

Tabla de Contenido

1. Introducción	1
1.1. Contexto y problema	1
1.2. Objetivos	3
1.2.1. Objetivo General	3
1.2.2. Objetivos Específicos	3
2. Estado del Arte	4
2.1. Servicios similares	4
2.2. Web Scraping	5
2.2.1. Prácticas comunes	5
2.2.2. Herramientas de web scraping	6
2.2.3. Legalidad	7
2.3. Metodologías ágiles	8
2.4. Regresión	11
2.4.1. Árboles de regresión	11
2.4.2. Boosting	12
2.5. Backend	13
2.5.1. Data Pipelines	13
2.5.2. Data Lake	13
2.5.3. Data Warehouse	14
2.5.4. Desarrollo de API	15

2.5.5. Bases de datos	15
3. Solución Propuesta	17
3.1. Scrapers	18
3.2. Limpieza y procesamiento	20
3.3. Procesamiento	20
3.4. Entrenamiento y predicción	21
3.5. API	22
4. Implementación	23
4.1. Scrapers	23
4.1.1. Comportamiento general	24
4.1.2. Implementación	26
4.1.3. Datos generados	28
4.1.4. Validación	29
4.1.5. Despliegue	32
4.1.6. Consideraciones adicionales	32
4.2. Data Pipelines	33
4.2.1. Limpieza y formateo de datos	33
4.2.2. Datos históricos	34
4.2.3. Datos adicionales	34
4.3. Sugerencia de precio	36
4.3.1. Idea general	36
4.3.2. Preprocesamiento de datos	37
4.3.3. Predicción	39
4.3.4. Mantención de los modelos	40
4.4. Backend	41
4.4.1. API	41

4.4.2. Base de datos	42
4.5. Interfaz gráfica	44
5. Resultados y Validación	47
5.1. Modelos de predicción	47
5.1.1. Selección de modelos	47
5.1.2. Resultados	49
5.2. Validación con usuarios	50
6. Conclusión y Trabajo Futuro	58
Bibliografía	62
ANEXO A. Figuras	63
A.1. Scrapers	63
A.2. Figuras H3	63