

Métodos de aprendizaje de máquina para inferir el nivel de cobertura de banda ancha fija en municipios de México

César Zamora Martínez^a

^a Alumno de Maestría en Ciencias de Datos (ITAM)

This manuscript was compiled on November 18, 2019

Aunque en fechas recientes se reconoce el impacto benéfico que la banda ancha tienen sobre en entorno económico y social, la penetración de tales servicios en los municipios obedece a múltiples factores que inciden en el despliegue de la infraestructura que permite su prestación. Motivado por ello, en este trabajo se plantea el uso de métodos basados en aprendizaje de máquina que permitan clasificar los municipios conforme a su nivel de cobertura a través de indicadores de penetración y establecer factores que propician o desincentivan los despliegues de banda ancha fija.

Aprendizaje de Máquina | Banda Ancha | Telecomunicaciones | ITAM

Durante las últimas tres décadas las telecomunicaciones han tenido un avance sin precedentes en el mundo, posicionándose como herramientas que potencian el desarrollo económico y social, pues, como ha sido ampliamente documentado en la literatura (Pradhan et al. (2014)), permiten crear oportunidades, reducir la pobreza e impulsar el progreso económico y social para el bienestar de la población*. Uno de los ejes que permiten explicar lo anterior es el impacto benéfico de los servicios de banda ancha en los procesos productivos, financieros y en general el bienestar de la población (Katz (2012)).

En México, a cerca de cuatro años de la reforma a de telecomunicaciones en 2013, que llevo a la promulgación de la Ley Federal de Telecomunicaciones y Radiodifusión (LFTyR) junto con la creación del Instituto Federal de Telecomunicaciones (IFT), se estimó un crecimiento superior al 37% en las conexiones de banda ancha fija (BAF), traduciéndose a que para entonces casi la mitad los hogares contaban con servicios de Internet (IFT (2017) e IFT (2018a)).

Aunque este fenómeno revela una tendencia favorable con respecto al entorno internacional[†], la Encuesta sobre Disponibilidad y Uso de Tecnologías de la Información y la Comunicación en los Hogares 2018 (ENDUTIH 2018, INEGI (2018)) mostró la existencia de una brecha en la adopción de estos servicios para la población mexicana (por tanto en sus conductores benéficos), pues sólo cerca de 65.8% de la población con seis años o más es usuario de servicios de Internet en los hogares del país, además de este es un fenómeno urbano, puesto el 73.1% del total de la población urbana son usuarios de este servicio en contraste con la población conectada en zonas rurales que es cercana a 40.6%.

A efecto de explicar el entorno de la penetración de servicios

* Por ejemplo, Katz (2018) muestra que un avance del 1% en un índice que mide el grado de desarrollo de digitalización, genera un incremento de la productividad que se traduce en un crecimiento económico de un 0.3% del PIB.

[†] A finales de 2018 México fue el cuarto país con mayor crecimiento de penetración de banda ancha fija, sólo por debajo de Turquía, Polonia y Eslovaquia, entre los países de la Organización para la Cooperación y Desarrollo Económicos (OCDE); y mostró un crecimiento de 17.9% en la penetración de accesos por medio de fibra óptica (IFT (2019))

de Internet a nivel municipal, en adelante nos centraremos en los servicios de banda ancha fija[‡], los cuales son servicios de acceso a Internet y transmisión de datos orientados a usuarios finales (personas físicas o empresas), que se brindan a través de equipos terminales (módems, terminales ópticas y demás) que tienen una ubicación geográfica determinada y fija (IFT (2018c)). Ello obliga a los operadores de telecomunicaciones interesados en prestar tales servicios a realizar inversiones en recursos que les permitan alcanzar los puntos geográficos en donde se localizan los clientes potenciales, esto es, cerca de hogares y edificios de empresas, aprovechando las capacidades de las tecnologías en las que se basan sus redes.

Dicho contexto les condiciona a establecer un circuito físico o virtual a través del cual se pueda conectar la ubicación del usuario a la red del operador y a través del que se prestarán los servicios ("Acceso de datos" o simplemente como "acceso", IFT (2018c)). Por ende, dado que afrontan costos considerables en infraestructura, equipos, permisos y recursos humanos para poder brindar servicios[§], típicamente los operadores concentran su oferta en zonas densamente pobladas donde existe suficiente capacidad económica para asegurar no solo que recuperarán sus inversiones sino que serán rentables desde la visión de negocio.

Por otro lado, un indicador ampliamente usado en el sector de telecomunicaciones (Hanafizadeh et al. (2009), Escobar (2017), IFT (2018c)) para cuantificar la cobertura de banda ancha fija es medir la penetración en una zona con la cantidad de accesos en ella por cada 100 hogares:

$$PenBAFHogares = \frac{Accesos}{Hogares} \times 100 \quad [1]$$

Este indicador tiene las siguientes limitaciones 1) la presencia de accesos en un municipio no implica necesariamente que las localidades que lo conforman cuentan con infraestructura para brindar estos servicios[¶], 2) la penetración se puede subestimar si no se tienen una división explícita entre accesos de usuarios residenciales y empresariales.

En vista de este último punto, la Organización para la Cooperación y el Desarrollo Económicos (OCDE) define el

[‡] Ello dado que desafortunadamente, el Banco de Información de Telecomunicaciones (BIT) del IFT sólo posee el detalle desagregado de servicios de Internet de banda ancha móvil para nivel estado, sin que se hayan podido localizar fuentes con datos precisos al respecto.

[§] En línea con Escobar (2017), no sólo se enfrentan costos directos, sino oportunidad y de transacción; así como el riesgo de afrontar costos hundidos.

[¶] A guisa de ejemplo, puede ser el caso que únicamente existan servicios de BAF en la cabecera municipal, pero no en el resto de sus comunidades

² E-mail: czamora5email.itam.mx

parámetro de penetración por cada 100 habitantes como un proxy del indicador de suscriptores por cada 100 habitantes:

$$PenBAFHabitantes = \frac{Accesos}{Habitantes} \times 100 \quad [2]$$

A través de tales indicadores, se puede categorizar a cada municipio del país acuerdo a su nivel de penetración de banda ancha fija como sigue:

Table 1. Niveles de penetración en un municipio

Nivel de penetración	Rango de penetración
Muy Alta	$Penetracion > 100$
Alta	$75 < Penetracion \leq 100$
Media	$50 < Penetracion \leq 75$
Baja	$25 < Penetracion \leq 50$
Muy Baja	$0 < Penetracion \leq 25$
Nula	$Pen = 0$

Clasificación de municipios según su rango de penetración de BAF

Aunado a ello, para BAF los accesos se basan típicamente en las tecnologías[†] (Moya (2014)): 1) DSL: tecnología de transmisión por cable trenzado de cobre, su disponibilidad y la velocidad dependen de la distancia; 2) Cable coaxial: se forma por dos hilos de cobre (uno en el centro y otro alrededor de una malla) cuya estructura permite banda ancha, entrega de imágenes y sonido. En general, permiten más capacidad para transmitir información que el par trenzado de cobre; 3) Fibra óptica: formados por un medio de vidrio o polímero que permite el paso de haces de luz, pueden transmitir más de 10 Gbit/s hasta a 10 kilómetros de distancia, 4) Otras: incluye uso de ondas electromagnéticas como microondas, señales satélites, tecnologías Wi-Max; usualmente las velocidades de transmisión para servicios de datos tienen menor desempeño comparado a 1), 2) o 3), pero son viables en regiones rurales o desérticas.

Con todo lo anterior, el objetivo del presente documento será plantear un modelo con métodos de aprendizaje de máquina que permitan identificar variables útiles para explicar el nivel de penetración de BAF (por ejemplo, datos geográficos y demográficos) así como de los factores que inciden en los despliegues de tecnologías capaces de dar servicios de Internet de alta velocidad.

En este sentido, con motivo de estudiar a los municipios que cuentan penetración de BAF basada en tecnologías capaces de dar servicios de velocidad alta, se enfocarán los indicadores de penetración presentados previamente sobre accesos correspondientes a tecnologías de cable coaxial o fibra óptica (es decir, se calcularán con respecto la cantidad de accesos resultado de sumar de los que correspondan a cable coaxial y aquellas de fibra óptica en cada municipio).

1. Revisión sobre fuentes de datos asociados a banda ancha fija

A continuación se resumen las fuentes de información consultadas con relación a servicios de BAF, junto con las consid-

[†] Dependiendo de la configuraciones de la red de un operador, se pueden emplear tramos híbridos con diferentes medios de transmisión que permiten aumentar la velocidad y desempeño del acceso; por ejemplo tecnologías FTTC que acercan la centra empleado fibra óptica y re-usan la planta de cobre existente en una zona.

eraciones particulares derivadas de su exploración^{**}.

A. Identificación de municipios. Dado que la disponibilidad de información social y demográfica en fuentes públicas con desglose municipal se encuentra limitada a ejercicios estadísticos que abarcan hasta el año 2015 (INEGI (2015), CONAPO (2015), ONU (2015)), la identificación de los municipios se hizo de manera congruente el marco metodológico de la Encuesta Intercensal 2015 en donde se contabilizaron un total de 2,457 municipios.

B. Datos de accesos de banda ancha. El Banco de Información de Telecomunicaciones (IFT (2018b)) posee datos históricos (de 2013 a mediados 2019) sobre los accesos de banda ancha móvil y fija de México; sin embargo únicamente en el segundo caso se ofrece el detalle a nivel municipio. Dicha fuente provee datos de 29 empresas a las que pertenecen los accesos de BAF junto tecnología correspondiente (DSL, cable coaxial, fibra óptica, satelital y otras), sin proveer el desglose entre accesos residenciales o no residenciales.

Con base en ella se observó que a junio de 2019, existían 18.85 millones de accesos de BAF distribuidos sobre 1,604 municipios el país (de un total de 2,457), para los que en un porcentaje de 2.76% de los accesos no había datos de su ubicación. En complemento, el 37.1% de tales accesos son DSL, 38.3% cable coaxial, 21.8% fibra óptica, 0.125% satelitales, mientras que para el cerca de 2.6% se desconoce la tecnología.

En términos de los principales grupos de telecomunicaciones de México a los que pertenecen los accesos de BAF, se tiene la siguiente distribución: América Móvil^{††} 51.6%, Grupo Televisa^{‡‡} 23.3%, Megacable-MCM^{§§} 16.1% y Totalplay^{¶¶} 7.35%

Otro punto a destacar es que únicamente los empresas que pertenecen a dichos grupos han desplegado accesos basados en cable coaxial o fibra óptica en el país^{***}.

Table 2. Distribución de accesos de BAF por tecnología para los principales grupos de telecomunicaciones en México (Junio 2019)

Grupo	Coaxial	DSL	Fibra	Satelital	No especificado
América Móvil		71.9%	28.1%		
Grupo Televisa	95.6%			0.1%	4.3%
Megacable-MC	99.9%				0.1%
TotalPlay			100%		

No se omite destacar la presencia de municipios con una cantidad de accesos inusualmente baja (e.g. hay 31 municipios con un sólo acceso, 7 que poseen únicamente dos accesos).

C. Datos socio-económicos. A través de la Encuesta Intercensal 2015 (INEGI (2015)), el INEGI reúne información de componentes que describen la evolución de la población en

^{**} El procesamiento de la información se llevó a cabo a través de scripts en Bash, R y Python, disponibles en el Anexo A de este documento

^{††} Agrupa a las empresas Telmex y Telcel

^{‡‡} Se refiere de manera conjunta a las empresas Cablecom, Cablemas, Cablevision, Cablevision Red, Television Internacional y Sky

^{§§} Se refiere de forma conjunta a las empresas Megacable y MCM

^{¶¶} Empresa Totalplay de Grupo Salinas

^{***} Los datos del BIT no reportan explícitamente datos de fibra óptica para Grupo Televisa, lo que contrasta con la oferta comercial de "Izzi" que brinda servicios de este tipo: "...hoy izzi ofrece servicio en más de 60 ciudades en 29 estados de la República Mexicana, mediante una red de más de 30,000 kilómetros de fibra óptica y 77,000 kilómetros de cable coaxial", consultado el 18 de Noviembre de 2019 en <https://www.izzi.mx/nosotros#infraestructura>

México y de sus condiciones socio-económicas. Ello se realiza para inferir el volumen, la composición y la distribución de la población y de las viviendas del territorio nacional.

Tras estudiar esta fuente, los datos que se han considerado de interés para el estudio a nivel municipal la penetración de BAF son: 1) número de hogares, 2) número de hogares, 3) porcentaje de viviendas que cuentan con disponibilidad de servicios de telecomunicaciones (es decir, a través de telefonía fija, telefonía celular, televisión de paga e Internet). Dicha elección se apoya en que, como se ha mencionado previamente, los despliegues de redes fijas se realizan alrededor de donde se ubican los clientes potenciales, y a su vez en el hecho de que la presencia de servicios de telecomunicaciones (incluso distintos a BAF) constituye una señal de que existen condiciones positivas para que los operadores desarrollen una cadena de elementos de infraestructura y operación que le permitan atender en una zona específica ofreciendo servicios a la población.

D. Datos asociados a marginación. La Comisión Nacional de Población (CONAPO)

E. Datos asociados a desarrollo humano. El Programa de las Naciones Unidas para el Desarrollo (PNUD) es una organización orientada a generar soluciones a los países que buscan alcanzar sus metas de desarrollo y lograr los objetivos compartidos con la comunidad internacional. Como parte de sus actividades en México ONU (2015), periódicamente evalúan el nivel de desarrollo de los municipios a través de la construcción del Índice de Desarrollo Humano (IDH), el cual considera los ejes de salud, educación e ingreso.

Al respecto, puesto que se considera que el ingreso de la población es un componente relevante en el acceso a servicios de telecomunicaciones, se considerarán los datos de ingreso anual

F. Otros datos socio-demográficos. hola

Sections: Theory, Methods, Results, Discussion...

Use section and subsection commands to organize your document. \LaTeX handles all the formatting and numbering automatically.

1. Figures and Tables

Figures and Tables should be labeled and referenced in the standard way using the `\label{}` and `\ref{}` commands.

Figure 1 shows an example of how to insert a column-wide figure. To insert a figure wider than one column, please use the `\begin{figure*}...\end{figure*}` environment. Figures wider than one column should be sized to 11.4 cm or 17.8 cm wide. Use `\begin{SCfigure*}...\end{SCfigure*}` for a wide figure with side captions.

Tables. In addition to including your tables within this manuscript file, PNAS requires that each table be uploaded to the submission separately as a “Table” file. Please ensure that each table .tex file contains a preamble, the `\begin{document}` command, and the `\end{document}` command. This is necessary so that the submission system can convert each file to PDF.

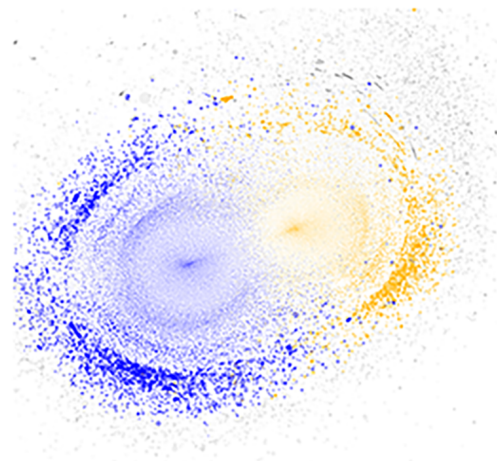


Fig. 1. Placeholder image of a Network with a long example caption to show justification setting.

$$\begin{aligned}(x+y)^3 &= (x+y)(x+y)^2 \\ &= (x+y)(x^2 + 2xy + y^2) \\ &= x^3 + 3x^2y + 3xy^2 + y^3.\end{aligned}\quad [3]$$

Equations. Authors may use 1- or 2-column equations in their article, according to their preference.

To allow an equation to span both columns, use the `\begin{figure*}...\end{figure*}` environment mentioned above for figures. Using only `\begin{figure*}...\end{figure*}` keeps the equation in a two column format

References

References should be cited in alphabetical order; this will be done automatically via bibtex. All references should be included in the main manuscript file.

ACKNOWLEDGMENTS. Please include your acknowledgments here, set in a single paragraph. Please do not include any acknowledgments in the Supporting Information, or anywhere else in the manuscript.

References

- CONAPO (2010-2015). Índice de marginación por entidad federativa y municipio 2010 y 2015. *Página electrónica de la Comisión Nacional de población*.
- Escobar, R. (2017). Disponibilidad municipal de la Infraestructura de Telecomunicaciones. *Página electrónica del Centro de Estudios del Instituto Federal de Telecomunicaciones*.
- Hanafizadeh, P., Hanafizadeh, M., and Khodabakhshi, M. (2009). Extracting core ict indicators using entropy method. *The Information Society*, 25:236–247.
- IFT (2017). Las telecomunicaciones a 4 años de la Reforma Constitucional en México. *Página electrónica del Instituto Federal de Telecomunicaciones*.

- IFT (2018a). Accesos a Banda Ancha Fija crecen 37% a partir de la Reforma en Telecomunicaciones (Comunicado 019/2018) 4 de marzo. *Página electronica del Instituto Federal de Telecomunicaciones*.
- IFT (2018b). Banco de Información de Telecomunicaciones. <https://bit.ift.org.mx/BitWebApp/>.
- IFT (2018c). Manual de definiciones de los indicadores estadísticos de telecomunicaciones. *Página electronica del Instituto Federal de Telecomunicaciones*.
- IFT (2019). México entre los países con mayor crecimiento de Banda Ancha Fija: OCDE (Comunicado 32/2019) 9 de julio. *Página electronica del Instituto Federal de Telecomunicaciones*.
- INEGI (2015). Encuesta Intercensal 2015. <https://www.inegi.org.mx/programas/intercensal/2015/>.
- INEGI (2018). Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH) 2018. *Página electronica del Instituto Nacional de Estadística y Geografía*.
- Katz, R. (2012). Impact of broadband on the economy. *International Union of Telecommunications*.
- Katz, R. (2018). La digitalización: Una clave para el futuro crecimiento de la productividad en América Latina. *Centro de Estudios de Telecomunicaciones de America Latina*.
- Moya, J. (2014). *Telecomunicaciones. Tecnologías, Redes y Servicios. 2ª edición actualizada*. Profesional. Grupo Editorial RA-MA.
- ONU (2010-2015). <https://www.mx.undp.org/content/mexico/es/home/library/poverty/informe-de-desarrollo-humano-municipal-2010-2015-transformando-.html>. *Página electrónica de la Comisión Nacional de población*.
- Pradhan, R. P., Arvin, M. B., Norman, N. R., and Bele, S. K. (2014). Economic growth and the development of telecommunications infrastructure in the g-20 countries: A panel-var approach. *Telecommunications Policy*, 38(7):634 – 649.