

# Tabla de contenido

<b>1. Introducción</b>	<b>1</b>
1.1. NotCo: motivación y problemas . . . . .	1
1.1.1. The Not Company: <i>raison d'être</i> . . . . .	1
1.1.2. La Combinación Perfecta de Ingredientes Vegetales . . . . .	1
1.2. Giuseppe: IA aplicada a alimentos . . . . .	2
1.2.1. Giuseppe: Un explorador Inteligente . . . . .	2
1.2.2. El Proceso de Selección de fórmulas . . . . .	3
1.2.3. Situación Actual . . . . .	3
1.2.4. Problemas de la Situación Actual . . . . .	3
1.3. Mejorando el proceso de Giuseppe . . . . .	4
1.3.1. Objetivos Especificos del Trabajo . . . . .	4
1.3.2. Descripción de la Solución Desarrollada . . . . .	4
1.3.3. Resultados Obtenidos . . . . .	5
<b>2. Marco Teórico</b>	<b>6</b>
2.1. <i>Curse of Dimensionality</i> . . . . .	6
2.2. Reducción de Dimensionalidad . . . . .	7
2.3. UMAP: Uniform Manifold Approximation and Projection . . . . .	7
2.3.1. Aproximación del <i>Manifold</i> . . . . .	8
2.3.2. Construcción de la representación Topológica . . . . .	10
2.3.3. Optimización de la Representación en baja dimensión . . . . .	12
2.4. Análisis de Clusters . . . . .	13
2.4.1. Taxonomía de Algoritmos de Clustering . . . . .	14
2.4.2. HDBSCAN : Hierarchical Density-based Clustering for Applications with Noise . . . . .	15
2.4.3. Tendencia de Puntos a formar Clusters . . . . .	16
2.4.4. Estadístico de Hopkins . . . . .	17
2.5. Métricas de Validación de un <i>Clustering</i> . . . . .	17
2.5.1. Índice de Jaccard . . . . .	17
2.5.2. DBCV: Density Based Clustering Validation . . . . .	18
2.6. Evaluación de un <i>Clustering</i> con expertos . . . . .	19
<b>3. Problema Abordado</b>	<b>21</b>
3.1. Descripción del Proceso de Selección de fórmulas . . . . .	21
3.2. Problemas Identificados . . . . .	21
3.3. Ventajas de un nuevo Proceso de Selección . . . . .	22
3.4. Requisitos . . . . .	23
<b>4. Solución Desarrollada</b>	<b>24</b>
4.1. Metodología de Trabajo . . . . .	24
4.2. Selección de espacio de características . . . . .	24
4.3. Reducción de Dimensión . . . . .	25
4.4. Tendencia de fórmulas a formar <i>Clusters</i> . . . . .	26
4.5. Comprobación Visual de existencia de <i>Clusters</i> . . . . .	27
4.6. Análisis de <i>Clusters</i> obtenidos . . . . .	28

4.7. Validación de <i>Clusters</i> con Chefs . . . . .	31
4.7.1. Algoritmo de Construcción de Conjunto de Fórmulas para Validación . . . . .	32
4.7.2. Resultados de Validación . . . . .	34
4.8. Algoritmo de Selección de Fórmulas . . . . .	36
4.8.1. Comparación cuantitativa de procesos de selección de fórmulas . . . . .	37
4.9. Arquitectura de la Solución . . . . .	37
<b>5. Conclusiones</b>	<b>40</b>
<b>6. Trabajo Futuro</b>	<b>41</b>
<b>Bibliografía</b>	<b>43</b>
<b>Apendice</b>	<b>44</b>
<b>A. Arquitectura de la Solución</b>	<b>44</b>
A.1. <i>Parser</i> de Fórmulas . . . . .	44
A.2. <i>Embedder</i> de Fórmulas . . . . .	45
A.3. <i>Clusterer</i> de Fórmulas . . . . .	46
A.4. Seleccionador de Fórmulas . . . . .	46
A.5. Herramienta de Línea de Comando . . . . .	47