

Técnicas de análisis de datos en la industria alimentaria

Héctor Olivera Pazó, Rebeca P. Díaz Redondo, Ana Fernández Vilas

Information & Computing Lab., AtlantTIC Research Center
Escuela de Ingeniería de Telecomunicación. Universidade de Vigo
{h.olivera, rebeca, avilas}@det.uvigo.es

Oportunidad

midía día

<http://www.midiadia.com/>



La oportunidad de negocio

midíaadía

<http://www.midiadia.com/>



Colaboración y objetivos

midíadía

<http://www.midiadia.com/>



<http://iclab.det.uvigo.es/>

- * **Automatizar la categorización de productos en la taxonomía**
- * **Refinar la taxonomía de *midiadia***
- * Creación automática de diccionarios de ingredientes
- * ...
- * **Recomendador avanzado (objetivo último)**

Categorización automática

- * 202 variedades + 8795 productos
- * Para cada producto: listado de ingredientes + listado de características agregadas por *mediadia*
- * 1ª aproximación: Categorización basada únicamente en el listado de ingredientes

Dificultades

- * Univocidad de los ingredientes
 - * p. ej. “*aroma natural*” vs. “*aromas naturales*”
 - * Asignación semiautomática de identificadores unívocos
 - * Otros no tan claros: p. ej. “*vitamina C*” & “*ácido ascórbico*”

Alimentació Bebidas | Zumos | Multifrutas / Otros sabores n Pago | Zumo de n: azúcar, aromas naturales, zumo concentrado de zanahoria (5%), vitamina c, provitamina a, zumo concentrado de limón (5%), agua, vitamina e, zumo concentrado de naranja (20%)

Alimentació Bebidas | Zumos | Multifrutas / Otros sabores n Granini | Zumo multi Agua, zumo de multifrutas concentrado a partir de (uva(18%), piña (14%), naranja (8%), , puré de melocotón (17%), pulpa de naranja (5%), acidulante: ácido cítrico, aroma natural,

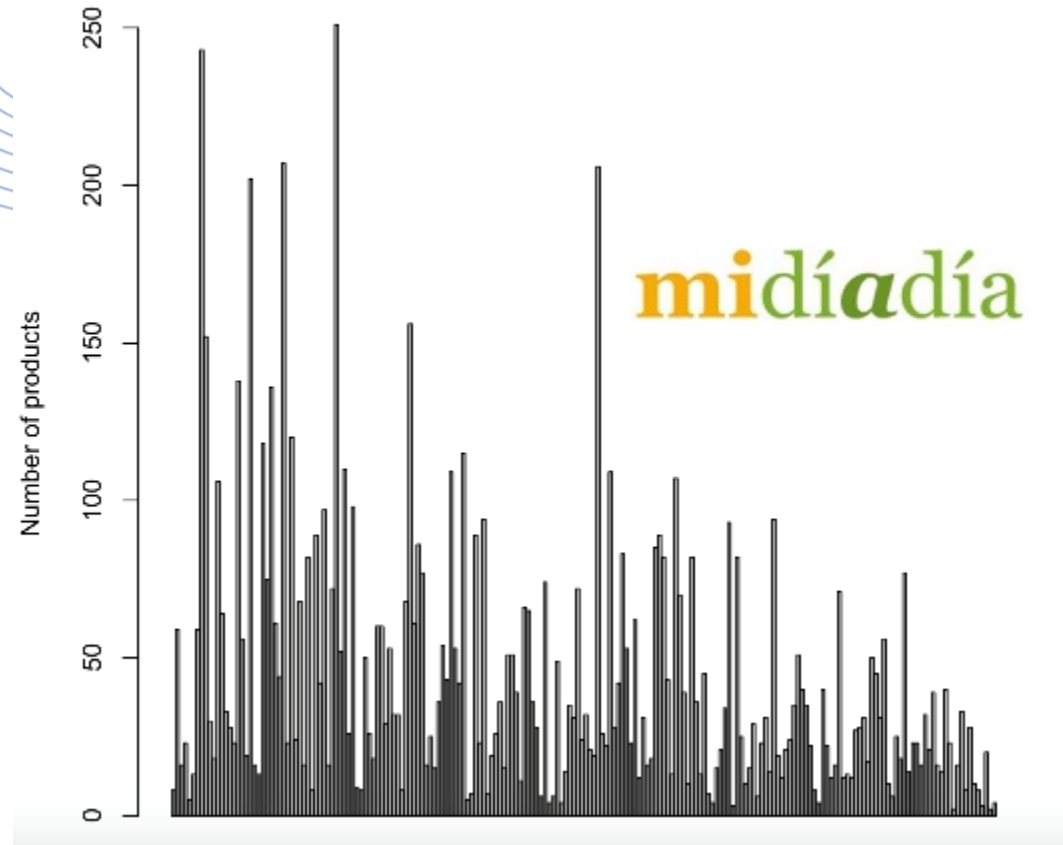
- * Productos con ingredientes similares para variedades diferentes:
 - * p.ej. “*Snacks varios*” & ”*Cocktails – Mezclas*”

Alimentació Aperitivos | Frutos secc Cocktails - Mezclas | Sekitos | Cocktail mi cacahuete, maíz, haba, harina de trigo, aceite de girasol, grasa vegetal, sal, azúcar, almidón de maíz, harina de arroz, fécula de patata,

Alimentació Aperitivos | Snacks | Snacks varios (patata, maíz, Alipende | Cocktail sn Harina de trigo, harina de patata, fécula de patata, aceite de girasol, harina de arroz, harina de maíz, sémola de maíz, harina de centeno, sal, azúcar,

Categorización automática

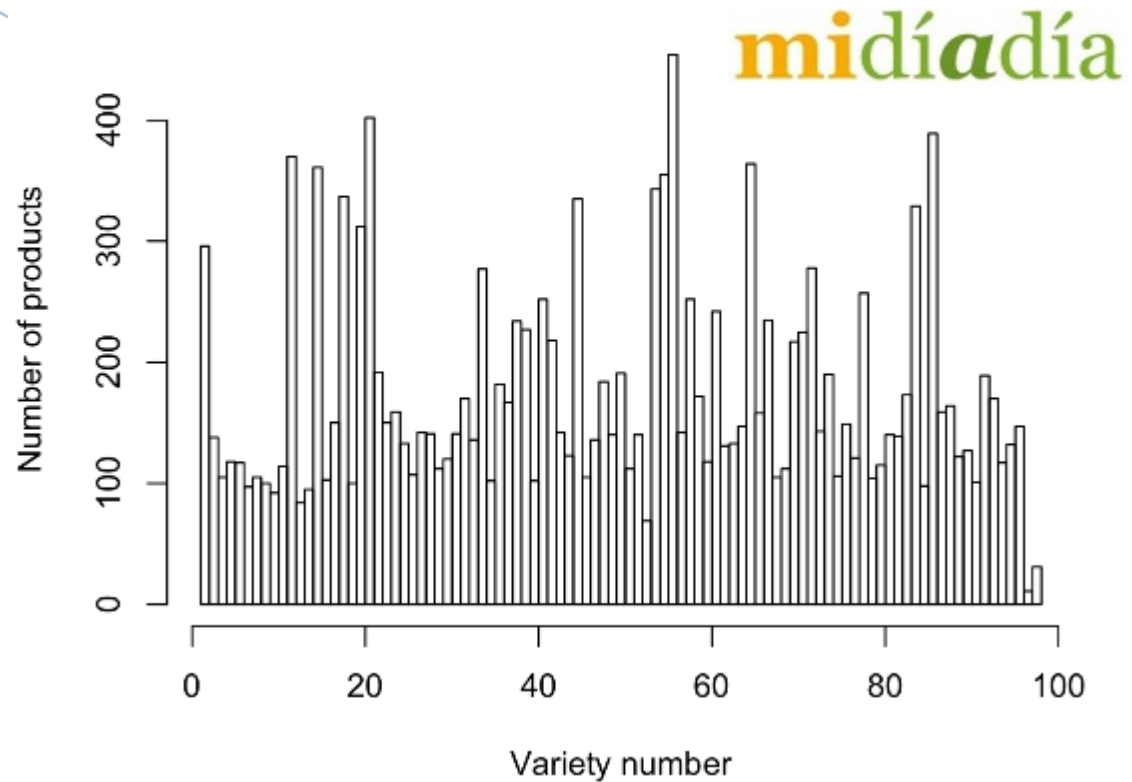
- Basada únicamente en el listado de ingredientes
- Definición de métrica que permitan comparar productos y, por tanto, productos y variedades



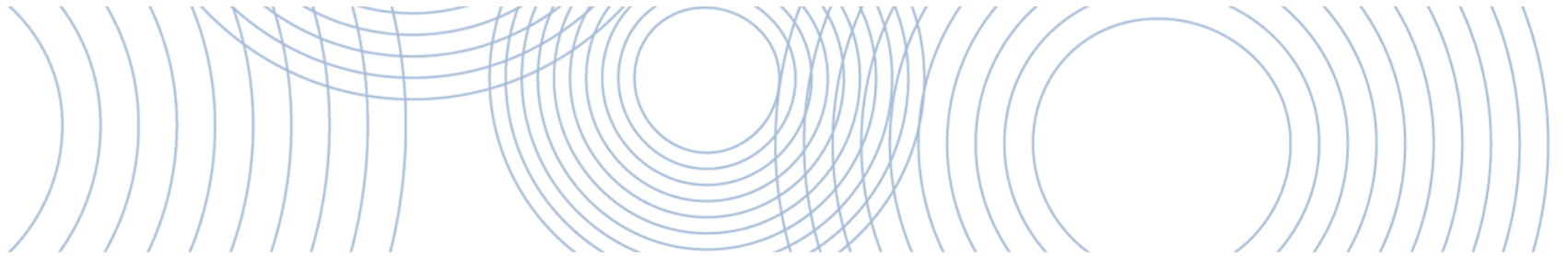
Threshold	> 150 products	> 100 products	> 50 products
Precision	78,95%	73,47%	64,49%
Number of varieties	7	17	62
Training values	1282	2335	5265
Testing values	135	250	616

Actualmente...

Products distribution



- Dataset mayor
- 2ª aproximación: Utilizar características adicionales
- Refinar la taxonomía



<http://iclab.det.uvigo.es/>
rebeca@det.uvigo.es

Técnicas de análisis de datos en la industria alimentaria

Héctor Olivera Pazó, Rebeca P. Díaz Redondo, Ana Fernández Vilas

Information & Computing Lab., AtlantTIC Research Center
Escuela de Ingeniería de Telecomunicación. Universidade de Vigo
{h.olivera, rebeca, avilas}@det.uvigo.es