

Laboratoire 1: Extraction de primitives et préparation de données

Cours	GTI770 – Systèmes intelligents
Session	Hiver 2017
N° de laboratoire	1
Groupe	
Étudiant 1 (nom et CP)	Frédéric-Simon Mimeault MIMF24049308
Étudiant 2 (nom et CP)	Shawn-David Merrifield MERS12079102
Étudiant 3 (nom et CP)	François Laplante
Chargé de cours	Alessandro Lameiras Koerich & Hervé Lombaert
Chargé de laboratoire	Rafael Menelau Oliveira e Cruz
Date	31 mai 2017

Rapport	(30 pts)
Classification binaire	/ 10
Classification à trois catégories	/ 10
Classification à quatre catégories	/ 10
Orthographe et présentation	Pénalité jusqu'à 4 pts
Code source	(20 pts)
Clair et efficace	/ 10
Fichiers des primitives et figures correspondants aux trois parties	/ 10
Total	/ 50

1 Problème de classification binaire : Homer X Bart / 10

1.1 Choix des primitives (/4points)

La couleur orange

Nous avons choisi la couleur orange, car c'est l'une des couleurs qui caractérise le plus Bart dans l'ensemble des images disponibles. En plus, elle ne se trouve généralement pas sur les autres personnages, car elle se trouve principalement sur le chandail de Bart.

La couleur bleue

La couleur bleue sert à identifier les pantalons de Bart ainsi que ceux de Homer dans les images de notre ensemble de données. Nous croyons que ces attributs nous permettront d'identifier l'absence de Homer et de Bart dans des images qui n'en possèdent pas.

La couleur blanche

Le blanc est la couleur du chandail de Homer et ne se retrouve pas sur les vêtements des autres personnages, elle est donc indiquée pour identifier celui-ci.

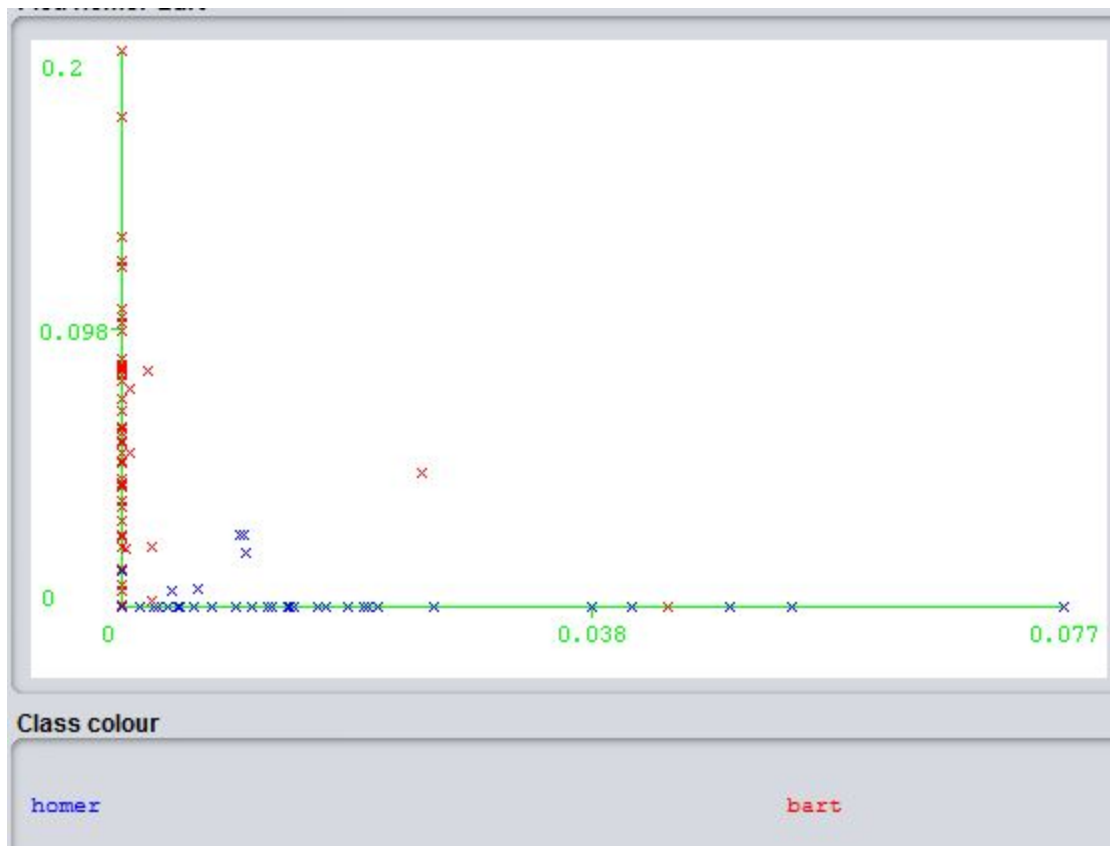
La couleur brune

Comme le blanc n'est pas suffisant pour identifier Homer, nous avons décidé de trouver un autre attribut le représentant. Nous avons donc remarqué que la couleur de la barbe de Homer était de couleur brune et que celle-ci ne se trouve pas ailleurs, alors nous avons décidé de la prendre, car elle était très spécifique à Homer.

La couleur verte

Nous avons choisi comme dernière primitive la couleur verte qui représente la couleur du skateboard de Bart. Le skateboard se retrouve dans plusieurs images de Bart. Cette primitive supplémentaire permettrait d'avoir une autre contrainte pour différencier Bart de Homer.

1.2 Les primitives les plus efficaces (/3 points)



X : Couleur Brun

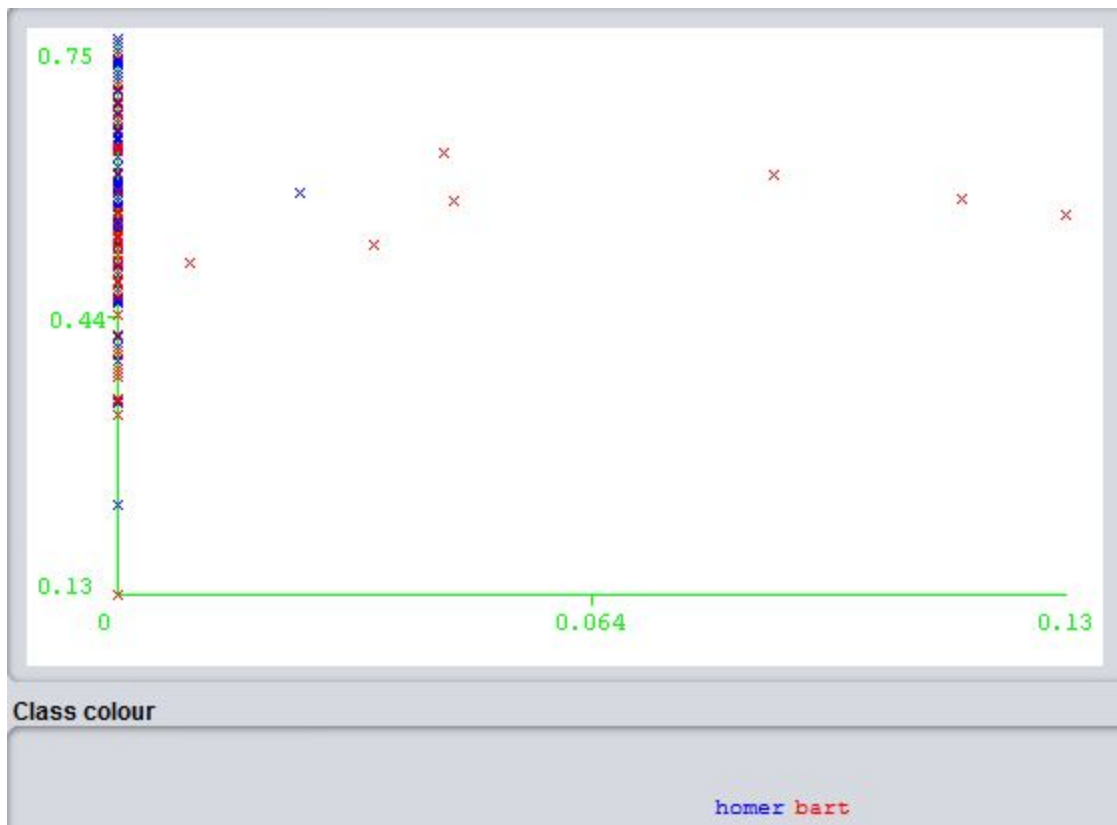
Y: Couleur Orange

Dans le graphique ci-dessus, nous avons la répartition des données en fonction de l'axe horizontal (la couleur brune) et l'axe de vertical (la couleur orange).

Nous constatons que les données ayant un haut pourcentage de brun sont majoritairement associées au personnage Homer, en plus, on voit que plus les données possédant cet attribut sont situées presque exclusivement sur l'axe horizontal ce qui en fait un attribut fortement discriminant, car les données possédant un haut niveau de cet attribut ne possèdent pas un haut niveau de l'autre. De manière plus claire, la raison est que la barbe de Homer est de couleur brune et que Bart n'en a pas.

Pour identifier Bart, nous avons choisi comme attribut la couleur de son chandail soit l'orange. Comme nous pouvons le constater dans le graphique il s'agit d'un attribut discriminant, car cette couleur se trouve uniquement dans l'ensemble de données appartenant à Bart.

1.3 Les primitives les moins efficaces (/3 points)



X : Couleur vert Y: Couleur blanc

Le blanc s'est avéré être moins efficace pour distinguer Homer et Bart, car le blanc est la couleur de fond des images de notre ensemble de données et donc se retrouve dans presque chacune d'elle. Elle n'est donc pas très discriminante pour notre ensemble de données, car elle ne fait pas vraiment de distinction entre les deux classes de sortie

On trouve rarement du vert dans les images de Homer, mais malheureusement on en trouve que très peu souvent dans les images de Bart. Cette primitive peut être utile en tant que primitive secondaire, mais elle n'est que très peu utile comme primitive primaire pour différencier Bart et Homer.

2 Problème de classification à trois catégories : Homer x Bart x Lisa / 10

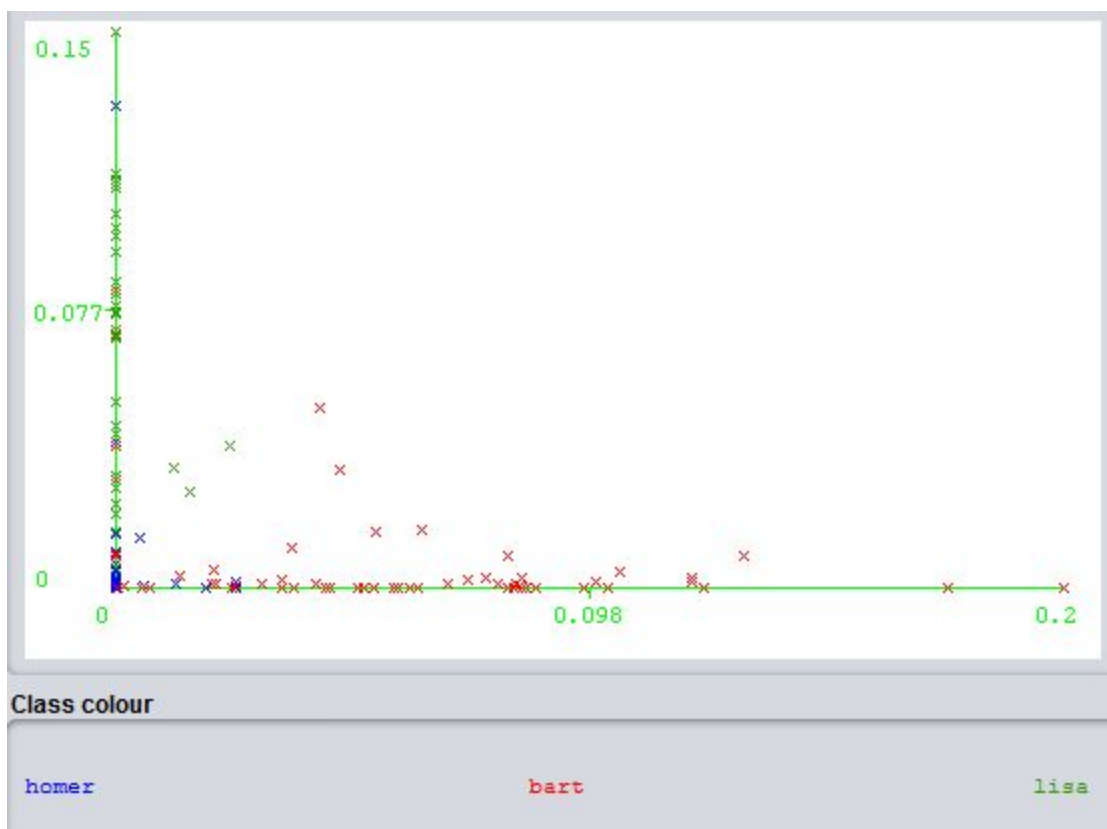
2.1 Choix des primitives (/4points)

Nous avons conservé les mêmes primitives que précédemment et nous avons ajouté la couleur rouge.

La couleur rouge

La raison qui nous a fait choisir la couleur rouge pour représenter Lisa est que la couleur de sa robe est le rouge. En plus, nous avons remarqué que cette couleur ne se retrouve généralement pas dans les mêmes proportions dans les autres images de l'ensemble de données.

2.2 Les primitives les plus efficaces (/3 points)

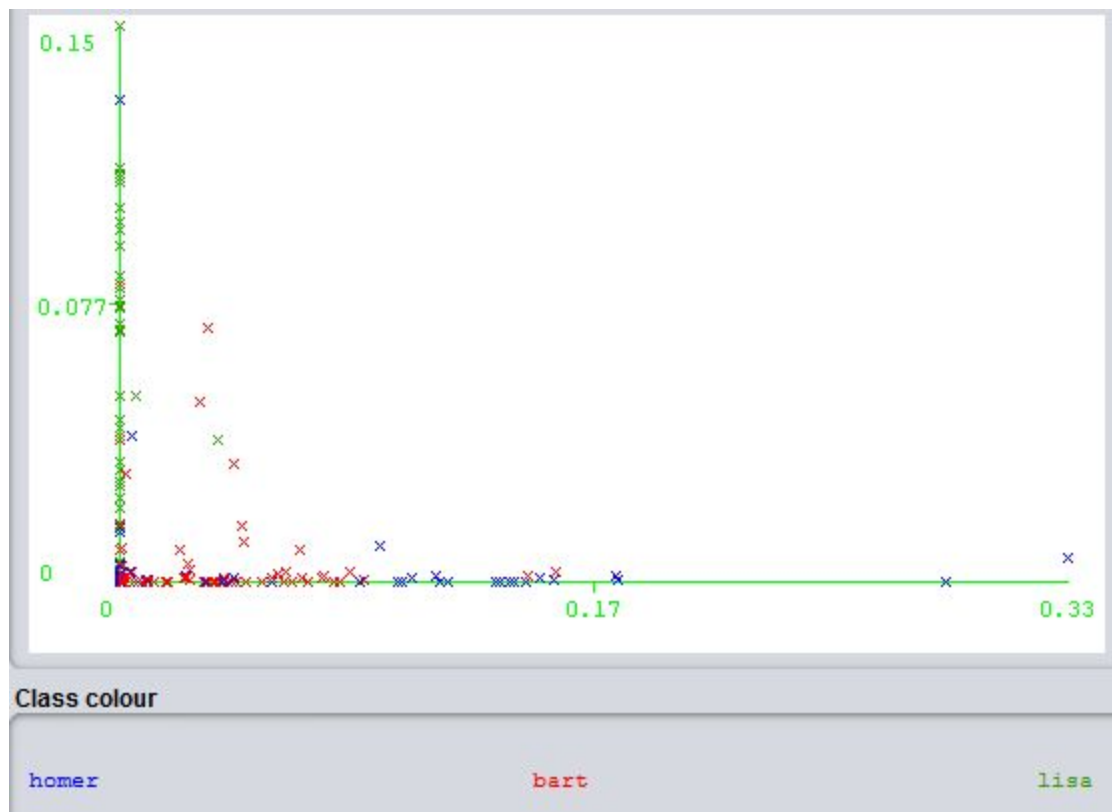


X : Couleur orange Y: Couleur rouge

Tout comme dans la comparaison entre Homer et Bart, le orange permet de bien distinguer Bart. Encore une fois, cette couleur le démarque des deux autres classes. Alors que le rouge permet de bien distinguer la nouvelle classe de Lisa. En comparant ces deux primitives, nous pouvons bien séparer Lisa

et Bart de Homer. Puisque celui-ci n'a très peu d'affinité avec ces deux primitives. Donc, comme on peut voir dans le graphique, Bart est bien représenté sur l'axe x, de l'orange, et Lisa sur l'axe y, le rouge. Finalement, Homer se retrouve proche du zéro. Ces deux primitives sont très efficaces

2.3 Les primitives les moins efficaces (/3 points)



X : Couleur bleu

Y: Couleur rouge

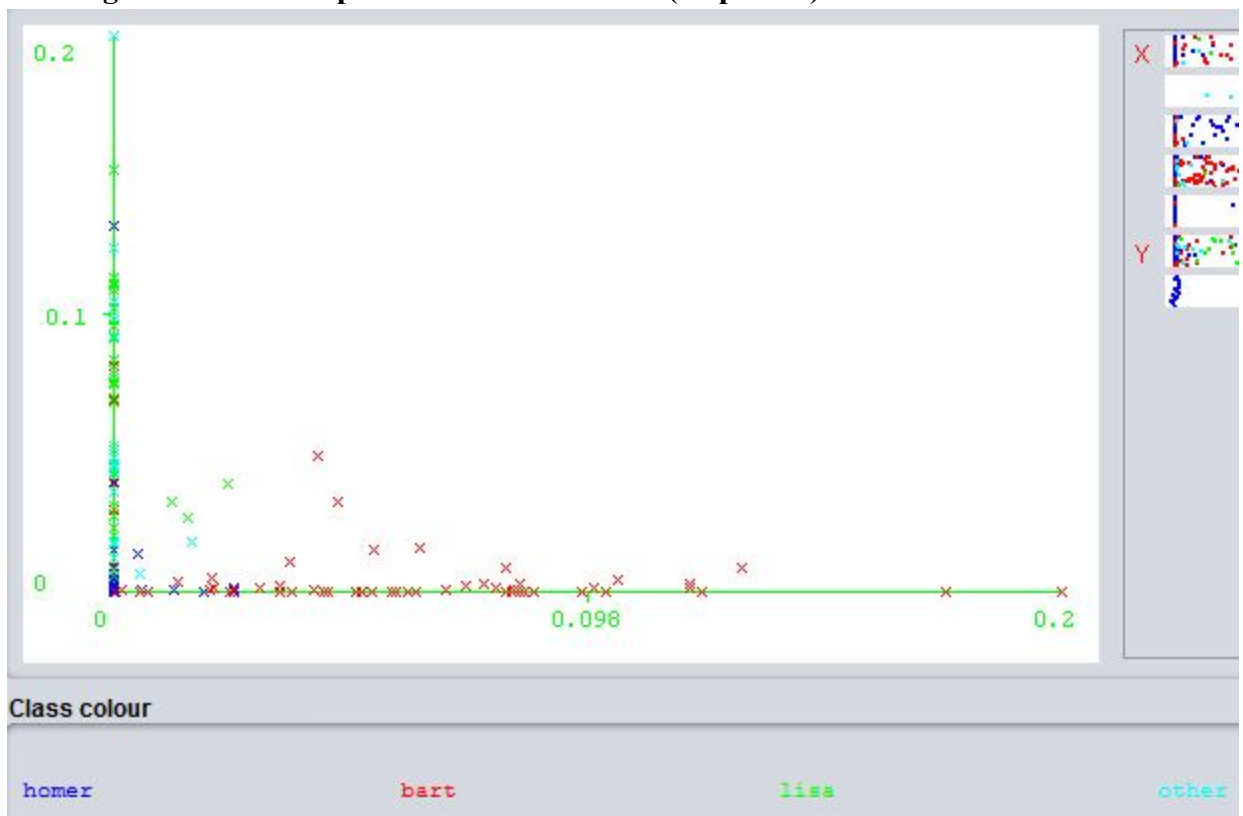
On peut voir par le graphique ci-dessus, que les deux primitives de couleur bleu et rouge sont moins efficaces. La nouvelle primitive rouge permet de bien distinguer la nouvelle classe de Lisa. Par contre, la primitive de couleur bleu ne permet pas de bien distinguer les deux autres classes que sont Bart et Homer. Le bleu est contenu dans les deux classes ce qui ne permet pas de bien les distinguer. Donc ces deux primitives permettent de faire ressortir la nouvelle classe de Lisa mais ne permettent pas de distinguer les 3 en même temps.

3 Problème de classification à quatre catégories : Homer x Bart x Lisa x Others /10

3.1 Choix des primitives (/2 points)

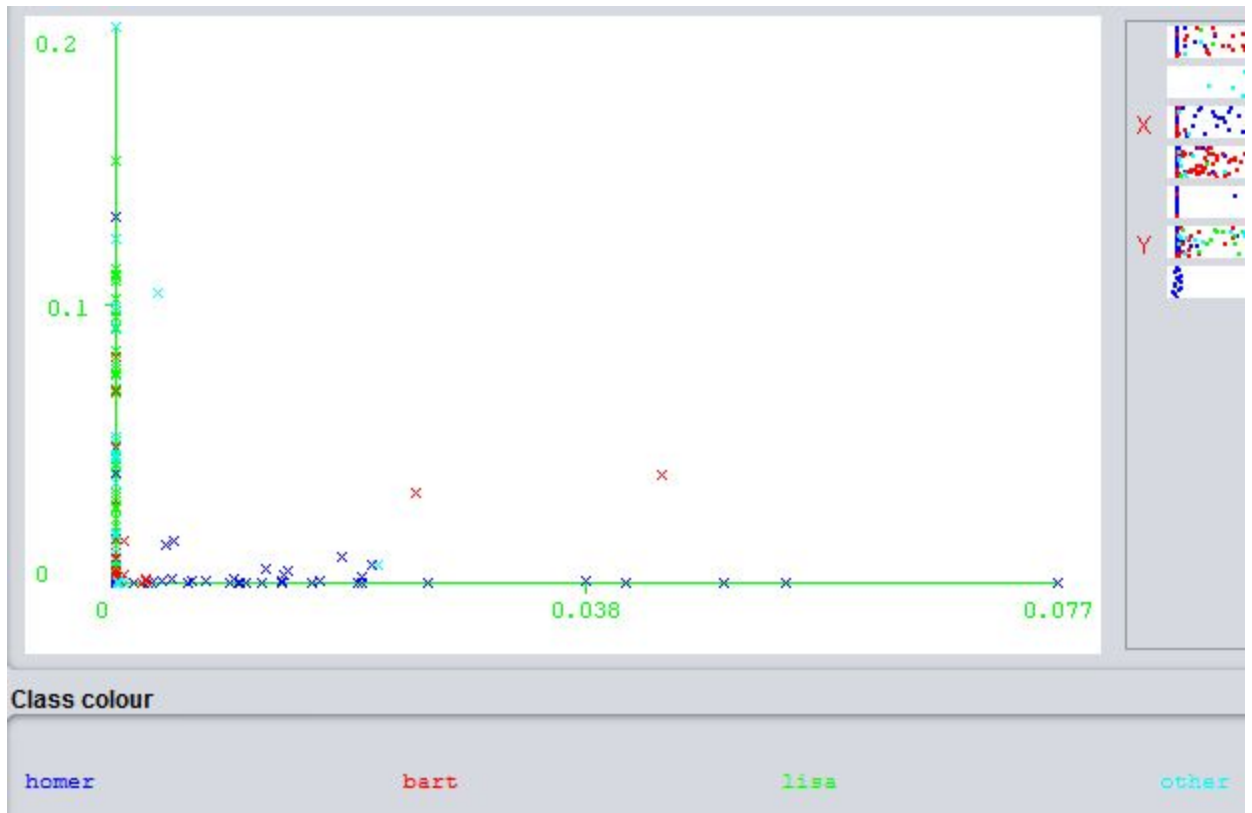
Nous avons conservé les mêmes primitives que précédemment, mais nous avons mis l'emphase sur les couleurs rouge, brun et orange qui permet de distinguer Lisa, Homer et Bart. Les Other seront distingués par leur manque ou leur insuffisance de ces 3 couleurs principalement.

3.2 Avantages et limites des primitives sélectionnées (/4 points)



X : couleur orange

Y : couleur rouge



X : couleur brun

Y : couleur roug



X : couleur brun Y : couleur orange

Le problème c'est qu'avec Weka, nous ne pouvons que comparer que 2 primitives à la fois, donc il est compliqué de bien différencier 4 classes différentes avec seulement 2 primitives principales.

Par exemple, si on pouvait comparer en même temps les 3 primitives montrées dans les graphiques ci-dessus, dans un graphique 3D, nous pourrions avoir une meilleure distinction des 4 classes.

3.3 Selon-vous quelles primitives on pourrait inclure pour avoir une meilleure distinction des 4 classes? (/ 4 points)

On pourrait utiliser les primitives des contours ou d'autre primitive plus précise que les couleurs. Malheureusement, avec le temps que nous avons pour le laboratoire et nos connaissances des outils opencv, nous avons seulement utilisé les couleurs pour différencier les 4 classes.

De plus, pouvoir utiliser 3 primitives au lieu de 2 primitives dans Weka nous aurait permis avec les 3 couleurs citées plus haut (rouge, brun et orange) de distinguer les 3 classes de Lisa, Homer et Bart et ainsi de classer la classe Other dans la région du zéro du graphique.