

DV

About

This application displays n-Dimensional data in 2D using GLC-L coordinates. For better class separation, linear discriminant analysis is used to get the optimal angles and threshold for a visualization.

Adjustments to the angles and threshold can be done by using the related sliders. Graphs produced by this program can be panned, zoomed in/out, and scaled. Graph order can be rearranged. Analytics generated by this program include the "All Data," "Data Without Overlap," "Overlap Data," and "Worst Case," confusion matrices as well as k-fold cross validation.

Dataset Information

- Dataset must be in .csv format
- Dataset must include a reader row
- If there is an ID column, it must be first,
- If there is a class column it must be last
- Dataset features besides "class" must be numeric

Example Dataset

ID	feature1	feature2	feature3	class
1	5	1	1	dog
2	5	4	1	dog
3	3	1	1	dog
4	6	8	1	cat
5	4	1	1	cat
6	8	10	1	cat
7	1	1	1	bird
8	2	1	1	bird
9	2	1	5	bird

Requirements

- Windows 10

Install and Run

1. Download DV.zip
2. Unzip DV.zip
3. Run "DV.exe"
4. Follow instruction in "Run Instructions.pdf"

Dataset Links

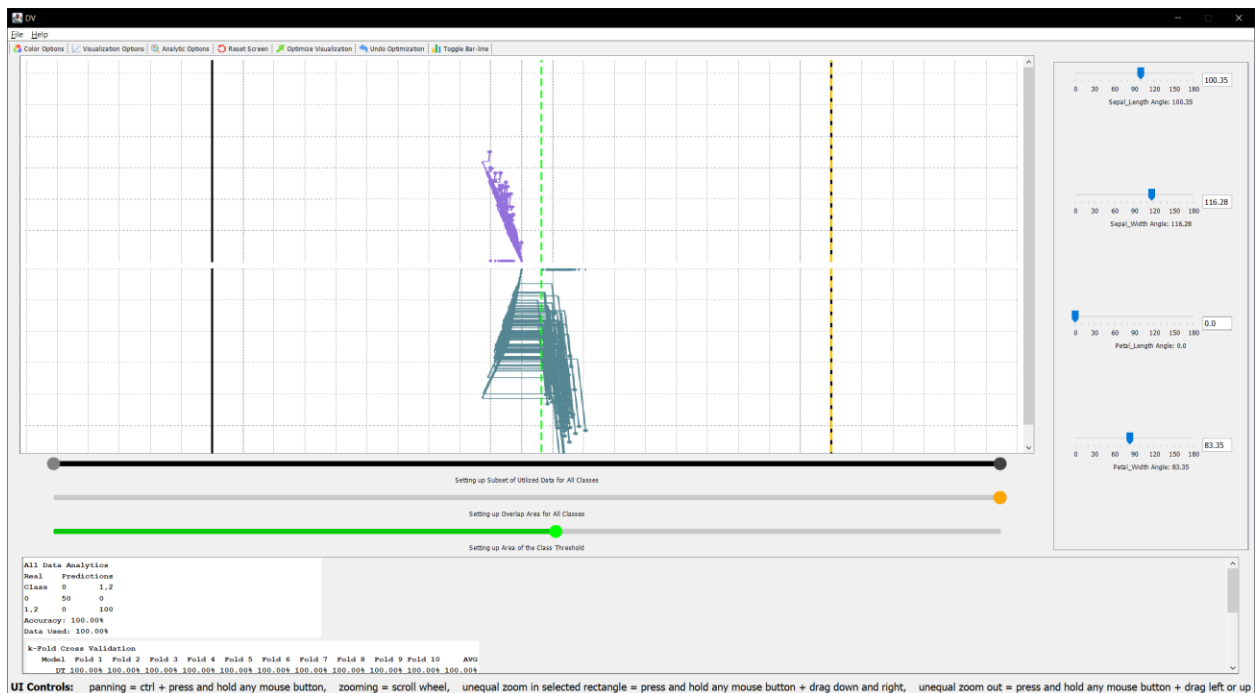
[Iris dataset](#)

[Breast Cancer Wisconsin \(Original\) dataset](#)

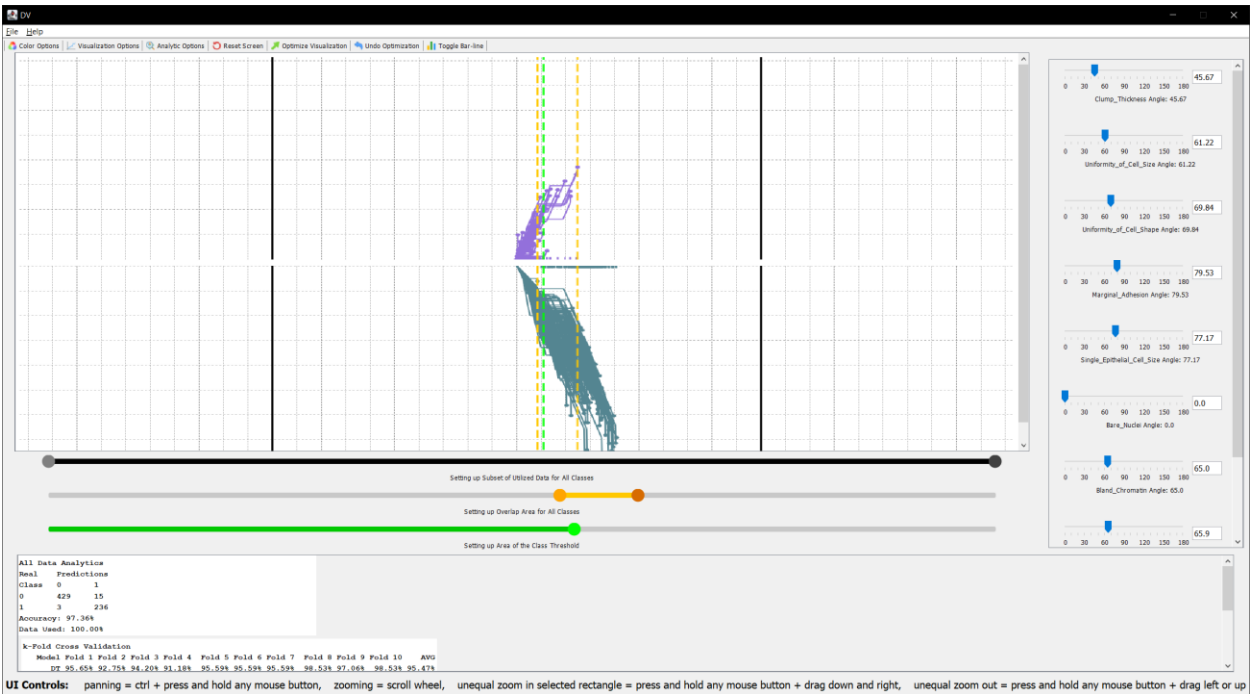
[Seeds dataset](#)

Visualizations

Iris dataset: Iris-Setosa (upper graph) vs Iris-Versicolor and Iris-Virginica (lower graph)



Breast Cancer Wisconsin: Benign (upper graph) vs malignant (lower graph)



Seeds: class 1 (upper graph) vs classes 2 and 3 (lower graph)

