

1. OAA: (A)

```

citrus?
-N-> lime?
|   -N-> gooseberry?
|   |   -N-> class 0 (356 for class 0, 10 for class 1)
|   |   -Y-> class 1 (0 for class 0, 4 for class 1)
|   -Y-> hint?
|   |   -N-> class 1 (1 for class 0, 15 for class 1)
|   |   -Y-> class 0 (2 for class 0, 0 for class 1)
-Y-> grapefruit?
|   -N-> flavors?
|   |   -N-> class 1 (4 for class 0, 12 for class 1)
|   |   -Y-> class 0 (11 for class 0, 5 for class 1)
|   -Y-> lingering?
|   |   -N-> class 1 (0 for class 0, 14 for class 1)
|   |   -Y-> class 0 (1 for class 0, 0 for class 1)

cherry?
-N-> raspberries?
|   -N-> strawberry?
|   |   -N-> class 0 (225 for class 0, 58 for class 1)
|   |   -Y-> class 1 (0 for class 0, 4 for class 1)
|   -Y-> full?
|   |   -N-> class 1 (0 for class 0, 12 for class 1)
|   |   -Y-> class 0 (1 for class 0, 0 for class 1)
-Y-> cassis?
|   -N-> petit?
|   |   -N-> class 1 (36 for class 0, 68 for class 1)
|   |   -Y-> class 0 (8 for class 0, 0 for class 1)
|   -Y-> allspice?
|   |   -N-> class 0 (21 for class 0, 0 for class 1)
|   |   -Y-> class 1 (0 for class 0, 2 for class 1)

```

When “Citrus” presents, 31/47 were classified as Sauvignon-Blanc, when it absents, 359/388 were classified as not Sauvignon-Blanc. When “grapefruit” presents, 14/15 were classified as Sauvignon-Blanc. So “Citrus” and “grapefruit” are most indicative of being Sauvignon-Blanc. When “hint” presents, 2/2 were classified as not Sauvignon-Blanc, and when “flavors” presents, 11/16 were classified as not Sauvignon-Blanc, so “hint” and “flavors” are most indicative of not being Sauvignon-Blanc. Words that most indicative of being Pinot-Noir are “cherry” (70/135) and “raspberries” (12/13). Words that most indicative of not being Pinot-Noir are “cassis” (21/23) and “petit” (8/8).

(B) The accuracy is about 0.3692. It took about 0.363 seconds. Words that most indicative of being Viognier are “peaches” and “milk”.

(C) The accuracy using zero-one predictions is about 0.25, it is less than using confidence.

AVA:

(A)

```
citrus?
-N-> lime?
|   -N-> melon?
|   |   -N-> class 0 (187 for class 0, 9 for class 1)
|   |   -Y-> class 1 (0 for class 0, 5 for class 1)
|   -Y-> class 1 (0 for class 0, 15 for class 1)
-Y-> class 1 (0 for class 0, 31 for class 1)
-----
crisp?
-N-> lime?
|   -N-> lemon?
|   |   -N-> class 0 (141 for class 0, 9 for class 1)
|   |   -Y-> class 1 (0 for class 0, 8 for class 1)
|   -Y-> grilled?
|   |   -N-> class 1 (0 for class 0, 13 for class 1)
|   |   -Y-> class 0 (1 for class 0, 0 for class 1)
-Y-> red?
|   -N-> class 1 (0 for class 0, 30 for class 1)
|   -Y-> class 0 (2 for class 0, 0 for class 1)

thai?
-N-> very?
|   -N-> between?
|   |   -N-> class 1 (4 for class 0, 56 for class 1)
|   |   -Y-> class 0 (1 for class 0, 0 for class 1)
|   -Y-> ripe?
|   |   -N-> class 1 (1 for class 0, 4 for class 1)
|   |   -Y-> class 0 (4 for class 0, 0 for class 1)
-Y-> class 0 (5 for class 0, 0 for class 1)
-----
apple?
-N-> pasta?
|   -N-> warm?
|   |   -N-> class 1 (11 for class 0, 56 for class 1)
|   |   -Y-> class 0 (3 for class 0, 0 for class 1)
|   -Y-> class 0 (4 for class 0, 0 for class 1)
-Y-> bright?
|   -N-> class 0 (10 for class 0, 0 for class 1)
|   -Y-> reminiscent?
|   |   -N-> class 1 (0 for class 0, 4 for class 1)
|   |   -Y-> class 0 (1 for class 0, 0 for class 1)
```

By analyzing these trees, words that most indicative of being Sauvignon-Blanc are “citrus”, “crisp” and “lime”. Words that most indicative of not being Sauvignon-Blanc are “thai” and “pasta”.

```

crisp?
-N-> lime?
|   -N-> lemon?
|   |   -N-> class 0  (141 for class 0, 9 for class 1)
|   |   -Y-> class 1  (0 for class 0, 8 for class 1)
|   -Y-> meats?
|   |   -N-> class 1  (0 for class 0, 13 for class 1)
|   |   -Y-> class 0  (1 for class 0, 0 for class 1)
-Y-> red?
|   -N-> class 1  (0 for class 0, 30 for class 1)
|   -Y-> class 0  (2 for class 0, 0 for class 1)
-----

```

```

cassis?
-N-> acidity?
|   -N-> duck?
|   |   -N-> class 1  (92 for class 0, 129 for class 1)
|   |   -Y-> class 0  (11 for class 0, 0 for class 1)
|   -Y-> tannins?
|   |   -N-> class 0  (22 for class 0, 0 for class 1)
|   |   -Y-> class 0  (15 for class 0, 11 for class 1)
-Y-> tea?
|   -N-> 100?
|   |   -N-> class 1  (1 for class 0, 47 for class 1)
|   |   -Y-> class 0  (1 for class 0, 0 for class 1)
|   -Y-> class 0  (2 for class 0, 0 for class 1)
-----

```

```

crisp?
-N-> peach?
|   -N-> pear?
|   |   -N-> class 1  (3 for class 0, 142 for class 1)
|   |   -Y-> class 0  (2 for class 0, 0 for class 1)
|   -Y-> class 0  (3 for class 0, 0 for class 1)
-Y-> red?
|   -N-> class 0  (7 for class 0, 0 for class 1)
|   -Y-> class 1  (0 for class 0, 2 for class 1)
-----

```

```

straw?
-N-> crisp?
|   -N-> shellfish?
|   |   -N-> class 1  (8 for class 0, 142 for class 1)
|   |   -Y-> class 0  (2 for class 0, 0 for class 1)
|   -Y-> red?
|   |   -N-> class 0  (7 for class 0, 0 for class 1)
|   |   -Y-> class 1  (0 for class 0, 2 for class 1)
-Y-> class 0  (12 for class 0, 0 for class 1)

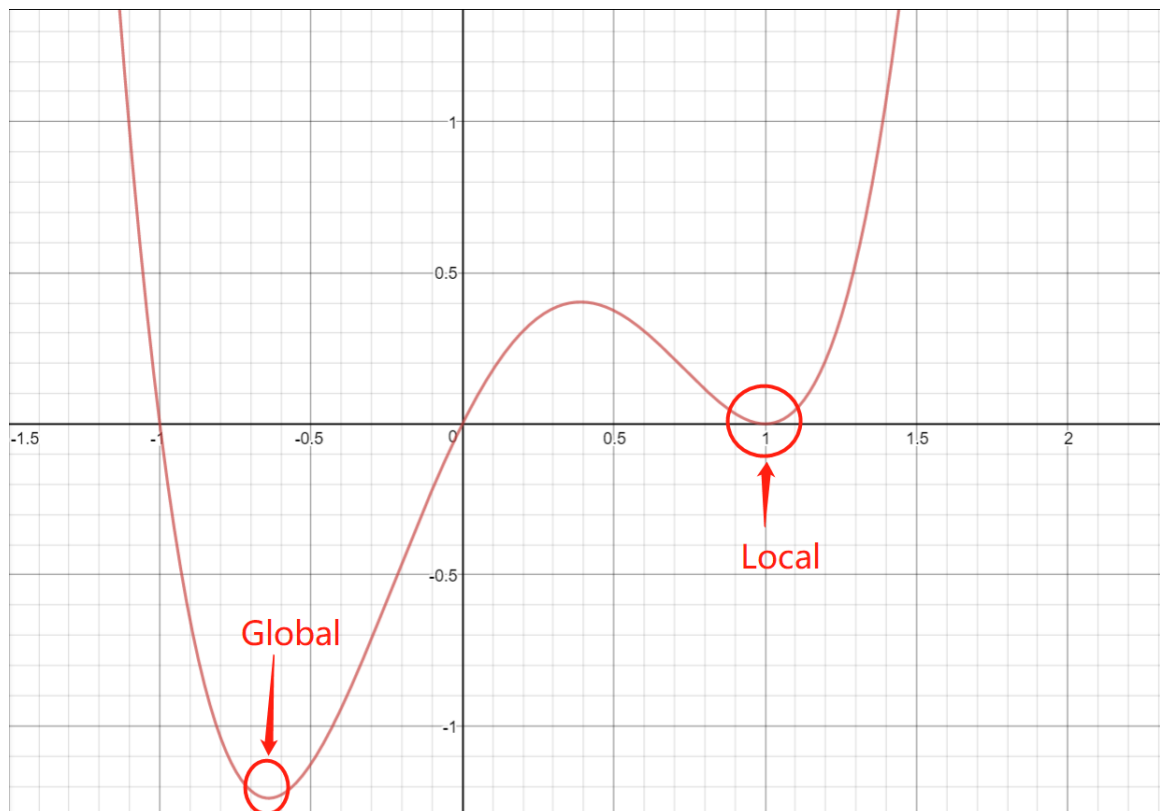
```

Words that most indicative of being Pinot-Noir are “cassis” and “red”. Words that most indicative of not being Pinot-Noir are “duck”, “tea”, “peach”, “shellfish”.

(B) The accuracy is about 0.2680. It took about 0.365 seconds. Words that most indicative of being Viognier are “peaches” and “milk”.

(C) The accuracy using zero-one predictions is 0.269, about the same as confidence prediction.

2. The test accuracy with a balanced tree on the WineData using a DecisionTreeClassifier with max depth 3 is 0.30890538033395176.
3. The step size may affect the speed of converging or diverging. When the step size is negative, the gradient descent algorithm diverges, which means it will never lead you to the correct value. When the step size is 0, there is no change for the approximate value, which means the algorithm has no progress. For step sizes of 0.2, 0.5 and 1, the global minimum can be found at different rounds of the iteration. When the step size is greater than 1, it tends to diverge then converge, but when the step size goes up to ten, it doesn't converge within these 100 iterations.
- 4.



Global minimum: -0.6403882032022079

Local minimum: 0.9999999752397253

We started two calculations, one started at -0.5, and the other one started at 0.8. For the first one, we got -0.6403882032022079, and for the second one, we got 0.9999999752397253, which are referring to global and local minimums we have marked in the plot.

5.

Squared Loss

Training accuracy 0.242915, test accuracy 0.313653

Logistic Loss

Training accuracy 0.995951, test accuracy 0.97417

Hinge Loss

Training accuracy 0.753036, test accuracy 0.686347

The logistic loss function worked the best. It has the highest training and test accuracy.

Top 5 positive words and their weights:

Citrus: 0.8832897531176551

Crisp: 0.7701247691557759

Lime: 0.7108905521536923

Acidity: 0.6891990079029966

Tropical: 0.6064232619016866

Top 5 negative words and their weights:

Blackberry: -0.5321916724675304

Cherry: -0.6295907281434366

Dark: -0.6835931677893792

Black: -0.7653093906427083

Tannins: -1.1695212164040438