Predicting wine properties with language: A machine learning approach

Ilja Croijmans, Iris Hendrickx, Els Lefever, Asifa Majid & Antal van den Bosch

Background

- Predicting smell and flavor profiles from descriptions or ratings can inform the underlying dimensions of perceived olfactory space.
- Despite recent successes to predict odor space using machine learning [1, 2], there are still challenges since odors, and consequently flavors, are difficult to name and rate consistently for most people [3].
- Wine experts give the impression that they can generate descriptions of the aroma and flavor of wine in wine reviews:

“Dusty apple and white peach aromas are basic and clean. The palate offers bright acids, citrus-driven fruit and whacks of oak-based cinnamon and resin, along with laser-like lemon notes. Mild on the finish, with flavors of vanilla and oak spice.”

The research questions:

1) Can wine experts, with their experience in smelling and describing odors and flavors, review wine distinctively and consistently?

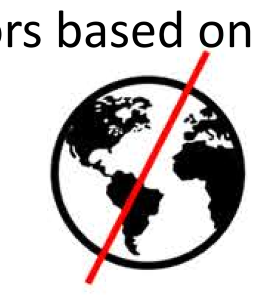
2) What domain-specific words do the experts use in their wine reviews?

Methods

An implementation of Support Vector Machines (i.e., LIBSVM) was used to do classification experiments, to test the prediction that wine experts would talk consistently about wine.

These classification experiments were done on a corpus of 76410 wine reviews written by 13 authors. The paradigm learned wine characteristics using bag-of-word vectors based on reviews from 12 authors based on:

- **color** (red/white/rosé)
- **grape variety** (31 types)
- **origin** (old/new world)



After this, reviews from the 13th author were classified into these classes. Precision (hits / total number of reviews) and recall (hits / number of reviews belonging to a class) scores were calculated, and averaged into F-scores to estimate succes:

$$F = (2 \times \frac{(Precision * Recall)}{(Precision + Recall)}$$

Next, a termhood analaysis was conducted to uncover the terms (n-grams) experts used more frequently in wine reviews compared to an English reference corpus, i.e., what terms are domain-specific?

Results classification experiments

Color

Author	red wine: Number of reviews	red wine: F-score	white wine: Number of reviews	white wine: F-score	rosé wine: Number of reviews	rosé wine: F-score	Average F-score
Author 1	13050	98.1	5364	96.9	302	44.0	97.1
Author 2	5569	92.3	4108	90.7	455	26.1	90.0
Author 3	6882	97.9	2261	94.8	231	33.9	96.0
Author 4	5287	97.2	2098	94.8	157	53.3	95.7
Author 5	4127	97.1	1648	93.9	61	46.6	95.7
Author 6	2831	96.6	2403	97.0	189	41.3	95.5
Author 7	2578	98.3	1011	95.9	39	39.3	97.2
Author 8	1147	99.5	476	99.3	14	78.6	99.3
Author 9	1042	98.3	400	97.8	44	60.6	97.3
Author 10	828	97.1	395	95.7	23	35.7	96.0
Author 11	345	94.8	754	98.5	42	50.0	96.1
Author 12	621	97.5	412	96.8	33	58.3	96.3
Author 13	572	97.0	353	98.5	72	63.0	95.7
Total	44879	97.0	21683	95.2	1662	41.4	96.0

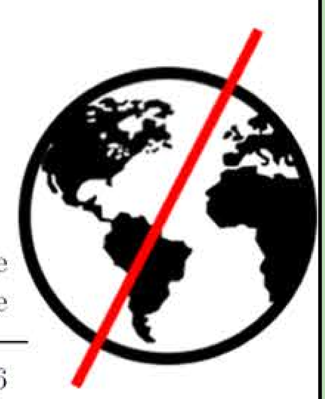
Majority baseline for color: F = 66%

Grape variety

Author	Number of reviews	F-score
Author 1	16268	69.1
Author 2	4796	42.0
Author 3	6457	35.0
Author 4	4547	45.6
Author 5	4661	60.3
Author 6	4015	64.8
Author 7	3010	53.9
Author 8	759	49.1
Author 9	1193	77.0
Author 10	896	56.4
Author 11	986	79.0
Author 12	618	70.6
Author 13	554	61.2
Total	48760	57.4

Majority baseline for grape variety: F = 14%

Origin

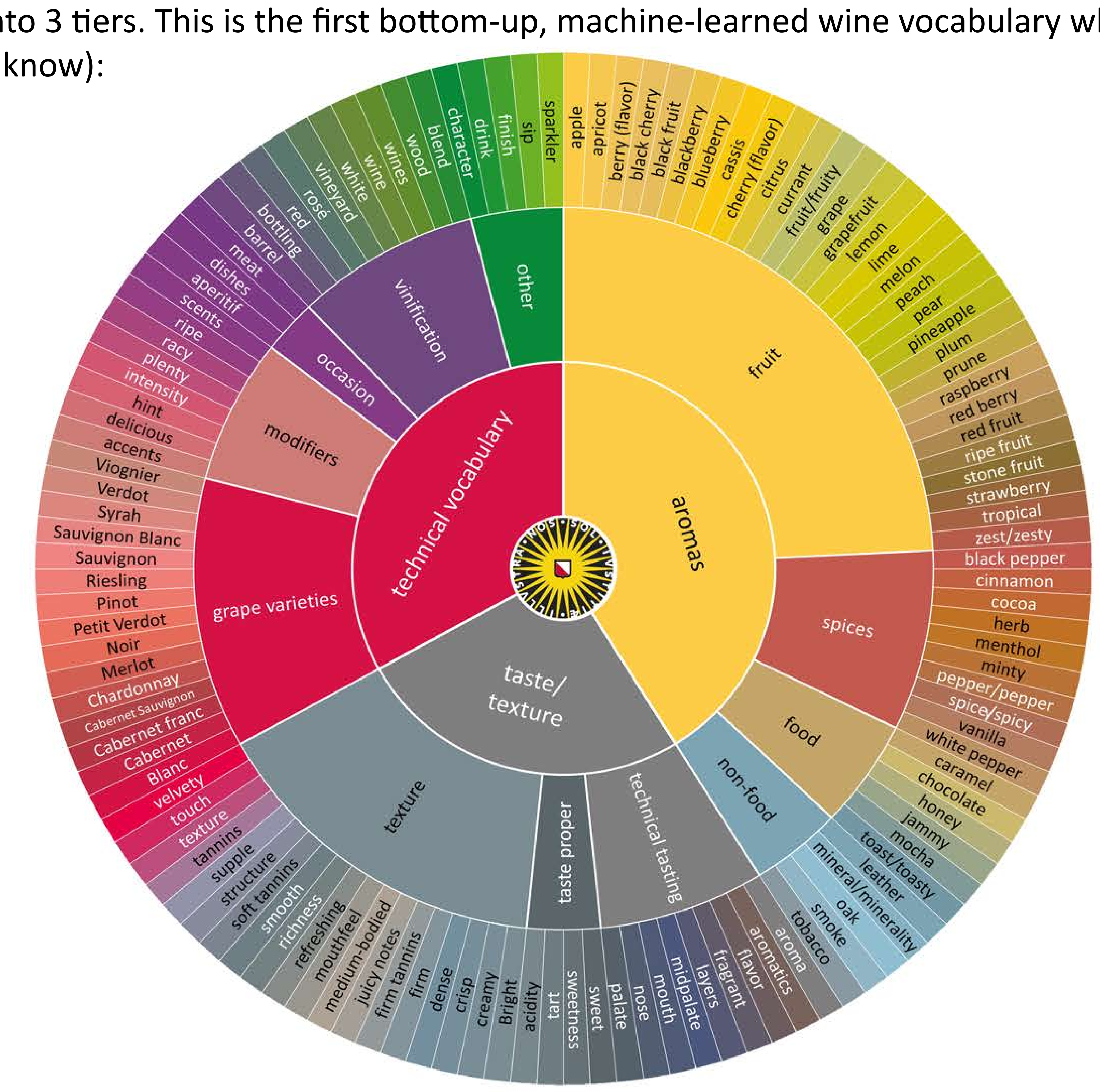


Author	*New world* class: Number of reviews	*New world* class: F-score	*Old world* class: Number of reviews	*Old world* class: F-score	Average F-score
Author 1	19266	86.7	90	2.0	76.6
Author 2	65	1.3	11879	31.0	18.8
Author 3	5428	69.2	4408	44.3	60.3
Author 4	0.0	0.0	8622	77.9	63.8
Author 5	5877	88.8	87	6.7	79.9
Author 6	3075	71.4	2695	47.1	62.9
Author 7	2306	77.8	1456	60.9	71.7
Author 8	0.0	0.0	1770	90.1	81.9
Author 9	1548	89.7	0.0	0.0	81.3
Author 10	1190	85.8	87	21.9	76.0
Author 11	844	72.1	301	41.0	62.1
Author 12	794	73.1	260	43.7	63.6
Author 13	450	64.1	427	62.5	63.3
Total	40843	69.8	32082	48.7	62.0

Majority baseline for origin: F = 56%

Results domain-specific wine vocabulary

- We found 146 terms used more frequently by all 13 authors than in the reference corpus.
- Of these, 89 did not occur in previously established wine vocabulary lists (e.g., Robert Parker’s list of wine vocabulary, Le Nez du Vin aroma terms & Ann Nobel’s wine wheel [5]).
- After minimally pre-processing (e.g., zest and zesty were collapsed), 146 terms were categorized into 3 tiers. This is the first bottom-up, machine-learned wine vocabulary wheel (as far as we know):



Take home messages

- Wine experts describe wine along specific dimensions (color and grape type) in a consistent and distinctive manner.
- Experts used domain-specific language in their wine reviews: this resulted in a new machine-learning informed wine wheel.
- This suggests properties of odor and flavor spaces can be predicted using (expert) language, given the right data.

Acknowledgements

References:

1. Keller, A., Gerkin, R. C., Guan, Y., Dhurandhar, A., Turu, G., Szalai, B., ... & Vens, C. (2017). Predicting human olfactory perception from chemical features of odor molecules. *Science*, 355(6327), 820-826.
2. Lötsch, J., Kringel, D., & Hummel, T. (2018). Machine learning in human olfactory research. *Chemical senses*, 44(1), 11-22.
3. Olofsson, J. K., & Gottfried, J. A. (2015). The muted sense: neurocognitive limitations of olfactory language. *Trends in cognitive sciences*, 19(6), 314-321.
4. Quandt, R. E. (2007). On wine bullshit: Some new software?. *Journal of Wine Economics*, 2(2), 129-135.
5. Noble, A. C., Arnold, R. A., Buechsenstein, J., Leach, E. J., Schmidt, J. O., & Stern, P. M. (1987). Modification of a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture*, 38(2), 143-146.

This poster is based on: Croijmans, I., Hendrickx, I., Lefever, E., Majid, A., & Van den Bosch (submitted). Uncovering the language of wine experts.

Any questions may be directed to: i.m.croijmans@uu.nl

Utrecht University
University of York
Radboud University
NWO Netherlands Organisation for Scientific Research