EE 608 - Applied Modeling and Optimization

Stellar Classification: A Particle Swarm Optimization Approach

Team members:
Akhil Prabhu
Shashankk Shekar Chaturvedi

December 19th, 2023

Problem Description

Understanding the Cosmos: Enhancing Stellar Classification through Innovative Techniques

The primary challenge addressed in this research is the spectral classification of celestial bodies, a task that involves the analysis of light spectra emitted by stars, galaxies, and quasars. The spectral characteristics of these entities are unique and provide a wealth of information about their composition, age, and evolutionary stages. However, the process of extracting and interpreting this information is complex and computationally intensive. Traditional methods of spectral classification often involve manual intervention and are not scalable for large astronomical datasets. This study proposes an innovative approach using Particle Swarm Optimization (PSO), a bio-inspired computational method, to automate and optimize the process of stellar classification.

Source:https://science.nasa.gov/missions/chandra/nasa-to-announce-new-discoveries-at-annual-astronomy-meeting/

Problem Description

Why does it matter?

The importance of this research is multifold. Firstly, it contributes to the field of astronomy by providing a more efficient and accurate method for stellar classification, thereby facilitating the study of celestial bodies on a larger scale. Secondly, by unlocking the secrets hidden in the spectral signatures of stars, we can gain a deeper understanding of the universe and its evolution.

This knowledge is fundamental to various areas of scientific inquiry, from the formation of galaxies to the possibility of extraterrestrial life. Lastly, the proposed method, with its roots in artificial intelligence, represents a significant advancement in the application of computational techniques to astronomical research. It underscores the potential of interdisciplinary approaches in pushing the boundaries of our understanding of the cosmos.

Optimization Model

The goal of Particle Swarm Optimization (PSO) is to optimize the parameters (weights and biases) of the neural network to order to minimize the objective function mentioned in the previous slide

The objective function:

$$\underset{min}{J(\theta)} = -\frac{1}{N}\sum_{i=1}^{N}\log\left(h_\theta(z)_i\right)$$

$$h_\theta(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \longrightarrow \text{Softmax}$$

Where z is a vector of logits for a sample

K is total number of classes (K=3)

N is total number of training samples (N = 152,082)

Θ represents the set of parameters (weights and biases) of the neural net

Optimization Model: Particle Swarm Optimization

We'll be using a standard feed forward algorithm for our neural network. However instead of backpropagation, we shall be using PSO. Here are the steps to execute this:

1. Initialize

$$\text{Position}: \theta_i = [\theta_{i1}, \theta_{i2}, \ldots, \theta_{iD}]$$
$$\text{Velocity}: v_i = [v_{i1}, v_{i2}, \ldots, v_{iD}]$$

where D is total number of neural network model parameters

2. Update Personal Best

$$\text{if } J_i < J_{best_i}, \text{ then } J_{best_i} = J_i, \ \& \ \theta_{best} = \theta_i$$

3. Update Global Best

$$\text{if } J_i < J_{global\,best}, \text{ then } J_{global\,best} = J_i, \ \& \ \theta_{global\,best} = \theta_i$$

4. Update Velocity and Position

$$V_{ij} = \omega V_{ij} + c_1 r_1 (\theta_{best_{ij}} - \theta_{ij}) + c_2 r_2 (\theta_{global\,best_j} - \theta_{ij})$$

where ω is the inertia weight
c_1 is the individual acceleration coefficient (Cognitive)
c_2 is the global acceleration coefficient (Social)
r_1 and r_2 are random values

5. Repeat Step 2 to 4 until you reach convergence or maximum iterations

Optimization Model: Constraints

The following are design constraints that will help us converge to a solution faster

Constraint	Reasoning
$0 < \omega < 1$	Ensure the model converges
$c_1 + c_2 = 4$	To strike a balance between personal best and global best components. For stability and convergence.
$0 < r_1 < 1, 0 < r_2 < 1$	To maintain scale between each iteration
Length of $\Theta = 588$	Number of nodes for input, hidden and output layers are 9, 45 and 3, respectively. Total number of parameters = $(9x45) + 45 + (45x3) + 3$
Number of particles = 500	Close to length of Θ. Can be reduced on increased based on computational resources available.

Proposed Solution

Convexity: Our objective function is convex (negative log-likelihood) and the constraints define a complex, potentially multimodal landscape with multiple local minima. However, PSO optimization part of this solution introduces non-convexity to this problem

Feasible Set: The feasible set is the set of all possible parameter values (Θ) that satisfy the constraints mentioned in the previous slide.

A solution may exist in the feasible set since there is a combination of parameters (Θ, ω, c_1, c_2) that satisfies all the constraints.

Closed-Form Solution: Finding a closed-form solution for the PSO optimization problem is challenging due to its non-convex nature.

Validating Solution

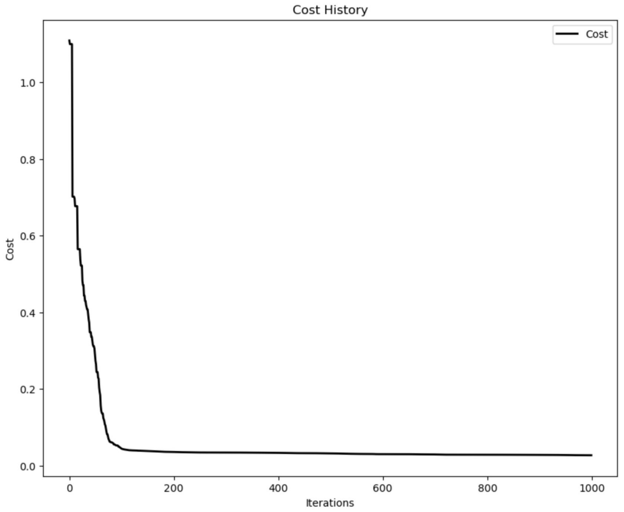
In our study, we utilized the "Stellar Classification Dataset SDSS17" from Kaggle. This dataset, derived from the Sloan Digital Sky Survey (SDSS), is a rich collection of 100,000 celestial observations. Each observation is described by a set of features capturing various spectral and observational characteristics. The dataset also includes a class label identifying the celestial body as a star, galaxy, or quasar. Our analysis leverages these features to classify these celestial bodies, thereby contributing to the broader understanding of our universe.

Image Source - [CMRLS News: Star Tours-Astronomy at It's Finest (cmrlslibrarynews.blogspot.com)]

Validating Solution: Feature Engineering

In our study, we undertook several feature engineering steps to refine the dataset and enhance the performance of our model: outlier removal, feature selection, handling data imbalance (SMOTE). These steps played a pivotal role in enhancing the robustness and accuracy of our stellar classification model. These steps are not merely procedural but strategic, each contributing to the model's ability to learn from the data and generalize to unseen instances. They helped transform raw data into a form that the model could learn from more effectively, and fine-tuned the model's parameters to better capture the underlying patterns in the data.

We iteratively adjusted the parameters to minimize the loss function. As seen in the graph below, after 1000 iterations, we observed a significant reduction in the loss, indicating that our model was learning effectively from the data.



Key Insights

Our solution for stellar classification using PSO demonstrated promising results. With optimized weights and biases (best cost function was 0.027), our model achieved an accuracy of 84.3% on a validation dataset, indicating its effectiveness in classifying celestial bodies (star, galaxy or quasar). This suggests that PSO is a viable approach for classification problems in machine learning and warrants further exploration in other domains. However, we observed a significant increase in computation time with only a slight increase in the number of particles, highlighting the need for careful parameter tuning.

Major Conclusions

Our approach, employing Particle Swarm Optimization for stellar classification, proved to be viable, as evidenced by the high accuracy achieved. However, we encountered challenges in the process, particularly in tuning the cognitive and social scaling factors (c_1 and c_2). Striking the right balance was critical as a minor alteration could lead to non-convergence or getting stuck in the local minima of the objective function. It required several trials to identify the optimal combination of these values.

Interestingly, the inertia term (ω) exhibited less influence on the results than anticipated. This could potentially be attributed to the substantial number of particles used and the overall quantity of parameters in our model. These observations underscore the intricate dynamics of parameter tuning in swarm intelligence algorithms and highlight areas for further investigation in future work.

Major Conclusions

In our study, we employed a neural network architecture comprising a single input layer, a single hidden layer, and a single output layer. However, the potential for enhancing the performance of our model exists. This could be achieved by experimenting with the architecture of the neural network, specifically by varying the number of hidden layers and the number of nodes within these layers.

It is crucial to note that an increase in the complexity of the model, through the addition of more layers or nodes, will inevitably lead to an increase in the computational expense of the PSO algorithm. This is a factor that must be carefully considered during the model design process.

However, a more complex model could potentially harness the full potential of larger datasets, thereby improving the overall performance of the classification task. Future work could explore this trade-off between model complexity, computational cost, and performance enhancement.

Suggestions for Improvement

In the current implementation of our PSO algorithm, the cognitive and social scaling factors (c_1, c_2) and the inertia weight (ω) are held constant across all iterations until convergence. While this approach has proven effective in our specific classification task, it may limit the algorithm's adaptability and convergence speed.

As a potential enhancement, we propose dynamic scaling factors and inertia weight. Specifically, the values of c_1, c_2, and ω could be adjusted at each iteration based on the algorithm's performance or the iteration number. This dynamic adjustment could allow the PSO algorithm to balance exploration and exploitation more effectively, potentially leading to faster and more accurate convergence.

In the broader perspective, given more time and resources, we would also consider testing our approach on larger and more diverse datasets, experimenting on more complex neural networks and exploring other swarm intelligence techniques. These considerations could add to the practicality and effectiveness of our approach, potentially leading to its adoption in astronomical research and beyond. Ultimately, our goal is to contribute to the ongoing efforts to decipher the cosmic narrative, one star at a time.