**Question # 1:**

**Using all 13 samples, decide whether or not you will play gold if the sample vector is [Sunny, Mild, High, TRUE]. That is, what is the class label for this sample?**

Class label for the given sample “Yes” as its probability (0.0075) is greater than probability of “No” (0.0072)

**Program Output:**



**Question # 2:**

**Using first 10 samples, predict the Yes or No label for the samples 11, 12 and 13. Write your results in the form of a table. Comment on the performance of NB classifier. What would be the class label for the sample given in 1. Does the class label change?**

**Table**

|  |  |  |
| --- | --- | --- |
| **Sample** | **Actual Label** | **Predicted/estimated Label** |
| **11** | Yes | Yes |
| **12** | No | Yes (without smoothing)  No (with smoothing) |
| **13** | Yes | Yes (without smoothing)  Yes (with smoothing) |
| **[Sunny, Mild, High, TRUE]** | **NO** | No |

**Program Outputs:**

1. Sample 11

****

1. Sample 12 (Without Smoothing)



1. Sample 12 (With Smoothing)



1. Sample 13 (Without Smoothing)



1. Sample 13 (With Smoothing)



1. Sample: [Sunny, Mild, High, TRUE]



**Comment on Performance of Classifier:**

* Classifier performed well with 3 out of 4 answers correct in case of no smoothing applied to the probability of feature of Outlook, and 4 out of 4 answers correct in case of smoothing applied to feature of Outlook.
* Furthermore, I believe if there was a larger dataset, the classifier could have performed better.

**Class Label for Sample given in Question 1:**

* Yes the class label has changed as it was Yes with all 13 samples, but it changed to No when the classifier was provided only 10 samples to compute the probability.
* Whereas the actual class label was the same as the calculated label by the classifier program.

**Question # 3**

**Using first 11 samples, predict the Yes or No label for the samples 12 and 13. Write your results in the form of a table. Comment on the performance of NB classifier. What would be the class label for the sample given in 1. Does the class label change?**

|  |  |  |
| --- | --- | --- |
| **Sample** | **Actual Label** | **Predicted/estimated Label** |
| **12** | No | Yes (without smoothing)  No (with smoothing) |
| **13** | Yes | Yes (without smoothing)  Yes (with smoothing) |
| **[Sunny, Mild, High, TRUE]** | **NO** | No |

**Program Outputs:**

1. Sample 12 (Without Smoothing):



1. Sample 12 (With Smoothing):



1. Sample 13 (Without Smoothing):



1. Sample 13 (With Smoothing):



1. Sample: [Sunny, Mild, High TRUE]



**Comment on the performance of Classifier Program:**

* The classifier performed quite well with 3 out of 3 answers correct with smoothing and 2 out of 3 answers correct without any smoothing of the probability of feature of Outlook.

**Class Label for Sample given in Question 1:**

* Yes the class label has changed as it was Yes with all 13 samples, but it changed to No when the classifier was provided only 10 samples to compute the probability.
* Whereas the actual class label was the same as the calculated label by the classifier program.