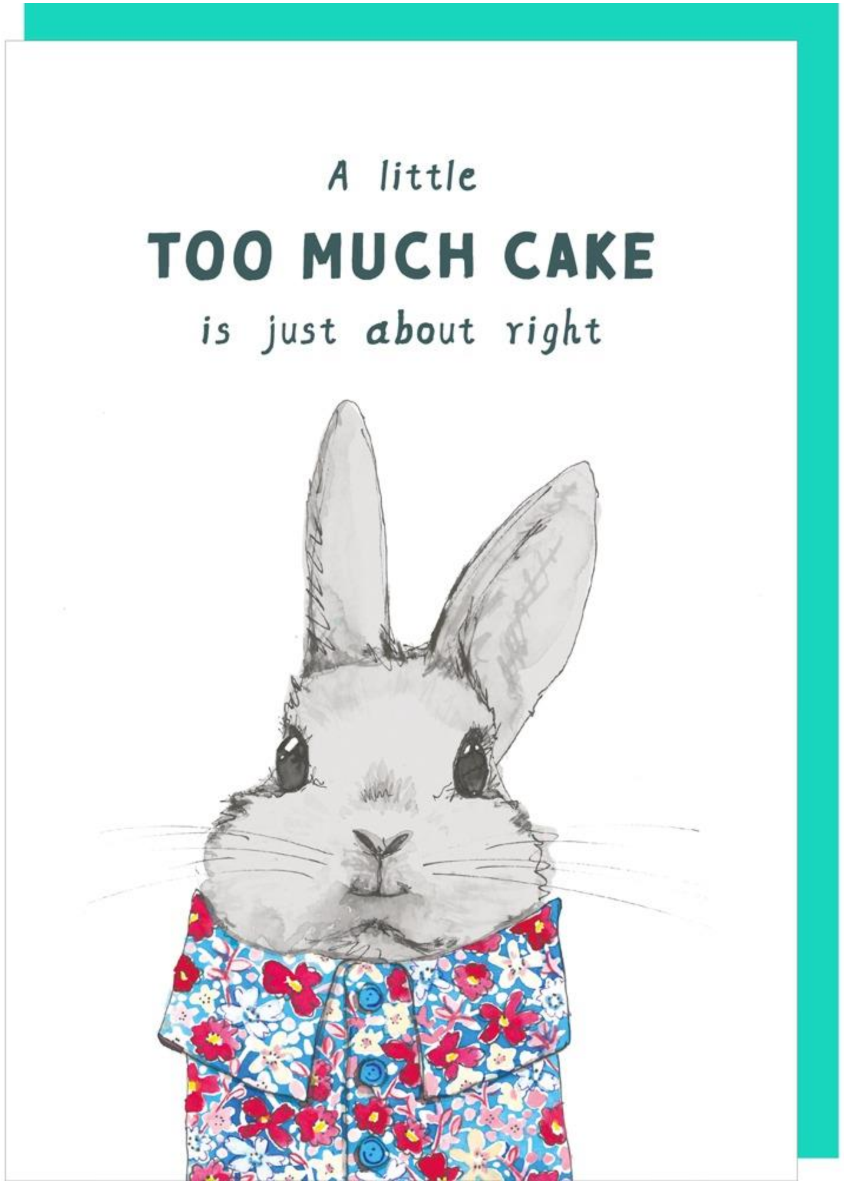
Free Just – About – Right Methodology

Lê Tuấn Phúc - HUST



Let's start with what a JAR is

- (JAR) scaling is widely applied in the food industry for product development
- The JAR scale is a bipolar measurement. In JAR scaling, two semantically opposite anchors and the midpoint is labeled "Just About Right"
- JAR scales are very to be an easy way to determine if an attribute's intensity is at an optimal level

JUST ABOUT RIGHT SCALES

Category:

____ Very much too sweet
____ Too sweet
____ Slightly too sweet
____ Just about right
____ Slightly not sweet enough
____ Not sweet enough
____ Very much not sweet enough

Line Scale:

Not Nearly Sweet enough	Just about Right	Much too Sweet
I	I	I

Directional Scale:

____ Increase it a lot
____ Increase it moderately
____ Increase it slightly
____ Don't change, leave it the same
____ Decrease it slightly
____ Decrease it moderately
____ Decrease it a lot

Limitation

- Predefined, uniform attribute lists are a limitation in sensory evaluation.
- Key attributes may be overlooked if they are not included in the list.
- Consumers may interpret attributes differently, leading to inconsistency.
- Presenting attributes to consumers can bias their evaluations.
- Attributes on the list may be given more importance than warranted, skewing focus.

A flexible method to capture authentic consumer insights without relying on predefined attribute lists

Free JAR Methodoly

I like this product because

Make a nudge for focus on strengths and weaknesses of product

Attribute

????

JAR scales

????

Too much

Not enough

JAR

Freely express

The **Free JAR** process uses **structured** text to capture consumer insights **without predefined attributes**, leveraging the JAR scale to **highlight** product **strengths and weaknesses.**

Free JAR Protocol

The **Free JAR** is a two-step methodology:

STEP 1
Categorization

I DON'T LIKE
P3 P7 P1 P9

I LIKE MODERATELY
P5 P2

I LIKE VERY MUCH
P8 P4 P6

STEP 2
Description with JAR qualifiers

*This product is **not firm enough***

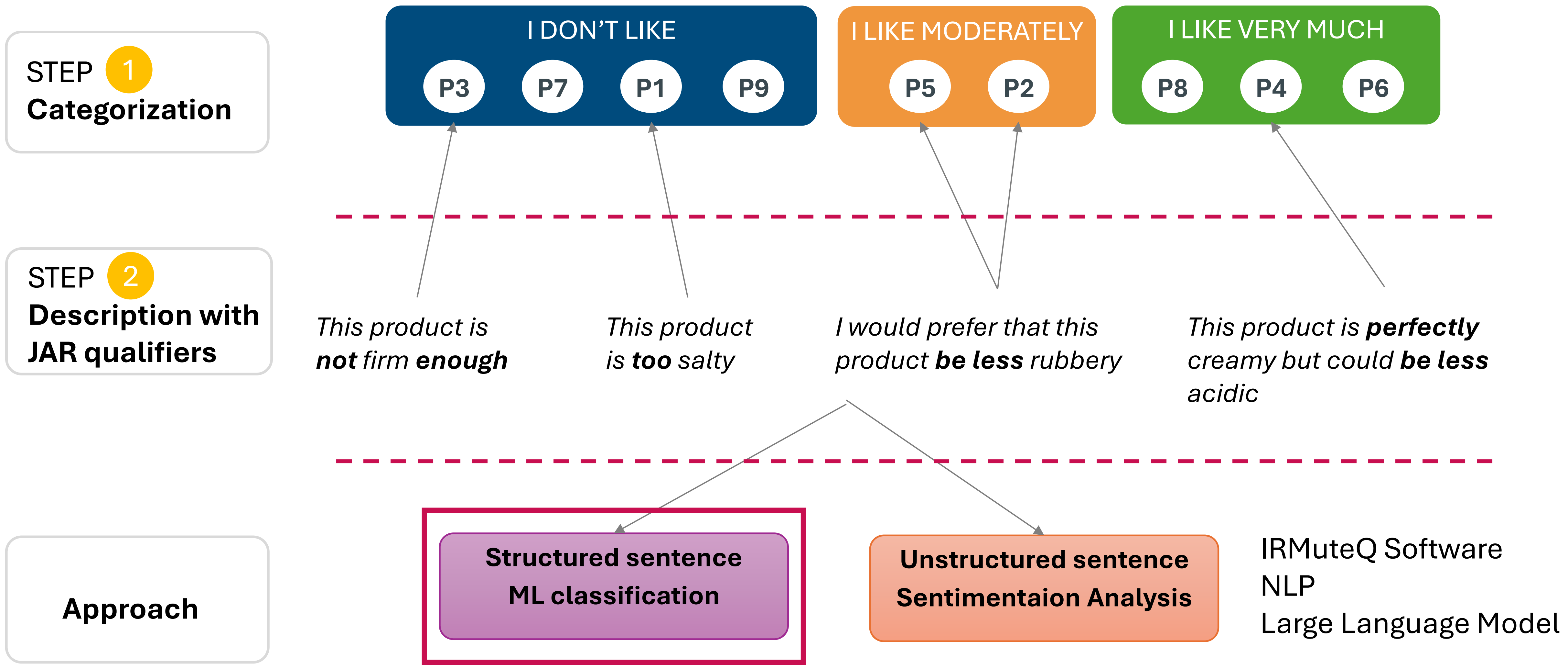
*This product is **too** salty*

*I would prefer that this product **be less** rubbery*

*This product is **perfectly** creamy but could **be less** acidic*

(Luc et al., 2022)

Free JAR Protocol



Free JAR Procedure

DATA STRUCTURE

Judge	Product	Hedonic category	Free JAR comment
1	709	B	ngọt_không_đủ,mùi_thơm_vừa_đủ,màu_quá_đậm
1	495	C	ngọt_không_đủ,mùi_thơm_không_đủ,màu_không_đủ
1	913	C	đắng_quá_nhiều,chát_quá_nhiều,màu_vừa_đủ
1	582	B	mùi_thơm_vừa_đủ,chát_vừa_đủ,đắng_không_đủ,màu_không_đủ
1	136	C	mùi_thơm_vừa_đủ,chát_vừa_đủ,đắng_quá_nhiều,màu_vừa_đủ

Association between a text and a hedonic category

- n be the number of judge
- p the number of products

→ The data resulting from a Free JAR procedure can be stored as a table with 4 columns and p∗n rows

A I like very much
B I like moderately
C I don't like

Free JAR Analysis

Y X

Hedonic	Free JAR comment	chua_vừa_đủ	chát_không_đủ	chát_quá_nhiều	chát_vừa_đủ
1	chát_vừa_đủ,đắng_không_đủ,mùi_thơm_vừa_đủ	0	0	0	1
2	chát_vừa_đủ,mùi_thơm_vừa_đủ	0	0	0	1
1	chát_quá_nhiều,mùi_thơm_vừa_đủ	0	0	1	0
3	chát_không_đủ,mùi_thơm_vừa_đủ	0	1	0	0
1	chát_quá_nhiều,đắng_quá_nhiều,mùi_thơm_vừa_đủ	0	0	1	0
2	ngọt_vừa_đủ,chát_quá_nhiều,mùi_thơm_vừa_đủ,màu_vừa_đủ	0	0	1	0

Label Encoding

3 I like very much

2 I like moderately

1 I don’t like

Supervised classification → $Y = f(X)$

Free JAR Analysis

Implementation of a **Random Forest** classifier:

- Modeling of the **link** between the **attributed** presence and the **hedonic** category
- Define the valency score of a given comment:

Valency score = p (« I like very much ») – p (« I don't like »)

-1 ≤ Valency score ≤ 1

→ **Quantitative** analyses

Luc et al., 2022b

Valency score high
→ The comment is Positive
(I like this product very much)

Valency score low
→ The comment is Negative
(I don't like this product)

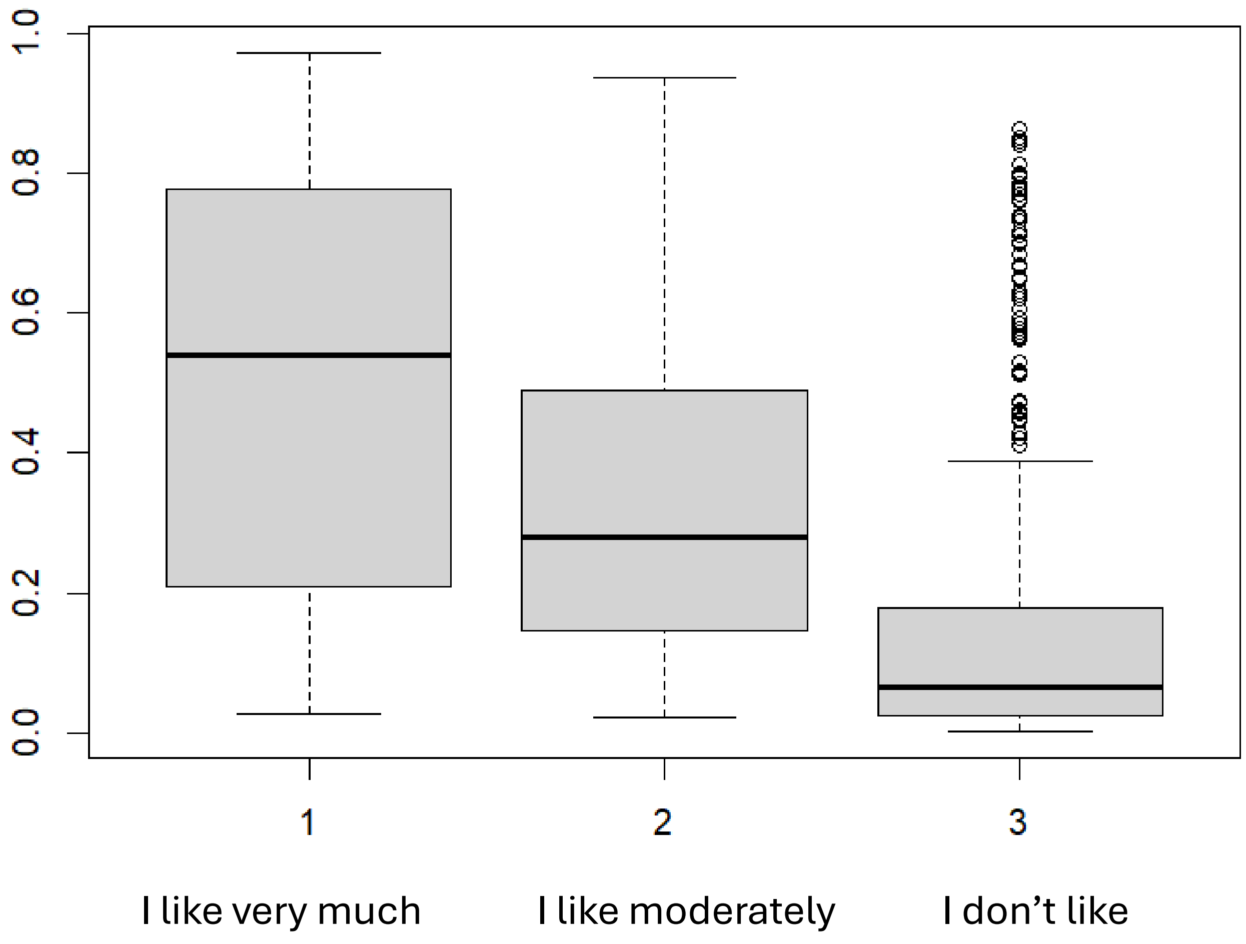
Free JAR Results

```
best_params_

'criterion': 'entropy',
'max_depth': 15,
'max_features': 'log2',
'n_estimators': 100
```

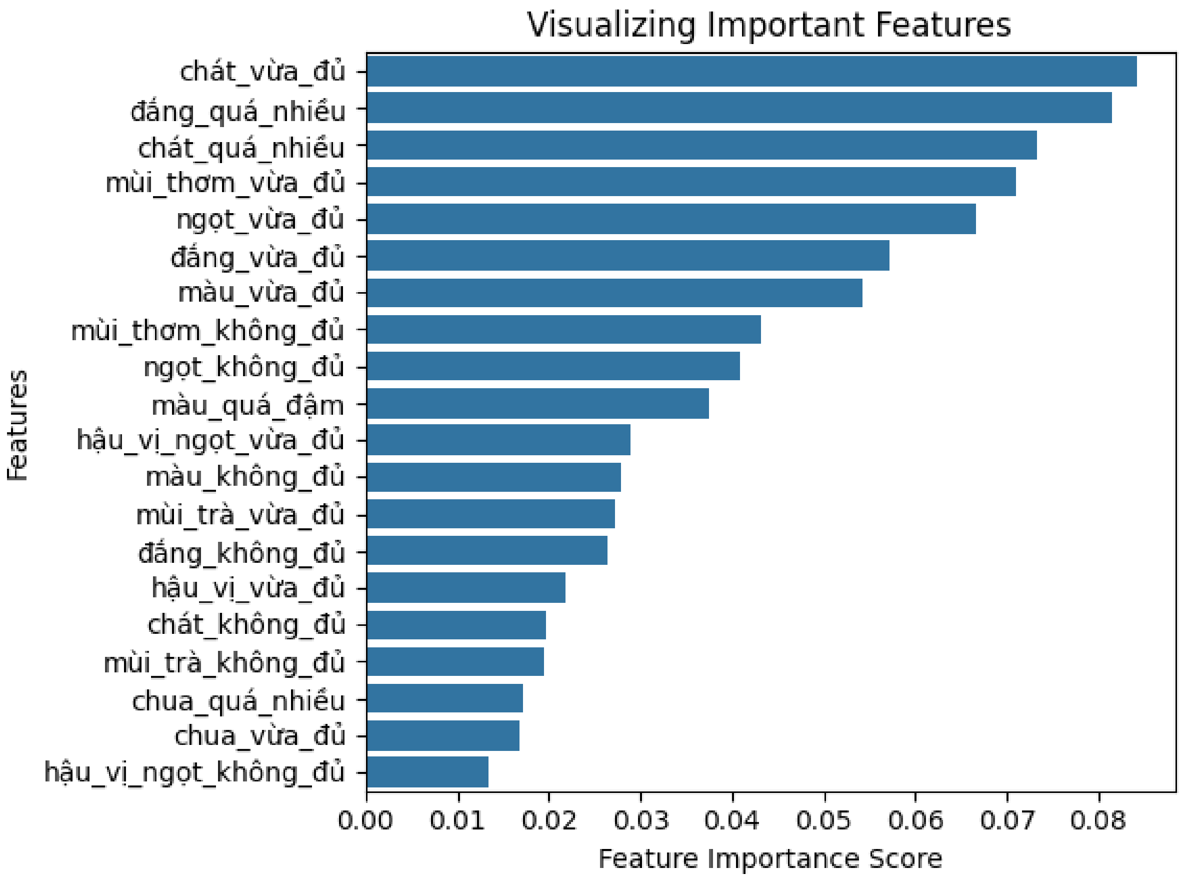
	precision	recall	f1-score	support
1	0.85	0.95	0.9	304
2	0.86	0.79	0.82	213
3	0.94	0.76	0.84	99
accuracy	**0.87**			616
macro avg	0.88	0.83	0.85	616
weighted avg	0.87	0.87	0.86	616

With selected optimal parameters.
The model has an accuracy of 87%



Free JAR Results

20 feature importance variable



Feature importance refers to techniques for determining the degree to which different features, or variables, impact a machine learning model's predictions

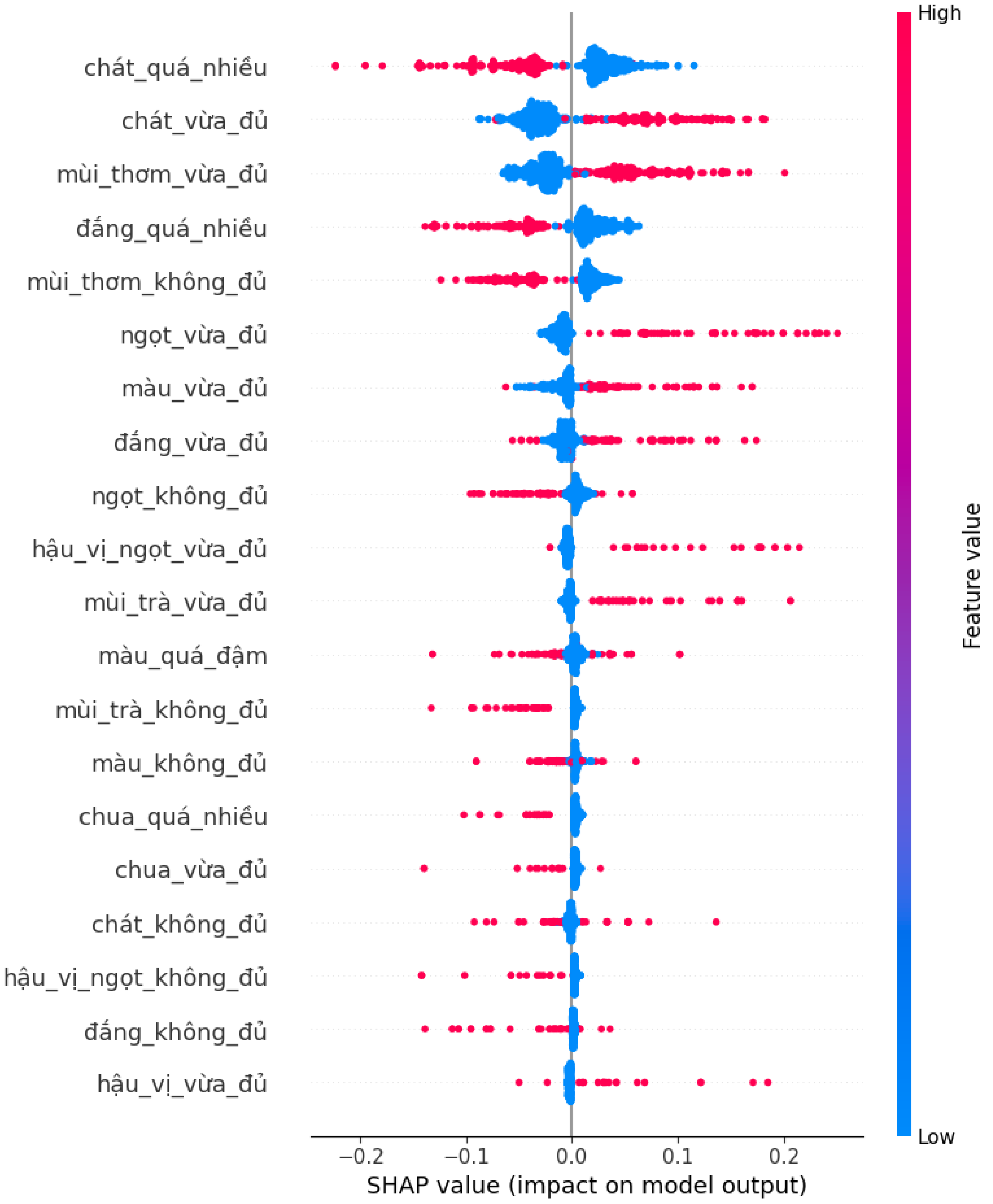
Astringency just right, Astringency too much, Bitter too much, Aroma just right is the main attribute for classify hedonic categories

Free JAR Results

SHAP (SHapley Additive exPlanations) is theoretic approach to explain the output of any machine learning model

Apply SHAP to interpret model

- Astringency too much, Bitter too much, Aroma not enough is the main attribute that negatively affects the occurrence of "I like product".
- Astringent just right and Aroma just right have a positive influence on the occurrence of higher “I like product”.



Free JAR Results

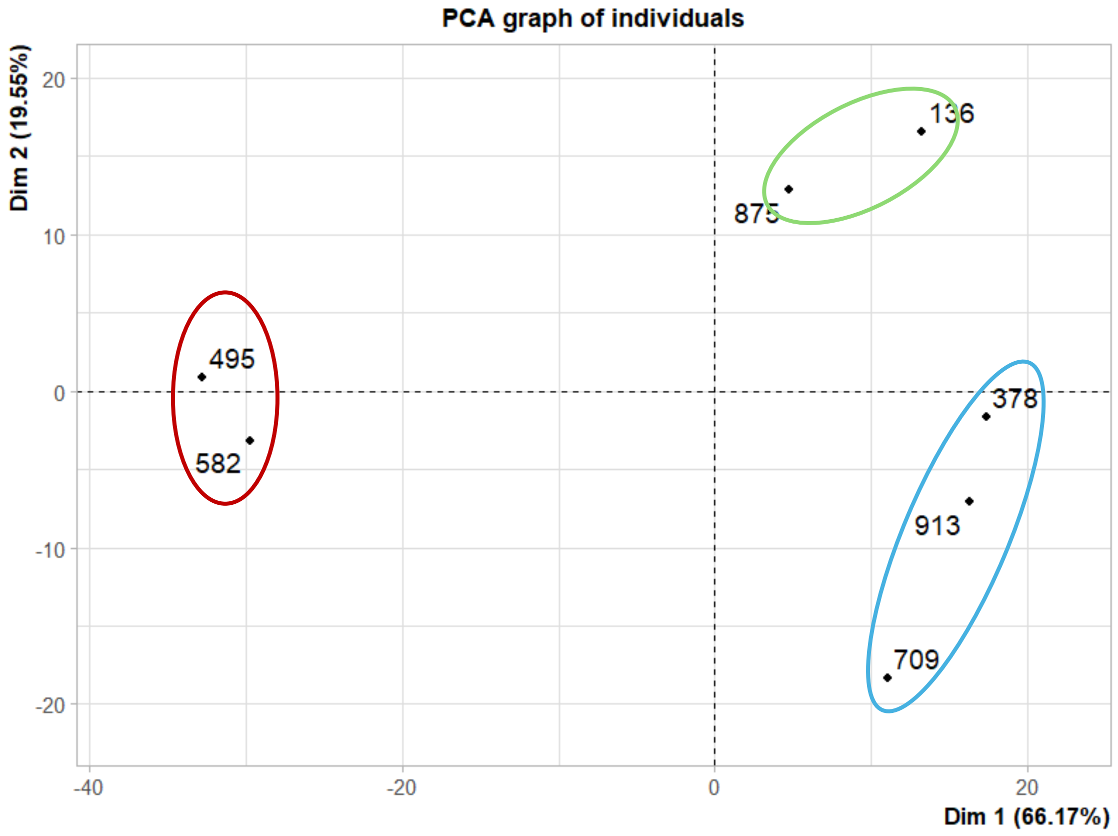
88 Jugde and 7 Product. In the maximum case, that 1 attribute can receive 88x7 = 616

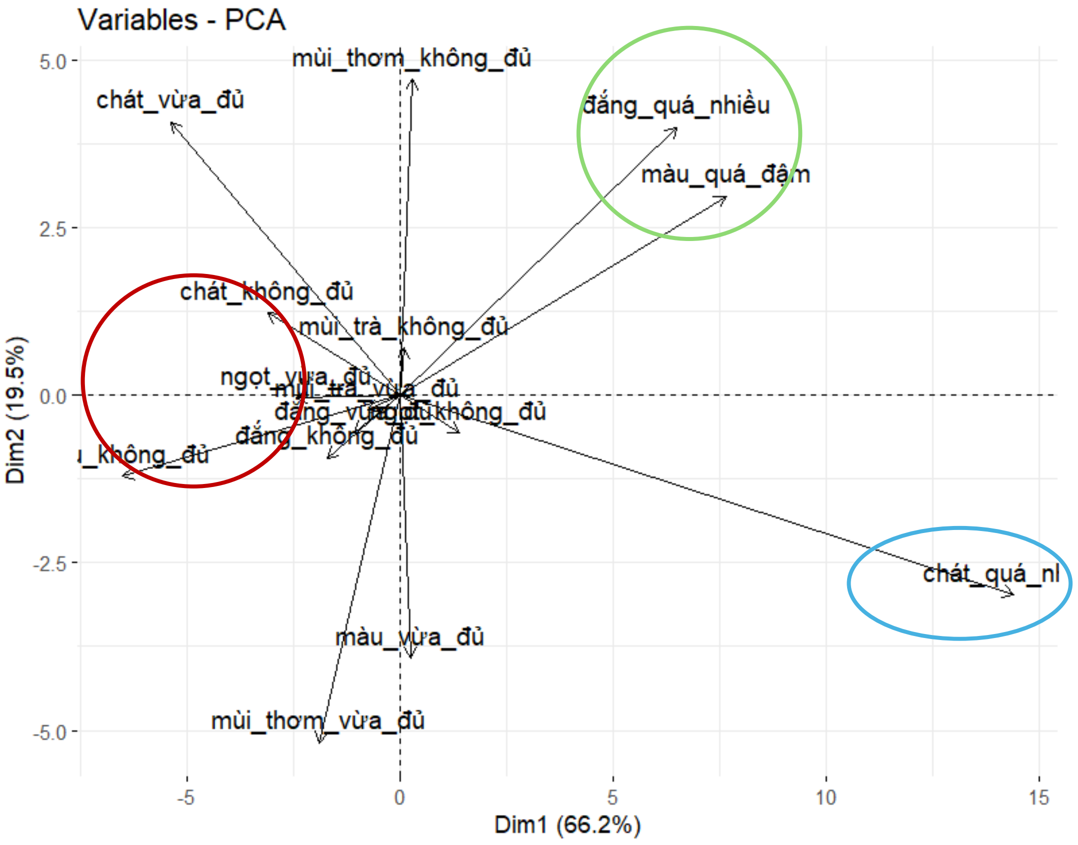
Each attribute was count and calculated its % with maximum case.

Terms that are too few (<5%) can be considered as not contributing much to the product description

	709	495	913	582	136	378	875	total_count	percentage
thanh_vừa_đủ	0	1	1	1	0	2	1	6	0.9740260
thơm_vừa_đủ	2	0	0	0	0	0	0	2	0.3246753
vị_hoa_quả_không_đủ	0	1	0	0	0	0	0	1	0.1623377
vị_không_đủ	0	1	0	0	0	0	0	1	0.1623377
vị_kim_loại_quá_nhiều	0	0	0	0	0	0	1	1	0.1623377
vị_nhạt_không_đủ	1	1	1	1	1	1	0	6	0.9740260
vị_thanh_vừa_đủ	0	0	0	1	1	1	0	3	0.4870130
vị_trà_không_đủ	1	1	0	0	0	0	0	2	0.3246753
vị_trà_quá_nhiều	2	0	0	0	0	0	0	2	0.3246753
vị_trà_vừa_đủ	0	1	0	0	0	0	0	1	0.1623377
đắng_không_đủ	4	6	5	10	2	5	3	35	5.6818182
đắng_quá_nhiều	18	12	23	13	37	31	21	155	25.1623377
đắng_vừa_đủ	11	13	13	15	7	13	15	87	14.1233766
đặc_quá_nhiều	0	0	0	0	1	0	0	1	0.1623377

Free JAR Results





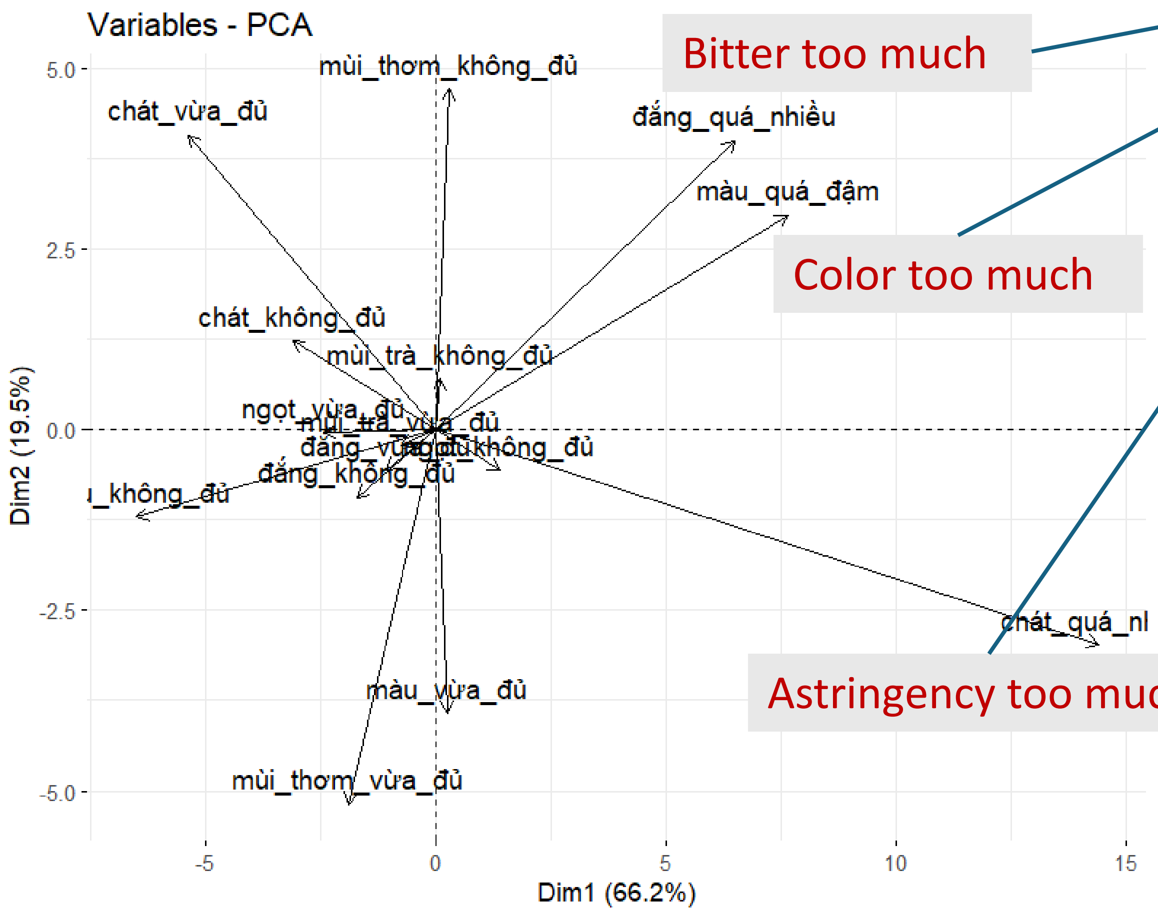
Free JAR Results

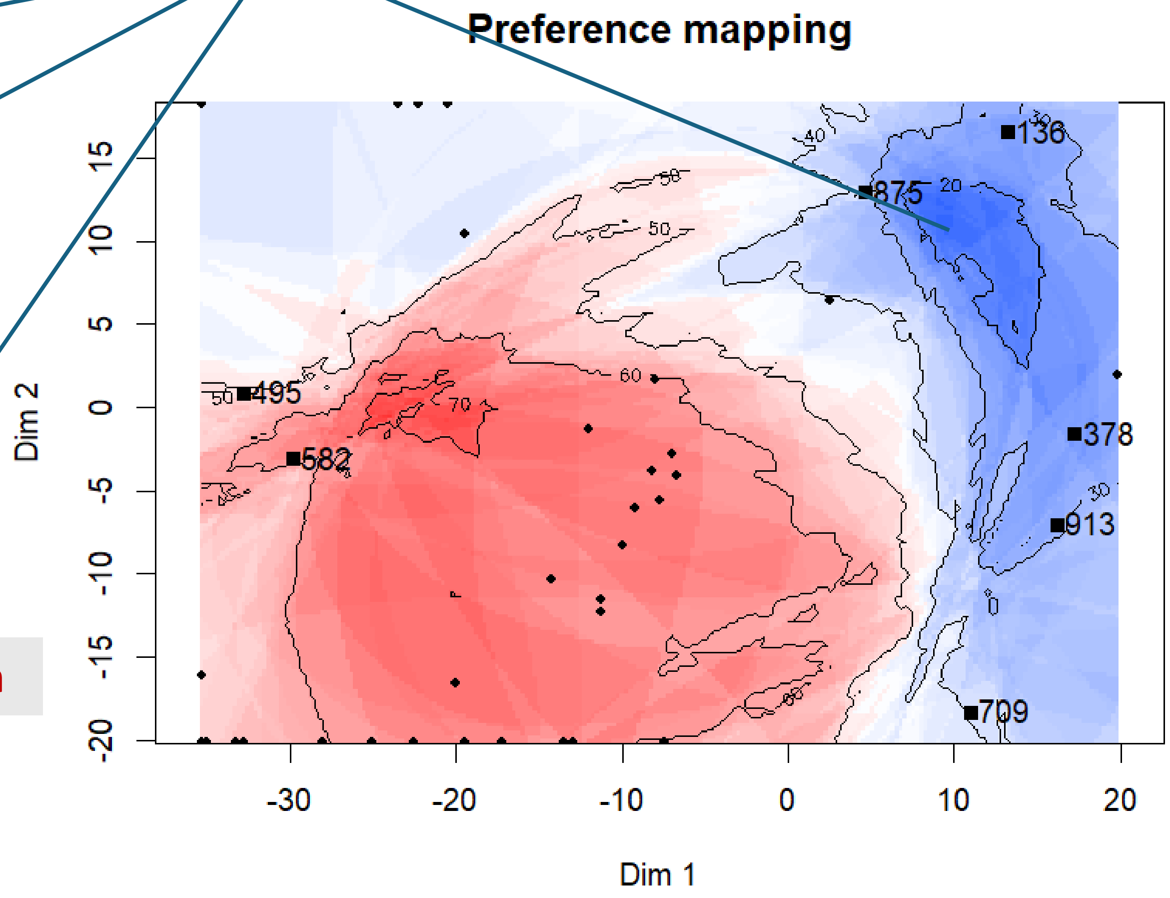
Make product that are not like

Bitter too much

Color too much

Astringency too much





Free JAR Results

Each entry of this table is the valency score associated with the description of a given respondent and product.

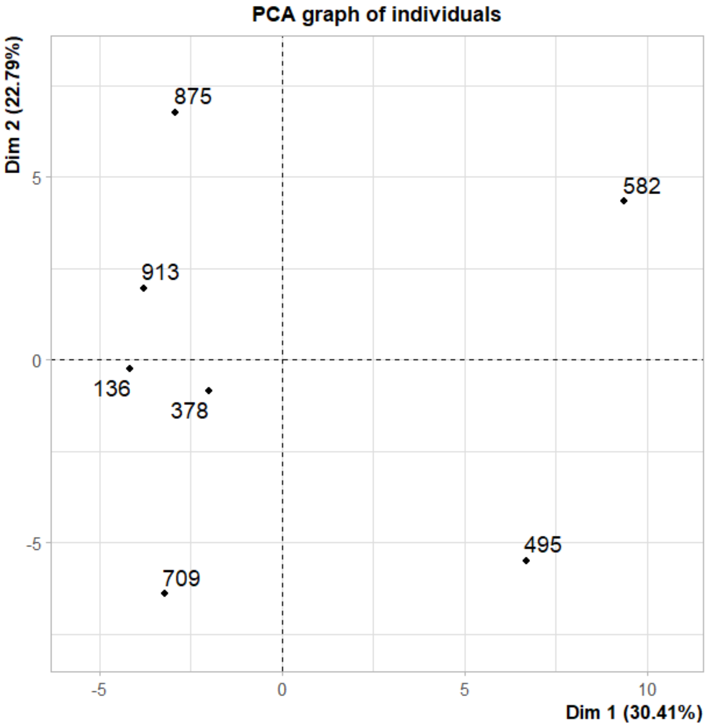
→ Make an **internal preference mapping** by PCA to identifying which products correspond to groups of consumers.

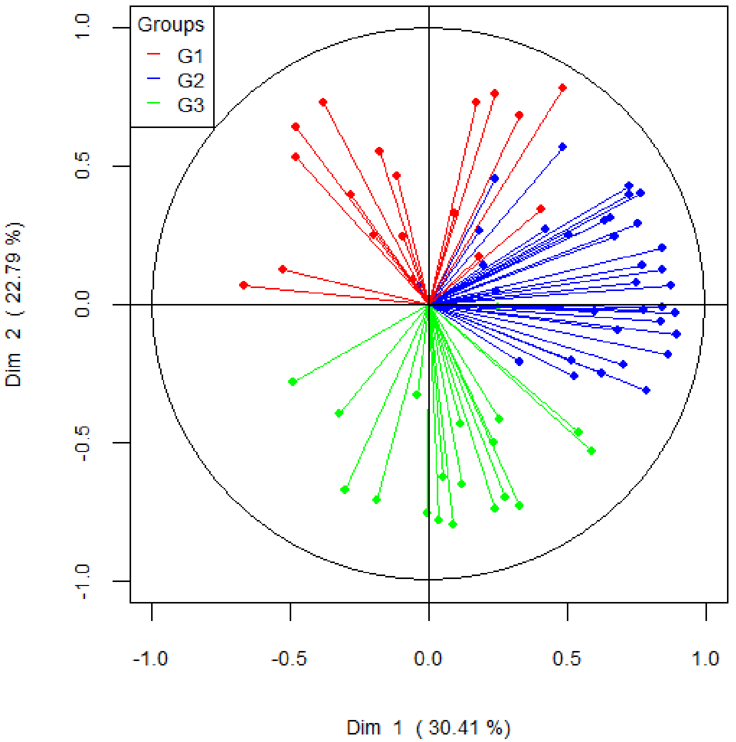
	54	55	56	57	58	59	60
136	-0.66609015	-0.73301933	-0.72880867	-0.5861055056	0.2356101	-0.4405406	-0.1568030
378	-0.08197511	-0.51613823	-0.90255604	-0.7714961745	-0.7531895	-0.6098790	-0.7679865
495	0.38372698	-0.09356122	-0.87081891	-0.0004591183	0.4615966	-0.4125666	-0.9226852
582	-0.68007309	0.80580570	-0.07598381	0.5850284798	0.3295531	-0.6098790	-0.6580867
709	-0.32190269	0.34016227	-0.81504989	-0.6594971668	-0.3682083	-0.4125666	0.6722394
875	-0.92960400	0.80580570	-0.90608316	0.0152609364	-0.7419421	-0.6098790	-0.8504142
913	-0.66361385	-0.49371008	-0.87081891	-0.6651309553	-0.4228199	-0.6098790	0.5820871

Free JAR Results

The two first components explained 53% of the variability in the data

- Group1 is the group of consumer prefer 875 or 913
- Group2 is the group of consumer like 582
- Group3 is the group of consumer who like 709 or 495





Conclustion

Free JAR methodology:

- Through a nudge approach, bring the consumer to provide product improvement keys
- Rich data

The Machine Learning and SHAP approach for FreeJAR is an easy approach and the explanation is good:

→ Get the most out of the data

→ From structured textual data to quantitative data

Interpretability for sensory data: highlight drivers of liking and disliking.