

Investigación Preliminar

R.B. — I.B. — X.R. — A.M. — A.D. — J.P.R.
Dr. Iván Ongay

25 de Marzo, 2025

Resumen

Se condujo una breve investigación preliminar para identificar factores significativos cuyos tendrán un impacto notable a lo largo de el desarrollo de el proyecto en cuestión.

Índice

1. Introducción	2
2. Características de Mercado	2
3. Forecast	4
4. Optimización	6

1. Introducción

Pinchi investigación mamalona que nos aventamos alv

2. Características de Mercado

En esta sección exploramos algunas características de el mercado en cuestión (Emisiones de tarjetas crediticias por Fintechs), para así, conocer algunos aspectos relevantes a el proyecto en desarrollo.

Resumen Rappicard

RappiCard ha emergido como un actor significativo en el mercado mexicano de tarjetas de crédito desde su lanzamiento en 2021. A continuación, se presentan datos clave que ilustran su crecimiento y posición en el sector:

- **Cartera de crédito:** Para agosto de 2024, RappiCard alcanzó una cartera de crédito de 5,107 millones de pesos, consolidándose entre las fintech más destacadas del país [8].
- **Número de clientes:** En abril de 2024, la empresa superó el millón de clientes en México, reflejando una rápida adopción de sus servicios [2].
- **Líneas de crédito:** Gracias al análisis de datos de la plataforma Rappi, RappiCard ofrece líneas de crédito iniciales promedio de 18,000 MXN, un 53 % superiores a las de otras fintechs [11].
- **Segmento demográfico:** Aproximadamente el 75 % de los clientes de RappiCard tienen menos de 35 años, indicando un enfoque exitoso en atraer a la población joven y digitalmente activa [2].
- **Proceso de aprobación y entrega:** RappiCard se distingue por su proceso ágil, con aprobaciones en dos minutos y entrega de la tarjeta física en un promedio de 30 minutos en áreas con cobertura de Rappi [8].
- **Beneficios adicionales:** La tarjeta ofrece recompensas como 5 % de cashback en compras a través de RappiTravel y 1 % en otros comercios, además de no cobrar anualidad y permitir diferir compras mayores a 500 MXN a meses sin intereses [11].
- **Índice de Morosidad (IMOR) de Banorte:** En el informe financiero del cuarto trimestre de 2023, se reportó un incremento en el índice de morosidad (IMOR) del segmento de tarjetas de crédito de Banorte, alcanzando el 3.3 %. Este aumento se atribuye a la consolidación de Banorte con RappiCard, ya que el IMOR era de 2.4 % en el mismo período del año anterior. El IMOR es una métrica clave en el análisis de riesgo crediticio, que mide el porcentaje de clientes que no han realizado pagos en sus tarjetas de crédito dentro de los plazos establecidos [4].
- **Visión de José Antonio Murillo, CEO de RappiCard:** En cuanto al futuro de la alianza entre RappiCard y Banorte, José Antonio Murillo comentó: “Hay mercado para todos”. Esta declaración refleja la confianza de RappiCard en su modelo de negocio y en la demanda de servicios digitales en el sector financiero, especialmente ante el auge de las fintechs como la propia RappiCard [9].
- **Modelo de operación digital de RappiCard:** El CEO de RappiCard también destacó que su modelo de operación digital, combinado con atención personalizada, ha sido un factor clave para su éxito. RappiCard ha logrado un Net Promoter Score (NPS) superior a los 80 puntos, comparado con el promedio de 37 puntos en la banca tradicional según encuestas de Banxico. Esto indica que los usuarios están muy satisfechos con el servicio y están dispuestos a recomendarlo, lo que refuerza su posición en el mercado [9].
- **Crecimiento de RappiCard:** Desde su lanzamiento en enero de 2021, RappiCard ha emitido más de 700,000 tarjetas de crédito y cuenta con un portafolio de crédito que supera los 4,000 millones de pesos. A tan solo dos años de su entrada al mercado, RappiCard ya ha superado en emisión de tarjetas a bancos internacionales establecidos desde hace décadas, lo que destaca su rápido crecimiento y aceptación [9].

Estos indicadores reflejan el crecimiento acelerado y la sólida posición de RappiCard en el mercado financiero mexicano, especialmente entre los consumidores jóvenes y familiarizados con la tecnología.

- El ecosistema fintech en México creció a una tasa del 18 % anual entre 2019 y 2023 [3].
- Existen 844 startups fintech activas en el país [6].

- En 2023, las fintechs otorgaron más de 3,000 millones de dólares en préstamos, beneficiando a más de 5 millones de usuarios [12].
- En 2021, solo el 32.7 % de la población entre 18 y 70 años tenía al menos un crédito formal [7].
- Los ingresos del sector fintech crecieron a una tasa compuesta del 22 % anual entre 2021 y 2024, y 31 % en el último periodo (2023-2024) [14].
- El Índice de Maduración del Ecosistema Fintech (INFIN) se sitúa en 48 %, indicando una etapa de desarrollo aún temprana [10].
- Solo el 20 % de la población mexicana tiene una tarjeta de crédito; el efectivo sigue siendo el medio de pago principal [5].

En conjunto, estos indicadores reflejan un sector en plena transformación, con un impacto creciente en el acceso al crédito y la inclusión financiera en México.

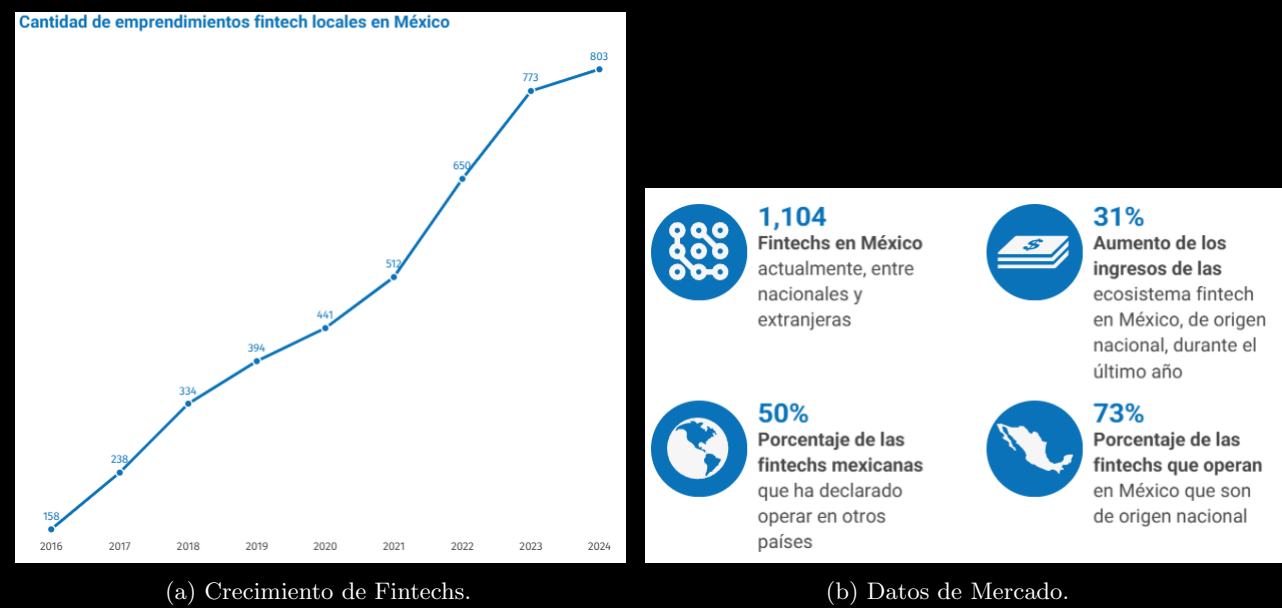


Figura 1: Fuente: Finnovista, Fintech Radar México, 2025

Mercado: Tarjetas y Créditos

De acuerdo con las bases de datos de la Comisión Nacional Bancaria y de Valores (CNBV) [13], se construyeron las siguientes tablas que se presentan a continuación.

Número de tarjetas de crédito por tipo de institución (trimestral)

Institución	2023 T1	2023 T2	2023 T3	2023 T4	2024 T1	2024 T2
Banca Múltiple	32,538,425	33,380,788	34,006,416	34,603,588	35,110,103	35,843,245
Banca de Desarrollo	20,362	20,495	20,680	20,601	20,756	21,037
Sofipos	3,487,589	3,155,757	4,011,168	3,427,085	3,955,533	5,194,963
Socaps	4,544	6,372	9,544	13,508	17,955	22,901

Número total de créditos otorgados por tipo de institución (trimestral)

Institución	2023 T1	2023 T2	2023 T3	2023 T4	2024 T1	2024 T2
Banca Múltiple	59,480,857	60,704,810	61,731,925	62,709,348	63,108,527	64,949,134
Banca de Desarrollo	809,487	824,525	821,385	843,350	838,168	888,626
Sofipos	3,911,704	3,719,181	4,677,062	4,223,454	4,953,483	5,852,773
Socaps	2,814,646	2,852,568	2,861,541	2,844,703	2,830,927	2,898,419

Población: Jóvenes $\in [18, 30]$

Población: Económicamente activos

Hábitos de consumo (México)

Población: Lineas telefonicas (no prepago)

Población: Personas que cumplen con req. de ingresos

3. Forecast

Predecir valores futuros usando datos del pasado —lo que llamamos *forecasting de series de tiempo*— es algo clave en muchas áreas como economía o negocios. En el caso de RappiCard, por ejemplo, saber cuántas tarjetas de crédito se necesitarán en el futuro ayuda a tomar mejores decisiones sobre cuántas producir y cuándo hacerlo.

Normalmente, este tipo de predicciones se hace con modelos estadísticos como ARIMA o modelos de machine learning como LSTM. Pero en este proyecto queremos probar algo distinto:

Un modelo basado en geometría, llamado *Geometrical Realization for Time Series Forecasting* [1].

Este modelo funciona de otra manera: en lugar de ver la serie de tiempo como una simple lista de números, la convierte en una especie de “figura” o “curva” en un espacio de más dimensiones. Luego, sobre esa figura aplica una transformación que permite hacer la predicción. El objetivo de esta sección es explicar bien cómo funciona ese enfoque y por qué creemos que tiene potencial para este proyecto. Además de la implementación de este modelo, se ha decidido implementar un modelo estadístico, de esta manera se genera una comparación entre dos perspectivas distintas de forecasting.

¿Qué es una serie de tiempo?

Una serie de tiempo es una colección de observaciones tomadas en momentos consecutivos, a intervalos regulares. Por ejemplo, el número de tarjetas emitidas por RappiCard cada trimestre es una serie de tiempo univariada. En general, las series pueden ser:

- **Univariadas:** una sola variable (e.g. ventas mensuales).
- **Multivariadas:** varias variables relacionadas medidas en paralelo (e.g. ventas + gasto en marketing).

Preliminares al modelaje geométrico

Antes de implementar modelos geométricos sobre una serie de tiempo, es necesario transformar esta serie en una estructura que capture su evolución temporal. A este proceso se le conoce como **embedding**, y es la piedra angular del enfoque propuesto en el artículo *Geometrical Realization for Time Series Forecasting* [1].

Delay Embedding

Dado un conjunto de observaciones x_1, x_2, \dots, x_n de una serie temporal univariada, se define el *embedding de dimensión d y retardo τ* como la colección de vectores:

$$\mathbf{x}^{(i)} = [x_i, x_{i-\tau}, x_{i-2\tau}, \dots, x_{i-(d-1)\tau}]$$

para i tal que todos los índices sean positivos (es decir, $i \geq (d-1)\tau$).

Este procedimiento transforma la serie original en una secuencia de puntos en \mathbb{R}^d , permitiendo así analizar su comportamiento como una **curva** o **superficie** en un espacio de dimensión superior. Esta idea está inspirada en el **Teorema de Takens**, el cual garantiza (bajo ciertas condiciones) que esta transformación preserva las propiedades dinámicas del sistema original.

Ejemplo de Embedding: Supongamos que tenemos la serie:

$$[5, 8, 12, 18, 25, 35]$$

Si usamos $d = 3$ y $\tau = 1$, el embedding produce:

$$[[12, 8, 5], [18, 12, 8], [25, 18, 12], [35, 25, 18]]$$

Cada uno de estos vectores puede verse como un punto en \mathbb{R}^3 .

HD-Embedding (High-Dimensional Embedding)

El embedding tradicional puede ser limitado si la serie no tiene una dinámica demasiado compleja. Sin embargo, cuando se desea capturar comportamientos no lineales y patrones más sutiles, se puede aumentar la dimensión d y ajustar el retardo τ estratégicamente. A este proceso se le conoce como **embedding de alta dimensión (HD-Embedding)**.

La idea detrás del HD-Embedding es que, al representar los datos en un espacio más rico, se facilita el aprendizaje de estructuras geométricas subyacentes (como curvas o superficies algebraicas) que pueden capturar las reglas implícitas de evolución temporal.

Esta técnica es la base para los modelos implementados en los scripts que se presentan a continuación.

Modelo `polynomial_model.py`

Este modelo implementa una aproximación algebraica a la serie embebida. El objetivo es ajustar una función polinomial multivariada sobre los puntos del embedding para luego predecir valores futuros. A continuación se explica su funcionamiento:

1. `build_polynomial_model(degree, alpha)`: Esta función construye un pipeline de `scikit-learn` compuesto por:

- `StandardScaler()`: Estandariza los datos.
- `PolynomialFeatures(degree)`: Genera todas las combinaciones polinomiales de las variables hasta el grado especificado.
- `Ridge(alpha)`: Aplica una regresión Ridge (con penalización L2) para evitar sobreajuste.

2. `fit_model(X, y, degree, alpha)`: Ajusta el modelo sobre el conjunto de entrenamiento. Aquí:

- $X \in \mathbb{R}^{m \times d}$: puntos del embedding.
- $y \in \mathbb{R}^m$: valores futuros a predecir (típicamente, el siguiente valor en la serie original).

3. `predict(model, x)`: Predice el siguiente valor en la serie, dado un vector de embedding.

4. `predict_recursive(model, init_point, steps)`: Realiza predicciones a múltiples pasos en el futuro. Usa la predicción anterior como entrada para la siguiente, imitando una simulación del sistema dinámico.

5. `search_best_degree_alpha`: Prueba múltiples combinaciones de grado polinomial y parámetros de regularización y selecciona la que minimiza el *sMAPE* (symmetric Mean Absolute Percentage Error), útil para comparar rendimiento entre modelos con diferente escala.

Visualización: Se genera un gráfico 3D del embedding original y otro con los puntos predichos. Se usa color y conectividad para mostrar claramente la evolución temporal y cómo la predicción se alinea (o no) con el embedding real.

Modelo `geo_realization.py`

Este script implementa una lógica distinta: en vez de ajustar un modelo simple sobre el embedding, construye una **superficie polinomial** sobre el embedding completo, tratando de parametrizarlo globalmente.

1. `fit_surface_to_embedding(X, degree)`: Ajusta una superficie polinomial en \mathbb{R}^3 sobre los puntos del embedding. Internamente:

- Se usa `PolynomialFeatures` para transformar los datos.
- Se ajusta un modelo `LinearRegression` (sin regularización) sobre el embedding para predecir la tercera coordenada a partir de las dos primeras.
- El resultado es una "variedad" (una especie de superficie algebraica) que intersecta todos los puntos del embedding.

2. `predecir_desde_superficie(punto_actual, model)`: A partir de un nuevo punto (x_0, x_1) del embedding, genera la coordenada x_2 usando la superficie ajustada. Esta es la lógica de predicción: buscar puntos en el embedding que sigan la dinámica estimada.

3. `predecir_n_puntos(X, model, n)`: Predice n puntos futuros a partir del último punto en el embedding original. Cada nuevo punto predicho se vuelve entrada para el siguiente paso.

4. `plot_embedding_with_surface(X, model, puntos_predichos=None)`: Visualiza:

- El embedding original (línea verde).
- La superficie ajustada (malla azul translúcida).
- Opcionalmente, los puntos predichos (línea roja).

Esto permite ver si los puntos futuros siguen la forma del embedding o se desvían.

5. `evaluar_modelo_fragmentado(X, idx_start, idx_end)`: Permite seleccionar un fragmento del embedding para ajustar la superficie y evaluar predicciones hacia adelante y hacia atrás. Se reportan métricas como MAE y MSE entre los puntos predichos y los reales no usados en el entrenamiento.

Esta lógica nos permite evaluar si la superficie está capturando correctamente la forma global del embedding, incluso más allá del rango de datos usados para el ajuste. Es una herramienta que nos ayuda a entender si el modelo generaliza bien o si está sobreajustado.

4. Optimización

Entendiendo: ¿Qué es un *stock* en tarjetas de crédito?

El objetivo es predecir cuántas tarjetas físicas (aún no activadas) deben fabricarse en cada periodo. Para esto, debemos tener en cuenta tres factores fundamentales:

1. **Tarjetas físicas sin activar**: Son los objetos físicos que se deben tener listos para satisfacer la demanda futura. Su producción debe anticiparse al uso, pero sin caer en el exceso.
2. **Demanda proyectada**: Depende de múltiples variables, como campañas de marketing (e.g. tarjetas temáticas), comportamiento regional y perfil de los usuarios (por edad, ingresos, historial).
3. **Costos de sobreproducción o desabasto**: Sirven como función objetivo en la optimización. También se deben considerar costos de almacenamiento, distribución y tiempos de entrega.

Esto nos lleva naturalmente a un problema de optimización bajo incertidumbre, con un componente predictivo fuerte (forecast) y decisiones estructurales que se prestan al modelado matemático.

Referencias

- [1] Ali Bayeh, Malek Mouhoub, and Samira Sadaoui. Geometrical realization for time series forecasting. In *Proceedings of the 36th International Conference on Industrial, Engineering Other Applications of Applied Intelligent Systems (IEA/AIE)*, Lecture Notes in Computer Science, Regina, Canada, 2024. Springer. doi: 10.1007/978-3-031-50724-1_23. URL https://link.springer.com/chapter/10.1007/978-3-031-50724-1_23.
- [2] Expansión Digital. Rappicard alcanza los 5,107 millones de pesos en su cartera de crédito. 2024. URL <https://expansion.mx/economia/2024/09/18/rappicard-alcanza-5107-mdp-cartera-de-credito>.
- [3] Dock. El ecosistema de fintechs en México ha crecido notablemente. 2024. URL <https://dock.tech/es/fluid/blog/financiero/servicios-financeiros-en-mexico>.
- [4] Fintech Expert. Efecto fintech: El impacto de rappicard en el mercado financiero. *Fintech Expert*, 2024. URL <https://www.fintechexpert.mx/p/efecto-fintech-el-impacto-de-rappicard>.
- [5] El Financiero. México, un mercado atractivo para las fintechs. 2024. URL <https://www.elfinanciero.com.mx/bloomberg/2024/12/22/hay-tiro-mercado-pago-y-nu-se-alistan-para-dominar-el-mercado-fintech-en-america>.
- [6] Finnovating. El sector fintech de México: principales startups y tendencias. 2024. URL <https://finnovating.com/news/el-sector-fintech-de-mexico-principales-startups-y-tendencias>.

- [7] INEGI. Encuesta nacional de inclusión financiera 2021. 2021. URL <https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2022/enif/ENIF21.pdf>.
- [8] Forbes México. Rappicard supera los 5,000 millones de pesos en cartera de crédito. 2024. URL <https://www.forbes.com.mx/ad-rappicard-supera-5000-millones-pesos-cartera-credito/>.
- [9] Forbes México. “hay mercado para todos”, dice rappicard sobre el banco digital de banorte. *Forbes México*, 2024. URL <https://forbes.com.mx/hay-mercado-para-todos-dice-rappicard-sobre-el-banco-digital-de-banorte/>.
- [10] Santander. El horizonte de las fintech en México. 2024. URL <https://www.santander.com/es/sala-de-comunicacion/insights/el-horizonte-de-las-fintech-en-mexico>.
- [11] José Ignacio Stark. Rappicard logra rentabilidad y supera el millón de clientes en México: José Antonio Murillo. 2024. URL <https://chocale.mx/2024/05/rappicard-tarjeta-de-credito-1-millon-de-clientes-mexico/>.
- [12] UFlow. El crecimiento del mercado crediticio en México y la influencia de las fintech. 2024. URL <https://uflow.biz/blog/el-crecimiento-del-mercado-crediticio-en-mexico-y-la-influencia-de-las-fintech>.
- [13] Comisión Nacional Bancaria y de Valores (CNBV). Bases de datos de inclusión financiera. <https://www.gob.mx/cnbv/acciones-y-programas/bases-de-datos-de-inclusion-financiera>, 2024. Consultado en marzo de 2025.
- [14] Reporte Índigo. El mercado de fintechs en México ha madurado: Finnovista. 2024. URL <https://www.reporteindigo.com/dinero/El-mercado-de-fintechs-en-Mexico-ha-madurado-Finnovista-20250225-0175.html>.