**Skincare Recommender**

**1. Introducción**

Este proyecto consiste en un sistema de recomendación de productos de skincare. Como se trata de un MVP, el análisis se ha limitado exclusivamente a productos que apuntan al cuidado facial. La misma mecánica es extrapolable a productos de cuidado corporal y capilar.

* Recomienda productos según el tipo de piel (normal, grasa, seca, mixta), condiciones específicas (rosácea, dermatitis, psoriasis) y preocupaciones (acné, arrugas, calmar, colágeno). Tanto las condiciones específicas como las preocupaciones son a modo de ejemplificar que el modelo funciona, pero podrían considerarse muchas opciones más (como piel madura, piel sensible, antioxidantes, pigmentación, etcétera).
* Compara productos europeos y coreanos, mostrando alternativas similares entre ambas regiones.
* Analiza ingredientes clave para garantizar que los productos recomendados sean adecuados para el usuario.

**2. Datos de Partida**

Se utilizaron dos conjuntos de datos obtenidos mediante web scraping:

* 🇪🇺 Productos Europeos (PromoFarma)

Columnas: nombre, marca, precio (€), ingredientes, url\_producto, imagen\_url.

* 🇰🇷 Productos Coreanos (Jolse)

Columnas: name, brand, price (KRW), ingredients, product\_url, image\_url.

Ambos datasets fueron procesados para que presenten la misma información en las mismas unidades y bajo columnas de igual nombre. Los datos que no aplicaran, fueron directamente eliminados en cada caso.

**3. Modelos y Técnicas Utilizadas**

Es un sistema de recomendación basado en contenido. Consta de distintas etapas:

* Filtrado por reglas de ingredientes: se definieron listas de ingredientes clave para cada tipo de piel, condición y preocupación. Esto es para simplemente reducir el universo de productos a comparar entre sí más adelante. Por ejemplo, en principio, nunca una piel con dermatitis debería utilizar productos que contengan ácido salicílico. De modo que, para este caso, productos que contengan dicha sustancia directamente no se tienen en cuenta para recomendar ni para comparar luego.
* Modelo de similitud semántica (Sentence Transformers): se usó el modelo all-MiniLM-L6-v2 para calcular similitud entre la consulta del usuario (por ejemplo, piel grasa con acné), los ingredientes y nombres de productos. La métrica es Cosine similarity, para ordenar productos por relevancia. Para cada producto seleccionado, el sistema busca alternativas en la otra región basándose en ingredientes comunes (60% del peso) y similitud semántica (40% del peso).

**4. Pantallazos de la aplicación**

La interfaz principal es Streamlit:

<https://skinseoulutions.streamlit.app/>

Interfaz de usuario gráfica, Aplicación

El contenido generado por IA puede ser incorrecto.

**5. Conclusiones y mejoras futuras**

En líneas generales, al aplicación funciona bien. Algunas posibles mejoras son lógicamente incluir más datos, de más sitios webs, con categorías (limpiador, sérum, crema, protector solar, etcétera), incluir filtro de precio inferior y superior, cuantificar el número de ingredientes de cada fórmula para la comparación (se espera que los productos coreanos tengan menos ingredientes), utilizar un modelo de machine learning más avanzado (fine-tuning del modelo de embeddings para skincare) o recurrir a la API de OpenAI (es de pago).