

# Sistema híbrido de recomendación. Un sistema multi-objetivo no convencional.

Matías Bavera  
Universidad Católica de Asunción  
Asunción, Paraguay  
matiasbavera@gmail.com

Benjamín Barán  
Universidad Nacional de Asunción y  
Universidad Comunera  
Asunción, Paraguay  
bbaran@pol.una.py

Uri Yael  
Universidad Católica de Asunción  
Asunción, Paraguay  
uyael@cba.com.py

**Resumen**—Actualmente, los *Sistemas de Recomendación* juegan un papel fundamental en el comercio; por ello, este trabajo implementa un innovador sistema de recomendación para una empresa paraguaya. Este sistema de recomendación se vuelve híbrido al combinar varios métodos de recomendación. El sistema busca recomendaciones personalizadas para sus usuarios, que además de ser precisas, sean diversas y novedosas a la vez. En este trabajo, el problema de brindar recomendaciones personalizadas, así como el problema de la diversidad máxima están modelados como un problema de optimización de múltiples objetivos. La propuesta es maximizar 5 objetivos: precisión, novedad y 3 métricas de diversidad halladas con diferentes definiciones de distancia. Para realizar las recomendaciones se proponen cuatro modelos híbridos de recomendación. Se diseñó un algoritmo de toma de decisiones personalizadas para seleccionar una solución del conjunto Pareto óptimo. Resultados experimentales, utilizando la base de datos de una empresa comercial paraguaya, demuestran que las soluciones brindadas por las opciones propuestas son preferidas a las soluciones tradicionales, vistas desde un punto de vista puramente multi-objetivo. Este sistema de recomendación fue implementado en una empresa comercial del Paraguay.

## I. INTRODUCCIÓN

Los sistemas de recomendación son sistemas capaces de generar recomendaciones significativas de una colección de artículos a un conjunto de usuarios potencialmente interesados [1]. Para implementar estos sistemas, es fundamental tener información sobre los clientes, ya sean sus preferencias o su historial de compras, entre otros datos relevantes. La importancia de esta información radica en el conocimiento que puede ser adquirido sobre el comportamiento del cliente. Dicho conocimiento de los clientes contribuye a aumentar la proporción de aciertos, entre las ofertas realizadas por la empresa y las aprovechadas por un cliente. Los métodos utilizados para crear o seleccionar una recomendación pueden ser divididos en dos grupos: (1) métodos tradicionales, como serían las técnicas basadas en contenido, filtrado colaborativo, conocimiento específico de dominio, entre otras técnicas basadas en un historial de los datos [2]; y (2) métodos no tradicionales o no convencionales que no se basan necesariamente en un historial de los datos disponibles para recomendar productos. Ejemplos de estos métodos no convencionales podrían ser: (i) métodos basados en opiniones de expertos de la empresa; (ii) métodos impulsados por eventos (como cambios climáticos, feriados, etc.); y (iii) métodos impulsados por el departamento

comercial (por ejemplo para incentivar el consumo de un producto con sobre-stock).

Es factible conformar un sistema híbrido de recomendación mediante la complementación de dos o más técnicas de recomendación. El fin de la construcción de estos sistemas híbridos es el de brindar sugerencias cuya utilidad general sea mayor a la obtenida por una sola técnica de recomendación [3].

Existen distintos criterios de evaluación para analizar un sistema de recomendación. El criterio más popular y relevante es la *precisión* [4]. Sin embargo, estudios más recientes han demostrado que la precisión por sí sola no garantiza una alta satisfacción del usuario [5]. Propiedades adicionales tales como diversidad, novedad, interpretabilidad, ocurrencia, entre otras, también cumplen roles relevantes en el aumento de la satisfacción de usuarios [6]. En consecuencia, este trabajo enfocará sus criterios de evaluación en las siguientes métricas: precisión, diversidad y novedad. Se propone entonces maximizar simultáneamente los valores de cada métrica, en un contexto de múltiples objetivos, del inglés *Multi Objective Problem* (MOP) [7]. El problema así planteado será resuelto con un *Multi Objective Evolutionary Algorithm* (MOEA) especialmente diseñado para el efecto.

Como elemento innovador, este trabajo propone utilizar por primera vez 3 métricas distintas de diversidad. Esto se realiza con el objetivo de resolver el problema de la dispersión máxima o diversidad máxima del inglés *Maximum Diversity Problem* (MDP) [8] al escoger la lista de sugerencias propuesta por un recomendador.

## II. ESTADO DEL ARTE

### II-A. *Sistemas de Recomendación*

Los sistemas de recomendación, del inglés *Recommender Systems* (RS), han sido reconocidos como un tema relevante por distintas comunidades de investigación [9]. Existen múltiples implementaciones de RS para distintos tipos de enfoques, útiles en el mundo real. Para este proyecto, resultan relevantes trabajos anteriores que traten sobre implementaciones de RS realizadas para negocios *retail*. Geuens et al. [10] proponen un *framework* para configurar las recomendaciones basadas en filtrado colaborativo. Las recomendaciones son derivadas de un historial de compras. Sano et al. [11] proponen un sistema

de recomendación para una tienda *retail* teniendo en cuenta la escasez de datos. Las evaluaciones de la métrica de precisión son realizadas con las categorías de los productos y no con los productos en sí.

Los métodos utilizados para realizar recomendaciones en los trabajos anteriores, en su mayoría, hacen uso solamente de un tipo de recomendador y solo alternativamente de un sistema híbrido. Tradicionalmente, las estrategias híbridas de recomendación consisten en la combinación de dos familias diferentes de algoritmos, el filtrado basado en el contenido y el colaborativo [12], [13]. Tanto en [12] como en [14], se encuentra una recopilación de datos muy interesante sobre RS que incluyen: (1) métodos para realizar las recomendaciones; (2) algoritmos de predicción y; (3) varias medidas utilizadas para evaluar los RS utilizados en la actualidad. Incluso, en [15] y [16] se propusieron enfoques para RS multi-objetivo o multi-criterio. Nikzad-Khasmakhi et al. [17] presentan el estado del arte de los sistemas expertos de recomendación, sus ventajas y desventajas, considerando también las diferentes técnicas empleadas en la actualidad.

Según Burke [18], los sistemas híbridos de recomendación pueden ser divididos en las siguientes categorías: (1) ponderado (*weighted*); (2) conmutado (*switching*); (3) mixto (*mixed*); (4) por combinación de características (*feature combination*); (5) por aumento de funciones (*feature augmentation*); (6) cascada (*cascade*); y (7) meta-nivel (*meta-level*). Según Ribeiro et al. [7], de estos siete tipos de técnicas híbridas, tres no son apropiadas para combinar muchos algoritmos de recomendación. Por ejemplo, el tipo de combinación de características es un híbrido que ni siquiera combina los algoritmos de recomendación. El tipo cascada, es apropiado para pocos recomendadores, debido a la desvinculación existente entre un algoritmo y el siguiente. Finalmente, el tipo de meta nivel requiere secuencias muy específicas de algoritmos. Estas secuencias son difíciles de implementar para dos algoritmos, y mucho más para varios algoritmos.

Se han propuesto varios enfoques de recomendación con el objetivo específico de proporcionar al usuario las sugerencias más precisas posibles [19]. Algunos ejemplos de trabajos realizados sobre sistemas híbridos de recomendación que buscan aumentar la precisión son los siguientes: Kaššák et al. [20] y Kardan y Ebrahimi [21]. En general, como métrica de evaluación de las recomendaciones, solamente se toma en cuenta la precisión. Estudios más recientes han demostrado que la precisión por sí sola no garantiza una alta satisfacción del usuario [5]. Como resultado de esta constatación, la atención también se ha dedicado a otras propiedades asociadas con elementos sugeridos, como la diversidad y la novedad [22]. Kunaver y Pozrl [23] presentan el estado del arte de distintas métricas de diversidad utilizadas en los RS. En Zhou et al. [24], los autores demostraron que la elección entre precisión y diversidad no es necesariamente un dilema. En dicho artículo se concluyó que es posible lograr simultáneamente ganancias en precisión y diversidad al proponer enfoques híbridos.

Generalmente, la novedad recomienda artículos de cola larga, es decir, aquellos artículos a los que pocos usuarios

han accedido. Vargas y Castells [25] presentan estrategias híbridas que combinan el método de recomendación de filtrado colaborativo con técnicas de dispersión en grafos, para mejorar la novedad. Celma y Herrera [26] en lugar de evaluar la novedad en términos de los ítems de cola larga, adoptan un enfoque alternativo. Este enfoque evalúa la novedad siguiendo los caminos que van desde las recomendaciones hasta la cola larga, usando enlaces de similitud. El trabajo de Celma y Herrera [26] es el primero (hasta donde se conoce) que propone un método híbrido multi-objetivo en término de tres métricas: precisión, diversidad y novedad.

Ribeiro et al. [9] proponen (i) la combinación de listas clasificadas (con un *ranking*) producidas por algoritmos existentes, en una sola lista. A este método llaman *ranking* eficiente de Pareto, y (ii) la combinación ponderada de algoritmos existentes que resulta en un recomendador híbrido. A este método llaman hibridación Pareto-eficiente. Esta hibridación Pareto-eficiente es encontrada mediante un EA (*Evolutionary Algorithm*) conocido como *Strength Pareto Evolutionary Algorithm, Version 2* (SPEA-II). Wang et al. [6] proponen un algoritmo evolutivo multi-objetivo (MOEA), basado en descomposición, del inglés *Multi Objective Evolutionary Algorithm Based on Decomposition* (MOEA/D), para sugerir listas de recomendaciones diversificadas para cada usuario. Notablemente, la información del contenido de un ítem se toma en cuenta cuando se diseña la función objetivo de diversidad. Esta propuesta no utiliza un enfoque híbrido de recomendación. Oliveira et al. [27] proponen utilizar optimización multi-objetivo basada en aspectos de diversificación, del inglés *Multi Objective Aspect Diversification* (MOAD), para generar listas de recomendaciones. La innovación del mencionado trabajo consiste en la admisión de una especificación explícita de los aspectos que pueden ser diversificados. El MOEA utilizado fue el *Non-dominated Sorting Genetic Algorithm, Version 2* (NSGA-II) y no se presenta un enfoque híbrido de recomendación. Los objetivos a optimizar son precisión y diversidad. Algunos de los trabajos mencionados anteriormente, como [6], [9], [27] y [16], se centraron en optimizar los siguientes objetivos: (1) precisión; (2) novedad y; (3) diversidad.

A diferencia de los trabajos anteriores, en este proyecto se utilizará un método que hace uso de un MOEA para combinar múltiples métodos de recomendación. El método fue inicialmente propuesto por [18] y tuvo una implementación previa en [9]. Por el contrario, proyectos como [6], [27] y [28] no han implementado un sistema híbrido. Este trabajo se centrará en optimizar los siguientes objetivos: (1) precisión; (2) novedad y (3) diversidad, así como lo hicieron [6], [9], [27] y [16].

Como elemento innovador, en este trabajo se implementa por primera vez un método que no fue probado hasta ahora en el entorno de los RS (hasta donde se conoce). Este método es la dispersión máxima en un entorno multi-objetivo [8], [29], dado que hasta hoy todas las implementaciones de diversidad en los RS fueron en un entorno mono-objetivo (solo considerando una única métrica de diversidad).

## II-B. Dispersión máxima

El problema de la diversidad máxima o dispersión máxima, del inglés *Maximum Diversity problem* (MDP), presenta un gran número de aplicaciones prácticas que surgen de la búsqueda de los  $k$  elementos mas disímiles de un conjunto de datos. Este tipo de problema utiliza modelos de diversidad y definiciones de distancia o disimilitud como forma de medir que tan diferentes son los elementos de un conjunto dado.

En un contexto multi-objetivo, Vera et al. [29] presentan por primera vez enfoques multi-objetivos para el problema de diversidad máxima. Se proponen dos modelos de optimización diferentes: (i) el modelo de diversidad máxima multi-objetivo, donde el número  $k$  de elementos a ser seleccionados se define a priori, y (ii) el modelo de diversidad media máxima multi-objetivo, donde la cantidad de elementos  $k$  a seleccionar también es una variable de decisión. Para resolver los problemas formulados, se presenta un MOEA.

Machuca et al. [8] proponen la utilización de un algoritmo evolutivo multi-objetivo (el reconocido NSGA-II), presentando varios casos de prueba que demuestran la eficiencia del algoritmo propuesto en comparación con la búsqueda exhaustiva. Como diversos expertos pueden preferir diferentes definiciones de distancia para problemas específicos, el mencionado trabajo propone resolver el MDP con un enfoque multi-objetivo, considerando la utilización simultánea de múltiples definiciones de distancia o disimilitud.

## III. SISTEMAS DE RECOMENDACIÓN

Un sistema de recomendación es un conjunto de herramientas y técnicas de software que proporcionan sugerencias para que los artículos recomendados sean de utilidad para un usuario. Las sugerencias se refieren a varios procesos diferentes, tales como qué artículos comprar, qué música escuchar, o qué noticias existen en línea para leerlas [30].

Según Isinkaye et al. [12] existen tres tipos de retroalimentación en el proceso de recomendación: (a) implícita; (b) explícita; e (c) híbrida. En este trabajo se utiliza la *retroalimentación implícita*. En este tipo de retroalimentación el sistema infiere automáticamente las preferencias del usuario mediante el seguimiento de sus diferentes acciones, pero sin su intervención explícita. En el caso de este trabajo, los *feedbacks* implícitos se pueden conseguir analizando el comportamiento de los clientes después de su compra.

En este trabajo se combinan dos tipo de recomendadores: los tradicionales y los no tradicionales. Como recomendador tradicional se utilizará el *sistema de recomendación basado en filtrado colaborativo*. El filtrado colaborativo es una técnica de predicción para contenido, independiente del dominio. La técnica de filtrado colaborativo funciona construyendo una base de datos (matriz de elementos de usuario, del inglés *user-item matrix*), de los ítems preferidos (o comprados) por los clientes. Luego, el algoritmo utilizado por el filtrado colaborativo compara a los usuarios con intereses y preferencias de modo a calcular similitudes entre sus perfiles para formular recomendaciones [12].

Como recomendadores no tradicionales se proponen los recomendadores por *empresa* y por *experto*. El recomendador por *empresa* se define como el que sugiere productos en base a intereses propios de la empresa, como vender artículos específicos; por ejemplo, para incentivar el consumo de un producto con sobre-stock. El recomendador por *experto* se define como el que sugiere productos en base al conocimiento y experiencias de expertos en ventas de la empresa. Al agregar los recomendadores no tradicionales al esquema, se podrá contar con los siguientes beneficios: (1) servirá para enfrentar el problema del *cold start* [2], (2) aumentará el nivel de diversidad en las recomendaciones propuestas a un usuario, y (3) sería de utilidad para guiar al módulo tomador de decisiones.

En este trabajo se utilizarán las *calificaciones únicas*, del inglés *unary ratings*. Este sistema de puntuación permite al usuario  $u$  especificar una preferencia positiva para un ítem  $j$  [2]. Es decir, la interacción de un cliente con un artículo se representa con el valor “1”, mientras que la falta de interacción, por practicidad, se representa con “0”.

**III-1. Objetivos a optimizar:** Con la motivación de tener buenas listas de recomendación para los usuarios, los objetivos a optimizar son la precisión, novedad y diversidad. A continuación se presenta una breve explicación de cada uno de los objetivos: (1) *precisión*, indica la forma en que los elementos recomendados satisfacen las necesidades del usuario; (2) *novedad*, está intrínsecamente vinculada a la noción de descubrimiento y representa lo novedoso para el usuario que son los elementos recomendados; y (3) *diversidad*, muestra cuán diferentes son entre sí los elementos recomendados de una lista. Los algoritmos para realizar recomendaciones como el *ALS*, del inglés *Alternating Least Squares* y *KNN*, del inglés *K Nearest Neighbors* se encargarán de asignar un puntaje  $p_{uj}$  a cada ítem  $j$  en relación al usuario  $u$ . El puntaje  $p_{uj}$  sirve para conocer el nivel de relevancia que tiene un ítem  $j$  para un usuario  $u$ , según el algoritmo utilizado. Es imposible calcular las verdaderas preferencias de los usuarios en la etapa de entrenamiento [31]. Por lo tanto, se utilizan las calificaciones estimadas de los artículos. El promedio de precisión a utilizar en este trabajo es el siguiente [32]:

$$eprecision(L) = \frac{1}{k} \sum_{l \in L} p_{ul} \quad (1)$$

donde  $L$  es una lista de recomendación,  $k$  es la longitud de la lista de recomendación,  $l$  es un elemento de la lista de recomendación,  $u$  es el usuario al que le corresponde las recomendaciones y  $p_{ul}$  es el puntaje asignado para un artículo  $l$  de la lista  $L$ , para un usuario  $u$ .

La *novedad* en RS se refiere a cuán diferentes son los ítems recomendados con respecto a los que un usuario ya ha visto antes. Una posibilidad para medir el nivel de novedad de una recomendación es utilizar la novedad basada en popularidad que fue propuesta por Castells et al. [22], expresada como la probabilidad de que un ítem  $j$  sea visto por un usuario seleccionado aleatoriamente, que es estimada como:

$$P(visto|j) = \frac{|u \in C | r(u, j) \neq \phi|}{|C|} \quad (2)$$

donde  $C$  es el conjunto de todos los usuarios,  $u$  es un usuario,  $j$  es un ítem y  $r(u, j)$  es la valoración de un usuario a un ítem. Entonces se puede definir novedad como [22]:

$$novedad(L(u)) = \sum_{j \in L(u)} (1 - P(visto|j)) \quad (3)$$

donde  $L(u)$  es una lista de recomendación para un usuario  $u$ .

La *diversidad* en RS se refiere a cuán diferentes (o distantes) son los objetos recomendados unos de otros. La diversidad media de los ítems de una lista  $L$  está dado por [32]:

$$div(L) = \frac{1}{k(k-1)} \sum_{g \in L} \sum_{l \in L, l \neq g} d(g, l) \quad (4)$$

donde  $k$  es la longitud de la lista de recomendación,  $g$  es un elemento de la lista de recomendación,  $l$  es otro elemento de la lista de recomendación y  $d(g, l)$  es una función de distancia que asegura la simetría, como:

$$d(g, l) = 1 - sim(g, l) \quad (5)$$

teniendo en cuenta que  $sim(g, l) \in [0, 1]$  es alguna medida de similaridad (como por ejemplo la de coseno). Entonces,  $div(L)$  está dado por el promedio de disimilaridad entre todos los pares de ítems contenidos en  $L$ . Un valor mayor corresponde a más diversidad.

### III-A. Métricas de evaluación de las recomendaciones

Se utilizarán las métricas de ranking basadas en la característica operativa del receptor, del inglés *Receiver Operating Characteristic* (ROC) [2]. Se puede suponer que se desea seleccionar un sub-conjunto de  $k$  ítems de una lista  $L$ , para recomendar al usuario. El conjunto de elementos recomendados se denota por  $S(k)$ , por lo que  $|S(k)| = k$ . Se propone que  $G$  represente el verdadero conjunto de elementos relevantes (positivos verdaderos) que consume el usuario. Luego, para cualquier tamaño  $k$  dado de la lista recomendada, la precisión se define como el porcentaje de elementos recomendados que verdaderamente resultan ser relevantes (es decir, consumidos por el usuario) [2].

$$precision(k) = \frac{|S(k) \cap G|}{|S(k)|} \quad (6)$$

La recuperación, del inglés *recall* se define como el porcentaje de positivos verdaderos  $G$ , que se han recomendado como positivos para una lista de tamaño  $t$ .

$$recall(k) = \frac{|S(k) \cap G|}{|G|} \quad (7)$$

Una forma de crear una medida única que resuma tanto la precisión como la recuperación es la medida  $F1$  [33], que es la media armónica entre la precisión y la recuperación:

$$F1(k) = \frac{2 \cdot precision(k) \cdot recall(k)}{precision(k) + recall(k)} \quad (8)$$

Las listas finales de recomendación también serán evaluadas en este trabajo en términos de novedad y diversidad, arriba ya definidas.

## IV. PROBLEMAS MULTI-OBJETIVO

El problema que se trata en este trabajo es multi-objetivo (MOP). Muchas veces estos problemas tienen objetivos conflictivos entre sí. Cuando se tiene más de un objetivo, se vuelve usual la necesidad de elegir entre un conjunto de soluciones no comparables (igualmente buenas desde un punto de vista matemático), debido a que puede no existir una única solución óptima del problema [34]. Se presentan a continuación las definiciones usadas en este trabajo.

*Dominancia Pareto:* Dados dos vectores de decisión  $\hat{q}$ ,  $\hat{z}$ , se dice que  $\hat{q}$  domina a  $\hat{z}$  lo que se denota como  $\hat{q} \succ \hat{z}$ , si y solo si  $\hat{q}$  es mejor o igual a  $\hat{z}$  en todos los objetivos y es estrictamente mejor en al menos un objetivo [35]. *Conjunto Pareto Óptimo:* el Conjunto Pareto Óptimo, representado como  $\mathcal{P}^*$ , se define como el conjunto de todas las soluciones factibles no dominadas en el espacio de decisión  $X$ . *Frente Pareto Óptimo:* el Frente Pareto Óptimo, que se expresa como  $\mathcal{PF}^*$ , se define como la imagen en el espacio objetivo del Conjunto Pareto Óptimo  $\mathcal{P}^*$  [34]. *Relacion de preferencia:* Dados dos vectores de decisión  $\hat{q}$  y  $\hat{z} \in \mathcal{X}$ , se dice que  $\hat{q}$  es preferido a  $\hat{z}$  si  $\hat{a}$  objetivos de  $\hat{q}$  son mejores que los correspondientes objetivos de  $\hat{z}$  y solo  $\hat{a}'$  ( $\hat{a}' < \hat{a}$ ) objetivos de  $\hat{z}$  son mejores que los objetivos correspondientes en  $\hat{q}$  [36]. En otras palabras,  $\hat{q}$  es preferido a  $\hat{z}$  si  $\hat{q}$  tiene más objetivos que resulten mejores que  $\hat{z}$ . Un ejemplo podría ser el siguiente,  $\hat{z}$  y  $\hat{q}$  tienen 3 objetivos cada uno,  $\hat{z}$  tiene 2 objetivos mejores ( $\hat{a} = 2$ ) y uno peor que  $\hat{q}$  ( $\hat{a}' = 1$ ), entonces  $\hat{z}$  es preferido porque  $2 > 1$ .

## V. PROBLEMA DE DIVERSIDAD MÁXIMA

El problema de la Diversidad Máxima (MDP) consiste en seleccionar un subconjunto de elementos de forma que la medida de diversidad del conjunto sea maximizada [37]. Es decir, consiste en seleccionar un subconjunto  $\hat{K} = \{\hat{i}\}$ ,  $\hat{i} = 1, \dots, |\hat{K}|$  de un conjunto  $J = \{\hat{j}\}$ ,  $\hat{j} = 1, \dots, |J|$  ( $|\hat{K}| < |J|$ ), de forma a maximizar la diversidad entre los elementos del subconjunto  $\hat{K}$  seleccionado. La medida de diversidad o medida de dispersión (*div*) está relacionada a una definición de distancia  $d$ , por lo tanto, para resolver el MDP necesitamos generar la matriz de distancias (o disimilitudes) inter-elemento, para luego maximizar la diversidad.

### V-A. Resolución del problema de la diversidad máxima en un contexto multi-objetivo

Según Machuca et al. [8], cada especialista puede preferir una definición de distancia  $d$  diferente. En consecuencia, surge la necesidad de resolver este problema como un problema

de optimización multi-objetivo, dado que diferentes investigadores pueden preferir definiciones distintas de distancia (o disimilitud).

El problema MDP multi-objetivo se plantea como [8]:

$$\text{Maximizar } \vec{div} = [div^1(\dot{K}) \dots div^{\hat{k}}(\dot{K})]^T \in \mathbb{R}^{\hat{k}} \quad (9)$$

donde  $\hat{k}$  representa la cantidad de métricas de distancia consideradas y la  $i$ -ésima función objetivo  $div^i(\dot{K})$  se define como:

$$div^i(\dot{K}) = \sum_{g < l} d_{g,l}^i; \text{ donde } i \in 1 \dots \hat{k}, \quad g, l \in \dot{K} \quad (10)$$

$d_{g,l}^i$  representa la  $i$ -ésima definición de distancia considerada entre 2 ítems  $g$  y  $l$ , pertenecientes a  $\dot{K}$ .

## VI. PROPUESTA DEL TRABAJO

### VI-A. Caso de estudio

Se considera una empresa de electrodomésticos, que denominamos “Empresa 1”. La Empresa 1 desea ofrecer recomendaciones personalizadas de productos a sus clientes por diferentes medios: (1) telefónico; (2) mensajes de texto; (3) aplicaciones para teléfonos celulares; entre otros medios. El negocio en cuestión, está interesado en conocer cuáles podrían ser los productos de interés para cada consumidor en particular. La información que se tiene disponible es el historial de compras de sus clientes. Según la Empresa 1, el comportamiento de compra de sus clientes tiene dos características: (1) compras de un artículo por vez, es decir que la cantidad de artículos en el carro de compras de un cliente, en general, es uno; y (2) compras de artículos no repetidos, es decir, que los usuarios, en general, no realizan la compra de un artículo ya adquirido con anterioridad.

Este comportamiento podría explicarse en parte por los tipos de artículos vendidos, ya que son productos de costos elevados y de prolongada duración, por lo que no hay necesidad de una compra recurrente de un mismo artículo. Un ejemplo podría ser una heladera de un precio elevado que no se necesita renovarla constantemente.

### VI-B. Descripción general del proyecto

Este proyecto trata de la implementación de un innovador sistema de recomendación, basado en la extracción de datos necesarios del ERP de la empresa en cuestión. Se define el tipo de retroalimentación a ser utilizada como implícita. Se seleccionan los recomendadores y los métodos a utilizar para realizar las recomendaciones de entre las opciones propuestas en este trabajo. Se proponen 4 opciones de solución detalladas más abajo. Estas opciones generarán como resultado conjuntos Pareto  $\mathcal{P}^*$  que serán analizados por un módulo tomador de decisiones o una persona física (experto) para seleccionar un único individuo (o lista  $L$  de ítems a ser finalmente recomendada a un usuario  $u$ ) del conjunto de posibles soluciones  $\mathcal{P}^*$ .

### VI-C. Diseño de la solución

En el proceso general de recomendación en un entorno multi-objetivo se tiene un universo  $U$ , formado por ítems  $j$ . Este universo  $U$  es analizado con un optimizador multi-objetivo al momento de proponer una lista  $L$  de recomendación. El optimizador forma una población, donde cada individuo representa a una lista de recomendación  $L$ , por lo que una población de  $z$  individuos es denotada como  $L_1, L_2, \dots, L_z$ . Como siguiente paso, se calculan las funciones de aptitud, del inglés *fitness*, para cada individuo. Mediante operaciones de selección, cruzamiento y mutación se evoluciona la población hasta obtener las soluciones buscadas. Como resultado de este proceso se obtiene un conjunto Pareto  $\mathcal{P}^*$ , en el cual están contenidas las mejores soluciones no dominadas. Un tomador de decisiones seleccionará una lista de las contenidas en el conjunto Pareto  $\mathcal{P}^*$  utilizando la estrategia que considere adecuada. Las 4 opciones presentadas implementan variaciones en cualquiera de los pasos del proceso de recomendación arriba descrito. Por limitaciones de espacio, solo se explica brevemente cada una de las 4 opciones.

**VI-C1. Opción 1:** Para esta opción se utilizan distintos tipos de recomendadores tradicionales de manera a formar un sistema híbrido de recomendaciones. Podrán ser agregados también recomendadores no tradicionales si se desea aumentar el nivel de diversidad en los resultados, pero a costa de perder personalización. El conjunto de los recomendadores utilizados es denotado como  $M_1, M_2, \dots, M_q$ , siendo  $M$  un recomendador y  $q$  la cantidad de recomendadores. Cada uno de estos recomendadores expondrá como resultado una única lista de recomendación con artículos ya ordenados en un *ranking*. El *ranking* se obtiene mediante un valor de predicción asignado por cada recomendador a cada artículo. Los recomendadores tienen distintas formas de asignar valores a las predicciones, por ende los resultados deberán pasar por un proceso de normalización *Norm*. El proceso de normalización *Norm* es necesario para poder clasificar las recomendaciones a un mismo nivel y así poder compararlas. Siendo  $v \in \mathbb{R}^k$  el vector de las valoraciones o *ratings* de cada elemento,  $j \in L_m$ ,  $m \in \{1, 2, \dots, q\}$ , se define  $\hat{v} \in \mathbb{R}^k$  como el vector de rating normalizado  $\hat{v}_j = v_j * Norm_m \quad \forall j \in \{1 \dots k\}$ , donde  $k$  es el tamaño de la lista de recomendación.  $\hat{L}_m$  es la resultante de la lista de recomendación normalizada, correspondiente a un recomendador  $M_m$ . Las listas resultantes serán utilizadas para crear un nuevo universo  $U$  conformado por todos los artículos de cada una de las  $q$  listas y solo por estos artículos. El *fitness* de cada individuo del algoritmo de optimización se calcula en base a la diversidad y la novedad. La precisión no es calculada en esta opción debido a que se asume que todos los artículos que fueron incluidos en el nuevo universo  $U$  tienen un buen nivel de personalización. Se tendrán como objetivos del optimizador, maximizar las distintas métricas de diversidad y la novedad, arriba descritas.

**VI-C2. Opción 2:** Esta opción se diferencia de la *Opción 1* en que no se permite utilizar los recomendadores no tradicionales, debido a que se hace uso del *rating* calculado por cada

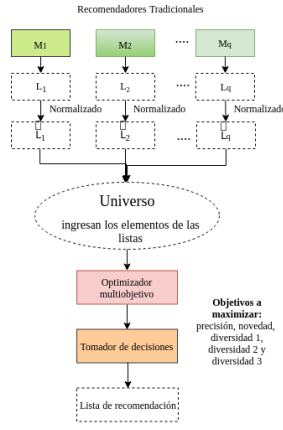


Figura 1. Esquema del recomendador *Opción 2*.

recomendador tradicional, considerando que los recomendadores no tradicionales no poseen un *rating* asociado. El *fitness* se calcula considerando la precisión, diversidad y novedad de la lista. Se utiliza la precisión como estimación de la exactitud de la lista en la función de *fitness*. Las estimaciones de predicciones brindadas por los recomendadores son utilizadas como precisión, dado que en las simulaciones se asume no conocer los datos del conjunto de prueba [31]. Se puede observar el diseño de esta opción en la Figura 1.

**VI-C3. Opción 3:** Esta opción es un caso especial de la *Opción 2* arriba presentada anteriormente. La diferencia se centra en la asignación de ponderaciones (o pesos) a los *ratings* de los artículos, dependiendo de cual fue el recomendador tradicional que lo sugirió. Los ítems  $j$  del recomendador más relevante serán ponderados de una mejor manera que el de los otros recomendadores. La relevancia de los recomendadores se mide con la exactitud, obtenida en este caso mediante un conjunto de validación. Una vez que los *ratings* de los ítems obtengan sus nuevos valores, pasarán a formar parte del universo  $U$ .

**VI-C4. Opción 4:** Para esta opción se utilizará todo el conjunto de artículos, para conformar el universo  $U$ . Este universo será tratado con un optimizador multi-objetivo que calcula listas de recomendaciones  $L_1, L_2, \dots, L_z$  donde  $z \in \mathbb{N}$  indica el número de soluciones del conjunto Pareto  $\mathcal{P}^*$ . El *fitness* utilizado considera la exactitud (en función al conjunto de validación), diversidad y novedad de la lista. Claramente, el espacio de búsqueda es mucho mayor por lo que se esperaría un tiempo de computo considerablemente mayor para obtener buenas soluciones.

#### VI-D. Tomador de decisiones

En este proyecto se propone utilizar el término tomador de decisiones, en adelante TD, como un módulo encargado de seleccionar una solución de entre todas las soluciones no comparables que resulten de la optimización multi-objetivo. El TD elige una solución (lista  $L$ ) del conjunto Pareto retornado por el módulo optimizador multi-objetivo. Se deberá elegir la política por defecto a utilizar en la elección de solución. Las

estrategias para la selección de la solución podrían ser: (1) el punto más lejano al origen de coordenadas; (2) la solución más cercana a una lista de recomendación no tradicional (se propone esta opción para dar una inclinación comercial a las recomendaciones); (3) por pesos asignados a los objetivos (se propone esa solución de manera a definir explícitamente la importancia de cada objetivo); y (4) orden lexicográfico (se debe tener pre-definido explícitamente el orden de importancia de los objetivos).

#### VI-E. Problema de recomendación multi-objetivo y problema de la dispersión máxima

Este trabajo propone resolver el MDP multi-objetivo dentro de las listas de recomendación, utilizando varias matrices de distancia, obtenidas cada una de ellas con una definición diferente de distancia, sobre la misma base de datos, convirtiéndose así cada una, en la matriz de entrada que permite calcular cada función objetivo del problema MDP, como se propone en [8]. Existen definiciones diferentes de distancia, de entre las cuales este trabajo en principio solo considera las 3 siguientes: coseno, euclidiana y de Jaccard. Se pueden encontrar 65 definiciones de distancia diferentes, en [8].

Los objetivos a optimizar serían los siguientes: (1) *exactitud*; (2) *novedad*; (3) *diversidad*<sub>1</sub> (coseno); (4) *diversidad*<sub>2</sub> (euclidiana) y; (5) *diversidad*<sub>3</sub> (jaccard). Entonces, el problema puede ser planteado como:

$$\text{Maximizar } \vec{Y}(x) = \{F_e(x), F_n(x), F_{d1}(x), F_{d2}(x), F_{d3}(x)\} \quad (11)$$

siendo  $F_e$  la función de exactitud,  $F_n$  la función de novedad y  $F_{d1}$  la función de *diversidad*<sub>1</sub>,  $F_{d2}$  la función de *diversidad*<sub>2</sub>,  $F_{d3}$  la función de *diversidad*<sub>3</sub>.

## VII. IMPLEMENTACIÓN

#### VII-A. Conjunto de datos

Luego de un análisis del contenido de la base de datos de la Empresa 1, se extrajo el historial de ventas de los usuarios, los artículos, clasificaciones (de estos artículos) y cantidad de compras de los mismos. Este conjunto de datos implica 17.080 clientes, 2.319 artículos y 33.422 transacciones. Estos datos representan las compras realizadas por los clientes. La información obtenida fue tratada como *feedback* implícito.

**VII-A1. Preprocesamiento y transformación de datos:** El primer paso fue el filtrado de la información correspondiente. Mediante este filtrado sobre el conjunto total de datos, se pudo remover datos conflictivos. Si los datos no son preprocesados, la información con la cual se trabaja podría estar “sucia” o distorsionada, pudiendo resultar en inconsistencia de datos. Los datos se filtran de la siguiente manera: (1) se identificaron los datos nulos y faltantes, tomando la decisión de eliminarlos; (2) se identificaron artículos duplicados (ya sea por error de escritura del usuario que lo cargó u otro motivo) y se procedió a unificarlos.

Como segundo paso se recalculó la cantidad de usuarios, cantidad de ítems, transacciones y se computó la dispersión

de la matriz usuario-item formada [2]. Se utilizaron las clasificaciones de los ítems y no los ítems en sí. Se implementó esta técnica con el objetivo de disminuir la dimensión de la matriz de recomendaciones  $R$  [2], disminuyendo el tamaño de  $b$  (cantidad de ítems) a  $c$  (cantidad de clasificaciones). Un ejemplo simple sería el siguiente: un celular Samsung galaxy y un celular Huawei P9, forman parte de la clasificación denominada “Celular”. Esta técnica fue implementada por Sano et al. [11] que indica que si se entrenan y se evalúan con artículos se observa que el promedio de exactitud es de 1.8 %. Si en cambio se entrenan y se evalúan con clasificaciones el promedio de exactitud asciende a 22 %. Por lo expuesto, se decidió utilizar las clasificaciones de los artículos.

Posteriormente se agruparon todas las compras de los usuarios, de manera a conocer la frecuencia de compra de los productos, por cada usuario. Se optó por realizar una transformación de los datos originales a datos únicos con lo cual podemos representar por “1” al producto adquirido por un usuario y “0” en caso contrario.

Con el conjunto de datos convertidos a unarios, se construye una matriz dispersa de dimensión 932 x 215 siendo 932 la cantidad de clientes, 215 la cantidad de clasificaciones de artículos y su contenido las transacciones obtenidas de los datos seleccionados que suman 5.136.

#### VII-B. Recomendadores y configuraciones

Los datos iniciales se dividieron en tres grupos: datos de entrenamiento, datos de validación y datos de prueba. Para realizar la configuración de algoritmos se utilizaron datos de validación de modo a sobre-alimentar al modelo. La configuración para el algoritmo ALS fue:  $factores = 10$ ,  $alpha = 15$ ,  $regularizacion = 0,1$  e  $iteraciones = 40$ . La configuración para la realización del entrenamiento del modelo utilizando el algoritmo KNN fue de  $K = 10$ , siendo  $K$  el número de vecinos definidos. La configuración para la realización del entrenamiento del modelo utilizando el algoritmo BM25 (KNN con variación *Best Matching 25*) fue de  $K = 90$ .

Los recomendadores multi-objetivo utilizan un MOEA conocido como *Non-dominated Sorting Genetic Algorithm, Version-2* NSGA-II [38]. Se eligió este algoritmo porque es reconocido como el MOEA más representativo del estado del arte [39]. Para las opciones 1, 2 y 3 se eligieron 150 generaciones. Para la Opción 4 se eligió inicialmente el mismo número de generaciones para que los tiempos de corrida sean similares a pesar de tener una entrada  $U$  considerablemente mayor. Para las opciones 1, 2 y 3 se eligió una población inicial de 300 individuos debido a que la cantidad de elementos de entrada es considerablemente menor que para la Opción 4. En efecto, para la Opción 4 se eligió una población inicial de 500 individuos debido a que posee una mayor cantidad de elementos de entrada. Para todas las opciones se definió un porcentaje de mutación alto de 0,3. Las listas de recomendación generadas son de 6 elementos en todos los casos presentados en este trabajo.

#### VII-C. Herramientas utilizadas

Para la construcción y programación del sistema de recomendación se utilizó el lenguaje de programación Python. Para la implementación de la optimización multi-objetivo dentro del sistema de recomendación se utilizó la librería Deap [40]. Para la implementación de los algoritmos dentro del sistema de recomendación se utilizó la librería Implicit (<https://github.com/benfred/implicit>). En esta librería se encuentran implementados los algoritmos ALS y KNN (con variaciones). Estos algoritmos fueron modificados para aceptar y operar con datos implícitos.

#### VII-D. Métricas de evaluación

Las métricas utilizadas para medir el impacto de las recomendaciones son: exactitud, novedad y 3 métricas de diversidad. Para medir la exactitud, se utilizó la relación  $F1$  debido a que las listas de recomendación no se evalúan a través de un rating, sino a través de sus aciertos sobre el conjunto de datos de prueba. Para medir la diversidad se codificó e implementó la métrica definida en secciones anteriores, como también las definiciones de distancia (Coseno, Jaccard y Euclidiana [8]), y de igual manera se implementaron la métrica novedad y las definiciones de dominancia y preferencia.

Para evaluar las recomendaciones en un entorno multi-objetivo se utilizaron las definiciones de dominancia y preferencia de la sección IV.

### VIII. RESULTADOS EXPERIMENTALES

El *hardware* utilizado es un computador con procesador Intel® Core™i5 4570 con 24 GB RAM y sistema operativo Deepin 15.5. El lenguaje de programación Python fue utilizado para la implementación de los algoritmos y el cálculo de las métricas experimentales.

A continuación se presentan la notación utilizada para analizar los resultados.  $Knn$  es el *K Nearest Neighbors*,  $Bm25$  es el *K Nearest Neighbors* con variación utilizando  $Bm25$ ,  $Als$  es el *Alternating Least Squares*.  $Op1$ ,  $Op2$ ,  $Op3$  y  $Op4$  son las opciones de recomendador propuestas en este trabajo.  $Op1O$ ,  $Op2O$ ,  $Op3O$ ,  $Op4O$  representan las opciones de recomendador combinado con la política de selección de la distancia más lejana al punto de origen.  $Op1L$ ,  $Op2L$ ,  $Op3L$ ,  $Op4L$  representan las opciones de solución con la política de selección de orden lexicográfico.  $Op1P$ ,  $Op2P$ ,  $Op3P$ ,  $Op4P$  representan las soluciones con la política de selección por pesos.

Para realizar los siguientes estudios se dividieron los datos en 10 cortes, partiendo del conjunto inicial de datos. Esta división de datos se realizó utilizando la técnica *K-fold* [2], [41] y [42]. En los siguientes estudios se muestran los promedios de los valores de exactitud, novedad y diversidad en cada una de las 10 divisiones mencionadas. Para conocer el número de corridas realizadas para obtener los resultados experimentales que siguen, se debe considerar: (a) 4 opciones de solución; (b) 3 políticas de selección de decisión; (c) 3 corridas para encontrar los mejores valores de exactitud, novedad y diversidad individualmente; (d) 3 recomendadores

tradicionales; y (e) 10 conjuntos de datos. Realizando el siguiente cálculo  $((a*b*c)+d)*e = ((4*3*3)+3)*10 = 390$  se tiene que los resultados que siguen generaron 390 modelos. Cada modelo tiene una lista de recomendación para cada uno de los 932 usuarios estudiados, esto es  $390 * 932 = 363,480$  listas.

Los valores generados por los recomendadores tradicionales serán utilizados como datos de referencia para realizar las comparaciones respectivas. Estos valores base se encuentran también en las tablas donde se realizarán las comparaciones.

#### VIII-A. Comparación de los recomendadores en un entorno multi-objetivo

A continuación se evaluarán los resultados de los recomendadores en un entorno puramente multi-objetivo. Se compararán las recomendaciones brindadas en término de los 5 objetivos considerados: exactitud (utilizando *F1*), novedad y las 3 métricas de diversidad. En una primera comparación de dominancia no se pudo determinar cual recomendador utilizar, porque la mayoría de las soluciones son no comparables entre sí. Entonces, con el objetivo de seleccionar un recomendador de entre los diferentes propuestos se utilizará el concepto de preferencia. En este caso se buscará el recomendador que presente el mayor número de soluciones preferidas respecto a los otros recomendadores.

**VIII-A1. Comparación de preferencia:** En las tablas I y II se presenta la cantidad de veces que fue preferida la solución presentada por cada recomendador, en 10 cortes. En la primera fila de las tablas I y II así como en la primera columna se encuentran los recomendadores utilizados en el proyecto. Las celdas de la tabla representan la intercepción entre dos recomendadores. En estas celdas se muestran la cantidad de veces que la solución de un recomendador fue preferida por sobre la solución de otro recomendador. Como ejemplo se podría tener el siguiente caso en la Tabla I: el *Knn* (fila 4) es preferido a *ALS* (columna 2) 10 veces, por lo que *ALS* nunca es preferido a *Knn* (fila 2, columna 4). El total que se encuentra en la penúltima columna de la Tabla II representa las veces que un recomendador brindó una solución que fue preferida sobre otra solución, un ejemplo sería que *Op1O* (fila 5) fue preferida 77 veces. El total de la penúltima columna se calcula sumando la fila correspondiente a cada recomendador. En la última columna se muestra un *ranking* de los recomendadores preferidos, medidos con la relación de preferencia. En la última fila de cada tabla se encuentra la cantidad de veces que un recomendador brindó una solución no-preferida, un ejemplo de ello en la Tabla I es *Op1O* (columna 5) que dio 62 soluciones no-preferidas. El total de la última fila se calcula sumando la columna correspondiente a cada recomendador.

Para tener una métrica de preferencia de los recomendadores, se opta por hallar una relación preferencia/no-preferencia. La relación preferencia/no-preferencia es hallada dividiendo la cantidad de veces que un recomendador fue preferido sobre las veces que no fue preferido más uno. La relación preferencia/no-preferencia se define como:

Tabla I  
Preferencia de recomendadores comparando otras políticas estudiadas.

Rec	Als	Bm25	Knn	Op1O	Op1L	Op1P	Op2O	Op2L	Rank
Als	X	0	0	0	0	0	0	0	13 <sup>o</sup>
Bm25	10	X	3	0	0	0	0	0	14 <sup>o</sup>
Knn	10	0	X	0	0	0	0	0	12 <sup>o</sup>
Op1O	10	10	10	X	0	0	1	0	9 <sup>o</sup>
Op1L	10	10	10	10	X	0	0	1	7 <sup>o</sup>
Op1P	10	10	10	9	0	X	1	0	5 <sup>o</sup>
Op2O	10	10	10	9	10	9	X	10	1 <sup>o</sup>
Op2L	10	10	10	10	9	10	0	X	3 <sup>o</sup>
Op2P	10	10	10	3	10	2	0	0	6 <sup>o</sup>
Op3O	10	10	10	9	10	9	4	10	2 <sup>o</sup>
Op3L	10	10	10	10	9	10	0	1	4 <sup>o</sup>
Op3P	10	10	10	2	10	1	0	1	8 <sup>o</sup>
Op4O	10	10	10	0	0	0	0	0	10 <sup>o</sup>
Op4L	0	0	0	0	0	0	0	0	15 <sup>o</sup>
Op4P	10	0	10	0	0	0	0	0	11 <sup>o</sup>
Total	130	100	113	62	58	51	6	23	-

Tabla II  
Preferencia de recomendadores comparando con políticas estudiadas.  
(Continuación)

Rec	Op2P	Op3O	Op3L	Op3P	Op4O	Op4L	Op4P	Total	Rank
Als	0	0	0	0	0	10	0	10	13 <sup>o</sup>
Bm25	0	0	0	0	0	10	10	33	14 <sup>o</sup>
Knn	0	0	0	0	0	10	0	20	12 <sup>o</sup>
Op1O	7	1	0	8	10	10	10	77	9 <sup>o</sup>
Op1L	0	0	1	0	10	10	10	82	7 <sup>o</sup>
Op1P	8	1	0	9	10	10	10	88	5 <sup>o</sup>
Op2O	10	5	10	10	10	10	10	133	1 <sup>o</sup>
Op2L	10	0	9	9	10	10	10	117	3 <sup>o</sup>
Op2P	X	1	2	6	10	10	10	84	6 <sup>o</sup>
Op3O	9	X	10	10	10	10	10	131	2 <sup>o</sup>
Op3L	8	0	X	9	10	10	10	107	4 <sup>o</sup>
Op3P	4	0	1	X	10	10	10	79	8 <sup>o</sup>
Op4O	0	0	0	0	X	8	10	48	10 <sup>o</sup>
Op4L	0	0	0	0	2	X	0	2	15 <sup>o</sup>
Op4P	0	0	0	0	0	10	X	30	11 <sup>o</sup>
Total	56	8	33	61	92	138	110	X	-

$$rel-preferencia = \frac{preferidos}{no-preferidos + 1} \quad (12)$$

Mediante los resultados presentados en las tablas I y II se puede observar que los recomendadores multi-objetivo (menos *Op4L* y *Op4P*) obtienen consistentemente más soluciones preferidas que los recomendadores tradicionales. Las mejores soluciones en términos de relación de preferencia son *Op2O*, *Op3O*, *Op2P* y *Op3L* en el orden presentado. Desde un punto de vista multi-objetivo se concluye entonces que *Op2O* sería el recomendador preferido, ya que presentó la mayor cantidad de soluciones preferidas. También, se puede observar que la política de selección de solución del punto más lejano al origen tiene las mejores soluciones para cada opción de solución (menos *Op1*). Es decir, *OpXO* es el recomendable de entre las políticas presentadas para *OpX* donde  $X \in \{2, 3, 4\}$ .

#### VIII-B. Tasa de aciertos

Esta sección se centra exclusivamente en la exactitud de cada técnica de recomendación. La exactitud se mide por los aciertos que tuvieron los recomendadores en el conjunto de prueba. Para cuantificar los aciertos se utiliza la definición de *F1* (Sección III-A). En la Tabla III se presenta el promedio de exactitud obtenido por cada recomendador. En la primera columna de la Tabla III se encuentran los recomendadores utilizados en el proyecto.

De los resultados presentados en la Tabla III se puede apreciar que *Op1P*, *Op1L* y *Bm25* en este orden, son las



Tabla III  
EVALUACIÓN DE LA EXACTITUD

Recomendador	F1	Nov	Div 1	Div 2	Div 3	Rank.
Als	0.0685	0.9950	0.2086	0.1868	0.1695	12 <sup>o</sup>
Bm25	0.1440	0.9937	0.2311	0.4503	0.3239	3 <sup>o</sup>
Knn	0.1325	0.9933	0.2614	0.5865	0.3516	5 <sup>o</sup>
Op1O	0.0005	0.9238	0.2935	0.9142	0.5297	15 <sup>o</sup>
Op1L	0.2042	0.9998	0.1877	0.0013	0.0036	2 <sup>o</sup>
Op1P	0.2051	0.9998	0.1877	0.0013	0.0037	1 <sup>o</sup>
Op2O	0.0591	0.9231	0.2956	0.9253	0.5344	9 <sup>o</sup>
Op2L	0.0852	0.9998	0.1875	0.0012	0.0033	7 <sup>o</sup>
Op2P	0.0445	0.9260	0.2887	0.8841	0.5059	13 <sup>o</sup>
Op3O	0.0528	0.9232	0.2947	0.9233	0.5329	10 <sup>o</sup>
Op3L	0.0828	0.9998	0.1868	0.0012	0.0034	8 <sup>o</sup>
Op3P	0.0494	0.9242	0.2901	0.9049	0.5145	11 <sup>o</sup>
Op4O	0.0123	0.9572	0.2225	0.4550	0.2955	14 <sup>o</sup>
Op4L	0.1333	0.9963	0.1824	0.0220	0.0190	4 <sup>o</sup>
Op4P	0.1321	0.9963	0.1822	0.0220	0.0190	6 <sup>o</sup>

mejores opciones en términos de exactitud. *Op1L* y *Op1P* muestran una mejoría por encima de los recomendadores tradicionales, lo cual claramente es deseable.

### VIII-C. Discusión

Mediante la comparación de preferencia realizada en la Sección VIII-A1 se puede concluir que los recomendadores propuestos (*Op1*, *Op2*, *Op3* y *Op4*) son mejores en general que los recomendadores tradicionales, visto desde un punto de vista puramente multi-objetivo. *Op2O* fue el recomendador con más soluciones preferidas y por ende el sugerido entre las opciones comparadas. Se destaca que para *Op2*, *Op3* y *Op4* la política de selección del punto más lejano al origen de coordenadas es la que más soluciones preferidas generó. *Op4* queda en los últimos lugares del ranking en la Tabla II debido a que 150 generaciones no parecen suficientes para converger el MOEA utilizado. No se aumentó el número de generaciones con el objetivo de hacer una comparación contra las otras opciones con tiempo de cómputo similares. Sin embargo, posteriormente se realizó otro experimento aumentando el tiempo de computo para *Op4* y los resultados mejoraron considerablemente volviendo competitiva a esta opción a costa de un mayor tiempo de computo.

En un principio se propuso obtener mayor diversidad en las recomendaciones, como la Empresa 1 lo sugirió. Sin embargo, un alto nivel de diversidad no genera la tasa de aciertos esperada. En la etapa de calibración se pudo observar que el aumento de diversidad influía negativamente en la tasa de aciertos de las recomendaciones. Luego de obtener los resultados de la etapa de calibración, se detectó que la diversidad no aumenta la tasa de satisfacción de los usuarios. Esto fue evaluado fuera de línea, con datos implícitos del cliente. Aún así, más diversidad en las recomendaciones puede desempeñar un rol importante estando en contacto directo con el cliente, lo cual puede no reflejarse en los datos implícitos.

Si únicamente se busca optimizar la exactitud, se podría utilizar solamente los métodos de recomendación tradicionales. Los recomendadores tradicionales de por sí obtienen una buena tasa de aciertos y tienen un menor costo computacional. En contrapartida, los recomendadores tradicionales por sí solos no están preparados para obtener un resultado diverso y novedoso, menos si se los compara en un entorno multi-objetivo. Para el caso de la Empresa 1, con *Op1P* y *Op1L* se

logró conseguir un aumento del 7 % en las tasa de acierto y un mejor nivel de novedad con respecto a los recomendadores tradicionales. Lógicamente, en el caso de dar prioridad a un objetivo como exactitud en las recomendaciones, la política de selección del punto más lejano al origen de coordenadas no resultó ser el mejor criterio. Esto se debe a que esta política de decisión elige una solución generalmente en medio de todas las soluciones en el espacio objetivo, no asignando más relevancia a ningún objetivo en particular. En contrapartida, para la evaluación en un entorno multi-objetivo, es la política que brinda las mejores soluciones de preferencia.

Las recomendaciones realizadas para la Empresa 1 en la práctica son de las clasificaciones de un artículo. Un ejemplo de clasificación es “celular de gama alta” y esto deja libre la posibilidad de seleccionar la marca del “celular de gama alta” al momento de ofrecer al usuario. Esto puede no ser conveniente en el caso de que un cliente tenga preferencia por una marca en particular y la empresa le ofrezca otra marca.

Finalmente, cabe destacar que para realizar las recomendaciones comerciales correspondientes a la Empresa 1 se utiliza el recomendador *Op1P*, el cual ha demostrado ser el más exacto en comparación a otras opciones presentadas. En caso de necesitar diversidad, se utiliza el recomendador *Op3P* u *Op1P* pues estos han demostrado ser los recomendadores con mayor cantidad de soluciones preferidas en cuanto a diversidad. En caso de priorizarse la novedad, se utiliza *Op1P*, el cual ha demostrado obtener las soluciones más novedosas. En el caso de obtenerse por recomendaciones que maximicen los 5 objetivos en un contexto puramente multi-objetivo, se recomienda utilizar el *Op2O*. Para más información sobre los experimentos, consultar [43].

## IX. CONCLUSIÓN

El objetivo principal de este trabajo fué diseñar e implementar un sistema híbrido de recomendación en un entorno multi-objetivo, integrando el sistema mencionado al ERP de una empresa paraguaya. El sistema recomendador cuenta con recomendadores tradicionales, no tradicionales, así como 4 opciones de recomendadores (3 de ellos híbridos). Estos recomendadores híbridos se logran combinando los resultados de otros recomendadores mediante un módulo optimizador multi-objetivo. Este sistema recomendador es independiente pero a la vez integrado al ERP. También se logró implementar recomendadores no tradicionales y la utilización de tres métricas de diversidad para realizar recomendaciones. Como aporte innovador de este trabajo se confirmó que es posible implementar la dispersión máxima en un entorno multi-objetivo utilizando métricas distintas de distancia, en el proceso de estructuración de listas de recomendación. Otro resultado relevante es que las soluciones brindadas por las opciones propuestas en este trabajo son preferidas a las obtenidas con recomendadores tradicionales, visto desde un punto de vista puramente multi-objetivo. Para avanzar y mejorar características en la línea de este trabajo, se consideran los siguientes trabajos futuros: (a) agregar más recomendadores, es decir, diferentes algoritmos de recomendación; (b) evaluar otros casos de estudios, es decir,

otras empresas cliente; (c) utilizar otras métricas de diversidad; (d) medir la exactitud en un entorno multi-objetivo utilizado diferentes métricas alternativas; (e) formar un conjunto de datos con *feedbacks* negativos; (f) agregar una interfaz para poder utilizar *feedbacks* explícitos; y (g) agregar más opciones de solución.

## REFERENCIAS

- [1] P. Melville y V. Sindhwani, "Recommender systems", en *Encyclopedia of machine learning*, Springer, 2011, págs. 829-838.
- [2] C. C. Aggarwal, *Recommender Systems*, Springer, 2016.
- [3] M. Sridevi, R. R. Rao y M. V. Rao, "A Survey on Recommender System", *International Journal of Computer Science and Information Security*, vol. 14, n.º 5, pág. 265, 2016.
- [4] J. L. Herlocker, J. A. Konstan, L. G. Terveen y J. T. Riedl, "Evaluating collaborative filtering recommender systems", *ACM Transactions on Information Systems (TOIS)*, vol. 22, n.º 1, págs. 5-53, 2004.
- [5] M. Ge, C. Delgado-Battenfeld y D. Jannach, "Beyond accuracy: evaluating recommender systems by coverage and serendipity", en *Proceedings of the fourth ACM conference on Recommender systems*, ACM, 2010, págs. 257-260.
- [6] J. Wang, Y. Liu, J. Sun, Y. Jiang y C. Sun, "Diversified Recommendation Incorporating Item Content Information Based on MOEA/D", en *System Sciences (HICSS), 2016 49th Hawaii International Conference on*, IEEE, 2016, págs. 688-696.
- [7] M. T. Ribeiro, A. Lacerda, A. Veloso y N. Ziviani, "Pareto-efficient hybridization for multi-objective recommender systems", en *Proceedings of the sixth ACM conference on Recommender systems*, ACM, 2012, págs. 19-26.
- [8] M. Machuca Cabral, B. Barán y F. Sandoya, "El problema de la dispersión máxima en un entorno multi-objetivo", en *XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017)*, 2017.
- [9] M. T. Ribeiro, N. Ziviani, E. S. D. Moura, I. Hata, A. Lacerda y A. Veloso, "Multiobjective pareto-efficient approaches for recommender systems", *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, n.º 4, pág. 53, 2015.
- [10] S. Geuens, K. Coussement y K. W. De Bock, "A framework for configuring collaborative filtering-based recommendations derived from purchase data", *European Journal of Operational Research*, vol. 265, n.º 1, págs. 208-218, 2018.
- [11] N. Sano, N. Machino, K. Yada y T. Suzuki, "Recommendation system for grocery store considering data sparsity", *Procedia Computer Science*, vol. 60, págs. 1406-1413, 2015.
- [12] F. Isinkaye, Y. Folajimi y B. Ojokoh, "Recommendation systems: Principles, methods and evaluation", *Egyptian Informatics Journal*, vol. 16, n.º 3, págs. 261-273, 2015.
- [13] G. Adomavicius y A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions", *IEEE transactions on knowledge and data engineering*, vol. 17, n.º 6, págs. 734-749, 2005.
- [14] J. Bobadilla, F. Ortega, A. Hernando y A. Gutiérrez, "Recommender systems survey", *Knowledge-based systems*, vol. 46, págs. 109-132, 2013.
- [15] G. Adomavicius e Y. Kwon, "Multi-criteria recommender systems", en *Recommender Systems Handbook*, Springer, 2015, págs. 847-880.
- [16] P. E. Osche, S. Castagnos y A. Boyer, "AntRS: Recommending Lists Through a Multi-objective Ant Colony System", en *European Conference on Information Retrieval*, Springer, 2019, págs. 229-243.
- [17] N. Nikzad-Khasmakhi, M. Balafar y M. R. Feizi-Derakhshi, "The state-of-the-art in expert recommendation systems", *Engineering Applications of Artificial Intelligence*, vol. 82, págs. 126-147, 2019.
- [18] R. Burke, "Hybrid recommender systems: Survey and experiments", *User modeling and user-adapted interaction*, vol. 12, n.º 4, págs. 331-370, 2002.
- [19] A. Bellogin, P. Castells e I. Cantador, "Precision-oriented evaluation of recommender systems: an algorithmic comparison", en *Proceedings of the fifth ACM conference on Recommender systems*, ACM, 2011, págs. 333-336.
- [20] O. Kaššák, M. Kompan y M. Bieliková, "Personalized hybrid recommendation for group of users: Top-N multimedia recommender", *Information Processing & Management*, vol. 52, n.º 3, págs. 459-477, 2016.
- [21] A. A. Kardan y M. Ebrahimi, "A novel approach to hybrid recommendation systems based on association rules mining for content recommendation in asynchronous discussion groups", *Information Sciences*, vol. 219, págs. 93-110, 2013.
- [22] P. Castells, S. Vargas y J. Wang, "Novelty and diversity metrics for recommender systems: choice, discovery and relevance", en *International Workshop on Diversity in Document Retrieval (DDR 2011) at the 33rd European Conference on Information Retrieval (ECIR 2011)*, 2011.
- [23] M. Kunaver y T. Požrl, "Diversity in recommender systems—A survey", *Knowledge-Based Systems*, vol. 123, págs. 154-162, 2017.
- [24] T. Zhou, Z. Kuscik, J. G. Liu, M. Medo, J. R. Wakeling e Y. C. Zhang, "Solving the apparent diversity-accuracy dilemma of recommender systems", *Proceedings of the National Academy of Sciences*, vol. 107, n.º 10, págs. 4511-4515, 2010.
- [25] S. Vargas y P. Castells, "Rank and relevance in novelty and diversity metrics for recommender systems", en *Proceedings of the fifth ACM conference on Recommender systems*, ACM, 2011, págs. 109-116.
- [26] Ö. Celma y P. Herrera, "A new approach to evaluating novel recommendations", en *Proceedings of the 2008 ACM conference on Recommender systems*, ACM, 2008, págs. 179-186.
- [27] R. S. Oliveira, C. Nóbrega, L. B. Marinho y N. Andrade, "A Multiobjective Music Recommendation Approach for Aspect-Based Diversification", en *ISMIR*, 2017.
- [28] C.-S. Hwang, "Genetic algorithms for feature weighting in multi-criteria recommender systems", en *Journal of Convergence Information Technology*, Citeseer, 2010.
- [29] K. Vera, F. Lopez-Pires, B. Baran y F. Sandoya, "Multi-objective maximum diversity problem", en *Computer Conference (CLEI), 2017 XLIII Latin American*, IEEE, 2017, págs. 1-9.
- [30] F. Ricci, L. Rokach y B. Shapira, "Introduction to recommender systems handbook", en *Recommender systems handbook*, Springer, 2011, págs. 1-35.
- [31] Y. Zuo, M. Gong, J. Zeng, L. Ma y L. Jiao, "Personalized recommendation based on evolutionary multi-objective optimization [research frontier]", *IEEE Computational Intelligence Magazine*, vol. 10, n.º 1, págs. 52-62, 2015.
- [32] B. Geng, L. Li, L. Jiao, M. Gong, Q. Cai e Y. Wu, "NNIA-RS: A multi-objective optimization based recommender system", *Physica A: Statistical Mechanics and its Applications*, vol. 424, págs. 383-397, 2015.
- [33] Z. C. Lipton, C. Elkan y B. Naryanaswamy, "Optimal thresholding of classifiers to maximize F1 measure", en *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, 2014, págs. 225-239.
- [34] C. Von Lüken, B. Barán y C. Brizuela, "A survey on multi-objective evolutionary algorithms for many-objective problems", *Computational Optimization and Applications*, vol. 58, n.º 3, págs. 707-756, 2014.
- [35] E. Zitzler y L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach", *IEEE transactions on Evolutionary Computation*, vol. 3, n.º 4, págs. 257-271, 1999.
- [36] N. Drechsler, "Self-adaptive Evolutionary Many-Objective Optimization Based on Relation Preferred", en *Computational Intelligence*, Springer, 2016, págs. 23-37.
- [37] C. C. Kuo, F. Glover y K. S. Dhir, "Analyzing and Modeling the Maximum Diversity Problem by Zero-One Programming", *Decision Sciences*, vol. 24, n.º 6, págs. 1171-1185, 1993.
- [38] K. Deb, A. Pratap, S. Agarwal y T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II", *IEEE transactions on evolutionary computation*, vol. 6, n.º 2, págs. 182-197, 2002.
- [39] N. Riquelme, C. Von Lüken y B. Baran, "Performance metrics in multi-objective optimization", en *Computing Conference (CLEI), 2015 Latin American*, IEEE, 2015, págs. 1-11.
- [40] F. A. Fortin, F. M. De Rainville, M.-A. Gardner, M. Parizéau y C. Gagné, "DEAP: Evolutionary Algorithms Made Easy", *Journal of Machine Learning Research*, vol. 13, págs. 2171-2175, 2012.
- [41] C. Schaffer, "Selecting a classification method by cross-validation", *Machine Learning*, vol. 13, n.º 1, págs. 135-143, 1993.
- [42] P. Refaeilzadeh, L. Tang y H. Liu, "Cross-validation", en *Encyclopedia of database systems*, Springer, 2009, págs. 532-538.
- [43] M. Bavera, "Sistema híbrido de recomendación. Un sistema multi-objetivo no convencional.", Trabajo de final de carrera, Universidad Católica de Asunción, oct. de 2018.