

Cosine Similarity based TopoMap

Angle dependent Data Reduction Algorithm

Modifications

Cosine Similarity and Tree based hull alignment

- Uses Cosine Similarity Matrix instead of conventional Euclidian Distance Matrix.
- Implements tree based hull alignment method (custom) instead of Convex Hull Alignment.
- Encompasses the contact point in the component with linear placement along the y-axis.

Spanning Tree Algorithm

Based on Cosine Similarity

```
Vnew[] = {x}
Enew[] = {}
while Vnew is not equal to V
    u → a node from Vnew
    v → a node that is not in Vnew such that
        edge u-v has the angle between them
    add v to Vnew
    add edge (u, v) to Enew
end while
Return Vnew and Enew
```

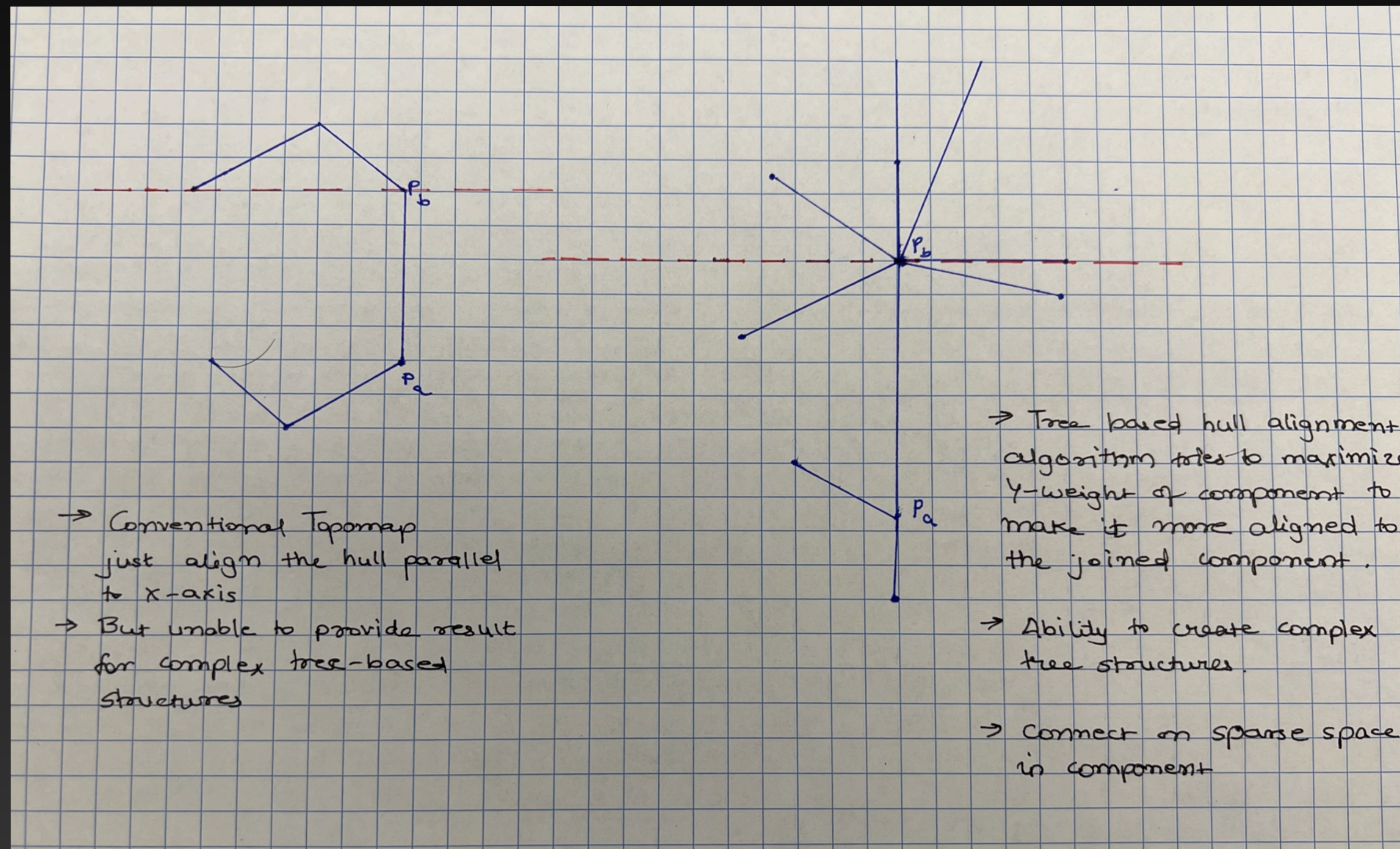
Hull Alignment Algorithm

Uses Custom tree-based Algorithm

```
Cset = [{P1}, {P2}, ..., {Pi}]
Pa = point of interest in Cset
weight = null
for Pi in Cset:
    Pi = Pi - Pa \\ subtract point of interest
                    from each point
make weight as max(concave) or min(convex):
    updated_component, weight = rotate(component) \\ rotate by 1 deg
                                                    in each iteration

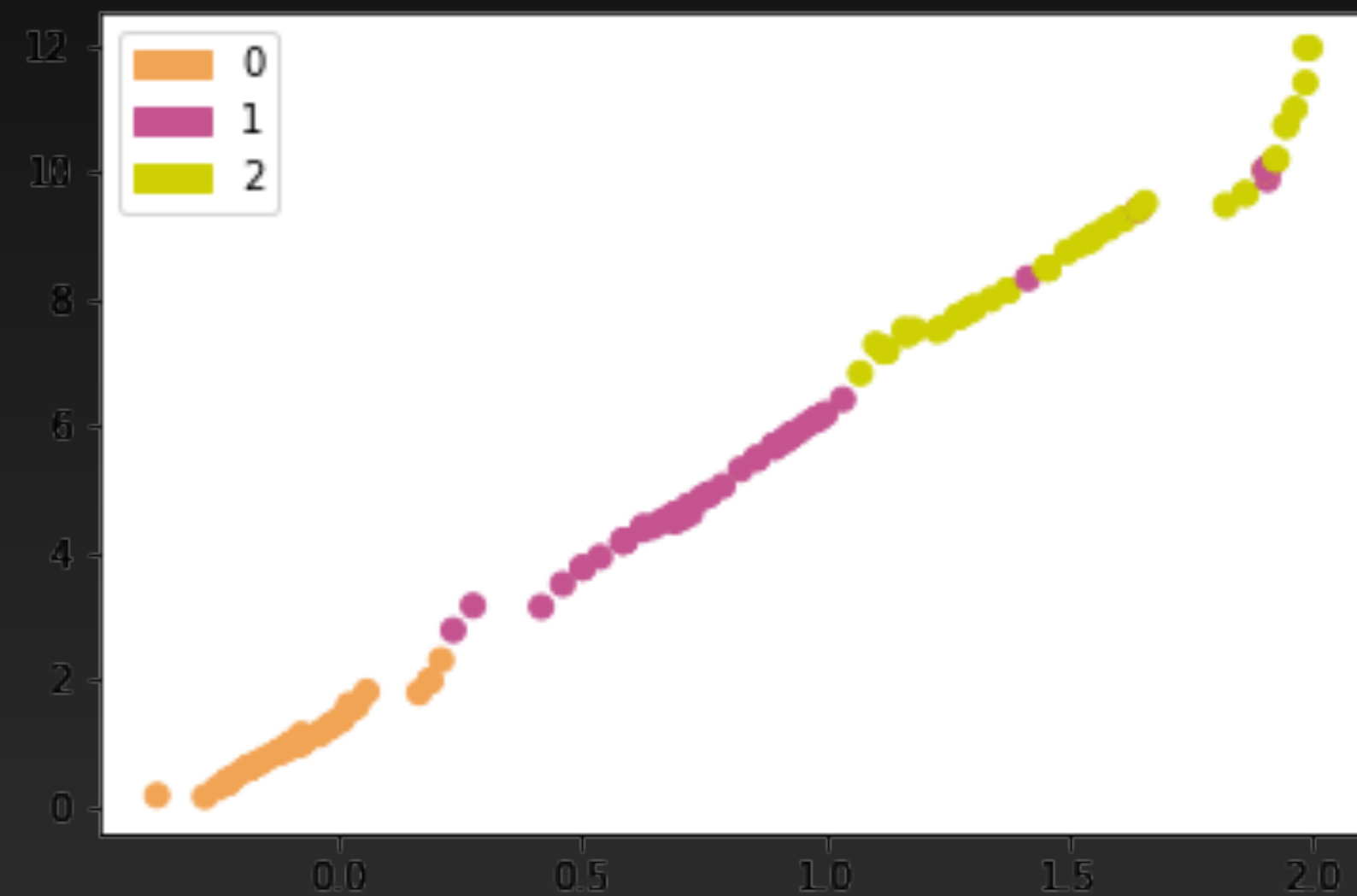
for Pi in Cset:
    Pi = Pi + Pa \\ get the original coordinates back
return updated_component
```


Hull Alignment Algorithm Comparison

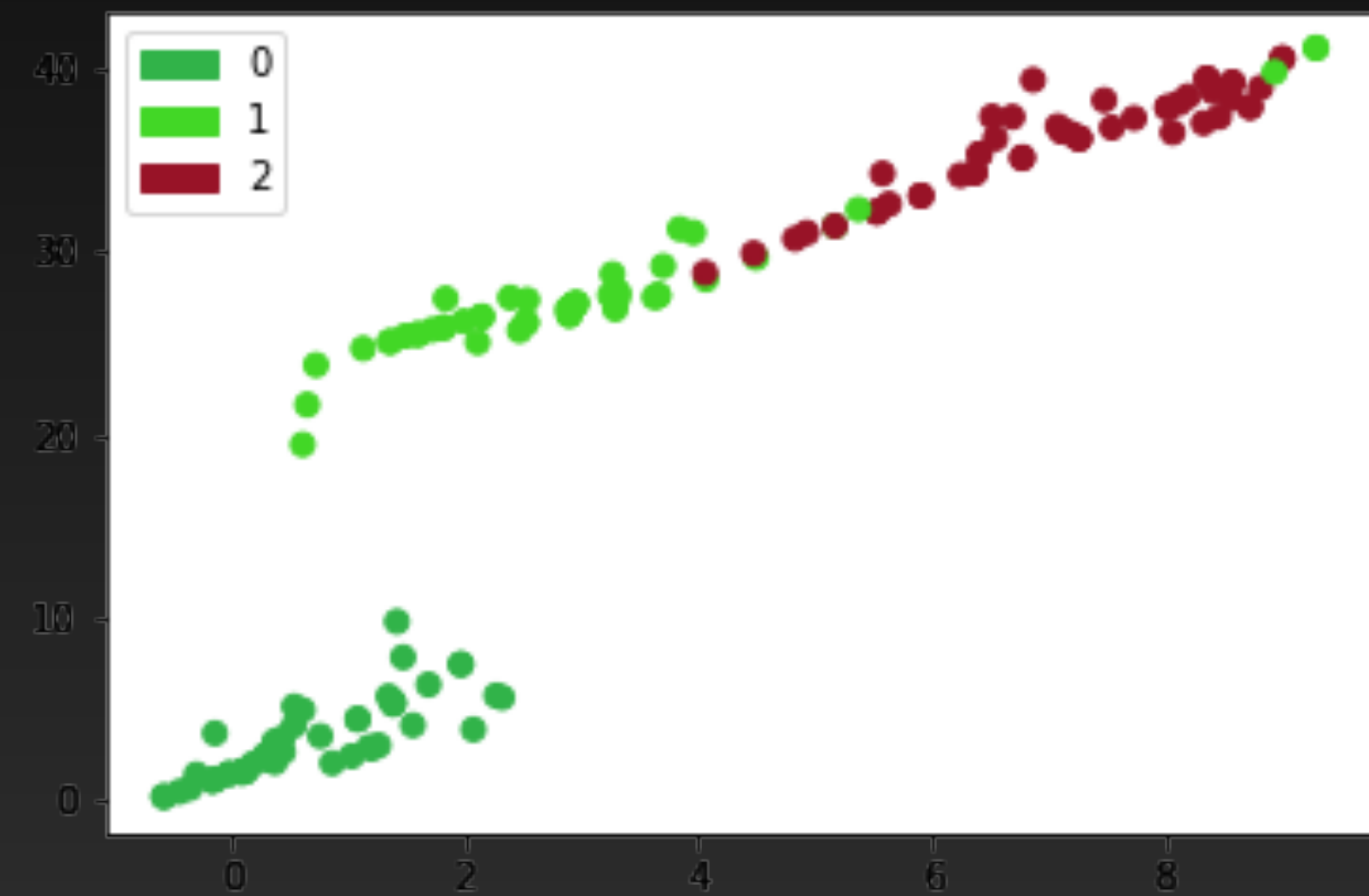


Result Comparison (Iris)

Comparison between Euclidian and Cosine Similarity



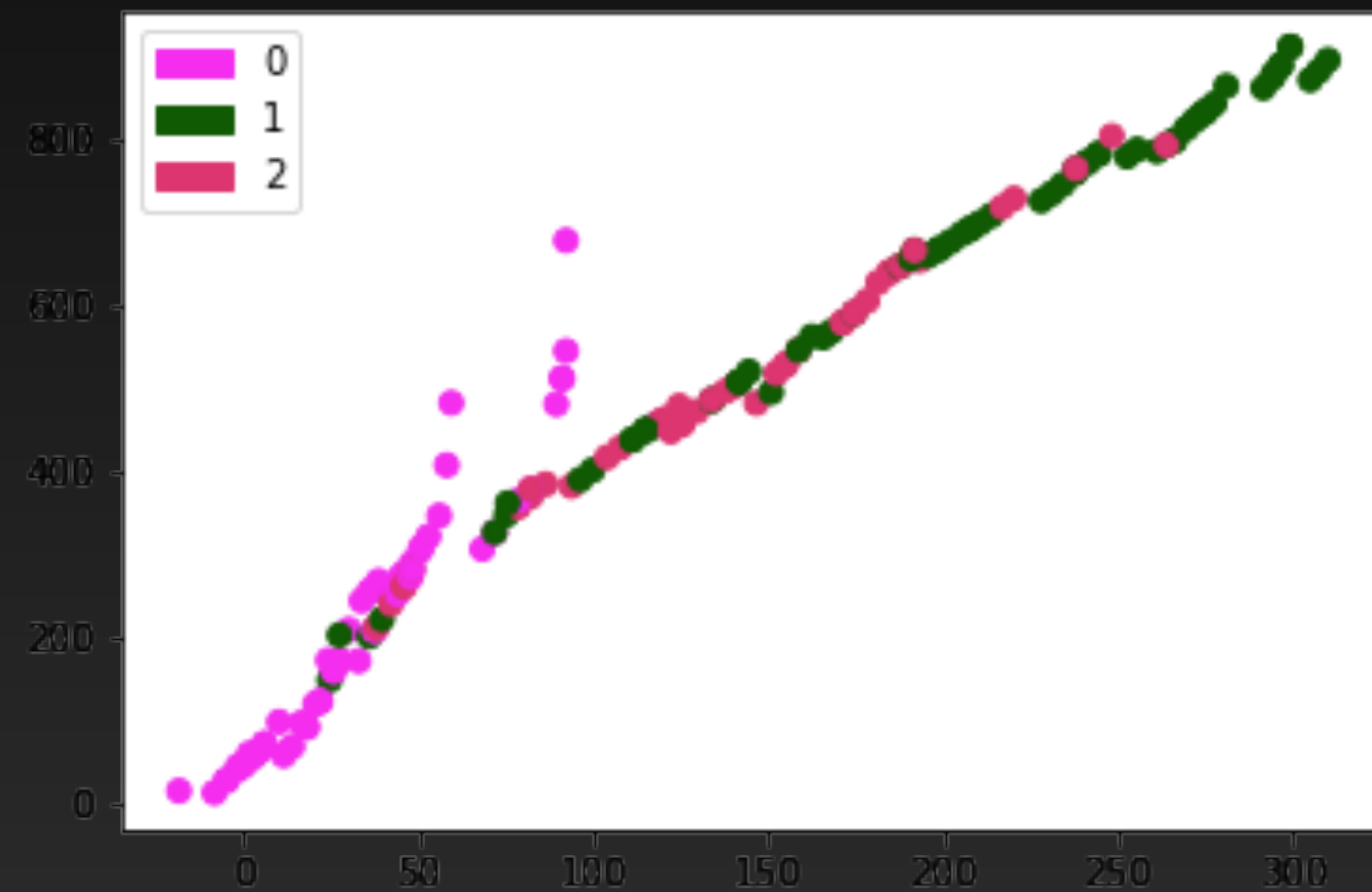
Euclidian Distance Matrix based Result



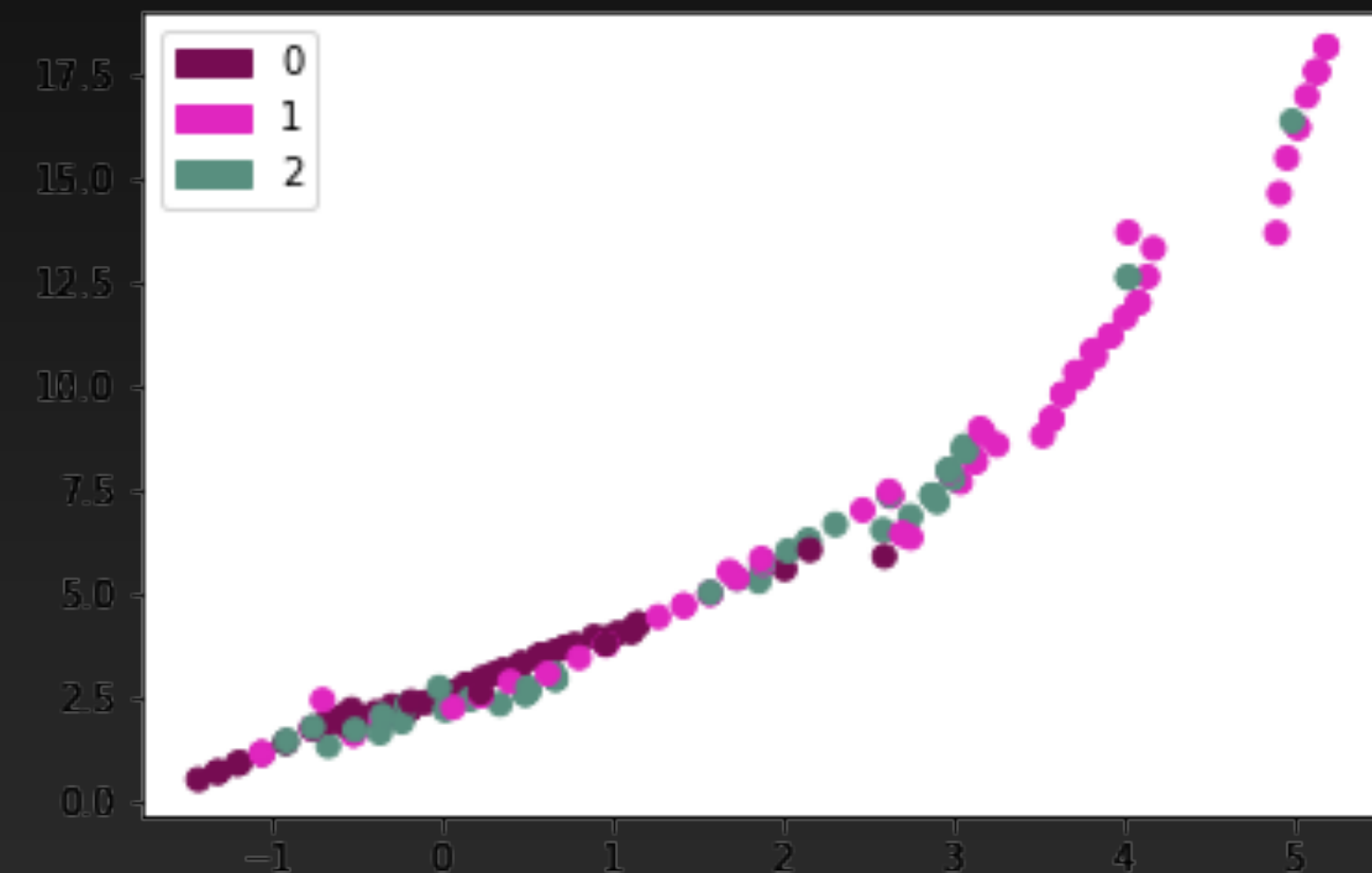
Cosine Similarity Matrix based Result

Result Comparison (Wine)

Comparison between Euclidian and Cosine Similarity



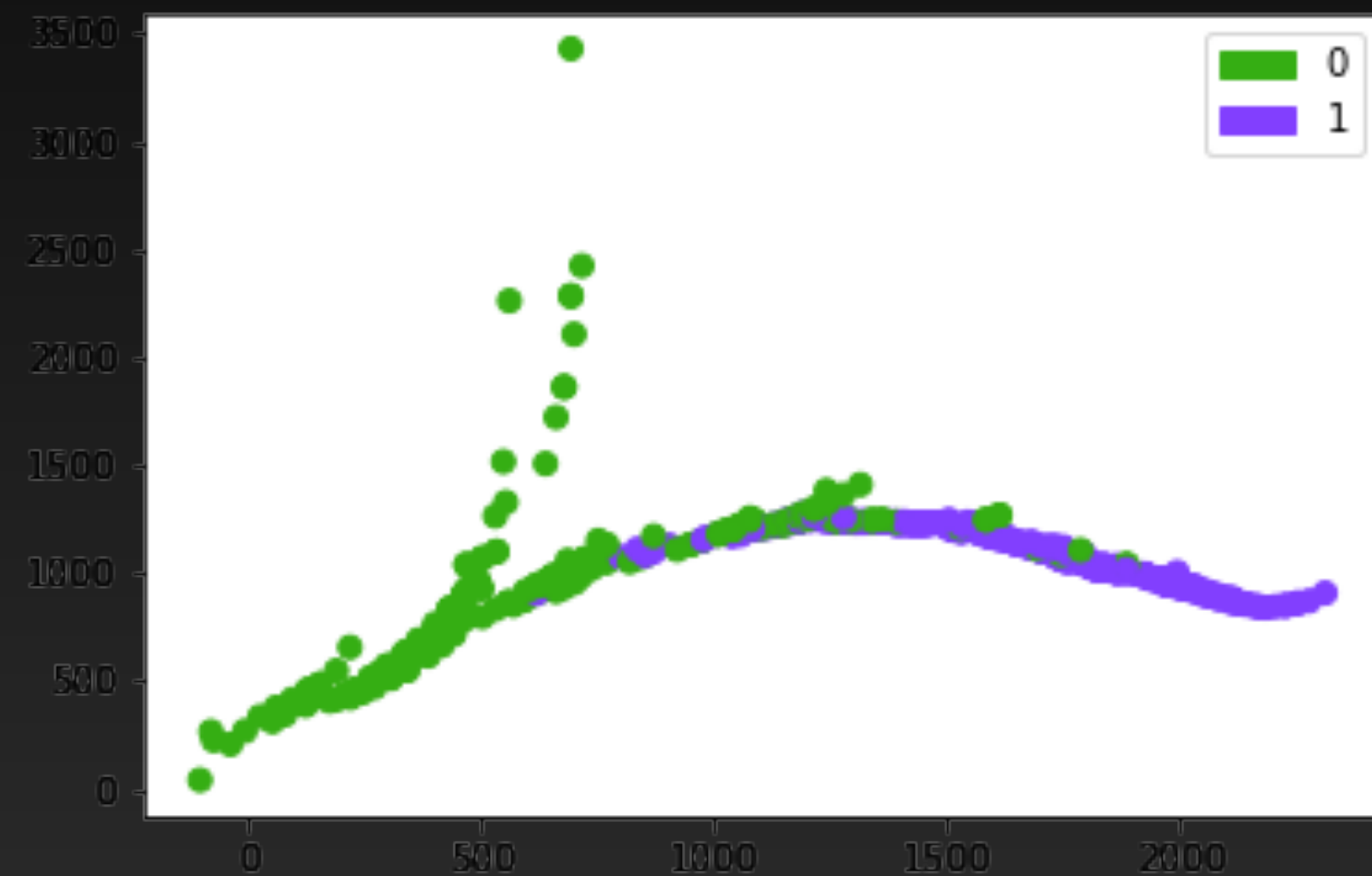
Euclidian Distance Matrix based Result



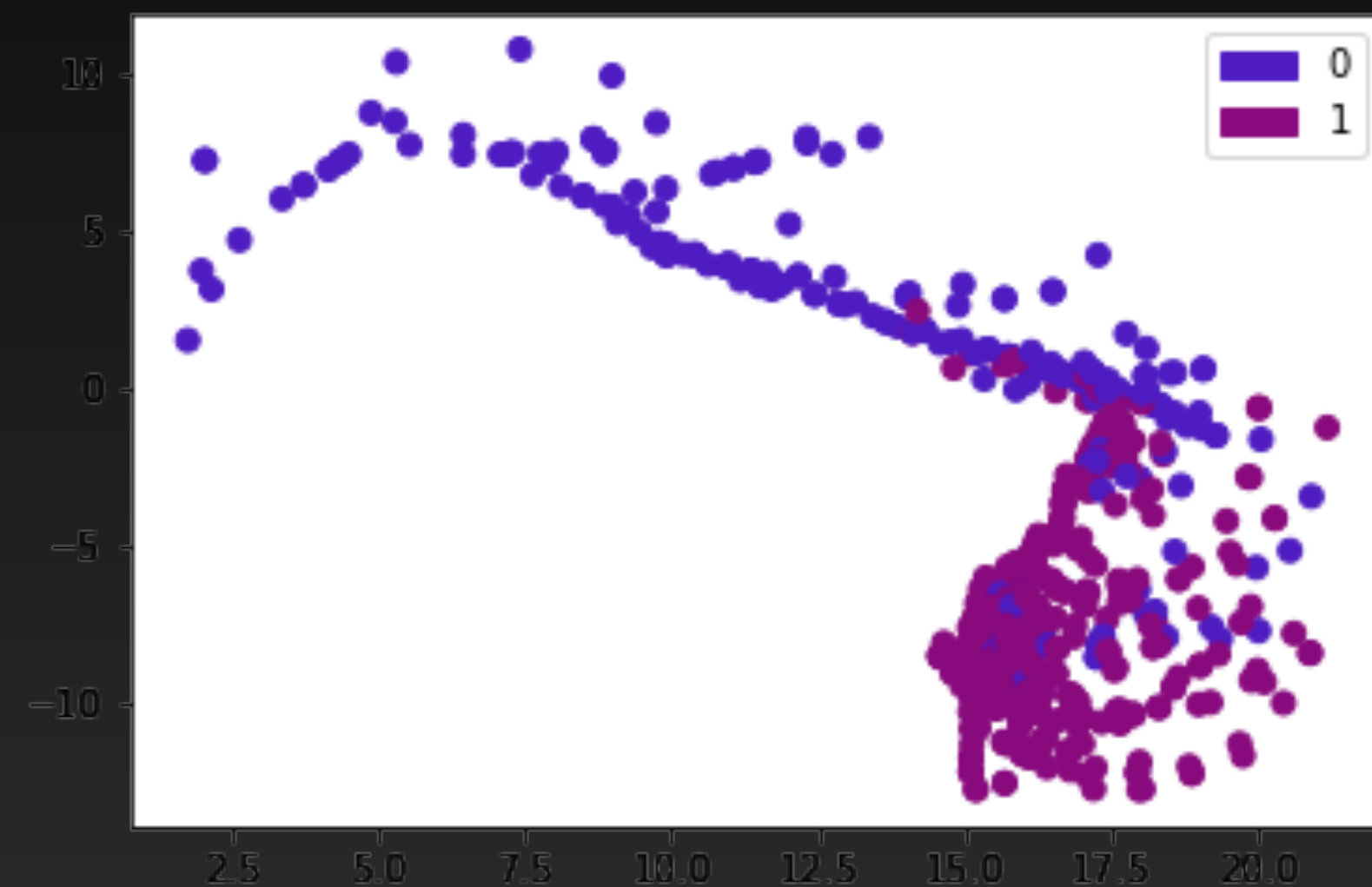
Cosine Similarity Matrix based Result

Result Comparison (Breast Cancer)

Comparison between Euclidian and Cosine Similarity



Euclidian Distance Matrix based Result



Cosine Similarity Matrix based Result

Conclusion

Based on Results

- TopoMap using Cosine Similarity was able to preserve the relation between points more clearly than euclidian distance matrixes.
- The hull alignment algorithm was clearly able to provide the relation of a points with its complete cluster.
- More clear separation between the points and their respective clusters.
- Increase time efficiency with increase computation complexity and minimized computation processing time.