

Estrategias de análisis automático de datos urbanos: movilidad, turismo y desigualdad

Mauricio Savarro

Dpto. de Informática - Facultad de Ingeniería
UNPSJB
Puerto Madryn, Argentina
savarromauricio@gmail.com

Leo Ordinez

Laboratorio de Investigación en Informática
(LINVI) - Facultad de Ingeniería - UNPSJB
Puerto Madryn, Argentina
leo.ordinez@gmail.com

RESUMEN

Los territorios deben enfrentar nuevos desafíos para el diseño de estrategias de desarrollo dentro de un contexto de mayor complejidad, incertidumbre y velocidad de cambios, así como adquirir mayores competencias, adaptarse a las exigencias del mercado y avanzar hacia el desarrollo del mismo. El trabajo planteado aquí se asienta en la ciudad de Puerto Madryn, como caso de estudio de una ciudad intermedia. El objetivo general del proyecto es la investigación en técnicas, métodos, estrategias y teorías que asistan en la obtención y procesamiento de datos e información para la construcción de conocimiento de los procesos involucrados en el planeamiento urbano. Para el mencionado fin, se propone el estudio de la literatura relacionada, la construcción de aplicaciones de software que permitan la captura de datos por parte de los ciudadanos, el análisis de dichos datos, la validación de los resultados obtenidos y la sintetización de los mismos.

Palabras Clave: Scrapping - Análisis de Datos - Aprendizaje Automático - Ciudades inteligentes - Informática urbana.

CONTEXTO

El presente forma parte del trabajo en curso dentro del marco de la beca EVC CIN 2019 otorgada a Mauricio Savarro. La misma forma parte de una serie de líneas de investigación recientes desarrolladas en Laboratorio de Investigación en Informática y

en articulación con otros centros y facultades de la UNPSJB. En particular, la beca se enmarca en el proyecto de investigación UNPSJB-PI 1494 “Datos espacio-temporales en entornos urbanos” (2019/2020, prorrogado), cuyo Director es el Dr. Ing. Leonardo Ordinez. Este proyecto se propone investigar y aplicar técnicas, métodos y teorías que asistan en la obtención de información del ámbito urbano, caracterizada por ser heterogénea, proveniente de diferentes orígenes y especialmente ligada a datos espacio-temporales. Vale resaltar que este proyecto se complementa con el UNPSJB-PI 1541, titulado “Desigualdad, experiencia y movilidad urbana en una ciudad turística, Puerto Madryn, Chubut (2008-2018)”, el cual es dirigido por el Dr. Sergio Andrés Kaminker y co-dirigido por el Dr. Ordinez. En este caso, el abordaje es desde las ciencias sociales

1. INTRODUCCIÓN

Durante gran parte del siglo XX, la idea de que una ciudad podía ser inteligente era una ciencia ficción. Manuel Castells (1995) define a la sociedad actual en un imbricamiento interdependiente e inseparable de los procesos sociales y aquellos que se dan en el mundo digital, constituyendo una nueva forma de relacionamiento en lo que él denomina Era Informacional. La convergencia de las tecnologías de la información y la comunicación está produciendo entornos urbanos que son bastante diferentes de todo lo que hemos experimentado hasta ahora. Las ciudades se están volviendo inteligentes no

sólo en términos de la forma en que podemos automatizar funciones rutinarias que atienden a personas individuales, edificios, sistemas de tráfico, sino de forma que nos permitan monitorear, comprender, analizar y planificar la ciudad para mejorar la eficiencia, equidad y calidad de vida para sus ciudadanos en tiempo real (M. Batty et al., 2012). Las ciudades inteligentes a menudo se representan como constelaciones de instrumentos a través de muchas escalas, que están conectados a través de múltiples redes que proporcionan datos continuos sobre los movimientos de personas y materiales, en términos del flujo de decisiones sobre la forma física y social de la ciudad. Sin embargo, las ciudades solo pueden ser inteligentes si pueden integrar y sintetizar estos datos para algún propósito; así mejorar la eficiencia, la equidad, la sostenibilidad y la calidad de vida en las ciudades (Michael Batty, 2013).

La propuesta exploratoria de diferentes enfoques de Brigitte Lamy (2006), resulta básica para comprender las distintas aproximaciones al objeto de estudio de “lo urbano”. Lo trascendental, se encuentra en la mirada compleja sobre un fenómeno multicausal, dialéctico, imbricado y altamente dinámico, que dificulta la comparación o universalización dogmática de teorías. La creciente urbanización está haciendo que las autoridades de las ciudades, los gobiernos, las empresas e incluso los ciudadanos comiencen a pensar en formas alternativas de administrar los recursos dentro de una ciudad. Este objetivo se logra no solo buscando más eficiencia, es decir, hacer más con menos, sino también buscando formas de lograr un mayor nivel de satisfacción entre los ciudadanos y los agentes económicos dentro de una ciudad (Aguilera, Lopez-de-Ipina, & Perez, 2016).

Como explica Urry (2003), la mayoría de los estudios de transporte se han realizado desde la economía o la ingeniería, prestando poca atención a la dimensión subjetiva o la experiencia de la movilidad. Sin embargo, en las últimas décadas existió un cambio de paradigma, a partir del giro que las ciencias

sociales dieron hacia el estudio de los flujos, desplazamientos y movibilidades en general como crítica a una visión más estática de distintos fenómenos, como las distintas expresiones de la desigualdad urbana (Segura, 2012).

Hannerz (1986) introduce cinco dominios en el estudio de la movilidad urbana: doméstico, aprovisionamiento, recreación, vecindad y tránsito. Los estudios de la movilidad de las últimas décadas van desde los micromovimientos corporales, la movilidad cotidiana, residencial y profesional de las personas, hasta los flujos globales de bienes y servicios (Cresswell, 2006; Urry, 2007; Jensen, 2009; Gutiérrez, 2012). Allí se identifican tiempos, rutas, ritmos, experiencias, velocidades, etc. (Cresswell, 2010). Esto implica una diversidad de análisis que importan a esta investigación sobre movilidad urbana, por un lado, y turismo, por el otro.

Como explican Jirón e Imilán Ojeda (2018), en el caso de la movilidad urbana, analizar, tanto el movimiento como las experiencias y los significados que emergen, ha sido una forma de ampliar e incluso cuestionar las nociones clásicas sobre el transporte, centradas en la eficiencia del desplazamiento desde un punto A a un punto B.

Por su parte, los estudios de turismo relativos a transporte suelen referirse al vínculo entre los sistemas de transporte y las vías de circulación de larga o media distancia. En general, la gran mayoría de éstos analizan el vínculo entre transporte aéreo y turismo (Knobel, 2009; Wallingre, 2010, 2013), transporte terrestre y turismo (Ballent y Gorelik, 2001; Wallingre, 2015), transporte marítimo y turismo (Losano et al. 2008; Martínez, 2012). Por otro lado, recientemente se le ha comenzado a prestar atención a la movilidad urbana de las ciudades relativa al turismo y la capacidad de la población residente de los destinos de acceder a determinados bienes y servicios.

En cuanto a estudios de la desigualdad urbana, en América Latina, se ha prestado poca atención a la movilidad. Asimismo, se debe recordar que las movilidades son muy diversas y se experimentan dependientes de una serie de desigualdades (Jirón, 2007), de género, socioeconómica, de edad, origen, etc.

De lo anterior surge la necesidad de destacar la importancia de los datos en este proceso. En este escenario, las personas participan como sensores sociales que proporcionan voluntariamente datos que capturan sus experiencias de la vida cotidiana (Silva et al., 2016). Ello supone una posibilidad de análisis de información para la toma de decisiones de política pública inédita (De Dios Ortúzar y Willumsen, 2008). El entendimiento de los datos generados por la población en el medio urbano, permite corregir una gran cantidad de problemas que afectan la vida diaria y perjudican la eficiencia de las ciudades (BID, 2016), en tanto degradan la calidad de vida y promueven la inequidad. En este sentido, desde la Informática, se han planteado diferentes estrategias, técnicas y herramientas para la obtención de ese conocimiento sobre lo urbano de forma automática (Batty, 2018; Laurini, 2017; Bavsar, 2017; Abduljabbar, 2019; Zhang, 2011).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La masiva disponibilidad de incontables cantidades de datos y el gran poder de cómputo deslocalizado han sido factores determinantes para el surgimiento y auge de la denominada ciencia de datos. Este paradigma se basa en técnicas probadas, que se apoyan en modelos estadísticos de regresión. En virtud de ello, este proyecto de iniciación a la investigación busca profundizar en técnicas y tecnologías, como la extracción automática de datos, el aprendizaje de máquina y los análisis de regresión, con el objetivo de obtener conocimiento sobre el entorno urbano. En particular, se toma como caso de estudio la ciudad de Puerto Madryn. En este sentido, se

busca comprender cómo se expresa y experimenta la *desigualdad* en la *movilidad urbana* en la ciudad de Puerto Madryn, integrando a la diversidad socioeconómica y residencial propia, la mirada de los y las *turistas* y comprender cómo se entrelazan estas dimensiones con las políticas públicas relativas a movilidad y turismo en una ciudad costera patagónica en expansión.

3. RESULTADOS OBTENIDOS

Desde lo metodológico, se busca adquirir las habilidades básicas para el desarrollo de la actividad investigativa en informática. En términos generales, se busca investigar diferentes estrategias, técnicas y modelos, basados en el procesamiento de datos de fuentes heterogéneas, para el análisis de la movilidad urbana en ciudades intermedias y su relación con el turismo y la desigualdad. Más específicamente, se pretende: 1) relevar técnicas y métodos para la sistematización de datos, mediante diferentes tipos de esquemas relacionales y no relacionales; 2) estudiar y desarrollar técnicas para el análisis automático de datos de disponibles en servicios externos de datos abiertos y fuentes primarias (encuestas, relevamientos, datos propios generados en el proceso de investigación); 3) profundizar el conocimiento en el análisis y modelado de datos espacio-temporales.

En esta primera etapa del proyecto se decidió reorientar el trabajo, postergando aspectos de movilidad y comenzando por el análisis de precios de inmuebles (casas y departamentos, en particular) en la ciudad de Puerto Madryn, Chubut. Esto se debió a que el período de trabajo del becario inició en julio de 2020, en el marco de la pandemia por COVID-19. Con este objetivo, el primer punto de este trabajo consistió en obtener información actualizada de valores inmobiliarios. Mediante el uso de scraping se obtuvo información de diferentes webs de diferentes sitios inmobiliarios. La herramienta utilizada fue Web Scraping, un plugin para navegador gratuito con componentes de pago.

Aquí surgió la primera complejidad del trabajo. Ya que la mayoría de las webs poseen información incompleta e inconsistente entre ellas. Es decir, la falta de algunos parámetros útiles para la valuación de un Inmueble como número de ambientes, si posee cochera, etc. Con lo cual el set de datos inicial debió ser trabajado arduamente hasta conseguir un punto de partida para la implementación de los diferentes algoritmos de Aprendizaje Automático. Para esto se revisaron los datos para corregir cualquier fallo en el Scrapping o faltante de información en las webs y se seleccionaron los parámetros finales que formarán parte del conjunto de entrenamiento.

El siguiente punto fue el análisis de los datos donde destaca la generación de una Matriz de Correlación, para ver rápidamente correlaciones entre los diferentes parámetros del modelo, aspecto que sirvió para depurar aún más los datos, junto con la eliminación de *Outliers*, Valores Atípicos que no representan la realidad.

A partir de aquí, se decidió implementar algoritmos de Aprendizaje Automático Supervisados, para generar un modelo de predictor. Se dividieron los datos en dos conjuntos, de entrenamiento para y de prueba. Los algoritmos elegidos fueron Regresión Lineal Múltiple, Regresión Keras (Redes Neuronales), K-Vecinos más cercanos y *Random Forest*.

Como conclusión general de esta primera etapa se destaca la experiencia adquirida en el análisis y depuración de la datos, para una posterior implementación de algoritmos de Machine Learning. Eran campos completamente nuevos que fueron explorados a través de una inmersión directa, que sin duda sirvió para motivar y marcar los aspectos más débiles de cara a continuar en futuras etapas con otros ámbitos de estudio dentro del marco del proyecto. Los siguientes puntos de análisis serán sobre datos recopilados a través encuestas realizadas en la ciudad sobre el turismo y un conjunto de datos provistos por el gobierno nacional sobre el uso de la tarjeta Sube en la ciudad, con el

objetivo de obtener información sobre los recorridos más comunes realizados por los ciudadanos.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo primario sobre esta temática está conformado por los autores del presente. No obstante, debido a los proyectos de investigación en los que se enmarca el trabajo de beca EVC CIN de Mauricio Savarro, las y los integrantes de dichos proyectos hacen las veces de soporte y expertos a disposición para consultas. En particular, entre ambos proyectos se contabilizan siete investigadores/as, cuatro becarios/as y siete estudiantes, que se encuentran haciendo pasantías o tesinas de grado.

5. BIBLIOGRAFÍA

1. Abduljabbar, Rusul; Dia, Hussein; Liyanage, Sohani; Bagloee, Saeed A. (2019). "Applications of Artificial Intelligence in Transport: An Overview." *Sustainability* 11, no. 1: 189.
2. Aguilera, U., Lopez-de-Ipina, D., & Perez, J. (2016). Collaboration-Centred Cities through Urban Apps Based on Open and User-Generated Data. *Sensors*, 16(7), 1–25. <https://doi.org/10.3390/s16071022>
3. Ballent A. y A. Gorelik (2001): "País urbano o país rural: La modernización territorial y su crisis", en Alejandro Cattaruzza, Crisis económica, avance del Estado e incertidumbre política, Sudamericana, Buenos Aires.
4. Batty, M. (2013). Urban Informatics and Big Data, 1–36.
5. Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Portugali, Y. (2012). Smart cities of the future. *European Physical Journal: Special Topics*, 214(1), 481–518. <https://doi.org/10.1140/epjst/e2012-01703-3>
6. Batty M. (2018) "Artificial intelligence and smart cities". *Environment and Planning B: Urban Analytics and City Science*, SAGE Journals, (2018), Vol. 45(1) pp. 3-6.
7. Castells, Manuel (1995). "La ciudad informacional. Tecnologías de la

- comunicación, reestructuración económica y el proceso urbano-regional”
8. Cresswell, T. (2006) *On the Move: Mobility in the Modern Western World*, Routledge, New York.
 9. Cresswell, Tim (2010) “Towards a politics of mobility”, *Environment and Planning D*, N° 28, pp. 17-31.
 10. De Dios Ortúzar, J. y L. Willumsen (2008): *Modelos de transporte (Vol. 1)*. Ed. Universidad de Cantabria.
 11. Gutiérrez, A. (2012): “¿Qué es la movilidad? Elementos para (re) construir las definiciones básicas del campo del transporte”, *Bitácora*, N° 21, pp. 61-74.
 12. Hannerz, U. (1986). *La exploración de la ciudad. Hacia una antropología urbana*. Buenos Aires, Fondo de Cultura Económica.
 13. Jensen, O. (2009): “Urban Mobility as Meaningful Everyday Life Practice”, *Mobilities*, N°4, pp. 139-158.
 14. Jirón, P. (2007): “Prácticas de Movilidad Cotidiana Urbana: Un Análisis para Revelar Desigualdades en la Ciudad”, en Tironi, M. y Pérez, F. *Espacios, Prácticas y Cultura Urbana*, ARQ Ediciones, Escuela de Arquitectura, pp. 176-189.
 15. Jirón, P. y W. Imilán Ojeda (2018), “Moviendo los estudios urbanos. La movilidad como objeto de estudio o como enfoque para comprender la ciudad contemporánea” *Quid 16*, Revista del Área de Estudios Urbanos del Instituto Gino Germani, N°10 – Dic. 2018-Mayo 2019, Facultad de Ciencias Sociales, UBA, pp.17-36.
 16. Knobel, Horacio (2009). “El transporte aéreo de pasajeros y sus equipajes” Editorial Ladevi, Buenos Aires.
 17. Lamy, Brigitte (2006). “Sociología urbana o sociología de lo urbano”, en “Estudios Demográficos y Urbanos”, Vol. 21, No 1.
 18. Laurini, Robert. (2017). *Geographic Knowledge Infrastructure: Applications to Territorial Intelligence and Smart Cities*.
 19. Losano, P., Trifaro, N., Arévalo, A. y A. Tagliorette, (2008): “Implicancias económicas, sociales y ambientales derivadas de la actividad de los cruceros turísticos en la ciudad de Puerto Madryn, Chubut, Patagonia Argentina”, Trelew, Facultad de Ciencias Económicas, Universidad Nacional de la Patagonia San Juan Bosco.
 20. Martínez, C. (2012). “Perspectivas del Turismo de Cruceros en Argentina en el Marco de las Tendencias Mundiales”, *Notas en Turismo y Economía*, Año III. Nro. IV, Universidad del Caribe, México, pp. 44-71.
 21. Segura, Ramiro (2012) “Elementos para una crítica de la noción de segregación residencial socio-económica: desigualdades, desplazamientos e interacciones en el periferia de La Plata”, *Quid 16*, N° 2, pp. 106-132.
 22. Silva, Thiago & Celes, Clayson & Borges, João & Mota, Vinícius & Felipe, Cunha & Ferreira, A.P.G. & Ribeiro, A.I.J.T. & Vaz de Melo, Pedro & Almeida, J.M. & Loureiro, A.A.F.. (2016). *Users in the urban sensing process: Challenges and Research Opportunities*. 10.1016/B978-0-12-803663-1.00003-6.
 23. Urry, J. (2003): *Global Complexity, Polity*, Cambridge.
 24. Urry, J. (2007): *Mobilities*, Polity Press, Cambridge.
 25. Wallingre, N. (2010): “Limitaciones del transporte aerocomercial al desarrollo del turismo. Análisis del caso Transporte interno en Argentina”, *Anuario Turismo y Sociedad*, Vol. X.
 26. Wallingre, N. (2013): “Transporte aéreo, su relación con el turismo y los servicios”, Ediciones Turísticas Fundación ProTurismo, Buenos Aires.
 27. Wallingre, N. (2015): *Transporte carretero argentino*, 4ª edición. Ediciones Turísticas LADEVI, Buenos Aires.
 28. Zhang, J. Wang, F. Wang, K. Lin, W. Xu, X. Chen, C. (2011)“Data-Driven Intelligent Transportation Systems: A Survey,” in *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624-1639.