoost* dan *Improved Firefly Algorithm*, digunakan perhitungan *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE).

4. Evaluasi Pembentukan Portofolio

Untuk mengukur kinerja pembentukan portofolio dilakukan dengan perhitungan *mean return*, standar deviasi dan *Sharpe Ratio*. Fungsi standar deviasi *return* yaitu sebagai pertimbangan risiko karena dapat menunjukkan ketidakpastian dari nilai *return* yang ada. Perhitungan *mean return*, standar deviasi dan *Sharpe Ratio* masing – masing diberikan oleh persamaan (11), (12) dan (13).

$$R = \frac{1}{T} \sum_{t=1}^T R_t \quad (11)$$

$$\sigma = \sqrt{\frac{\sum_{t=1}^T (R_t - \mu)^2}{T}} \quad (12)$$

$$SR = \frac{R}{\sigma} \quad (13)$$

di mana T adalah jumlah data saham, dan R_t adalah *return* portofolio.

D. HASIL DAN PEMBAHASAN

A. Skenario Pengujian

1. Melakukan optimasi *hyperparameter* dengan menggunakan *Improved Firefly Algorithm*
2. Membangun model prediksi *return* saham dengan menggunakan *XGBoost*
3. Menentukan nilai konstanta *threshold* sebesar 0.005 untuk menyeleksi saham berdasarkan hasil prediksi *return*. Saham akan dimasukkan ke dalam portofolio apabila nilai prediksi melebihi *threshold*.
4. Membangun portofolio dengan menggunakan metode *Equal Weight* dimulai dari 3 saham, 5 saham, dan 7 saham (berdasarkan urutan *market cap* terbesar) berdasarkan hasil seleksi pada poin (3).
5. Bobot portofolio dihitung menggunakan persamaan (8).
6. Membandingkan hasil kinerja portofolio dengan dan tanpa seleksi untuk 3 saham, 5 saham, dan 7 saham berdasarkan *mean return*, standar deviasi, dan *Sharpe Ratio*.

B. Hasil Pengujian

Hasil yang didapatkan setelah pengujian pada *hyperparameter* dengan dan tanpa optimasi *Improved Firefly Algorithm* (IFA) ditunjukkan pada Tabel 2.

TABEL 2
EVALUASI HASIL PERBANDINGAN HYPERPARAMETER

Saha m	Tanpa optimasi <i>hyperparameter</i>			Dengan optimasi <i>hyperparameter</i>		
	Parameter	RM SE	M AE	Parameter	RM SE	M AE
BBC A.JK	'n_estimators': 100,	0.0 109	0.0 075	'n_estimators': 400,	0.0 106	0.0 072
BBRI .JK	'learning rate': 0.01,	0.0 164	0.0 115	'learning rate': 0.05,	0.0 153	0.0 110
TLK M.JK	'max_depth': :10,	0.0 140	0.0 094	'max_depth': : 17,	0.0 128	0.0 087
BMR L.JK	'gamma': 0.01,	0.0 143	0.0 104	'gamma': 0.01,	0.0 126	0.0 091
ASII .JK	'random_sta te': 42,	0.0 135	0.0 102	'random_sta te': 42,	0.0 132	0.0 096
UNV R.JK	'subsample': : 0.8	0.0 134	0.0 086	'subsample': : 0.8,	0.0 144	0.0 090
BBNI .JK		0.0 183	0.0 128	'colsample_ bytree': 1, 'colsample_ bylevel': 1	0.0 173	0.0 122

Setelah mendapatkan hasil pengujian pada *hyperparameter* dengan dan tanpa optimasi IFA, selanjutnya didapat hasil pengujian pada *splitting* data dengan mempertimbangkan nilai RMSE dan MAE pada data *training*, data *validation*, dan data *testing* yang ditunjukkan pada Tabel 3.

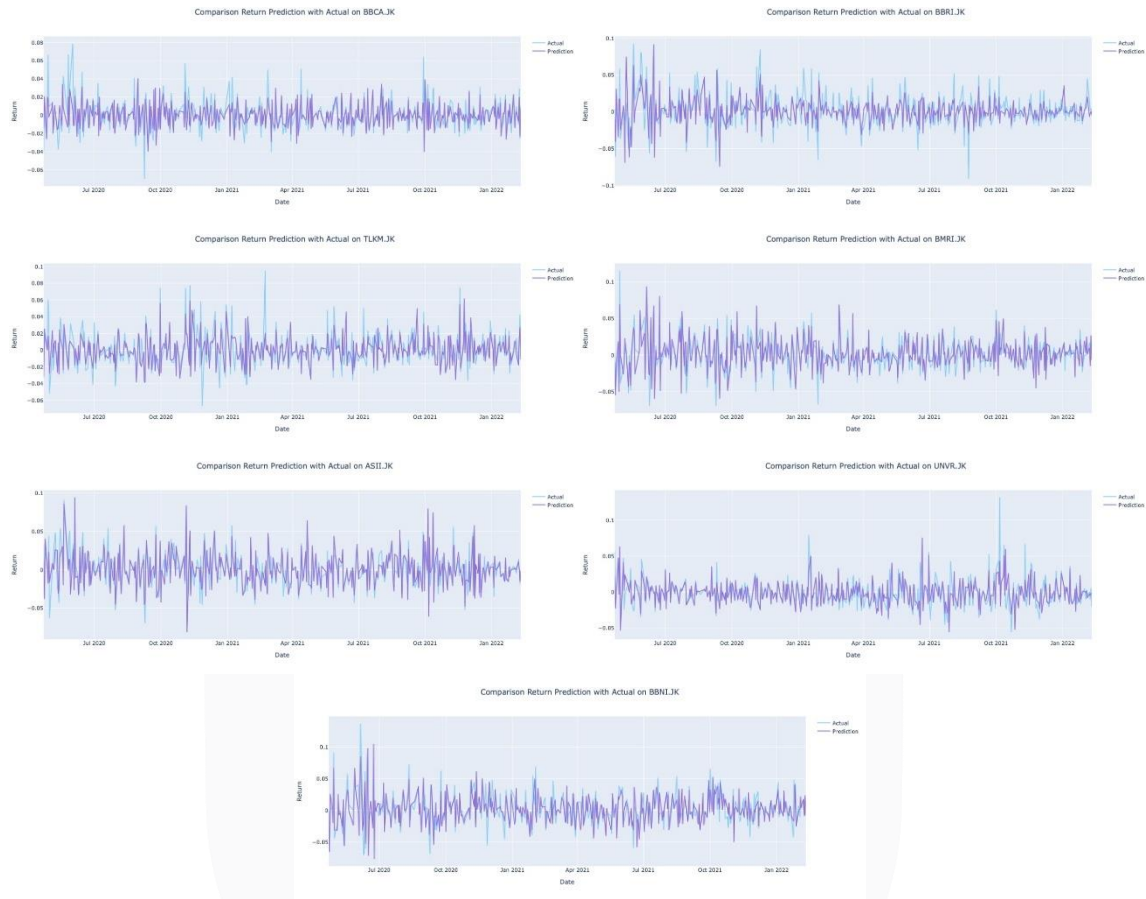
TABEL 3
EVALUASI HASIL PERBANDINGAN DATA SPLITTING

Saham	Data training	Data	Data testing
-------	---------------	------	--------------

			validation			
	RMS E	MAE	RMS E	MAE	RMS E	MAE
BBCA.JK	39.914	29.428	0.012	0.006	0.010	0.007
K	μ	μ	4	1	6	2
BBRI.JK	53.886	39.079	0.015	0.009	0.015	0.011
K	μ	μ	7	7	3	0
TLKM.JK	41.040	29.426	0.013	0.008	0.012	0.008
K	μ	μ	2	6	8	7
BMRI.JK	37.391	26.618	0.020	0.011	0.012	0.009
K	μ	μ	3	1	6	1
ASII.JK	20.804	14.564	0.012	0.008	0.013	0.009
K	μ	μ	8	7	2	6

UNVR.JK	26.105	17.943	0.015	0.009	0.014	0.009
K	μ	μ	3	1	4	0
BBNI.JK	36.015	16.978	0.016	0.011	0.017	0.012
K	μ	μ	9	0	3	2

Kemudian hasil prediksi *return* dan evaluasi model untuk 3 saham, 5 saham dan 7 saham, ditunjukkan pada Gambar 3 dan Tabel 4.



GAMBAR 3
HASIL PREDIKSI UNTUK 3 SAHAM, 5 SAHAM, DAN 7 SAHAM

TABEL 4
EVALUASI HASIL PREDIKSI

Saham	RMSE	MAE
BBKA.JK	0.0106	0.0072
BBRI.JK	0.0153	0.0110
TLKM.JK	0.0128	0.0087

BMRI.JK	0.0126	0.0091
ASII.JK	0.0132	0.0096
UNVR.JK	0.0144	0.0090
BBNI.JK	0.0173	0.0122

Dan hasil perbandingan nilai portofolio ditunjukkan pada Gambar 4, 5, 6.



GAMBAR 4
PERBANDINGAN NILAI PORTOFOLIO UNTUK 3 SAHAM



GAMBAR 5
PERBANDINGAN NILAI PORTOFOLIO UNTUK 5 SAHAM



GAMBAR 6
PERBANDINGAN NILAI PORTOFOLIO UNTUK 7 SAHAM

Evaluasi dari ketiga skenario perbandingan nilai portofolio dapat ditunjukkan pada Gambar 7 dn Tabel 5.

TABEL 5.
EVALUASI PERBANDINGAN NILAI PORTOFOLIO UNTUK
3, 5, 7 SAHAM

Evaluasi	3 Saham		5 Saham		7 Saham	
	Tanpa prediksi	Dengan prediksi	Tanpa prediksi	Dengan prediksi	Tanpa prediksi	Dengan prediksi
Mean Return	0.0011	0.0019	0.0012	0.0027	0.0009	0.0029
Standa	0.015	0.015	0.015	0.016	0.014	0.015

r Devias i	3	6	6	2	8	8
Sharpe Ratio	0.074 8	0.125 6	0.076 9	0.170 7	0.062 2	0.183 7



GAMBAR 7
PERBANDINGAN KINERJA PORTOFOLIO UNTUK 3, 5, 7 SAHAM

B. Analisis Hasil Pengujian

Berdasarkan Tabel 2, perbandingan nilai *error* antara *hyperparameter* dengan dan tanpa optimasi tidak terlalu signifikan, namun tetap nilai *error* untuk *hyperparameter* dengan optimasi lebih rendah dibandingkan tanpa optimasi. Sehingga optimasi *hyperparameter* dengan menggunakan *Improved Firefly Algorithm* berkontribusi untuk meningkatkan kinerja model prediksi walau tidak terlalu besar. Kemudian model dengan *hyperparameter* yang sudah dioptimasi tersebut dilatih dengan data *training*, data *validation*, dan juga data *testing*. Pada Tabel 3, nilai *error* data *training* sangatlah kecil, ini terjadi karena model dilatih menggunakan data yang sama sehingga model sudah mengingat data tersebut. Kemudian nilai *error* data *validation* sangat tinggi apabila dibandingkan dengan data *training*, ini disebabkan karena model melakukan validasi dengan data *validation* untuk mencegah overfitting. Untuk data *testing* terlihat nilai *error* yang didapat lebih kecil dari data *testing* karena model susah mulai memahami data – data yang diberikan. Jika diamati maka dapat diketahui bahwa model mengalami proses pembelajaran dengan data yang ada baik dari data *training*, data *validation*, dan juga data *testing*.

Untuk analisis hasil perbandingan nilai portofolio 3 saham dapat dilihat pada Gambar 4, bahwa pembentukan portofolio dengan menggunakan prediksi menghasilkan nilai portofolio 2.2 yang berbeda 0.6 dengan portofolio yang dibentuk tanpa menggunakan prediksi. Meskipun jika diperhatikan pada Tabel 5, *mean return* nilai portofolio 3 saham tidak terlalu jauh berbeda untuk portofolio dengan pertimbangan prediksi dengan yang tidak. Namun tetap pada pengujian portofolio 3 saham, portofolio dengan prediksi memberikan *mean return*, standar deviasi dan *Sharpe Ratio* yang lebih baik daripada portofolio tanpa prediksi.

Pada kasus pengujian portofolio 5 saham, berdasarkan Gambar 5, pembentukan portofolio dengan mempertimbangkan prediksi menghasilkan nilai portofolio 2 kali lebih besar dibandingkan tanpa pertimbangan prediksi. Dan apabila merujuk pada Tabel 5, baik dari *mean return* ataupun *Sharpe Ratio* memiliki nilai yang

cukup tinggi untuk portofolio dengan prediksi. Sehingga pada pengujian portofolio 5 saham, kinerja portofolio dengan prediksi memberikan *mean return*, standar deviasi dan *Sharpe Ratio* yang lebih baik.

Begitu pula untuk hasil perbandingan nilai portofolio 7 saham. Pada Gambar 6, terlihat nilai portofolio dengan pertimbangan prediksi *return* saham menghasilkan *gap* yang besar untuk portofolio tanpa pertimbangan prediksi. Apabila dilihat melalui Tabel 5, *mean return* dan *Sharpe Ratio* portofolio dengan prediksi pun memiliki nilai yang amat tinggi, perbandingan nilainya bisa sampai sepertiga. Ini sudah sangat membuktikan untuk pengujian portofolio 7 saham, portofolio dengan prediksi memberikan hasil kinerja yang lebih baik.

E. KESIMPULAN

Dari hasil pengujian yang dilakukan dapat dapat ditarik kesimpulan sebagai berikut :

1. Hasil pengujian untuk optimasi *hyperparameter* menunjukkan bahwa *hyperparameter* dengan optimasi menggunakan *Improved Firefly Algorithm* (IFA) menghasilkan perbedaan yang tidak terlalu signifikan apabila dibandingkan dengan *hyperparameter* tanpa optimasi dengan perbedaan nilai RMSE dan MAE sekitar 0.003 sampai 0.0020, hal ini bisa disebabkan karena model yang dibuat sebelumnya sudah cukup baik.
2. Hasil pengujian *data splitting* menunjukkan bahwa model mengalami perubahan dan pembelajaran berdasarkan data – data yang diberikan.
3. Hasil pengujian pembentukan portofolio 3 saham, 5 saham dan 7 saham menunjukkan bahwa portofolio dengan adanya pertimbangan prediksi *return* saham memiliki

kinerja yang lebih baik dari sisi *mean return*, standar deviasi maupun *Sharpe Ratio* dibandingkan portofolio tanpa pertimbangan prediksi dengan portofolio 3 saham memperoleh nilai *mean return* 0.0019, standar deviasi 0.0156, dan *Sharpe Ratio* 0.1256; portofolio 5 saham memperoleh nilai *mean return* 0.0027, standar deviasi 0.0162, dan *Sharpe Ratio* 0.1707; dan portofolio 7 saham memperoleh nilai *mean return* 0.0029, standar deviasi 0.0158, dan *Sharpe Ratio* 0.1837. Sehingga dapat disimpulkan juga banyaknya saham mempengaruhi nilai portofolio yang ada, semakin banyak saham yang ada semakin tinggi nilai *mean return*, standar deviasi dan juga *Sharpe Ratio*. Hal ini juga menandakan bahwa pertimbangan prediksi *return* saham sangat mempengaruhi portofolio.

and Applications SAGA 2009, pp. 169-178, 2009, doi: 10.1007/978-3-642-04944-6_14

- [8] S. Benartzi and R. H. Thaler, "Naive diversification strategies in defined contribution saving plans," *Am. Econ. Rev.*, vol. 91, no. 1, 2001, doi: 10.1257/aer.91.1.79.

A. Saran

Dari hasil pengamatan, beberapa saran untuk penelitian selanjutnya yaitu selain membangun portofolio dengan *Equal Weight* (EW) bisa digunakan metode lain seperti *Mean Variance* (MV) yang sudah banyak diimplementasi. Kemudian pembentukan portofolio bisa ditambahkan dengan seluruh saham – saham yang ada di indeks LQ45.

REFERENSI

- [1] P. Meesad and R. I. Rasel, "Predicting stock market price using support vector regression," 2013 *Int. Conf. Informatics, Electron. Vision, ICIEV 2013*, 2013, doi: 10.1109/ICIEV.2013.6572570.
- [2] D. Selvamuthu, V. Kumar, and A. Mishra, "Indian stock market prediction using artificial neural networks on tick data," *Financ. Innov.*, vol. 5, no. 1, 2019, doi: 10.1186/s40854-019-0131-7.
- [3] C. Peng, Z. Yin, X. Wei, and A. Zhu, "Stock Price Prediction based on Recurrent Neural Network with Long Short-Term Memory Units," 2019 *Int. Conf. Eng. Sci. Ind. Appl. ICESI 2019*, pp. 1–5, 2019, doi: 10.1109/ICESI.2019.8863005.
- [4] H. Markowitz, "Portfolio Selection," *J. Finance*, vol. 7, no. 1, pp. 77–91, 1952, doi: 10.1111/j.1540-6261.1952.tb01525.x.
- [5] Z. Dai and F. Wang, "Sparse and robust mean–variance portfolio 18 optimization problems," *Phys. A Stat. Mech. its Appl.*, vol. 523, pp. 1371–1378, 2019, doi: 10.1016/j.physa.2019.04.151.
- [6] T. Chen, C. Guestrin, "XGBoost: A scalable tree boosting system," 2016 *Int. Conf. Knowledge Discovery Data Mining. ACM SIGKDD*, pp. 785–794, 2016, doi: 10.1145/2939672.2939785.
- [7] X. -S. Yang, "Firefly Algorithm for Multimodal Optimization," *Stochastic Algorithms. Foundations*

