

Modelado del Riesgo de la Vivienda Urbana para la Salud y el empleo de Redes Neuronales Artificiales para su Estimación

María del C. Rojas¹, Juan C. Vázquez²,
Julio J. Castillo², Marina E. Cardenas²

¹ *Instituto de Investigaciones Geohistóricas (IIGHI-CONICET), Chaco, Argentina.*

² *Universidad Tecnológica Nacional – Facultad Córdoba – Dpto. Ing. en Sistemas de Información, Córdoba, Argentina.*

Resumen:

El presente proyecto se enmarca en una línea de investigación relacionada con la aplicación de técnicas de aprendizaje automático, a la resolución de problemas de las ciencias sociales. El grupo de investigación nuclea a investigadores de la UTN y IIGHI-Conicet, y tiene como uno de sus logros la creación de un software y un modelo computacional que implementa la metodología de estimación del riesgo para la salud de la vivienda urbana desde el enfoque de la vulnerabilidad social, propuesta por el IIGHI-CONICET. Este software se ha denominado RVS (por sus siglas de Riesgo, Vivienda y Salud) y está basado en redes neuronales artificiales. El dominio del problema del riesgo de la vivienda urbana se supone complejo y no lineal, y ha sido modelado empleando redes neuronales artificiales que emula una red conceptual de relación de factores sociales, económicos y demográficos que no se ajusta a los modelos clásicos. El cálculo de un índice de riesgo permitiría a las autoridades de salud dirigir más acertadamente los fondos disponibles, en la medida que la metodología sea confirmada por trabajo de campo, tarea que ya está en marcha en varios países de América Latina.

Palabras clave: riesgo en salud; redes neuronales artificiales; sistema complejo; sistema no lineal.

1. Introducción

Las autoridades gubernamentales ocupadas en temas de salud, tienen siempre un presupuesto acotado para una problemática sensible y compleja, por lo que la decisión de dónde, cuándo y cuántos fondos asignar en salud, no es una tarea trivial y es de gran relevancia para el bienestar de la población que dicha tarea sea llevada a cabo de manera efectiva.

Un grupo de investigadores del Instituto de Investigaciones Geohistóricas (IIGHI) del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), bajo la dirección de la Dra. Norma Meichtry y la Dra. María del Carmen Rojas, han desarrollado una metodología para la estimación del riesgo de la vivienda urbana para la salud humana [1],[2]; utilizando un enfoque holístico, se estudian los factores que determinan la vulnerabilidad social, considerando a la vivienda como un todo con su entorno, los servicios públicos disponibles, los aspectos sociales y económicos de sus residentes, la capacidad de reacción frente a emergencias, las situaciones de riesgo, entre otros. Los factores involucrados en la determinación del riesgo se han seleccionado premeditadamente de aquellos que pueden encontrarse usualmente en los censos de población, con lo cual el cálculo puede realizarse hacia atrás en el tiempo para estudiar evolución y verificar el ajuste del modelo frente a estadísticas de salud existentes.

Como es usual en las Ciencias Sociales, es prácticamente imposible efectuar mediciones del impacto sobre la salud de un determinado factor de riesgo en presencia de todas las otras variables constantes, por lo que es difícil atacar el modelado del dominio con herramientas

matemáticas (por ejemplo, ecuaciones diferenciales o algebraicas); por otro lado, si bien la metodología propone la incidencia de los factores tenidos en cuenta sobre conceptos demográficos no mensurables como resiliencia, exposición y fragilidad social [2], los expertos del IIGHI-CONICET estiman que estas relaciones son en general complejas y no lineales, aunque por su experiencia pueden armar “ejemplos” de valores de los factores y los resultados esperados que los mismos tienen sobre el riesgo de salud.

Por otra parte, el modelo computacional para la adquisición de datos y el cálculo del riesgo según esta metodología ha sido desarrollado por un grupo de investigadores del Laboratorio de Investigación de Software del Departamento de Ingeniería en Sistemas de Información de la UTN-Facultad Regional Córdoba.

El presente artículo describe brevemente la metodología del IIGHI-CONICET desde el punto de vista informático/computacional, se comentan las posibles mejoras al esquema desarrollado y los próximos pasos previstos en ese sentido. En la sección 2 se expone la Metodología de Estimación del Riesgo para la Salud, la sección 3 se muestra el modelo computacional desarrollado y sus supuestos, en la sección 4 se comentan las pruebas, discusión y difusión de los resultados alcanzados, y finalmente la sección 5 se resumen las conclusiones.

2. Metodología de estimación del riesgo para la salud

La metodología pertenece al campo del conocimiento y a la práctica de la Medicina Social en América Latina y reconoce como fuente a la tesis doctoral sobre Estimación holística del riesgo sísmico utilizando sistemas dinámicos complejos¹, desarrollada por Cardona Arboleda [3].

Actualmente constituye uno de los elementos básicos para el desarrollo de la línea de investigación sobre vigilancia ambiental que se está trabajando conjuntamente con investigadores de Paraguay, Brasil, Cuba y Venezuela en el marco de la Red Interamericana de Vivienda Saludable de la Organización Panamericana de la Salud (OPS).

Pensando en espacios de intervención donde poder crear herramientas sistemáticas de valoración de procesos protectores y deteriorantes de las condiciones de salud-enfermedad, se presenta como un interesante ámbito de trabajo el estudio y la valoración de la vivienda y su influencia en la salud de sus residentes.

La idea básica surge de la necesidad de fortalecer los sistemas nacionales y locales de vigilancia de los factores de riesgo y protección a la salud asociados con la vivienda, a través del diseño de nuevos modelos que contribuyan a generar alternativas para el desarrollo saludable y sustentable regional latinoamericano.

Se intenta así mutar el concepto de factores linealmente relacionados con el riesgo en salud, por el de proceso dinámico; se podría decir entonces, que los modos de devenir que determinan la salud se desarrollan mediante un conjunto de procesos y estos procesos adquieren proyecciones distintas frente a la salud, de acuerdo a los condicionantes sociales de cada espacio y tiempo, es decir de acuerdo a las relaciones sociales en que se desarrollan.

Consecuentemente, es necesario aproximarse a la vivienda no como un reservorio estático de contaminantes, parásitos, y vectores de transmisión de enfermedades infecciosas, sino como un espacio históricamente estructurado donde también se expresan las consecuencias benéficas y destructivas de la organización social, donde los procesos del espacio construido

¹ Omar Darío Cardona Arboleda es ingeniero de la Universidad Nacional de Colombia. Fue distinguido por el premio Sasakama de Prevención de desastres de Naciones Unidas.

llegan a ser mediadores necesarios y donde se transforman las condiciones de reproducción social dominantes en bienes o soportes que favorecen la salud, o en fuerzas destructivas que promueven la enfermedad.

Los investigadores del IIGHI-CONICET han propuesto un modelo que trabaja con información de censos de población para la evaluación sociodemográfica y planillas de recuento de datos de gobiernos locales, para la valoración de la resiliencia. Se procura establecer una tipología de viviendas urbanas en relación con la salud humana, con validez para los diferentes niveles sociales, pero su aplicación concreta inmediata, se referirá a las viviendas urbanas marginales. Se presentan los componentes de la amenaza y la vulnerabilidad a partir de las variables de la vivienda que constituyen la amenaza para la salud y variables sociodemográficas que conforman la vulnerabilidad.

El modelo de determinación del índice de riesgo para la salud que implica la vivienda urbana se muestra en la figura 1.

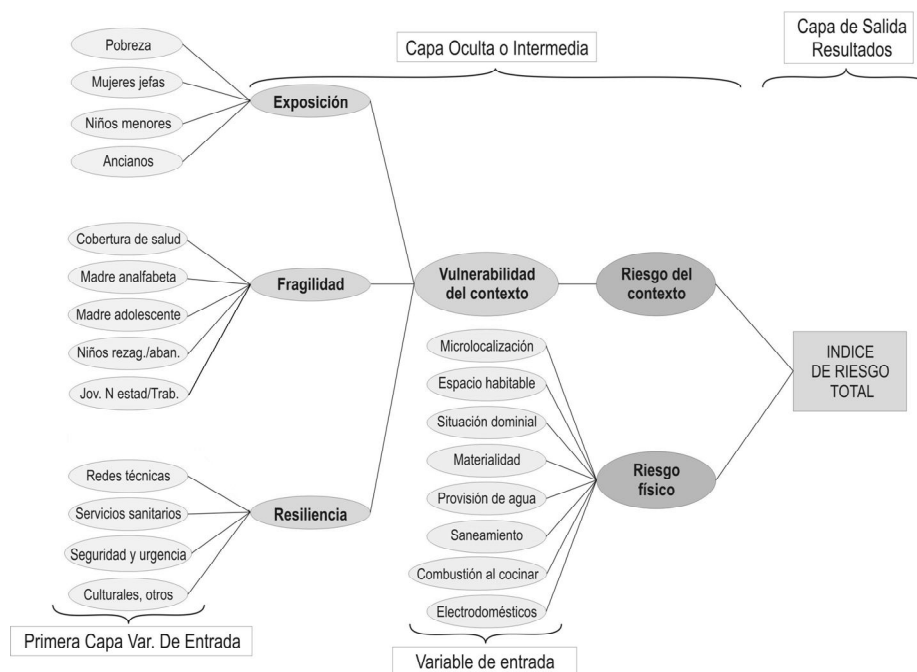


Figura 1 – Modelo conceptual de estimación del riesgo propuesto por IIGHI-CONICET.

Como se puede observar, los demógrafos han pensado el esquema teniendo en mente la teoría de redes neuronales artificiales (por la nomenclatura de capas empleada) pero sin ajustarse a un modelo conocido; se incluyen por ejemplo, entradas intermedias dentro de la denominada capa oculta para el cálculo del riesgo físico y se omiten capas intermedias reales necesarias para el cálculo de sub-conceptos.

3. Modelo computacional desarrollado

3.1. Supuestos.

Si bien al parecer los factores propuestos como entrada no son absolutamente independientes, los expertos del IIGHI consideran que la influencia mutua que pudiera existir se encuentra reflejada en las interrelaciones planteadas en el modelo conceptual. Por otra parte, los valores de los factores tomados como entrada siempre se serán porcentajes, ya que sobre un total de habitantes de una región, se informan cantidad de hogares que detentan el cumplimiento de determinado factor, por lo cual todos los casos son porcentajes (enteros de cero a cien).

Para los conceptos intermedios y finales no mensurables (resiliencia, exposición, fragilidad, vulnerabilidad, riesgo físico y riesgo total), los expertos indicaron que podrían aproximar una estimación porcentual basada en su experiencia, según los valores de factores mensurables tabulados de cero a cien. Sin embargo, esta estimación sería difusa, ya que proporcionaron intervalos en vez valores individuales, por lo que se procedió a establecer una escala de conjuntos borrosos triangulares, basados en los datos suministrados.

3.2. Diseño del modelo computacional.

Ante la imposibilidad de establecer un modelo matemático mediante el uso de ecuaciones algebraicas o diferenciales, ya que se desconoce la relación formal existente entre los factores y los subconceptos y entre los subconceptos y el índice total, se pensó en modelar computacionalmente el modelo conceptual demográfico con redes neuronales artificiales (RNA).

Esta decisión se basó también en el formato del modelo propuesto por el IIGHI-CONICET y por la complejidad y no linealidad declarada por sus investigadores, respecto del dominio de problema.

En este sentido se pensaron dos alternativas:

- a) Una sola red neuronal en la que todos los factores mensurables fueran puestos al mismo nivel, a modo capa de entrada, con una sola salida compuesta por el riesgo total (figura 2).
- b) Un conjunto de varias redes neuronales que calculasen en forma independiente cada subconcepto y finalmente el índice de riesgo total, cada una con sus capas intermedias. (figura 3)

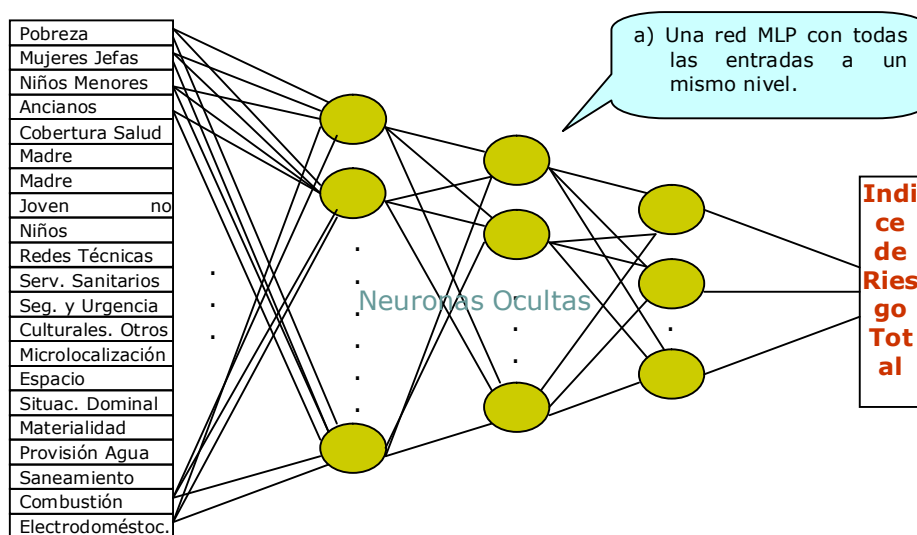


Figura 2 – Primera propuesta para el modelo computacional.

La primera alternativa fue descartada debido a que se perdían los conceptos intermedios (resiliencia, exposición, fragilidad y vulnerabilidad) que son considerados de importancia por los investigadores de IIGHI y que podrían ser utilizados para otras investigaciones posteriores.

Por otro lado, los expertos tenían dificultades para estimar los valores de los subconceptos dados un conjunto de valores de prueba para los factores de entrada, por lo que la estimación del índice de riesgo total teniendo que considerar todos los factores sociodemográficos al mismo tiempo se complicaba o al menos era muy poco confiable.

La segunda alternativa (figura 3) resultó más conveniente debido a que:

- Los conceptos intermedios podían ser calculados para su posterior tabulación y utilización en otras investigaciones.
- El criterio de los expertos resultaba más claro teniendo que hacer estimaciones de problemas más acotados; estas estimaciones serían de vital importancia para el entrenamiento de las redes neuronales ya que proporcionarían los casos de entrenamiento necesarios para el método de retro-propagación de errores.
- El modelo computacional se asemeja más al modelo conceptual del IIGHI.

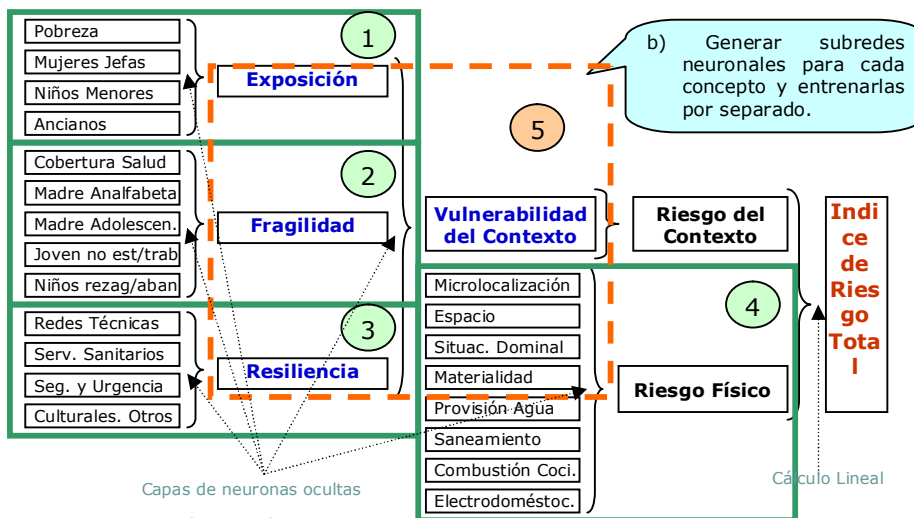


Figura 3 – Segunda propuesta para el modelo computacional.

Se diseñaron finalmente, cinco redes neuronales artificiales tipo perceptrón multicapa que se entrenarían con el algoritmo de retropropagación de errores; el cálculo final del índice de riesgo se haría en forma lineal según ponderaciones indicadas por los expertos.

Las redes neuronales se diseñaron como perceptrón multicapa puro, con neuronas de capa oculta gobernadas por tangente hiperbólica y neuronas de salida lineales; las neuronas de la capa de entrada sólo normalizarían los valores de ingreso para que pertenezcan al rango real [0,1].

El desarrollo del software fue realizado con tecnología Microsoft usando Visual Studio .Net para la implementación y el modelado orientado a objetos para el diseño programas. Este desarrollo fue llevado a cabo por el equipo de trabajo del proyecto RNA-AC – 25/E078 de la Facultad Córdoba de la UTN [4],[5], y [6].

3.3. Módulos construidos

En principio, se construyeron dos módulos bien diferenciados:

1. **Módulo de entrenamiento:** este módulo configura las redes neuronales en base a archivos de configuración en formato texto, que permiten definir la cantidad de capas y la cantidad de neuronas por capa para cada subred. Se fijan además por programación, el error medio cuadrático aceptado como razonable para terminar los entrenamientos. Al finalizar, el módulo crea un archivo descriptivo de la topología neuronal utilizada, el tipo de subred, la fecha de entrenamiento, la cantidad de ejemplos tomados y el valor de los pesos sinápticos obtenidos como resultado del entrenamiento.
2. **Módulo de producción:** este módulo (RVS 2.1) es el que se dispone para transferencia a distintos centros de investigación de estudios de población y salud comunitaria. Leyendo el archivo descriptivo generado por el módulo de entrenamiento, configura la *red entrenada* en memoria y permite el ingreso de valores reales de poblaciones; efectúa el cálculo por propagación sobre la red de cada subconcepto y del índice total, informando en cada caso los porcentajes de participación según los conjuntos borrosos definidos. Genera además informes de valores exportables a planillas Excel para su registro y posterior estudio.

No mostramos aquí las distintas pantallas que demuestran el funcionamiento de estos módulos, pero los mismos pueden consultarse en el informe final del proyecto 25/E078, disponible en la Secretaría de Ciencia, Técnica e Innovación Productiva (hoy Ministerio) de la República Argentina o en la Secretaría de Ciencia y Técnica del Rectorado de la Universidad Tecnológica Nacional en Buenos Aires, Argentina.

4. Pruebas, discusión y difusión

El software desarrollado está siendo probado por los investigadores en demografía del IIGHI-CONICET utilizando datos reales censados en regiones de las ciudades de Córdoba y Resistencia de Argentina, con resultados favorables hasta el momento. En las pruebas iniciales se detectó un problema con el tratamiento de valores nulos de entrada que fue reparado en el software; las neuronas de la red perceptron no aprendían su tarea adecuadamente cuando propagaban sus entradas en cero, por lo que procedió a desfasar los valores en 0,01 en caso de valores nulos sin pérdida de las cualidades del modelo conceptual.

El software y la metodología base han sido presentado en distintos foros:

- Congresos de la Asociación de Estudios de Población de Argentina.
- Seminario Iberoamericano de Hábitat Popular, Centro de Estudios Avanzados de la Universidad Nacional de Córdoba.
- II Congreso de la Asociación Latinoamericana de Estudios de Población, Guadalajara, México.
- Primer Congreso Internacional Mexicano de Vivienda Saludable, Pachuca de Soto Hidalgo, México.

Por otro lado, la Organización Panamericana de la Salud ha manifestado su interés en efectuar pruebas de metodología y programas en su red latinoamericana de vivienda saludable; la primer transferencia fue realizada en la Facultad de Arquitectura y Urbanismo

de la Universidad Católica de Nuestra Señora de la Asunción, en Paraguay y están en previstos similares convenios con Universidades de Brasil, Cuba y Venezuela.

En los casos estudiados, los resultados del software coinciden con la opinión de los expertos en cuanto al riesgo previsible para la salud. Sin embargo es temprano aún para declarar éxito; la metodología debe verificar estadísticamente sus previsiones, el software debe ser llevado a situaciones límites para comprobar su corrección, incluso los ejemplos propuestos por los demógrafos están bajo revisión permanente.

5. Conclusiones

En variados campos de las Ciencias Sociales, un tablero de comandos que brinde información de variables críticas o algún sistema de decisión a medida, siempre pueden ser útiles, requiriendo un grupo de expertos que interprete esas informaciones y tome las acciones preventivas o correctivas que sean necesarias.

Pero ¿qué hacer cuando se desconoce la relación existente entre los factores y sus determinados?, cuando se supone complejo y no lineal el dominio de problema.

Las redes neuronales artificiales pueden emplearse en estos casos, pudiendo aprender el criterio de los expertos en el campo mediante el aprendizaje a partir de ejemplos concretos. De esta manera, los resultados obtenidos pueden ser puestos a disposición de las autoridades gubernamentales que deben decidir sobre la asignación de fondos, con la posibilidad de efectuar simulaciones y comparaciones pertinentes.

Finalmente, las pruebas de campo servirán para someter a validación esta metodología, y efectuar los ajustes que sean necesarios al modelo teórico-computacional, y de esta manera aportar una herramienta más a este proceso de toma de decisiones en salud colectiva y familiar.

Referencias

1. Rojas, MC. La vivienda precaria urbana marginal y su relación con la salud de la población en el proceso de sustentabilidad. Un enfoque teórico para la estimación del riesgo y la vulnerabilidad. I Congreso de la Asociación Latinoamericana de Población -ALAP-, CD-ROM. Asociación Brasileira de Estudos de Población -ABEP-. [Ponencia] Caxambú -MG- Brasil; 2004
2. Rojas, MC. Población, vivienda salud y vulnerabilidad global. Propuesta teórico-metodológica para la estimación del riesgo de la vivienda urbana para la salud basada en el análisis de la vulnerabilidad sociodemográfica. [tesis doctoral] Facultad de Ciencias Económicas de la Universidad Nacional de Córdoba; Córdoba, Argentina; 2006
3. Cardona A. Estimación holística del riesgo sísmico utilizando sistemas dinámicos complejos. Tesis Doctoral, Barcelona, España, 2001.
4. Rojas, M del C, Meichtry, N. C, Vázquez, J. Castillo, M. B, Ciuffolini. Repensando de manera holística el riesgo de la vivienda urbana precaria para la salud. Un análisis desde el enfoque de la vulnerabilidad sociodemográfica. Revista de Salud Colectiva. Universidad Nacional de Lanás, Argentina
5. Rojas, M del C, Meichtry, N. C, Vázquez, J. Castillo, M. B, Ciuffolini. Crítica de la noción de vivienda saludable desde el realismo dialéctico. Modelización matemática basada en redes neuronales y variables difusas para la estimación del riesgo de la vivienda urbana para la salud a partir del análisis de la vulnerabilidad sociodemográfica. XXVI Encuentro de Geohistoria Regional. Universidad Nacional de Asunción. Facultad de Filosofía. 16, 17 y 18 de agosto de 2007, Asunción, Paraguay

6. Rojas, M del C, Meichtry, N. C, Vázquez, J. C, Marciszack M., Modelización Numérica de la Vivienda Precaria Urbano Marginal a partir de la Estimación Holística del Riesgo para la Salud de la Población considerando Variables de Vulnerabilidad Sociodemográficas, VIII Jornadas Argentinas de Estudio de Población (AEPA 2005), CD ISBN 950-658-158-4, 27 pp., Tandil, Bs. As., Argentina
7. Rojas María del Carmen, Meichtry Norma C., Vazquez Juan C., Castillo Julio J., Ciuffolini M. B.. “Repensando de manera holística el riesgo de la vivienda urbana precaria para la salud. Un análisis desde el enfoque de la vulnerabilidad sociodemográfica”. Revista de Salud Colectiva. Universidad Nacional de Lanús.2008.