Peerj

FEATURE SELECTION USING A MULTI-STRATEGY IMPROVED PARROT OPTIMIZATION ALGORITHM IN SOFTWARE DEFECT PREDICTION

Reviewers Remarks

The paper presents a novel, well-structured, and practically significant methodology with clear experimental results. It is suitable for publication with minor changes to enhance clarity and comprehensiveness.

Reviewers Questions

1) While the parameters of the Parrot Optimization Algorithm are outlined, providing a sensitivity analysis or rationale for their selection would enhance the clarity of the methodology.

2) A brief discussion of why the PROMISE datasets were selected and their relevance to real-world software projects would improve context for the reader.

3) The paper introduces a multi-strategy improved Parrot Optimization Algorithm (POA) with enhancements like Tent chaotic mapping and t-distribution mutation. Could the authors mathematically validate why these enhancements specifically improve the POA's ability to avoid local optima and premature convergence compared to other swarm intelligence methods?
   (e.g., "Benchmark tests confirm the efficacy of the enhanced parrot optimizer in function optimization.")

4) Although AUC is an excellent metric, the inclusion of additional metrics such as precision, recall, or F1-score could strengthen the evaluation, especially for imbalanced datasets.

5) The paper could benefit from a more detailed comparison with state-of-the-art feature selection and defect prediction methods, particularly in terms of computational efficiency.

6) The suggestion of applying the algorithm to other domains (e.g., network intrusion detection) is promising. Briefly elaborating on how the algorithm could handle domain-specific challenges would further emphasize its versatility.

7) The paper suggests applying the BMEPO algorithm to network intrusion detection and disease factor analysis. Can the authors elaborate on how the optimization algorithm would handle the different feature spaces and data dynamics in these domains? *(e.g., "In the future, the BMEPO algorithm can be applied to other fields such as network intrusion system category feature selection.")*

8) Can the authors expand the discussion on the limitations of the proposed algorithm and potential scenarios where it might underperform?

9) The heterogeneous data stacking ensemble framework shows superior performance. Can the authors discuss the potential impact of using other base learners or meta-models in this framework?