**GOAL**

Understanding associations between words and nutritional values to come up with new models and to improve practical applications effectiveness.

**CONTRIBUTIONS**

(i) a study of a large-scale online food community in terms of relations between nutritional values and textual descriptions

(ii) the introduction of a new topic model combining text with several outputs (nutrient facts)

(iii) an evaluation of efficacy in discovering recipe topics and predicting nutritional values

**MODEL INCENTIVES**

- nutritional values strongly correlate
- similar words are associated with all nutrient facts

**EVALUATION**

Two practical applications:

1. Recipe topics identification (clustering):
   - our model vs. LDA+LM, 5 topics
   - more consistent weights (e.g., all positive/negative) of nutritional values
   - more focused (e.g., less mixing) topics

2. Prediction of nutritional values from text:
   - our model vs. LDA+LM and LDA+GBT
   - SHARED - representation shared between all outputs
   - SEP - separate model for each of outputs
   - O1 - one of the nutrients is known
   - X1 - all but one nutrient are known
   - i-th output error [%]:

\[
\text{sMAPE}_i = \frac{2}{|\text{doc}|} \sum_{d \in \text{doc}} \left| \frac{f_{i,d} - \hat{f}_{i,d}}{|f_{i,d}|} \right|
\]

**APPLICATION 1: RECIPE TOPICS IDENTIFICATION**

Comparison of topics found by standard LDA (bottom) and our model (top).

**APPLICATION 2: NUTRITIONAL VALUES PREDICTION**

Nutrients prediction avg performance.

Prediction improvement when additional outputs values are included.

**REFERENCES**


