



Paakat: Revista de Tecnología y Sociedad
e-ISSN: 2007-3607
Universidad de Guadalajara
Sistema de Universidad Virtual
México
paakat@udgvirtual.udg.mx

Año 12, número 23, septiembre 2022-febrero 2023

Estimación de la disponibilidad de internet en casa por medio de una función de predicción score

Estimation of Internet availability at home by a score prediction function

Djamel Toudert*

<http://orcid.org/0000-0003-2833-4128>

El Colegio de la Frontera Norte, México

[Recibido 21/01/2022. Aceptado para su publicación 3/06/2022]

DOI: <http://dx.doi.org/10.32870/Pk.a12n23.735>

Resumen

Disponer de internet en casa es uno de los pasos más importantes para gozar de los beneficios que se desprenden de su uso. En este sentido, conocer el peso que implica la disponibilidad o la carencia de la red se vuelve estratégico en el diseño de políticas y acciones que permitan intensificar la penetración socioterritorial de la web y el apoderamiento por parte de los usuarios. Esta investigación tuvo por objetivo elaborar una función de predicción de la disponibilidad de internet en los hogares, con base en los microdatos del censo de población y vivienda de 2020 realizado por el Instituto Nacional de Estadística y Geografía (INEGI). Este tipo de predicciones, además de aclarar las condiciones socioterritoriales que intervienen en la disponibilidad física de la red en las viviendas, contribuyen a la estructuración de un contexto de intervención para los factores públicos y privados del desenclave digital. En términos metodológicos, la función de predicción se fundamenta en una muestra aleatoria de 11 000 entrevistados del censo, repartidos en dos grupos de igual consistencia: el primero cuenta con internet en casa y el segundo carece de este servicio. Para la generación de la predicción se utilizó una función lineal discriminante valorada por una función score. La calidad de la función fue estimada a través de una evaluación interna con Bootstrap (75% de la muestra) y otra externa (25% de la muestra), ambas fueron complementadas con un análisis de la curva ROC y una gráfica de LIFT. La función

resultante exhibe una sensibilidad de 76.9%, una especificidad de 20.5% y una capacidad discriminante de 86.7% en un intervalo de confianza del 95%. Con estos parámetros de eficiencia, la función encontrada se convierte en un posible instrumento de predicción eficaz tanto para la reflexión como para la acción en el combate a la brecha digital en su etapa inicial.

Palabras clave:

Internet en casa; brecha digital; función predictiva.

Abstract

The availability of Internet at home is one of the most important steps to enjoy the benefits that come from its use. In this sense, knowing the weight implied by the availability or lack of the network becomes strategic in the design of policies and actions to intensify the socio-territorial penetration of the Internet and the empowerment of its users. The aim of this research was to develop a function for predicting Internet availability in households, based on microdata from the 2020 population and housing census conducted by the National Institute of Statistics and Geography (INEGI). This type of prediction, in addition to clarifying the socio-territorial conditions involved in the physical availability of the Internet in homes contribute to the structuring of an intervention context for the public and private factors of digital disengagement. In methodological terms, the prediction function is based on a random sample of 11 000 census interviewees, divided into two groups of equal consistency: the first has Internet at home and the does not. To generate the prediction, a linear discriminant function was used and valued by a score function. The quality of the prediction was estimated through an internal evaluation with Bootsrap (75% of the sample) and an external one (25% of the sample), both of which were complemented with an analysis of the ROC curve and LIFT graph. The resulting function exhibits a sensitivity of 76.9%, a specificity of 20.5% and a discriminant capacity of 86.7% in a 95% confidence interval. With these efficiency parameters, the found function becomes a possible effective prediction tool for both reflection and action against the digital divide in its initial stage.

Keywords:

Internet at home; digital divide; predictive function.

Introducción

De manera similar a lo que ocurrió con la difusión del teléfono fijo, los actores sociales y económicos, en el sentido más extenso de la palabra, avistaron en el internet un servicio trascendental que podría proyectarse con rapidez en el marco de un consumo universal (Graham, 2008; Toudert, 2013; Warf, 2001).

Este objetivo de globalizar el acceso a la red fue elevado por una legitimidad adquirida gracias a la premura de la masificación de internet, servicio que condiciona el acceso a los beneficios de la "revolución digital" y su corolario: la sociedad de la información (Castells, 2002; Graham, 1998; Eynon *et al.*, 2018). Independientemente de la pertinencia de los postulados sociotecnológicos asociados a las tecnologías de la información y la comunicación (TIC), para el discurso dominante en estos contextos, la disponibilidad de la web se ha vuelto un paso imprescindible para la integración social.

La creciente contracción del mundo análogo visibiliza la fractura social entre los que cuentan con el acceso a los canales de la oferta digital, los que

están a la expectativa de ingresar y los que inevitablemente se quedarán fuera (Ghobadi & Ghobadi, 2015; Graham & Marvin, 2001). El cruce entre las lógicas de interés –relacionadas con el desarrollo del mercado de provisión de la red– y de valores –asociadas con la inclusión social a los servicios digitales– interviene con fuerza en la primera etapa de la brecha digital: la disponibilidad material de internet, que constituye el fundamento básico de toda reflexión o acción encaminada a fomentar la apropiación de las TIC y el acceso a sus beneficios (Ali *et al.*, 2020; Toudert, 2019). Esta última afirmación se corroboró de sobra durante el largo confinamiento debido a la pandemia de la covid-19, período en el que los escolarizados tenían que contar con el servicio de internet en casa para continuar con sus estudios (Özge *et al.*, 2022).

La disponibilidad de internet, que con frecuencia parecería ser una situación superada en el discurso dominante, encuentra serias limitaciones en algunas realidades individuales, sociales y territoriales. En la actualidad, una proporción significativa de la población no cuenta con este servicio en sus hogares, lo que limita sus posibilidades de conectividad a las redes en espacios públicos, de trabajo y escolares.

En este sentido, que en las viviendas se cuente con acceso a internet todavía constituye en los contextos nacionales un objetivo central para las políticas públicas y privadas involucradas en el desenclave digital (Middleton & Chambers, 2010; Pick *et al.*, 2014; van Deursen & van Dijk, 2019). Cuando en una residencia convergen la accesibilidad, la privacidad y la asesoría, se transforma en el espacio predilecto para sacarle provecho a todos los tipos de contenidos de la red (Loh & Chib, 2021; Zhao *et al.*, 2010).

Así, contar con internet en casa para una adecuada apropiación de las TIC se convierte en un objetivo dentro de las iniciativas y los programas en el ámbito público y privado (Toudert, 2015; Zeqi *et al.*, 2019). El conocimiento de los contextos socioterritoriales que arrojan la indisponibilidad de la red adquiere un valor estratégico en el diseño de políticas apropiadas para enfrentar esta problemática.

La predicción de la disponibilidad de internet en los hogares por medio de la información sociodemográfica y condiciones de vivienda de los habitantes se encuentra en el marco de los instrumentos que permiten consolidar la comprensión de los diferentes escenarios, y puede ser de gran apoyo en los procesos de reflexión y toma de decisiones. Conocer el estado de acceso a la red de individuos y familias es un paso necesario para emprender en el ámbito del desenclave digital de forma exitosa.

El presente estudio tuvo como objetivo elaborar una función discriminante que por medio de un puntaje *score* permita predecir, con un alto

margen de confianza, el estado de disponibilidad de internet en casa, con base en los microdatos del censo de población y vivienda de 2020 realizado por el Instituto Nacional de Estadística y Geografía (INEGI, 2020).

La función de predicción resultante se describe en el apartado de resultados y se desarrollada con la intención de convertirse de manera inmediata en una prueba piloto, al permitir detectar principalmente la indisponibilidad de la web en las residencias en el marco de los diferentes contextos socioterritoriales relacionados.

Marco teórico

Brecha digital y disponibilidad de internet en casa

Desde su aparición en la década de los noventa en el marco de los cuatro estudios *Falling Through the Net*, publicados por el departamento de comercio estadounidense en 1999, el concepto de brecha digital ha venido cambiando alrededor de un significado invariable (Toudert, 2015; van Dijk, 2006). Esta constancia conceptual proviene del posicionamiento de su significado en las desigualdades sociales y territoriales del acceso a las oportunidades que ofrecen las TIC y a los beneficios de la sociedad de la información (Castells, 2002; Ghobadi & Ghobadi, 2015).

La perspectiva dinámica de la brecha digital alude a la fenomenología, incluyendo tanto la evolución de los artefactos y servicios de las TIC, como las etapas de su apropiación individual y social (Dupuy, 2007; Toudert, 2018; van Deursen & van Dijk, 2015). De esta manera, la brecha digital aparece como un esfuerzo de conceptualización complejo que involucra aspectos materiales, cognitivos y sociales (De Haan, 2004).

Desde la perspectiva de la apropiación de internet, la brecha digital interviene de forma multifacética: inicia con las barreras de accesibilidad al servicio, continúa con la carencia de habilidades para el acceso a los artefactos y contenidos digitales, y finaliza con las limitaciones que intervienen en el uso provechoso de la red (Toudert, 2018; van Deursen & van Dijk, 2015). A la par de esta perspectiva gradual de la brecha digital, e independientemente de si la motivación de los individuos para usar las TIC es importante o no (Reisdorf & Grošelj, 2017), la disponibilidad del servicio constituye una condición *sine qua non* para la accesibilidad a la red (Martínez Domínguez, 2020; Deursen & van Dijk, 2015).

Disponer de internet sugiere una situación física donde existe la posibilidad de conectar un artefacto de intermediación, como una computadora,

al servicio de red, ya sea que este se suministre a través de línea, cable, aire o satélite. Aunque esta afirmación puede parecer obvia en la actualidad, la confusión entre *disponibilidad* y *accesibilidad* llevó a la perspectiva liberal a considerar la brecha digital como un producto de la falta de motivación por parte de la población digitalmente marginada (Graham & Marvin, 2001).

En este sentido, la disponibilidad constituye la etapa inicial de la brecha digital, pues involucra la carencia física del servicio de internet en el lugar en donde se pretende acceder a la red (Toudert, 2015; Zeqi *et al.*, 2019). Esta carencia se debe a diversos factores que pueden actuar de manera individual o en conjunto; algunos de los más importantes son aquellos relacionados con el mercado de la provisión de internet, las condiciones socioeconómicas de las personas y el interés para contratar el servicio cuando está disponible (Barzilai-Nahon, 2006; Reisdorf & Grošelj, 2017; Zeqi *et al.*, 2019). En efecto, como afirma Barzilai-Nahon (2006) no perdimos de vista que la red y las tecnologías asociada, tanto en su estructura como en su contenido, son un producto social y político.

La incapacidad del mercado para proporcionar un servicio eficiente de internet a todos los segmentos sociales, independientemente de su ubicación, constituye una de las principales barreras para su disponibilidad en puntos de mayor estancia como el hogar, la escuela y el trabajo (Ali *et al.*, 2020; Schleife, 2010; Toudert, 2013), lugares trascendentes para la apropiación de las TIC (Martínez Domínguez, 2020; Toudert, 2018). La indisponibilidad de la red parece prevalecer en los contextos de pobreza y marginación, así como en las zonas rurales de baja densidad poblacional, las regiones remotas y los espacios en las periferias urbanas, lo que confirma un suministro todavía relacionado con las grandes carencias de consumo (Graham & Marvin, 2001; Pick *et al.*, 2014; Toudert, 2013).

Desde luego, el desarrollo tecnológico permite la erupción de plataformas alternas a la conectividad por línea en estos espacios marginados, como la conexión móvil de bajo débito o por satélite, no obstante, la baja calidad y su elevado costo siguen siendo factores negativos que entran en juego al considerar la contratación de este tipo de servicios (Graham, 2008; Toudert, 2013; Martínez Domínguez, 2020; Zeqi *et al.*, 2019). Bajo esta perspectiva, pareciera que la carencia de internet en casa depende de la estratificación poblacional, que se relaciona con las localidades de poca densidad de habitantes donde la consistente representación de los grupos demográficos y socioeconómicos menos proclives a usar internet incide en la tasa local de penetración social y apropiación de las TIC (Pick *et al.*, 2014; Toudert, 2013).

En estos escenarios, las tasas de jóvenes, estudiantes, migrantes, trabajadores y habitantes que usan de manera intensiva la tecnología en sus

labores suelen ser poco representados (Büchi *et al.*, 2016; van Deursen *et al.*, 2015). A esto se le suma el rechazo de los proveedores de internet de suministrar el servicio en este tipo de localidades, la baja motivación de algunos residentes para contratar el servicio e inclusive el desarrollo de una actitud refractaria hacia las TIC, situaciones que pueden observarse como barreras significativas al considerar la disponibilidad de la red en los hogares (Eynon *et al.*, 2018; Lindblom y Räsänen, 2017).

En México, Mora-Rivera y García-Mora (2021) hallaron que la falta de internet en los espacios rurales se traduce en dificultades para la reducción de la pobreza, en comparación con los espacios urbanos. En términos concretos, en el caso mexicano el conjunto de estas limitaciones se traduce durante el 2020 en un diferencial relativo de 27% entre las zonas urbanas y rurales respecto al uso de internet y la computadora (INEGI, 2020).

En este comparativo nacional, el desfase de las tasas de uso de internet entre el estrato socioeconómico alto y bajo alcanza 56.6%, 54.4% para el uso de la computadora y 33.5% para la telefonía celular. Estas cifras por sí solas son indicadoras de una carencia en el uso de tecnologías que aún se encuentra tipificada por características sociodemográficas, que suelen confundirse con pobreza y marginación socioterritorial (Toudert, 2013, 2015; Martínez Domínguez, 2020). La indisponibilidad de internet en estos contextos es un hecho observado desde los inicios del acceso público a esta red de servicios digitales (Castells, 2002; Graham & Marvin, 2001).

Si bien en estos escenarios se han adquirido formas de acceder internet conforme a los avances tecnológicos y a la reducción de los costos de una conectividad alterna –que aún no es adecuada para disponer de todos los contenidos disponibles en la red (Graham, 2008; Toudert, 2013)–, la tasa de vivienda particulares habitadas que no cuentan con internet en México sigue siendo alta, cerca de 48% en 2020 (INEGI, 2020). Esta limitación se vuelve más discriminante al considerar el papel preponderante que tiene la conectividad doméstica en el acceso a los diferentes contenidos que se albergan en la web.

Esto debido a que la consulta de internet en el trabajo, la escuela o en espacios públicos suele estar acompañada de restricciones objetivas y subjetivas que reducen el tiempo disponible para la interacción, la posibilidad de contar con algún tipo de asesoría y, sobre todo, la libertad de acceder a toda la oferta digital de la red (Loh y Chib, 2021; Zhao *et al.*, 2010). Estas deficiencias de la disponibilidad de internet son agravadas por condiciones de servicio, como el ancho de banda y la regularidad de la señal, condicionantes

que se toman en cuenta en la fase de accesibilidad de la brecha digital (Dupuy, 2007; van Deursen *et al.*, 2015).

Con todo lo anterior en cuenta, tanto la disponibilidad como la indisponibilidad del servicio de internet parece estar tipificada por características contextuales, sociodemográficas, económicas y de condiciones de vivienda con el potencial de predecirse para generar, entre otros avances, políticas públicas encaminadas a mejorar los mecanismos de apropiación socioterritorial de las TIC.

La predicción de internet: una herramienta para la reflexión y la acción

La predicción de la disponibilidad de internet en casa adquiere un valor estratégico en un contexto general convencido de que la cobertura universal de este servicio es eminente para la superación de las limitaciones presenciales del desarrollo e integración socioterritorial. En efecto, estas visiones fundadas en una supuesta oportunidad digital que vendría a rescatar a las comunidades dejadas al margen por la "vieja economía" y el mundo análogo son consideraciones que suelen exagerar los posibles impactos sociales y territoriales de las TIC (Graham, 1998; Toudert, 2015).

Este tipo de perspectivas son respaldadas por un discurso estructurado de manera imperante alrededor de la trascendencia tecnológica, del determinismo del paisaje digital y de la sustitución del espacio real por el virtual que ha venido saturando la reflexión académica y de los operadores públicos y privados involucrados en el desarrollo de la apropiación de las TIC (Graham, 1998; Graham & Marvin, 2001).

Es precisamente este tipo de consenso, que agrupa a factores provenientes de diferentes ámbitos e intereses, el que proporciona a la universalidad del acceso a internet un sustento moral y legítimo que entusiasma las cruzadas emprendidas en el nombre de la brecha digital (Bertrand, 2001; Toudert, 2014). Desde otra perspectiva, la masificación del servicio permite afrontar las iniquidades que pueden surgir en momentos de crisis, como los vividos durante los primeros momentos de la pandemia de la covid-19, cuando una parte de la población escolarizada no pudo seguir con sus actividades en línea (Özge *et al.*, 2022).

La universalidad del acceso a las TIC y la integración en masa a los beneficios de la sociedad de la información componen un nuevo paradigma para la convergencia de diferentes agentes sociales que fomenten la acción pública en dirección del desenclave digital (Castells, 2002). En el marco de

estas acciones, por lo general voluntaristas, queda claro que la disposición de la red es la base elemental para la adopción y el desarrollo de las tecnologías.

Es por esto que en los lugares de mayor estancia de la población la disponibilidad de la web se posiciona como una tarea trascendente para el conjunto de los actores involucrados en el combate a la brecha digital (Ali *et al.*, 2020; Toudert, 2013). Contrariamente a la complejidad de concretar iniciativas colectivas en dirección de las demás limitaciones de la brecha digital –como son la accesibilidad, las habilidades digitales y el uso provechoso de la red (Deursen & van Dijk, 2015; Martínez Domínguez, 2020)–, estas acciones públicas se caracterizan por una operatividad que permite generalizar la provisión del servicio.

En una reciente revisión de literatura, Lythreatis *et al.* (2022) menciona que la mayoría de los últimos estudios se centraron en el nivel dos de la brecha digital, lo que deja con poca investigación a la cuestión relacionada con la disponibilidad de la red, a pesar de que la población todavía no conectada sigue siendo muy importante en países como México (cerca del 40% de las viviendas habitadas, INEGI, 2020).

Los proyectos y programas de desenclave digital que suelen ser emprendidos por iniciativas gubernamentales o mixtas se han caracterizado por sostenerse generalmente en predicciones que permiten pronosticar sus impactos y oportunidades de éxito. En el marco de estas dinámicas, la predicción de las trayectorias de apropiación de las TIC por medio de funciones de uso de estas herramientas y servicios es una práctica común tanto para la reflexión como para la acción.

En este sentido, el advenimiento del modelo unificado para la adopción y uso de la tecnología (UTAUT, por sus siglas en inglés) es el instrumento de síntesis conceptual más empleado en la predicción de comportamientos de adopción y uso de las TIC (Venkatesh *et al.*, 2012).

En el mismo grupo de herramientas de predicción aparecen también algunas variantes del UTAUT y, sobre todo, del modelo de accesibilidad en etapas a la tecnología (AET) que pretende explicar la trayectoria de uso de las TIC a través de las cuatro etapas de la brecha digital: la motivación, el acceso, las habilidades y el uso (Middleton & Chambers, 2010; Toudert, 2019; van Deursen & van Dijk, 2015). En este último modelo, van Deursen y van Dijk (2015) incluyen la disponibilidad del internet como dimensión intrínseca de la accesibilidad material de los usuarios a las TIC.

Recientemente, Yu *et al.* (2018) propusieron, a partir de una amplia reflexión crítica, un modelo integral que permite el análisis de las inequidades

digitales que generan las diferentes facetas de la brecha digital; en esta, la disponibilidad de internet se etiqueta como micro causa que restringe la apropiación social de las TIC. De hecho, la conceptualización misma de las barreras y limitaciones de la apropiación de las tecnologías, como se afirma en Castells (2002), van Deursen y van Dijk (2015) y Toudert (2019), se refiere a las diferentes fases de la brecha digital cuya etapa inicial se caracteriza por la disponibilidad del servicio de internet, que hizo posible las TIC de cara a los demás artefactos y servicios.

A la par de las modelaciones a nivel del usuario por medio de ecuaciones estructurales, como en los casos de Toudert (2019) y van Deursen y van Dijk (2015), se han desarrollado modelos regresivos o factoriales con un interés predictivo de la disponibilidad, adopción y uso de internet. Estos modelos por lo general son elaborados para contextos territoriales, en busca de entender la apropiación global y regional de las TIC por medio de variables explicativas del ámbito sociodemográfico, económico y político (Mubarak *et al.*, 2020; van Deursen & van Dijk, 2019), y en términos generales, subrayan la incidencia en los procesos de apropiación de las TIC según la edad, el género, el nivel educativo, los ingresos, entre otros.

No obstante, a pesar de la existencia de las experiencias de modelación mencionadas, son escasos los estudios que se enfocaron de manera concreta a la predicción de la disponibilidad de internet en los hogares a partir de las variables que son incluidas de manera recurrente en los censos de población y vivienda del INEGI, inclusive cuando estos microdatos permiten llevar a cabo procesos exploratorios enfocados a generar funciones predictivas robustas para determinar a nivel de las personas el estado de la disponibilidad de internet en las casas.

Datos y metodología

Generación de la función básica de predicción

En términos generales, en este proyecto se aplicaron métodos estadísticos exploratorios y de predicción probabilística, ampliamente conocidos y utilizados en el área de las ciencias sociales. Para llevar a cabo la investigación, se usaron microdatos del censo de población y vivienda de 2020 del INEGI. A partir del universo general de entrevistados por el censo, se eligieron de manera aleatoria un total 11 000 casos, 5 500 entrevistados que expresaron contar con internet en sus viviendas y otros 5 500 que afirmaron no disponer del mismo servicio. Esta limitación de los datos permite operar con una muestra

representativa y suficiente, sin tener que cargar con los defectos estadísticos y operativos que trae considerar una elevada cantidad de datos.

La selección de variables que integraron la investigación se realizó por medio de un análisis discriminante sobre un extenso conjunto de variables que presentan una cercanía temática y epistemológica con la problemática de la disponibilidad de internet (Toudert, 2013). A partir de este, se eligieron para los siguientes pasos solo las variables discriminantes que mostraron un X^2 con $P < 0.05$ y un valor-test ampliamente superior a dos, que constituye un piso mínimo para la significación. Las variables retenidas son exhibidas en la tabla 1, en conjunto con su cuantificación individual: los que cuentan con internet en casa en la columna A y los que no disponen de este servicio en la columna B.

A continuación, las variables retenidas fueron incorporadas en un análisis de correspondencias múltiples con la finalidad de transformar sus modalidades discretas en factores continuos, un imperativo para poder operar posteriormente el cálculo de la función lineal discriminante de Fisher (Droesbeke *et al.*, 2005; Steyerberg *et al.*, 2010). De esta manera se obtuvieron cerca de 46 factores, de los cuales se conservaron 32 que mostraron ser estadísticamente significativos con $P < 0.05$ y que suman un 75.11% de la varianza total explicada, lo que caracteriza un importante nivel de extracción para las ciencias sociales.

Estos factores fueron la base para el cálculo de la función lineal discriminante que dio sustento a la función de predicción y a la asignación, por medio de una función *score* de coeficientes a las modalidades de las variables involucradas en el estudio (Steyerberg *et al.*, 2010). En efecto, los factores determinados con anterioridad desempeñan únicamente un papel transitorio de transformación de las modalidades de las variables involucradas en el estudio en valores continuos que exige el cálculo de la función lineal discriminante de Fisher para la asignación de un puntaje *score*. Para esta etapa, la función de predicción básica fue calculada con 75% de los casos de la muestra general utilizada en la investigación y después se procedió a la validación externa con 25% de los casos restantes.

Para la transformación de la función de predicción, en una prueba piloto se aplican los coeficientes transformados (columna D) que corresponden a cada modalidad que describe la persona entrevistada. Los coeficientes resultantes serán sumados de manera lineal a través de una expresión de tipo: $S = \sum_{k=1}^n kxz$, con S caracterizando el puntaje *score* resultante, x la variable en cuestión y z la modalidad de la variable que corresponde a cada persona entrevistada. Cabe señalar que esta misma prueba puede montarse en un formulario digital para llenarse en línea o en modalidad presencial.

Evaluación de las cualidades de la función de predicción

La evaluación de la eficiencia predictiva de la función básica se realizó a través de una validación interna y externa. La evaluación interna consistió en efectuar una estimación por medio de la técnica de Bootstrap utilizando 2 000 muestreos que arrojaron diferencias marginales con el cálculo básico que permitió aceptar la función de predicción (Bleeker *et al.*, 2003; Wahl *et al.*, 2016). Además de la evaluación anterior, es ineludible operar una validación externa que contemple confrontar la función básica elaborada con 75% de los casos con el resultado de la misma función alimentada con 25% de los casos restantes de la muestra total (Bleeker *et al.*, 2003). Ambos procedimientos de validación se complementan con indicadores de eficiencia derivados de la curva ROC (*receiver operating characteristic*) y la gráfica de LIFT (Vuk & Curk, 2006; Xie & Qiu, 2007).

Tabla 1. Características sociodemográficas y coeficientes discriminantes y transformados

Variables	A	B	C	D	Variables	A	B	C	D
V5: ¿Dónde se atiende su salud?					V8: Nivel escolar aprobado				
No se entiende	0.83	1.38	3.64	2.23	Ninguno	2.33	6.75	-1.41	2.49
IMSS, ISSSTE, ISSSTE estatal	18.16	7.48	18.53	5.97	Prescolar	2.41	3.20	-3.57	1.94
PEMEX, Defensa o Marina	0.52	0.15	-3.99	0.31	Primaria	12.74	19.79	0.78	3.04
SSA, IMSS-BIEN-PROSP	14.72	30.29	4.82	2.52	Secundaria	12.10	11.74	5.05	4.11
Clínica, hospital privado	10.56	6.14	1.05	1.57	Preparatoria	9.80	5.78	6.01	4.35
Consultorio de farmacia	4.55	3.92	-2.06	0.79	Universidad	9.43	2.16	-11.3	0.00
Otro lugar	0.66	0.65	-4.03	0.78	Especialidad	0.16	0.01	2.97	3.59
V6: ¿Se considera indígena?					V9: ¿Sabe leer y escribir?				
Sí	11.30	23.26	-4.06	0.00	Sí	48.14	40.96	0.75	1.91
No	39.03	26.41	2.05	1.54	No	2.55	8.36	-6.85	0.00
V12: Posición en el trabajo					V13: ¿Cuenta con servicio médico?				
Empleada(o) u obrera(o)	33.15	18.48	-0.78	1.03	Sí	31.87	10.42	1.25	1.20
Jornalera(o) o peón(a)	2.08	5.76	-4.88	0.00	No	25.61	32.10	-3.52	0.00
Ayudante con pago	2.17	3.35	-2.67	0.56	V17: Tamaño de localidad (número de habitantes)				
Patrón(a) o empleador(a)	2.72	0.86	-1.17	0.93	Menos de 2 500	11.30	28.17	-3.61	0.00
Trabajador(a) por su cuenta	13.12	10.99	5.48	2.61	De 2 500 a 14 999	13.52	13.38	-1.56	0.51
Trabajador(a) sin pago	2.08	5.23	-3.54	0.34	De 15 000 a 49 999	10.25	5.19	0.17	0.95
V18: Material del piso					De 50 000 a 99 999	3.37	0.95	8.76	3.11
Tierra	0.85	5.62	-5.80	29.90	100 000 y más	11.55	2.30	10.39	3.52
Cemento o firme	25.13	36.69	-3.34	30.52	V19: Número de cuartos en vivienda				
Madera, mosaico u otro	24.02	7.69	7.72	33.30	1 a 2	5.92	16.82	-5.98	0.00
V20: ¿Cuenta con agua entubada?					3 a 4	24.84	25.46	-1.95	1.01
Dentro de la vivienda	39.48	21.86	3.58	2.53	5 a 6	15.03	6.71	7.92	3.50
Solo en el patio o terreno	9.51	23.45	-6.50	0.00	7 a 8	3.21	0.83	12.46	4.64
Solo en el patio o terreno	9.51	23.45	-6.50	0.00	> a 8	1.01	0.18	9.65	3.93
No tienen agua entubada	1.01	4.68	-0.90	1.41	V22: ¿Cuenta con eliminación de basura?				
V21: ¿Cuenta con drenaje o desagüe?					Por camión o carrito	42.48	29.58	1.97	3.63
Red pública	39.15	23.12	2.40	5.75	En un contenedor	2.77	1.42	-6.27	1.56
Fosa o tanque sépticos	8.73	16.42	-2.56	4.51	La queman	3.71	16.75	-6.45	1.51
Va a dar a una barranca	0.39	1.14	-5.41	3.79	La entierran	0.13	0.55	8.72	5.33
Va a dar a un río, lago, mar	0.14	0.26	-20.5	0.00	V24: ¿Cuenta con televisor?				
No tiene drenaje	1.60	9.06	-6.46	3.52	Al basurero público	0.75	1.00	10.03	5.66
V24: ¿Cuenta con televisor?					La tiran en otro lugar	0.16	0.71	-7.57	1.23
Sí	47.72	39.18	1.23	2.36	V33: Ingresos mensuales en el hogar (pesos)				
No	2.28	10.82	-8.16	0.00					

V25: ¿Cuenta con computadora, <i>laptop</i> o <i>tablet</i> ?					No recibe ingresos	2.05	6.05	-	-
Sí	26.51	3.67	10.00	3.60	De 1 a 5 000	7.61	16.26	-3.32	0.02
No	23.49	46.33	-4.32	0.00	5 001 a 10 000	15.16	15.08	-3.07	0.08
V26: ¿Cuenta con línea telefónica fija?					10 001 a 15 000	9.48	5.44	2.24	1.42
Sí	24.68	3.95	9.29	3.27	15 001 a 20 000	7.06	2.33	8.82	3.07
No	25.32	46.05	-3.73	0.00	20 001 a 25 000	3.75	1.16	8.09	2.89
V27: ¿Cuenta con teléfono celular?					Más de 25 000	7.37	1.21	10.38	3.47
Sí	48.38	35.07	1.47	2.24	V28: ¿Cuenta con servicio de televisión de paga?				
No	1.62	14.93	-7.42	0.00	Sí	28.65	13.67	5.27	2.30
V30: ¿Cuenta con consola de videojuegos?					No	21.35	36.33	-3.87	0.00
Sí	7.25	0.32	18.21	4.96					
No	42.75	49.68	-1.49	0.00					

Notas: (A) % de la modalidad en el total de la variable para los que cuentan con internet. (B) % de la modalidad en el total de la variable para los que no cuentan con internet. (C) Coeficientes de la función discriminante. (D) Coeficientes transformados.
Fuente: elaboración propia con datos de INEGI (2020).

Resultados

Las variables sociodemográficas de vivienda y de posesiones en casa habitación, consignadas en la tabla 1, resultaron todas discriminantes de la función de predicción (*valor-test* superior a 2, $P < 0.05$) dejando entrever una clara segmentación de la población entre los que sí cuentan con internet y los que no. En efecto, conforme a lo exhibido en las columnas A y B, la población que tiene internet en casa parece gozar de mejores condiciones de integración socioeconómica en comparación con los que no cuentan con este servicio. De hecho, los que cuentan con acceso a la red, además de revelar mayores ingresos, se distinguen también por mejores condiciones y posesiones en cuestión de vivienda, lo que suele caracterizar a contextos sociodemográficos más favorables.

El nivel alto de educación, la posición de liderazgo en el trabajo y la residencia en localidades de mayor tamaño son contextos que parecen incentivar la disponibilidad de internet en los hogares. La estructura de esta disponibilidad parece todavía tipificada en México, independientemente de si ha ido disminuyendo las barreras de edad y género para la adopción de las TIC (Toudert, 2015, 2019).

La elaboración de una función de predicción de la disponibilidad de internet en casa, como la descrita en los párrafos anteriores, se sustenta en la discriminación contextual y debe cumplir con ciertos parámetros de validación que le den capacidad y robustez predictiva. De esta manera, el desarrollo de la función de predicción pasó por pruebas de validación interna y externa que permitieran considerarla como un instrumento confiable para la estimación de la disponibilidad de internet en las viviendas.

Validación interna de la función de predicción

La validación interna de la función de predicción consistió en comparar los resultados del cálculo básico y la estimación resultante por medio del método de Bootstrap de 2 000 muestreos (Bleeker *et al.*, 2003; Wahl *et al.*, 2016). Los resultados que se alistan en la tabla 2 muestran una capacidad de predicción total de la función elaborada con el cálculo básico de 78.65% para los casos bien clasificados y 21.37% para los mal clasificados. A la luz de estos hallazgos, se percibe que el cálculo básico acertó en casi ocho casos de diez, con una mayor capacidad para los que no tienen internet, con una tasa diferencial de 5.26% de los casos.

La estimación con Bootstrap de la función de predicción produjo diferencias marginales comparado con el cálculo básico; estas son del orden de 0.09% por los bien y mal clasificados y una desviación estándar de bajo de medio punto

porcentual. A nivel de los dos grupos analizados, las diferencias de estimación con el cálculo básico son ligeramente más altas para los que no cuentan con internet y estas no sobrepasen la décima de un punto porcentual. Bajo esta perspectiva, las diferencias halladas son muy marginales y, por lo tanto, todo parece indicar que la función de predicción del cálculo básico corresponde a la estimación con Bootstrap cumpliendo de esta manera con las exigencias de la validación interna.

Validación externa de la función de predicción

La validación externa que involucra el cálculo de prueba con el 25% de la muestra general exhibe pocas diferencias comparado con el cálculo básico de la función (ver tabla 2); estas son del orden de 0.05% para los bien clasificados totales y de 0.06% para los mal clasificados totales. Para el segmento de los que cuentan con internet, la diferencia con el cálculo general fortalece los clasificados de manera positiva con 0.19%, mientras que con los que no cuentan con servicio de red se fortalecen los mal clasificados con 0.17%.

Tabla 2. Validación interna y externa de la función lineal discriminante de Fisher

Elaboración de la función	Cálculo básico		Estimación con Bootstrap		Comprobación de la función	Cálculo de prueba	
	Bien clasificados	Mal clasificados	Bien clasificados	Mal clasificados		Bien clasificados	Mal clasificados
Con internet n %	3136 76.02	989 23.98	3132.44 75.95 [0.57]	992.52 24.06 [0.57]	Con internet n %	3144 76.21	981 23.78
Sin internet n %	3353 81.28	772 18.72	3349.17 81.19 [0.55]	775.83 18.81 [0.55]	Sin internet n %	3349 81.18	776 18.89
Total n %	6489 78.65	1761 21.35	6481.61 78.56 [0.44]	1768.39 21.44 [0.44]	Total n %	6493 78.70	1757 21.29

Nota: entre corchetes se indica la desviación estándar.

Fuente: elaboración propia con datos de INEGI (2020).

Tomando en cuenta el conjunto de estas diferencias entre el cálculo de prueba y el cálculo básico de la función que varían en todos los casos entre 0.05 y 0.19%, lo que se percibe son discrepancias marginales que no alteran

significativamente los resultados básicos. De esta manera, la función de predicción propuesta pasa de manera satisfactoria la prueba de validez externa y se procede en lo que va a seguir a la revisión del nivel de eficiencia de la predicción.

Eficiencia de la función de predicción y asignación de scoring

El análisis de la curva ROC que se muestra en la figura 1 indica un *score* límite de 61, caracterizado por una sensibilidad de 76.9% y una especificidad de 20.5% que corresponden a una capacidad discriminante de 0.867. Esta última define el poder discriminatorio de la función de predicción estimada con el área bajo la curva (AUC) en un intervalo de confianza del 95%, exhibiendo una capacidad predictiva para diferenciar los que cuenten con internet de los que no con una probabilidad del 86.7%.

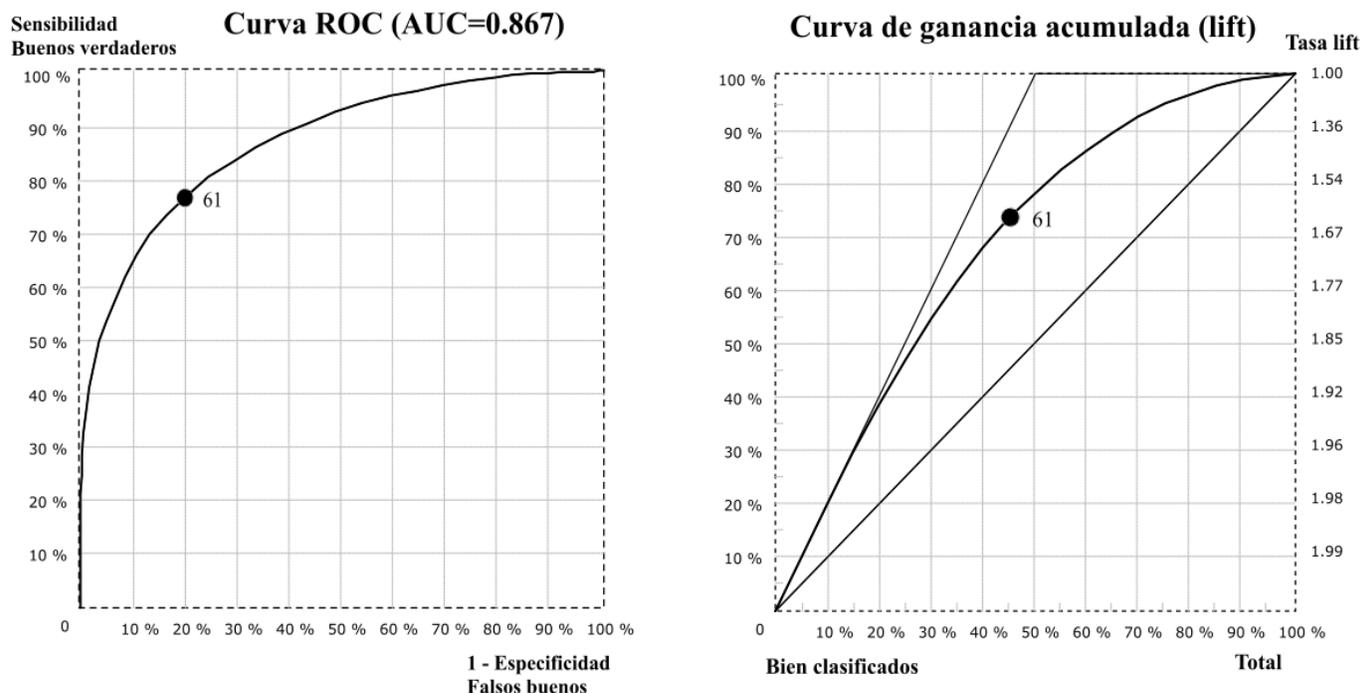
Se considera que esta capacidad de discernimiento parece aceptable para sustentar decisiones de política pública o empresarial, esta afirmación se confirma por la calidad de predicción de la gráfica LIFT (Vuk & Curk, 2006), que permite corroborar que la función de predicción pudo acertar en 77.98% de las predicciones con solamente la mitad de los casos analizados, logrando mostrar un ratio LIFT de 1.63 en el punto del *score* límite de 61.

Al fijar para la función de predicción una tasa de error tolerada de 10%, el resultado del cálculo de la función discriminante fue el aporte para determinar la función *score* para cada una de las modalidades de las variables involucradas (ver columna D de la tabla 1). Por medio de esta asignación, los casos analizados se consignaron a tres intervalos en función del *scoring* total obtenido. El *score* más elevado (entre 64 y 100) corresponde a los casos que fueron pronosticados de contar con internet en sus casas, un *score* bajo oscilante entre 0 y 56 indica a una predicción de casos sin el servicio mencionado mientras, un *score* variable entre 56 y 64 caracteriza los casos de asignación indecisa que no logran integrar ninguna de las dos primeras categorías mencionadas (ver tabla 3).

Estos mismos intervalos definen tres zonas de predicción que arrojan, en un contexto de una tasa de error tolerada de 10%, una aglutinación de 37% de casos en zona verde, 33.56% en zona rojo y los casos restantes en zona naranja de indecisión (ver tabla 3). Con la misma tasa tolerada de error, la predicción de disponer de internet en casa es de casi siete puntos porcentuales superior a la predicción de los casos que no cuentan con el servicio. Esta diferencia, cuando la tasa tolerada de error es fijada a un 10%, caracteriza una predicción con mayor

eficiencia en el pronóstico de casos que no cuentan con internet comparado con los que sí tienen el servicio.

Figura 1. Calidad y eficiencia de la función de predicción



Fuente: elaboración propia.

Tabla 3. Distribución de casos con una tasa tolerada de error de 10%

Zonificación	Zona roja [0, 56[Zona naranja [56, 64[Zona verde [64, 100]
Con internet			
n	551	1428	3521
%	10.02	25.96	64.02
Sin internet			
N	3143	1807	550
%	57.15	32.85	10
Total			
n	3694	3235	4071
%	33.56	29.41	37.01

Fuente: elaboración propia.

Discusión de resultados

Los hallazgos de la presente investigación confirman la racionalidad de la predicción de internet en los hogares, que puede transformarse en una herramienta que permita determinar a nivel de los individuos las necesidades de

conectividad a la red. Esta misma herramienta da lugar a una reflexión encaminada hacia una práctica de planeación que permita consolidar las condiciones de inclusión para nuevos beneficiarios de los servicios digitales.

Desde la perspectiva de la función score, un mayor puntaje atribuido por cada uno de los casos de estudio corresponde a una creciente posibilidad de contar con el servicio de internet en las viviendas (ver tabla 3). Dicho pronóstico se vuelve positivo a partir del valor 64 de un máximo de 100 puntos, indicando la disponibilidad del servicio con una efectividad de un poco más de 76% de los casos (ver tabla 2).

Por su parte, para los puntajes inferiores a 64, la predicción indica una indisponibilidad del servicio con un acierto en más de 80% de los casos. La diferencia de efectividad del pronóstico a favor de los casos que no cuentan con el servicio de red en el hogar incrementa el valor de utilidad de la función de predicción que se convierte en un posible instrumento para las políticas públicas y empresariales en temas de inclusión a los servicios digitales en línea (Martínez Domínguez, 2020; Toudert, 2018).

Lo anterior se vuelve estratégico en momentos en los que acceder a internet desde la casa se vuelve una necesidad urgente para llevar a cabo actividades escolares y profesionales (y otra índole), como ocurrió de improviso en el auge de la crisis de la covid-19 (Özge *et al.*, 2022). Estos contextos en particular no disminuyen la utilidad de predecir aquellos hogares que cuentan con el servicio de internet cuando se toma en cuenta la incidencia de otras vertientes de la brecha digital que pueden afectar usuarios de la red (van Deursen *et al.*, 2015).

Es cierto que la falta de artefactos adecuados para la conexión, la carencia de habilidades y de condiciones para un uso satisfactorio reducen de manera significativa la experiencia en la red y su uso provechoso (Loh y Chib, 2021; van Deursen & van Dijk, 2019). Sin embargo, no debe perderse de vista que cerca de 48% de las viviendas en México no cuentan todavía con internet, por lo que las cuestiones relacionadas a la conectividad se ubican como una de las principales problemáticas nacionales en la primera etapa de la brecha digital (INEGI, 2020). En estos escenarios marcados por carencias para la apropiación del instrumento digital, el planteamiento de acciones de apoyo que permiten superar estas limitaciones le otorga a la conexión a la red un significado de accesibilidad a la sociedad de la información (Castells, 2002).

La potencial asequibilidad a crecer los niveles de apropiación social de internet y de promover su uso provechoso se encuentra condicionado a la disponibilidad de la conectividad física a la red. Desde esta perspectiva, la sola disponibilidad del servicio de internet parece estructurar a una condición inicial

de brecha digital, misma que suele fluctuar en función de la posición social y la ubicación territorial, definiendo diversos contextos para la apropiación de las TIC (Ghobadi & Ghobadi, 2015; Graham & Marvin, 2001).

Una vez asegurada la disponibilidad de internet intervienen otras condicionantes técnicas y contextuales para la accesibilidad a las tecnologías. En el ámbito de las limitaciones técnicas de la accesibilidad, podemos enumerar un bajo ancho de banda, la inadecuación de los artefactos de consulta, que puede frenar la interacción con los diferentes tipos de contenidos disponibles en la red, entre otros factores (Barzilai-Nahon, 2006; Middleton & Chambers, 2010). En lo que respecta a las limitaciones contextuales, la facilidad para acceder a los artefactos conectados, la frecuencia y el tiempo de uso son aspectos señalados de manera recurrente en la literatura de la brecha digital (De Haan, 2004; Zeqi *et al.*, 2019). En resumen, se requiere superar las limitaciones a la accesibilidad a las TIC y contar con suficientes conocimientos y habilidades para interactuar adecuadamente con los contenidos digitales en línea (Toudert, 2019; van Deursen & van Dijk, 2015).

En investigaciones anteriores se ha puesto en evidencia que contar con internet en diferentes ubicaciones del espacio de vida (casa, trabajo, escuela, espacio público) abre la posibilidad para una mayor apropiación de la red (Middleton & Chambers, 2010; Toudert, 2019). De hecho, para Zhao *et al.* (2010) la disponibilidad de la conexión en casa permite contar con mayor privacidad para el desarrollo de actividades personales en la web, además de la posibilidad de acudir a familiares cercanos para solicitar apoyo para asesoría en el uso de la misma. Ahora bien, la disponibilidad de internet en la vivienda no implica sensu estricto la accesibilidad y su uso provechoso por todos y cada uno de los integrantes del hogar (Toudert, 2019; van Deursen *et al.*, 2015).

Además de las limitaciones técnicas y contextuales, mencionadas en párrafos anteriores, otras variables sociodemográficas, económicas, culturales y territoriales suelen incidir tanto en la disponibilidad como la accesibilidad a las tecnologías (Martínez, 2020; Mubarak *et al.*, 2020). En la columna D de la tabla 1 se exhiben los coeficientes transformados de la función discriminante que conforman los valores de la función *score* otorgados a cada una de las modalidades de variables participantes en el estudio. Cada vez que el *score* realizado por una modalidad es alto indicaría que el entrevistado recibe más puntos que le permiten pretender a los 64 de entrada a la zona verde, característica de los que cuentan con internet en casa.

El puntaje *score* que viene a caracterizar las modalidades de las variables involucradas en la función de predicción no es la traducción de una correlación

directa con la disponibilidad de internet en la vivienda. En términos específicos, el puntaje de estas variables es indicador de un contexto de marginación social y territorial que parece condicionar la disponibilidad del servicio.

Las modalidades de la variable que más puntaje aportan a la función es el material de cobertura del piso con *score* por modalidad fluctuante entre 29.90 para los pisos de tierra y 33.30 para los cubiertos de materiales de mayor elaboración (ver tabla 1). Tomando en cuenta que la diferencia de puntos entre modalidades no es amplia, el *score* total es indicador del importante valor discriminante de la variable con una distribución que favorece los materiales de mejor elaboración para la cobertura del piso de las casas.

En términos de importancia de puntaje otorgado, la segunda variable se encuentra relativamente lejos de la primera, apuntando el tipo de drenaje y desagüe de la casa habitación, que oscila entre un *score* de 0 para un drenaje que va a dar a un cuerpo de agua y 5.75 para las casas conectadas a la red pública. Aquí también la distribución del puntaje más alto parece seguir una lógica cualitativa creciente. En términos generales, para las variables indicadoras de las modalidades de servicios públicos en vivienda, prevalece en gran medida una distribución creciente conforme el nivel cualitativo del servicio disponible.

No obstante, esta lógica no parece confirmarse para el servicio de eliminación de la basura, en donde modalidades menos apropiadas son flanqueadas de un mayor *score*, indicando una probable distribución independiente del contexto de marginación –para este tipo de servicios municipales la prestación depende en gran medida del alejamiento y la antigüedad de los barrios y fraccionamientos–. Para las demás características de la vivienda, un creciente número de cuartos juega a favor de un mayor *score*, lo mismo ocurre también con las posesiones en casa, en donde contar con una consola de videojuegos es tributario de un mayor puntaje (4.96).

Desde la perspectiva del posicionamiento social de los casos de estudios, contar con servicio médico preferentemente con las instituciones públicas del IMSS, ISSSTE y ISSSTE estatal (5.97) parece otorgar un mayor *score* debido tal vez a la importante densidad de derechohabientes que pertenecen al grupo con disponibilidad de internet en casa. Del mismo modo, son los individuos que trabajan por su cuenta los que sacan un mejor puntaje (2.61), seguidos por empleados y obreros (1.03) que conforman el paisaje laboral en México.

La disponibilidad de la red en el caso de los trabajadores independientes podría traducirse en el uso profesional del internet y las tecnologías, que parece incrementarse con la población en general en función del nivel escolar aprobado, en particular para secundaria, preparatoria y especialidad (4.11, 4.35 y 3.59,

respectivamente). Contar con un nivel universitario no parece estimular la disponibilidad de la red en casa y en el estado del conocimiento actual es difícil avanzar una explicación inconcusa al respecto, aunque puede especularse con el hecho de que esta franja social pose internet en sus lugares de trabajo y se mantiene conectada todo el día con datos móviles.

La disponibilidad web en las viviendas parece recibir un mayor puntaje según el incremento del ingreso en el hogar, que llega hasta 3.47 para los ingresos superiores a los 25 000 pesos mensuales. De la misma manera, el *score* otorgado crece de manera lineal con el incremento del tamaño poblacional de las localidades, lo que indica que la disponibilidad de internet en casa continúa siendo un hecho principalmente urbano, dominados por la densidad de mercado (Graham & Marvin, 2001; Pick *et al.*, 2014; Schleife, 2010).

En conjunto, la variación del puntaje *score* por modalidad de las variables involucradas en el estudio parece incriminar casi en todos los casos las condiciones de marginación socioterritorial que condicionen la baja penetración social de la red en los hogares.

En términos generales, los hallazgos del estudio dejan entrever que el servicio de la red sigue siendo tipificado socialmente incluso después de más de dos décadas de abrirse al acceso público. Esta disponibilidad parece relacionarse con mayores niveles de educación, ingresos consecuentes, condiciones óptimas y posesiones específicas en vivienda. Estas características de segmentación social parecen expresarse con mayor fuerza en el marco de la densificación poblacional que se refleja en el tamaño de las localidades.

Conclusiones

En un contexto caracterizado por la creciente importancia del servicio de internet dentro de las casas particulares, comprender las lógicas de su disponibilidad y lograr predecirla se convierte en una oportunidad para la promoción de la integración digital, desde el ámbito público y privado. Al mismo tiempo, es un avance en el conocimiento empírico que puede nutrir la reflexión sobre las dinámicas de la apropiación social de las TIC. Desde una perspectiva práctica, la función de predicción permite ubicar a nivel personal la disposición de la web en los hogares, lo que contribuye en la elaboración de las políticas para la provisión del servicio, así como para brindar apoyo técnico y didáctico para un mejor aprovechamiento de la red cuando se cuente con el servicio.

En la misma medida, la función de predicción puede potenciar proyectos de expansión de cobertura, políticas de inversión en infraestructura pública y el

diseño de programas específicos para el desenclave digital. En este sentido, usar una base de datos empresarial o gubernamental en la función de predicción generada permite determinar con una precisión alta si los clientes o beneficiarios de los programas sociales cuentan o no con internet en casa.

Esta predicción se sustentó en una función *score* que arrojó tres zonas de pronóstico: los negativos *qui* se encuentren entre el valor 0 y 56, los casos indecisos entre 56 y 64, y finalmente los positivos entre 64 y 100 puntos. La caracterización por un puntaje *score* de las modalidades de variables deja entrever variaciones dispares. Por ejemplo, las modalidades de variables con mayor puntaje *score* que determinen el pronóstico de disponibilidad de internet en las viviendas se relacionan con una mejor calidad en las condiciones, pertenencias y servicios dentro de la misma. En el marco de esta lógica, sobresalen también las modalidades asociadas con altos ingresos del hogar, niveles de educación, ocupación profesional y acceso a seguridad social.

Para finalizar, se señala que la dinámica cualitativa vertical de estas variables por lo general se encuentra acompañada de un aumento en el puntaje. Desde otra perspectiva, tanto el incremento del tamaño de las localidades como la ascensión de los estratos sociales parece incidir positivamente en el pronóstico de contar con red en los hogares. Tomando en cuenta el conjunto de estos hallazgos puede concluirse que la disponibilidad del servicio de internet en casa todavía parece estar influido por las condiciones socioterritoriales que limitan la apropiación de las TIC.

Referencias

- Ali, M.; Alam, K. & Taylor, B. (2020). Do social exclusion and remoteness explain the digital divide in Australia? Evidence from a panel data estimation approach. *Economics of Innovation and New Technology*, 29(6), 643-659. <https://doi.org/10.1080/10438599.2019.1664708>
- Barzilai-Nahon, K. (2006). Gaps and Bits: Conceptualizing Measurements for Digital Divide/s. *The Information Society*, 22(5), 269-278. <https://doi.org/10.1080/01972240600903953>
- Bertrand, N. (2001). Technologies d'information et de communication: quel rôle dans les dynamiques territoriales de développement? *Revue d'Economie Regionale et Urbaine*, (1), 135-152. <https://www.cairn.info/revue-d-economie-regionale-et-urbaine-2001-1-page-135.htm>
- Bleeker, S.; Moll, H.; Steyerberg, E.; Donders A.; Derksen-Lubsen, G.; Grobbee, D. & Moons, K. (2003). External validation is necessary in prediction research: a clinical example. *Journal of Clinical Epidemiology*, 56(9), 826-32. [https://doi.org/10.1016/s0895-4356\(03\)00207-5](https://doi.org/10.1016/s0895-4356(03)00207-5)
- Büchi, M.; Just, N. & Latzer, M. (2016). Modeling the second-level digital divide: A five-country study of social differences in Internet use. *New Media & Society*, 18(11). <https://doi.org/10.1177/1461444815604154>

- Castells, M. (2002). *The Internet Galaxy. Reflections on the Internet, business, and society*. Oxford University Press.
- De Haan, J. (2004). A multifaceted dynamic model of the digital divide. *Information Technology & Society*, 1(7), 66-88.
- Department of Commerce. (1999). Falling Through the Net: Defining the digital divide. <http://www.ntia.doc.gov/legacy/ntiahome/fttn99/contents.html>
- Droesbeke J.; Lejeune M. y Saporta G. (eds.). (2005). *Modèles statistiques explicative pour données qualitatives*. Editions Technip.
- Dupuy, G. (2007). *La fracture numérique*. Ellipses Marketing.
- Eynon, R.; Seetjen, U. & Malmberg, L. (2018). Moving on up in the information society? A longitudinal analysis of the relationship between Internet use and social class mobility in Britain. *The Information Society*, 34(5), 316-327. <https://doi.org/10.1080/01972243.2018.1497744>
- Ghobadi, S. & Ghobadi, Z. (2015). How access gaps interact and shape digital divide: a cognitive investigation. *Behaviour & Information Technology*, 34(4), 330-340. <https://doi.org/10.1080/0144929X.2013.833650>
- Graham, M. (2008). Warped geographies of development: The Internet and theories of economic development. *Geography Compass*, 2(3), 771-789. <https://doi.org/10.1111/j.1749-8198.2008.00093.x>
- Graham, S. (1998). The end of geography or the explosion of place? Conceptualizing space, place and information technology. *Progression Human Geography*, 22(2), 165-185. <https://doi.org/10.1191/030913298671334137>
- Graham, S. & Marvin, S. (2001). *Splintering urbanism: Networked infrastructure technological mobilities and the urban conditions*. Routledge.
- Instituto nacional de estadística y geografía (INEGI). (2020). Censo de población y vivienda, 2020. INEGI.
- Lindblom, T. & Räsänen, P. (2017). Between class and status? Examining the digital divide in Finland, the United Kingdom, and Greece. *The Information Society*, 33(3), 147-158. <https://doi.org/10.1080/01972243.2017.1294124>
- Loh, Y. & Chib, A. (2021). Reconsidering the digital divide: an analytical framework from access to appropriation. *Information Technology & People*, 35(2), 647-676. <https://doi.org/10.1108/ITP-09-2019-0505>
- Lythreathis, S.; Singh, S. & El-Kassar, A. (2022). The digital divide: A review and future research agenda. *Technological Forecasting and Social Change*, 175. <https://doi.org/10.1016/j.techfore.2021.121359>
- Martínez Domínguez, M. (2020). La desigualdad digital en México: un análisis de las razones para el no acceso y el no uso de internet. *Paakat*, 10(19). <http://dx.doi.org/10.32870/Pk.a10n19.519>
- Middleton, K. & Chambers, V. (2010). Approaching digital equity: is wifi the new leveler? *Information Technology & People*, 23(1), 4-22. <https://doi.org/10.1108/09593841011022528>
- Mora-Rivera, J. y García-Mora, F. (2021). Internet access and poverty reduction: Evidence from rural and urban Mexico. *Telecommunications Policy*, 45(2). <https://doi.org/10.1016/j.telpol.2020.102076>
- Mubarak, F.; Suomi, R. y Kantola, S. (2020). Confirming the links between socio-economic variables and digitalization worldwide: the unsettled debate on digital divide. *Journal of Information Communication and Ethics in Society*, 18(3), 415-430. <https://doi.org/10.1108/JICES-02-2019-0021>

- Özge, K.; Elif, E. & Deniz, E. (2022). Internet access and its role on educational inequality during the COVID-19 pandemic. *Telecommunications Policy*, 46(5). <https://doi.org/10.1016/j.telpol.2022.102353>
- Pick, J.; Nishida, T. y Sarkar, A. (2014). Broadband utilization in the Indian states: socio-economic correlates and geographic aspects. En J. Choudrie, y C. Middleton (eds.), *Management of broadband technology innovation* (269-296). Routledge.
- Reisdorf, B. & Grošelj, D. (2017). Internet (non-)use types and motivational access: Implications for digital inequalities research. *New Media & Society*, 19(8), 1157-1176. <https://doi.org/10.1177/1461444815621539>
- Schleife, K. (2010). What really matters: Regional versus individual determinants of the digital divide in Germany. *Research Policy*, 39(1), 173-185. <https://doi.org/10.1016/j.respol.2009.11.003>
- Steyerberg, E.; Vickers, A.; Cook, N.; Gerds, T.; Gonen, M.; Obuchowski, N.; Pencina, M. & Kattan, M. (2010). Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*, 1(1), 128-38. <https://doi.org/10.1097/EDE.0b013e3181c30fb2>
- Toudert, D. (2013). La brecha digital en los contextos de marginación socioterritorial de las localidades mexicanas de más de 2500 habitantes: Exploración y discusión. *Comunicación y Sociedad*, (19), 153-180. <http://www.scielo.org.mx/pdf/comso/n19/n19a7.pdf>
- Toudert, D. (2014). Evolución de la polarización territorial de la producción web: ¿para cuándo la esperada descentralización? *Revista Economía, sociedad y territorio*, XIV(45), 549-580. <http://www.scielo.org.mx/pdf/est/v14n45/v14n45a9.pdf>
- Toudert, D. (2015). Brecha digital y marginación socioterritorial: el caso de México. En C. Rangel y G. Buzai (eds.), *Geografía Aplicada en Iberoamérica. Avances, retos y perspectivas* (343-370). El Colegio Mexiquense, AC.
- Toudert, D. (2018). Connectivité et isolement dans le monde scolaire mexicain. *La Revue Communication, Technologies et Développement*, 5, 116-126. <https://doi.org/10.4000/ctd.374>
- Toudert, D. (2019). Brecha digital, uso frecuente y aprovechamiento de Internet en México. *Convergencia revista de ciencias sociales*, (79), 1-27. <https://doi.org/10.29101/crcs.v0i79.10332>
- van Deursen, A.; van Dijk, J. y Klooster, P. (2015). Increasing inequalities in what we do online: A longitudinal cross-sectional analysis of Internet activities among the Dutch population (2010 to 2013) over gender, age, education, and income. *Telematic and Informatic*, 32(2). <https://doi.org/10.1016/j.tele.2014.09.003>
- van Deursen, A. & van Dijk, J. (2015). Toward a Multifaceted Model of Internet Access for Understanding Digital Divides: An Empirical Investigation. *The Information Society*, 31(5). <https://doi.org/10.1080/01972243.2015.1069770>
- van Deursen, A. & van Dijk, J. (2019). The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New media & society*, 21(2), 354-375. <https://doi.org/10.1177/1461444818797082>
- van Dijk, J. (2006). Digital divide research, achievements and shortcomings. *Poetic*, 34(4), 221-235. <https://doi.org/10.1016/j.poetic.2006.05.004>
- Venkatesh, V.; Thong, J. & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157-178. <https://doi.org/10.2307/41410412>
- Vuk, M. & Curk, T. (2006). Lift Chart and Calibration Plot. *Metodoloski zvezki*, 3(1), 89-108. <http://mrvar.fdv.uni-lj.si/pub/mz/mz3.1/vuk.pdf>

- Wahl, S.; Boulesteix, A.; Zierer, A.; Thorand, B. & van de Wiel, M. (2016). Assessment of predictive performance in incomplete data by combining internal validation and multiple imputation. *BMC Medical Research Methodology*, 16(1). <https://doi.org/10.1186/s12874-016-0239-7>
- Warf, B. (2001). Segueways into cyberspace: Multiple geographies of the digital divide. *Environment and Planning B: Planning and Design*, 28(3), 3-19. <https://doi.org/10.1068/b2691>
- Xie, J. & Qiu Z. 2007. Bootstrap technique for analysis: A stable evaluation of Fisher classifier performance. *Journal of Electronics*, 24(4), 523-527. <https://doi.org/10.1007/s11767-005-0261-0>
- Yu, B.; Ndumu, A.; Mon, L. & Fan, Z. (2018). E-inclusion or digital divide: an integrated model of digital inequality. *Journal of Documentation*, 74(3), 552-574. <https://doi.org/10.1108/JD-10-2017-0148>
- Zeqi, Q., Shuqin, Z. y Shiding, L. (2019). From the Digital Divide to the Connectivity Dividend Difference: A Connectivity Capital Perspective. *Social Sciences in China*, 40(1), 63-81. <https://doi.org/10.1080/02529203.2019.1556475>
- Zhao, L.; Lu, Y.; Huang, W. y Wanga, Q. (2010). Internet inequality: The relationship between high school students' Internet use in different locations and their Internet self-efficacy. *Computers & Education*, 55(4), 1405-1423. <https://doi.org/10.1016/j.compedu.2010.05.010>

Este artículo es de acceso abierto. Los usuarios pueden leer, descargar, distribuir, imprimir y enlazar al texto completo, siempre y cuando sea sin fines de lucro y se cite la fuente.

CÓMO CITAR ESTE ARTÍCULO:

Toudert, D. (2022). Estimación de la disponibilidad de internet en casa por medio de una función de predicción score. *Paakat: Revista de Tecnología y Sociedad*, 12(23). <http://dx.doi.org/10.32870/Pk.a12n23.735>

* Doctor en Geografía, urbanismo y ordenamiento territorial por el Instituto de Altos Estudios de América Latina, Universidad de Paris III. Profesor-Investigador del Departamento de Estudios Urbanos y Medio Ambiente de El Colegio de la Frontera Norte, Especialista en las cuestiones de la brecha digital, ordenamiento territorial y ecológico y el desarrollo de herramientas para el análisis y diagnóstico de los contextos socioambientales.