Swarm Behaviour Classification

Statistical Models for High Dimensional Data
Project Presentation

Horatiu Palaghiu, Leonardo Schiavo

January 9th, 2024

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Outline

Introduction

Recent studies on various animal species have highlighted the development of swarm intelligence, emphasizing the distinction between random and organized actions. The focus of this study is to identify the characteristics of organized swarm behavior.

Dataset

The dataset is a syntetic one composed of 24017 snapshots of

- 200 boids (bird androids)
- moving freely in a box of dimensions 1400x1000
- 16 different handcrafted behaviours
- target: classification between grouped, flocking and aligned or none of them.
  - Aligned refers to their movement in the same direction
  - Flocking behaviour refers to the way that groups of birds, insects, fish or other animals, move close to each other without ever colliding
  - Grouped refers the way that they clustered together

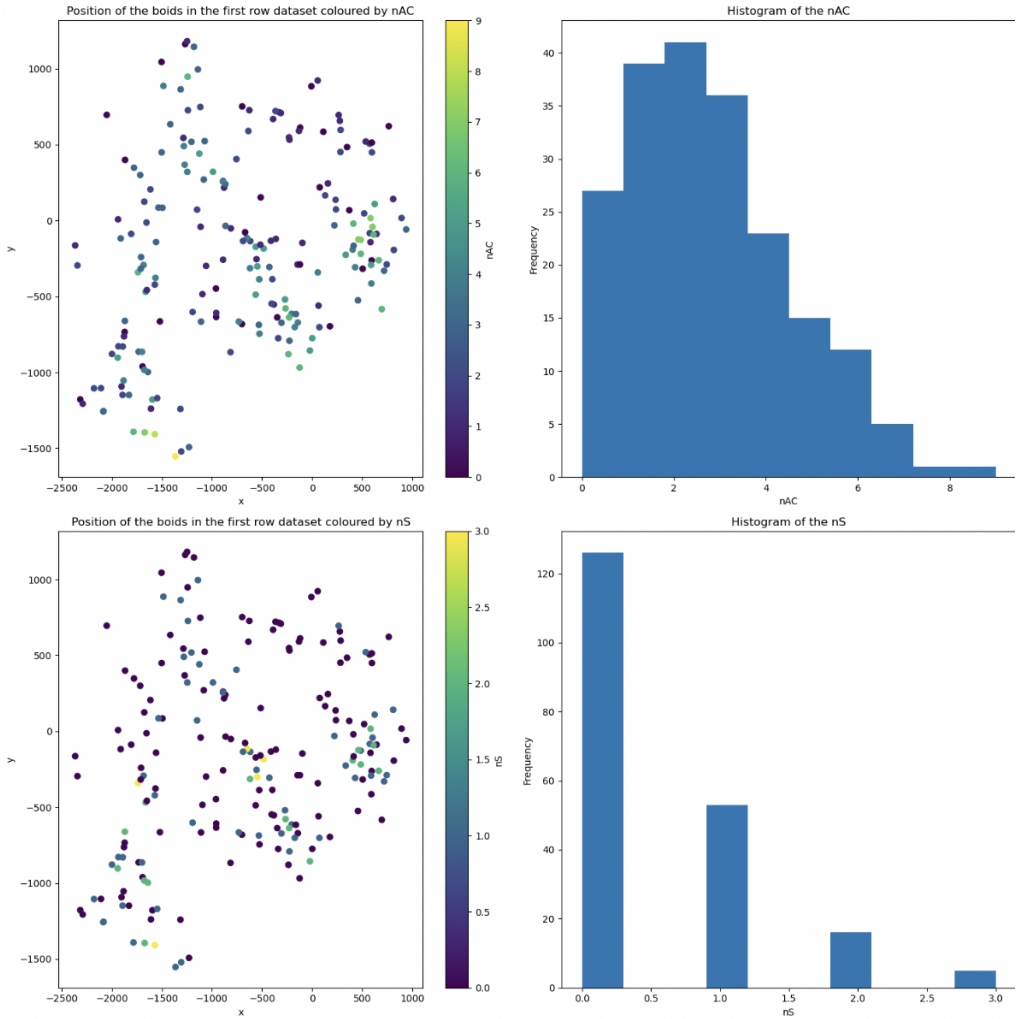
Dataset

Every raw of the dataset corresponds to a snapshot of the 200 boids, for every boid have been measured:

- x_m and y_m as the (X, Y) position in the grid
- $xVel_m$ and $yVel_m$ as the velocity vector
- xA and yA as the alignment vector
- xS and yS as the separation vector
- xC and yC as the cohesion vector
- nAC as the number of boids in the radius of Aligmennt/Cohesion
- nS as the number of boids in the radius of Separation

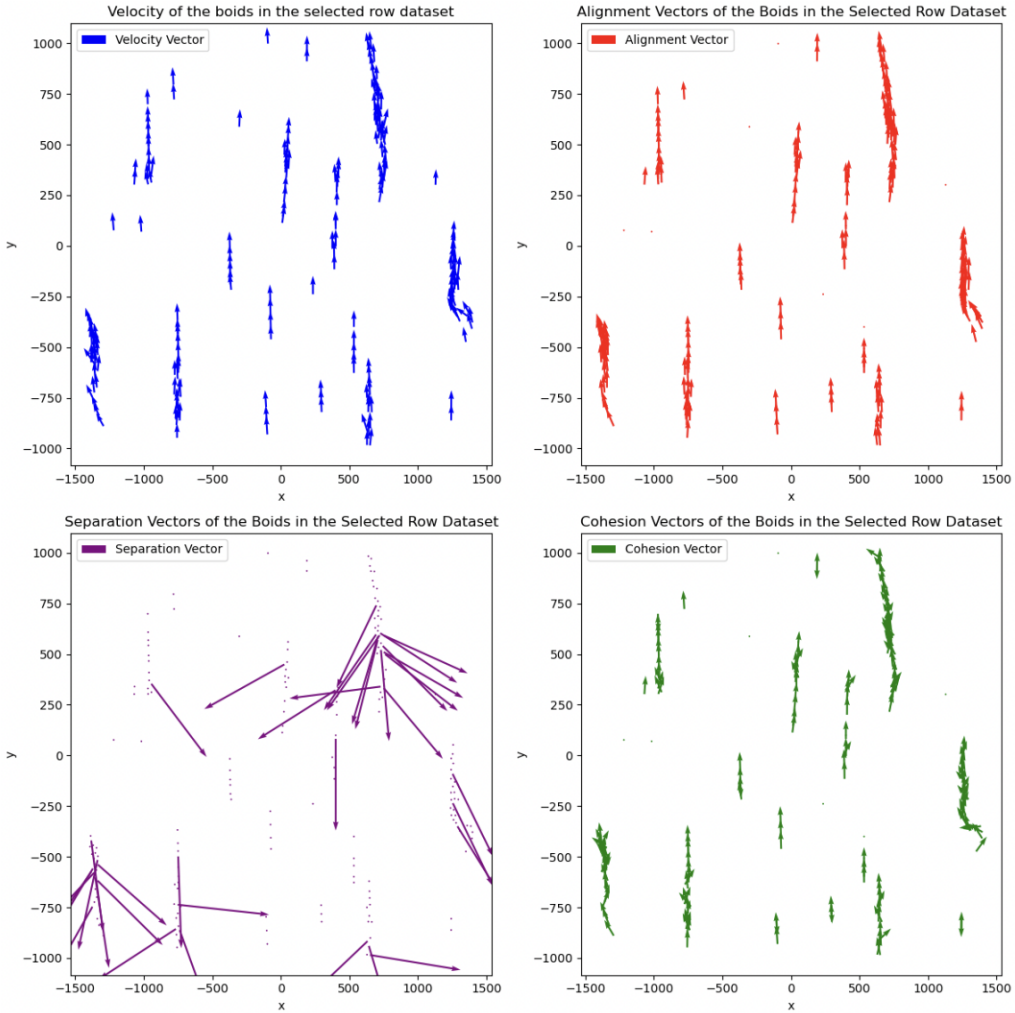
Dataset

The features nAC and nS:

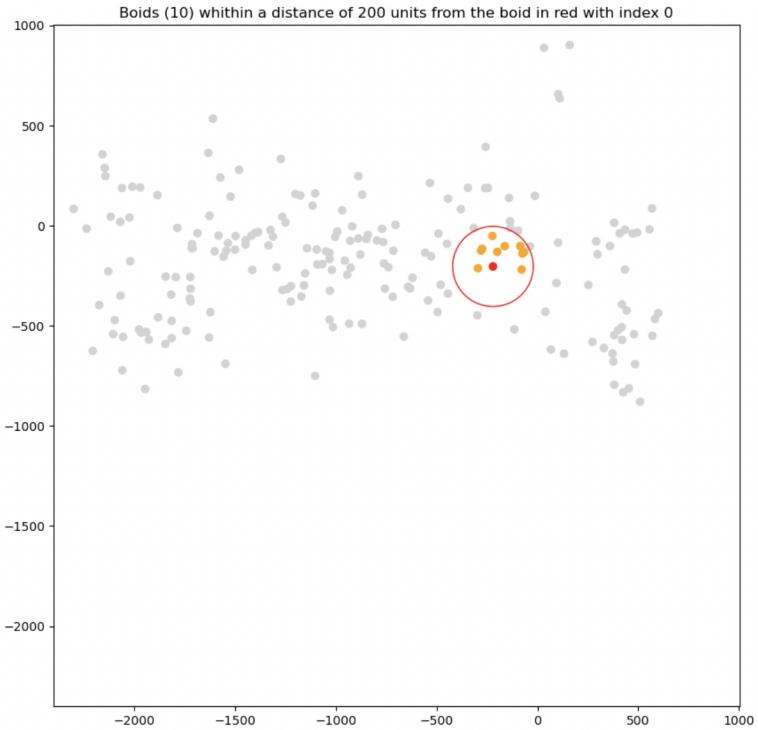


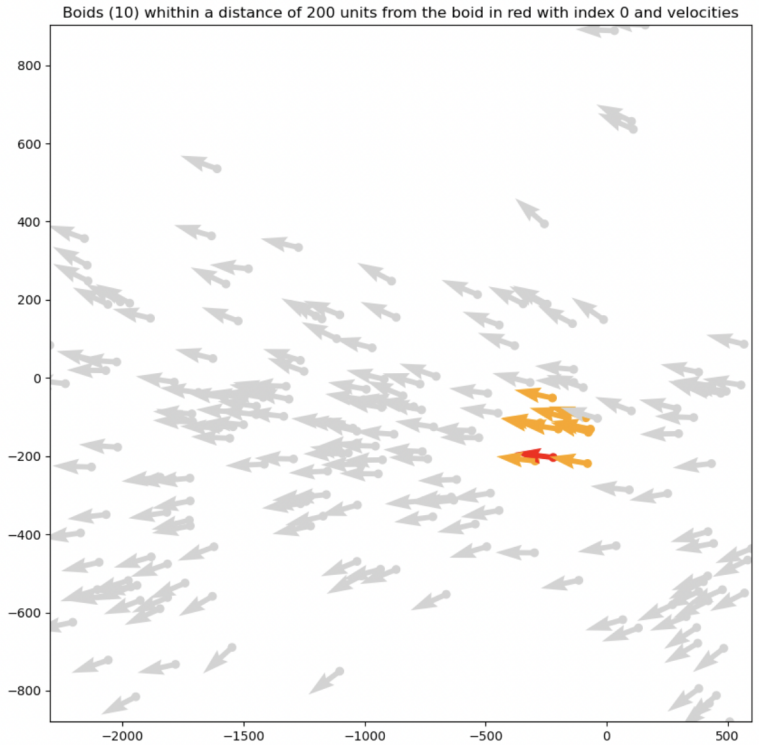
Dataset

The velocity, alignment, separation and cohesion vectors:



Dataset

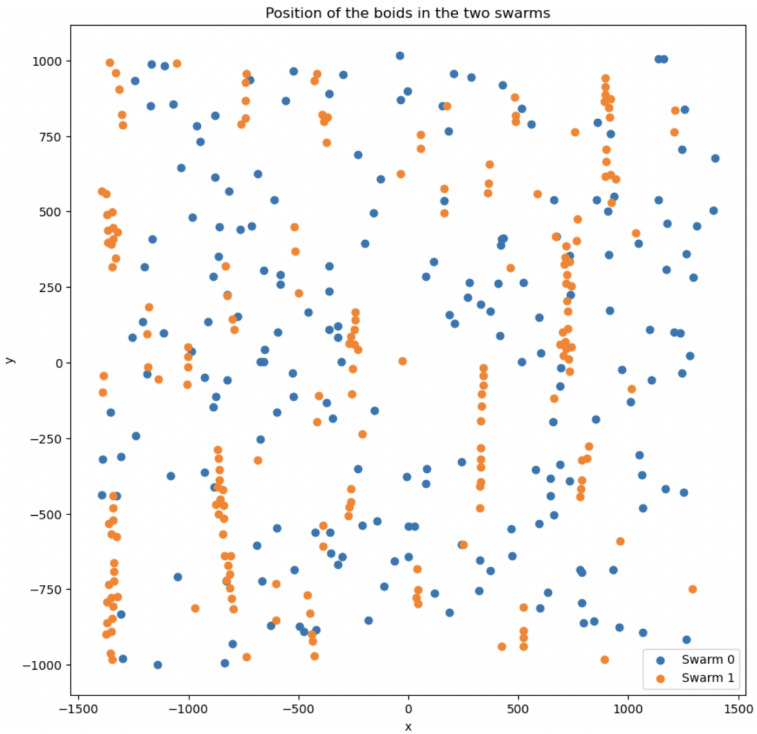




The alignment vector of the red boid is the mean of the orange velocities

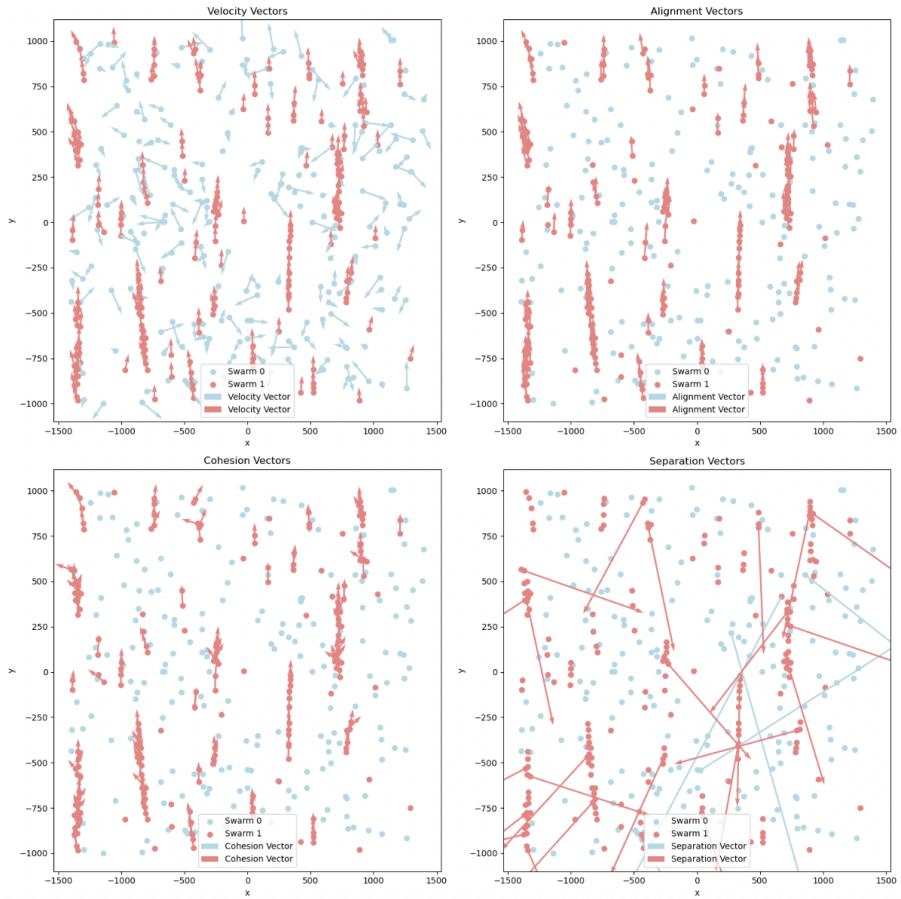
Dataset

Are the classes well separated?



Positions of the boids from two different classes

Dataset



Vectors of the boids from two different classes

Models

These are the models we tested:

- Linear Regression
- Logistic Regression
- Lasso Classifier - Linear Model
- Ridge Classifier - Linear Model
- Elastic Net Classifier
- Group Lasso Classifier - Logistic Model
- Sparse SVM Classifier
- Generalized Lasso Classifier

Methodology

- Most of the methods we performed were tuned by a cross validation with 5 to 10 folds
- In those cases the training set was reduced to 1000 instances
- In the case of Generalized Lasso we have been forced by computational limitation to skip it and use a validation set instead.

Baseline Results

Model	Accuracy	Precision	Recall
Linear Regression	90.45 %	88.76 %	92.63 %
Logistic Regression	90.74 %	89.08 %	92.88 %

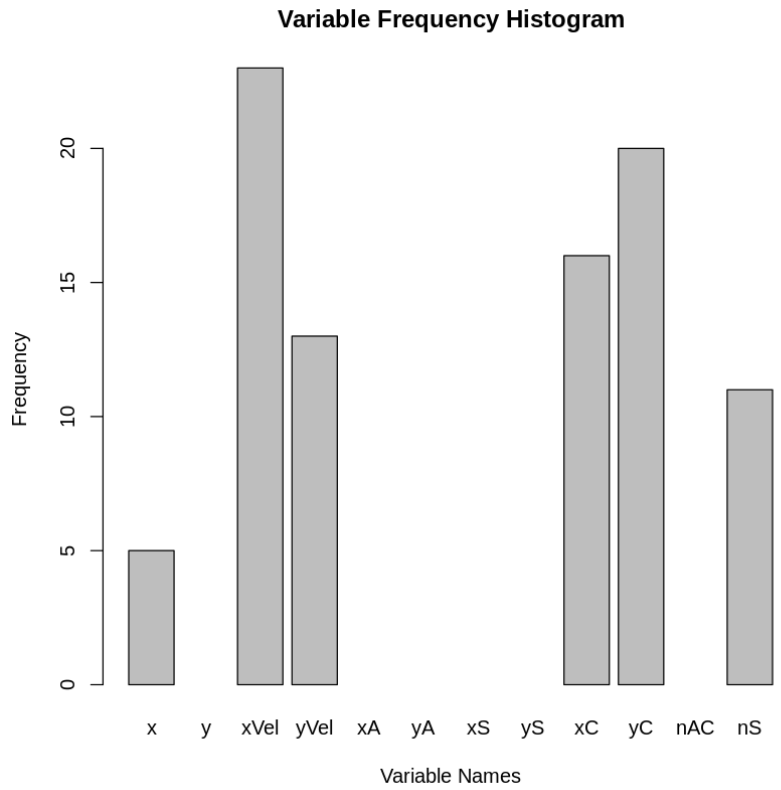
Basic Models Results

Results

Model	Accuracy	Precision	Recall
LC	91.00 %	92.55 %	89.76 %
RC	88.53 %	89.68 %	87.68 %
EN	91.11 %	93.00 %	89.62 %
GL	91.50 %	88.21 %	95.82 %
sSVM	86.06 %	88.32 %	84.51 %
GenL	70.50 %	95.06 %	64.17 %

Models Results

Vabiale analysis



Lasso Classifier - linear model - variables selected

Conclusion

- Lasso Classifier, Elastic Net and Group Lasso were the best models
- The variables with most impact were the velocity and coherence vector
- The overall accuracy increased using those techniques

Thanks for your attention!