

Professional networks in online learning processes / Redes profesionales en procesos educativos en línea

Sofía García¹, Antonio López², Nicolás Rubido³, Pablo Massaferrero⁴, and Alvaro Margolis²

¹Flacso Uruguay & Fundación Ceibal, Uruguay

²Evimed, Uruguay

³ Universidad de la República, Inst. de Física. Facultad de Ciencias, Uruguay.

⁴ Universidad de la República, Inst. de Ingeniería Eléctrica, Uruguay.

Resumen En el presente estudio, se analiza un curso de educación médica continua en línea. Se han investigado los beneficios de utilizar técnicas de Análisis de Redes Sociales y Minería de Datos para predecir y mejorar el desempeño de los estudiantes, así como para optimizar la división en grupos de estudiantes y docentes, para mejorar los resultados del curso. Específicamente, se utiliza información de los participantes, compartida antes del curso, como el motor para crear una red de conocidos, que evoluciona a medida que el curso evoluciona. Estos vínculos sociales entre los participantes permiten caracterizar sus interacciones en relación con el curso, por lo tanto, estableciendo el papel que desempeñan en su red social, por ejemplo, como hubs o outliers. En particular, esto permite agrupar a los participantes según su cercanía social. Además, al utilizar técnicas de minería de datos educativos, puede pronosticarse el rendimiento potencial de los estudiantes en el curso y diseñar estrategias educativas en consecuencia. Esta investigación muestra la importancia de trabajar en equipos multidisciplinarios para abordar los problemas sociales, y también para brindar a la educación conocimientos y técnicas de investigación que se han utilizado con éxito en otros campos de investigación.

Abstract. In the present study, an online continuing medical education course is analyzed as it evolves. We investigate the benefits of using Social Network Analysis and Data Mining techniques to predict and improve the students' performance, as well as to tailor the course into suitable groups of students and teachers, in order to enhance the course results. Specifically, we use information from participants, shared prior to the course, as the motor to create an acquaintance network, which evolves as the course evolves. These social links between participants allow us to characterize their interactions in relationship to the course, hence, establishing the role they play in their social network, for example, as hubs or outliers. In particular, this allows us to group participants according to their social closeness. Moreover, using Educational Data Mining techniques, we can forecast the students' potential performance in the course and design instructional strategies accordingly. This research shows the importance of working in multidisciplinary teams to address

social problems, and also, to bring forth to education research knowledge and techniques that have been successfully used in other research fields.

Keywords: Social Network Analysis · Online education · Professional development · Educational Data Mining.

1. Introducción

En la era de la información y la comunicación, los estudios relacionados a las redes sociales y profesionales han adquirido cada vez más relevancia. En particular, se destaca la línea de investigación conocida como Social Network Analysis (SNA), que estudia los vínculos entre las personas basándose en información empírica y utilizando modelos matemáticos para el análisis (Luke, 2007). Se define una red como un conjunto de nodos conectados entre sí por medio de un conjunto de enlaces (Easley, 2010). El análisis de los datos que se obtienen mediante SNA busca determinar las propiedades de nodos, subgrupos de nodos y enlaces, y las características generales de la red (Luke, 2007). En el área de la salud, el SNA se utiliza fundamentalmente para comprender la importancia de las redes en cuestiones de salud pública, tales como la propagación del VIH-SIDA o para investigar los modos en que los médicos adoptan las innovaciones científicas.

El acceso a las tecnologías de la información y la comunicación (TIC) por parte de la población en general ha cambiado radicalmente en los últimos 20 años. Particularmente, en el ámbito de la medicina el crecimiento del uso de Internet ha sido notorio, de tal forma que ya en el año 1997 el 54 % de los médicos de Estados Unidos utilizaban Internet. Este porcentaje creció 24 puntos para el año 2001, alcanzando más del 86 % en 2012 (Cooper, 2012). Asimismo, el incremento en el uso de las TIC es acompañado por un crecimiento en la generación y disponibilidad de datos sobre las personas. Sin embargo, para procesar estos grandes volúmenes de información se hacen necesarias técnicas de minería de datos (*data mining*), que permiten el procesamiento automático o semi-automático de los datos para modelar procesos complejos y extraer información relevante (Mohamad, 2013). En la última década, estas técnicas han tomado fuerza en el análisis de los procesos educativos en general, y en la educación virtual, en particular: surge la minería de datos educativa (EDM, *educational data mining*). EDM, así como SNA, son dos campos disciplinares que tienen gran potencial para la mejora de las experiencias educativas en entornos virtuales de aprendizaje. Su crecimiento se ha visto potenciado por la capacidad computacional de almacenar y analizar grandes cantidades de datos (Calvet, 2015). “Dos aproximaciones separadas, que son efectivas individualmente en la identificación de variables, son aún más efectivas si se utilizan juntas en un enfoque combinado” (Miller, 2015).

Se ha destacado la importancia de la Educación Médica Continua (EMC) como un aspecto fundamental en la mejora de la atención de la salud de la población. En este contexto, la educación a distancia se ha convertido en un pilar

fundamental de muchas propuestas educativas en esta área. Si, como menciona Castells (2001), el aprendizaje en comunidades pre-existentes aumenta la motivación y potencia el impacto de las acciones educativas, ¿ocurre esto en las actividades de EMC a distancia? Históricamente los médicos se han agrupado en sociedades científicas o colectivos profesionales, ¿en qué medida las interrelaciones que establecen inciden en los procesos de aprendizaje en la educación a distancia? (Margolis, 2015). Partiendo de estas cuestiones, el objetivo general del presente trabajo es analizar la incidencia de los vínculos profesionales en los procesos de aprendizaje en un curso de educación médica continua en línea.

2. Marco conceptual

Dado el carácter social de los procesos educativos mediados por tecnología, surge el concepto de aprendizaje en red, que es aquel que “se produce en el marco de un entramado de vínculos sociales. Cuando este entramado se encuentra orientado a la construcción colaborativa de conocimiento adquiere el formato de una red de aprendizaje” (Caldeiro, 2014).

Por un lado, las redes de aprendizaje son compuestas por miembros interdependientes que comparten su compromiso con metas específicas y trabajan colaborativamente para lograrlas (Aveling, 2012), son constituidas por un “grupo de personas que aprende en común, utilizando herramientas comunes en un mismo entorno” (García Fernández, 2002). Una de las herramientas por excelencia en educación en línea para el aprendizaje en interacción con otros es el foro de intercambio académico. Las técnicas de SNA permiten investigar las relaciones entre los individuos de una red, así como la estructura de la misma, y cómo esto da soporte al aprendizaje. Puede incluir un análisis individual de cada participante o uno general de la red en su totalidad.

Indicadores de red – El grado (en inglés *degree*) se refiere a la cantidad de vínculos que tiene cada nodo. O sea, dada una persona (nodo de la red), el grado establece con cuántas personas está vinculada. Las personas con las que este nodo se vincula se denominan “vecinos”. Este indicador se puede estudiar con o sin direccionalidad. Cuando se trabaja con redes no direccionadas, como en el presente estudio, se asume que si una persona conoce a otra, se conocen mutuamente. La intermediación del nodo (*node betweenness centrality*) refiere a la cantidad de caminos más cortos entre cada par de nodos de la red que pasan por un determinado nodo. Para calcular la intermediación del nodo debe primero hallarse es el camino más corto que une a cada par de nodos. Aplicando este procedimiento para todos los nodos de la red se busca establecer si hay nodos que recurrentemente se encuentran en esos caminos, es decir, nodos que poseen una intermediación alta. En general, un nodo con intermediación alta puede tener mayor importancia para la diseminación de información en la red.

La centralidad de vector propio (en inglés *eigenvector centrality*) trasciende la centralidad dada por los propios vínculos, ya que se fija en los vínculos de

los vecinos y en lugar de otorgar un punto por cada vecino, este indicador define un valor proporcional al promedio del grado de los vecinos. En definitiva mide la influencia potencial de un participante en la red, teniendo en cuenta si está vinculado con participantes importantes o si él mismo tiene gran cantidad de contactos: “un individuo en una red social, puede ser importante, según esta medida, porque conoce a mucha gente (incluso si esa gente no es importante en sí misma) o conoce a poca gente en altos lugares” (Newman, 2010).

El coeficiente de agrupamiento (en inglés *clustering coefficient*) indica en qué medida una serie de nodos están vinculados entre sí. Mide la probabilidad de que un par de nodos conectados entre sí tengan una conexión en común, permitiendo visualizar la disgregación de una población heterogénea en subpoblaciones más cohesivas (Kleinberg, 1998). En otras palabras, responde a la pregunta de cuántos de los vecinos que posee un nodo se conocen entre sí con respecto a la posibilidad de que todos sus vecinos se conozcan entre sí. Según Newman (2010): “en promedio los nodos de mayor grado tienden a tener menor agrupamiento local”. En este sentido, la información brindada por este indicador es complementaria a la de los indicadores vistos anteriormente, ya que no reviste tanta importancia la cantidad de vínculos de un participante, sino cuánto se conocen entre sí una porción de los participantes. Este es un dato sumamente relevante para la potencial conformación de comunidades de aprendizaje.

En los casos en los que tres o más nodos de la red están todos interrelacionados entre sí, se forman *cliques*. Los *cliques* son grupos de nodos que forman estructuras geométricas (como las que se exhiben en la Fig. 1), por lo que encontrar *cliques* permite visualizar si dentro de la red hay subgrupos de nodos que estén más conectados entre sí que con el resto.

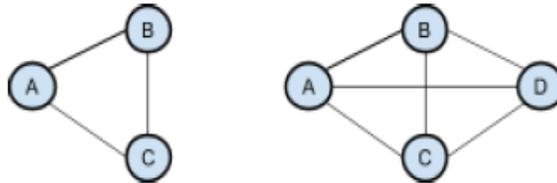


Figura 1. Ejemplo de *cliques* de 3 (panel izquierdo) y 4 (panel derecho) nodos.

Estos *cliques* se enmarcan o conectan con la red a través de los vínculos que uno o varios de ellos tenga con otros nodos que no pertenezcan al *clique*. Este proceso de vinculación genera pequeñas comunidades, que son subgrupos de nodos interconectados más densamente entre ellos que con el resto de la red (Danon et al, 2008). La estructura de comunidades puede favorecer el proceso de aprendizaje tanto en el espacio virtual como fuera de él, ya que los participantes se conocen y es posible incluso que trabajen juntos en la atención de pacientes en un mismo centro de salud.

3. Metodología

La red que se analiza en este estudio es la conformada por los participantes del curso “Cuidados continuos en la Enfermedad Renal Crónica” de la Sociedad Latinoamericana de Nefrología e Hipertensión (SLANH). El curso se realizó en 2016 durante ocho semanas en forma virtual y estaba dirigido a médicos nefrólogos, enfermeras, nutricionistas y otros profesionales de la salud de toda América Latina que atienden pacientes con enfermedad renal crónica, siempre en clave de atención integral interdisciplinaria, marcada por la propia especificidad de las patologías tratadas. Cada módulo incluyó materiales teóricos (textos breves o videos) preparados por expertos especialmente para el curso y un foro de caso clínico, en el que se planteaban preguntas buscando promover el diálogo entre participantes. Por un lado, para el análisis de datos de red y su relación con el proceso educativo, se utilizó una muestra no probabilística de participación voluntaria, obteniendo información de 423 profesionales de un total de 601 participantes del curso. Según los antecedentes relevados en la literatura, para recabar información de las redes sociales pueden utilizarse básicamente dos metodologías: 1) cuestionarios diseñados para conocer los vínculos entre los miembros de la red, 2) análisis computacional de las relaciones a través de dispositivos electrónicos. En la recolección de datos a través de cuestionarios puede obtenerse información en diferentes niveles, según requiera la investigación. Para la presente investigación, se desarrolló una herramienta que permite una recolección de datos de red semi-automatizada, en la que los participantes pueden realizar el proceso con poco esfuerzo, dado que, por la cantidad de contactos potenciales, el llenado manual de un formulario resultaría tedioso.

Recolección de datos de red – Para la recolección de datos de la red (vínculos entre los participantes) se solicita a los participantes del curso que autoricen al sistema a acceder a sus contactos de Gmail, mediante el uso de un botón dispuesto en el campus virtual para ello, como se observa en la Fig. 2.



Figura 2. Herramienta de recolección de datos de red

Cuando el cursante otorga la autorización, el sistema se conecta con Gmail, cruzando los contactos de la persona con la base de participantes registrados en

el curso. La información que el sistema trae se refleja en una lista de participantes (con nombre y foto) ordenados por país, siendo el primer país en aparecer el del propio usuario que realiza el proceso. El participante debe confirmar dicha lista, pudiendo agregar o quitar contactos si lo desea. Esta información da lugar a una red que representa los participantes y sus contactos conectados por enlaces no direccionados.

Indicadores de red – Al utilizar la clasificación propuesta por Newman (2010) hay ciertas propiedades de las redes que resultan importantes a la hora de realizar un análisis de las mismas. Para este caso, se ha optado por los siguientes indicadores del nivel microscópico: grado, intermediación, centralidad de vector propio y coeficiente de agrupamiento local; y del nivel mesoscópico la conformación de *cliques*. Son los indicadores de red los que pueden proveer información para la mejora de los modelos pedagógicos y tutoriales de cursos de EMC en línea.

Clasificación – Para las clasificaciones se utiliza *Random Forest*, entrenando el algoritmo a partir de una validación cruzada con cinco *folds*. El sistema ajusta los parámetros del clasificador de forma de maximizar la correcta clasificación de los cursantes, generando un modelo que contempla los indicadores más relevantes del proceso. Se obtienen la precisión por clase y la matriz de confusión. Pueden obtenerse además el ranking de correlación de cada atributo con las clases definidas, así como gráficas ilustrativas de estas relaciones.

4. Resultados

Características de la red – Con la información obtenida a partir de 370 respuestas se ha constituido una red de 423 nodos (el 70% de los participantes activos), con un total de 1148 enlaces entre ellos.

Para iniciar el análisis de los aspectos microscópicos de la red, se relevan los datos relativos a los nodos individuales. En primera instancia, se asigna a cada uno el grado, es decir, la cantidad de contactos que tiene en la red profesional del curso. Se obtiene que en esta red de 425 el grado promedio es de 6 (o sea que en promedio, cada participante conoce a 6 personas) con una distribución que se asemeja a una libre de escala (*power law*), como la que Barabási (2002) ha descrito para distintas redes sociales (ver histograma en Fig. 4).

El coeficiente de agrupamiento indica en qué medida un nodo está vinculado con pares de vecinos que se conocen entre sí. Resulta de interés, que 92 participantes tengan un coeficiente de agrupamiento 1 (el máximo posible). Asimismo, se encontraron 279 *cliques*. Este es un aspecto que reviste interés práctico para la mejora de las propuestas pedagógicas. Los *cliques* de una red y los vínculos de cada nodo, son datos que pueden facilitar la división en grupos de trabajo en entornos seguros que permitan una apertura de los participantes a contar sus historias de la práctica para mejorar su desempeño futuro. Tradicionalmente la

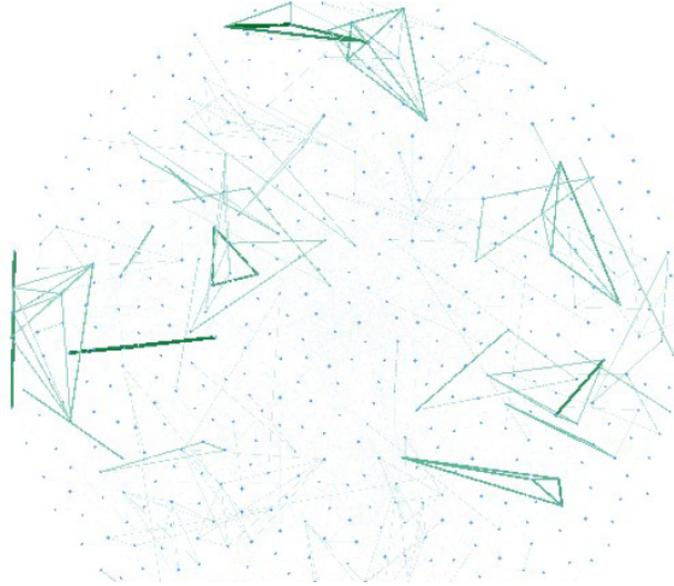


Figura 3. Grafo de la red

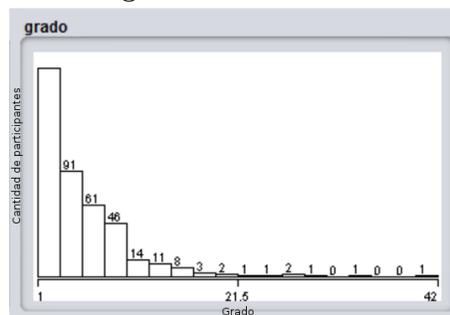


Figura 4. Histograma grado

división en grupos en un curso online masivo como el analizado se realiza en forma aleatoria. Sin embargo, la información provista por la red de participantes permitiría una división en grupos de participantes relacionados por afinidad, principalmente, a partir de la información que brinda el análisis de los *cliques*.

Clasificación y predicción – Para la predicción de la actividad en foros se utilizan 10 atributos que permiten clasificar a los cursantes entre aquellos con participación baja en los espacios de intercambio (entre 0 y 3 comentarios) y aquellos con participación alta (más de 4 comentarios), obteniendo una precisión del 65%. Como se observa en la Tabla 1, la precisión para ambas clases es similar, generando un alto porcentaje de verdaderos positivos y un bajo porcentaje de falsos positivos.

Clase	Verdaderos positivos	Falsos positivos	Presición
Entre 0 y 3 comentarios	0,584	0.291	0,638
Más de 4 comentarios	0,709	0,415	0,661
Promedio ponderado	0,651	0,357	0,650

Tabla 1. Precisión por clase

En el análisis sin atributos de red, se obtiene un nivel de clasificaciones correctas del 62 %, siendo la eficacia del sistema menor que con los datos de red. Si bien al analizar la matriz de confusión (Tabla 2) se observa que la clasificación es relativamente buena sin los datos de red (clase a= 0-3 comentarios / clase b= Más de 4 comentarios), los resultados obtenidos muestran que hay una serie de participantes para los que los datos de red son decisivos para su correcta clasificación.

Valor real	Valor predicho	
	Clasificados como "a"	Clasificados como "b"
Verdaderos "a"	106	87
Verdaderos "b"	71	149

Tabla 2. Matriz de confusión para predicción actividad foros sin datos de red

El nivel de clasificaciones correctas obtenido a partir de los datos de red, relaciona la posición de los individuos en la red con su nivel de participación en los espacios de interacción, retomando así los conceptos planteados por Parboosingh acerca de la importancia de la red para establecer espacios de intercambio basados en la confianza para compartir las opiniones, casos y dudas sobre los temas académicos. En este sentido, la información de la red puede, por un lado, permitir la separación en grupos que sean considerados entornos seguros por los participantes, y por otro lado, brindar información para la motivación de aquellos que se predice que no participarán activamente.

5. Discusión

En esta investigación se analiza la incidencia de los vínculos profesionales en los procesos de aprendizaje en un curso de educación médica continua en línea. En este caso, la aplicación de técnicas de SNA para la caracterización de la red y la extracción de indicadores, ha mostrado ser un insumo relevante para predecir la participación, aplicando técnicas de EDM. Por tanto, conocer la red establecida entre los cursantes de una actividad educativa de corta duración puede resultar un aporte útil para generar estrategias pedagógicas que permitan establecer procesos dialógicos para la consecución de cambios concretos en la práctica profesional. La combinación de técnicas de SNA y EDM, es no solamente relevante, sino oportuna en el contexto de cursos de educación en línea masivos

en el área de la medicina y cuyos resultados serán pasibles de ser analizados en otras áreas del conocimiento.

La revisión teórica indica que las actividades de EMC que generan más cambios en la práctica (y por ende, mejoras en la atención de la salud de los pacientes), son aquellas que se basan en las necesidades, incluyen actividades de interacción con otros y propuestas que contemplen las diferentes modalidades de aprendizaje (Lockyer, 2007). Es en las actividades que propongan interacción entre los participantes donde surge la oportunidad de generar diálogos que partan de la propia práctica y puedan volver a esta promoviendo mejoras en el cuidado de la salud de los pacientes. Establecer un diálogo requiere confianza, entornos seguros, en los que los participantes puedan exponer sus inquietudes, dudas y errores. La generación de estos entornos puede propiciarse desde el comienzo del propio curso, pero puede potenciarse si se basa en los vínculos profesionales ya existentes entre los participantes: la red profesional.

En el caso de la red analizada, se concluye que la metodología de obtención de los datos sobre los vínculos de los participantes fue adecuada para un curso masivo. Resulta interesante que el 21% de los nodos tengan un coeficiente de agrupamiento local de uno, siendo este el máximo valor esperado. Del mismo modo, el hallazgo de 279 cliques resulta significativo para la realización actividades grupales que contemplen estas estructuras pre-existentes. Contando con los datos de la red -y específicamente quién se vincula con quién- es posible establecer grupos pequeños de trabajo como entornos seguros para la interacción educativa.

En los entornos virtuales de aprendizaje, la colección de datos y huellas que dejan los cursantes es inmensa y está en crecimiento permanente. La forma en la que se utilizan estos datos, marca un diferencial en las etapas de diseño pedagógico (antes y durante el curso), así como en las estrategias de tutoría, seguimiento y motivación. Así, los vínculos entre los cursantes y los indicadores que estos vínculos proveen, pueden aportar herramientas para la mejora de los procesos educativos, por ejemplo, configurando alertas personalizadas por usuario. Por otra parte, la verificación de la existencia de una relación entre los indicadores de red y la participación activa en los foros, refuerza la propuesta de utilizar estos datos para la generación de grupos de trabajo dentro del curso, que promuevan el aprendizaje activo, la interacción y la generación colectiva de conocimiento. La utilización de la red pre-existente permite una mejor optimización de los recursos y una potenciación de vínculos que se han generado en la práctica previa.

6. Conclusiones

Las técnicas utilizadas en el presente trabajo no son resorte de una única disciplina sino que requieren de la mirada y pericia de profesionales de diferentes áreas, potenciando así, desde la multidisciplinariedad, el alcance que estas pueden tener en el ámbito de la educación en línea, en general, y de la educación médica continua virtual, en particular. La gestión de la abundante información que provee un entorno virtual de aprendizaje, en tanto registra todas las acciones

de los cursantes, se hace no solo oportuna sino necesaria, para procurar procesos formativos cada vez más individualizados, a la vez que en diálogo profundo con otros. Esta dualidad que no es incompatible sino complementaria constituye el principal desafío de la educación en línea masiva. Valorar la importancia de la red profesional en el proceso educativo de un participante centra la atención en la generación de propuestas que promuevan la interacción, con un fuerte anclaje en la práctica, permitiendo que las interacciones se valgan de las relaciones pre-existentes, potencien el proceso de aprendizaje, y vuelvan al ámbito profesional en términos ya no solamente individuales, sino con equipos que se dispongan a mejorar su práctica clínica, con acciones concretas, apoyándose mutuamente en este proceso.

En el contexto actual, los datos de la red de cursantes y las predicciones del comportamiento en actividades en línea, son aspectos que pueden aportar a la gestión personalizada de estudiantes, tanto en pequeños cursos como en cursos masivos. A partir de estos datos, puede generarse un sistema de alertas dinámico, que opere en las diferentes etapas de una propuesta educativa en línea, para trabajar en la motivación y el seguimiento individualizado, sea por parte de un tutor-docente en un curso pequeño, como por un sistema automatizado de tutoría en un curso masivo.

Referencias

1. Aveling, E., et al: Quality improvement through clinical communities: eight lessons for practice. *Journal of Health Organization and Management*, Vol. 26. 158-174 (2012).
2. Barabási, A.: *Linked*. Perseus, USA (2002).
3. Caldeiro, G.: *Dinámicas colaborativas y tensiones de la EeL*. FLACSO Argentina - Homosapiens Ediciones, Buenos Aires (2016).
4. Calvet, L.: Educational Data Mining and Learning Analytics: differences, similarities, and time evolution. *RUSC. Universities and Knowledge Society Journal*, Vol 12 (N° 3) (2015).
5. Castells, M.: *Lección inaugural del programa de doctorado sobre la sociedad de la información y el conocimiento*. UOC, (2001).
6. Cooper, C. et al: Physicians Who Use Social Media and Other Internet-based Communication Technologies. *J Am Med Inform Assoc*, Vol. 19 (N° 6), Págs. 960-964 (2012).
7. Danon, L. et al: Impact of community structure on information transfer. *Physical Review*, N° 77 (2008).
8. Easley, D., Kleinberg, J.: *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press (2010).
9. García Fernández, N.: Sistema de trabajo con las TIC en el Sistema educativo y en la formación de profesionales. *Revista de educación a distancia*, N° 6 (2002).
10. Luke, D., Harris, J: Network analysis in public health: history, methods, and applications. *Annual Review of Public Health*, Vol. 28., 69-93 (2007).
11. Mansouri, M. y Lockyer, J.: A meta-analysis of continuing medical education effectiveness. *JCEHP*, Vol. 27. 6-1 (2007).

12. Margolis A, Parboosingh J.: Networked Learning and Network Science: Potential Applications to Health Professionals' Continuing Education and Development. JCEHP, Vol. 35. 211-9 (2015).
13. Miller: A comparison of educational statistics and data mining approaches to identify characteristics that impact online learning. Journal of Educational Data Mining, Vol. 7 (N° 3) (2015).
14. Mohamad, S., Tasir, Z.: Educational Data Mining: A Review. Procedia - Social and Behavioral Sciences, Vol. 97 (2013).
15. Newman, M.: Networks, an introduction. Oxford University Press, New York (2010).
16. Parboosingh, J., et al: Enhancing practice improvement by facilitating practitioner interactivity: New roles for providers of continuing medical education. JCEHP. Vol. 31 (N° 2), 122-127 (2011).