

# Formulación de un modelo de negocio basado en ciencia de datos: Preciario

Sebastian Corzo Penha  
Director: León Darío Parra Bernal  
Maestría en Ciencia de Datos — 2026

Perteneces  
**AQUÍ.**

SOMOS LA U. DEL  
**EMPRENDIMIENTO  
SOSTENIBLE.**

Acreditaciones:



# El Problema

## Asimetría de información en precios

- Hogares destinan 35% de ingresos a canasta básica.
- Inflación 5,2% con dispersión de precios entre retailers.
- No existen comparadores multi-tienda automatizados en Colombia.
- Rappi y similares resuelven logística, NO transparencia de precios.



# Objetivos

## General y Específicos

General: Formular un modelo de negocio para la creación de una empresa basada en una aplicación que permita comparar los precios de diferentes supermercados, facilitando la toma de decisiones de compra en los consumidores.

1. Analizar mercado y tendencias de consumo
2. Diseñar propuesta de valor diferenciada
3. Definir fuentes de ingresos y estructura de costos
4. Evaluar aspectos legales y tecnológicos del scraping
5. Proponer plan de implementación y escalabilidad

# Análisis del Sector

## RetailTech en Colombia

- Penetración de Internet 75,7%.
- E-commerce retail +26,7% (CCCE, 2025).
- **PESTEL:** Sin competidores directos en comparación con unificación SKUs.
- **Porter:** NLP como ventaja tecnológica.
- **Océano Azul:** transparencia + optimización canasta.



SOMOS LA U. DEL  
**EMPENDIMIENTO**  
SOSTENIBLE.

# Validación de Mercado

77 encuestas + 5 entrevistas

| Indicador           | Resultado       | Implicación     |
|---------------------|-----------------|-----------------|
| Comparación precios | 80,5% prioridad | Función core    |
| Alertas ofertas     | 61% demanda     | Proactivo       |
| Disposición pago    | 67% dispuesto   | Freemium        |
| Segmento target     | 26-40 años E2-4 | Adultos activos |

| Indicador | Resultado        | Implicación       |
|-----------|------------------|-------------------|
| TAM       | \$27,4B COP      | Sector retail COL |
| SAM       | \$3,6B COP       |                   |
| SOM Año 1 | 2.900 usuarios   | 0,01%             |
| SOM Año 5 | 104.322 usuarios | 0,26%             |

# Modelo de Negocio Freemium

B2C + B2B

## 1 Propuesta de Valor

1. Comparador inteligente que unifica SKUs entre retailers - NLP
2. Optimizador canasta completa.
3. Detector ofertas Prophet/SARIMA.

## 2 3 Líneas de Ingreso

1. Gratuito: \$5.000/año (publicidad)
2. Premium: \$10.000/mes
3. Corporativo: \$50.000/mes

B2B = 5,8% usuarios, 64% ingresos

# Arquitectura Técnica + NLP

Pipeline validado con prototipo

## 1 Pipeline 4 Etapas

- 1) Scraping 6h (5 retailers)
- 2) Procesamiento + NLP unificación SKUs
- 3) API REST: comparador + knapsack + predictor
- 4) React Native + React.js + Dashboard B2B

## 2 Prototipo NLP: F1=75,9%

Corpus: 1.717 productos, 3 retailers  
30 Pares etiquetados manualmente (15+15)  
Precision 78,6% | Recall 73,3%  
Accuracy 76,7% | Threshold 75%

FuzzyWuzzy - Fuzzy String Matching  
Tokenización con distancia de Levenshtein  
Grid search thresholds 70-90%

'Naranja Tangelo 1 und' de Éxito VS 'Naranja Tangelo Sel Car 1 und' de Carulla = similitud del 84%

'Papa 1000 gr' versus 'Papa Criolla 1 kg' = 74%

# NLP

## Pipeline validado con prototipo

1

### Resultados

- Precision: 78,6%
- Recall: 73.3%
- F1-Score del 75.9%

2

### Interpretación

- 8/10 clasificados equivalentes OK
- 7/10 equivalentes OK
- Balance adecuado

'Naranja Tangelo 1 und' de Éxito VS 'Naranja Tangelo Sel Car 1 und' de Carulla  
'Papa 1000 gr' versus 'Papa Criolla 1 kg'

Embeddings semánticos Word2Vec Wang et al. y Köpcke et al.: 85 al 92%

# Indicadores Financieros

- Inversión total: \$403M COP (90% deuda, 10% aporte)
  - VPN: \$540M (1,38x inversión) — Proyecto crea valor
  - TIR: 40,91% (2,7x costo de oportunidad del 15%)
  - Recuperación: 2,0 años
  - Punto de equilibrio: 730 usuarios - con distribución entre los diferentes tipos de usuario
- 
- Año 5: Ingresos \$6.483M | Utilidad neta \$2.503M
  - Retorno 621% sobre inversión inicial
  - Exit estratégico: 48-80x (múltiplos 3-5x revenue)
  - Comparables LATAM: Rappi 10,9x, Cornershop \$3B (Uber)

# Hoja de Ruta



# Conclusiones

Hallazgos en los 5 objetivos específicos :

1. Mercado: inflación + internet 75,7% + sin competidores = oportunidad validada
2. Propuesta: NLP + comparador + optimizador = barrera de entrada tecnológica
3. Modelo: Freemium B2C+B2B viable. B2B genera 64% ingresos con 5,8% usuarios
4. Legal: scraping permitido (Ley 1480 + Ley 1581)
5. Financiero: VPN \$540M, TIR 40,91%, recuperación 2 años

**Preciario es un proyecto viable y rentable.**

Genera impacto social al democratizar la información de precios.



acreditada  
en alta calidad  
res. N° 023654 del Mineducación  
10/12/21, vigencia 10/12/27

# [www.universidadean.edu.co](http://www.universidadean.edu.co)

Universidad Ean: SNIES 2812 Personería Jurídica Res. 2898 del  
Minjusticia - 16/05/69 - Vigilada Mineducación.

Ean Legacy: Carrera 11 No. 78-47 - Fundadores: Calle 79 # 11 - 45 - Av.  
Chile: Calle 71 # 9 - 84 Bogotá D.C., Cundinamarca, Colombia, Suramérica.

Línea gratuita nacional: +57-01-8000-931000  
Centro de contacto en Bogotá: +(57-601) 593 6161 - +(57-601) 593 6464  
E-mail: [informacion@universidadean.edu.co](mailto:informacion@universidadean.edu.co)

© 2024 Todos los derechos reservados - Universidad Ean

