**RANDOM FOREST (BAGGING) RESULTS**

Notes:

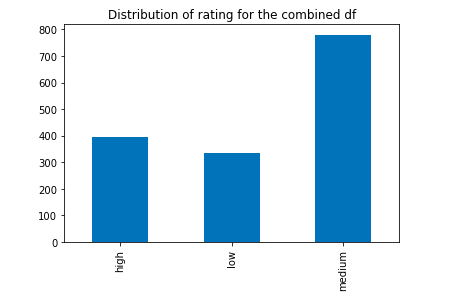
* When predicting properties, the order of the labels for the confusion matrix are

1. Smoothness
2. Thickness
3. Warmth
4. Flexibility
5. Softness

* When predicting rating, the order of the labels for the confusion matrix are

1. low
2. medium
3. high

* Generated random seed for random forest is 183
* When predicting properties, the macro F1 score was used as the data was balanced (same number of observations for each property)
* When predicting ratings, the weighted F1 score was used as the data was unbalanced (refer figure below)



**(1) Estimating the PROPERTY based on available data:**

* Chance classification accuracy: 20% (1/5)

**Method 1: Leave One Participant Out CV**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Average Micro F1 score** | **Average Macro F1 score** | **Overall accuracy** |
| **Leave One Participant Out CV** | 0.39 | 0.34 | 39.13% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Participant left out** | **Number of trees in rf** | **Confusion matrix** | **Micro F1 score** | **Macro F1 score** | **Classification accuracy** |
| 1 | 500 | [[13 3 2 0 0]  [ 0 11 7 0 0]  [ 6 0 12 0 0]  [ 2 12 0 0 4]  [ 2 16 0 0 0]] | 0.4 | 0.32 | 40.00% |
| 2 | 1000 | [[ 6 4 0 5 3]  [ 0 16 0 0 2]  [ 0 0 18 0 0]  [ 0 0 0 17 1]  [ 2 7 0 7 2]] | 0.66 | 0.61 | 65.56% |
| 3 | 1000 | [[15 0 3 0 0]  [ 0 8 0 9 1]  [ 7 0 1 10 0]  [ 1 0 0 17 0]  [ 0 11 0 7 0]] | 0.46 | 0.36 | 45.56% |
| 4 | 1500 | [[12 0 3 0 3]  [ 0 8 2 1 7]  [ 4 1 4 2 7]  [ 1 1 0 14 2]  [ 2 0 0 6 10]] | 0.53 | 0.52 | 53.33% |
| 5 | 1500 | [[13 0 5 0 0]  [ 3 5 0 0 10]  [13 0 1 1 3]  [ 0 0 1 17 0]  [ 0 0 0 0 18]] | 0.6 | 0.55 | 60.00% |
| 6 | 500 | [[ 3 3 1 1 10]  [ 0 7 0 0 11]  [ 0 7 3 0 8]  [ 0 1 0 5 12]  [ 0 0 0 1 17]] | 0.39 | 0.36 | 38.89% |
| 7 | 1500 | [[ 4 12 2 0 0]  [ 1 15 0 0 2]  [ 7 6 5 0 0]  [ 0 16 0 0 2]  [ 2 14 0 2 0]] | 0.27 | 0.2 | 26.67% |
| 8 | 1500 | [[ 6 0 1 11 0]  [ 5 1 4 8 0]  [ 9 0 1 8 0]  [ 3 1 0 13 1]  [ 1 0 1 12 4]] | 0.28 | 0.24 | 27.78% |
| 19 | 1000 | [[ 0 11 1 5 1]  [ 0 6 1 11 0]  [ 0 4 4 10 0]  [ 0 1 0 17 0]  [ 0 10 1 7 0]] | 0.3 | 0.21 | 30.00% |
| 21 | 500 | [[13 1 2 1 1]  [11 5 0 2 0]  [ 4 5 0 7 2]  [ 7 3 2 3 3]  [15 1 0 1 1]] | 0.24 | 0.19 | 24.44% |
| 22 | 1500 | [[10 5 2 0 1]  [ 1 10 6 0 1]  [ 3 6 6 2 1]  [ 1 7 5 4 1]  [ 6 2 5 4 1]] | 0.34 | 0.32 | 34.44% |
| 23 | 1000 | [[14 2 0 0 2]  [ 7 4 0 6 1]  [ 8 2 1 7 0]  [ 3 2 1 10 2]  [12 2 0 2 2]] | 0.34 | 0.29 | 34.44% |
| 24 | 1000 | [[15 3 0 0 0]  [ 8 3 1 0 6]  [ 5 3 1 2 7]  [ 8 0 4 2 4]  [16 1 1 0 0]] | 0.23 | 0.18 | 23.33% |
| 25 | 1500 | [[ 5 1 9 3 0]  [ 0 6 4 8 0]  [ 0 0 17 1 0]  [ 0 0 8 10 0]  [ 1 2 9 5 1]] | 0.43 | 0.39 | 43.33% |

**Method 2: Leave One Cloth Out CV**

Made sense to do LOCOCV for the sock data as every participant touched all socks. However, the clothes that the participants in Lili’s experiment touched were different. So thought LOCOCV might not make sense.

Chart, bar chart, histogram

Description automatically generated

**(2) Estimating the RATING based on data:**

As the data is imbalanced used the weighted F1 score instead of the macro F1 score

* Chance classification accuracy: 33.33% (1/3)

**Method 1: Leave One Participant Out CV**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Property** | **Confusion matrix** | **Average micro F1 score** | **Average weighted F1 score** | **Average classification accuracy** |
| Smoothness | [[ 0 55 2]  [ 11 120 10]  [ 2 49 3]] | 0.49 | 0.39 | 48.81% |
| Thickness | [[ 6 65 1]  [ 23 122 2]  [ 3 30 0]] | 0.51 | 0.44 | 50.79% |
| Warmth | [[ 0 32 7]  [ 0 108 30]  [ 0 67 8]] | 0.46 | 0.38 | 46.03% |
| Flexibility | [[26 31 15]  [21 41 46]  [ 8 42 22]] | 0.35 | 0.34 | 35.32% |
| Softness | [[ 5 36 19]  [15 69 24]  [12 29 43]] | 0.46 | 0.38 | 46.43% |

**Method 2: Leave One Sock Out CV**

Didn’t do this for the combined dataset for the reason stated above