

## Registros sociales y aprendizaje estadístico: respuestas oportunas para gestantes y niñez temprana en Uruguay

Registros sociais e aprendizagem estatístico: respostas oportunas para gestantes e primeira infância no Uruguai

 **Elina Gómez Bonaglia**

Universidad de la República, Facultad de Ciencias Sociales, Unidad de Métodos y Acceso a Datos. Montevideo, Uruguay.

 **Gustavo Méndez Barbato**

Universidad de la República, Facultad de Ciencias Sociales, Departamento de Ciencia Política. Montevideo, Uruguay.

**Philippe Rimoli Rimbaud**

Ministerio de Desarrollo Social, Sistema de Información Integrada del Área Social. Montevideo, Uruguay.

 **Natalia da Silva**

Universidad de la República, Facultad de Ciencias Económicas y de Administración, Instituto de Estadística. Montevideo, Uruguay.

 [gustavo.mendez@cienciassociales.edu.uy](mailto:gustavo.mendez@cienciassociales.edu.uy)

 <https://doi.org/10.28917/ism.2025-v6-id152>

**Resumen:** Este artículo analiza el potencial de la integración de registros sociales y la capacidad predictiva de los modelos basados en aprendizaje estadístico para mejorar la oportunidad y efectividad de las políticas sociales dirigidas a gestantes y niñez temprana en Uruguay. Ante la persistente pobreza infantil y los problemas de focalización (errores de exclusión, inclusión y *timing*) el estudio propone la integración del Sistema Informático Perinatal (SIP) y el Sistema de Información Integrada del Área Social (SIAS) para mejorar las respuestas de las Asignaciones Familiares (AFAM-PE). A partir de esta integración, se desarrolla un modelo predictivo de vulnerabilidad social utilizando el algoritmo Stochastic Gradient Boosting (GBM). El modelo, entrenado con datos de las cohortes 2015-2023, identifica con un 85,7% de precisión a las gestantes en hogares vulnerables, utilizando variables clave que se relevan en la primera consulta antenatal. El estudio demuestra, mediante un ejercicio con la cohorte 2023, que este modelo permitiría reducir significativamente el error de *timing* y el error de exclusión, al posibilitar una captación temprana durante el embarazo. Se concluye que, superando desafíos como la calidad de los datos y la interoperabilidad en tiempo real, Uruguay tiene la oportunidad de utilizar sus registros para transformar la protección social hacia respuestas más oportunas y efectivas. También identifica la escalabilidad de esta herramienta a otros países de la región en los cuales se utiliza el SIP.

**Palabras clave:** aprendizaje estadístico, registros sociales, pobreza infantil, políticas sociales, vulnerabilidad.

Revista MERCOSUR de políticas sociales  
Revista MERCOSUL de políticas sociais  
Vol. 6 - 2025

ISSN 2523-0891 (impreso)  
ISSN: 2663-2047 (online)



Publicado en acceso abierto bajo la  
Licencia Creative Commons.

**Resumo:** Este artigo analisa o potencial da integração de registros sociais e da capacidade preditiva de modelos baseados em aprendizado estatístico para melhorar a oportunidade e a efetividade das políticas sociais direcionadas a gestantes e à primeira infância no Uruguai. Diante da persistente pobreza infantil e dos problemas de focalização (erros de exclusão, inclusão e *timing*),

o estudo propõe a integração do Sistema Informático Perinatal (SIP) e do Sistema de Informação Integrada da Área Social (SIAS) para aprimorar as respostas das Asignações Familiares (AFAM-PE). A partir dessa integração, desenvolve-se um modelo preditivo de vulnerabilidade social utilizando o algoritmo Stochastic Gradient Boosting (GBM). O modelo, treinado com dados das coortes 2015-2023, identifica com 85,7% de precisão as gestantes em lares vulneráveis, utilizando variáveis-chave levantadas na primeira consulta pré-natal. O estudo demonstra, por meio de um exercício com a coorte de 2023, que esse modelo permitiria reduzir significativamente o erro de timing e o erro de exclusão, ao possibilitar uma captação precoce durante a gravidez. Conclui-se que, superados desafios como a qualidade dos dados e a interoperabilidade em tempo real, o Uruguai tem a oportunidade de utilizar seus registros para transformar a proteção social em respostas mais oportunas e efetivas. Também se identifica a escalabilidade dessa ferramenta para outros países da região nos quais o SIP é utilizado.

**Palavras-chave:** aprendizagem estatístico, registros sociais, pobreza infantil, políticas sociais, vulnerabilidade.

## I. Introducción

La pobreza infantil es uno de los desafíos principales de Uruguay en materia de políticas sociales. De acuerdo a la nueva metodología de medición del Instituto Nacional de Estadística (INE), para 2024, la incidencia de la pobreza en el país se estimó en 17,3% y en 32,2% si se consideran únicamente los menores de 6 años. La sobrerrepresentación de niños, niñas y adolescentes en la pobreza es un fenómeno identificado hace ya varias décadas en Uruguay (Katzman, Filgueira, Rodríguez & Lijtenstein, 2001; Terra, 1979, 1988 y 1990).

El sesgo etario del bienestar se explica, en parte, por procesos estructurales como el cambio demográfico, las transformaciones del mundo del trabajo y la desigualdad de género (Filgueira, 2007; Rossel, 2013). Otra parte se debe a factores de economía política, principalmente a la forma en la que el estado, mercado y familias se combinan para proveer bienestar y protección ante los riesgos sociales, es decir, el tipo de régimen de bienestar (Esping-Andersen, 1990).

Uruguay ha sido clasificado como un universalismo estratificado (Filgueira, 1998; Midaglia et al., 2017), una variante regional de los regímenes conservadores-corporativos europeos, donde la protección social está asociada a la pertenencia al mercado de trabajo formal. En estas arquitecturas, la protección social está estratificada según ocupación y el cuidado de personas dependientes recae principalmente en las familias, lo que en los hechos significa, en las mujeres (Martínez Franzoni, 2008). Este esquema se edifica sobre la asunción, explícita o implícita, de un modelo de familia biparental de hombre proveedor que percibe un salario familiar (Sátyro & Midaglia, 2021).

El problema radica en que estas premisas sobre las cuales se edificaba el modelo de bienestar uruguayo dejaron de cumplirse -si es que alguna vez se cumplieron a cabalidad- y la resultante fue una arquitectura de protección fuertemente desfasada de la estructura de riesgos (Filgueira, 2007). No es casual que

la imagen prototípica de hogar vulnerable en Uruguay sea el compuesto por una madre soltera con muchos hijos/as. Si la protección social está anclada al empleo formal y los cuidados de personas dependientes recaen en las familias, no es difícil adivinar que las mujeres solteras con hijos encontrarán muchos obstáculos para insertarse en el mercado de trabajo. Así, a los déficits de ingresos se le suman los déficits de protección social.

Sin embargo, el paisaje de las políticas sociales no ha quedado congelado y algunos cambios han ido al encuentro de la estructura de riesgos (Rossel, 2013). En el marco de lo que se ha dado en llamar la expansión segmentada de las políticas sociales en América Latina a inicios del siglo XXI (Arza et al., 2022; Ferre, 2023; Garay, 2016), Uruguay ha incorporado modificaciones en su sistema de protección social. Ejemplos de ello son la expansión de los centros de cuidados de primera infancia, la incorporación de licencias parentales y la creación y expansión de las transferencias monetarias no contributivas con foco en hogares con menores de 18 años.

En efecto, los programas de transferencias monetarias condicionadas (PTMC) fueron el ejemplo paradigmático de políticas de expansión segmentada en la región (Arza et al., 2022). En Uruguay, las Asignaciones Familiares del Plan de Equidad (AFAM-PE), creadas en 2008 por Ley 18.227, son la principal transferencia monetaria no contributiva del país, con el objetivo de atender a los hogares con niños, niñas y adolescentes en situación de vulnerabilidad. Desde entonces se han consolidado como un pilar central del sistema de protección social uruguayo, alcanzando una amplia cobertura de la población infantil y adolescente y funcionando, en los hechos, como una de las principales herramientas de política pública para reducir la pobreza en este grupo etario (MIDES, 2017; Atuesta & Cecchini, 2017). Sin embargo, a pesar de su importancia, la prestación ha estado atravesada por problemas persistentes de suficiencia, vinculados a los bajos montos transferidos en relación con las líneas de pobreza, así como por limitaciones en la oportunidad y

efectividad de la llegada de los apoyos (Gómez Bonaglia, 2022).

Estos desafíos no son exclusivos de Uruguay, sino que forman parte de los dilemas estructurales de los programas focalizados en América Latina. Tal como señala la literatura, los mecanismos de focalización, aun cuando buscan asignar los recursos de manera eficiente hacia los hogares más vulnerables, generan errores de exclusión (hogares que cumplen criterios, pero quedan fuera) y de inclusión (hogares que no cumplen con dichos criterios y aun así reciben la prestación) (Cecchini & Madariaga, 2011; Coady et al., 2004). La mayoría de las investigaciones han centrado su atención en estos problemas (Gómez Bonaglia, 2022; Lavalleya & Tenenbaum, 2020 y 2022; Perazzo, Rivero & Vigorito, 2021; Stampini, Medellín & Ibarrarán, 2023) y también en la evaluación del cumplimiento de las condicionalidades (Baráibar Ribero, 2023; Rossel, Antía & Manzi, 2022). Algo menos estudiado es el problema del error de *timing*, es decir, cuando el Estado llega, pero no de manera oportuna. Esta cuestión está asociada directamente a las capacidades institucionales (Altman & Luna, 2012; Bertranou, 2015), dentro de las cuales la generación de información y su utilización para el diseño de políticas públicas es un asunto medular.

En efecto, un aspecto clave en el proceso de expansión de las políticas sociales ha sido el desarrollo de sistemas de monitoreo y evaluación de los programas de transferencias condicionadas en la región, que permitieron medir cobertura, impactos y desempeño de los mecanismos de focalización (BID, 2017; Cecchini & Martínez, 2011). Hoy, a esas capacidades históricas de relevamiento se suma la posibilidad de procesar grandes volúmenes de información mediante registros administrativos, y en particular registros sociales, que constituyen un insumo estratégico para mejorar el diseño e implementación de políticas públicas (Ibarrarán et al., 2017). La integración de dichos registros abre la puerta a explorar métodos estadísticos más flexibles, como los métodos de aprendizaje estadístico, que pueden colaborar en la reducción de los errores de inclusión, exclusión y *timing*, y, por tanto, fortalecer la respuesta estatal oportuna.

Desde el año 2020, el Fondo de Población de las Naciones Unidas en Uruguay (UNFPA) en acuerdo con el gobierno uruguayo, ha llevado adelante una serie de diagnósticos sobre los registros administrativos de mujeres gestantes y niñez temprana, enfatizando la necesidad de avanzar en la integración e interoperabilidad en tiempo real de las distintas fuentes de información existentes. Con base en dichos antecedentes, desde el año 2021 se está desarrollando un proyecto para evaluar las potencialidades de integrar el Sistema Informático Perinatal (SIP) y el Sistema de Información Integrada del Área Social (SIAS), ambos sistemas de información en la órbita del Estado. En ese marco, una línea de trabajo

se concentra en el desarrollo de modelos predictivos basados en aprendizaje estadístico a partir de la integración de los registros sociales de SIP y SIAS para predecir la vulnerabilidad de las personas gestantes con el objetivo de mejorar la respuesta estatal.

Este artículo tiene el objetivo de presentar algunos de los avances de esta línea, organizados de la siguiente manera. En primer lugar, se describe el proceso de integración de registros sociales del SIP y SIAS. En segundo lugar, se documenta el diseño de un modelo predictivo de vulnerabilidad social de personas gestantes a partir de aprendizaje estadístico con los registros sociales integrados. En tercer lugar, se presenta un ejercicio con la cohorte de nacimientos de 2023, para evaluar la potencialidad del modelo para mejorar las respuestas estatales oportunas y disminuir el error de exclusión. Finalmente, el artículo se cierra destacando algunos desafíos para la implementación.

## 2. Integración de registros sociales de salud y protección social: SIP y SIAS

El SIP fue creado por el Centro Latinoamericano de Perinatología – Mujer y Salud Reproductiva (CLAP/SMR) en 1983 y actualmente es un estándar de la Organización Panamericana de la Salud (OPS) para el registro clínico de la atención en servicios de salud dirigidos a la mujer, embarazo, parto y recién nacido, en la mayor parte de los países de la región<sup>1</sup>. Además de los datos clínicos, en el SIP también se releva información sociodemográfica de las mujeres desde la primera consulta antenatal. Uruguay comenzó a utilizarlo hace ya más de cuatro décadas y actualmente está implementando el SIP Plus, una versión web del sistema para registrar información en tiempo real.

El SIAS es un organismo interinstitucional conformado actualmente por más de 30 instituciones públicas y presidido por el Ministerio de Desarrollo Social (MIDES) desde 2010 (Ley N° 18.719), que integra datos de casi medio centenar de registros administrativos (Salud y Alimentación, Educación, Soluciones habitacionales, Seguridad y Protección Social, Trabajo y Empleo, Discapacidad y Atención integral)<sup>2</sup>. El intercambio sistemático de datos apunta a fortalecer las capacidades de diseño, gestión, evaluación y monitoreo de las políticas públicas, y de este modo contribuye a mejorar los niveles de eficiencia y eficacia en la gestión. Además del propósito de producir (recolectar, integrar y disponibilizar) información como insumo de calidad para las funciones propias de los organismos efectores de políticas sociales, se encuentra el de contribuir a la generación de conocimiento sobre el Estado y la

1 Por más información visitar <https://www.sipplus.org/>

2 Por más información visitar <https://www.gub.uy/ministerio-desarrollo-social/sias>

dinámica de provisión pública en materia de bienestar.

El SIP y el SIIAS son buen ejemplo de que el Estado uruguayo cuenta con registros administrativos y sistemas de información cuya integración permitiría dar un salto cualitativo en la mejora del diseño e implementación de políticas sociales para gestantes y niñez temprana. Sin embargo, a pesar de contar con este importante activo todavía persisten desafíos importantes en la comunicación de datos entre los sistemas de información de los organismos.

**Tabla 1.** Variables de SIP.

Abortos	Condición al egreso	Etnia	Parto / Aborto
Alimento al alta	Consultas prenatales	Fallece lugar de parto	Partos previos
Ant. nacidos muertos	Edad gestacional al parto	Fecha de egreso materno	Peso anterior
Ant. nacidos vivos	Edad materna	Fecha de nacimiento	Peso RN al egreso
Antecedentes Infertilidad	EG confiable por Eco < 20s	Fecha embarazo anterior	Referido
Antecedentes Violencia	Egreso RN - fecha	Fecha probable de parto	RN que viven
Años estudios mayor nivel	Embarazo planeado	Fecha última menstruación	Talla madre
Atendió Parto	Enfermedades RN	Fracaso método anti-conceptivo	Vive sola
Cervix PAP	Estado civil	Gestas previas	
Cesáreas	Estudios	Ligadura cordón	

Fuente: Elaboración propia.

Con el objetivo de visualizar la potencialidad de la integración entre SIP y SIIAS -siempre ajustados a la normativa vigente en materia de tratamiento de datos personales y clínicos-, se procedió a la construcción de tablas de datos anonimizados para todas las cohortes nacidas entre 2013 y 2023<sup>3</sup>. Este proceso consistió

3 La primera prueba de concepto se realizó con base en información para las cohortes 2011, 2012 y 2014. La siguiente fase permitió trabajar con las cohortes nacidas de 2013 a 2023. En el proceso de integración de registros se detectaron diferencias entre el número de nacimientos por cohorte registrado en las bases de SIP y la cantidad de nacimientos totales ocurridos en el país. Para dar noción de magnitudes, si se considera la cantidad de casos que pudieron integrar información de SIP y SIIAS sobre la cantidad de nacidos vivos para cada cohorte, se observa

en integrar datos provenientes de los registros de SIP (unas 50 variables correspondientes al Nivel I que se presentan en la Tabla 1) y de registros de SIIAS (variables sociodemográficas y de prestaciones sociales vinculadas a la protección social de gestantes y niñez temprana que se presentan en la Tabla 2). A partir de esa información, se consolidó una base anonimizada con 433.662 registros, con datos de la primera consulta antenatal de SIP, sociodemográficos y de prestaciones sociales<sup>4</sup>.

**Tabla 2.** Resumen de variables de SIIAS.

Prestaciones	Alimentación	Programa de Riesgo Nutricional (PRIN)
		Canasta de Emergencia Alimentaria
	Transferencias	TUS
		AFAM-PE
	Trabajo	Subsidio por maternidad
		Subsidio por desempleo
		Subsidio por enfermedad
	Educación	CAIF
		Educación Inicial ANEP
	Primera infancia	UCC
Sociodemográficas	Pensiones	Pensión Invalidez
	Salud	Cobertura de salud (tipo de institución prestadora)
		Departamento
		Localidad
		Proxy de trabajo formal
		Primera consulta antenatal

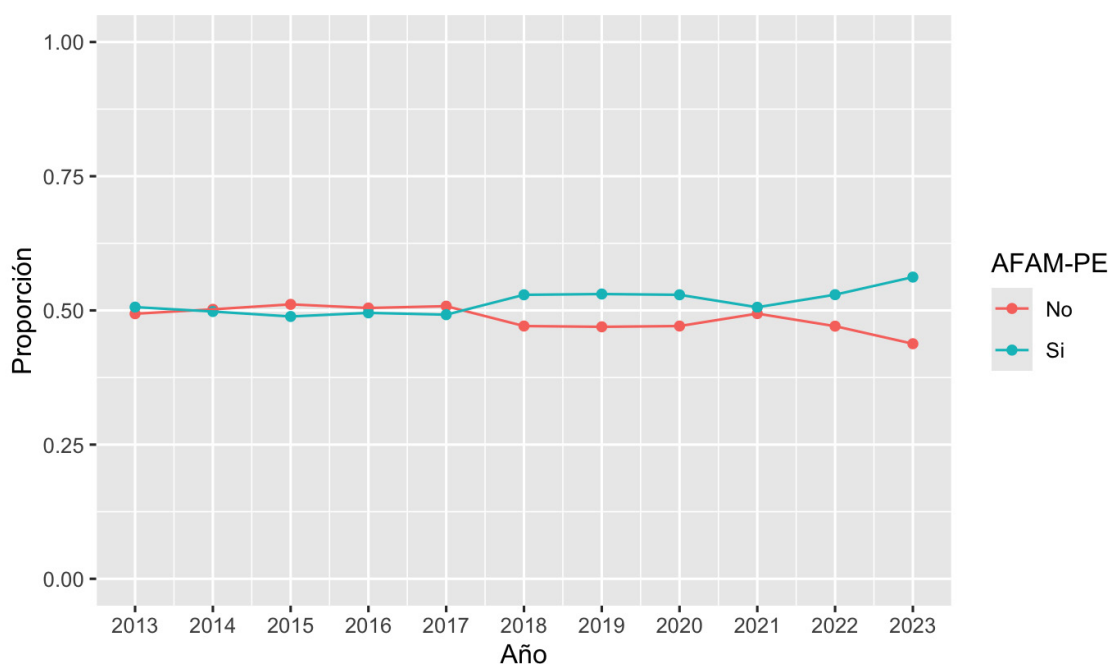
Fuente: Elaboración propia.

A partir de esta información en el marco general del proyecto mencionado, se procedió a la construcción de distintas herramientas de gestión y análisis, a saber: i) el diseño de una versión beta de Observatorio sobre Madres Gestantes e Infancia Temprana (con distintas posibilidades en relación al cruzamiento de datos, construcción de indicadores y visualización), ii) la creación de un sistema de novedades de embarazos entre el Ministerio de Salud y el Ministerio de Desarrollo Social, iii) estimaciones basadas en micro-simulaciones sobre el impacto de los cambios en distintos componentes de la matriz de protección

que la cobertura de información se ubicó en el entorno del 75% para las primeras cohortes del periodo considerado y en el entorno del 95% para las cohortes más recientes.

4 La integración de los datos requirió del siguiente trabajo: limpieza de datos anómalos, homogeneización de nombres de variables, recodificación, identificación de datos faltantes y creación de variables resumen a partir de la combinación de variables de la base integrada.

**Figura 1.** Evolución de la proporción de beneficiarios de AFAM-PE según cohorte de nacimiento.



Fuente: Elaboración propia.

social en la situación de los hogares pobres con niños/as; iv) el desarrollo de un predictor de vulnerabilidad, construido en base a métodos de aprendizaje estadístico, que agrega al sistema de novedades la posibilidad de generar una alerta temprana, y que se presenta en el apartado siguiente.

### 3. Modelo predictivo de vulnerabilidad con aprendizaje estadístico

Con el objetivo de predecir la vulnerabilidad o riesgo social de las personas gestantes se exploraron métodos de aprendizaje estadístico supervisado usando como insumo la base de datos integrada con información proveniente de SIP y de SIAS (2013-2023). Para el ajuste de estos modelos primero fue necesario definir la variable de respuesta que mida vulnerabilidad. Posteriormente se realizó un análisis exploratorio de datos para evaluar la calidad de los mismos, entender sobre el problema de interés e identificar posibles variables predictoras asociadas con la respuesta para incorporar en el modelo.

Se definió la variable de respuesta en base a la percepción de AFAM-PE, de forma que una persona gestante es vulnerable si pertenece a un hogar que ha recibido la prestación alguna vez en el periodo analizado. Esto implica que los hogares tengan menores de edad, hayan sido visitados y se les haya aplicado un formulario de ingreso al programa y superado el umbral de elegibilidad establecido en la definición de la población objetivo a partir del cálculo del Índice de

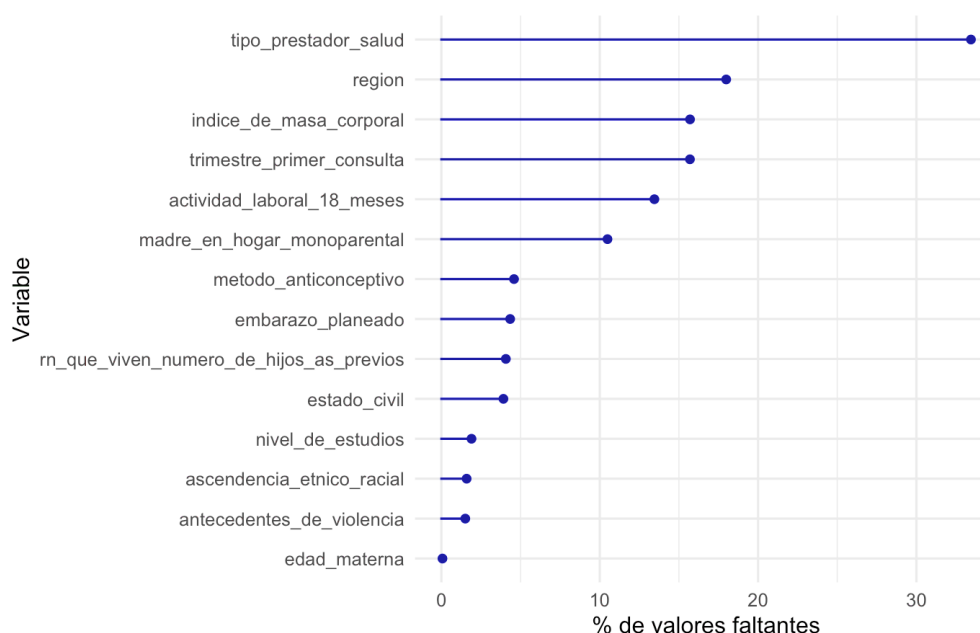
Carencias Críticas (ICC)<sup>5</sup>. Como se observa en la Figura 1, considerado el universo de cohortes de nacidos (2013-2023), la proporción de niños/as en hogares beneficiarios de AFAM-PE se ha mantenido estable en el tiempo con un leve incremento en los últimos años y representa aproximadamente la mitad de los casos en el período analizado.

La distribución y balance de la cobertura en un escenario en que se cuenta con información de todas las cohortes de nacidos durante un período amplio de tiempo, hace que dicha información pueda ser utilizada como indicador de vulnerabilidad del hogar. Contar con este indicador de vulnerabilidad permite utilizarlo como variable de respuesta, o dependiente, en un modelo que tiene como objetivo predecir si una persona (en nuestro caso persona gestante) pertenece a un hogar vulnerable o no. Es decir que el resultado de los modelos que se ensayaron tienen el objetivo de predecir la vulnerabilidad de un hogar (y por tanto

5 El Índice de Carencias Críticas (ICC) es el instrumento que se utiliza para dar cuenta del nivel de vulnerabilidad de los hogares y evaluar si estos forman parte de la población objetivo. Fue creado en el año 2008 en convenio entre el Mides y la IECON-Udelar, y como se menciona en un informe interno que detalla su diseño (DAES-Dinem y Mides, 2013): «La metodología propuesta para captar a la población objetivo combina un abordaje multidimensional de la situación socioeconómica de los hogares, con la concepción más clásica de pobreza asociada a los ingresos. Como resultado surge el ICC, un indicador que sintetiza información respecto a la situación de los hogares en distintas dimensiones entre las que no se incluye el ingreso, aunque este sí se utiliza como referencia en la elaboración de dicha síntesis».



**Figura 2.** Porcentaje de valores faltantes por variable entre 2013 y 2023.



Fuente: Elaboración propia.

del contexto donde el/la niño/a nacerá), a partir de la asignación de una probabilidad que permite orientar la protección social hacia aquellas situaciones prioritarias.

En la primera fase del proceso, se realizó una exploración predictiva inicial considerando el universo de 433.662 casos correspondientes a niños/as que nacieron entre 2013 y 2023 en Uruguay. En una primera instancia se probaron modelos considerando 24 variables clínicas y sociodemográficas del SIP y SIAS<sup>6</sup>. Este subconjunto de variables fue seleccionado considerando la calidad de los datos, el conocimiento experto sobre su relevancia para el problema planteado y el análisis exploratorio de estos.

Luego, se entrenaron modelos de aprendizaje estadístico supervisado para el problema de clasificación binaria, utilizando la variable dependiente de pertenencia a un hogar beneficiario de AFAM-PE en los datos de entrenamiento (70% de las observaciones) y la evaluación del modelo se realiza en el conjunto de testeo (30% de las observaciones). Los modelos considerados para comparar su desempeño predictivo están dentro de las siguientes familias: Generalized

Linear Models (GLM), Distributed Random Forest (DRF), Gradient Boosting Machine (GBM) y XGBoost<sup>7</sup>.

Esta primera exploración permitió evaluar resultados iniciales y métricas de importancia de las variables predictoras utilizadas, lo cual resultó un insumo relevante para la siguiente fase de selección de variables priorizando aquellas que fueran relevadas en la primera consulta antenatal y no post parto, de modo que el modelo permite la captación oportuna desde el inicio de la gestación. A su vez, en base a la exploración de datos faltantes se decidió excluir los años 2013 y 2014 por su cantidad de faltantes. Por ejemplo, el porcentaje de faltantes para la variable `tipo_prestador_salud` es del 100% en ambos años y para `región` es aproximadamente 30% que es mayor al porcentaje considerando todos los años como se presenta en la Figura 2.

Por último, en la selección se tuvieron en cuenta aquellas que en la exploración inicial presentaban una mayor asociación con la respuesta. En la Tabla 3 se

<sup>6</sup> Las variables consideradas en este primer ensayo fueron: Edad materna, ascendencia étnico racial, nivel de estudios, estado civil, madre en hogar monoparental, antecedentes de infertilidad, antecedentes de violencia, gestas previas, abortos, partos previos, cesáreas, número de hijos/as previos, antecedentes de RN muertos, embarazo planeado, edad gestacional al parto, nivel de instrucción madre, cantidad de consultas prenatales, método anticonceptivo, grupos de edad materna, índice de masa corporal, trimestre de primera consulta, actividad laboral 6 meses, actividad laboral 18 meses, y región.

<sup>7</sup> Los modelos fueron entrenados utilizando la librería `h2o` (Fryda et al., 2024), que permite realizar un entrenamiento eficiente y escalable de modelos de aprendizaje automático. Esta librería presenta algunas ventajas en cuanto al preprocesamiento de datos como ser la imputación de datos faltantes, recodificación y estandarización a su vez el ajuste de modelos individuales se realiza mediante validación cruzada y realiza una búsqueda eficiente de los hiperparámetros de los modelos seleccionados. El resultado de todos los modelos queda disponible y ordenado en base a la métrica de rendimiento seleccionada optimizando el punto de corte para la probabilidad de la variable de respuesta. Adicionalmente permite exportar el modelo seleccionado para su puesta en producción (Fryda et al., 2024).

**Tabla 3.** Diseño y entrenamiento de modelo final.

<b>1. Definición de universo</b>	Cohortes 2015-2023 de niños/as nacidos en Uruguay (CNV-SIP) N= 339119
<b>2. Variables seleccionadas</b>	<b>Variable dependiente</b> (categórica): Pertenencia a hogar AFAM-PE (al menos un mes de percepción) Si = 174675 / No = 164444 <b>Variables predictoras (14):</b> Actividad laboral en los últimos 18 meses Número de hijos/as previos (RN que viven) Nivel de estudios Edad materna Estado civil Región Índice de masa corporal (IMC) Tipo de prestador de salud Embarazo planeado Ascendencia étnico-racial Uso de método anticonceptivo Madre en hogar monoparental Antecedentes de violencia Trimestre de la primera consulta prenatal
<b>3. Modelado y evaluación:</b> a. Construcción de muestras de entrenamiento (70%) y testeo (30%) b. Entrenamiento de modelos de clasificación en muestra de entrenamiento (validación cruzada e iteración de modelos) con librería H2o (R CRAN) c. Evaluación de modelo en muestra de testeo	

Fuente: Elaboración propia.

presenta el flujo de trabajo para el ajuste y selección del modelo final y se presentan a su vez las 14 variables seleccionadas para entrenar el modelo final. Con este nuevo subconjunto de variables se entrenaron 127 modelos dentro de las familias antes mencionadas en el conjunto de entrenamiento y ajuste de hiperparámetros basado en técnicas de validación cruzada<sup>8</sup>. La selección del tipo de modelo final utilizado consideró criterios de precisión y de interpretabilidad.

El modelo que tuvo el mejor rendimiento predictivo, considerando las 14 variables predictoras, fue el Stochastic Gradient Boosting (GBM)<sup>9</sup> (Friedman, 2002), en línea con antecedentes en el mismo campo de exploración asociado a la orientación de la focalización de transferencias monetarias (Gómez Bonaglia & da Silva, 2024; Noriega-Campero et al., 2020).

En términos de resultados, el modelo entrenado y ajustado considerando datos de las cohortes de nacimientos entre los años 2015 y 2023, cuenta con una

precisión de 85,7% en el conjunto de testeo, es decir considerando observaciones que no fueron utilizadas para entrenar el modelo. Así también se ha optimizado el valor de corte de la probabilidad en base a la métrica FI que mide la calidad global del modelo para predecir los casos positivos. El modelo optimizado, en lugar de considerar el umbral usual de clasificación binaria (0.5), utiliza como punto de corte una probabilidad mayor o igual a 0.41 para clasificar un caso como vulnerable. En la Tabla 4 se presenta la matriz de confusión con los datos de testeo para el mejor modelo predictivo. Otro dato que se desprende de la matriz de confusión es que un 57.2 % del conjunto de testeo se considera en situación de vulnerabilidad según nuestro modelo predictivo, y de ellos hay un 17,5% (10188 niños/as) que no cuentan con la prestación AFAM-PE.

**Tabla 4.** Matriz de confusión (Conjunto de testeo).

Tabla 11. Matriz de confusión (Se, James de 0,8555).				
		Predicción		
		No	Sí	Total
AFAM-PE	No	39061	10188	49249
	Sí	4314	47999	52313
	Total	43375	58187	101562
	Precisión	85,7%		

Fuente: Elaboración propia.

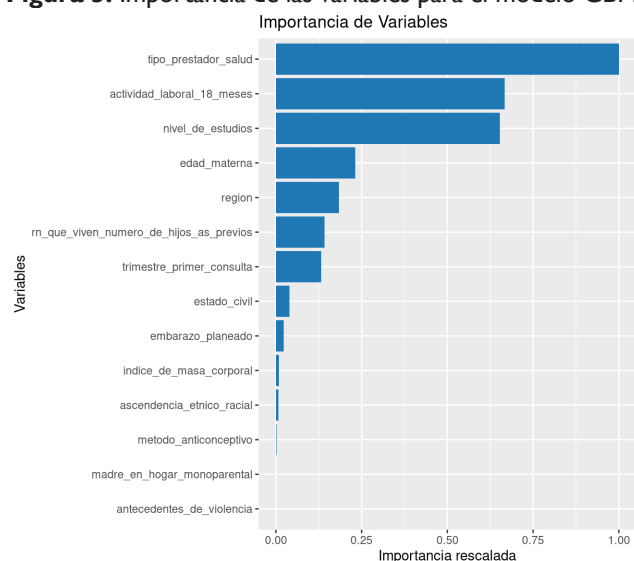
La Figura 3 muestra la importancia<sup>10</sup> de las variables para el modelo GBM, podemos observar que

8 La validación cruzada es una técnica que evalúa la capacidad predictiva del modelo repitiendo el entrenamiento en distintos subconjuntos de datos para estimar su rendimiento fuera de muestra y evitar el sobreajuste. Toda la implementación de los modelos fue realizada utilizando la librería h2o (Fryda et al., 2024).

9 Los métodos de tipo boosting son métodos de agregación al igual que bagging que combina modelos individuales débiles, árboles de clasificación en este caso, cuyo poder predictivo individualmente es limitado para construir un predictor general más fuerte. A diferencia del bagging, donde los modelos individuales se entrenan de forma independiente, en el boosting los modelos se entrenan secuencialmente, de manera que cada modelo se enfoca en los errores cometidos por el anterior. Los métodos de tipo boosting reducen el sesgo al refinar iterativamente modelos base con baja varianza y alto sesgo mediante actualizaciones secuenciales (Friedman, 2002).

10 La importancia de una variable refleja cuánto contribuye al desempeño predictivo del modelo.

**Figura 3.** Importancia de las variables para el modelo GBM.



Fuente: Elaboración propia.

las 5 variables más relevantes son tipo de prestador de salud, actividad laboral en los últimos 18 meses, nivel de estudios finalizados, edad materna y región.

En la Figura 4 se observa la relación entre la predicción del modelo para el conjunto de testeo y la cobertura efectiva de AFAM-PE, a partir de visualizar la probabilidad de vulnerabilidad predicha por el modelo para cada grupo de interés. La línea punteada es el umbral óptimo de probabilidad para este modelo predictivo basado en FI ( $\text{prob} = 0.41$ ).

Una manera de aproximarse a la evaluación de la precisión del modelo predictivo, es cotejarla con la performance de los actuales mecanismos de focalización en base a trabajos recientes. Para la cohorte 2019, Lavalleya y Tenenbaum (2022), estimaron un error de exclusión del 39,7% y señalan que esa métrica se ha

mantenido estable en los años previos. Para la cohorte 2023, el error de exclusión de AFAM-PE se ha estimado en 27.5% del universo de población que teóricamente cumple con las condiciones de elegibilidad, es decir, hogares en situación de vulnerabilidad con menores de 18 años (DINTAD-MIDES, 2025). Si bien las estimaciones no son estrictamente comparables, se considera que la precisión del 85,7% del modelo predictivo, presenta una buena performance, además de tener la ventaja de que es posible obtener la probabilidad al momento de la primera consulta antenatal.

El desarrollo del modelo predictivo presenta importantes desafíos. Actualmente, el trabajo está enfocado en la mejora en la calidad de los datos para obtener mejores resultados en los modelos predictivos. A su vez, aunque la implementación realizada permite predecir observaciones con datos faltantes, ya que se realizan imputaciones de los mismos, es necesario un análisis más profundo del efecto de distintos métodos de imputación en la predicción del modelo.

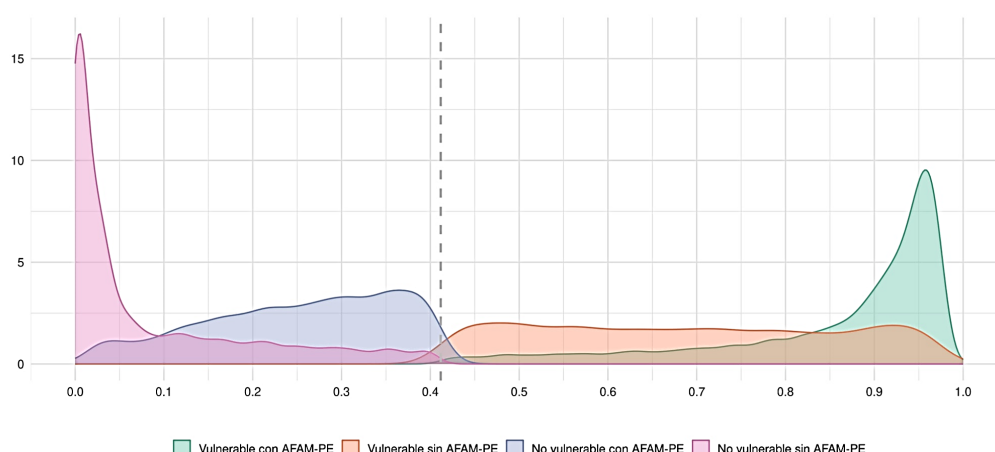
#### 4. Potencialidad del modelo predictivo para mejorar las respuestas estatales

En esta sección se presenta un ejercicio para visualizar la potencialidad del modelo predictivo detallado anteriormente, para mejorar las respuestas estatales para gestantes y niñez temprana. El ejercicio consiste en analizar el problema de *timing* en la percepción de AFAM-PE en la cohorte 2023, la identificación de oportunidades de captación temprana desaprovechadas (consultas antenatales) y el impacto que podría tener la implementación del modelo predictivo para reducir el problema de *timing* y de exclusión.

##### Descripción del problema de timing (cohorte 2023)

El problema del *timing* de percepción representa una falla en el mecanismo de selección de una política

**Figura 4.** Distribución de probabilidad de Modelo predictivo (GBM) por grupos - para el conjunto de testeo.



Fuente: Elaboración propia.



**Tabla 5.** Resumen percepción AFAM-PE según momento (Cohorte 2023).

Momento	Percepción de AFAM-PE			
	No		Si	
6 meses antes del nacimiento	23857	76,0%	7524	24,0%
Nacimiento	18871	60,1%	12507	39,9%
6 meses después del nacimiento	17789	56,7%	13502	43,3%
12 meses después del nacimiento	17252	55,0%	14118	45,0%

Fuente: Elaboración propia.

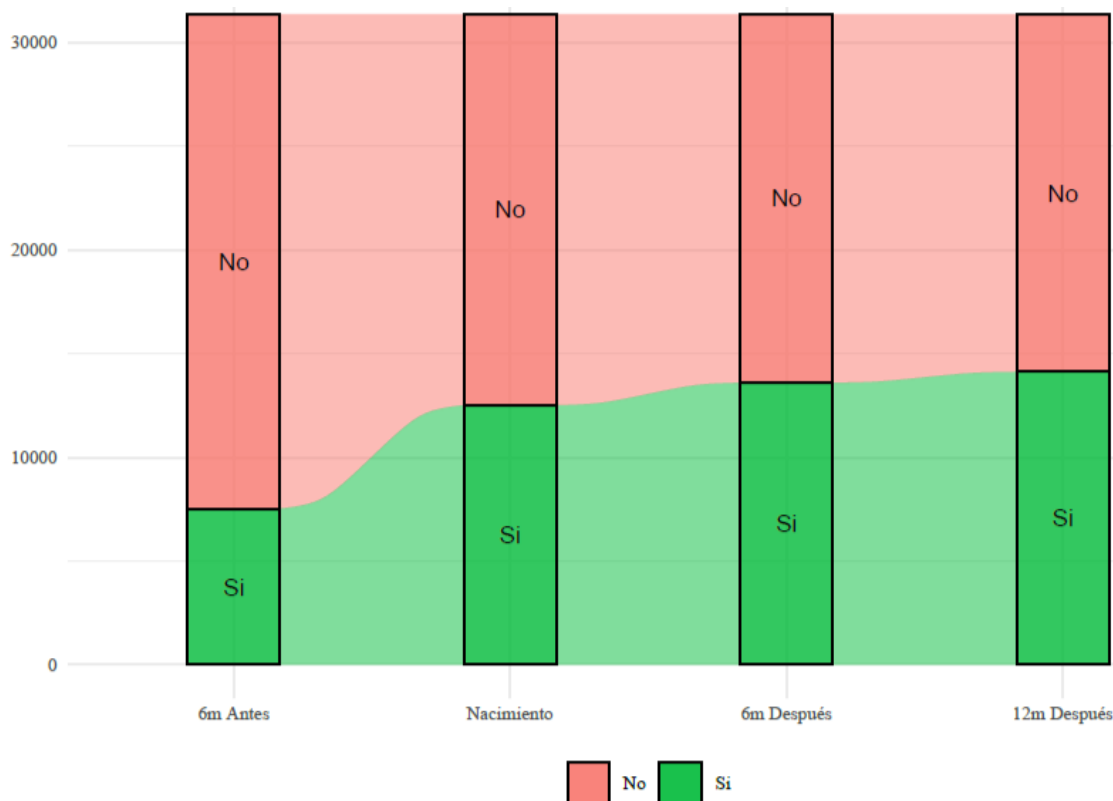
que deriva en una demora en la percepción del beneficio por parte de hogares o personas que cumplen los criterios de elegibilidad de una prestación. Este problema implica que personas que deberían acceder a una prestación social en determinado momento, lo hacen más tarde.

Para realizar la descripción del problema de *timing* de percepción de AFAM-PE se trabajó con los datos anonimizados de la cohorte de nacidos en 2023. Estos datos permiten conocer para cada nacimiento el dato de percepción mensual de la prestación AFAM-PE. Se seleccionaron cuatro momentos en el tiempo: seis meses antes del nacimiento; mes de nacimiento; seis meses posteriores al nacimiento; doce meses posteriores al nacimiento. A partir de ello, se calcularon

las diferencias entre la percepción en cada momento. El principal supuesto que se asume es que los hogares no modifican su situación entre el embarazo y el año de nacido del niño/a. Es decir, se asume que si un hogar es beneficiario de AFAM-PE a los doce meses del nacimiento, también cumpliría con las condiciones de elegibilidad durante el embarazo.

En la Tabla 5 se presentan los porcentajes y cantidades de personas según percepción de AFAM-PE en cada momento del tiempo. Se constata una importante diferencia (21 puntos porcentuales o 6594 personas) entre la cobertura durante la gestación (seis meses antes del nacimiento) y a los doce meses luego del nacimiento. Una diferencia más reducida se constata al comparar la cobertura al nacimiento y a los doce

**Figura 5.** Flujo de percepción AFAM-PE según momento - Cohorte 2023.



Fuente: Elaboración propia.

meses del nacimiento (5.1 puntos porcentuales o 1611 personas). Una diferencia aún más reducida se constata al comparar la diferencia de cobertura a los seis meses de nacimiento y a los doce meses del nacimiento (2.7 puntos porcentuales o 616 personas).

A partir de los registros administrativos disponibles de la cohorte 2023, considerando la ventana temporal que va desde los seis meses antes del nacimiento y los doce meses posteriores al nacimiento, estimamos que el estado tiene un problema de *timing* que abarca (con diferentes temporalidades) al 21% de la cohorte. Esta brecha temporal en la percepción de la prestación, cuya visualización gráfica se presenta en la Figura 5, es la que podría reducirse a su mínima expresión integrando SIP y SIAS.

#### Oportunidades de captación temprana (cohorte 2023)

El SIP proporciona información relevante que permite dar cuenta de las oportunidades desaprovechadas para captar situaciones de vulnerabilidad de manera oportuna. La información se construye cotejando el momento de percepción de AFAM-PE y dos indicadores elaborados a partir de los registros de SIP: el trimestre de la primera consulta antenatal y la cantidad de consultas antenatales.

En la Tabla 6, se puede visualizar un porcentaje importante de situaciones de percepción tardía de AFAM-PE (un 16,54% de la cohorte 2023 si se suma quienes perciben al nacimiento, a los seis meses o a los doce meses desde el nacimiento), en casos de embarazos captados en el primer trimestre. Estos datos permiten evidenciar de forma empírica situaciones que potencialmente podrían ser captadas de manera oportuna. Para la cohorte 2023, aproximadamente el 86% de los nacimientos son de embarazos captados en el primer trimestre y aumenta al 97% si se incluye el segundo trimestre. Esto supone una ventana de oportunidad clave para poder captar situaciones de vulnerabilidad de manera temprana.

En la Tabla 7, se presenta el porcentaje de situaciones de percepción tardía de AFAM-PE (un 19,04% de la cohorte 2023 si se suma quienes perciben al nacimiento, a los seis meses o a los doce meses desde el nacimiento), en casos de madres que concurren a 5 o más consultas antenatales. Esto supone que el sistema de protección social tuvo 5 o más oportunidades de lograr una captación temprana de una situación de vulnerabilidad, a partir de registros del sistema de salud

Finalmente es importante señalar que de acuerdo a Estadísticas Vitales del Ministerio de Salud Pública

Tabla 6. Combinaciones de percepción AFAM-PE de embarazos captados en el primer trimestre (Cohorte 2023).

6 meses antes del nacimiento	Nacimiento	6 meses después del nacimiento	12 meses después del nacimiento	Cantidad	% Cohorte
No	No	No	No	13437	42.82
Si	Si	Si	Si	5175	16.49
No	Si	Si	Si	3764	<b>11.99</b>
No	No	Si	Si	958	<b>3.05</b>
No	No	No	Si	471	<b>1.5</b>

Fuente: Elaboración propia.

Tabla 7. Combinaciones de percepción AFAM-PE de embarazos con 5 o más controles antenatales (Cohorte 2023).

6 meses antes del nacimiento	Nacimiento	6 meses después del nacimiento	12 meses después del nacimiento	Cantidad	% Cohorte
No	No	No	No	14550	46,37
Si	Si	Si	Si	5980	19,06
No	Si	Si	Si	4302	13,71
No	No	Si	Si	1111	3,54
No	No	No	Si	562	1,79

Fuente: Elaboración propia.

**Tabla 8.** Resumen percepción AFAM-PE según momento (Cohorte 2023: Pred. Si  $\geq 0.5$ ).

Momento	Percepción de AFAM-PE			
	No		Si	
6 meses antes del nacimiento	10077	65,2%	5381	34,8%
Nacimiento	6682	43,2%	8776	56,8%
6 meses después del nacimiento	5923	38,3%	9535	61,7%
12 meses después del nacimiento	5582	36,1%	9876	63,9%

Fuente: Elaboración propia.

(MSP), en Uruguay el 100% de los nacimientos son asistidos por personal de salud capacitado y más del 99% de los partos ocurren en hospitales. Estos datos dan cuenta de que todos los nacimientos de cada cohorte hacen contacto con el sistema de salud, al momento del nacimiento o antes.

*Potencialidades del modelo predictivo: error de timing y error de exclusión*

El modelo predictivo construido para predecir la probabilidad de ser beneficiario/a de AFAM-PE a partir de la información que puede obtenerse al momento de la primera consulta antenatal combinando información SIP y del SIAS, permitiría reducir de manera notable el error de *timing* de percepción e incluso podría avanzar de manera sustantiva en reducir el error de exclusión. Si el error de *timing* supone que el estado llega, pero no de manera oportuna, el error de exclusión implica que el estado no llega a la población beneficiaria.

En este ejercicio se replica el análisis descriptivo de percepción de AFAM-PE para aquellos nacimientos de la cohorte 2023 que una calibración específica del modelo predictivo arroja que son vulnerables. Para realizarlo, se entrenó nuevamente el modelo predictivo pero sin los datos de la cohortes 2023, dado que el objetivo era predecir la vulnerabilidad para la totalidad de esta cohorte. Asimismo, a los efectos del ejercicio se trabajó con el umbral usual de clasificación binaria (probabilidad de 0.5) como criterio de elegibilidad. Es decir, a todos aquellos nacimientos cuya probabilidad de ser beneficiarios de AFAM-PE era igual o mayor a 0.5, se los considera cómo población que debería ser beneficiaria.

En la Tabla 8 se presentan los datos que muestran claramente la brecha temporal y de cobertura que podría ser eliminada o significativamente reducida. En este caso la brecha del *timing* de cobertura entre la gestación y los doce meses luego del nacimiento es de 29.1 puntos porcentuales (4495 personas). Por otra parte, el modelo predictivo también permitiría aproximar el error de exclusión y reducirlo. En nuestro ejercicio con la cohorte 2023, si se asume que todo

niño/a cuya probabilidad de ser beneficiario de AFAM-PE es igual o mayor que 0.5, obtenemos que a los 12 meses de nacidos hay un 36.1% (5582 personas) que no percibe AFAM-PE. Este valor es consistente con el error de exclusión identificado por Tenenbaum y Lavalleja (2022) para la cohorte 2019, aunque algo mayor que el identificado por MIDES-DINTAD (2025) para la cohorte 2023.

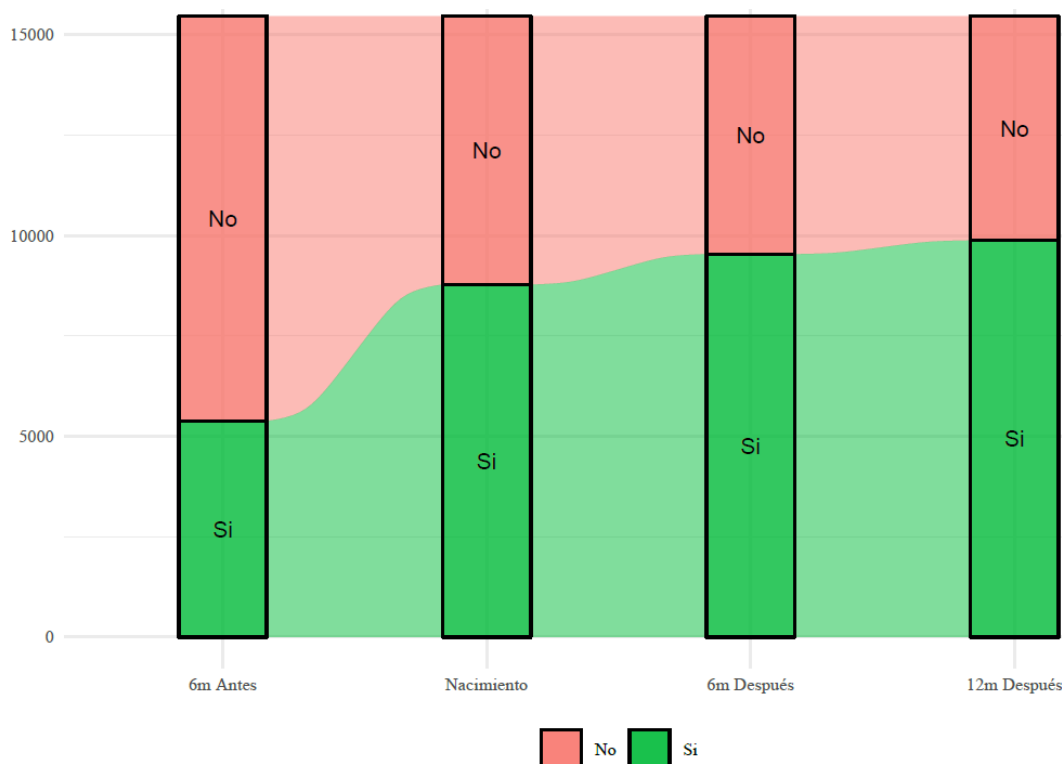
En la Figura 6 se pueden visualizar el error de *timing* siguiendo la evolución del área verde y el error de exclusión en el área roja de la última barra (12m después), correspondiente a la categoría “No”. Si la carga de datos de SIP fuese perfectamente oportuna (en el momento de las consultas) y completa (sin faltantes), y se tomara la decisión de dar el beneficio a todas las madres gestantes cuya probabilidad de ser beneficiaria de AFAM-PE supere el 0.5 (u otro umbral definido), sería posible eliminar virtualmente el error de *timing* y también reducir de manera importante el error de exclusión durante la gestación, dado el alto porcentaje de captación temprana de embarazos y la cantidad de consultas antenatales. El error podría reducirse a prácticamente cero al momento del nacimiento.

## 5. Desafíos

Este artículo presentó evidencia de la potencialidad que tiene Uruguay para mejorar las respuestas estatales a gestantes y niñez temprana utilizando modelos predictivos de vulnerabilidad con aprendizaje estadístico a partir de la integración de registros sociales (SIP y SIAS). En términos sustantivos, lo que subyace es la oportunidad que tiene Uruguay de utilizar sus buenos registros sociales y la oportuna captación de embarazos para mejorar las respuestas de protección social para atacar la pobreza infantil, uno de los desafíos más importantes que tiene en materia de políticas sociales. Si bien ya es posible implementar un sistema de estas características, es oportuno identificar algunos de los principales desafíos para que su implementación práctica sea exitosa.

Primero, la necesidad de carga de datos oportuna y completa. Como se especificó en los apartados de

**Figura 6. Flujo de percepción AFAM-PE según momento - Cohorte 2023: (Pred. Si  $\geq$  0.5).**



Fuente: Elaboración propia.

integración de la base y construcción del modelo predictivo, persisten algunos problemas de calidad de datos y oportunidad de su carga, pese a que hubo avances significativos. Tres parecen ser los problemas principales. Primero, existe una brecha entre los nacimientos ocurridos y los ingresados en SIP (en las últimas cohortes, es inferior al 5% de los nacimientos). Segundo, en muchos casos la carga del SIP se realiza luego del nacimiento y no al momento de la primera consulta. Tercero, la no completitud de los datos de algunas de las variables del modelo predictivo. En todos estos planos, se han realizado avances significativos y estos continúan realizándose a medida que el SIP Plus (la aplicación web del SIP) se implementa en los diferentes centros de salud materno-infantil. Por otra parte, también existieron decisiones del MSP estableciendo la obligatoriedad de carga de SIP en forma oportuna, aunque el *enforcement* de esta medida no es riguroso aún. En tanto, la completitud de los datos viene mejorando de forma continua en cada nueva cohorte.

Segundo, interoperabilidad en tiempo real. La potencia de estas herramientas está directamente vinculada a la oportunidad con la que se cuenta la información. Las mejoras de respuestas oportunas parten de la base que la información crítica para realizar la predicción de vulnerabilidad debe estar disponible al momento de la primera consulta. Esa es la finalidad del sistema de novedades generado entre MSP y MIDES para intercambiar de manera oportuna información de SIP y SIIAS acompañado de una alerta temprana. El SIP

ha evolucionado de manera tal que puede integrarse a las más diversas plataformas e interoperar con distintos sistemas de gestión y registros clínicos de cualquier prestador de salud y así obtener un repositorio de información con actualización al instante cuyos servidores se alojan en el organismo rector de las políticas de salud (MSP).

Tercero, la definición respecto al rol que un modelo predictivo de estas características podría tener es algo que merece discusión. Las alternativas polares parecen ubicarse entre servir como alerta temprana o convertirse en criterio para otorgar una prestación de carácter preventiva, que podría complementarse con un mecanismo de certificación posterior. Un aspecto clave y oportunidad de mejora es la combinación de este tipo de herramientas con las estrategias que ha implementado MIDES para lograr una mayor penetración territorial y una reducción de los tiempos de espera para las visitas de potenciales beneficiarios de AFAM-PE, como, por ejemplo, la ventanilla única para acercar los puntos de trámites a la población beneficiaria.

Cuarto, la definición de un umbral de probabilidad es una decisión que siempre encierra algo de arbitrariedad. En el entrenamiento del modelo predictivo con las cohortes de nacimiento entre 2015 y 2023 se presentó una métrica optimizada que se ubica en una probabilidad del 0.41. En el ejercicio realizado para la cohorte 2023, lo establecimos en 0.5. Esto sirve para dar cuenta que el umbral puede ser definido con arreglo a diferentes criterios, por ejemplo,

un nivel de vulnerabilidad absoluto definido, un nivel de cobertura (en cantidad o en porcentaje de la cohorte) o incluso un parámetro presupuestal.

En tanto, los ejercicios realizados en este artículo con AFAM-PE en Uruguay son sólo un ejemplo de la potencialidad que tiene la elaboración de modelos predictivos con aprendizaje estadístico a partir de registros sociales para mejorar las políticas sociales dirigidas a maternidad y niñez temprana. Estas herramientas tienen un enorme potencial de escalabilidad y pueden utilizarse para una batería de políticas de protección social, por ejemplo, sistema de cuidados, educación, servicios públicos subsidiados, atención prioritaria en salud, transporte, conectividad a internet, acceso a créditos, etc.

Además, la escalabilidad también puede pensarse en clave regional. Dado que el SIP es un estándar de OPS que se utiliza en buena parte de América Latina y el Caribe, su vinculación con los sistemas de información de protección social podría impulsar experiencias similares en otros países de la región.

Finalmente, es preciso remarcar que los desafíos que tienen las matrices de protección social de la región para revertir las condiciones de vulnerabilidad de niños, niñas y adolescentes, requieren -además de mejorar los mecanismos de focalización- una sustantiva mejora en la suficiencia de las prestaciones, así como el fortalecimiento y diversificación de los apoyos a las familias, para posibilitar el acceso a estándares mínimos de bienestar y la apertura de una estructura de oportunidades de movilidad social que hoy parece estar trunca.

## Referencias

- Altman, D., & Luna, J. P. (2012). Introducción: El Estado latinoamericano en su laberinto. *Revista de Ciencia Política (Santiago)*, 32(3), 521–543. <https://doi.org/10.4067/S0718-090X2012000300001>
- Arza, C., Castiglioni, R., Martínez Franzoni, J., Niedzwiecki, S., Pribble, J., & Sánchez-Ancochea, D. (2022). *The political economy of segmented expansion: Latin American social policy in the 2000s*. Cambridge University Press.
- Atuesta, B., & Cecchini, S. (2017). *Programas de transferencias condicionadas en América Latina y el Caribe: Tendencias de cobertura e inversión (CEPAL - Serie Políticas Sociales N° 224)*. CEPAL.
- Bertranou, J. (2015). Capacidad estatal: revisión del concepto y algunos ejes de análisis y debate (dossier). *POSTData: Revista de Reflexión y Análisis Político*, 20(2), 501-530.
- Banco Interamericano de Desarrollo, BID. (2017). *Así funcionan las transferencias condicionadas. Buenas prácticas a 20 años de su implementación*.
- Cecchini, S., & Madariaga, A. (2011). *Programas de transferencias condicionadas: Balance de la experiencia reciente en América Latina y el Caribe (Cuadernos de la CEPAL, N° 95)*. CEPAL.
- Cecchini, S., & Martínez, R. (2011). *Protección social inclusiva en América Latina: Una mirada integral, un enfoque de derechos (Libros de la CEPAL, N° 111)*. CEPAL.
- Coady, D., Grosh, M., & Hoddinott, J. (2004). *La focalización de las transferencias en los países en desarrollo: Revisión de lecciones y experiencias*. Banco Mundial.
- Esping-Andersen, G. (1990). *The three worlds of welfare capitalism*. Polity Press.
- Ferre, J. C. (2023). Social policy expansion and segmentation in the first two decades of the 21st century in Latin America. *International Journal of Social Welfare*, 32(4), 486–505. <https://doi.org/10.1111/ijsw.12579>
- Filgueira, F. (1998). El nuevo modelo de prestaciones sociales en América Latina: eficiencia, residualismo y ciudadanía estratificada. En *Centroamérica en reestructuración: Ciudadanía y política social* (pp. 71–116). Facultad Latinoamericana de Ciencias Sociales (FLACSO).
- Filgueira, F. (2007). Cohesión, riesgo y arquitectura de protección social en América Latina (No. 6141). CEPAL.
- Friedman, J. H. (2002). Stochastic gradient boosting. *Computational statistics & data analysis*, 38(4), 367–378.
- Fryda, T., LeDell, E., Gill, N., Aiello, S., Fu, A., Candel, A., Click, C., Kraljevic, T., Nykodym, T., Aboyoun, P., Kurka, M., Malohlava, M., Poirier, S., & Wong, W. (2024). *h2o: R Interface for the 'H2O' Scalable Machine Learning Platform (Versión 3.44.0.3)* [Paquete de R]. H2O.ai. <https://h2o.ai>
- Garay, C. (2016). *Social policy expansion in Latin America*. Cambridge University Press.
- Gómez Bonaglia, E. (2022). Programas de transferencias condicionadas en la región. Mecanismos de focalización desde una perspectiva comparada entre Uruguay y Costa Rica [Manuscrito no publicado].
- Gómez Bonaglia, E., da Silva, N. (2024). *Programas de transferencias condicionadas. Focalización basada en Aprendizaje Estadístico (Serie Documentos de Trabajo; n° 1/24)*. Universidad de la República. Facultad de Ciencias Económicas y de Administración, Instituto de Estadística.
- Ibarrarán, P., Medellín, N., Regalia, F., & Stampini, M. (2017). *Así funcionan las transferencias condicionadas: Buenas prácticas a 20 años de implementación*. Banco Interamericano de Desarrollo.
- Kaztman, R., Filgueira, F., Rodríguez, F., & Lijtenstein, S. (2001). *Panorama de la infancia y la familia en Uruguay*. Universidad Católica del Uruguay.
- Lavalleja, M. y Tenenbaum, V. (2020). *Mecanismos de focalización de los programas de transferencias monetarias en América Latina: el caso del Uruguay. Serie Estudios y Perspectivas, N° 43*. Oficina de la CEPAL en



- Montevideo.
- Lavalleja, M., & Tenenbaum, V. (2022). *Análisis y recomendaciones sobre el diseño de los programas de transferencias monetarias en el Uruguay*. CEPAL.
- Martínez Franzoni, J. (2008). Welfare regimes in Latin America: Capturing constellations of markets, families, and policies. *Latin American Politics and Society*, 50(2), 67–100.
- Midaglia, C., Antía, F., Carneiro, F., Castillo, M., Fuentes, G., & Villegas Plá, B. (2017). *Orígenes del bienestar en Uruguay: explicando el universalismo estratificado (Documento de Trabajo N° 01117)*. Facultad de Ciencias Sociales - Instituto de Ciencia Política.
- MIDES-DINTAD. (2025). *Informe de Transferencias*. Ministerio de Desarrollo Social.
- Noriega-Campero, A., Garcia-Bulle, L. F. B., Cantu, M., Bakker, L., Tejerina, A., & Pentland, A. (2020). Algorithmic targeting of social policies: Fairness, accuracy, and distributed governance. En *FAT\*20: Conference on Fairness, Accountability, and Transparency*.
- Perazzo, I., Rivero, A., & Vigorito, A. (2021). ¿Qué sabemos sobre los programas de transferencias no contributivas en Uruguay? Una síntesis de resultados de investigación disponibles sobre el PANES, AFAM-PE y TUS (No. 21-33).
- Rossel, C. (2013). *Desbalance etario del bienestar: el lugar de la infancia en la protección social en América Latina (No. 6190)*. Comisión Económica para América Latina y el Caribe (CEPAL).
- Rossel, C., Antía, F., & Manzi, P. (2022). The politics of conditionality in Latin America's cash transfer reforms. *The world politics of social investment*, 2, 379-401.
- Sátyro, N., & Midaglia, C. (2021). Family policies in Latin American countries: Re-enforcing familialism. En C. Midaglia & N. Sátyro (Eds.), *Latin American social policy developments in the twenty-first century* (pp. 287–314). Springer International Publishing.
- Stampini, M., Medellín, N., & Ibarrarán, P. (2023). *Transferencias monetarias, pobreza y desigualdad en América Latina y el Caribe (No. IDB-WP-01531)*. IDB Working Paper Series.
- Terra, J. P. (1979). *Situación de la infancia en América Latina y el Caribe*. UNICEF.
- Terra, J. P. (1988). Estado nutricional y desarrollo psicomotor en los niños de las familias pobres. *Cuadernos del CLAEH*, 12(47), 71-94.
- Terra, J. P. (1990). *Población en riesgo social. Infancia y políticas públicas en el Uruguay*. Instituto Nacional del Libro.

## Social administrative records and statistical learning: timely responses for pregnant women and early childhood in Uruguay

**Abstract:** This article analyzes the potential of integrating social administrative records and statistical learning methods to improve the effectiveness of social policies aimed at pregnant women and early childhood in Uruguay. In light of persistent child poverty and targeting problems (exclusion errors, inclusion errors, and timing issues), the study proposes integrating the Perinatal Information System (SIP) and the Integrated Social Area Information System (SIAS) to improve the delivery of Family Allowances (AFAM-PE). Based on this integration, a predictive model of social vulnerability is fitted using Stochastic Gradient Boosting (GBM) algorithm. Trained on 2015-2023 cohort data and using key variables from the first prenatal visit, the model achieves 85.7% accuracy in identifying pregnant women in vulnerable households. The study shows that, by overcoming challenges such as data quality and real-time interoperability, Uruguay has the opportunity to use its administrative records to transform social protection toward more timely and effective responses. It also identifies the potential to scale this tool to other countries in the region where SIP is used.

**Keywords:** statistical learning, social registries, child poverty, social policies, vulnerability.

## Resumen Biográfico

*Elina Gómez Bonaglia:* Socióloga, profesora de la Unidad de Métodos y Acceso a Datos de la Facultad de Ciencias Sociales y coordinadora del Centro de Información Socioeconómica y Comportamental del Instituto de Investigación en Justicia Social y Desigualdades. Es magíster en Estudios Contemporáneos de América Latina y candidata a Doctora en Sociología por la Universidad de la República. Sus temas de investigación se centran en políticas públicas, metodología y métodos computacionales en ciencias sociales.

*Gustavo Méndez Barbato:* Profesor del Departamento de Ciencia Política de la Facultad de Ciencias Sociales de la Universidad de la República, Uruguay. Es magíster en Historia Política y candidato a Doctor en Ciencia Política por la Universidad de la República. Sus temas de investigación se centran en economía política, políticas sociales, estado y políticas públicas.

*Philippe Rimoli Rimbaud:* Economista, Asesor del Sistema de Información Integrada del Área Social (SIIAS, Ministerio de Desarrollo Social), y Docente del Centro Interdisciplinario en Cognición para la Enseñanza y el Aprendizaje (CICEA, Universidad de la República). Sus temas de interés son el desarrollo infantil, sistemas de información y políticas sociales.

*Natalia da Silva:* Profesora del Departamento de Métodos Cuantitativos en el Instituto de Estadística de la Facultad de Ciencias Económicas y de Administración de la Universidad de la República, Uruguay. Realizó su doctorado y maestría en Estadística en Iowa State University. Sus temas de investigación se basan en aprendizaje estadístico, estadística computacional, visualización estadística e investigación reproducible.

## Contribución de los Autores

Todos los autores han contribuido de manera sustancial e igualitaria a la conceptualización, redacción, revisión y edición de este manuscrito.

## Financiamiento

Este artículo surge a partir de un proyecto en curso del Fondo de Población de Naciones Unidas (UNFPA) en Uruguay, en acuerdo con el Gobierno de Uruguay (Ministerio de Desarrollo Social y Ministerio de Salud Pública, en el marco del Programa de País 2021- 2025. El equipo técnico está conformado por: Alejandra Erramuspe (coord.), Daniel Macadar (coord. UNFPA), Philippe Rimoli (MIDES), Natalia da Silva, José Díaz Rosello, Elina Gómez, Luis Mainero, Gustavo Méndez, Cristina Tello, Matías Brum.

## Conflicto de Interés

El autor no presenta conflicto de interés.

## Cita

Gómez Bonaglia, E., Méndez Barbato, G., Rimoli Rimbaud, P., y da Silva, N. (2025). Registros sociales y aprendizaje estadístico: respuestas oportunas para gestantes y niñez temprana en Uruguay. *Revista MERCOSUR de políticas sociales*, 6, 32-46. <https://doi.org/10.28917/ism.2025-v6-id152>