

## ABSTRACT

In this research, Non Dominated Sorting Particle Swarm Optimization – Sequential Clustering Classification (NSPSO-SCC) was proposed for revealing the hidden patterns on one set of data, selecting data features and training a model which can predict the pattern of the data. In order to perform these tasks, the proposed approach combines clustering and classification method with multi-objective PSO method. The Non Dominated Sorting PSO was utilized to perform feature selection on both clustering and classification task, and at the same time it also maintained the quality of clustering and classification result in terms of compactness of clusters and classification accuracy. The clustering and classification procedures are conducted using two types sub-dataset i.e. **Q** and **X** dataset. For conducting clustering and classification task, hierarchical clustering and decision tree method are used. The performance of NSPSO-SCC was compared with PSO-SCC, traditional-SCC and Non Dominated Sorting Genetic Algorithm – Sequential Clustering Classification (NSGAI-SCC). The experimental result shows that NSPSO-SCC achieves better performance than PSO-SCC and traditional-SCC methods. Although the statistical test shows there is significant difference between NSPSO-SCC and NSGAI-SCC in terms of MSE and  $1/acc$ , there is no obvious evidence showing NSPSO-SCC is better than NSGAI-SCC because their Pareto-front lines are overlapped. However, the diversity metric shows that NSPSO-SCC performance is better than NSGAI-SCC since the diversity metric of NSPSO-SCC is lower than diversity metric of NSGAI-SCC. This result suggests that the proposed method can achieve promising clustering and classification results at one time.

**Keywords:** Clustering, Classification, Multi-Objective Optimization, Non-Dominated Sorting Algorithm, Particle Swarm Optimizer.

## ABSTRAK

Dalam penelitian ini, Non Dominated Sorting PSO – Sequential *Clustering Classification* (NSPSO-SCC) diusulkan sebagai suatu metode yang mengidentifikasi pola tersembunyi dalam data, menyeleksi fitur data yang penting dan melatih model yang dapat memprediksi pola dari data. Agar dapat melakukan keseluruhan proses tersebut, pendekatan yang diusulkan menggabungkan metode *clustering* dan *classification* dengan metode *multi-objective PSO*. *Non Dominated Sorting PSO* diaplikasikan untuk melakukan fitur seleksi pada kedua proses, yaitu *clustering* dan *classification*, dan pada saat yang bersamaan juga mempertahankan kualitas hasil dari *clustering* dan *classification*, yang mana dalam hal ini adalah kerapatan dari cluster dan keakuratan prediksi klasifikasi. Prosedur *clustering* dan *classification* dilakukan dengan menggunakan 2 jenis *sub-dataset*, yaitu dataset Q dan X. Metode *hierarchical clustering* dan *decision tree* digunakan untuk metode/teknik *clustering* dan klasifikasi. Performansi NSPSO-SCC dibandingkan dengan PSO-SCC, *traditional-SCC* dan *Non Dominated Sorting Genetic Algorithm – Sequential Clustering Classification* (NSGAI-SCC). Hasil percobaan menunjukkan bahwa NSPSO-SCC memiliki performansi yang lebih baik dibandingkan PSO-SCC dan *traditional-SCC*. Meskipun hasil uji statistik menunjukkan adanya perbedaan yang signifikan antara NSPSO-SCC dan NSGAI-SCC dalam hal MSE dan  $1/acc$ , tetapi tidak ada bukti lebih lanjut yang menunjukkan bahwa NSPSO-SCC lebih baik daripada NSGAI-SCC karena garis *pareto front* keduanya saling overlapped. Akan tetapi, *diversity metric* menunjukkan bahwa performansi NSPSO-SCC lebih baik daripada NSGAI-SCC. Hasil ini menunjukkan bahwa metode yang diusulkan dapat memberikan hasil *clustering* dan klasifikasi yang menjanjikan.

**Kata Kunci** : *Clustering, Classification, Multi-Objective Optimization, Non-Dominated Sorting Algorithm, Particle Swarm Optimizer.*