Explanation of the Diagram

Phase 1: Preprocessing Phase

The first step is the Preprocessing Phase. This phase involves preparing the data and performing feature engineering before we start training. It consists of the steps:

1. **Removing Unwanted Features**: We eliminate unnecessary features from our dataset to reduce complexity and improve training efficiency.
2. **Encoding Categorical Features**: We convert features into representations so that they can be easily used with machine learning models.
3. **Utilizing the Criteria File**: We employ a criteria file to determine the "drinkability" of each sample X_i This file defines the target variable based on factors such as properties, geographic coordinates and sensory characteristics like smell, color and taste.

Phase 2: Initialization of PSO Parameters

After completing the Preprocessing Phase we move on to initializing PSO parameters that drive the particle swarm optimization algorithm;

- **Setting up Model Parameters**; We initialize network models (particles) with weight matrices and bias vectors. These initial configurations serve as starting points, for optimization.
- **Initializing Particle Parameters**: Each particle, which represents a network model possesses attributes such, as position, velocity, personal best (pbest) and global best (gbest).
- **Initializing Global Parameters**: Global parameters like max_v; Wmax; Wmin; C_1 ;C_2 and others have an impact on the movement of particles and the optimization process.

Phase 3: Training Loop

This phase involves updates to particle positions and velocities in order to optimize the network models according to the "drinkability" criteria outlined in the criteria file along with its associated features (X_i).

- **Updating Particle Positions**:The new position (X_i) is calculated by adding the position (X_i) to the velocity (V_i).
- **Updating Particle Velocities:** The new velocity (V_i) is determined based on factors such, as inertia weight (W) cognitive component (C_1) social component (C_2) as random values (rand_1 and \(rand_2).

Fitness Function (Accuracy)

The fitness function used to assess how well each particle (network model) performs is accuracy. It measures the models ability to correctly classify the data;

- **Calculating Accuracy**: Accuracy is determined by calculating the ratio of Correct Predictions, to Total Predictions expressed as a percentage.
- Evaluating Performance: During the evaluation phase we select the performing model based on its best position. This position represents the configuration that aligns closely with our criteria, for "drinkability."
- **Tracking Training Progress**: Throughout the training process we record both training and testing accuracy histories to monitor how well our model performs in terms of "drinkability" and other relevant attributes.