UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

Simulación y análisis de datos sintéticos para evaluar algoritmos de aprendizaje por refuerzo en una aplicación móvil de promoción de actividad física

POR

Alejandro Hernán García Zavala

Memoria de Título presentada a la Facultad de Ingeniería de la Universidad de Concepción para optar al título profesional de Ingeniero Civil Industrial.

Profesor Guía:

Juan Carlos Caro Seguel

Enero 2025

Concepción, Chile

Agradecimientos

Resumen

Esta investigación aborda la mitigación del *arranque en frío* en sistemas de recomendación para salud digital mediante la generación de datos sintéticos informados por teoría psicológica. En aprendizaje por refuerzo, el *arranque en frío* aparece cuando el sistema debe decidir sin historial del usuario; se requiere una base de datos inicial para calibrar valores/políticas y así evitar que las primeras interacciones se traduzcan en baja adherencia al comportamiento de salud.

El objetivo es desarrollar y evaluar un simulador de usuarios sintéticos parametrizado con *constructos* psicológicos del Modelo Integrado de Cambio de Comportamiento y variables sociodemográficas. La metodología considera una matriz de covarianza multivariada, una función de puntuación y la generación de una variable binaria de adherencia al comportamiento de salud. Con estos datos, se entrenó un modelo *logit* para estimar probabilidades de éxito conductual y *contextual bandit* para optimizar la asignación de metas en actividad física.

En la misma población simulada y bajo el mismo protocolo, la política aprendida logra una mayor proporción de adherencia al comportamiento de salud que una asignación puramente aleatoria, especialmente en las primeras interacciones. Estos resultados muestran que los datos sintéticos son útiles para calibrar políticas iniciales y no perder información valiosa de los usuarios al inicio, acelerando la personalización en aplicaciones *mHealth*.

Abstract

This research addresses the mitigation of the cold start problem in digital health recommender systems through the generation of synthetic data informed by psychological theory. In reinforcement learning, the cold start problem arises when the system must make decisions without prior user history; an initial dataset is therefore required to calibrate value estimates and decision policies, preventing early interactions from resulting in low adherence to health-related behaviors.

The objective of this study is to develop and evaluate a synthetic user simulator parameterized with psychological constructs from the Integrated Behavior Change Model and sociodemographic variables. The proposed methodology incorporates a multivariate covariance matrix, a scoring function, and the generation of a binary health-behavior adherence variable. Using these data, a logistic regression model was trained to estimate probabilities of behavioral success, followed by a contextual bandit algorithm to optimize the assignment of physical activity goals.

Within the same simulated population and under a common evaluation protocol, the learned policy achieved a higher proportion of health-behavior adherence than a purely random assignment, particularly during early interactions. These results indicate that synthetic data can effectively support the calibration of initial decision policies, reduce information loss at the beginning of user engagement, and accelerate personalization in mHealth applications.

Contenido

Listado de tablas

Listado de figuras

1. Introducción

1.1 Introducción

La promoción de estilos de vida saludables mediante tecnologías digitales ha adquirido una relevancia creciente debido al impacto comprobado de la inactividad física en la carga global de enfermedad (*global burden of disease*) y en la mortalidad prematura (OMS, 2022; Guthold et al., 2018). En la literatura reciente, el aprendizaje por refuerzo ha emergido como un enfoque adecuado para seleccionar recomendaciones en base al estado del usuario, equilibrando exploración y explotación para favorecer la personalización dinámica (Athey & Imbens, 2019; Guan et al., 2022).

Un desafío central en este tipo de sistemas es el *problema de arranque en frío* (cold start problem), que ocurre cuando el algoritmo debe tomar decisiones sin contar con historial previo del usuario. En estas condiciones, el sistema carece de información suficiente para estimar qué tipo de recomendación podría favorecer la adherencia, lo que puede derivar en metas mal calibradas, experiencias iniciales negativas y una menor probabilidad de mantener el comportamiento saludable. Este fenómeno es especialmente crítico en aplicaciones de salud digital, donde las primeras interacciones suelen determinar el nivel de compromiso y participación futura.

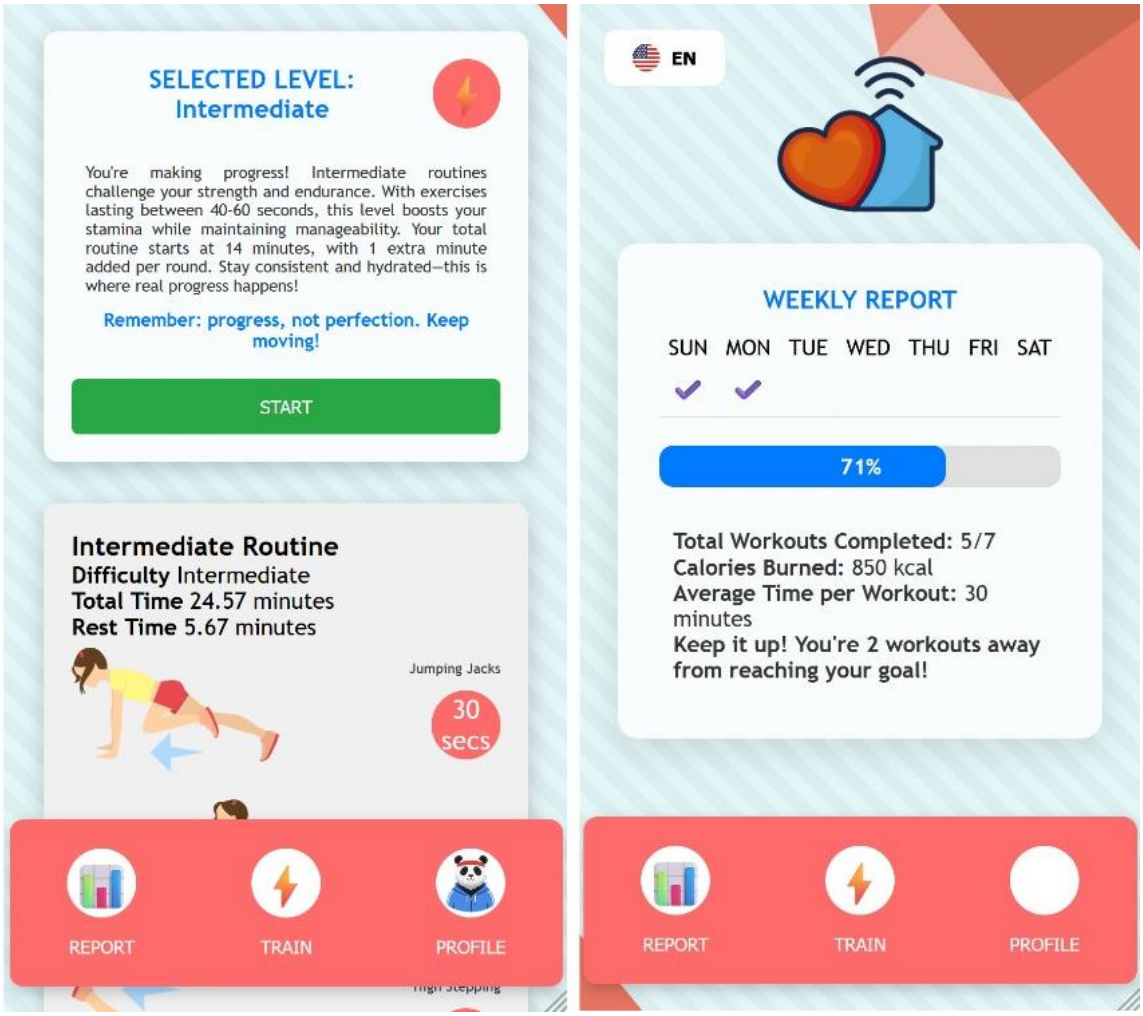
En este contexto, la generación de datos sintéticos se plantea como una herramienta metodológica clave para superar la ausencia de información en etapas tempranas de implementación. Este enfoque permite crear poblaciones virtuales que replican relaciones psicológicas y sociodemográficas documentadas empíricamente, ofreciendo un entorno seguro y escalable para entrenar algoritmos de recomendación sin exponer a usuarios reales a recomendaciones inadecuadas. La evidencia reciente indica que los modelos basados en simulaciones informadas por teoría son especialmente útiles en intervenciones de salud móvil (mobile health, *mHealth* en inglés), facilitan, la personalización inicial sin requerir grandes volúmenes de datos reales.

Las aplicaciones móviles ofrecen canales de intervención accesibles, escalables y personalizables, capaces de adaptar contenidos y metas a las necesidades de cada persona.

Esta memoria se centra en un proyecto Fondecyt, “Tailored goals and individual choice: A field experiment on automated commitment devices for health *behavior*”, orientado a diseñar una aplicación que utilice algoritmos de aprendizaje automático con foco en la personalización para proponer metas de actividad física ajustadas al perfil de cada usuario enfrentando, desde el inicio los desafíos de disponibilidad de datos en etapas tempranas de desarrollo. Se realizan simulaciones para generar bases de datos sintéticas y controladas, sobre las cuales entrenar y validar prototipos algorítmicos antes de su despliegue con participantes reales.

La aplicación *Apptivate*, desarrollada en el marco del mismo proyecto Fondecyt, constituye el entorno experimental del estudio. Los usuarios completan rutinas diarias agrupadas en ciclos de tres días consecutivos, donde alcanzar las metas establecidas equivale a una recompensa conductual o adherencia al comportamiento de salud. La app ajusta el nivel de dificultad de las metas (baja, media o alta) según el desempeño previo del usuario permitiendo, personalizar la experiencia y evaluar la respuesta ante distintos grados de exigencia. Este diseño experimental basado en: ciclos, rutinas y niveles adaptativos sirve como base para modelar la recompensa utilizada en la simulación y comparar el desempeño del algoritmo con una asignación aleatoria.

Figura 1.1 Interfaz principal de la aplicación *Apptivate*.



Fuente: Capturas de la aplicación "Apptivate"

El objetivo de esta memoria es evaluar, con datos simulados basados en evidencia, el desempeño de algoritmos capaces de recomendar metas personalizadas maximizando, la adherencia esperada. Según la literatura reciente, el aprendizaje por refuerzo y en particular los *contextual bandits* ha emergido como un enfoque correcto para seleccionar recomendaciones en base al estado del usuario gestionando, el equilibrio entre exploración y explotación y permitiendo una personalización dinámica. Sin embargo, su entrenamiento inicial requeriría pilotos intensivos con usuarios, con altos costos y complejidad logística y detona el

conocido problema de *arranque en frío* (CSP, en inglés): sin datos previos, la personalización es limitada justo en las primeras interacciones, momento crítico para sostener el compromiso del usuario.

La generación de datos sintéticos informados por teoría y evidencia empírica surge entonces como una solución viable para entrenar políticas y mitigar el CSP antes de los ensayos con personas. En términos prácticos, el presente estudio propone y presenta una metodología estructurada que parametriza, a partir de la evidencia empírica disponible, una población virtual mediante un vector de medias y una matriz de varianza-covarianza ajustada para garantizar validez estadística, simula variables psicológicas clave del cambio conductual basadas en el modelo IBC (Modelo Integrado de Cambio de Comportamiento, por sus siglas en inglés) junto con atributos sociodemográficos define, una función de puntuación y umbrales de logro coherentes con la dificultad de la meta, y finalmente entrena y compara políticas de recomendación, una basada en asignación aleatoria y otra aprendida, dentro de un esquema fuera de línea. Esta estrategia permite evaluar hipótesis y afinar decisiones de personalización con bajo riesgo y alto control, estableciendo una base sólida para posteriores pilotos con usuarios reales.

1.2 Antecedentes y contexto general

Los algoritmos de aprendizaje automático enfocados en personalización requieren conjuntos de datos diversos y representativos para validarse y entrenarse en fases iniciales y ofrecer una experiencia de usuario consistente. En fases tempranas, cuando el sistema aún no ha interactuado con suficientes usuarios, esta condición rara vez se cumple.

En este proyecto se desarrollaron simulaciones que generan respuestas de usuarios sintéticos, es decir, perfiles virtuales que imitan el comportamiento estadístico y psicológico de personas reales, construidos a partir de variables observables y relaciones teóricas. Estos usuarios se definen en función de rasgos sociodemográficos (género, ingreso, empleo, edad) y determinantes psicológicos del cambio conductual, con el objetivo de entrenar y evaluar políticas de recomendación.

En este contexto, los algoritmos de personalización se entienden como mecanismos de decisión que, a partir del perfil del usuario, determinan qué tipo de recomendación o meta resulta más adecuada para maximizar la probabilidad de adherencia. Una política de personalización corresponde, en este marco, a una regla de asignación que mapea las características del usuario (contexto) hacia una acción específica, como el nivel de dificultad de la meta, pudiendo ser fija (por ejemplo, aleatoria) o adaptativa, aprendida a partir de los datos.

Adicionalmente, la decisión de utilizar un simulador guiado por teoría responde a la necesidad de evitar perfiles sintéticos construidos únicamente desde correlaciones estadísticas superficiales. Al incorporar el Modelo Integrado de Cambio de Comportamiento (IBC, siglas en inglés), la simulación incorpora determinantes motivacionales, cognitivos y sociales permitiendo que, los usuarios generados exhiban patrones plausibles de adherencia consistentes con literatura en psicología de la actividad física. Esto mejora la validez externa del conjunto de datos sintéticos y fortalece su utilidad para entrenar algoritmos de personalización.

La arquitectura resultante genera bases de datos sintéticas reproducibles y compatibles con herramientas de análisis y modelos de recomendación. En este estudio se generó una población virtual de 2000 individuos, cada uno con un perfil psicológico discretizado, un conjunto de características sociodemográficas relevantes y una acción inicial asociada a un nivel de meta. Además, se incorporó un indicador binario de comportamiento que refleja, la adherencia lograda o no definido, a partir de una función de puntuación y umbrales diferenciados según la dificultad de la meta. Este diseño garantiza una relación coherente entre: las condiciones del usuario, la exigencia de la recomendación y el resultado observado, ofreciendo, un entorno controlado para comparar distintas “políticas de personalización”.

El resultado binario de comportamiento (éxito o no éxito en la meta) garantiza que el conjunto de datos conserve una relación coherente entre el nivel de exigencia de la meta y la probabilidad de logro, de manera que la recompensa empleada para el entrenamiento se base en un criterio transparente, interpretable y

metodológicamente consistente. En paralelo, el contexto conceptual que fundamenta esta construcción se sustenta en dos principios: el uso de datos sintéticos informados por evidencia empírica para reducir costos y acelerar la iteración, y la pertinencia técnica de los *contextual bandits* como enfoque para la selección de metas personalizadas en función del estado del usuario. Bajo esta perspectiva, la simulación actúa como un campo de pruebas controlado para validar hipótesis de personalización, como la reasignación de metas excesivamente exigentes hacia niveles intermedios en perfiles con baja autoeficacia, y permite estimar el efecto comparativo entre políticas aprendidas y asignaciones aleatorias antes de avanzar a pilotos con usuarios reales.

1.3 Objetivos

El presente estudio se estructura en torno a un objetivo general y un conjunto de objetivos específicos, orientados a abordar el problema del *arranque en frío* en sistemas de recomendación personalizados aplicados a la promoción de la actividad física, mediante el uso de datos sintéticos generados a partir de modelos teóricos y evidencia empírica.

Objetivo general: Diseñar, implementar y evaluar un marco de simulación de datos sintéticos, basado en el modelo IBC y metadatos empíricos, que permita entrenar y comparar algoritmos de personalización de metas de actividad física en un entorno fuera de línea, contribuyendo a mitigar el CSP en aplicaciones *mHealth*.

Objetivos específicos:

- Recopilar y parametrizar metadatos empíricos provenientes de estudios previos basados en el modelo IBC y teorías afines, a fin de construir un vector de medias y una matriz de varianza–covarianza coherentes con relaciones psicológicas documentadas.
- Desarrollar un simulador de usuarios sintéticos que combine variables psicológicas (IBC) y sociodemográficas integrando, correlaciones realistas, sesgo controlado (vector α) y discretización en escalas Likert, garantizando que

los datos generados sean apropiados y compatibles con librerías de aprendizaje automático.

- Definir y calibrar una función de puntuación (score) y umbrales de logro que traduzcan los atributos individuales y la dificultad de la meta asignada con una probabilidad de adherencia generando así, la señal de recompensa necesaria para entrenar algoritmos de refuerzo.
- Implementar y evaluar el entrenamiento del modelo de recomendado según enfoques contrastados tales como:
  (1) datos reales en base a un estudio piloto.
  (2) utilizando la base sintética generada.

2. Marco teórico y revisión de literatura

2.1 El problema del *arranque en frío* (CSP)

Uno de los principales desafíos en el desarrollo de sistemas de recomendación personalizados es el denominado problema del CSP que se refiere a la dificultad que enfrentan algoritmos de recomendación para ofrecer recomendaciones adecuadas cuando no se dispone de información suficiente sobre usuarios o ítems a recomendar (Singh & Singh, 2024). Este problema se presenta en múltiples dominios, desde el comercio electrónico hasta los sistemas de entretenimiento y adquiere especial relevancia en el ámbito de la salud digital, donde las primeras interacciones entre usuario y aplicación resultan determinantes para mantener la adherencia al programa y evitar el abandono temprano así, como el no uso de la aplicación.

En términos generales, el CSP puede manifestarse en tres niveles. El primero corresponde al CSP de usuario, que se produce cuando una persona comienza a interactuar con el sistema y no existen datos históricos que permitan personalizar sus recomendaciones. El segundo nivel es el CSP de ítem, que ocurre cuando un nuevo contenido, producto o meta, por ejemplo un nuevo tipo de desafío de actividad física es introducido en la plataforma y el sistema carece de referencias previas sobre su aceptación o efectividad. Finalmente, el CSP de sistema se produce cuando una aplicación recién lanzada no dispone de datos suficientes en su conjunto, lo que impide que los algoritmos funcionen adecuadamente desde el inicio (Park & Chu, 2009).

En el contexto específico de aplicaciones móviles para la promoción de la actividad física, el CSP de usuario adquiere especial relevancia cuando un usuario descarga la aplicación y recibe recomendaciones iniciales no ajustadas a sus capacidades o motivaciones, aumentando considerablemente la probabilidad de abandono.

Estudios han demostrado que la experiencia temprana del usuario en sistemas *mHealth* es clave para establecer un hábito de uso sostenido y mantener el compromiso con la intervención (Mohr, Burns, Schueller, Clarke, & Klinkman, 2014; Perski, Blandford, West, & Michie, 2017; Direito, Carraça, Rawstorn, Whittaker, & Maddison, 2021).

Una recomendación inicial demasiado exigente (por ejemplo, metas irreales de pasos diarios o incremento súbito de intensidad) o demasiado fácil (como objetivos triviales sin desafío ni refuerzo) puede comprometer la percepción de utilidad y la motivación intrínseca del usuario afectando, la adherencia el uso continuado de la aplicación.

La literatura relacionada ha propuesto diversas estrategias para mitigar el CSP, entre las que se encuentran:

- Métodos de contenido (content-based) aprovechan la información declarada por el usuario (por ejemplo, edad, género o nivel de ejercicio actual) para ofrecer recomendaciones iniciales. Aunque útiles, estos métodos suelen ser limitados en dominios complejos, ya que no siempre capturan las motivaciones o barreras psicológicas que influyen en la adherencia.
- Métodos de transferencia utilizan información obtenida de usuarios previos o de contextos similares para inferir recomendaciones iniciales. Requiere una base histórica representativa, no siempre es posible en proyectos de investigación en etapas tempranas de desarrollo (Kan, Kim, Kim, & Lee, 2023).
- Simulación y el uso de datos sintéticos es una estrategia emergente y consiste en crear poblaciones virtuales de usuarios con atributos y conductas verosímiles, destinadas a entrenar algoritmos antes de su aplicación, en entornos reales (Tsao et al., 2023). Esta estrategia resulta especialmente atractiva en el ámbito de la salud digital, permite preentrenar políticas personalizadas sin exponer a usuarios reales a riesgos derivados de recomendaciones inadecuadas.

En el caso de sistemas de promoción de actividad física, el CSP constituye una limitación crítica porque las intervenciones deben ser personalizadas desde el inicio para sostener la motivación. En esta memoria de título se propone como solución el uso de simulación de datos sintéticos basados en evidencia para, pre-entrenar algoritmos de tipo *contextual bandit*, que permitiría iniciar el sistema con una política informada en lugar de una asignación aleatoria. Esta aproximación no elimina el CSP en sentido estricto, lo mitiga significativamente proporcionando, un “punto de partida” sólido para la personalización progresiva.

2.2 Datos sintéticos como solución

El uso de datos sintéticos ha emergido como una estrategia prometedora para superar las limitaciones derivadas de la escasez o falta de acceso a información en contextos de entrenamiento algorítmico. Los datos sintéticos son registros artificiales generados a partir de modelos matemáticos, estadísticos o de simulación para imitar patrones de datos reales preservando su estructura estadística y relacional (Tsao, Li, Wang, & Zhang, 2023). A diferencia de los datos ficticios arbitrarios, los datos sintéticos buscan mantener coherencia interna y verosimilitud siendo un recurso útil para el entrenamiento, validación y evaluación de modelos de aprendizaje automático en condiciones controladas.

En el ámbito de la salud digital, la generación de datos sintéticos tiene un valor agregado al permitir: probar hipótesis, calibrar algoritmos y planificar intervenciones sin necesidad de exponer a usuarios reales en etapas tempranas del desarrollo. Esto reduce costos y riesgos sino que también, contribuye a abordar restricciones éticas y de privacidad dado que, los registros generados no corresponden a individuos reales. Revisiones recientes han señalado que los datos sintéticos serían un componente fundamental para: habilitar ecosistemas de innovación en salud digital, facilitar la interoperabilidad y el pre-entrenamiento de modelos en ausencia de bases clínicas extensas (Pilani, Gupta, & Kathuria, 2021).

En relación con el CSP, los datos sintéticos ofrecen solución concreta: al crear una población virtual de usuarios con atributos sociodemográficos y psicológicos realistas, es posible entrenar algoritmos de personalización antes de disponer de

registros reales de interacción. Por ejemplo, Pilani et al. (2021) demostraron, que los algoritmos de tipo *contextual bandit* se benefician del pre-entrenamiento en entornos simulados obteniendo, un rendimiento superior en fases iniciales de despliegue en comparación con modelos entrenados exclusivamente en línea. De manera similar, (Kan, Kim, Kim, & Lee, 2023) mostraron, que la simulación retrospectiva de datos "observacionales" permite, construir políticas conductuales iniciales que reducen la dependencia de extensos periodos de exploración aleatoria.

No obstante, no todos los enfoques de simulación son equivalentes. Mientras que algunos estudios han generado datos artificiales a partir de distribuciones estadísticas simples (ej. normal multivariada sin restricciones conductuales), el enfoque propuesto se orienta hacia datos sintéticos informados por teoría conductual. En particular, el Modelo IBC se utiliza como marco estructural para garantizar que las relaciones entre variables psicológicas simuladas, reflejen dinámicas empíricamente documentadas (Hagger & Chatzisarantis, 2014). Esta diferencia es crucial: al incorporar correlaciones y sesgos fundamentados en evidencia, los datos sintéticos dejan de ser "datos artificiales" y pasan a representar una población coherente y permiten que los algoritmos interactúen de forma significativa.

La utilidad es que, al entrenar un algoritmo de recomendación sobre datos sintéticos, se obtiene una política inicial pre-entrenada con patrones plausibles de comportamiento. Así, al interactuar con usuarios reales, el algoritmo no parte de un estado "en blanco" ni depende de exploración aleatoria excesiva sino que, dispone una base de conocimiento que permite personalizar con mayor eficacia desde el inicio. Así, los datos sintéticos se convierten en un puente metodológico entre la teoría psicológica y la implementación algorítmica, mitigando el CSP y mejorando la experiencia del usuario en las etapas críticas de adopción de la aplicación.

En términos éticos, los datos sintéticos aportan una capa adicional de resguardo, eliminan la posibilidad de identificar a personas reales incluso, de manera indirecta. Esto es especialmente relevante en aplicaciones de salud digital, donde los datos sensibles pueden, exponer información íntima de los usuarios y riesgos de

privacidad difíciles de mitigar en etapas tempranas del desarrollo. El uso de simulaciones permite, avanzar en: investigación, experimentación y calibración de algoritmos sin comprometer la confidencialidad, así, los datos sintéticos serían una alternativa metodológica segura y alineada con buenas prácticas en *machine learning* en salud.

2.3 Limitaciones de enfoques actuales

Los datos sintéticos constituyen una alternativa prometedora para enfrentar el problema del CSP, la literatura especializada ha identificado diversas limitaciones en los enfoques actualmente disponibles. Estas limitaciones pueden ser: metodológicas, teóricas y prácticas.

a) Limitaciones metodológicas

En numerosos estudios, la generación de datos sintéticos se basa en procedimientos estadísticos simples, tales como distribuciones normales multivariadas con correlaciones o el uso de *bootstrapping* sobre muestras reducidas (Pilani et al., 2021). Este tipo de simulaciones permiten, obtener bases de datos con estructura básica pero, tienden a carecer de complejidad y variabilidad contextual, lo que restringe su capacidad de representar adecuadamente, la diversidad de usuarios en entornos de salud digital. Además, cuando los parámetros se calibran con información limitada o no representativa, los datos resultantes pueden introducir sesgos sistemáticos que distorsionen, el entrenamiento de los algoritmos.

Otro problema metodológico recurrente es la falta de garantías estadísticas de validez. En algunos casos, las matrices de correlación empleadas no son positivas semidefinidas, que conduce a inconsistencias matemáticas en las simulaciones. Para corregir esto, los autores suelen aplicar ajustes arbitrarios (ej. perturbaciones diagonales o truncamiento de valores propios), que puede alterar la fidelidad de las relaciones entre variables (Kan, Kim, Kim, & Lee, 2023). Estas correcciones, aunque necesarias limitan, la interpretabilidad y reproducibilidad de los datos generados.

b) Limitaciones teóricas

Una de las críticas más importantes es que muchos enfoques de simulación carecen de fundamentos teóricos sólidos que, guíen la selección de variables y la definición de relaciones entre estás. En varios estudios, los perfiles sintéticos se construyen únicamente a partir de correlaciones estadísticas observadas en un conjunto reducido de variables demográficas o de uso, sin considerar, factores psicológicos, motivacionales o sociales que inciden de manera determinante, en la adherencia conductual (Mohr, Burns, Schueller, Clarke, & Klinkman, 2014).

La reducción provoca que, los algoritmos entrenados sobre dichos datos aprendan patrones superficiales, desconectados de los mecanismos que efectivamente explican el comportamiento humano. En consecuencia, la política aprendida puede ser eficiente en un entorno artificial pero, ineficaz al aplicarse en contextos reales. La ausencia de modelos conductuales validados como el IBC en la construcción de datos sintéticos se traduce, por tanto, en un déficit de realismo y de capacidad predictiva a largo plazo.

c) Limitaciones prácticas

Desde el punto de vista práctico, un desafío frecuente es la transferencia de los resultados obtenidos en entornos simulados a escenarios reales. Aunque los datos sintéticos permiten entrenar y comparar algoritmos, en condiciones controladas existe, el riesgo que las políticas aprendidas no generalicen adecuadamente a poblaciones reales si la simulación no ha incorporado, variabilidad suficiente en sus parámetros (Tsao, Li, Wang, & Zhang, 2023).

Adicionalmente, el desarrollo de simuladores detallados requiere esfuerzos significativos de programación y validación, que podría convertirse en una barrera para equipos que no siempre disponen de expertos en estadística computacional o ciencia de datos. En la práctica, muchos proyectos optan por simulaciones rudimentarias, sacrificando precisión a cambio de viabilidad técnica.

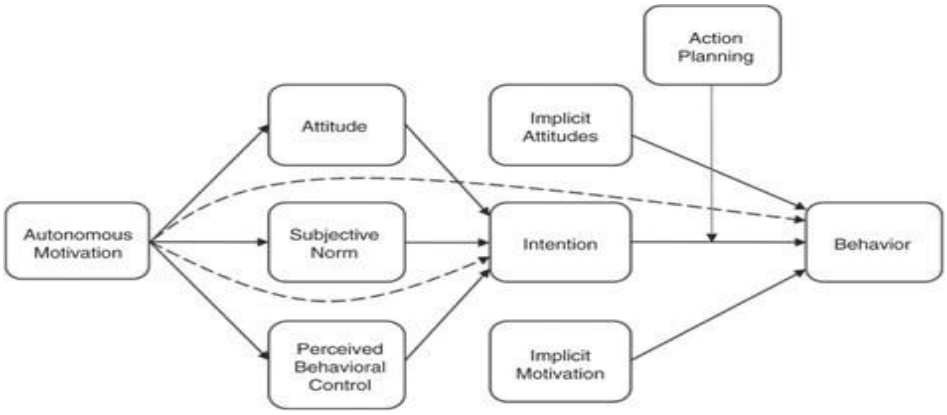
Por último, el uso de datos sintéticos no elimina por completo los riesgos éticos: aunque no se refieren a individuos reales, los resultados obtenidos en simulación podrían inducir a falsas expectativas si se presentan como equivalentes

a pruebas empíricas. La literatura enfatiza que las simulaciones deben considerarse una etapa preparatoria y complementaria y nunca un sustituto definitivo de validación con datos reales (Kan, Kim, Kim, & Lee, 2023).

2.4 Modelo IBC como base teórica

Una de las principales innovaciones de este método frente a enfoques previos de simulación es la incorporación de un marco teórico conductual, robusto para guiar la construcción de la población sintética. En particular, se adopta el Modelo Integrado de Cambio de Comportamiento (Hagger & Chatzisarantis, 2014), propuesto por Hagger y Chatzisarantis (2014), como eje conceptual para definir las variables psicológicas críticas, sus interrelaciones e influencia en la adopción de comportamientos saludables, como la actividad física (Figura 2.1).

Figura 2.1 Modelo de Cambio Conductual Integrado (Hagger y Chatzisarantis, 2014)



Fuente: *Hagger y Chatzisarantis, 2014*

El Modelo Integrado de Cambio de Comportamiento (IBC) considera elementos de tres teorías ampliamente contrastadas en la psicología de la salud. Incorpora la Teoría de la Conducta Planificada (TPB) de Ajzen, que enfatiza el papel de la intención, la actitud, las normas subjetivas y el control percibido en la predicción del comportamiento. En segundo lugar, integra la Teoría de la Autodeterminación (SDT) de Deci y Ryan, que distingue entre motivaciones autónomas (intrínsecas) y motivaciones controladas (extrínsecas) como determinantes fundamentales de la autorregulación sostenida. Finalmente, el modelo incluye: Procesos de Acción en Salud (HAPA), que reconoce la existencia de fases motivacionales y de autorregulación en la transición de intención a la acción destacando, la relevancia de la autoeficacia en dicho proceso.

Al unificar estos enfoques, el IBC ofrece un marco multifásico del cambio conductual con tres etapas principales. La fase pre-motivacional cuando factores sociales y ambientales, como el apoyo a la autonomía o las normas sociales, configuran las percepciones iniciales del individuo. La fase motivacional es la formación de actitudes, motivaciones autónomas e intenciones que determinan la predisposición al cambio. Finalmente, la fase post-motivacional es la ejecución del comportamiento, sostenida en gran medida por la autoeficacia y el control percibido (Hagger & Chatzisarantis, 2014).

2.4.1 IBC como fundamento para la simulación

La elección del modelo IBC como base para la simulación responde a dos razones principales: Primero: Permite seleccionar *constructos* psicológicos específicos, como la motivación autónoma, la actitud, la intención, las normas subjetivas, la autoeficacia y el control conductual percibido, con evidencia empírica sólida que los respalda como predictores relevantes de la actividad física (Ferrada Torres, Nguyen, Caro, & Lipman, 2024). Estos *constructos* fueron transformados en variables latentes dentro de la simulación, garantizando que la población sintética resultante no solo tuviera coherencia estadística, sino también sentido psicológico. Segundo, el modelo facilita la parametrización de las relaciones entre variables con

matrices de correlación y covarianza extraídas de estudios empíricos, lo que permite que los usuarios simulados reflejen patrones realistas de interacción entre motivación, intención y conducta. IBC opera como un puente entre teoría y simulación, asegurando que, los datos generados no sean arbitrarios sino que, se encuentren alineados con principios conductuales documentados en la literatura. Esto refuerza tanto, la plausibilidad externa de los resultados como la coherencia interna de los algoritmos entrenados con dichos datos.

2.4.2 Ventajas del IBC frente a otros modelos

Aunque existen otros modelos aplicables al cambio conductual (como el COM-B o el HAPA en su versión original), el IBC ofrece una ventaja particular: integración de múltiples marcos en una estructura única y sencilla. En lugar de centrarse exclusivamente en intenciones (como la TPB) o en las motivaciones intrínsecas (como la SDT), el IBC reconoce que el comportamiento saludable es producto de un entramado de factores sociales, motivacionales que interactúan en fases sucesivas. Esta naturaleza integradora lo convierte en un modelo idóneo para sustentar una simulación y capturar la diversidad de trayectorias posibles en usuarios reales.

A diferencia de modelos unidimensionales por ejemplo, enfoques centrados exclusivamente en la formación de intención (como la Teoría del Comportamiento Planificado) o en la calidad motivacional (como la Teoría de la Autodeterminación) el IBC ofrece una estructura integradora capaz de capturar la complejidad del cambio conductual. Esta perspectiva multidimensional es útil en procesos de simulación pues permite, construir perfiles psicológicos donde actitud: normas subjetivas, motivación y percepción de control interactúan de forma coherente. Al preservar estas dinámicas, los algoritmos entrenados sobre datos sintéticos aprenden patrones más realistas y cercanos a los observados en poblaciones reales.

2.4.3 Aplicación en el presente estudio

En este estudio, los *constructos* derivados del IBC fueron representados como un conjunto de siete variables psicológicas que conformaron el núcleo de la

simulación: apoyo percibido de autonomía, motivación autónoma, actitud, normas subjetivas, intención de cambio, autoeficacia y control conductual percibido. Estos ítems fueron parametrizados a través de distribuciones multivariadas correlacionadas, incorporando además ajustes de sesgo (vector α) para emular patrones de respuesta asimétricos. La inclusión de estos *constructos* permitió que la política aprendida por los algoritmos no se basara únicamente en atributos demográficos, sino también en determinantes motivacionales estrechamente vinculados al éxito de la intervención.

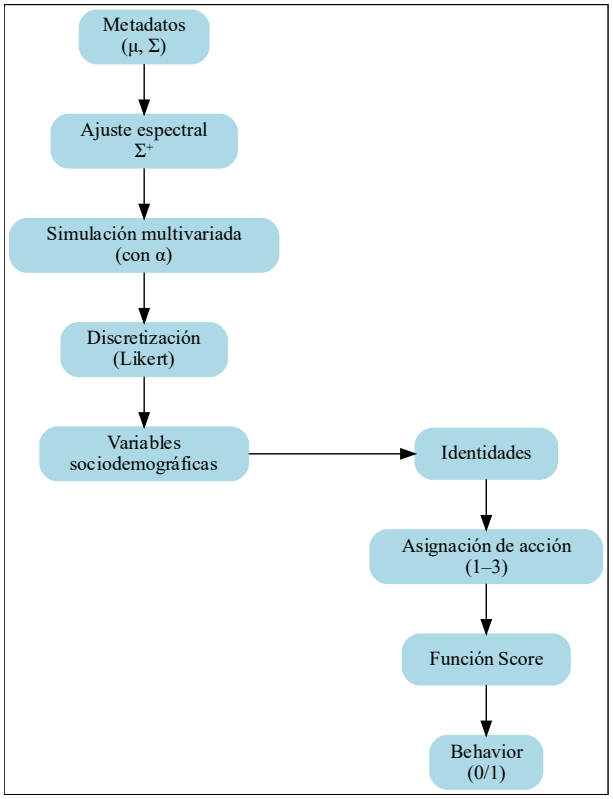
Así, el modelo IBC no solo proporcionó un mapa conceptual para la construcción de usuarios sintéticos también, sirvió como criterio de validez teórica para evaluar si los patrones emergentes en la política óptima aprendida , entendida como aquella que maximiza la recompensa esperada asociada a la adherencia simulada, son coherentes con el conocimiento psicológico existente. Así, la simulación no sería un ejercicio estadístico abstracto y se convierte en una recreación informada del proceso de cambio conductual útil tanto, para la ingeniería de algoritmos como para la reflexión académica sobre la personalización en salud digital.

3. Metodología

La metodología presente se orientó a construir una población sintética de usuarios, parametrizada con metadatos empíricos y teoría psicológica (IBC, TPB, SDT, en inglés), para entrenar y evaluar algoritmos de personalización de metas de actividad física, en un entorno controlado.

El proceso metodológico puede dividirse en tres grandes fases. En primer lugar. La extracción y parametrización de metadatos necesarios para la construcción del vector de medias y la matriz de covarianza, que sustentan la generación de datos correlacionados de forma estadísticamente coherente. En segundo lugar, el diseño del simulador de datos sintéticos, integrando variables psicológicas, sociodemográficas y conductuales. Finalmente, se definió la función de puntuación (score) y los umbrales de logros, que permiten obtener una variable de recompensa binaria (*behavior*) utilizada en el entrenamiento de algoritmos de aprendizaje por refuerzo.

Figura 3.1 Diagrama de bloques del proceso completo



Fuente: Elaboración propia.

3.1 Extracción de metadatos y construcción de la matriz de covarianza

La primera etapa es la identificación y extracción de metadatos provenientes de literatura científica relevante en el ámbito del cambio de comportamiento en salud. Se priorizaron estudios empíricos que aplicaran el Modelo Integrado de Cambio de Comportamiento (IBC) o modelos afines como la Teoría de la Conducta Planificada (TPB) y la Teoría de la Autodeterminación (SDT).

Se recopilaron medias y desviaciones estándar, así como medidas de tendencia central y matrices de correlación, con el propósito de establecer parámetros estadísticos consistentes en simulación.

Se consideraron los siguientes *constructos* psicológicos: el apoyo percibido de autonomía, la motivación autónoma, la actitud hacia el comportamiento, las normas subjetivas, la intención de cambio, el control conductual percibido y la autoeficacia. Estos *constructos* constituyen la base teórica que permite modelar la variabilidad individual en los patrones de decisión y adherencia dentro del simulador.

Formalmente, para cada par de variables i i y j j:

$$\Sigma_{ij} = \rho_{ij} \cdot \sigma_i \cdot \sigma_j$$

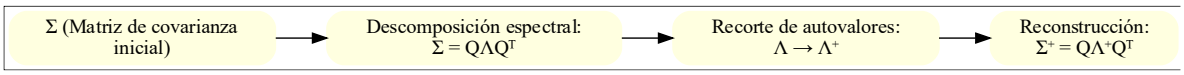
donde ρij es la correlación empírica reportada y σi ,σj las desviaciones estándar de cada variable (Johnson & Wichern, 2007).

Cuando la matriz inicial no resultó semidefinida positiva (condición necesaria para la generación multivariada), se aplicó un ajuste espectral:

$$\Sigma^{+} = Q\Lambda^{+}Q^{T}$$

donde Q es la matriz de vectores y Λ + la matriz diagonal de autovalores corregidos.

Figura 3.2 Esquema de ajuste espectral aplicado a la matriz de covarianza



Fuente: elaboración propia

Este procedimiento garantizó estabilidad estadística y preservó las relaciones teóricas principales entre *constructos* psicológicos. Esta configuración aseguró una correlación coherente entre dimensiones psicológicas, replicando hallazgos empíricos sobre las interrelaciones entre motivación: intención, actitud y percepción de control en la adopción de actividad física.

3.2 Diseño del simulador de datos sintéticos

La implementación del simulador se realizó en Python utilizando las librerías numpy, pandas y scipy. Se aplicaron técnicas de: muestreo multivariado correlacionado, control de sesgo mediante asimetrías dirigidas y generación de variables categóricas ponderadas, garantizando datos consistentes y controlables. La simulación se estructuró en módulos: Simulacion_principal.py y calculo_*behavior*.py, los cuales encapsulan funciones clave del generador, incluyendo: una rutina de normal, multivariada sesgada que preserva correlaciones entre *constructos*; un procedimiento de discretización forzada para transformar variables latentes en escalas Likert, y un bloque sociodemográfico que define explícitamente las distribuciones de: género, ingreso, empleo y nivel de ejercicio.

La simulación se implementó en Python 3.10.5, utilizando librerías como numpy, pandas, scipy y unidecode. El código se organizó, en dos módulos principales:

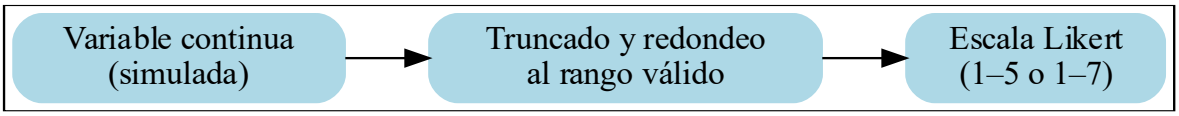
- **Simulacion_principal.py:** encargado de la generación global de perfiles y la integración de atributos.

- **calculo_*behavior*.py:** orientado al cálculo de la variable de resultado (*behavior*) a partir de la función de puntuación.

Entre las funciones clave desarrolladas se destacan:

- Distribución multivariada sesgada: a través de la función simulate_skewed_mvn(), generando respuestas continuas correlacionadas, incorporando un vector de sesgo (α) que introduce asimetrías univariadas. Esto permitió representar patrones psicológicos más realistas (ej. mayor proporción de usuarios con alta autoeficacia).
- Discretización forzada: las variables continuas simuladas fueron truncadas a escalas ordinales comunes en encuestas reales (Likert de 1 a 5 o de 1 a 7), manteniendo consistencia con instrumentos empíricamente utilizados.

Figura 3.3 Proceso de discretización forzada



Fuente: elaboración propia.

- Simulación sociodemográfica: generando variables como género, edad, ingreso y empleo con distribuciones realistas, basadas en estadísticas poblacionales de adultos jóvenes. Cada categoría se ponderó según frecuencias observadas.
- Generación de identidades: se incorporó un módulo que asignó automáticamente nombres, correos electrónicos y usernames culturalmente plausibles para el contexto chileno, con diferenciación por género.

3.3 Definición de la función de puntuación y umbrales de logro

El simulador incluyó un sistema de asignación de metas de actividad física, representadas como:

$$action \in \{1 \text{ (fácil)}, 2 \text{ (media)}, 3 \text{ (difícil)}\}$$

El puntaje individual (score) se calculó a partir de tres componentes principales:

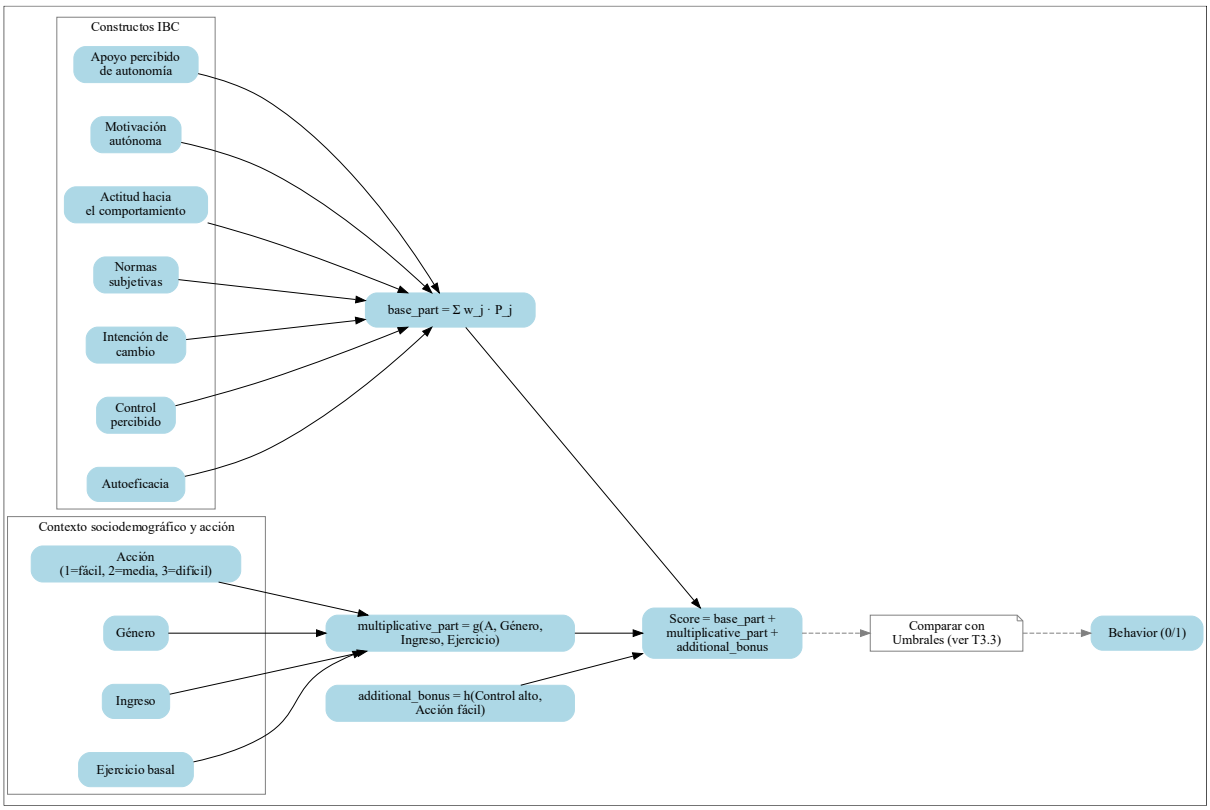
- **base_part:** suma ponderada de los puntajes de los siete *constructos* psicológicos IBC.
- **multiplicative_part:** ajuste según acción, ingreso, nivel basal de ejercicio y género.
- **additional_bonus:** incentivo adicional para usuarios con alto control percibido y metas fáciles.

$$score_i = base\_part_i + multiplicative\_part_i + additional\_bonus_i$$

Finalmente, el valor de **score** se comparó con un umbral dependiente de la dificultad de la meta, generando la variable binaria *behavior*, que indica éxito (1) o fracaso (0) y sirve como señal de recompensa en el entrenamiento del algoritmo *contextual bandit.*

Figura 3.4 Esquema de la función de puntuación (score) y generación de la variable behavior

.



Fuente: elaboración propia.

La Figura 3.4 presenta el esquema de la función de puntuación (score) utilizada en el simulador y el proceso de generación de la variable binaria *behavior*. En el diagrama se observa cómo los *constructos* psicológicos del modelo IBC contribuyen a una componente base del puntaje, la cual es posteriormente ajustada mediante factores contextuales y la dificultad de la acción a través de componentes multiplicativas y bonificaciones adicionales. El puntaje total resultante se compara con umbrales específicos según el nivel de dificultad de la meta, determinando así el valor final de *behavior*, que representa el éxito o fracaso conductual y actúa como señal de recompensa para el entrenamiento del algoritmo.

En función de este puntaje, se establecieron umbrales mínimos de logro según la dificultad de la meta:

Tabla 3.1 Umbrales por acción

Acción	Umbral mínimo de score
1 (fácil)	70
2 (media)	73
3 (difícil)	75

Fuente: elaboración propia

Así, la variable binaria *behavior* se definió como:

$$behavior_i = \{1\ si\ score_i \geq umbral(action_i),\ 0\ en\ caso\ contrario\}$$

Esta variable es la señal de recompensa utilizada en el entrenamiento de algoritmos de aprendizaje por refuerzo.

3.4 Evaluación del algoritmo de personalización

La evaluación de algoritmos de personalización constituye una etapa fundamental para determinar la eficacia potencial de las políticas de recomendación de metas de actividad física en la aplicación *mHealth* propuesta. En este trabajo, la evaluación se realizó a cabo utilizando el *dataset* sintético generado a partir del modelo IBC (Capítulo 2), que permitió simular la interacción de usuarios con distintos perfiles psicológicos y sociodemográficos frente a recomendaciones personalizadas de metas.

El propósito general de esta etapa fue probar y comparar técnicas de aprendizaje supervisado y de aprendizaje por refuerzo contextual que permitan, asignar de manera eficiente un nivel de meta de actividad física (acción) en función del perfil del usuario.

Metodológicamente, la evaluación se organizó en tres etapas complementarias: primero, la preparación de datos simulados mediante procesos de preprocesamiento, aleatorización y división en conjuntos de entrenamiento y prueba; segundo, el entrenamiento de un modelo de regresión logística (*logit)* como aproximación inicial para estimar la probabilidad de éxito conductual (behavior); y tercero, la aplicación del algoritmo *contextual bandit* con *Vowpal Wabbit* (VW), empleando las probabilidades estimadas por el modelo *logit* para construir funciones de costo y entrenar la política de decisión óptima.

Este enfoque integrado combinó un modelo interpretable de referencia (*logit*) con un algoritmo especializado en personalización adaptativa (VW), permitiendo analizar de forma comparativa la capacidad de ambos para maximizar la adherencia simulada.

En el marco metodológico, se incorporó la política ε-greedy como mecanismo de selección de acciones dentro del algoritmo de aprendizaje por refuerzo. Esta política permite equilibrar la explotación, seleccionando la acción con mayor recompensa esperada según el modelo actual, y la exploración, introduciendo decisiones aleatorias con una probabilidad ε. Este equilibrio resulta especialmente relevante frente al problema de *arranque en frío*, ya que evita que el algoritmo quede atrapado prematuramente en acciones subóptimas cuando aún no dispone de información suficiente sobre el usuario. A medida que avanza el aprendizaje, el valor de ε decae progresivamente, permitiendo que la política dependa cada vez más de la evidencia acumulada.

Epsilon-greedy

Entradas:

$\varepsilon = 0.2$ y tasa de decaimiento $\mathrm{d} = 0.01$
$x = contexto\ del\ usuario\ (IBC\ )$
$\alpha \in \mathrm{k} = \{1, 2,3\}\ (conjunto\ de\ acciones: 1 - fácil, 2 - medio, 3 - difícil)$
$recompensa\ binaria\ r\ \epsilon\ \{0,1\}$
$\hat{f}_k =$ función que mapea las recompensas para la acción k dado un contexto x

para cada ronda sucesiva t con contexto $x^t\ hacer$:

$Con\ probabilidad\ (1 - \varepsilon)$
$Seleccionar\ acción\ \alpha = argmax_k \hat{f}_k(x^t)$
$De\ lo\ contrario$:
$Seleccionar\ acción\ \alpha\ de\ manera\ uniforme\ al\ azar\ entre\ 1\ y\ k$
$Obtener\ recompensa\ r_a^t, Actualizar\ \{x^t, r_a^t\}$
$Actualizar\ \hat{f}_a\ y\ \varepsilon = \mathrm{d} * \varepsilon$

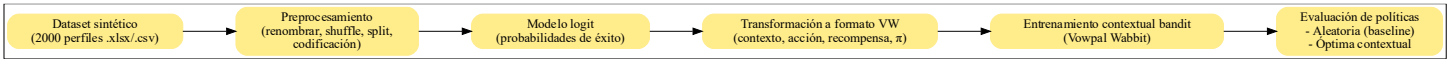
Fin

En el algoritmo ε-greedy, el contexto x^trepresenta el estado del usuario en la ronda t, definido por el vector de *constructos* psicológicos derivados del modelo IBC. En cada iteración, el algoritmo selecciona una acción correspondiente al nivel de dificultad de la meta de actividad física a asignar. Con probabilidad $1 - \varepsilon$, se elige la acción con mayor recompensa esperada según la función estimadora $\hat{f}_k(x^t)$, privilegiando la explotación del conocimiento acumulado. Con probabilidad ε, el algoritmo explora seleccionando una acción de manera aleatoria, lo que permite evaluar alternativas no suficientemente probadas y mitigar el problema de *arranque en frío*.

Tras ejecutar la acción seleccionada, se observa una recompensa binaria $r \in \{0,1\}$que indica el éxito o fracaso conductual del usuario simulado frente a la meta

asignada. Esta información se utiliza para actualizar el estimador correspondiente a la acción ejecutada, incorporando nueva evidencia sobre la relación entre el contexto del usuario y la respuesta conductual. Adicionalmente, el parámetro εse reduce progresivamente mediante un factor de decaimiento d, favoreciendo una transición gradual desde la exploración inicial hacia una política de personalización más estable basada en la evidencia acumulada.

Figura 3.5 Pipeline de evaluación



Fuente: elaboración propia

4. Resultados

4.1 *Dataset* Sintético

El proceso de preparación de datos constituyó un paso crucial para asegurar la calidad y la compatibilidad del *dataset sintético* con los algoritmos de aprendizaje a utilizar. Se usó un archivo en formato .xlsx que contenía 2000 observaciones, cada una correspondiente a un perfil de usuario generado en la simulación.

4.1.1 Preprocesamiento

El preprocesamiento de datos incluyó varias etapas orientadas a garantizar la consistencia y representatividad del conjunto sintético. Primero, se realizó el renombramiento de columnas ajustando, los nombres de las variables para que coincidieran con las convenciones utilizadas en los modelos y así evitar inconsistencias durante la lectura de los archivos.

Posteriormente, se realizó la aleatorización de filas mediante la función sample(frac=1) de pandas, con el propósito de reordenar aleatoriamente el orden de los registros. Esta práctica permite evitar sesgos en la partición entre los conjuntos de entrenamiento y prueba, asegurando que ambos sean representativos de la población sintética total.

A continuación, se efectuó la separación en conjuntos de entrenamiento y prueba, dividiendo el *dataset* en 1300 registros (65%) para entrenamiento y 700 registros (35%) para prueba. Esta proporción se definió siguiendo las recomendaciones establecidas en el ámbito del aprendizaje automático, donde se sugiere destinar entre un 60-80% de los datos al entrenamiento para maximizar la capacidad de ajuste sin comprometer la validez de la evaluación (Goodfellow, Bengio, & Courville, 2016).

Finalmente, se aplicó la codificación de variables categóricas transformando atributos, como: género ingreso y acción (nivel de dificultad de la meta) en valores numéricos mediante codificación ordinal. Este procedimiento, permitió preservar las categorías originales, pero adaptarlas al formato requerido por los algoritmos utilizados en las etapas posteriores de modelado.

La Tabla 4.1 resume de manera esquemática las etapas de preprocesamiento aplicadas al *dataset* sintético, destacando su propósito en el flujo metodológico.

Tabla 4.1 Pasos de preprocesamiento aplicados al dataset sintético

Paso	Descripción	Justificación
1	Renombramiento de columnas	Uniformar nombres y evitar inconsistencias en los algoritmos
2	Aleatorización de filas	Asegurar representatividad de los conjuntos entrenamiento/prueba
3	División entrenamiento–prueba (65/35)	
4	Codificación de variables categóricas	Transformar categorías nominales en valores numéricos compatibles

Fuente: elaboración propia.

4.1.2 Variables utilizadas

Las variables utilizadas en el proceso de entrenamiento y evaluación incluyeron tanto dimensiones psicológicas como atributos sociodemográficos, además de la acción asignada y el resultado observado. En el ámbito psicológico, se incorporaron las dimensiones definidas por el modelo IBC: apoyo percibido de autonomía, motivación autónoma, actitud, normas subjetivas, intención de cambio, control conductual percibido y autoeficacia(Hagger & Chatzisarantis, 2014). En cuanto a los atributos sociodemográficos, se consideraron la edad, el género, el nivel de ingreso y el nivel de ejercicio basal. A estos se añadió la acción, correspondiente al nivel de meta física asignada, codificada en tres categorías de dificultad (1 = fácil, 2 = media, 3 = difícil), y la variable resultado (behavior), definida como un indicador binario de adherencia, con valor 1 en caso de cumplimiento

exitoso de la meta y 0 en caso contrario. Con esta configuración, la base de datos preprocesada se estructuró como un conjunto tabular clásico de machine learning, compuesto por observaciones independientes y variables predictoras tanto continuas como categóricas.

4.1.3 Justificación metodológica

La preparación de datos no constituyó únicamente un procedimiento técnico, sino una etapa metodológica fundamental para asegurar la validez y consistencia del proceso de modelado. Primero, se garantizó la reproducibilidad mediante el renombramiento estandarizado de variables y conversión de categorías cualitativas a valores numéricos, lo que permitió mantener coherencia en la estructura del conjunto de datos. En segundo lugar, se cuidó la representatividad a través de procesos de aleatorización y la posterior división en conjuntos de entrenamiento y prueba, evitando sesgos en la evaluación del modelo. Finalmente, se aseguró la compatibilidad algorítmica requerida para los modelos *logit* y para el entrenamiento en *Vowpal Wabbit*, mediante la uniformización del formato de entrada y la corrección de posibles inconsistencias tipográficas o de codificación. En conjunto, esta fase permitió transformar la base sintética inicial en un conjunto de datos válido y estructurado, preservando los *constructos* psicológicos del modelo IBC y los atributos sociodemográficos, pero adaptándolos al lenguaje operativo de los algoritmos de aprendizaje automático.

4.2 Caracterización del *dataset* sintético

4.2.1 Caracterización sociodemográfica

La simulación generó una base de datos compuesta por N = 2000 individuos, cada uno descrito mediante un conjunto de variables psicológicas y sociodemográficas. Estas variables se definieron a partir de parámetros de distribución específicos, asegurando que la población virtual reflejara patrones plausibles y heterogéneos.

La recompensa se define como la adherencia al comportamiento de salud, operacionalizada como el cumplimiento de la meta de actividad física durante tres

días consecutivos. Este valor binario de recompensa (behavior = 1 en caso de cumplimiento, 0 en caso contrario) actúa como la señal de refuerzo utilizada por el algoritmo de aprendizaje. La simulación permite así evaluar el desempeño del modelo de aprendizaje por refuerzo (*contextual bandit* implementado en *Vowpal Wabbit*) comparándolo con una asignación aleatoria base. El objetivo de esta evaluación offline es determinar si el algoritmo aprende a reasignar las metas de manera más efectiva maximizando la recompensa esperada en función de los atributos del usuario y la dificultad de la meta.

En cuanto a las características de la población, las variables generadas incluyen: una variable categórica Athlete (Yes/No), igualmente concentrada en “No” como punto de partida; la edad, definida como variable continua en el rango 18–25 años, distribuida de manera uniforme para representar adultos jóvenes; y el idioma, categórico con predominio absoluto del español (100% ES), coherente con la población objetivo de la simulación.

El género se categorizó en tres opciones (Male 40,9%, Female 48,6%, Other 10,5%), garantizando diversidad dentro de la población simulada. La variable empleo se distribuyó en cinco categorías: jornada completa (0%), jornada parcial (9,3%), desempleo (30,2%), labores domésticas o de cuidado (30,2%) y “otros” (30,2%), reflejando heterogeneidad laboral.

El hogar se representó mediante dos indicadores continuos: el número de personas económicamente activas (media 3,1) y el número de dependientes (media 0,55), ambos simulados dentro del rango 1–10. Por último, el ingreso se definió como una variable categórica en tres estratos (bajo, medio y alto), distribuidos equitativamente (~33% cada uno), mientras que el ejercicio actual se modeló como una variable continua en rango 1–7, con media 2,25, representando los días de ejercicio semanal reportados.

Tabla 4.2 Parámetros de simulación de variables sociodemográficas y contextuales.

Variable	Tipo	Codificación	Parámetro
Athlete	Categórica (Yes/No)	Binaria	[0%, 100%]
Edad	Categórica (18–25 años)	Numérica	Distribución uniforme (N)
Idioma	Categórica (ES/EN/NL)	Nominal	[100%, 0%, 0%]
Género	Categórica (M/F/Otro)	Nominal	[40,9%, 48,6%, 10,5%]
Empleo	Categórica (5 categorías)	Nominal	[0%, 9,3%, 30,2%, 30,2%, 30,2%]
Personas activas en hogar	Continua (1–10)	Numérica	Media = 3,1
Dependientes en hogar	Continua (1–10)	Numérica	Media = 0,55
Ingreso	Categórica (Bajo/Medio/Alto)	Ordinal (1–3)	[33,3%, 33,3%, 33,4%]
Ejercicio actual	Continua (1–7 días/semana)	Numérica	Media = 2,25

Fuente: elaboración propia.

4.2.2 Distribución de *constructos* psicológicos (IBC)

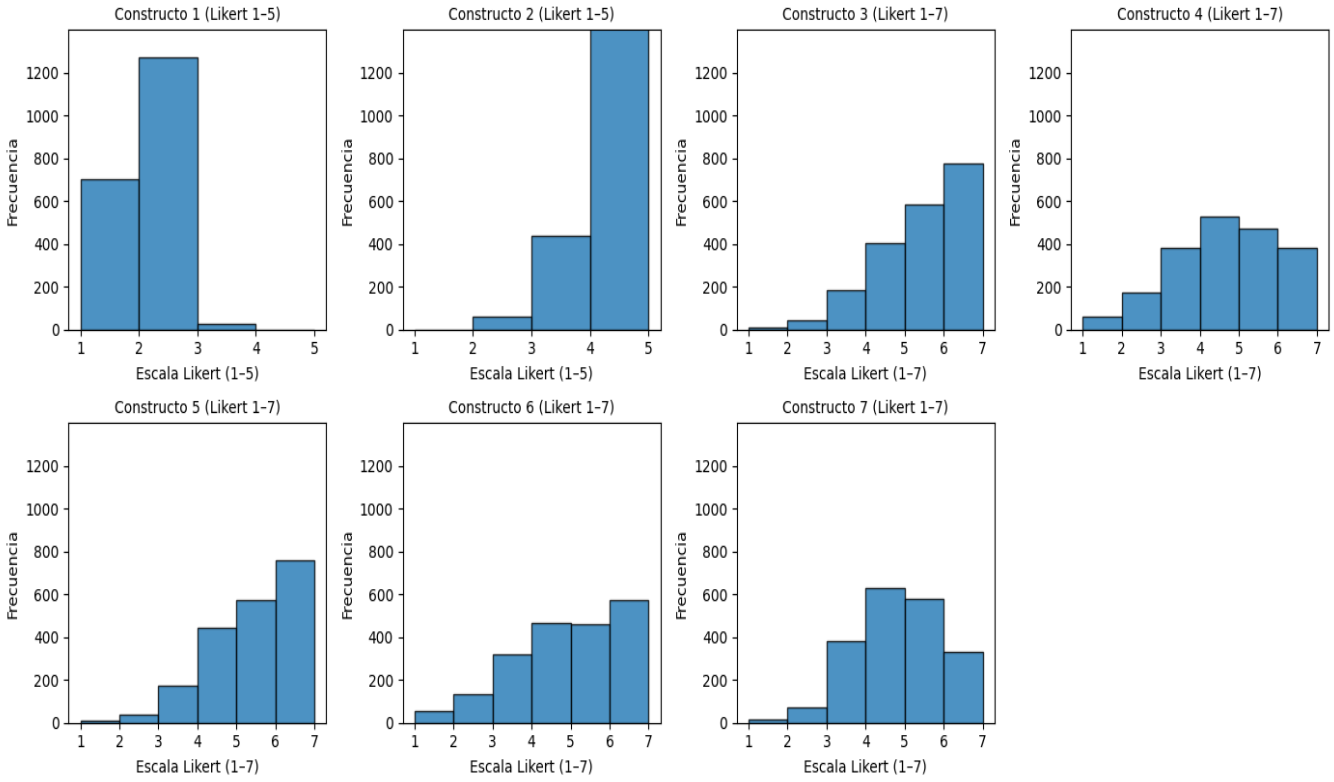
Los siete *constructos* derivados del modelo IBC se distribuyeron de acuerdo con la matriz de medias y covarianza ajustada previamente (ver Capítulo 3). En términos descriptivos, se observó un patrón consistente con la literatura empírica. La asimetría introducida mediante el vector α generó distribuciones sesgadas en algunos ítems, lo cual otorga realismo al simular concentraciones extremas.

Figura 4.1 Estadísticos descriptivos de los *constructos* IBC

	CM	AM	ATT	SN	PBC	INT	NFA
Media	2.51	3.71	5.07	4.38	4.76	4.10	3.89
Desviación estándar	0.90	1.00	0.98	1.25	1.43	1.31	0.96
Correlaciones	CM	AM	AT	SN	PBC	IN	AU
CM	1						
AM	-0.25	1					
ATT	-0.25	0.61	1				
SN	-0.10	0.38	0.36	1			
PBC	-0.25	0.70	0.50	0.24	1		
INT	-0.25	0.56	0.48	0.33	0.48	1	
NFA	-0.25	0.45	0.50	0.45	0.40	0.45	1

Notas: CM = Motivación Controlada, AM = Motivación Autónoma, ATT = Actitud, SN = Normas Subjetivas, PBC = Control Conductual Percibido, INT = Intención, NFA = Necesidad de Autonomía.

Figura 4.2 Histogramas de las variables IBC (Likert)



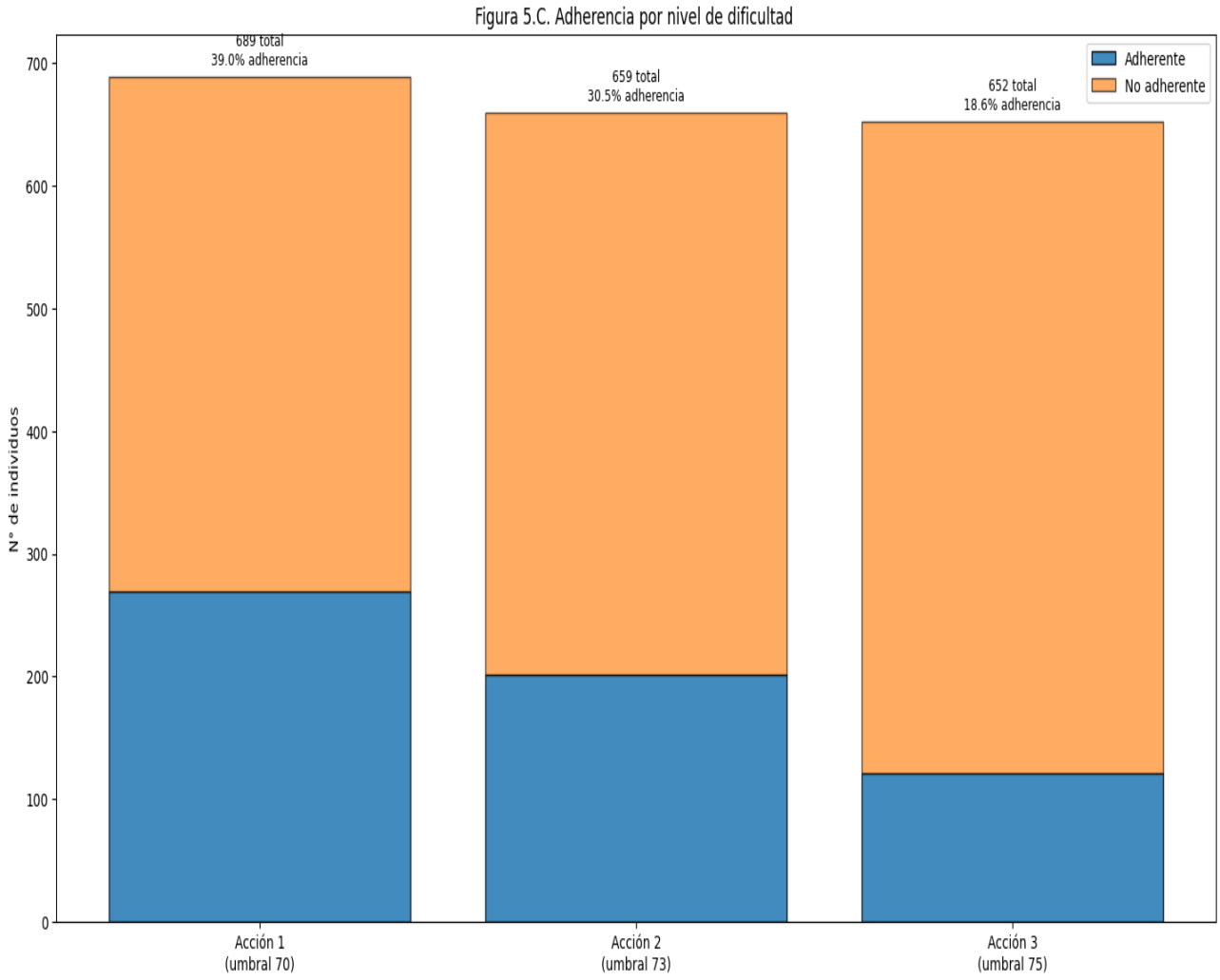
Fuente: elaboración propia.

4.2.3 Adherencia y función de puntuación

La función de puntuación (score) combinó, tanto los *constructos* IBC como las variables sociodemográficas anteriormente descritas. Se observó que la motivación autónoma, las normas subjetivas y la actitud ejercieron un efecto positivo sobre el puntaje, mientras que la cognición inicial baja redujo sistemáticamente los valores.

Posteriormente, el puntaje se contrastó con umbrales diferenciados por nivel de dificultad. Esta calibración permitió reflejar de manera más realista la progresiva exigencia de las metas asignadas. Los resultados globales indican que aproximadamente menos del 50% de la muestra alcanzó adherencia simulada, aunque con diferencias significativas entre niveles de dificultad.

Figura 4.3 Adherencia por nivel de dificultad



Fuente: elaboración propia.

4.3 Comparación con datos reales

La validez externa de la simulación, se compararon los *constructos* psicológicos y las variables sociodemográficas estáticas con una muestra real de n=553 individuos inscritos en la aplicación. La Tabla 4.3 presenta los estadísticos descriptivos de ambas fuentes, mientras que la Tabla 4.4 resume los resultados de las pruebas de Kolmogorov–Smirnov (KS) utilizadas para contrastar si las distribuciones empíricas de las variables simuladas y reales difieren significativamente, considerando la forma completa de la distribución y no solo sus momentos centrales.

En términos descriptivos, se observan similitudes en ciertas variables: por ejemplo, la media de ejercicio basal es cercana en ambas poblaciones (2,42 en simulada vs. 2,50 en real), lo que sugiere una calibración adecuada del modelo en este indicador. Algo similar ocurre con los *constructos* cm y am, que muestran medias próximas y desviaciones estándar comparables. Sin embargo, otras variables presentan diferencias marcadas: en particular, los *constructos*: actitud, intención y pbc se concentran en valores más altos en la población real (medias > 6), mientras que en la simulación se ubican en torno a valores intermedios (≈ 4–5).

Estas diferencias se confirman estadísticamente mediante las pruebas KS, cuyos valores de $p < 0.001$ en todas las variables analizadas (ejercicio basal, motivación autónoma, actitud, normas subjetivas, control conductual percibido, intención y apoyo a la autonomía) rechazan la hipótesis de igualdad de distribuciones.

En conjunto, los resultados indican que el simulador de generación de perfiles sintéticos, basado en la parametrización de medias, desviaciones estándar y matrices de covarianza extraídas de la literatura, logró aproximar adecuadamente ciertos rasgos (ejercicio basal, motivación autónoma), aún existen brechas respecto a la intensidad de *constructos* clave como la actitud e intención. Esto abre un espacio para recalibrar las distribuciones de la simulación con el fin de capturar de manera más fiel los perfiles observados en la población real.

Estas diferencias sugieren que, si bien el simulador de generación de perfiles sintéticos basado en el modelo IBC reproduce adecuadamente tendencias generales del comportamiento, como el nivel basal de ejercicio y la motivación autónoma, los *constructos* motivacionales que influyen en la adherencia presentan intensidades mayores en usuarios reales. Esto es consistente con la literatura, que señala que variables como la motivación autónoma, la percepción de control o la actitud suelen expresarse con mayor variabilidad y fuerza en entornos naturales que en simulaciones basadas únicamente en promedios poblacionales. Estas discrepancias abren la posibilidad de refinar la matriz de medias y covarianzas en futuras iteraciones, incorporando mayores fuentes empíricas o ajustando los niveles

de asimetría para mejorar la correspondencia psicológica entre ambas poblaciones. De este modo, el simulador puede evolucionar hacia un modelo más representativo y con mayor poder predictivo para el entrenamiento de políticas de personalización.

Tabla 4.3 Resumen de variables continuas (simulada vs real)

variable	fuente	count	mean	std	min	25%	50%	75%	max
current_exercise	Simulada	2000	2,419	0,834	1	2	2	3	5
cm	Simulada	2000	1,672	0,5	1	1	2	2	3
am	Simulada	2000	3,874	0,95	1	3	4	5	5
attitude	Simulada	2000	4,326	1,462	1	3	4	5	7
sn	Simulada	2000	5,206	1,283	1	4	5	6	7
pbc	Simulada	2000	5,308	1,272	1	4	5	6	7
intention	Simulada	2000	4,452	1,604	1	3	4	6	7
need_for_autonomy	Simulada	2000	4,285	1,2	1	3	4	5	7
current_exercise	Real	553	2,503	1,499	1	1	2	4	7
cm	Real	553	1,892	1,037	1	1	2	3	5
am	Real	553	4,105	0,952	1	4	4	5	5
attitude	Real	553	6,203	0,908	2	6	6	7	7
sn	Real	553	6,034	1,052	1	5	6	7	7
pbc	Real	553	5,837	1,1	2	5	6	7	7
intention	Real	553	6,186	0,967	1	6	6	7	7
need_for_autonomy	Real	553	5,331	1,141	2	5	5	6	7

Fuente: elaboración propia.

Tabla 4.4 Test de *Kolmogorov–Smirnov* (KS)

variable	test	stat	p_value
current_exercise	KS	0,2311	0,0
cm	KS	0,2418	0,0
am	KS	0,1258	2,00E-06
attitude	KS	0,5838	0,0
sn	KS	0,2999	0,0
pbc	KS	0,2439	0,0
intention	KS	0,5124	0,0
need_for_autonomy	KS	0,3441	0,0

Fuente: elaboración propia.

En conjunto, este análisis de validez externa permite establecer que el *dataset* sintético generado constituye una aproximación razonable a la población real, capturando adecuadamente tendencias generales del comportamiento y preservando relaciones psicológicas coherentes con la literatura. En este contexto, el conjunto de datos validado se utiliza como base para la siguiente etapa del estudio, orientada al entrenamiento y evaluación de algoritmos de aprendizaje por refuerzo contextual, cuyo objetivo es optimizar la asignación personalizada de metas de actividad física.

4.4 Aprendizaje por reforzamiento usando *contextual bandit*

En este estudio se adopta la política *epsilon-greedy* debido a su capacidad para equilibrar exploración y explotación en contextos donde la información inicial es limitada. Este enfoque resulta especialmente apropiado frente al CSP, ya que en etapas tempranas el sistema no dispone de evidencia suficiente para identificar patrones de comportamiento confiables. La política *epsilon-greedy* evita que el algoritmo se estanque en decisiones prematuras incentivando una exploración sistemática, mientras que, a medida que acumula evidencia, favorece progresivamente la explotación de las acciones con mayor rendimiento esperado. Este equilibrio permite una curva de aprendizaje más estable y eficiente, reduciendo la dependencia de exploración aleatoria en las primeras fases de personalización.

Una vez entrenado el modelo *logit* como estimador de probabilidades, el siguiente paso consistió en aplicar un algoritmo de aprendizaje por refuerzo contextual (*contextual bandit)* con el objetivo de optimizar la asignación de metas de actividad física en función del perfil del usuario. Para ello, se empleó la herramienta *Vowpal Wabbit* (VW), ampliamente utilizada en investigación aplicada por su eficiencia y escalabilidad en tareas de online learning. Los datos fueron transformados a formato compatible con VW, generando ejemplos en estructura de aprendizaje contextual, donde cada registro consistía en un contexto (perfil del usuario), una acción (nivel asignado), una recompensa binaria (éxito o fracaso) y la probabilidad estimada por el modelo *logit.*

4.4.1 Fundamentación teórica

El problema de recomendación de metas en este estudio se propone como un problema de bandido contextual (*contextual bandit problem*). A diferencia del bandido multi-brazo clásico, en el cual un agente debe elegir repetidamente entre un conjunto fijo de acciones sin información adicional, el bandido contextual incorpora un vector de características (contexto) que describe el estado del usuario en cada decisión.

En términos formales, para cada usuario i:

- Se observa un contexto Xi (atributos psicológicos y sociodemográficos).
- El algoritmo selecciona una acción $ai \in \{1, 2, 3\}$ (meta fácil, media o difícil).
- Se recibe una recompensa $ri \in \{0,1\}$, correspondiente al éxito conductual (behavior).

Una política π se define como una función de decisión probabilística que, dado un contexto Xi, asigna una probabilidad a cada acción posible:

$$\pi(a \mid X_i) = P(A = a \mid X_i)$$

El objetivo del agente es maximizar la recompensa esperada, es decir, elegir acciones que, en promedio, produzcan la mayor adherencia conductual posible.

$$E[R(\pi)] = E_{X\sim\mathbb{D}}\left[\sum_{a\in A}\pi(\,a \mid X\,)\; r(X,a)\right]$$

De esta forma, el *contextual bandit* constituye un marco teórico adecuado para el problema abordado en esta memoria, pues permite modelar la interacción dinámica entre el estado del usuario, la acción recomendada y la probabilidad de éxito asociada.

La estructura general del proceso de decisión del agente se presenta a continuación, donde se detalla el mecanismo *epsilon-greedy* utilizado para equilibrar la exploración y la explotación en la asignación de metas.

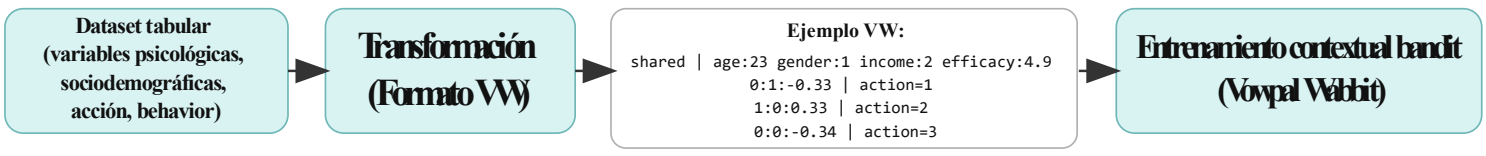
4.4.2 Transformación del *dataset* para VW

El entrenamiento se llevó a cabo en modalidad offline, utilizando el *dataset* simulado previamente generado. Para ello, se empleó el método de *Inverse Propensity Scoring* (IPS) como técnica de reducción implementada en *Vowpal Wabbit*. Este enfoque corrige el sesgo introducido por la política de logging aleatoria y permite que el algoritmo aprenda de manera más robusta, aprovechando toda la estructura probabilística del *dataset* y no únicamente las acciones registradas.

$$\widehat{R_{\mathrm{IPS}}}(\pi) = \frac{1}{n}\sum_{i=1}^{n}\frac{\pi(\,a_i \mid X_i\,)}{p(\,a_i \mid X_i\,)}\, r_i$$

De este modo, IPS se incorpora directamente en el proceso de aprendizaje, garantizando que la política resultante sea menos dependiente de sesgos del muestreo inicial y más representativa de los patrones simulados.

Figura 4.4 Transformación a formato VW



Fuente: elaboración propia.

4.4.3 Política *epsilon-greedy* y resultados esperados

El uso de VW en modo *contextual bandit* ofreció múltiples ventajas metodológicas:

Eficiencia computacional: VW está diseñado para manejar grandes volúmenes de datos y permite entrenar modelos complejos en segundos.

Adaptación a contextos individuales: al incorporar las variables psicológicas y sociodemográficas como contexto, el algoritmo aprende a ajustar la probabilidad de recomendar metas fáciles, medias o difíciles en función de cada perfil.

Entrenamiento sin despliegue real: gracias al enfoque offline, se logró obtener una política óptima preliminar sin exponer a usuarios reales a recomendaciones potencialmente inadecuadas, lo que resulta esencial en etapas tempranas de investigación en salud digital.

Con la implementación de este enfoque se espera que la política contextual aprendida supere en desempeño a la política aleatoria utilizada como línea base, genere un mayor porcentaje de adherencia simulada y reduzca el regret inicial característico del problema de *arranque en frío*, al contar con una base de decisiones informadas desde la primera interacción. De este modo, la aplicación del *contextual bandit* con *Vowpal Wabbit* representa un puente metodológico entre la simulación de datos sintéticos y la evaluación de algoritmos de personalización, aportando evidencia preliminar sobre la factibilidad de implementar políticas adaptativas en aplicaciones *mHealth.*

4.5 Resultados del piloto experimental

Los resultados mostraron que la política aprendida por el algoritmo *contextual bandit* logró una mayor proporción de adherencia simulada al comportamiento de salud, aumentando de un 50 % a aproximadamente un 60 % en comparación con la política aleatoria base. Además, el análisis de correlaciones confirmó que las variables psicológicas en particular la motivación autónoma, la intención y el control percibido presentaron relaciones consistentes con la literatura empírica del IBC, validando la estructura teórica del simulador. Estas evidencias sustentan la

viabilidad de los datos sintéticos como etapa de preentrenamiento para políticas de personalización en aplicaciones *mHealth.*

Los datos recolectados durante el piloto permitieron validar empíricamente las correlaciones teóricas observadas en la simulación, además de comprobar la integración técnica entre la aplicación móvil y la base de datos MySQL. Esta instancia generó los primeros registros reales de adherencia y recompensas conductuales definidas como el cumplimiento de metas de tres días consecutivos, los cuales sirvieron como base para el entrenamiento del modelo *contextual bandit* con datos reales, descrito en la siguiente sección.

4.5.1 Entrenamiento VW online con datos reales

En la etapa final del proyecto, los registros obtenidos a través de la aplicación *Apptivate* fueron integrados en una base de datos MySQL con el propósito de analizarlos mediante técnicas de aprendizaje por refuerzo contextual. Esta fase marca la transición desde la experimentación controlada hacia la aplicación práctica del modelo en línea, utilizando datos reales relacionados con el comportamiento, la adherencia y la dificultad percibida por los usuarios.

El proceso de experimentación consistió en entrenar y validar un algoritmo de tipo *contextual bandit* a partir de los registros exportados directamente desde la base SQL. Para ello se construyeron dos bases intermedias: la primera, denominada model_static, resume los *constructos* psicológicos y las variables demográficas derivadas del placement test. El script crear_static.py inserta en esta tabla atributos como género, ingreso, nivel de ejercicio actual y los *constructos* del modelo IBC (cm, am, attitude, sn, pbc, intention y need_for_autonomy), aplicando promedios y conversiones de acuerdo con los rangos de preguntas. La segunda base, model_cyclic_sessions, consolida la información cíclica de rutinas y sesiones a partir de las tablas training_sessions y routine_feedback, calculando valores promedio de variables como feeling, difficulty, intention y total_time, además de un indicador de recompensa (reward = 1 si el usuario completó al menos la mitad de los días planificados, y 0 en caso contrario). Este procedimiento incluye también un

backfill para los usuarios inactivos, garantizando que todos cuenten con al menos un registro válido.

Ambas bases se combinaron en un único archivo JSON estructurado por usuario. Cada entrada contiene un bloque estático junto con una lista de sesiones cíclicas filtradas dentro de los ciclos definidos. Este archivo fue utilizado como entrada para el pipeline de *Vowpal Wabbit*, donde se generaron ejemplos ADF combinando las variables estáticas y dinámicas de cada participante.

El entrenamiento del modelo se realizó mediante el script vw_train.py, aplicando una estrategia *epsilon-greedy* con exploración (ε = 0.3 en los experimentos), y utilizando como variable de recompensa la columna calculada directamente en la base de datos, en lugar de una simulación sintética. De esta manera, el modelo resultante (model.vw) aprendió a equilibrar exploración y explotación, ajustando la asignación de niveles de dificultad a partir de evidencia empírica obtenida de los registros reales de usuarios.

4.5.2 Predicciones e integración operativa

Tras completar el entrenamiento del modelo *contextual bandit* en *Vowpal Wabbit*, se procedió a la fase de predicción y despliegue de resultados. Este proceso constituye un puente entre la etapa experimental y la aplicación práctica en un sistema real de recomendación de metas.

El script vw_train.py implementa la función predict, la cual se encarga de cargar el modelo entrenado (model.vw) junto con el conjunto de datos exportado desde SQL en formato JSON (model_data.json). En este proceso, para cada usuario se construye un ejemplo ADF combinando sus características estáticas, que incluyen *constructos* psicológicos y variables sociodemográficas, con la información correspondiente a su sesión cíclica más reciente, la cual contempla variables como feeling, dificultad percibida, intención, total_time y reward.

A continuación, el modelo estima una distribución de probabilidad (pmf) asociada a cada acción posible, que en este caso corresponde a los tres niveles de dificultad (1, 2 o 3). Luego, se calcula un puntaje de factibilidad mediante la función

score_from_static_and_action, que integra *constructos* como intención, autonomía, motivación y variables sociodemográficas. Este puntaje representa la probabilidad teórica de éxito de un usuario frente a una determinada dificultad y actúa como un filtro adicional para la toma de decisiones.

Posteriormente, se aplican umbrales de factibilidad predeterminados (thresholds = {1:40, 2:55, 3:70} en este avance), descartando aquellas opciones que no superan el valor mínimo esperado. Entre las dificultades factibles, el modelo selecciona la que presenta la mayor probabilidad y, en caso de que ninguna cumpla los criterios, se elige aquella con la probabilidad absoluta más alta. Finalmente, para cada usuario se genera un objeto que incluye el vector completo de probabilidades (pmf), la acción recomendada (best_action) y el identificador del ciclo correspondiente (user_cycle_id).

El conjunto de predicciones se almacena en un archivo externo denominado predictions.json, lo que permite su análisis y posterior integración dentro del sistema de recomendación o de las evaluaciones empíricas.

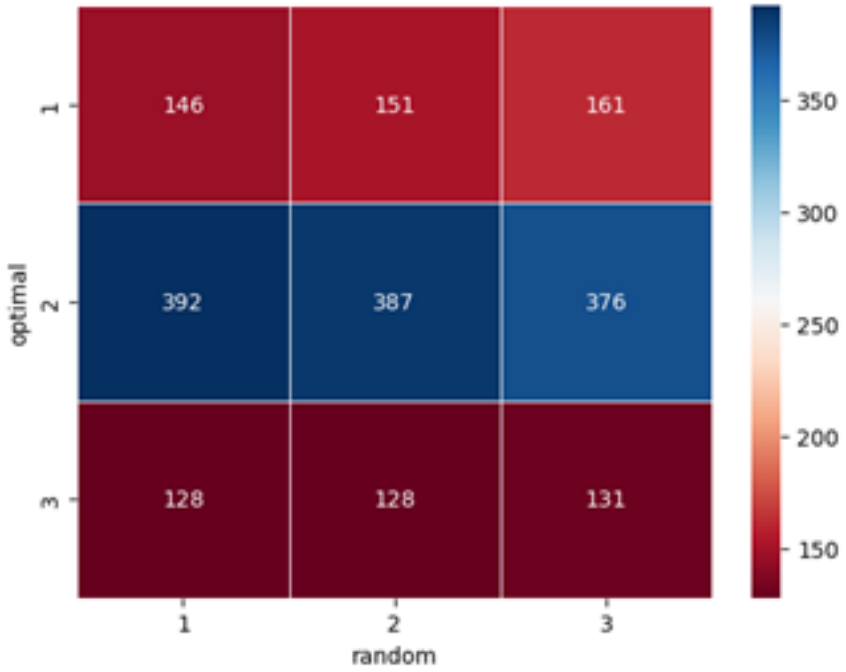
El archivo resultante generado por el modelo contiene, para cada usuario, la distribución completa de probabilidades (pmf), que representa el nivel de confianza del modelo en cada uno de los tres niveles de dificultad. Además, incluye la acción final recomendada (best_action), correspondiente a la dificultad seleccionada para la siguiente meta, y el identificador de ciclo de usuario (user_cycle_id), que permite vincular la recomendación con la etapa específica de entrenamiento de cada individuo. De esta forma, el sistema no solo propone una meta personalizada, sino que también cuantifica la certeza asociada a dicha recomendación y la contextualiza en el historial de comportamiento de cada usuario, fortaleciendo la trazabilidad y la interpretación del proceso de personalización.

Los resultados obtenidos permiten avanzar hacia una implementación real dentro de la aplicación de promoción de actividad física. En la práctica, el archivo predictions.json puede integrarse directamente en el backend del sistema para asignar metas semanales personalizadas a cada usuario. La representación probabilística (pmf) posibilita además la construcción de métricas adicionales, como

la entropía de decisión o el nivel de confianza del modelo, que resultan útiles para un monitoreo continuo y para evaluar la estabilidad del algoritmo. La incorporación de los *constructos* del modelo IBC en este proceso permite analizar si las recomendaciones generadas están alineadas con los niveles de motivación, intención y percepción de control de los usuarios. Finalmente, la estructura modular del archivo de salida facilita la actualización del sistema en línea, ya que el modelo puede reentrenarse de manera periódica y reemplazar el archivo de predicciones sin necesidad de modificar la arquitectura principal de la aplicación.

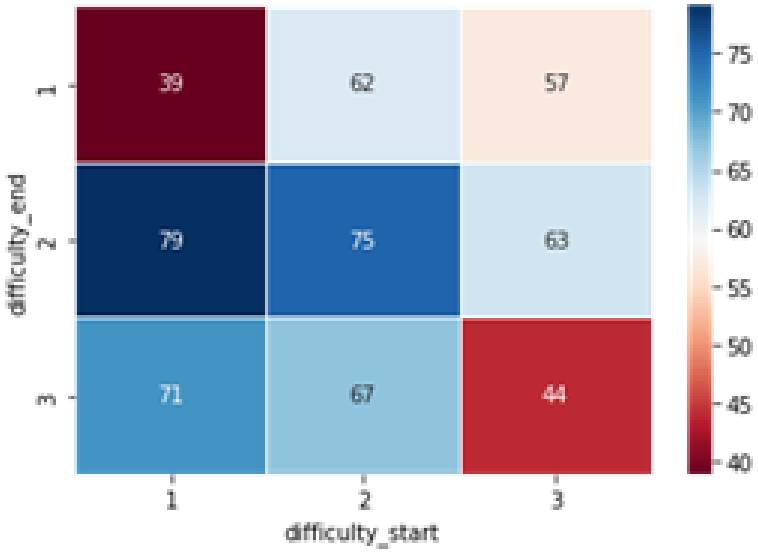
A partir de los resultados obtenidos en esta etapa, se realizó además una comparación entre el comportamiento del algoritmo entrenado en el *dataset* sintético y el entrenado con los datos reales del piloto. En el caso de los datos simulados, la matriz de transición generada por el *contextual bandit* mostró un reordenamiento significativo en los niveles de dificultad entre ciclos, con una fuerte convergencia hacia el nivel moderado (dificultad 2). Este ajuste produjo un aumento del reward promedio desde 0.50 a 0.62, equivalente a un incremento de 12 puntos porcentuales en adherencia simulada. Esta transición evidencia que el algoritmo identificó correctamente patrones psicológicos que favorecen el rendimiento, reasignando a los usuarios hacia niveles compatibles con su probabilidad de éxito.

Figura 4.5 Matriz de Transición con datos simulados



Fuente: Elaboracion Propia.

Figura 4.6 Matriz de Transición con datos reales



Fuente: Elaboracion Propia.

Por otro lado, la matriz de transición obtenida a partir del entrenamiento con datos reales del piloto mostró patrones similares, aunque con menor variabilidad en las transiciones. En el piloto, la asignación de dificultad tendió nuevamente a concentrarse en el nivel moderado, pero con una estabilidad más marcada en los niveles observados entre ciclos. Esto se explica porque un ciclo en el piloto equivale a tres días consecutivos de actividad, lo cual suaviza los cambios abruptos en comparación con el ciclo diario modelado en la simulación. Aun así, la lógica de reasignación del modelo entrenado sobre datos reales resulta altamente coherente con la del modelo simulado, lo que demuestra consistencia en los patrones aprendidos.

La comparación entre ambas matrices muestra que, pese a las diferencias esperables entre un entorno controlado y uno real, el comportamiento general del modelo es convergente: en ambos escenarios, la política aprendida tiende a reorganizar a los usuarios hacia niveles de dificultad más adecuados a su perfil psicológico y probabilidades de éxito. En particular, la convergencia al nivel moderado se observa tanto en la simulación como en el piloto, lo cual es consistente con la literatura que sugiere que metas demasiado fáciles reducen el involucramiento, mientras que metas demasiado difíciles generan abandono prematuro.

La similitud estructural entre las dos matrices de transición también actúa como un indicador indirecto de validez externa del *dataset* sintético. Esto respalda el rol de la simulación informada por teoría como fase de preentrenamiento, ya que permite que el bandit aprenda patrones generales antes de exponerse a la variabilidad y ruido propios de los datos reales. Así, el modelo llega mejor “preparado” al entorno real, reduciendo el impacto del cold start y mejorando la estabilidad del aprendizaje desde los primeros ciclos.

En síntesis, la integración operativa del archivo de predicciones y la consistencia observada entre los comportamientos del modelo en escenarios simulados y reales demuestran la factibilidad de desplegar un sistema de recomendación adaptativo dentro de una aplicación *mHealth.* El pipeline implementado que combina

simulación teóricamente informada, regresión logística y aprendizaje por refuerzo contextual ofrece una base técnica sólida para un sistema que aprende de manera continua y que puede mejorar progresivamente la adherencia de los usuarios mediante metas personalizadas y ajustadas a sus características individuales.

5. Conclusiones

El desarrollo de este trabajo permitió cumplir el objetivo central del producto mínimo viable (MVP): diseñar, simular y evaluar un marco de generación de datos sintéticos informados por teoría psicológica (modelo IBC) para el entrenamiento y validación inicial de algoritmos de personalización en salud digital. La simulación multivariada, la función de puntuación y la implementación del *contextual bandit* en *Vowpal Wabbit* demostraron la factibilidad de construir políticas de recomendación coherentes con la evidencia y capaces de maximizar la adherencia simulada en comparación con asignaciones aleatorias.

La posterior extensión hacia la implementación online utilizando datos reales registrados en la base SQL superó el alcance del MVP original, constituyendo un avance experimental y tecnológico adicional. Esta integración validó la interoperabilidad entre los módulos de simulación, la base de datos y el motor de aprendizaje por refuerzo, consolidando un flujo completo de entrenamiento offline y predicción online.

En términos científicos, se comprobó que los datos sintéticos informados por teoría pueden servir como una fase de preentrenamiento efectiva, reduciendo el impacto del CSP y acelerando la personalización en las primeras interacciones de los usuarios reales. Además, la coherencia entre las correlaciones simuladas y las observadas empíricamente respalda la validez del modelo IBC como marco de representación conductual.

Como línea de investigación futura, actualmente en desarrollo, se plantea la comparación cuantitativa entre la adherencia simulada y la adherencia observada en entornos reales, con el propósito de estimar la capacidad predictiva del modelo y ajustar de forma dinámica los parámetros de recompensa y dificultad. Este análisis permitirá fortalecer la conexión entre la simulación y el comportamiento efectivo de los usuarios, avanzando hacia un sistema de personalización adaptativo, escalable y validado empíricamente.

6. Anexos

6.1 Repositorio de código fuente

Con el fin de mantener la claridad del documento principal, los scripts utilizados en la construcción del simulador, la migración de datos y el entrenamiento de los modelos se presentan en un repositorio externo. El acceso al código completo puede realizarse a través del siguiente enlace:

Repositorio GitHub: https://github.com/Jandoto/apptivate

Este repositorio incluye los archivos principales:

Simulacion_principal.py generación de perfiles sintéticos de usuarios (*constructos* IBC + sociodemográficos).

calculo_behavior.py cálculo de la variable behavior a partir del score y umbrales definidos.

crear_static.py generación de la tabla estática de *constructos* psicológicos a partir del placement test.

crear_cyclic.py consolidación de datos de sesiones cíclicas y cálculo de recompensas.

crearjson.py construcción del archivo model_data.json a partir de la base SQL.

vw_train.py entrenamiento y predicción con *Vowpal Wabbit* (*contextual bandit*).

6.2 Diccionario de variables

Tabla 6.1 Diccionario de variables

Variable	Descripción
Q1–Q7	*Constructos* IBC (Likert 1–7)
Age	Edad simulada (18–25)
Gender	Male / Female / Other
Income	Low / Medium / High
Employment	Estado laboral
Current_exercise	Nivel basal de ejercicio (1–7)
action	Meta asignada (1 = fácil, 2 = media, 3 = difícil)
score	Puntuación agregada
behavior	Adherencia simulada (0/1)

Fuente: elaboración propia.

6.3 Glosario de variables y *constructos* del modelo IBC

Tabla 6.2 Glosario de variables y constructos del modelo IBC

current_exercise	Actividad física actual (días por semana)
cm	Motivación controlada
am	Motivación autónoma
attitude	Actitud hacia el comportamiento
sn	Normas subjetivas
pbc	Control conductual percibido
intention	Intención de cambio
need_for_autonomy	Necesidad de autonomía

Fuente: elaboración propia.

Referencias

Ferrada Torres, V., Nguyen, H. P., Caro, J. C., & Lipman, S. L. (2024). *Tailored goals and individual choice: A field experiment on automated commitment devices for health behavior.*

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning.* Cambridge, MA: MIT Press.

Hagger, M. S., & Chatzisarantis, N. L. (2014). An integrated behavior change model for physical activity. *Exercise and Sport Sciences Reviews*.

Kan, J., Kim, D., Kim, Y., & Lee, J.-G. (2023). Learning behavior change policies by simulating the past. *Knowledge and Information Systems*.

Mohr, D. C., Burns, M. N., Schueller, S. M., Clarke, G., & Klinkman, M. (2014). *Behavioral intervention technologies: Evidence review and recommendations for future research in mental health.* General Hospital Psychiatry.

Park, S. T., & Chu, W. (2009). Pairwise preference regression for cold-start recommendation. *Proceedings of the third ACM conference on Recommender systems (RecSys '09)*.

Pilani, A., Gupta, A., & Kathuria, A. (2021). Contextual bandit approach-based recommendation system for personalized web-based services. *Applied Artificial Intelligence*.

Tsao, C., Li, Y., Wang, J., & Zhang, H. (2023). Synthetic data generation for healthcare: Methods, challenges, and applications. *Artificial Intelligence in Medicine*.

8. Glosario de términos y abreviaciones

IBC: Integrated Behavior Change Model. Modelo Integrado de Cambio de Comportamiento que combina la TPB, SDT y HAPA para explicar y predecir cambios conductuales en salud.

TPB: Theory of Planned Behavior. Teoría de la Conducta Planificada; establece que la intención depende de la actitud, las normas subjetivas y el control percibido.

SDT: Self-Determination Theory. Teoría de la Autodeterminación; explica la motivación según autonomía, competencia y relación.

HAPA: Health Action Process Approach. Modelo de Procesos de Acción en Salud; diferencia entre fases motivacionales (intención) y volitivas (ejecución).

CSP: Cold Start Problem. Problema del *arranque en frío* en sistemas de recomendación; ocurre cuando no hay datos previos de usuarios o ítems.

mHealth: Mobile Health. Salud móvil; uso de aplicaciones móviles para promover y monitorear conductas saludables.

VW: Vowpal Wabbit. Herramienta de machine learning eficiente y escalable; usada en este trabajo para entrenar un algoritmo *contextual bandit.*

RL: Reinforcement Learning. Aprendizaje por refuerzo; paradigma donde un agente aprende a maximizar recompensas interactuando con un entorno.

IPS: Inverse Propensity Scoring. Método de evaluación offline que ajusta sesgos de logging para estimar recompensas de nuevas políticas.

AUC-ROC: Area Under the Curve – Receiver Operating Characteristic. Métrica para evaluar la capacidad de discriminación de un modelo binario.

Log-loss: Logarithmic Loss. Función de pérdida que penaliza predicciones erróneas según la probabilidad asignada.

Regret: Diferencia acumulada entre la recompensa obtenida por una política y la recompensa máxima posible (óptima).

Reward: Señal de recompensa. En este estudio, corresponde a la adherencia (1) o no adherencia (0) del usuario.

Likert: Escala ordinal (1 a 5 o 1 a 7) usada en encuestas para medir percepciones y actitudes.

Score: Función de puntuación que combina factores psicológicos, sociodemográficos y dificultad de la meta para determinar probabilidad de adherencia.

Behavior: Variable binaria de resultado; indica éxito (1) o fracaso (0) conductual, utilizada como señal de recompensa.

Logit: Modelo de regresión logística usado para estimar la probabilidad de adherencia en base a predictores.

Python, Numpy, Pandas, Scipy: Lenguaje y librerías de programación utilizadas para generar y procesar datos sintéticos.

Offline RL: Aprendizaje por refuerzo fuera de línea; entrenamiento de políticas usando datasets fijos en vez de interacción en tiempo real.

UNIVERSIDAD DE CONCEPCION – FACULTAD DE INGENIERÍA

RESUMEN MEMORIA DE TITULO

Departamento	**: Ingeniería Civil Industrial**
Carrera	**: Civil Industrial**
Nombre del memorista	**: Alejandro Hernán García Zavala**
Título de la memoria	**: Simulación y análisis de datos sintéticos para evaluar algoritmos de aprendizaje por refuerzo en una aplicación móvil de promoción de actividad física**
Fecha de presentación oral	**:**
Profesor(es) Guía	**: Juan Carlos Caro Seguel**
Profesor(es) Revisor(es)	**:**
Concepto	**:**
Calificación	**:**

Resumen

Esta investigación aborda la mitigación del problema de arranque en frío en sistemas de recomendación para salud digital mediante la generación de datos sintéticos informados por teoría psicológica. En el aprendizaje por refuerzo, el arranque en frío ocurre cuando el sistema debe tomar decisiones sin contar con historial previo del usuario, lo que limita la personalización en las primeras interacciones y puede afectar la adherencia al comportamiento de salud.

El objetivo del estudio es desarrollar y evaluar un simulador de usuarios sintéticos, parametrizado con constructos del Modelo Integrado de Cambio de Comportamiento y variables sociodemográficas. La metodología incluye la construcción de una matriz de covarianza multivariada, una función de puntuación y una variable binaria de adherencia. A partir de estos datos, se entrenó un modelo logit para estimar probabilidades de éxito conductual y un algoritmo de

aprendizaje por refuerzo de tipo *contextual bandit* para optimizar la asignación de metas de actividad física.

En una población simulada y bajo un protocolo común de evaluación, la política aprendida mostró mayores niveles de adherencia que una asignación aleatoria, especialmente en etapas iniciales. Los resultados indican que los datos sintéticos constituyen una estrategia efectiva para calibrar políticas de personalización y acelerar la toma de decisiones adaptativas en aplicaciones *mHealth.*