

## ABSTRACT

Gene expression programming (GEP) is capable of solving many prediction, classification, and optimization problem effectively. It uses a fixed-length chromosome representing a set of equations. A chromosome may contains one or more genes and the gene and chromosome length significantly affects the algorithm's performance. Different problems may require varied gene/chromosome lengths to achieve good results and user have to do trial and error to get best length. There is no study found to overcome gene length problem in GEP. Therefore, this study aimed to develop an adaptive GEP to find proper gene length during the evolutionary process, called adaptive length GEP (ALGEP).

In ALGEP, the gene length may be varied for each individual instead of using the same length. The evolutionary process would adjust the gene length and the chromosome with proper gene length will tend to survive. Furthermore, the study proposed a length adjustment operator that could delete or insert an allele in the chromosome to make it short or extended. This operator is expected to adjust the gene length to its optimal. A special slice crossover was also proposed to accommodate the crossover between parents with different gene lengths. The proposed algorithms' performance was investigated by solving three symbolic regression problems and the performance was compared to related previous gene expression programming algorithms which are original GEP and Robust GEP (RGEP). A constant creation method and a constant mutation operator were also proposed to deal with complex constants those may occur in the problems. In addition, the algorithm will be applied to solve a complex real problem that is stock market prediction problem. The performance of the algorithm in predicting Indonesian Stock Exchange Composite Index and an individual stock price will be compared with another well-known algorithm, deep learning artificial neural network.

The result show that ALGEP performs better than GEP and RGEP for solving the SRPs. Furthermore, the use of slice crossover in GEP make the algorithm perform significantly better for solving the SRP's; however, the use of length adjustment operator does not have significant impact. Finally the performance of ALGEP is significantly better for predicting a stock index and a stock price compared to artificial neural network.

**Keyword:** *gene expression programming, parameter setting, adaptive length, stock market prediction, artificial neural network.*



## ABSTRAK

*Gene expression programming* (GEP) banyak digunakan untuk melakukan prediksi, klasifikasi, dan optimasi secara efektif. GEP biasanya menggunakan kromosom dengan panjang tertentu. Dimana kromosom tersebut mewakili suatu program atau persamaan tertentu. Satu kromosom bisa terdiri atas satu atau lebih gen, Dimana panjang gen secara signifikan mempengaruhi kinerja algoritma. Masalah yang berbeda mungkin memerlukan panjang kromosom yang berbeda untuk mencapai hasil yang baik dan pengguna harus melakukan trial and error untuk mendapatkan panjang terbaik. Belum ditemukan penelitian yang dilakukan untuk menangani masalah panjang gen pada GEP tersebut. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan GEP dengan panjang gen yang adaptif, dimana panjang gen yang tepat akan secara otomatis didapatkan selama proses evolusi.

Untuk mengatasi masalah panjang gen/ kromosom, penelitian ini mengusulkan panjang gen yang bervariasi untuk setiap individu dalam satu populasi. Proses evolusi akan menyesuaikan panjang gen dan kromosom dengan panjang gen yang sesuai akan cenderung bertahan. Lebih lanjut, penelitian ini mengusulkan *length adjustment operator* yang dapat menghapus atau menambah satu alel (bit) ke dalam gen untuk menjadikan lebih pendek atau lebih panjang. Operator ini diharapkan dapat mengatur panjang gen hingga tercapai panjang optimal. Sebuah operator *crossover* yaitu *slice crossover* juga diusulkan untuk mengakomodasi persilangan antara dua individu dengan panjang gen/ kromosom yang berbeda. Kinerja algoritma yang diusulkan akan diuji untuk menyelesaikan tiga *symbolic regression problem* lalu kinerjanya akan dibandingkan dengan algoritma *gene expression programming* yang terkait yaitu GEP dan Robust GEP (RGEP). Sebuah metode pembangkitan konstanta dan operator mutasi konstanta juga diusulkan untuk menangani konstanta yang lebih kompleks yang mungkin diperlukan dalam permasalahan. Selain itu, algoritma ini juga akan diterapkan untuk menyelesaikan permasalahan nyata yang kompleks yaitu permasalahan prediksi pasar saham. Kinerja algoritma dalam memprediksi Indeks Harga Saham Gabungan (IHSG) dan harga saham individu akan dibandingkan dengan algoritma terkenal lainnya, yaitu jaringan syaraf tiruan.

Hasil penelitian menunjukkan bahwa ALGEP berkinerja lebih baik dibandingkan GEP dan RGEP dalam menyelesaikan SRP. Selain itu, penggunaan *slice crossover* di GEP membuat algoritme bekerja lebih baik secara signifikan dalam menyelesaikan SRP; Namun penggunaan *length adjustment operator* tidak memberikan dampak yang signifikan. Terakhir, kinerja ALGEP secara signifikan lebih baik dalam memprediksi indeks saham dan harga saham dibandingkan dengan jaringan syaraf tiruan.

**Kata kunci:** *gene expression programming, setting parameter, panjang gene adaptif, prediksi pasar saham, jaringan syaraf tiruan.*