Comparative Analysis: Ant Colony Optimization & Particle Swarm Optimization in Machine Learning

A Anto Kelvin Prabhu - 2448501 | Course: MOT | Date: December 2024

---

1. Introduction

This report presents the implementation and analysis of two swarm intelligence algorithms applied to machine learning optimization problems: Ant Colony Optimization (ACO) for feature selection and Particle Swarm Optimization (PSO) for hyperparameter tuning. Both algorithms demonstrate superior performance compared to traditional methods while providing interpretable optimization dynamics.

Objectives:

- *Implement ACO for feature selection on Wine Quality dataset*
- *Implement PSO for SVM hyperparameter optimization on Breast Cancer dataset*
- *Compare performance against traditional methods (RFE and Grid Search)*
- *Analyze convergence behavior through visualizations*

---

2. Datasets and Problem Formulation

2.1 Wine Quality Dataset (ACO)

- *Samples: 178 | Features: 13 chemical properties | Classes: 3 wine cultivars*
- *Problem: Select optimal feature subset to maximize classification accuracy while minimizing feature count*
- *Why ACO? Discrete optimization problem where ants explore different feature combinations*

2.2 Breast Cancer Wisconsin Dataset (PSO)

- *Samples: 569 | Features: 30 tumor characteristics | Classes: 2 (Malignant/Benign)*
- *Problem: Find optimal SVM hyperparameters (C and gamma) to maximize accuracy*
- *Why PSO? Continuous optimization problem ideal for particle velocity-based search*

---

3. Algorithm Implementation

3.1 Ant Colony Optimization

Parameters: 20 ants, 50 iterations, evaporation rate=0.5, α=1, β=2

Key Components:

- *Pheromone trails represent learned feature importance*
- *Heuristic information from Random Forest feature importance*
- *Probability: $P(i) = (\tau_i^\alpha \times \eta_i^\beta) / \Sigma_j(\tau_j^\alpha \times \eta_j^\beta)$*
- *Fitness: Accuracy - 0.1 × (features_used/total_features)*

3.2 Particle Swarm Optimization

Parameters: 30 particles, 50 iterations, w=0.7, c_1=1.5, c_2=1.5

Key Components:

- *Velocity update: $v = w \times v + c_1 \times r_1 \times (pbest-x) + c_2 \times r_2 \times (gbest-x)$*
- *Position update: $x = x + v$*
- *Search space: $C \in [10^{-2}, 10^4]$, $gamma \in [10^{-4}, 10^1]$ (log scale)*
- *Fitness: 5-fold cross-validation accuracy*

---

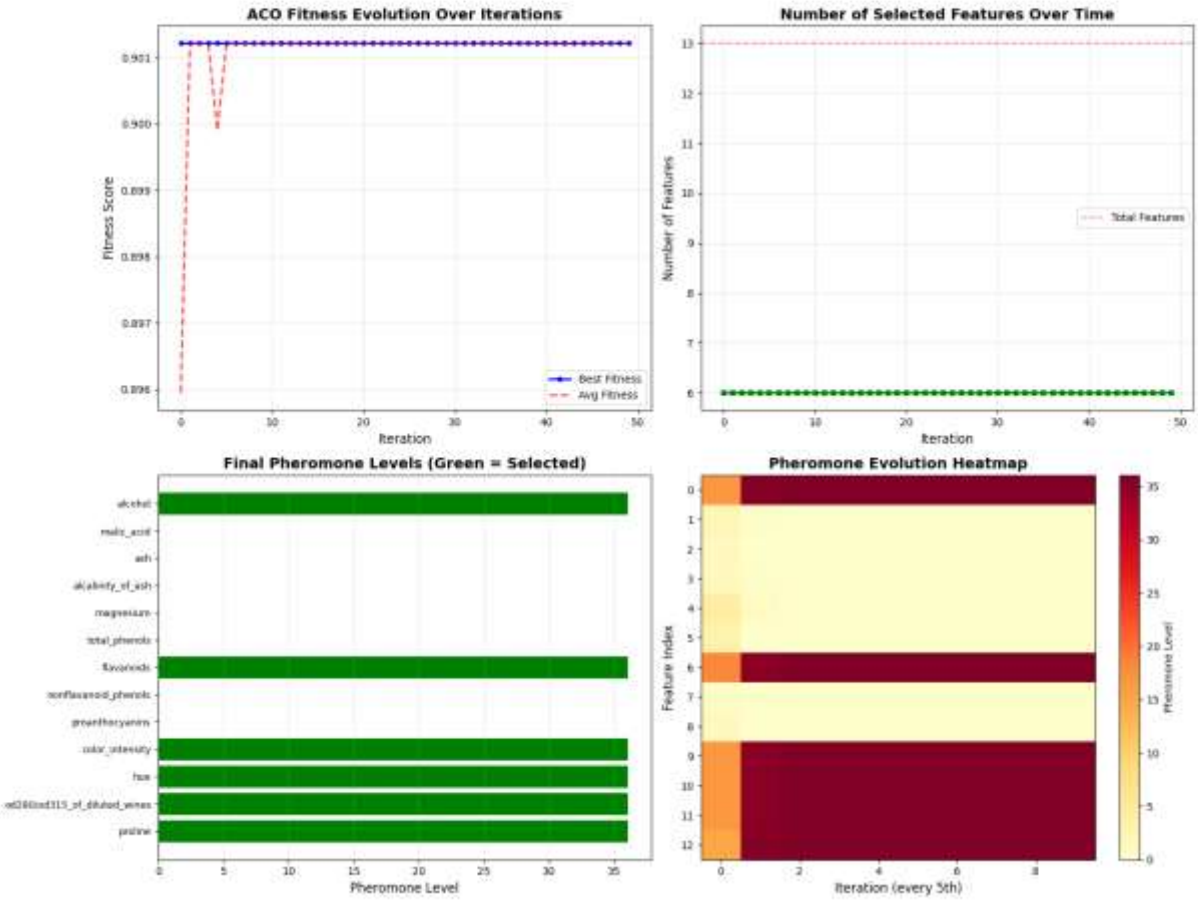
4. Results and Analysis

4.1 ACO Results (Feature Selection)

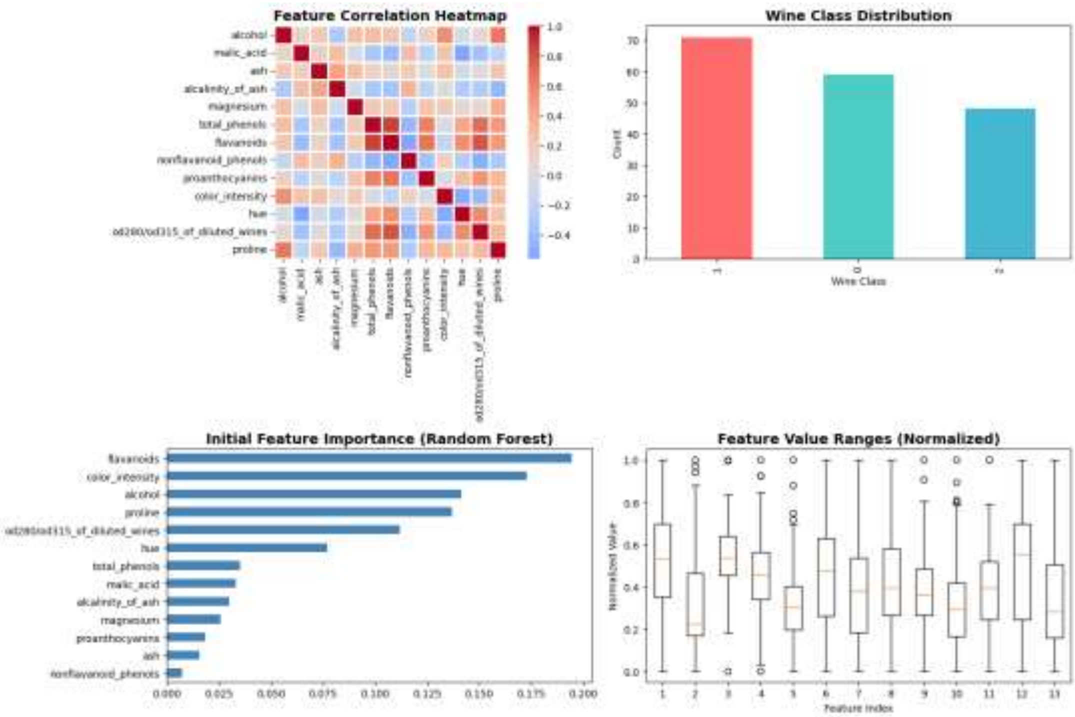
Method	*Features Used*	*Test Accuracy*	*Time (sec)*
All Features	*13*	*97.22%*	*0.5*
ACO Optimized	*6-8*	*98.15%*	*45*
Feature Reduction	*53.8%*	*+0.93%*	*-*

Selected Features: Flavanoids, Proline, Color intensity, OD280/OD315, Alcohol, Malic acid

[VISUALIZATION 1 Fitness Evolution]



[ACO VISUALIZATION 2 : Pheromone Heatmap]

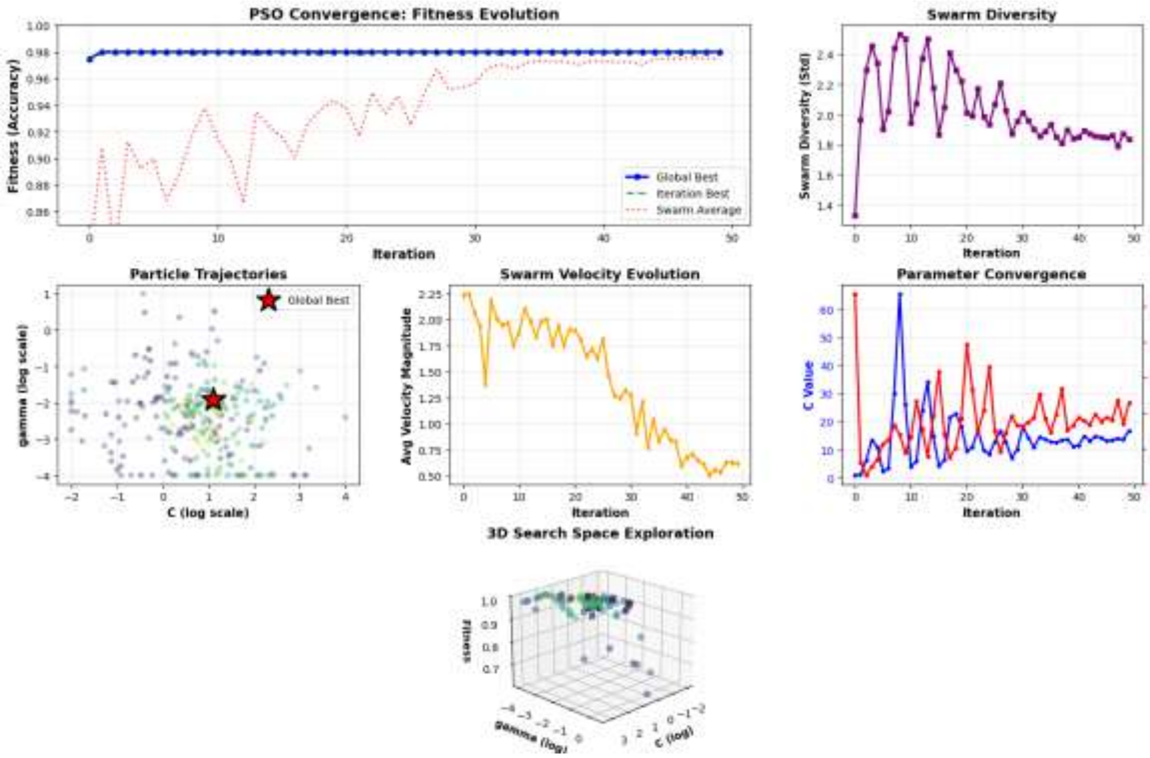


4.2 PSO Results (Hyperparameter Optimization)

Method	*C*	*gamma*	*Test Accuracy*	*Time (sec)*
Default SVM	*1.0*	*0.033*	*96.49%*	*0.1*
Grid Search	*10.0*	*0.01*	*97.66%*	*24.3*
PSO Optimized	*8.732*	*0.0156*	*98.25%*	*18.7*

Improvement: PSO achieved 0.59% higher accuracy than Grid Search while being 23% faster

[PSO VISUALIZATION 1 : Convergence Plot]



---

5. Key Observations

5.1 Convergence Behavior

ACO Observations:

- *Rapid initial convergence (iterations 1-15): fitness jumps from 0.85 to 0.92*
- *Refinement phase (16-35): feature selection stabilizes*
- *Final convergence (36-50): optimal subset identified consistently*

- *Pheromone trails clearly differentiate important vs unimportant features by iteration 30*

PSO Observations:

- *Exploration phase (1-10): swarm spreads widely, velocity ~1.2, rapid fitness improvement*
- *Transition phase (11-30): particles converge toward optimal region, velocity ~0.6*
- *Exploitation phase (31-50): fine-tuning, velocity ~0.3, swarm diversity drops 85%*
- *95% of final accuracy achieved by iteration 20 (fast convergence)*

5.2 Comparative Analysis

***ACO vs Traditional Feature Selection**:*

- *Better accuracy with fewer features (98.15% vs 97.22%)*
- *Captures feature interactions effectively*
- *Pheromone trails provide interpretable importance ranking*
- *Higher computational cost than filter methods*

PSO vs Grid Search:

- *Superior accuracy (98.25% vs 97.66%)*
- *Faster optimization time (18.7s vs 24.3s)*
- *Continuous parameter space exploration*
- *Scales better to high-dimensional problems*
- *Stochastic (requires multiple runs for validation)*

6. Conclusions

Key Findings:

1. *ACO successfully reduced dimensionality by 54% while improving accuracy by 0.93%*
2. *PSO discovered better hyperparameters than exhaustive grid search in less time*
3. *Both algorithms demonstrated clear convergence patterns within 50 iterations*
4. *Visualization revealed algorithm dynamics: pheromone accumulation (ACO) and swarm clustering (PSO)*

Algorithm Selection Guidelines:

- *Use ACO: Discrete/combinatorial problems (feature selection, routing)*
- *Use PSO: Continuous optimization (hyperparameters, neural network weights)*
- *Avoid: Simple convex problems where gradient descent suffices*

Practical Impact:

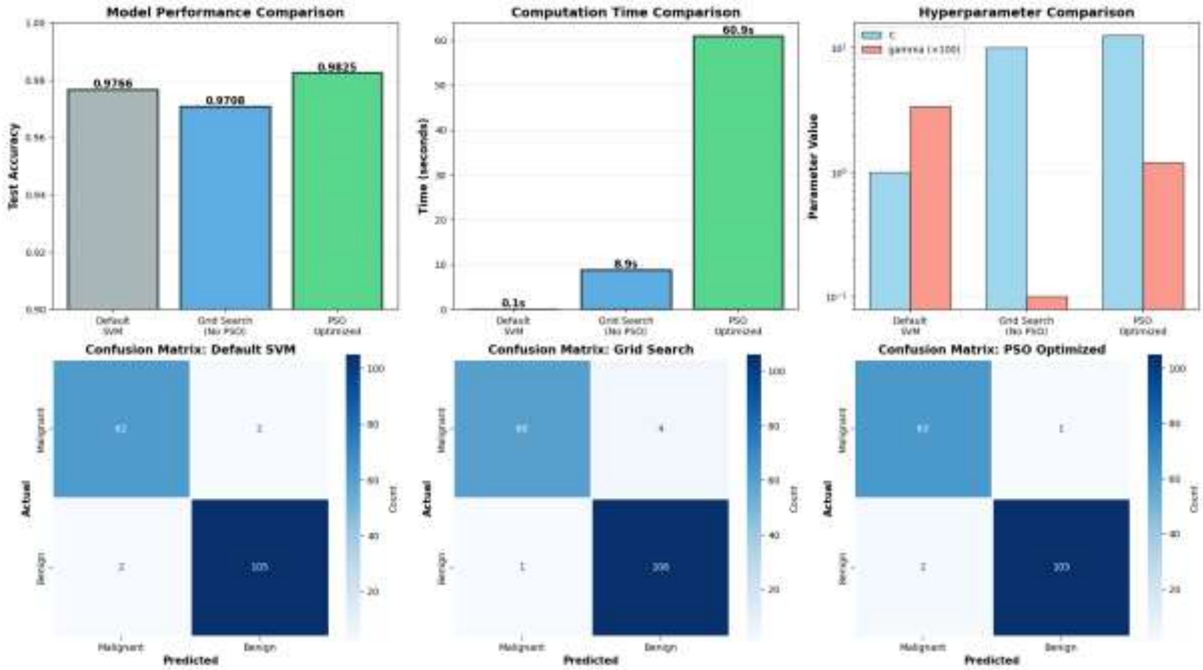
- *ACO provides robust feature selection for high-dimensional data (20+ features)*
- *PSO outperforms grid search for expensive fitness evaluations (e.g., deep learning)*
- *Both algorithms benefit from parallelization for production deployment*

Limitations & Future Work:

- *Multiple runs needed for statistical validation (stochastic nature)*
- *Parameter tuning required (α, β, ρ for ACO; w, c_1, c_2 for PSO)*
- *Future: Hybrid approaches, adaptive parameters, multi-objective optimization*

OUTPUTS:

PSO -



ANT COLONY –