

# Perfiles de tesis de la disciplina informática y su vinculación con los proyectos de investigación de la UM

Iris Sattolo<sup>1</sup>, Marisa Panizzi<sup>1</sup>, Javier Lafont<sup>1</sup>, Nicolás Armilla<sup>1</sup>

<sup>1</sup> Universidad de Morón. Facultad de Informática, Ciencias de la Comunicación y Técnicas Especiales. Cabildo 134. Pcia de Buenos Aires. Morón. Argentina.

[iris.sattolo@gmail.com](mailto:iris.sattolo@gmail.com); [marisapanizzi@outlook.com](mailto:marisapanizzi@outlook.com); [lafontjavier@hotmail.com](mailto:lafontjavier@hotmail.com); [nicolasarmilla@hotmail.com](mailto:nicolasarmilla@hotmail.com)

**Resumen.** La participación de estudiantes de carreras de grado en proyectos de investigación permite detectar vocaciones para que desarrollen su carrera como investigadores en el ámbito académico y/o reconozcan la importancia de la investigación en el contexto industrial. En este trabajo, el objetivo es descubrir la relación entre las características de los tesis de grado y su vinculación con los proyectos de investigación ejecutados en la UM, como así también la identificación de los tesis de grado que han experimentado de otras actividades vinculadas a la investigación. Para alcanzar el objetivo, se realizó un conjunto de experimentos utilizando los algoritmos Kmeans, Expectation Maximization, J4.8, Naives Bayes, y BayesNet. Se utilizaron los entornos WEKA y Elvira. Los resultados de los experimentos permitieron descubrir que el motivo de la selección de la tesis es el atributo de mayor incidencia hacia la participación en proyectos de investigación como así también en otras actividades de investigación propuestas por la cátedra.

**Palabras clave:** Tesis de grado, Perfiles, Proyectos de investigación, Minería de datos.

## 1 Introducción

La aplicación de la minería de datos en el ámbito de Educación ha adquirido un gran auge en la última década para resolver problemas de diferentes tipos. La Minería de Datos Educativos (Educational Data Mining - EDM) es una disciplina relacionada con el desarrollo de métodos para extraer información útil a partir de los datos que se generan en los entornos educativos, y utilizarla para mejorar dicho entorno. La información así obtenida se convierte en el insumo indispensable para la toma de decisiones [1].

En el contexto internacional, existen publicaciones previas, las cuales han demostrado que la minería de datos se puede utilizar para detectar estudiantes en riesgo de deserción. En su trabajo, Luan [2] aplicó exitosamente técnicas de minería de datos para predecir qué grupos de estudiantes podrían abandonar los estudios. En su trabajo, Lin [3] usó técnicas de minería de datos para optimizar los esfuerzos para retener estudiantes. Investigadores de la Universidad Estatal de Bowie, desarrollaron

un sistema basado en minería de datos que permite a dicha institución, identificar y atender a estudiantes en riesgo de abandono de estudios [4].

En los últimos años, la Universidad de Morón mediante la Secretaría de Ciencia y Tecnología (SeCyT-UM) ha priorizado la investigación dentro de su plan estratégico. La presentación y el desarrollo de los proyectos de investigación están orientados a estimular y afianzar en la UM actividades sistemáticas que están estrechamente relacionadas con la generación, el perfeccionamiento y la aplicación de los conocimientos en Ciencia y Tecnología. Se espera con esta acción promover dentro del campo científico y tecnológico la investigación e innovación, la formación de recursos humanos, las actividades de difusión, el desarrollo de servicios tecnológicos (bases de datos, biblioteca especializada, etc.) y la transferencia de conocimientos a la Comunidad [5].

Este trabajo tiene como objetivo definir los perfiles de los estudiantes que han participado en proyectos de investigación para la realización de su trabajo de fin de carrera. También se pretende lograr caracterizar a estudiantes que hayan vivenciado la escritura de artículos y realizado su presentación en eventos científicos, aunque no hayan formado parte de un proyecto de investigación de la UM.

En la sección 2 se describen los trabajos relacionados, la sección 3 presenta el desarrollo del proceso de minería de datos, se describe cada uno de los pasos realizados y finalmente en la sección 4, se presentan las conclusiones y futuros trabajos.

## 2 Trabajos relacionados

En trabajos anteriores [6], realizamos un estudio de mapeo sistemático siguiendo el proceso propuesto en [7]. El objetivo del SMS (en Inglés, Systematic Mapping Studies), las preguntas de investigación, el formulario de extracción utilizado, el esquema de clasificación de estudios primarios, la distribución de los estudios según la clasificación propuesta y el listado de los estudios primarios utilizados para el SMS, se encuentran en el Apéndice: <https://doi.org/10.6084/m9.figshare.9202445>

De un total de 110 estudios primarios, existen trabajos que definen perfiles de estudiantes; pero en la caracterización de este, no se considera la relación con proyectos de investigación de su universidad. En estos trabajos, tampoco se discute si los estudiantes vivenciaron la escritura de artículos y/o la participación en eventos científicos. Si bien reconocemos que nuestro SMS no es absoluto, nos ha permitido descubrir las características consideradas en otras investigaciones para la definición de perfiles de estudiantes.

## 3 Desarrollo

Para el descubrimiento de la información se aplican las fases del proceso KDD (en Inglés, Knowledge Discovery in Databases, KDD) [8] en correspondencia con trabajos anteriores [9] y [10]. A continuación, se presentan las fases y lo realizado en cada una de ellas.

### 3.1 Fase de Integración y recopilación

Para la realización de esta fase, se utilizó la planilla de cálculo que la cátedra posee, la cual permitió un registro sistemático de las tesis defendidas y el instrumento de recolección de datos construido denominado “TESISTAS-UM” [9]. En este trabajo, se utilizaron las encuestas de los estudiantes recibidos durante los últimos 10 años ya que, desde esa fecha se cuenta con un registro de los proyectos de investigación ejecutados en la UM., resultando una muestra de 97 registros compuesta por 47 graduados de la carrera Licenciatura en Sistemas y 50 graduados de la carrera Ingeniería en Informática.

### 3.2 Fase de selección, limpieza y transformación

En esta fase se identifican las características adecuadas que deben participar en la construcción del modelo. El objetivo propuesto en este trabajo fue determinar si existe relación entre los perfiles obtenidos de los tesistas de las carreras de informática de la Universidad de Morón con su participación en la investigación, tanto a nivel de proyectos como propuestas de cátedra. Por tal motivo se trabajó con los perfiles derivados de nuestros trabajos anteriores [9][10], donde los atributos son: carrera, edad, grupo familiar, área de trabajo (laboral), área de trabajo de la tesis. Se incorporaron los atributos: motivo de selección de tesis y, si resultó alguna publicación de su trabajo de tesis (ambos recuperados de la encuesta realizada).

Para el atributo clase, se utilizó la pregunta de la encuesta “De su trabajo de tesis ¿surgió alguna publicación?” y de las mismas se comprobó su veracidad con los registros de los proyectos de investigación.

La tabla 1, presenta los atributos a utilizar en la fase de minería de datos con sus valores asociados, luego de la transformación y limpieza de los datos.

**Tabla 1.** Atributos significativos para la fase de minería de datos.

Atributos	Valores
Área de tesis	Agentes y sistemas inteligentes/ Ingeniería de Software/ Base de Datos y Minería de datos/ Innovación en Sistemas de Software/ Arquitectura, Redes y Sistemas Operativos/ Seguridad Informática/ Tecnología y Educación/ Procesamiento de señales y sistemas en tiempo real.
Carrera	Licenciatura en Sistemas/ Ingeniería en Informática.
Área de trabajo	Análisis funcional y requerimientos/ Bases de datos y minería de datos/ Desarrollo/Infraestructura/ Procesos de negocio/ Seguridad informática/Testing/Varios/ No trabaja.
Edad	Menores de 25 años/ Entre 25 y 30 años/ Mayores a 30 años.
Grupo familiar	Con compromiso/ Sin compromiso.
Motivo de elección del tema	Vinculación con su trabajo/Interés personal/Propuesto por la cátedra/Líneas de investigación/Otros. “Propuesto por la cátedra”: significa que el estudiante puede seleccionar las áreas de trabajo que ofrece la cátedra en la primera reunión de clases. Para la definición de las áreas de trabajo, se consideraron las áreas

	<p>propuestas en el Congreso Argentino de Computación (CACIC), evento científico organizado por la Red de Universidades con carreras en informática (RedUNCI) [11].</p> <p>“<i>Líneas de investigación</i>”: son las líneas de investigación de los proyectos que se encuentran en ejecución cuando comienza el año lectivo; y se les ofrece a los estudiantes su incorporación para desarrollar su trabajo de fin de carrera en el marco del proyecto.</p>
Tiene Publicación	No, Si-con proyecto de investigación, Si-sin proyecto de investigación

En la tabla 2, se presentan los proyectos de informática de los últimos 10 años ejecutados en la UM y sus correspondientes áreas de investigación.

**Tabla 2** proyectos de investigación del área de informática de la UM.

Área de investigación	Nombre del proyecto de investigación
Ingeniería de Software	<p>Ingeniería de software en el desarrollo de aplicaciones sensibles al contexto. Validación de las herramientas desarrolladas para la implementación de Sistemas de Información Automatizados.</p> <p>Aproximación a una metodología para la Implementación de Sistemas de Información Basados en Computadoras.</p> <p>Análisis y fortalecimiento de la etapa de Implementación del Proceso Software.</p>
Agentes y Sistemas Inteligentes	<p>Aprendizaje de Robots por Refuerzo con supervisión y Control a Través de Bioseñales.</p> <p>Interfase Cerebro-máquina aplicada a la domótica y a la Robótica.</p> <p>Control de artefactos y robots centrado en emociones.</p> <p>Estudio exploratorio para el desarrollo de un framework en el contexto de la ciberdefensa y la ciberseguridad.</p> <p>Influencia en el estado biométrico emocional de personas.</p> <p>Explotación de datos EEG y parámetros fisiológicos de usuarios interactuando en contextos virtuales.</p>
Tecnología para la Educación	<p>Ambientes virtuales inmersivos de aprendizaje orientados a la capacitación y al desarrollo de simulaciones en situaciones de riesgo.</p> <p>Ambientes virtuales inmersivos- elementos interactivos.</p>
Minería de Datos	Aplicación de tecnologías inteligentes de explotación de información para el análisis de perfiles de tesis de grado de carreras informáticas de la UM.

### 3.3 Fase de minería de datos

La selección de características o atributos es una técnica necesaria en el pre procesamiento cuando se desea realizar tareas de minería de datos. Los algoritmos de selección de características tienen dos objetivos principales:

Reducir el costo computacional asociado tanto al aprendizaje como al propio modelo de conocimiento generado (eliminando atributos irrelevantes o redundantes).

Aumentar la precisión de dicho modelo (eliminando atributos perjudiciales para el aprendizaje) [12].

En general, las técnicas de selección de atributos pueden clasificarse en dos grandes grupos:

Técnicas de filtrado. Evalúan los atributos de acuerdo con heurísticas basadas en características generales de los datos e independientes del algoritmo de aprendizaje.

Técnicas envolventes. Trabajan juntamente con el algoritmo que va a usarse para el aprendizaje para determinar que atributos son más relevantes.

En esta instancia, se aplicaron distintos evaluadores para determinar con que atributos se trabajarán los nuevos experimentos En la tabla 3 se presentan los resultados obtenidos con los evaluadores utilizados [12].

De los evaluadores propuestos, todos seleccionan como atributo relevante a: motivo de selección de tesis con un 100% de correspondencia. Se observa que Grupo familiar, es el atributo menos relevante. Por este motivo se decidió eliminar el atributo grupo familiar para las siguientes fases.

**Tabla 3.** Resultados al aplicar los evaluadores propuestos.

Evaluadores	Modo de evaluación	Método	Evaluación de atributos
cfsSubSet Eval (filtro)	Cross-validation	Best first	Motivo de selección 100%
ChiSquareAttribute Eval (filtro)	Cross-validation	Ranker :	rank atributo 1 +- 0    6 motivo-selección- 2 +- 0    5 área-tesis 3.1 +- 0.3    4 Trabajaba 3.9 +- 0.3    2 Edad 5.1 +- 0.3    1 Carrera 5.9 +- 0.3    3 Grupo-fam.
ReliefAttributeEval (filtro)	Cross validation	Ranker	rank atributo 1 +- 0    6 motivo-selección 2.3 +- 0.64    5 área-tesis 3.1 +- 0.54    2 Edad 3.6 +- 0.66    4 Trabajaba- 5.4 +- 0.49    1 Carrera 5.6 +- 0.49    3 Grupo-fam.
WrapperSubsetEval weka.classifiers.bayes. BayesNet 5 padres envoltura	-B Cross-validation	Greedy- Stepwise	Motivo de selección de tesis 100%
WrapperSubsetEval weka.classifiers.bayes. BayesNet 5 padres envoltura	-B Cross validation	Best-first	1 Carrera 20 % 2 edad 10 % 3 grupo-familiar 0 % 4 trabajaba-en 10 % 5 área-tesis 0 % 6 motivo-selección 100%

### Experimento 1.

La pregunta de investigación que surgió en una primera instancia fue:

*¿Existe correlación entre los perfiles obtenidos de los tesis de informática, con el atributo planteado como clase: tiene -publicación?*

Se trabajó con el software WEKA [13], del cual se obtuvo un modelo de clasificación con técnicas de clustering. Estas tratan de encontrar grupos dentro de un conjunto de individuos, donde el principal objetivo reside en agrupar con características similares entre sí a los sujetos, y a su vez diferenciadas de los individuos pertenecientes al resto de los grupos. [8] Los clusters fueron observados en el SMS como tareas muy frecuentes en la minería de datos.

Para este trabajo se utilizaron los algoritmos de clustering: K-means o k-medias, basado en particiones y EM o Expectation Maximization basado en probabilidades. K-means necesita como dato de entrada el número de grupos en los que se segmentará la población. A partir de este número k de clusters, el algoritmo coloca primero k puntos aleatorios (centrodes). Luego asigna a cualquiera de esos puntos todas las muestras con las distancias más pequeñas. A continuación, el punto se desplaza a la media de las muestras más cercanas. Esto generará una nueva asignación de muestras, ya que algunas muestras están ahora más cerca de otro centroide. Este proceso se repite de forma iterativa y los grupos se van ajustando hasta que la asignación no cambia más moviendo los puntos. El algoritmo EM, empieza prediciendo, al azar, los parámetros de las distribuciones y los usa para calcular las probabilidades de que cada instancia pertenezca a un cluster y usa esas probabilidades para estimar nuevamente los parámetros de las probabilidades [13].

Se inicializó para ambos algoritmos el número k (cantidad de clusters) igual a 3 y se corrieron estos, utilizando la clase para su evaluación. En la tabla 4 se muestran los resultados obtenidos para ambos algoritmos, donde cluster 0 = “No tiene publicación”, cluster 1 = “Si tiene publicación, (sin proyecto de investigación)” y Cluster 2 = “Si tiene publicación (con proyecto de investigación)”.

**Tabla 4.** Comparación entre los algoritmos K-means y EM.

Algoritmo utilizado	Cantidad de instancias que se obtienen en cada clase			Incorrectas con respecto a la clase
	Cluster 0	Cluster 1	Cluster 2	
	64	13	20	
K-means	43	4	8	43,29%
EM	45	3	12	38,14%

De estos resultados, se interpreta que existe una baja correlación entre los clusters obtenidos y el atributo clase, *tiene-publicación*.

La tabla 5 muestra la distribución de los centroides obtenidos por K-means, donde el número, en la primera fila de cada cluster (59, 23,15), representa la cantidad de instancias que considera para el grupo.

**Tabla 5.** Centroides obtenidos por K-means.

Atributo	Cluster 0 (59 instancias)	Cluster 1 (23 instancias)	Cluster 2 (15 instancias)
Carrera	Ing-I	Lic-S	Lic-S
Edad	25-30	Mayor-30	Mayor-30
Trabaja-en	Varias	An-fun-req	Varias
Área-de-tesis	Ag.Sis-Int	Ing-Sof	Ag-Sis-Int
Motivo-selección	In-per	In-per	L-investigación

En este experimento se evidencia que, al utilizar algoritmos de agrupamiento, no existe correlación entre los perfiles obtenidos por el cluster y, si el tesista logró o no, una publicación científica.

### Experimento 2

¿Cómo reaccionan los algoritmos de clasificación ante el set de datos con los nuevos atributos: “*motivo de selección de tema y si tiene-publicación*”?

Entre los algoritmos de clasificación, se utilizó el J4.8, (algoritmo supervisado), en el cual se tiene una variable dependiente o clase y el objetivo del clasificador es determinar el valor de dicha clase para nuevos casos.

Se trabajó con los algoritmos J4.8, Naïves Bayes, y Bayes Net (con 5 padres y 3 padres). En todos los casos se utilizó cross validation con 10 carpetas. Los resultados al comparar dichos algoritmos con la herramienta WEKA, en modo experimenter, se muestran en la tabla 6, en porcentajes de atributos clasificados correctamente.

**Tabla 6** Porcentaje de atributos clasificados correctamente.

Algoritmo	J4.8	Naïves Bayes	Bayes Net (5 padres)	Bayes Net (3 padres)
Porcentaje de correctos	82,52%	78,07	76,30	76,30

Se observa que, al aplicar algoritmos probabilísticos, como Naïves Bayes o Bayes Net, en ningún caso mejora el rendimiento en la clasificación obtenida por el algoritmo J4.8

A continuación, se presenta el árbol obtenido por el algoritmo J4.8, donde se clasificaron correctamente el 82,48% instancias de un total de 97 instancias.

```

motivo-selección-tesis = In-per: No (50.0/9.0)
motivo-selección-tesis = Prop-cátedra
| Edad = 25-30: No (6.0/1.0)
| Edad = menor-25: Sí-SIN-PID (2.0)
| Edad = mayor-30: No (6.0/1.0)
motivo-selección-tesis = L-investigación: Sí-CON-PID (20.0/2.0)
motivo-selección-tesis = Vinc-trabajo: No (10.0/1.0)
motivo-selección-tesis = Otros: No (3.0/1.0)

```

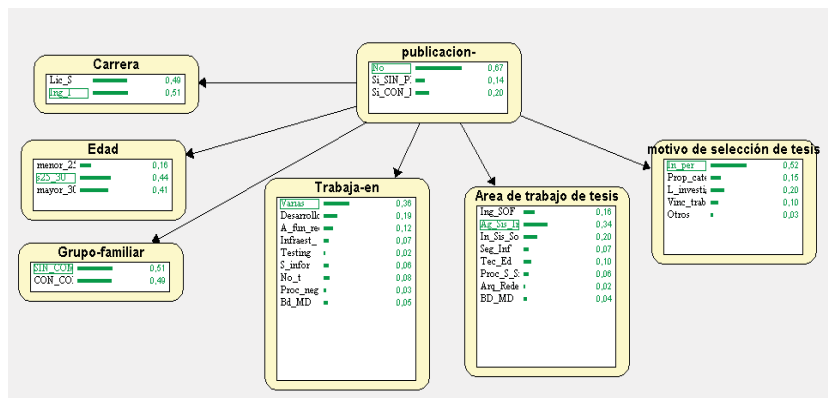
De este experimento, se interpreta que el atributo “*motivo de selección de tesis*” es el que tiene mayor incidencia en la clase.

### Experimento 3.

La pregunta de investigación fue: *¿es posible obtener un modelo predictivo confiable para nuestros datos?*

Uno de los problemas a los que se enfrenta la minería de datos es ocuparse de la incertidumbre. El que se soluciona al trabajar con métodos y técnicas bayesianas, ya que una de sus principales características es el uso explícito de la teoría de probabilidad.

La idea de usar el teorema de Bayes, en cualquier problema de aprendizaje automático, es que se puede estimar la probabilidad a posteriori de cualquier hipótesis consistente con el conjunto de datos. El algoritmo utilizado fue el clasificador Naïve Bayes realizado por Duda & Hart en el año 1973 [14], donde supone que todos los atributos son independientes entre sí, conocido el valor. Se utilizó para tal fin, el software Elvira [14]. Se obtuvo la red que se muestra en la Figura 1, donde la variable clase es: *tiene-publicación* (publicación) y los nodos hijos son las variables independientes: carrera, edad, grupo familiar, área de trabajo y motivo de selección de tesis. La red muestra la distribución conjunta de probabilidades obtenida por el teorema de Bayes.



**Fig. 1.** Red obtenida con el software Elvira aplicando Naïves Bayes. Distribución conjunta de probabilidades.

En la Figura 2, se muestra un nuevo caso en el cual la evidencia es: Carrera = Ingeniería, Edad = mayor a 30, trabaja en = Varios, Área de trabajo de tesis = Agentes y sistemas inteligentes, motivo de elección de tesis = interés personal (todos los nodos en gris) donde el nodo clase publicación tiene un 90 % de probabilidad que el tesista no realice una publicación.

Si a este caso se modifica sólo el motivo de selección de tesis por líneas de investigación, los valores de la clase cambian a “si- publica con proyectos en un 75 %” (ver Figura 3).



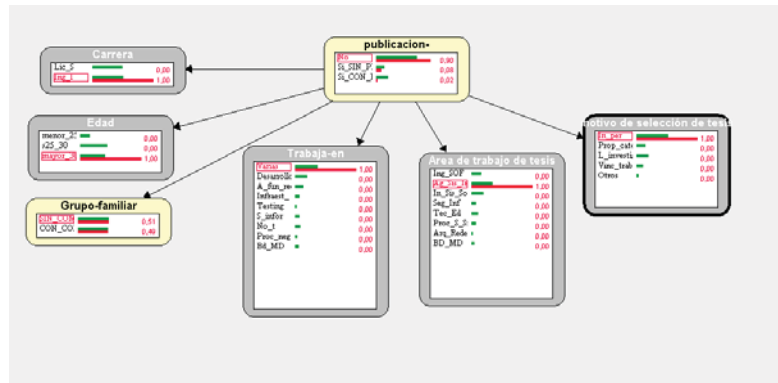


Fig. 2. Evidencia 1: Carrera. Ingeniería, Edad: mayor a 30, trabaja en Varios, Área de trabajo de tesis. Agentes y sistemas inteligentes, motivo de elección de tesis: interés personal.

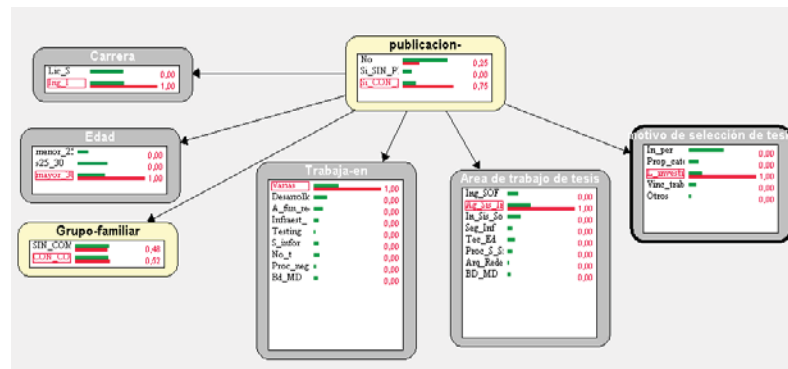


Fig. 3. Evidencia 2: Carrera. Ingeniería, Edad: mayor a 30, trabaja en Varios, Área de trabajo de tesis. Agentes y sistemas inteligentes, motivo de elección de tesis: líneas de investigación.

#### 4 Conclusiones y futuros trabajos

Se han presentado los trabajos relacionados que fueron revisados mediante un método de investigación, estudio de mapeo sistemático (SMS). Este permitió identificar que problemas se resuelven con la minería de datos educacional. Logramos evidenciar que los estudios primarios enfocados en la definición de perfiles de estudiantes, ninguno considera, en sus características, el acercamiento del estudiante a las actividades de investigación.

Los perfiles hallados de los tesisistas, utilizando clusters, no permitieron evidenciar su participación en un proyecto de investigación.

Se descubrió el grado de interés de los tesisistas por determinadas áreas de investigación, mediante el atributo “*motivo de selección de tesis*”. En el caso de interés personal o propuesta de cátedra, existe una tendencia hacia la escritura de

artículos y participación en eventos científicos. Si el motivo de selección es líneas de investigación, los tesistas desarrollan su trabajo de fin de carrera en el marco de un proyecto de investigación, esto permite motivar a los alumnos hacia la escritura de artículos y luego la presentación de estos en eventos científicos.

Los resultados de este trabajo permitieron generar una serie de propuestas a los graduados, como la posibilidad de obtener un subsidio para realizar su doctorado y/o la incorporación, como investigador-graduado UM, en un proyecto de investigación.

Como futuros trabajos se identifican: (a) continuar la experimentación con otros algoritmos (b) analizar las áreas de interés de aquellos alumnos que se han interesado por la escritura de artículos y/o participación en eventos científicos con el objetivo de viabilizar la presentación de proyectos de investigación en dichas áreas.

## Referencias

1. Romero, C., Ventura, S., Pechenizkiy, M. & Baker, R. S. Handbook of educational data mining. CRC Press. Eds.(2010).
2. Luan, J. Data mining and its applications in higher education. New directions for institutional research, (113), 17-36 (2002).
3. Lin, S. H. Data mining for student retention management. Journal of Computing Sciences in Colleges, 27(4), 92-99 (2012).
4. Chacon, F., Spicer, D., & Valbuena, A. Analytics in support of student retention and success. Research Bulletin, 3, 1-9. (2012).
5. Universidad de Morón. Secretaria de Ciencia y Tecnología (SeCyT-UM). Recuperado el 20 de julio de 2019, página web: <https://www.unimoron.edu.ar/area/cienciaytecnologia>
6. Marisa Daniela Panizzi. "Establecimiento del estado del arte sobre la Minería de Datos Educativa en el Nivel Superior: Un Estudio de Mapeo Sistemático". Revista de Investigaciones Científicas del Universidad de Morón. Vol. 3 nro. 1, 2019.
7. Kitchenham, B. and Charters, S. Guidelines for performing Systematic Literature Reviews in Software Engineering. Versión 2.3 EBSE-2007-01 (2007).
8. Hernández Orallo José, Ramírez Quintana María José, Ferri Ramírez César. Introducción a la Minería de Datos. Ed. Pearson Education S.A. Madrid. (2004).
9. Iris Sattolo, Gastón Alvarez, Nicolás Armilla, Oscar Bravo, Matias García, Javier Lafont, Gabriel Mariuz, Lucila Mira, Marisa Panizzi. Hacia la caracterización de perfiles de tesistas de Carreras de Informática de la Universidad de Morón. XIII Congreso Nacional de Tecnología en Educación y Educación en Tecnología (TE&ET 2018). Universidad Nacional de Misiones. Posadas, Misiones. Argentina. 14 y 15 de junio 2018.
10. Iris Sattolo, Gastón Alvarez, Matias Garcia, Javier Lafont, Lucila Mira, Gabriel Mariuz, Nicolás Armilla, Marisa Panizzi. Descubrimiento de las áreas de investigación seleccionadas por los tesistas de las carreras de informática de la UM mediante árboles de decisión. XXIV Congreso Argentino de Ciencias de la computación (CACIC 2018) Tandil Universidad Nacional del Centro de la Pcia. De Bs.As. ISBN 978-950-658-472-6
11. RedUNCI (Red de Universidades con carreras de Informática). Recuperado el 18 de julio de 2019, página web: <http://reduinci.info.unlp.edu.ar/>
12. Elvira soft. Recuperado:16-07-2019 web: <http://www.ia.uned.es/investig/proyectos/elvira/>
13. Weka. University of Waikato. Machine Learning Group. Página web:[www.cs.waikato.ac.nz/ml/Weka/downloading.html](http://www.cs.waikato.ac.nz/ml/Weka/downloading.html). Recuperado: 20 -06- 2019.
14. Castillo, Gutiérrez y Hadi. Sistemas Expertos y Modelos de Redes Probabilísticas. Recuperado:10-07-2019 web: <https://personales.unican.es/gutierjm/papers/BookCGH.pdf>