

Análisis del impacto de las redes de interacción de estudiantes en el desempeño académico en un sistema de tutoría inteligente



Colaboración

Itzel Berenice Guerrero Alfaro; María Andrade-Aréchiga; JRG Pulido; Silvia B. Fajardo-Flores, Universidad de Colima

RESUMEN: En este estudio, se analizan las redes de interacción académica de grupos de estudiantes universitarios para determinar el impacto que tienen los grupos de colaboración, que los mismos estudiantes integran, para realizar actividades académicas, así como entender la dinámica de dicha colaboración. Para este fin, se realizaron análisis estadísticos, así como un análisis de las interacciones de cada grupo de estudio basado en métricas derivadas de la teoría de grafos, nuestra principal contribución. Los resultados indican que existe una influencia estadísticamente significativa del uso de un Sistema de Tutoría Inteligente (STI) en el desempeño académico de los estudiantes. Además, se pudo corroborar que los alumnos dependen de sus redes de trabajo académico colaborativo para mantener el promedio aprobatorio de sus calificaciones, ambos aspectos de gran importancia en la vida académica de todo estudiante.

PALABRAS CLAVE: Desempeño académico, Grafos, Métricas, Redes de interacción, Sistema de tutoría inteligente.

ABSTRACT: In this contribution, the academic interaction networks of university student groups to determine the impact that team-collaboration, student groups, to carry out academic activities, and understand the dynamics of such a collaboration. To this end, a statistical analysis has been developed, as well as an analysis of group interaction based on graph metrics, our main contribution. The results indicate that there exists a statistically significant influence of the use of an Intelligent Tutoring System (ITS) in the academic performance of the students. Furthermore, it was corroborated that students rely on their academic networks in order to maintain good academic grades, both important aspects in the academic life of students.

KEYWORDS: academic performance, network interaction, Graphs, Metrics, Expert Tutoring System.

INTRODUCCIÓN

En la Universidad de Colima, México, se diseñó un Sistema de Tutoría Inteligente (STI) que se define como un sistema de cómputo basado en inteligencia artificial en el cual se maneja el contenido académico, brinda asesoría e interactúa con el estudiante [1, 2].

El STI fue realizado con el fin de apoyar y dar tutoría a los estudiantes, de tal manera que el alumno fortalezca el conocimiento a través de un sistema de inteligencia [3]; así como de apoyo para disminuir los índices de reprobación y deserción estudiantil, esto en concordancia con [4] que señala que uno de los principales problemas con los que se enfrentan las Universidades hoy en día es el bajo rendimiento académico de los alumnos. Esta situación se presenta normalmente durante los primeros semestres de los estudios universitarios, ya que gran parte del alumnado presenta distintas y diversas problemáticas, entre ellas un promedio bajo, si no es que reprobatorio; por lo tanto en la mayoría de los casos esto conlleva la deserción estudiantil. En esta etapa, es primordial

el trabajo académico colegiado entre los estudiantes y se intuye que las redes de trabajo académico influyen de manera importante en el rendimiento académico.

Este STI fue desarrollado por investigadores y profesores de la Universidad de Colima en conjunto con la Universidad Autónoma de Baja California. En la actualidad el STI se está implementando con grupos de la generación 2016-2020 de estudiantes de la primera institución. Para la obtención de los datos necesarios se aplicaron 4 encuestas en el módulo de diagnóstico del sistema (las cuales abarcan las áreas de: Hábitos de Estudio, Autoestima, Expresión Oral y Escrita, y Matemáticas [5]), además de una encuesta sobre las redes de trabajo académico y el registro de los promedios de las calificaciones de los participantes.

Diagnóstico en el STI

El STI desarrollado cuenta con cinco módulos de interacción [5], que son: Diagnóstico (consta de cuatro instrumentos para recabar los conocimientos con los que ingresa el estudiante), SITI (sistema integral de tutoría inteligente consistente de contenido educativo y ejercicios en matemáticas, comunicación oral y escrita, hábitos de estudio y autoestima), SEVI (sistema para la evaluación de vocabulario basado en imágenes), Moodle (manejador de contenido para cursos en línea) y ERúbrica (sistema de gestión de proyectos para redes de colaboración), cuya interfaz se muestra en la Figura 1.

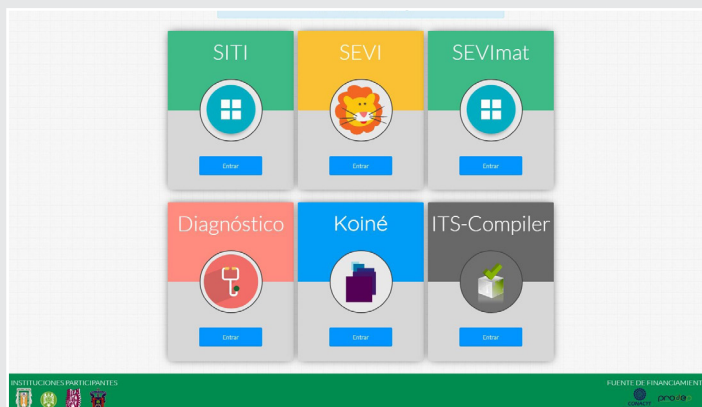


Figura 1. Interfaz del Sistema de Ambiente de Tutoría Inteligente.

Métricas

Una red se define como un conjunto de elementos llamados nodos o vértices con relaciones entre ellos llamados aristas o bordes.

Por medio de las aristas o bordes se pueden representar distintas relaciones entre las redes, las cuales pueden ser: amistad, conocimiento, relación profesional. Además, permiten representar distintos pesos entre las relaciones como por ejemplo: qué tan bien se conocen dos personas. Los diversos pesos en una arista se pueden representar por medio del grosor de la misma.

Para calcular estas relaciones se hace uso de la teoría de grafos [6] que incluye las métricas que se consideraron para analizar las distintas relaciones entre las redes de interacción. Considerando su relevancia, representatividad e importancia en el estudio se eligieron las siguientes:

Coefficiente de Agrupamiento (CA): Calcula la probabilidad de que al elegir un nodo al azar, éste pertenezca a un triángulo.

Densidad de la red (DR): Indica las conexiones potenciales que podrían existir en una red.

Centralización de la red (CR): Indica el tipo de topología, la cual a mayor centralidad se asemeja a una estrella.

Heterogeneidad de la red (HR): Indica el grado de dispersión de los elementos de una red.

Número promedio de vecinos compartidos (NVC): Se refiere al número de socios compartidos entre los nodos. Esta métrica está relacionada con la densidad de red, ya que la densidad es una versión normalizada de la misma.

Diámetro de la red (D): Indica cuáles son las redes con mayor distancia entre sus nodos por lo tanto son los grupos con mayor número de nodos aislados. De manera contraria, a menor diámetro de red, indica que tienen menor distancia entre sus nodos y por lo tanto mayor coeficiente de agrupamiento.

En todo diseño es de gran importancia conocer qué tanto influye en su progreso la contribución de las herramientas, para lo cual en la presente investigación se estudiaron las redes de interacción de estudiantes dentro del ambiente de tutoría inteligente, para determinar la influencia que tienen dichas redes en el rendimiento académico de los alumnos participantes, través de análisis estadístico y análisis automático de redes de interacción.

MATERIAL Y MÉTODOS

En esta sección se describe brevemente el proceso que se llevó a cabo para la obtención de los datos utilizados en esta investigación, del análisis estadístico y del análisis automático de las redes de interacción, mostrado a través de la Figura 2.

Encuesta

Para la recolección de los datos sobre las redes de estudiantes, se utilizó como herramienta una encuesta. Dicha encuesta es de autoría propia, tomando en cuenta criterios para respuestas en escala tipo Likert y validada por un grupo piloto de 5 usuarios potenciales y dos expertos en elaboración de instrumentos. Para su diseño, se consideraron los datos que se necesitaban obtener para evaluar y/o analizar las redes de interacción: elección de grupo de trabajo, grado de influencia en sus calificaciones, forma de trabajo, frecuencia para pedir ayuda, tipo de relación entre compañeros y calificaciones por periodo.

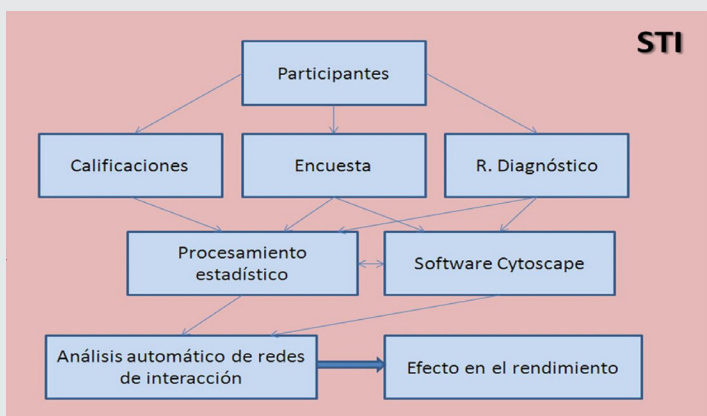


Figura 2. Metodología de la investigación.

Participantes en el estudio

Los participantes que fueron considerados en este proyecto son alumnos de la Facultad de Telemática, de las carreras de Ingeniería en Telemática e Ingeniería de Software. La población del caso de estudio se delimitó a los alumnos que hacen uso del sistema de tutoría inteligente (STI), de la generación 2016-2020, con un total de 73 estudiantes. La relación por grupo se muestra en la Tabla 1.

Tabla 1. Distribución de participantes en el estudio.

Grupo	Carrera	Participantes
1	Ing. en Telemática	13
2	Ing. en Telemática	12
3	Ing. de Software	8
4	Ing. de Software	12
5	Ing. en Telemática	9
6	Ing. de Software	12
7	Ing. de Software	7

Calificaciones

El nivel de desempeño del usuario es una valoración cuantitativa del logro de su aprendizaje en una escala numérica de 0 a 10, considerando como un rendimiento aprobatorio un valor arriba de 6 puntos.

Las calificaciones se obtuvieron por medio del departamento escolar de la Facultad de Telemática. Las que se consideraron para el análisis son las obtenidas tanto en el promedio de bachillerato como las del primer, segundo y tercer semestre de licenciatura de cada uno de los participantes, con las cuales se hizo el análisis de rendimiento.

Software

Se revisaron varias herramientas de software que se utilizan para el análisis de las redes de interacción, tales como Ucinet [7-8], R y Green software [9], Gephi [10], Cytoscape [11], Netmine [12], Pajek [12-13] y NodeXL [13].

Sin embargo, después de compararlos, evaluar el uso y las capacidades que tienen, el software que se decidió

utilizar para el análisis de las redes de interacción entre los estudiantes, fue el Cytoscape.

Cytoscape es una plataforma de software libre, de código abierto, en el que se visualizan redes de interacción de diferentes tipos. Originalmente este software fue diseñado para la investigación biológica. Sin embargo, ahora se conoce por ser una plataforma general para el análisis y visualización de redes complejas [13]. Cuenta también con APIs de desarrollo para varios lenguajes de programación.

Se optó por trabajar con este software por la sencilla capacidad de visualización de las redes, el tipo de análisis automático que realiza y los formatos de exportación e importación con los que dispone, a diferencia de las otras herramientas de software cuyo manejo es complejo.

Métricas consideradas

Las métricas que se consideraron para analizar las distintas relaciones entre las redes de interacción, por la representatividad e importancia en el estudio son: densidad de red, centralidad de la intermediación, coeficiente de agrupamiento, centralidad de proximidad, centralización de la red, vecinos compartidos, y coeficiente topológico [14].

Derivado de estas métricas se obtuvieron los resultados necesarios para determinar las relaciones e impactos entre las redes de interacción, las cuales se detallan en la siguiente sección.

Análisis Estadístico

Para la sección del análisis de la información recabada en el proyecto, se utilizó Excel, R y Minitab [15]. Estas herramientas nos permiten relacionar y administrar los datos obtenidos de las encuestas como de los resultados del STI. Se hicieron diversos análisis de correlación simple y múltiple, así como pruebas de hipótesis y pruebas de normalidad, con el propósito de encontrar la correlación entre rendimiento y los promedios de las evaluaciones en el STI [6].

RESULTADOS

En esta sección se presentan los resultados más relevantes del estudio. Para ello se describe el análisis de las redes de interacción, a través de sus métricas, y del procesamiento estadístico realizado.

Análisis de interacción

Cytoscape cuenta con NetworkAnalyzer que es un complemento de Java para este software. El complemento calcula parámetros específicos que describen la topología de la red [14].

En la Figura 3 se pueden observar las siete redes de interacción, las cuales corresponden a cada uno de los 7 grupos que participaron en el estudio. Cada uno de

los nodos corresponde a un estudiante con la matrícula del alumno. De la misma manera se puede identificar que no existe relación alguna entre los grupos, las relaciones se dan sólo entre los miembros del grupo al que corresponden.

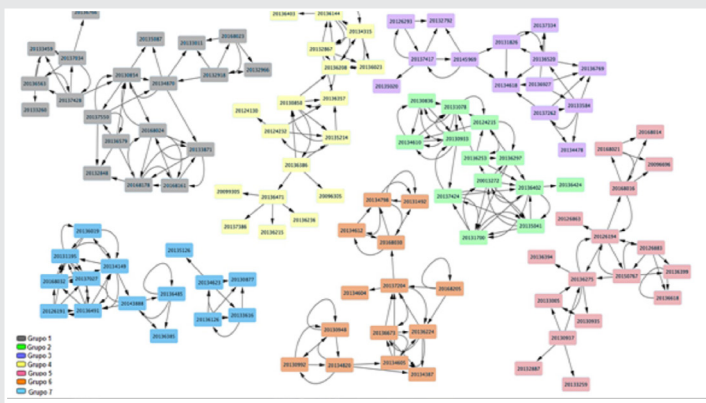


Figura 3. Redes de estudiantes de los grupos 1 al 7 identificadas por matrícula (elaboración propia).

En la Tabla 2 se muestran los resultados obtenidos en Cytoscape, la cual contiene las métricas mencionadas anteriormente y enseguida se hace la descripción de los valores más significativos en los grupos de estudio.

Tabla 2. Resultados de las métricas por grupo, del análisis en Cytoscape.

G	CA	DR	CR	HR	NVC	D
1	0.498	0.179	0.152	0.420	3.400	6
2	0.538	0.333	0.197	0.294	4.000	4
3	0.433	0.275	0.128	0.330	3.571	5
4	0.278	0.170	0.140	0.552	2.889	6
5	0.400	0.175	0.181	0.553	2.625	5
6	0.588	0.242	0.167	0.488	3.143	5
7	0.654	0.233	0.190	0.484	3.500	4

G	Grupo de estudio
CA	Coefficiente de agrupamiento
DR	Densidad de la red
CR	Centralización de la red
HR	Heterogeneidad de la red
NVC	Número promedio de vecinos compartidos
D	Diámetro

El grupo que tiene mayor coeficiente de agrupamiento es el Grupo 7, lo cual indica que este grupo presenta mayor interconexión entre sus compañeros, con un valor de 0.654, con respecto al resto de los grupos. Para la métrica de densidad de la red, la mejor fue la del Grupo 2, con un valor de 0.333, lo cual indica que en esta red existe un mayor potencial de conexiones.

En cuanto a la métrica de centralización de la red, el Grupo 2 fue el que obtuvo mayor puntuación con un valor de 0.197. Esto indica que la topología de esta red es la que más se asemeja a una estrella, por lo tanto presenta mayor centralidad entre los grupos estudiados.

El Grupo 5, fue el que obtuvo mayor puntuación en la métrica de heterogeneidad, con un valor de 0.553, lo cual refleja que tiene mayor número de nodos centrales en la red, esto se puede verificar en la Figura 2 la red de color rosa, ya que se puede observar que existen nodos claves para la relación entre los subgrupos formados dentro de la red. El número promedio de vecinos compartidos, Grupo 2 con un valor de 4, indica que esta red tiene una mayor conectividad. Las redes con mayor diámetro de la red son los grupos 1 y 4, por lo tanto son los grupos con mayor número de nodos aislados.

Análisis Estadístico

Para el análisis estadístico se utilizó el Minitab versión 13.20, particularmente para las pruebas de normalidad, homogeneidad de los grupos, pruebas de hipótesis y grados de correlación.

El Test de normalidad de los grupos se realizó para verificar que los datos que se manejaron provienen de una distribución normal. Esta característica es requisito indispensable para que los análisis posteriores tengan validez estadística, y para ello se aplicó la prueba de normalidad Anderson-Darling a cada grupo. Esta prueba se aplica solamente antes del análisis de las redes. Si "p" es mayor que 0.05 significa que los grupos provienen de una distribución normal. En el caso de que alguno de los grupos no cumplan esta condición, se sugiere utilizar estadística no paramétrica [15].

En la Tabla 3 se aprecia que los grupos presentan un comportamiento derivado de una distribución normal, por lo que los resultados que se obtengan pueden ser manejados con estadística paramétrica sin ningún problema.

La prueba t de muestreo pareado se utilizó para comprobar las hipótesis de la existencia de diferencias estadísticamente significativas entre las medias de dos variables, en este caso se consideró rendimiento contra resultados de las encuestas de diagnóstico en el STI. Si "p" es menor que 0.05 indica que los grupos presentan la intervención del efecto del uso de la plataforma y por lo tanto los grupos se ven favorecidos, o en otras palabras, hay una diferencia estadísticamente significativa quizá debido a la intervención del uso de la plataforma.

Al realizar la prueba "t" de muestreo pareado en este estudio, se obtuvieron los resultados que se muestran en la Tabla 3, lo cual nos indica que en todos los grupos hay una diferencia estadísticamente significativa. En este sentido, se puede considerar que la intervención del uso de la plataforma ha influido positivamente en el rendimiento académico.

La prueba de correlación de Pearson se realizó para determinar el grado de relación entre el rendimiento

(medido por el promedio de las calificaciones de los tres semestres) y el promedio del resultado de las cuatro áreas en el STI. Un valor de “r” mayor que 0.5 indica un alto grado de relación entre las variables [15].

En la Tabla 3 se pueden observar los resultados por grupo al realizar la Prueba de correlación de Pearson, lo cual nos indica que en la mayoría de los grupos existe un alto grado de relación entre las variables, ya que se obtuvo una correlación arriba de 0.5.

Tabla 3. Resultados de la prueba de Hipótesis y Correlación.

Grupo	Test de normalidad	Test de Hipótesis	Coefficiente de Correlación
	P- Value1	P- Value2	r
1	0.101	0.0010	0.538
2	0.932	0.0021	0.432
3	0.211	0.0037	0.499
4	0.171	0.0310	-0.220
5	0.291	0.0020	0.503
6	0.313	0.0050	0.542
7	0.173	0.0290	0.762
P-value1	Indica si los grupos son normales, $p > 0.05$		
P-value2	Indica la aceptación de la hipótesis, $p < 0.05$		
r	Indica el grado de relación entre las variables		

Como ya se mencionó, un aspecto importante en el desarrollo de este proyecto es analizar el efecto del STI en el desempeño de los estudiantes. Para ello se consideraron las calificaciones finales por semestre de cada uno de éstos y se calculó la ganancia en las calificaciones por cada estudiante.

Mientras que, el 59% del total de estudiantes tuvo una ganancia positiva de 0.65 en promedio de primero a segundo semestre, el 57% de los estudiantes registraron una ganancia promedio de 0.69 de primero a tercer semestre. El 8.9% presentó una ganancia negativa de 0.36 puntos, aquí se están incluyendo los alumnos que desertaron por reprobado materias, de primer a tercer semestre. En este sentido se considera que el efecto del uso de la plataforma contribuyó de manera significativa en el estimado de la ganancia promedio positiva.

Al revisar el rendimiento por grupo, el mayor fue el Grupo 7 con 8.81, seguido del Grupo 2 con 8.15, luego el Grupo 6 con 8.02. En cuarto lugar, el Grupo 1 con 7.92, quinto lugar el Grupo 3 con 7.88, después el Grupo 4 con 7.80 y finalmente el Grupo 5 con 7.65. Se aprecia que hay concordancia con las métricas obtenidas para el análisis de las redes de interacción.

CONCLUSIONES

El análisis de las redes de interacción permitió detectar su impacto en el rendimiento académico e identificar conductas tales como: conexiones entre los subgrupos, elección de compañeros, formas de trabajo aca-

démico o de interés personal, de amistad o afinidad en ciertas actitudes, con el fin de alentar y promover aquellas que benefician a los estudiantes y eliminar aquellas que no. Analizar las redes de interacción de estudiantes a través de métricas derivadas de la teoría de grafos es una contribución novedosa de este estudio.

Al hacer las pruebas de hipótesis, se observó que el impacto que tienen las redes de interacción en el rendimiento académico es estadísticamente significativo, dado que en todos los casos $p < 0.05$, y los valores de las métricas de grafos son concordantes con los grupos de acuerdo a su rendimiento. Además, el grado de relación entre las calificaciones y los resultados del sistema de tutoría es alto, en su mayoría.

El efecto del STI y la interacción de las redes académicas en el rendimiento de sus calificaciones fue positivo, reflejado en la satisfacción del trabajo en red de los alumnos participantes en la encuesta aplicada. Ambos aspectos son de suma importancia en la vida académica de todo estudiante por las diversas implicaciones que tiene en ámbitos, tales como el social, técnico, conceptual, y académico.

Como trabajo futuro, una de las acciones importantes por realizar es, si a través de los semestres las interacciones cambiaron o permanecen y cómo éstas pueden seguir influyendo en su rendimiento académico, en la motivación para el estudio, o bien si hay otros patrones de comportamiento social diferentes a las identificadas que se deriven de estas interacciones.

BIBLIOGRAFÍA

[1] Benmarrakchi, F., Ouherrou, N., Elhammoumi, O. & Kafi, J. (2019). An innovative approach to involve students with learning disabilities in intelligent learning systems. *Advances in Intelligent Systems and Computing, Springer*, 914, 39-50.

[2] Durango, J. & Pascuas, Y. (2016). Los Sistemas Tutores Inteligentes y su aplicabilidad en la educación. *Revista Horizontes Pedagógicos*, 17(2), 104-116.

[3] López, D., Álvarez, O., Jordán, A., & Sánchez, D. (2016). Sistema de Tutores Inteligentes para Ayudar a Estudiantes de Bachillerato en la Solución de Problemas de Matemáticas. *Lat. Am. J. Phys. Educ.*, 10(4), 4312-4319.

[4] Pérez Galmiche, G. (2017). La educación superior en México: el rendimiento académico versus deserción académica. *Revista Internacional de Estudios sobre Sistemas Educativos*, 2(6), 75-91.

[5] Damian-Reyes, P., Pulido, J., Andrade-Aréchiga, M., Ramos Michel, E., Acosta, R., Fajardo, S., Hernández, A., Tapia, M. & Benítez, A. (2016). Herra-

mienta de contextualización del sistema integral de tutores inteligentes. Reporte técnico, Universidad de Colima, México.

[6] Pulido, J., Andrade-Aréchiga, M., Ramos Michel, E. & Damián-Reyes, P. (2016). Tecnologías de información para el análisis exploratorio de redes de interacción. En R. Herrera y L. Barbosa (Eds.) *Tópicos introductorios a la gestión del conocimiento* (pp. 131-162). Ediciones de la noche, México.

[7] Zhang, S. & Fang, Y. (2018). Research on Construction Project Organization Based on Social Network Analysis. *Wireless Personal Communications*, 102 (2), 1867-1877.

[8] Marqués-Sánchez, P., Alfonso-Cendón, J., Fernández-Martínez, M., Liébana-Presa, C., Pinto-Carral, A. & Conde, M. (2015). Cooperative networks and their influence on engagement with nursing students. En *Proceedings of the 3rd International Conference on Technological Ecosystems for Enhancing Multiculturality, TEEM '15*, pp. 65-70, New York, NY, USA. ACM.

[9] Marimuthu, C. & Chandrasekaran, K. (2017). Software Engineering Aspects of Green and Sustainable Software: A Systematic Mapping Study. En *Proceedings of the 10th Innovations in Software Engineering Conference, ISEC '17*, pp. 34-44, New York, NY, USA. ACM.

[10] Smite, D., Moe, N., Sablis, A., & Wohlin, C. (2017). Software teams and their knowledge networks in large-scale software development. *Information and Software Technology*, 86(1), 71-86.

[11] Nepomnyachiy, S., Ben-Tal, N. & Kolodny, R. (2015). CyToStruct: Augmenting the Network Visualization of Cytoscape with the Power of Molecular Viewers. *Structure*, 23(5), 941-948.

[12] Dabkowski, M., Breiger, R. & Szidarovszky, F. (2015). Simultaneous-direct blockmodeling for multiple relations in Pajek. *Social Networks*, 40, 1-16.

[13] Domínguez-Rodríguez, R., Santana, J. & Royo, E. (2015). Monitoring technologies in Suricata model for the formal and informal learning areas. *Third International Conference on Technological Ecosystems for Enhancing Multiculturality, ACM*, 125-129.

[14] Wu, L., Li, M., Wang, J. & Wu, F. (2018). CytoCtrlAnalyser: A Cytoscape app for biomolecular network controllability analysis. *Bioinformatics*, 34(8), 1428-1430.

[15] Miller, I. (2000). *Estadística matemática con aplicaciones*. 6ª. Edición. Pearson educación, México.