

Factores que influyen en la titulación de los egresados de un programa académico de pregrado

TOSCANO, Beatriz†*, MARGAIN, Lourdes**, PONCE, Julio*** y PEÑA, Jimmy****

†Universidad Autónoma de Nayarit Unidad Académica de Economía

***Universidad Politécnica de Aguascalientes*

****Universidad Autónoma de Aguascalientes, Centro de Ciencias Básicas*

*****Instituto Tecnológico Superior del Sur del Estado de Yucatán*

Recibido 8 de Julio, 2016; Aceptado 3 de Noviembre, 2016

Resumen

El proceso de análisis de información en las Instituciones de Educación Superior requiere mantener estándares en los indicadores de resultados que manifiestan el nivel de logro o fracaso en la calidad educativa. El indicador de titulación manifiesta la capacidad de una institución de titular a los egresados. Los resultados emitidos de este indicador por las Instituciones, ofrecen informes descriptivos que reflejan un problema en la gestión educativa: bajos índices de titulación con respecto a los índices de egreso. El trabajo propuesto profundiza en el indicador de titulación de los egresados, en un primer momento para identificar los factores que influyen para que un egresado se titule en el nivel superior y en un segundo momento para diseñar estrategias que incrementen este indicador de calidad. Mediante la aplicación de técnicas de minería de datos, se identifican las características de los estudiantes, se determinan los factores que inciden en el proceso de titulación y el patrón de características que tiene un estudiante que se titula. Los resultados muestran algunos patrones que evidentemente pueden contribuir en el incremento del indicador y en la mejora de la gestión del proceso. La determinación de un patrón de características podrá ayudar con análisis predictivos, dados los resultados entre las variables académicas, personales y de contexto.

Calidad Educativa, Titulación, Minería de datos

Citación: TOSCANO, Beatriz, MARGAIN, Lourdes, PONCE, Julio y PEÑA, Jimmy. Factores que influyen en la titulación de los egresados de un programa académico de pregrado. *Revista Investigaciones Sociales*. 2016, 2-6: 73-93

Abstract

The process of analyzing information in Higher Education Institutions required maintaining standards in the performance indicators that show the level of achievement or failure in educational quality. The degree indicator shows the ability of a holder institution graduates. The results reported in this indicator by the institutions, provide descriptive reports that reflect a problem in educational management: low graduation rates with respect to graduation rates. The proposed work explores the indicator degree graduates, initially to identify factors that influence a graduate is titled at the top level and a second time to design strategies that increase the quality indicator. By applying data mining techniques, the characteristics of the students are identified, the factors affecting the titling process and pattern of characteristics that have a student who is entitled are determined. The results show some patterns that clearly can contribute to the increase of the indicator and improved process management. The determination of pattern features can help with predictive analysis, given the results among academic, personal and contextual variables.

Educational Quality

† Investigador contribuyendo como primer autor

* Correspondencia al autor: (email: angelica.delatorre@uan.edu.mx)

Introducción

La gestión de la calidad debe tener como principio fundamental, el enfoque objetivo hacia la toma de decisiones basadas en el análisis de la información disponible de una situación determinada. Los modelos de gestión de la calidad de la familia ISO 9000 y el EFQM (European Foundation for Quality Management, por sus siglas en inglés), tienen en común el Enfoque Basado en Procesos (EBP). El EBP es un principio de gestión de la calidad, que sostiene que "...el resultado que una organización persigue, se alcanza de manera más eficiente cuando las actividades y los recursos se gestionan como parte de un proceso". Articular las actividades, los procedimientos y los recursos entre sí constituyendo un proceso, permite a una organización centrar su atención sobre "áreas de resultados", de ahí que un elemento esencial en el EBP es el uso de indicadores.

Hablando de la gestión de la calidad en la educación, la eficiencia, la eficacia y la congruencia son los criterios que constituyen el marco aceptado para valorar la calidad educativa. Existen diversas opiniones sobre la forma como debe evaluarse o medirse la calidad educativa. En este sentido, la Organización para la Cooperación y el Desarrollo Económico (OCDE) emprendió desde el año 1973 la creación de un sistema internacional de indicadores de los sistemas educativos, proyecto que se cristalizó hasta el año 1988 (Kisilevsky & Roca, 2010). El modelo de indicadores de la OCDE consideraba catorce indicadores, que Scheerens agrupó y describió del siguiente modo: indicadores de contexto; indicadores de entrada; de proceso; de resultados y los indicadores de impacto o de resultados a largo plazo (Scheerens, 2004).

Retomando la clasificación de Scheerens (Scheerens, 2004), los llamados indicadores de resultados de las Instituciones de Educación Superior (IES), son una manifestación del logro o fracaso en la búsqueda de la calidad educativa. Estos indicadores hacen referencia al acceso a la educación, la proporción de titulados, las tasas de abandono en los distintos niveles del sistema educativo, el rendimiento escolar y las competencias obtenidas por los alumnos. Dicho así, el avance en el aprendizaje de los estudiantes es un reflejo del funcionamiento del sistema educativo en su conjunto, mientras que hablar de un bajo índice de titulación, deserción, reprobación y rezago, se traduciría en la falta de calidad en el proceso educativo.

En los últimos años se ha venido observando en México un acelerado crecimiento de la oferta educativa en la educación superior, esta situación ha impactado de manera inversa en los indicadores educativos, particularmente en los llamados indicadores de resultados. Los indicadores son un reflejo de la existencia de un problema en la gestión de la calidad educativa en las IES. Bajo la consideración de que la falta de calidad en una IES es un problema que atañe a todos los actores que participan en el proceso educativo, algunos docentes e investigadores, han venido trabajando investigaciones aplicadas a la generación y/o análisis de indicadores descriptivos de la calidad existente en las IES, algunos de estos trabajos tendientes a proponer estrategias para resolver esta situación.

Conforme lo anteriormente expuesto, este artículo presenta un estudio realizado sobre uno de estos indicadores, específicamente el de Titulación con el objetivo de identificar los factores que influyen en que un egresado se titule en el nivel superior en México. Con la finalidad de que esta información sea de utilidad para orientar el diseño de estrategias que incidan en el incremento de la titulación en pregrado.

Antecedentes

En México el título profesional es un certificado que se otorga por las IES a un estudiante egresado de un programa de estudios de nivel universitario. El título avala al poseedor para el ejercicio de una profesión, el documento representa el reconocimiento de la formación profesional que el titulado posee al término de sus estudios y del cumplimiento de los requisitos establecidos en los planes y programas de estudio correspondientes.

En los artículos 5 y 121 de la Constitución Política de los Estados Unidos Mexicanos, se establece que la ley de cada entidad federativa determinará cuáles son las profesiones que requieren título para su ejercicio, las condiciones para obtenerlo y las autoridades que han de expedirlo. Además, se establece que los títulos expedidos por las autoridades de una entidad federativa serán respetados en las otras (Cámara de Diputados del H. Congreso de la Unión. Secretaría General. Secretaría de Servicios Parlamentarios, 2016). En el caso particular del Estado de Nayarit, el artículo 3 de la Ley para el Ejercicio de las Profesiones y Actividades Técnicas en el Estado de Nayarit (Periódico Oficial del Estado de Nayarit), establece que las profesiones que requieren título para poder ejercer, son aquellas que reconocen respectivamente las leyes y que, de acuerdo con los programas de estudios y grados, han creado las instituciones educativas en la actualidad legalmente reconocidas y por aquellas instituciones educativas que en el futuro fueren establecidas y reconocidas. Por lo anterior, sólo aquellas profesiones reconocidas por alguna Ley (por ejemplo, la Ley General de Salud y la Ley de Aviación Civil), serán las que les requieran a los profesionistas la posesión del título respectivo para poder ejercer.

Por lo tanto, existen profesiones que no están sujetas a una regulación, por lo que podrían ser ejercidas por personas que no posean algún título profesional, aunado a esto, algunos empleadores consideran más importante la experiencia, la habilidad de “aprender a aprender” y la asertividad que el candidato posea en el ejercicio de las actividades relacionadas con la profesión, que el hecho de que este candidato posea un título universitario, lo que contribuye a que exista la percepción para los estudiantes egresados de que en algunos casos no es necesario realizar los trámites de titulación correspondientes al concluir los estudios profesionales, ya que no es determinante en la obtención de un empleo (Durán Ramos, 2002), (Díaz Pérez, 2012).

Por otro lado, debido a la diferenciación en la regulación sobre profesiones, hay una marcada diferencia en los índices de titulación entre carreras como Medicina, Odontología, Ingeniería y Derecho, con niveles mayores al 85%, y algunas carreras de las áreas de Humanidades, Artísticas y Ciencias Sociales, con niveles menores del 20% (Sánchez C., 2009).

Los bajos índices de titulación con respecto a los índices de egreso en el nivel superior han sido una de las preocupaciones de las universidades en México, ya que se ha percibido que la obtención del título profesional es un proceso burocrático, oneroso e innecesario para el ejercicio de la mayoría de las profesiones. Las universidades, tanto públicas como privadas, han establecido diferentes estrategias para incrementar los índices de titulación, creando nuevas modalidades para obtener el título profesional, tales como memorias de las prácticas profesionales o el servicio social, exámenes generales de conocimientos, reconocimiento a la excelencia académica, aprobación de cursos a nivel de posgrado o seminarios de titulación.

Omitir el proceso de titulación al egresar de un programa educativo no permite a las instituciones evaluar integralmente al egresado, ni a éste concluir con su formación académica (Martínez, 2015), (Vargas Pureko & Rivera Michelena, 2006).

En lo que se refiere a la evaluación de la gestión de la calidad educativa, en las últimas décadas se han propuesto estrategias de evaluación de las IES y de los programas de estudio en sus diferentes niveles, así como de los docentes y los estudiantes. Estas evaluaciones están reguladas por diferentes instancias como los Comités Interinstitucionales para la Evaluación de las Instituciones de Educación Superior (CIEES), el Centro Nacional de Evaluación (CENEVAL), el Programa de Mejoramiento del Profesorado de las Instituciones de Educación Superior (PRODEP), y el Consejo para la Acreditación de la Educación Superior (COPAES). En noviembre de 2011, se instaló la Comisión de Coordinación de los Organismos de Evaluación de la Educación Superior (COCOEEES) formada por representantes de la Secretaría de Educación Pública (SEP), de las asociaciones de universidades públicas y privadas, y de los diferentes organismos de evaluación y acreditación de la educación superior en el país. Entre sus objetivos, el COCOEEES pretende fortalecer las prácticas de evaluación en la educación superior conforme a los criterios aceptados nacional e internacionalmente, así como proponer adecuaciones a los marcos de referencia, indicadores y estándares que enfatizan sobre los resultados en el desempeño de estudiantes y del personal académico (Narro Robles, Marthuscelli Quintana, & Barzana García, 2012).

El proceso de acreditación de programas educativos por parte de los consejos de acreditación, consiste en la verificación del cumplimiento de un conjunto de criterios mínimos avalados por COPAES. La COCOEEES determinó que las medias nacionales actuales para los estándares de rendimiento escolar son: índice de reprobación 50%, índice de deserción 20%, eficiencia terminal 40% y titulación 30% (Vargas Leyva & Jiménez Hernández, 2015).

Por su parte, los organismos acreditadores y de evaluación de la educación superior establecen sus propios criterios y niveles de desempeño que las instituciones deben cumplir para otorgar la acreditación de los programas; y a partir de los resultados de la evaluación, los consejos de acreditación emiten un conjunto de observaciones y recomendaciones que serán la base del plan de mejora del programa educativo. Por ejemplo, con respecto al índice de titulación, el Consejo de Acreditación de la Enseñanza de la Ingeniería (CACEI) requiere el cumplimiento de al menos el 60%, y puntualiza que "las opciones de titulación deben ser lo suficientemente variadas y eficientes para lograr que se tittle el mayor número posible de egresados. Si el porcentaje promedio de titulados con respecto a los egresados es bajo (menor del 30%), debe haber programas específicos destinados a incrementar la titulación" (Consejo de Acreditación de la Enseñanza de la Ingeniería, A.C.). Para el Consejo de Acreditación en Ciencias Administrativas, Contables y Afines (CACECA) el índice de titulación por cohorte debe ser mayor al 90%, y en caso contrario, deben existir programas o estrategias para incrementar el porcentaje de titulación de los pasantes mediante diversas opciones de titulación (Consejo de Acreditación en Ciencias Administrativas, Contables y Afines).

El Consejo Nacional de Acreditación en Informática y Computación (CONAIC) solicita la presentación de las estadísticas respectivas de las últimas tres generaciones de estudiantes del programa, pero no establece un requerimiento mínimo para el índice de titulación. Sin embargo, en los criterios correspondientes se establece que la institución evaluada debe contar con una reglamentación disponible para los estudiantes donde se especifiquen las modalidades de titulación y los procedimientos respectivos para cada una, y que se garantice la calidad de los trabajos a través de la evaluación de grupos colegiados internos y/o externos a la institución (Consejo Nacional de Acreditación en Informática y Computación, A.C., 2013).

Los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES) recomiendan realizar un análisis de la pertinencia de las opciones de titulación que tome en cuenta los requisitos, costos y el tiempo promedio del proceso administrativo de cada opción. Además, recomiendan realizar un análisis de la eficiencia en la titulación que considere la proporción por cohorte de estudiantes titulados con respecto de los que egresaron y de los que ingresaron, junto con la proporción de titulados por opción de titulación (Comités Interinstitucionales para la Evaluación de la Educación Superior, A.C., 2015).

Problematización

Atendiendo lo anteriormente expuesto, se puede decir entonces que el indicador de titulación representa la capacidad de una institución educativa de titular a sus egresados.

Sin embargo, que un estudiante se titule o no, no es responsabilidad exclusiva de la institución educativa, la titulación del egresado de las IES puede estar relacionada no solo con aspectos relativos a los trámites administrativos onerosos que un egresado tiene que realizar para ello, sino también con factores extrínsecos a la institución, pero intrínsecos al estudiante y su entorno, tales como, su contexto social, familiar y económico. Otros factores que inmiscuyen a la institución, pero no solo a ella, sino a todos los componentes del proceso educativo, son aquellos relacionados con la trayectoria académica del egresado durante su estancia en la institución. La titulación puede incluso ser un reflejo de la dimensión oferta-demanda que se está viviendo entre el mercado de recursos humanos y el mercado laboral, esto puede generar en el egresado una incertidumbre de conseguir a través de un título la seguridad de un empleo en donde pueda aplicar las competencias profesionales adquiridas.

De acuerdo con la Asociación Nacional de Universidades e Instituciones de Educación Superior (ANUIES), del total de jóvenes que ingresan a la educación superior en México, solo el 50% logra titularse, y este porcentaje varía conforme sea el régimen de sostenimiento de la institución educativa, cinco de cada diez alumnos de instituciones privadas se titulan, mientras que en las universidades públicas estatales este índice de titulación es de 52%, es decir, que en promedio por cada cien alumnos que ingresaron en una misma generación, la mitad no obtuvo el título. Las cifras del estudio que hizo la ANUIES muestran que las universidades federales son las que tienen un rezago mayor en la titulación de sus egresados, pues sólo 54% de los pasantes de una misma generación obtiene ese documento que lo hace acreedor a una cédula profesional.

Mientras que en las universidades tecnológicas el porcentaje de titulados es de 58% en promedio, cifra que es superior a las universidades federales como la Universidad Nacional Autónoma de México, el Instituto Politécnico Nacional o la Universidad Autónoma Metropolitana (Asociación Nacional de Universidades e Instituciones de Educación Superior, 2013).

Debido a ese bajo nivel de titulados, la SEP ha señalado que es necesario que las instituciones públicas busquen opciones para que los alumnos se titulen y no egresen sólo como pasantes, ya que eso influye en la obtención de un mejor empleo, dado que en alguno de los casos los empleadores exigen el título profesional y que no lo tengan limita su inserción al mercado laboral. Esta preocupación no es exclusiva de la administración pública, las IES han asumido esta responsabilidad, a través de replantear estrategias que permitan incrementar este indicador y resolver las causas que pueden estar minándolo.

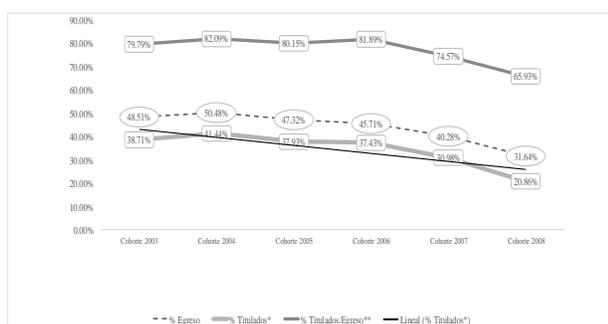
No ajena de esta responsabilidad, la Universidad Autónoma de Nayarit (UAN), la institución pública del estado de Nayarit que absorbe la mayor parte de la matrícula en educación media superior y superior, en el mes de diciembre del año 2002 como parte de su proceso de reforma universitaria, aprobó el documento “Bases Generales para la Titulación de Licenciatura por Áreas del Conocimiento”, dicho acuerdo tiene por objeto reconocer y acreditar los estudios de licenciatura realizados en los centros educativos de la UAN, para ejercer las profesiones contempladas dentro de los planes y programas de estudio. Esta normativa tuvo en sus inicios también la finalidad de otorgar un abanico de posibilidades a egresados que guardaban un rezago en la obtención de su título, para que pudieran titularse e incrementar así el indicador de titulación de la institución.

La UAN reconoce desde ese año como opciones de titulación: la Tesis; Generación o aplicación del conocimiento; Examen de Conocimientos (Pueden optar por el Examen General de Egreso de Licenciatura (EGEL) que aplica el CENEVAL); Promedio general; Curso de titulación; Curso de posgrado y Memoria de experiencia profesional (Universidad Autónoma de Nayarit, 2002). Cada Programa Académico de los que se ofertan en la UAN, decidió por acuerdo del órgano de Consejo de Unidad Académica, cuáles de estas opciones considerar atendiendo las condiciones propias del programa académico y su plan de estudios.

Sin embargo, aunque la diversidad de opciones de titulación ha impactado de manera significativa en el indicador de titulación por año, de la UAN, el índice de titulación por cohorte generacional esta por abajo del 50% y con una tendencia lineal a la baja desde el año 2003, tal como se puede observar en la figura 1. Cabe señalar que la flexibilidad del modelo académico de la UAN permite que los estudiantes avancen en ritmos diferenciados lo cual impacta en que el estudiante acorte o alargue su permanencia en el programa académico de origen, por ende, esto redundo tanto en el indicador de Eficiencia Terminal (ET) por cohorte generacional como en el porcentaje de titulación.

Atendiendo este panorama, esta investigación parte del hecho de que la diversidad en las modalidades por las que un egresado puede optar para titularse no han resuelto en mucho, el bajo índice de titulación que guarda la institución.

Los investigadores aquí reunidos consideran que es importante antes de trazar cualquier línea de acción, elaborar un estudio que permita identificar los causales del problema, a fin de que las soluciones sugeridas vayan encaminadas a la mejora de la calidad educativa de la institución, e impactar en segundo término, pero no menos importante, en el cumplimiento de los objetivos prioritarios establecidos en las políticas públicas del país.



* Este porcentaje se obtiene sobre el total de los estudiantes que ingresaron en cada cohorte generacional.

** Este porcentaje se obtiene de la relación que existe con el número de los que egresaron de cada cohorte generacional.

Figura 1 Evolución de la eficiencia terminal y titulación de los Programas Académicos de Licenciatura de la UAN, cohorte 2003 a la cohorte 2008. Elaboración propia. Fuente datos primarios: Unidad de Desarrollo Institucional de la UAN. Actualizado en diciembre, 2013

Este trabajo forma parte del proyecto denominado “La Aplicación de Técnicas de Minería de Datos para Determinar los Factores que Influyen en los Indicadores de Eficiencia Terminal”, el objetivo central del proyecto es contribuir a la mejora de la ET en programas educativos del área de las Tecnologías de Información y Comunicación TIC que se ofertan en las IES en México. El proyecto tomó como primer caso de estudio el Programa Académico de Informática (PAI) de la UAN, del que ya se realizó un primer diagnóstico sobre ET (Toscano de la Torre, Ponce Gallegos, Díaz Mendoza, & Carrillo Ortíz, 2015).

Hablar de mejorar la calidad de la educación tiene que ver con alcanzar una mayor congruencia entre lo planeado, la eficiencia para llevar a cabo los procesos para hacer lo planeado y la eficacia para lograrlo. La ET y el índice de titulación son indicadores que representan el número de estudiantes que logran terminar su trayectoria formativa para poder ejercer una profesión a través de la cual puedan transformar el entorno productivo, de tal forma que es muy común que estos indicadores sean considerados como un reflejo de la calidad de una institución educativa. Partiendo de esta premisa, este trabajo es el segundo estudio producto del proyecto arriba mencionado, plantea como objetivo identificar las características de los estudiantes del PAI e indagar los factores que inciden en su titulación, tomando en cuenta la dimensión personal (sexo, edad, estado civil, lugar de nacimiento); académica (promedio de calificaciones en la licenciatura, trayectoria académica, situación de egreso, resultado en el Examen General de Egreso) y de contexto (promedio de calificaciones en bachillerato, bachillerato según el tipo de sostenimiento, nivel de rezago económico del lugar en donde hizo su bachillerato, otros).

Titulación. El caso del Programa Académico de Licenciatura en Informática de la Universidad Autónoma de Nayarit

El PAI de la UAN inició operaciones en el año de 1991, a partir de esa fecha ha sufrido dos rediseños curriculares, uno de ellos en el año 2003 y otro en el año 2012 (para el 2016 se encuentran funcionando los planes curriculares 2003 y 2012). Los estudiantes inscritos en el plan de estudios 2003, podían elegir entre dos salidas terminales, Licenciado en Informática con énfasis en: administración de proyectos informáticos o énfasis en administración de la información en las organizaciones.

Para su egreso, atendiendo la flexibilidad curricular que promueve el modelo académico de la UAN, los estudiantes inscritos en el PAI pueden concluir su total de créditos en un periodo de 4.5 a 9 años, y se consideran egresados una vez que concluyen el total de los créditos dentro de los cuales se considera también la prestación del servicio social y prácticas profesionales, además de cumplir como requisito de egreso con la presentación del EXACRI⁴ y el CENEVAL⁵. La legislación universitaria no contempla un tiempo límite para la titulación de los egresados de los programas de pregrado.

En lo que se refiere al plan de estudios 2003 del PAI, este plan curricular articula su estrategia para la titulación al Acuerdo que establece las Bases Generales para la Titulación de Licenciatura por Áreas del Conocimiento (Universidad Autónoma de Nayarit, 2002), aprobado en diciembre de 2002, que establece las opciones o mecanismos de titulación para los programas académicos de licenciatura que oferta la UAN. De tal forma que los egresados de este plan de estudios pueden considerar como mecanismos de titulación, los siguientes: tesis; generación o aplicación del conocimiento; examen de conocimientos; promedio general; curso de titulación; curso de posgrado; y memoria de experiencia profesional.

Concerniente a la estructura formativa del plan de estudios 2003 del PAI, el área de formación disciplinar y profesionalizante se conforma de 8 líneas de formación que son: Entorno Social; Matemáticas; Programación e Ingeniería de Software; Arquitectura de Computadoras; Redes; Software de Base; Tratamiento de Información y por último la de Interacción Hombre-Máquina.

Este plan de estudios dejó de recibir estudiantes de nuevo ingreso a partir del año 2011, que fue cuando ingresó su última cohorte generacional y entró en vigor el plan curricular 2012 del cual aún no egresa la primera generación.

Para los fines del proyecto marco que dio pie a este estudio, se manejó como tiempo estándar de egreso, 6 años, de tal forma que para el cálculo del índice de ET solo se consideraron aquellos estudiantes que egresaron dentro de este límite de tiempo. Aquellos que no concluyeron en ese tiempo se cuantificaron como egreso extemporáneo. Atendiendo esta acotación, la población elegida para el presente estudio, corresponde a los estudiantes que ingresaron al PAI con el plan curricular 2003. Ahora bien, aunque la última cohorte de ingreso al PAI con este plan curricular en vigencia fue la que ingresó en el año 2011, solo se tomaron para el estudio las cohortes generacionales que ingresaron en los años 2003 al 2009 dado que cuando se inició este trabajo de investigación (julio, 2015) estas cohortes ya contaban con egresados que cumplían con el requisito del tiempo estándar de egreso, de 6 años, para poder ser considerados como parte de la población de estudio; mientras que para la cohorte 2010 estaban apenas por egresar los primeros estudiantes, y para el caso de la 2011 aún no se tenía ningún egreso.

Como un primer paso se consideró pertinente trabajar un análisis estadístico exploratorio para aportar datos que permitan caracterizar a los 245 estudiantes egresados que forman la población de estudio, estos representan el 35.71% de un universo de 686 estudiantes que ingresaron en estas 7 cohortes generacionales.

⁴ Examen de acreditación de Lengua Extranjera que promueve la UAN con el propósito de acreditar el dominio del idioma inglés como un requisito de egreso y de titulación para los estudiantes de licenciatura de los programas académicos que oferta la institución.

⁵ Es el examen general de egreso (EGEL) que aplica el Centro Nacional de Evaluación para la Educación Superior. Los estudiantes a egresar del PAI aplican el denominado EGEL-Info.

De los estudiantes que egresaron, 178 (72.65%) concluyó sus créditos en un tiempo⁶ menor o igual a los seis años ($t \leq 6$ años) y 67 (27.35%) en un tiempo mayor a los seis años, pero menor a los 9 años ($6 < t < 9$ años).

Los egresados corresponden en diferente proporción a las siete cohortes generacionales que se analizan en el trabajo, de ahí que la ET varía por cohorte. El comportamiento de este indicador en el PAI, conforme el análisis realizado, tal como se observa en la figura 2, tiene una tendencia lineal a la baja en las últimas cohortes que son parte de este estudio. Cabe señalar, que 154 estudiantes (22.45% de los 686), están pendientes aún por concluir el total de sus créditos, estos pertenecen a las cohortes 2007 a la 2009; dado que aún no exceden el número de años que establece la UAN como plazo para egresar, sin embargo, como ya excedieron el tiempo estándar de los 6 años que se consideró para obtener el indicador de ET, el resultado que aquí se muestra no se vería afectado aún cuando estos 154 estudiantes egresaran.

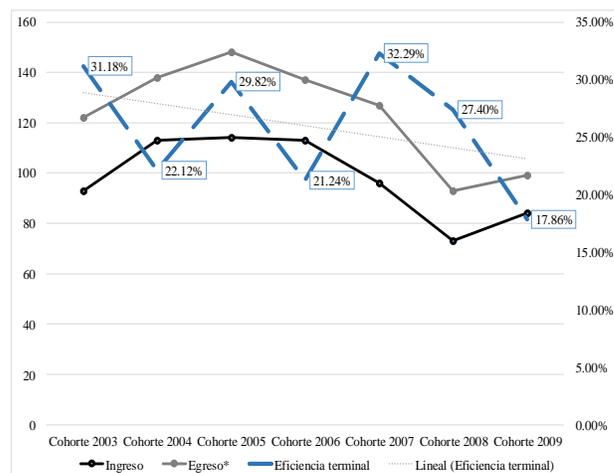


Figura 2 Eficiencia Terminal⁷ en el Programa Académico de Informática, Cohortes Generacionales 2003, 2004, 2005, 2006, 2007, 2008 y 2009

Conforme el análisis estadístico aplicado a la información recopilada, se observa que, del total de los egresados, 137 (55.92%) son hombres y 108 (44.08%) mujeres. El mayor porcentaje de los hombres que egresan, 94 (68.35%) concluyen su carrera de 4.5 a 6 años; en el caso de las mujeres que egresan el porcentaje de las que concluyen en ese mismo tiempo es mayor 84 (77.27%).

El proceso de selección para el ingreso a licenciatura en la UAN, vigente para las cohortes generacionales que se analizan, fue a través de la aplicación y aprobación de un examen único denominado EXHCOBA que evaluaba habilidades y conocimientos de los aspirantes.

⁶ El tiempo (t) de egreso se cuenta a partir de que el estudiante ingresa hasta que concluye el total de créditos del plan de estudios y cumple los requisitos de egreso.

⁷ Eficiencia terminal es la relación del número de estudiantes egresados entre el número de estudiantes que ingresaron de la misma cohorte generacional.

El PAI conforme los lineamientos institucionales maneja dos tiempos para el egreso (4.5 años, terminación normal y el extraordinario en 9 años), considerando los lineamientos de la SEP, se consideró para este estudio establecer un límite de egreso de 6 años, a los que cumplieron con esta característica se les contabilizó como ET y los estudiantes que egresaron después de ese periodo, se consideraron como egresos extemporáneos.

Pero, además, la Federación de Estudiantes de la Universidad (FEUAN), promovió la implementación de un programa adicional que permite ingresar algunos estudiantes que no lograron aprobar el examen. De tal forma que de la población de egresados que se analiza, 154 estudiantes (62.86%) ingreso por el EXHCOBA, mientras que 91 estudiantes (37.14%) lo hizo a través del programa adicional que promueve la FEUAN.

La edad promedio de ingreso de los estudiantes egresados de estas cohortes, fue de 19 años, mientras que su edad promedio de egreso fue de 24 años con 9 meses.

Con respecto al bachillerato de procedencia (clasificado por el tipo de sostenimiento), 81 egresados (33.06%) procedía de los bachilleratos de la UAN; 68 egresados (27.76%) de bachilleratos federales; 70 egresados (28.57%) de estatales; 24 egresados (9.80%) de bachilleratos particulares y 2 egresados ((0.81%) no lo indicaron. Sobre el grado de marginación⁸ de la localidad municipal en donde realizaron el bachillerato, 169 (68.98%) egresados lo hicieron en localidades con grado de marginación muy bajo, 41 (16.73%) egresados en localidades con un grado medio, 3 (1.22%) de egresados no se conoce en dónde lo hicieron, 30 (12.24%) egresados en localidades de un grado de marginación bajo, 1 egresado (0.41%) en localidad de un grado alto y 1 egresado (0.41%) en una localidad de un grado de marginación muy alto.

En lo referente al rendimiento académico de los egresados que se analizan, su promedio obtenido en educación media superior, el mínimo fue de 66 y el máximo de 100, con una media de 83.877; mientras que el promedio obtenido en la licenciatura el promedio mínimo fue de 68.35 y el máximo de 98.04, con una media de 82.283.

Atendiendo las líneas de formación del plan curricular al que pertenecen los estudiantes que forman parte de la población objeto de estudio, el comportamiento de el promedio de calificación obtenido por grupo de materias que conforman cada línea, fue: una media de calificación de 83.5 en Entorno Social; 77.30 en Matemáticas; 80.68 en Programación e Ingeniería de Software; 80.43 en Arquitectura de Computadoras; 83.37 en Redes; 81.93 en Software de Base; 83.63 en Tratamiento de Información y por último 77.97 en la de Interacción Hombre-Máquina.

Respecto a la aplicación del EGEL-CENEVAL aunque conforme la normatividad de la UAN éste es considerado como un requisito de egreso, alrededor de 27 egresados (11.02%) de estas cohortes generacionales no cumplieron con este requisito pero dado que el procedimiento administrativo para su aplicación no estaba bien establecido aun, se les hizo válido el egreso aunque no cumplieran con esta exigencia. De los que aplicaron el EGEL, 86 egresados (39.45%) no obtuvieron testimonio de desempeño.

Sobre la titulación, punto medular en este estudio, el 70.61% (173 personas) de los estudiantes egresados de las cohortes generacionales 2003 a la 2009, al momento de la realización del estudio, estaban tituladas.

⁸ Se consideró la clasificación del año 2015 del Consejo Nacional de Población (CONAPO).

De los egresados titulados, 133 (76.88%) lo hizo en el transcurso del año contando a partir de la fecha de su egreso; 26 (15.03%) realizó este trámite después de un año de su egreso, pero dentro de los dos años; el 8.09% restante lo hizo posterior a los dos años, pero dentro de los cinco años contando a partir de la fecha de su egreso. Cabe señalar que el cálculo del porcentaje de titulación arriba mencionado, es la relación de los titulados con los estudiantes que lograron su egreso, pero al establecer la relación de los titulados contra los estudiantes que ingresaron en las cohortes que se estudian, el porcentaje de titulación quedaría en 25.22%. En la figura 3 se puede observar el comportamiento de ingreso, egreso y titulación en las cohortes que se analizan.

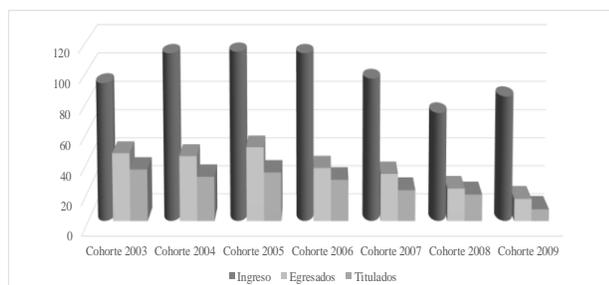


Figura 3 Evolución del ingreso, egreso y titulación del PAI de la UAN, cohorte 2003 a la cohorte 2009. Elaboración propia. Información actualizada a julio de 2015

En la figura 3 se observa una tendencia a disminuir del índice de titulación, cohorte a cohorte, motivo que dio pauta a esta investigación, ya que se consideró importante encontrar evidencias de la influencia de un conjunto de variables de diversa índole sobre el porcentaje de titulación de los egresados del PAI de la UAN. Convirtiéndose en el objetivo general de este trabajo de investigación.

Con base al objetivo general planteado, se determinaron como objetivos específicos, los siguientes:

- Caracterizar a través de un análisis estadístico exploratorio a la población de estudiantes que se inscriben al programa académico. Siendo una de estas características la situación de titulación individual y por cohortes generacionales.
- Determinar el patrón de características que tiene un estudiante que se titula.
- Establecer el grado de incidencia en que cada una de las variables independientes influyen y explican la variable dependiente Índice de Titulación.
- Identificar cuál es la variable independiente que guarda una correlación directa positiva mayor con la variable dependiente "Índice de Titulación".
- Se plantearon tres hipótesis para esta investigación, las cuales fueron:
 - Ho₁: La dimensión personal guarda una relación directa positiva con la titulación de un egresado mayor que la dimensión académica y/o de contexto.
 - Ho₂: La dimensión académica guarda una relación directa positiva con la titulación de un egresado mayor que la dimensión personal y/o la dimensión de contexto.
 - Ho₃: Las características previas del estudiante a su ingreso en el PAI, como su promedio obtenido en la educación media superior, el tipo de bachillerato del que procede, localidad en que hizo el bachiller y todas las que pertenecen a la dimensión de contexto del egresado, guardan una relación directa positiva con la titulación de un egresado mayor que la dimensión personal y/o la dimensión académica durante su estadía en el PAI.

Estas hipótesis nulas buscarán confirmarse o rechazarse a través del uso de técnicas de minerías de datos haciendo uso del software computacional WEKA.

Del método y los datos

De acuerdo a la metodología de Hernández, Fernández y Baptista (2003) hay estudios de tipo exploratorio, descriptivo, correlacional y explicativo.

Este estudio es de tipo descriptivo con enfoque cuantitativo. Para realizarlo, se analizó el contexto previo al ingreso del estudiante al pregrado; su historial académico durante su trayectoria en el programa académico; y los datos personales del estudiante. Además, se revisó la información sobre el plan de estudios y los reglamentos que rigen el currículo al que pertenece la población que es el objeto de estudio. Tomando como base el objetivo del estudio se trabajó con cohortes reales⁹ y solo con los 245 estudiantes que egresaron en el periodo permitido y, por tanto, son candidatos a titulación.

Esta investigación tiene como objetivo mejorar el indicador de titulación en el Programa Académico de Informática de la Universidad Autónoma de Nayarit, no solo conociendo el estado que guarda este indicador de calidad en el PAI sino, además, identificando a través de técnicas de minerías de datos los factores que ocasionan el bajo porcentaje y generando modelos para determinar el patrón de factores que influyen en que el estudiante se titule. Esto a fin de implementar estrategias para mejorar esa trayectoria académica.

Las técnicas de minería de datos permiten recopilar datos en grandes volúmenes de información, analizarlos, procesarlos, compararlos y tomar decisiones que permitan sacar beneficio al máximo de una situación, haciendo uso del análisis matemático. Además de ello estas técnicas, permiten deducir patrones y tendencias que existen en los datos.

En el ámbito educativo la Minería de Datos ha sido de gran utilidad para la predicción de fenómenos que inciden en el rendimiento escolar, como la deserción, reprobación, absorción, identificación del estilo de aprendizaje de los estudiantes o lograr la ET en cualquier nivel de educación.

Para este propósito se trabajó con la herramienta Weka. Weka es una conocida suite de software para el aprendizaje y la máquina que soporta varias tareas de minería de datos típicos, especialmente los datos del proceso previo, el agrupamiento, clasificación, regresión, visualización y selección de características, además de que tiene como ventaja que proporciona interfaces para la comunicación con el usuario y el software se encuentra disponible libremente bajo la licencia pública general de GNU, es muy portable porque está completamente implementado en Java y puede correr en casi cualquier plataforma y además contiene una extensa colección de técnicas para pre-procesamiento de datos y modelado.

A fin de dar cuenta del objetivo planteado para esta investigación, en la figura 4 se observa la propuesta metodológica utilizada para la obtención de conocimiento mediante la Minería de Datos.

⁹ Se define como Cohorte Real, al estudio de poblaciones homogéneas o pertenecientes a una misma generación.

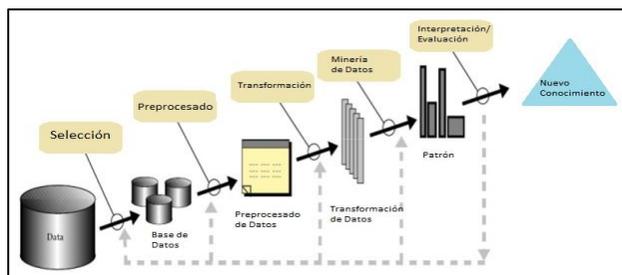


Figura 4 Propuesta metodológica para el proceso de obtención de conocimiento mediante la Minería de Datos

Del uso de la minería de datos

Data. El fichero de datos objeto de análisis contiene una población de 245 estudiantes egresados que corresponden a las cohortes generacionales de la 2003 a la 2009 del Plan de Estudios 2003 del Programa Académico de Informática de la Universidad Autónoma de Nayarit. Los datos fueron obtenidos del Área de Control Escolar de la Unidad Académica de Economía (dependencia que oferta el programa académico), por medio del Sistema de Administración Escolar (SADCE). Esta base de datos inicial se complementó con datos proporcionados por la Dirección de Control Escolar de la Universidad, las Bases de Datos de Nuevo Ingreso proporcionadas por el Centro de Documentación y Digitalización de Información (CEDDI) y la base de datos de resultados de la aplicación EGEL-Info que proporcionó la Secretaría de Docencia de la misma universidad. La información se trabajó en Excel y con el software WEKA para la Minería de Datos. El fichero de datos inicial contiene un total de 78 atributos que describen a las 245 instancias, que corresponde a datos generales, civiles, del bachillerato, de la trayectoria académica del alumno y de los resultados obtenidos en el EGEL-Info.

Selección. En esta fase se recuperaron de las diferentes bases de datos los datos más relevantes para llevar a cabo el análisis objetivo de esta investigación. Se eligieron aquellos datos que no fueran redundantes entre sí.

Preprocesado. Para introducir los datos en WEKA, fue necesario primero aplicar el proceso ETL (Extract, Transform and Load, según las siglas en inglés) a fin de seleccionar los datos útiles y homogeneizarlos. Se eliminaron los datos ruidosos o irrelevantes, caso específico de todas las instancias que no tenían el status como egresados fueron eliminadas a fin de hacer el estudio más específico. En el caso de las instancias con atributos incompletos, tal fue el caso de 19 instancias que no tenían el atributo de la edad de ingreso del estudiante, este atributo se completó con la edad promedio del total de la población.

Transformación de los datos. Una vez determinados los datos relevantes, se procedió a compilar las bases de datos obtenidas de las diferentes fuentes en una base de datos única, esta se trabajó primeramente en una hoja de cálculo Excel. La base de datos en Excel se transformó a un archivo ASCII que usa como delimitador la coma, en este caso se creó un fichero CSV que se pudiera trabajar con el software WEKA.

Weka permite realizar diversas operaciones antes de aplicar los algoritmos de minería de datos. Se pueden realizar diversos preprocesados entre los que destacan los filtros de atributos, con los que se pueden realizar transformaciones de los datos de todo tipo. En este caso aplicando el filtro *Remove*, se filtraron aquellos atributos que no son relevantes para este estudio, tal fue el caso de la matrícula ya que este atributo no era importante para el análisis de correlación entre variables.

El atributo de nombre dado que no se considera un factor que pueda incidir en la obtención del título profesional, de no hacerlo puede darse el caso de obtener como resultado que “Juan López” o “María González” tienen una alta probabilidad de no titularse dado que estos nombres son frecuentes entre los mexicanos. Finalmente quedaron un total de 29 atributos. Para el análisis de correlación los atributos fueron clasificados en dimensiones, denominadas dimensión personal; dimensión académica y dimensión de contexto. Cada atributo identifica un factor que puede tomar una variable determinada que caracteriza a un estudiante (véase la tabla 1).

Minería de Datos. Posterior al proceso de transformación, donde los datos redujeron de dimensionalidad y derivó en la agrupación por características, la etapa de modelado requiere esculpir el proceso. Para efectos de esta investigación, se implementaron técnicas de clasificación para conseguir patrones de comportamiento, que finalmente darán razón a la evaluación y entendimiento de la información procesada; precisión en las predicciones; facilidad de interpretación de los datos proporcionados por el modelo.

La clasificación permite encontrar las características que identifican a un grupo para ser clasificado dentro de cierta clase, esto permite comprender el comportamiento del sistema que genera los datos y en determinado momento predecir la clase a la que pertenecerá una nueva instancia. Mientras que, para evaluar los resultados de un clasificador, la evaluación se basa en el porcentaje de instancias clasificadas incorrectamente en la matriz de confusión. Ésta matriz ofrece información del tipo de error de clasificación de los algoritmos.

Las columnas de esta matriz indican las categorías clasificadas por el clasificador y las filas las categorías reales de los datos, por lo que los elementos en la diagonal principal se corresponden con las clasificaciones sin fallo y el resto de elementos son los errores que el algoritmo ha cometido.

Atributo	Descripción	Variable
<i>Dimensión personal</i>		
Sexo	Genero del estudiante. Posibles salidas: femenino o masculino	Nominal
Estado_Civil	Situación Civil del Estudiante: Casado, Divorciado, Soltero, Viudo, Unión Libre, Sin respuesta	Nominal
Edad	Edad del estudiante a la fecha de su ingreso	Numeric
Municipio_Nac	Localidad de nacimiento del estudiante	Nominal
Estado_Nac	Entidad federativa de nacimiento del estudiante	Nominal
<i>Dimensión de contexto</i>		
Bachillerato_Sost	Tipos de bachillerato de acuerdo al sostenimiento	Nominal
Municipio_Bachiller	Localidad donde se ubica la institución educativa en donde el estudiante realizó sus estudios de nivel medio superior	Nominal
Est_Bach	Entidad federativa donde se ubica la institución educativa en donde el estudiante realizó sus estudios de nivel medio superior	Nominal
Grado_Marg_Bach	Se refiere al nivel de marginación de la localidad en donde el estudiante realizó sus estudios de nivel medio superior. Los valores pueden ser: Bajo, Medio, Muy Alto, Muy Bajo, Alto, Se desconoce	Nominal
Prom_Bach	Promedio de calificaciones obtenido por el estudiante en el nivel medio superior	Numeric
Ayo_ingre	Es la cohorte de la generación en la que ingreso el estudiante. Posibles salidas: 2003, 2004, 2005, 2006, 2007, 2008 y 2009	Nominal
Opc_Ing	Es la opción para la cual ingreso el estudiante. Aceptado (fue aceptado por el examen de ingreso de la universidad), Programa Adicional (No fue aceptado en el examen de ingreso, pero ingreso a través del programa de apoyo que implementa la universidad para tal fin)	Nominal
<i>Dimensión académica</i>		
Var_Eficiencia	Clase o salidas esperadas: Aún sin egresar, No Egresó, Tiempo de Egreso <=6 años, Tiempo de Egreso >6 años. La salida "Aún sin egresar" se refiere a que el estudiante aún está inscrito en el PAI y no ha concluido sus créditos. La salida "No Egresó" se refiere a que el estudiante no concluyó sus créditos, pero ya no está inscrito en el PAI, sea el caso de baja definitiva o deserción.	Nominal
Time_egreso	Valor numérico que indica el número de años en que el estudiante concluyó el total de los créditos que corresponden al programa académico	Numeric
Promedio_Lic	Promedio de calificaciones obtenido por el estudiante al finalizar su trayectoria en el programa académico	Numeric
Por_Ordina	Porcentaje de Unidades de Aprendizaje que el estudiante acreditó durante su trayectoria formativa en el Programa Académico en el periodo ordinario	Numeric
Prom_entor	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Entorno Social (según el Plan de Estudios 2003)	Numeric
Prom_mat	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Matemáticas (según el Plan de Estudios 2003)	Numeric
Prom_arq	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Arquitectura de Computadoras (según el Plan de Estudios 2003)	Numeric
Prom_Redcs	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Redes (según el Plan de Estudios 2003)	Numeric
Prom_SB	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Software de Base (según el Plan de Estudios 2003)	Numeric
Prom_prog	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Programación e Ingeniería de Software (según el Plan de Estudios 2003)	Numeric
Prom_Titula	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Tratamiento de Información (según el Plan de Estudios 2003)	Numeric
Prom_inter	Promedio de calificaciones acumulado obtenido en las unidades de aprendizaje cursadas y acreditadas del Área de conocimiento de Interacción Hombre-Máquina (según el Plan de Estudios 2003)	Numeric
Status_Titu	Clase o salidas esperadas: No titulado, Titulado	Nominal
Time_Titula	Valor numérico que indica el número de años en que el egresado se titula	Numeric
Opcion_Titula	Es la opción por la cual el estudiante se tituló. Clase o salidas esperadas: Aún sin egresar; aún sin titularse; Ceneval; curso de titulación; no egresó; generación y ap; promedio; tesis; curso de posgrado; memoria de experiencia profesional;	Nominal
Status_CENEVAL	Clases o salidas esperadas: No presentó, Presentó	Nominal
Res_CENEVAL	Resultados obtenidos en el EGEL-I aplicado por el CENEVAL. Clase o salidas esperadas: Aprobado; No aprobado; No presentó	Nominal

Tabla 1 Atributos finales utilizados para el proceso de minería de datos

Para la clasificación se eligieron los algoritmos Naive Bayes y el KStar, haciendo uso de la variable dependiente Status_Titula (maneja dos salidas: no titulado y titulado) como el atributo class.

Ambos algoritmos se aplicaron sobre el mismo conjunto sobre el que se construye el modelo predictivo para determinar el error (use training set). La finalidad de usar estos dos, es aprovechar las ventajas que ofrece cada uno de ellos.

El algoritmo Naive Bayes, es un algoritmo de clasificación bayesiano basado en la teoría de la probabilidad, es un algoritmo sencillo y obtiene buenos resultados de clasificación, tiene la ventaja de que, al clasificar forma las clases tomando en cuenta las posibles salidas que tiene el atributo denominado class, en este caso clasificó en las clases: titulados y no titulados. La clasificación a través de este algoritmo permite conocer el patrón de factores que guarda un estudiante que se titulará. Al aplicar el algoritmo 243 instancias quedaron correctamente clasificadas, 2 de las instancias no se clasificaron correctamente en la matriz de confusión, lo que significa un error medio absoluto de 0.8163%. Obsérvese la figura 5 sección a).

La combinación de clasificadores permite tener una exactitud mayor en la clasificación de los datos por lo que se decidió aplicar también el algoritmo de clasificación KStar. El algoritmo KStar es un clasificador basado en las variables, lo que quiere decir que la clasificación de una variable está basada en variables de entrenamiento similares, determinadas por alguna función de similitud. Se diferencia de otros aprendizajes basados en lo mismo en que usa una función de distancia basada en entropía, la principal suposición que se hace es que instancias similares tendrán clasificaciones similares.

Este tipo de algoritmo realiza poco trabajo en la etapa de aprendizaje, en los casos más simples tan solo se almacenan los ejemplos en memoria, transfiriendo el esfuerzo al momento de clasificar una nueva instancia, cuando el sistema debe decidir cuáles de los ejemplos memorizados debe utilizar para hacer la clasificación. La ventaja que se observó es que hace la clasificación de manera más “limpia”, clasificando todas las instancias correctamente, tal como se observa en la figura 5 secciones b).

Una vez aplicados los algoritmos de clasificación, se llevó a cabo la fase de clustering o clusterización. El clustering es el proceso que consiste en la división de los datos en grupos de objetos similares. Representar los datos por una serie de clústers, consigue la simplificación de los mismos, dicho así el clustering es una técnica de aprendizaje automático, en la que el aprendizaje es no supervisado. Este proceso juega un papel importante en las aplicaciones de minería de datos (Garré, Cuadrado, & Sicilia). En este trabajo la clusterización se utilizó como un complemento para definir el patrón de los estudiantes egresados que sí se titulan y encontrar el patrón de las variables de los estudiantes egresados que no lo hacen.

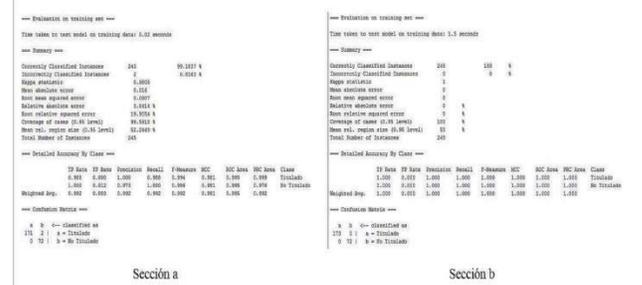


Figura 5 Matrices de Confusión de los algoritmos de clasificación. Sección a: Naive Bayes. Sección b: KStar

Se probaron varios algoritmos de clustering, sin embargo, dado que el SimpleKmeans, es de los más utilizados en aplicaciones científicas se optó por hacer uso de éste. Este algoritmo es clasificado como método de particionado y recolocación, la razón de su nombre es porque representa cada uno de los clústers por la media (o media ponderada) de sus puntos, es decir, por su centroide. Para desarrollar el experimento, se consideró el atributo class denominado Status_titula (con valores posibles: titulado o no titulado), que es el usado para identificar si el estudiante llevó a cabo su titulación. Con la finalidad de tener un porcentaje bajo de instancias mal clasificadas se tomó el número de clúster igual a 2, y seed (10). Los resultados podrán observarse en el apartado correspondiente.

A fin de explicar la variable dependiente, se realizó la etapa de minería de datos denominada selección de atributos. Antes de llevar a cabo la aplicación del algoritmo correspondiente se removieron los atributos que identifican a la opción de titulación elegida por el egresado y el atributo del tiempo de titulación, dado que podrían causar ruido en el resultado de la correlación. Una vez removidos los atributos, para establecer la correlación de variables se utilizaron dos algoritmos CfsSubsetEval y OneR.

CfsSubsetEval es un atributo evaluador que calcula la correlación de la clase con cada atributo, y elimina atributos que tienen una correlación muy alta como atributos redundantes. Para el estudio el evaluador CfsSubsetEval se combinó con el método de búsqueda denominado Bestfirst, este método busca primero el atributo que aparentemente tiene la mejor correlación con la variable dependiente.

El algoritmo se aplicó con el modo de selección cross-validation. Mientras que el algoritmo OneRAttributeEval utiliza el clasificador OneR que es de los más rápidos y sencillos, sus resultados pueden ser muy buenos a comparación de otros algoritmos más complejos. Lleva a cabo la correlación seleccionando el atributo que mejor explica la clase de salida. El algoritmo se aplicó con el método Ranker. Los resultados obtenidos de ambos algoritmos podrán observarse en el apartado de resultados.

Resultados y discusión

En este apartado se presentan los resultados obtenidos a través del uso de la minería de datos, aplicada para determinar el patrón de características que identifican al grupo de estudiantes que se titula y los cruces entre las variables académicas, personales y de contexto con la variable dependiente *Status_titula*.

Características que identifican al egresado que se titula

Una vez aplicados los algoritmos de clasificación y de clustering (figura 6), se pudo observar lo siguiente. A pesar de que los hombres ocupaban el 55.92% de la población de egresados, se observó que, de 137 hombres egresados, solamente se ha titulado el 32%, mientras que de las 108 mujeres que son parte de la población de egresados, el 61% de ellas ya se tituló. En la figura se observa en los resultados del clustering, que de los egresados que forman el clúster de titulados más del 50% son del género femenino.

Respecto a la opción por la cual el egresado hizo su ingreso a la universidad, de los 154 estudiantes egresados que ingresaron por el EXHCOBA, el 55.81% está titulado, mientras que de los 91 egresados que ingresaron por el programa adicional solo el 27.47% está titulado.

En la figura 6 se observa en el clúster de titulados que el 77.42% de los titulados ingresaron por el EXHCOBA que era el examen de admisión vigente para esas cohortes generacionales que la UAN aplicaba.

Cluster	Full Data (254)	C (152)	L (98)	
Age_ingre	Cohorte 2005	Cohorte 2005	Cohorte 2005	Time taken to build model: (full training data) : 6.42 seconds
Age_ing	Aprobado	Aprobado	Procesos Admisionales	
Edad_ing	18.349	18.1514	18.364	--- Model and evaluation on training set ---
Sex	Male	Male	Male	
Estado_Civil	Soltero	Soltero	Soltero	Clustering Iterations
Motivacion_2do	Tejico	Tejico	Tejico	
Estado_2do	Bayreuth	Bayreuth	Bayreuth	1 100 (424)
Matrícula_2do	Autónomo	Autónomo	Estatal	1 90 (354)
Motivacion_Bachiller	Tejico	Tejico	Tejico	
Est_Bach	Bayreuth	Bayreuth	Bayreuth	
Sex_Motiv_Bach	May 3420	May 3420	May 3420	Class attribute: Status_Titulo
Sex_Motiv_2do	62.0787	62.2884	62.2879	Clases no Clusters
Procedim_1to	82.2827	86.3488	78.9118	
Time_egreso	5.4242	5.4242	5.4242	0 1 0 ← assigned to cluster
Var_Eficiencia	Tiempo de Egreso <= 6 años	Tiempo de Egreso <= 6 años	Tiempo de Egreso <= 6 años	142 01 1 Titulado
Edad_2do	24.7970	24.7122	24.8247	10 42 10 Titulado
Sex_2do	160.304	160.304	160.304	
Sex_Motiv	62.0002	62.2147	62.2829	Cluster 0 ← Titulado
Sex_Sex	77.202	76.47	75.7607	Cluster 1 ← No Titulado
Sex_Acc	86.4252	83.2894	78.3004	
Sex_Motiv	80.8703	80.3900	80.1087	Incorrectly classified instances : 61.0 24.7341 %
Sex_2do	81.2429	81.2927	76.3641	
Sex_Sex	80.0784	80.100	75.6074	
Sex_2do	81.6249	81.2874	80.0944	
Sex_Sex	77.0470	77.0388	75.1029	
Time_Titulo	1 480	1 480	254 Titulados	
Quitar_Titulo	Clasific	Clasific	Aut sin Titulacion	
Status_Cursada	Excesivo	Excesivo	Excesivo	
Sex_2do	Aprobado	Aprobado	No Aprobado	

Figura 6 Resultado del Clustering usando el algoritmo SimpleKmeans

Hablando de la edad, el promedio de edad al momento del ingreso de los egresados titulados de estas cohortes, fue de 18 años con 7 meses, mientras que su edad promedio de egreso fue de 23 años con 9 meses. En el caso de los egresados que no se titularon su edad promedio de ingreso fue de 19 años y la edad promedio de egreso fue de 25 años con un mes.

Con respecto al bachillerato de procedencia de los 81 egresados que procedían de los bachilleratos de la UAN, se titularon el 43.20%; de los 68 egresados del sistema de bachilleratos federal, se titularon 52.94%; de los 70 egresados del sistema estatal se tituló el 44.29%; de los 24 egresados de bachilleratos particulares el 37.5%. Es decir, son los egresados que hicieron su bachillerato en el sistema federal los que tienen un más alto índice de titulación según se observa en la figura de la clustering. Sobre el grado de marginación de la localidad municipal en donde realizaron el bachillerato, de los 169 egresados que lo hicieron en localidades con grado de marginación muy bajo el 42.01% se tituló.

De los 41 egresados que hicieron el bachillerato en localidades con un grado medio, se tituló el 60.97%; de los que hicieron el bachillerato en localidades de un grado de marginación bajo se tituló el 43.33%, mientras que de los que lo hicieron en localidades de grado de marginación alto y muy alto, se tituló el 100%.

En lo referente al rendimiento académico que manejan los egresados que se titulan, la media para el promedio obtenido en educación media superior, fue de 87.36 y la media para el promedio obtenido en la licenciatura fue de 87.12. Para el caso de los egresados que no se titularon, la media del promedio de calificación en el bachillerato fue de 81.12, y la media del promedio de calificación en la licenciatura fue de 78.13. En relación al promedio por línea de formación los estudiantes egresados que se titulan tienen una media de calificación promedio más alta que los que no se titulan, salvo el caso de la línea de formación de interacción hombre máquina, en donde manejan una media de calificación más alta los egresados que no se titularon. De tal forma que se puede afirmar que el historial de calificaciones incide en que un estudiante se titule.

Hablando del tiempo de egreso, de los 178 que concluyeron sus créditos en un tiempo menor o igual a los seis años ($t \leq 6$ años) el 58.99% se tituló y de los 67 que egresó en un tiempo mayor a los seis años, pero menor a los 9 años ($6 < t < 9$ años) el 8.96% se tituló. Tal como se observa en la figura 6 de los egresados que forman el clúster de titulados egresaron en un tiempo promedio de 5 años 1 mes, mientras que los que componen el clúster de no titulados la media de tiempo de egreso fue de 6 años un mes.

De los 132 estudiantes egresados que aprobaron el EGEL-Info, el 51.16% se tituló y de los 86 egresados que lo presentaron y no lo aprobaron, solo el 38.37% se tituló.

Sobre la opción de titulación que mas eligen los estudiantes que se titulan, por orden de elección quedo ordenado así: Examen de conocimientos, que en este caso se aplica el EGEL-Info; Promedio General; Tesis; Curso de Titulación y Generación o Aplicación del Conocimiento (por esta opción solo han optado dos personas). En cuanto al tiempo de titulación según se observa en la figura, el 64.79% de los que se titulan lo hacen dentro de un año, contando a partir de su fecha de egreso.

Correlación de las variables

Una vez aplicado los algoritmos para la selección de atributos OneRAttributeEval (figura 7) y BestFirst (figura 8), se observa que todos los atributos manejados muestran una correlación positiva con la variable clase Status_titula, sin embargo se puede observar que los que tienen una mayor correlación, como lo muestra el resultado de la aplicación del algoritmo BestFirst son (se enlistan de mayor a menor grado de correlación): el resultado que el estudiante obtiene en la aplicación de el CENEVAL; el promedio de licenciatura; el grado de marginación de la localidad en la que el estudiante hizo su bachillerato; el estado civil (aunque este atributo se descarta porque casi el 100% de los estudiantes son solteros, lo cual hace que por eso se de este comportamiento) y en el mismo grado de correlación están la opción de ingreso y los promedios de calificaciones en el grupo de materias que conforman las líneas de formación de Entorno Social, Matemáticas, Redes, Programación (el orden del grado de correlación difiere conforme el algoritmo de correlación que se aplicó).

Ranked attributes:	
80.4082	26 Res_CENEVAL
71.8367	6 Estado_Nac
70.6122	2 Edad_Ing
70.6122	9 Est_bach
70.6122	10 Grad_Marg_Bach
70.6122	7 Bachillerato_Sost
70.6122	3 Sexo
70.6122	14 Var_Eficiencia
70.6122	25 Status_CENEVAL
70.6122	1 Opc_ing
70.2041	15 Edad_Eg
69.7959	21 Prom_SB
69.3878	23 Prom_Trata
68.9796	8 Municipio_Bachiller
68.5714	4 Estado_Civil
68.5714	16 For_Ordina
68.5714	13 Time_Egreso
68.1633	12 Promedio_Lic
68.1633	24 Prom_Inter
68.1633	5 Municipio_Nac
67.7551	22 Prom_Prog
67.3469	20 Prom_Redes
66.9388	19 Prom_Arq
65.3061	11 Prom_Bachiller
63.6735	17 Prom_Entor
62.0408	18 Prom_Mat

Selected attributes: 26,6,2,9,10,7,3,14,25,1,15,21,23,8,4,16,13,12,24,5,22,20,19,11,17,18 : 26

Figura 7 Resultados del algoritmo de selección de atributos OneRAttributeEval

Por lo que apoyados en los resultados que arrojaron los algoritmos de minería de datos para la selección de atributos explican la variable dependiente, no hay evidencia suficiente para aceptar la hipótesis nula H_{01} , de tal forma que se acepta la hipótesis alternativa H_{a1} : Las dimensiones académicas y/o de contexto guardan una relación directa positiva mayor que las dimensiones personales con la titulación de un egresado.

Asimismo hay evidencia suficiente para aceptar la H_{02} de tal forma que se puede afirmar que las dimensiones académicas guardan una relación directa positiva con la titulación de un egresado mayor que la dimensión personal y/o la dimensión de contexto.

Y en el caso de la hipótesis nula H_{03} , se descarta dado que las evidencias no son suficientes para aceptarla dado que el atributo que tuvo un mayor grado de incidencia, sobre el hecho de que un egresado se titule es el resultado obtenido en el CENEVAL y este atributo pertenece a la dimensión académica.

Attribute selection output	
=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===	
number of folds (%)	attribute
1 (10 %)	1 Opc_ing
0 (0 %)	2 Edad_ing
0 (0 %)	3 Sexo
5 (50 %)	4 Estado_Civil
0 (0 %)	5 Municipio_Nac
0 (0 %)	6 Estado_Nac
0 (0 %)	7 Bachillerato_Sost
0 (0 %)	8 Municipio_Bachiller
0 (0 %)	9 Est_bach
6 (60 %)	10 Grad_Marg_Bach
0 (0 %)	11 Prom_Bachiller
7 (70 %)	12 Promedio_Lic
0 (0 %)	13 Time_Egreso
0 (0 %)	14 Var_Eficiencia
0 (0 %)	15 Edad_Eg
0 (0 %)	16 Por_Ordina
1 (10 %)	17 Prom_Entor
1 (10 %)	18 Prom_Mat
0 (0 %)	19 Prom_Arq
1 (10 %)	20 Prom_Reddes
0 (0 %)	21 Prom_SB
1 (10 %)	22 Prom_Prog
0 (0 %)	23 Prom_Trata
0 (0 %)	24 Prom_Inter
0 (0 %)	25 Status_CENEVAL
10 (100 %)	26 Res_CENEVAL

Figura 8 Resultados del algoritmo de selección de atributos BestFirst

Por último, como se puede observar en el gráfico 1, se puede confirmar que la variable que tuvo mayor incidencia en que un estudiante se titulará según la información que arroja este estudio, es el resultado que el estudiante obtuvo en la aplicación de el CENEVAL conforme el tipo de evaluación que es aplicado para el PAI que en este caso es el EGEI-Info. El gráfico indica que el mayor número de titulados se concentra en el grupo de egresados que aprobaron esta evaluación.

Agradecimientos

A la Universidad Autónoma de Nayarit por las facilidades brindadas para el acceso a la información. Al Dr. Luis Fernando Curi Quintal de la Universidad Autónoma de Yucatán y a la Mtra. Edith Leticia Torres Arias de la Universidad Autónoma de Nayarit, por las observaciones vertidas a esta investigación que sirvieron para fortalecer los resultados obtenidos.

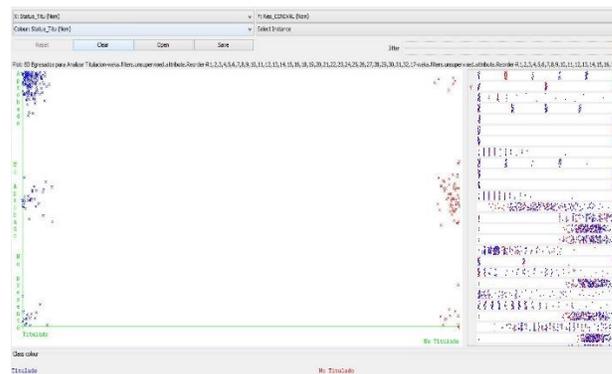


Gráfico 1 Visualización de los resultados obtenidos con los algoritmos de selección de atributos. Atributo: Res_CENEVAL

Conclusiones

La calidad de la educación ha sido un factor de comparación y medición entre instituciones de educación superior y México no es la excepción. Uno de los índices más observados en la calidad de la educación es el índice de titulación. La titulación de un egresado no es solamente responsabilidad de la institución educativa, pero si afecta al ser comparada y evaluada por una entidad acreditadora.

Para estudiar dicha problemática se realizó una investigación exploratoria, tomando como principal fuente de entrada los datos obtenidos de diversas bases de datos de control escolar, referente a 245 estudiantes egresados del Programa Académico de Informática de la Universidad Autónoma de Nayarit, pertenecientes a las cohortes generacionales en cada año del 2003 al 2009. De cada alumno se clasificó su información en dimensiones personales, académicas y de contexto.

Con los datos recopilados se realizó la selección, preprocesado y transformación de la información, para aplicar algoritmos de minería de datos, por medio del software de código abierto Weka.

Para explicar el comportamiento de la variable dependiente, titulación de un egresado, se calculó la correlación con cada atributo que describe al alumno, empleando el algoritmo CfsSubsetEval (con el método de búsqueda Bestfirst y el modo de selección cross-validation) y el algoritmo OneRAttributeEval (con el clasificador OneR y el método Ranker).

Del estudio estadístico se obtuvo que la dimensión académica es un factor preponderante en la titulación de un egresado del Programa Académico de Informática, de forma más específica, la variable asociada al índice de aprobación del examen general de egreso de licenciatura.

Como trabajo futuro se considera definir un conjunto de estrategias centradas en la dimensión académica para una mejora de resultados en el índice de titulación; además de definir un conjunto de indicadores para medir el impacto de la aplicación de las estrategias en las instituciones educativas.

Otro trabajo futuro es comparar los resultados del estudio con otros perfiles de egreso afines y no afines al programa educativo de informática, observando las variables que conservan el mismo comportamiento.

Por último, se propone extender el estudio a otros indicadores de resultados de las IES, generando bases para definir estrategias en busca de la calidad educativa.

Referencias

Asociación Nacional de Universidades e Instituciones de Educación Superior. (2013). Mercado Laboral de Profesionistas en México. Diagnóstico 2000-2009 y Prospectiva 2009-2015 y 2009-2020. México: ANUIES.

Cámara de Diputados del H. Congreso de la Unión. Secretaría General. Secretaría de Servicios Parlamentarios. (29 de enero de 2016). Obtenido de http://www.diputados.gob.mx/LeyesBiblio/pdf/1_29ene16.pdf

Comités Interinstitucionales para la Evaluación de la Educación Superior, A.C. (2015). Principios y estándares para la evaluación de programas educativos en las instituciones de educación superior de México. Recuperado el marzo de 2016, de <http://www.ciees.edu.mx/images/documentos/Estandares%20y%20principios%20para%20la%20evaluacion%20de%20programas%20de%20educacion%20superior%20VA%20MAYO%2016.pdf>

Consejo de Acreditación de la Enseñanza de la Ingeniería, A.C. (s.f.). Marco de Referencia para la Acreditación de los Programas de Licenciatura (2014). Recuperado el marzo de 2016, de <http://cacei.org.mx/index.php/acreditacion/formatos-y-manuales/manual-del-marco-de-referencia-2014>

Consejo de Acreditación en Ciencias Administrativas, Contables y Afines. (s.f.). Recuperado el marzo de 2016, de www.caceca.org

Consejo Nacional de Acreditación en Informática y Computación, A.C. (2013). Marco Referencia para la Acreditación.

Díaz Pérez, C. (2012). Tendencias y requerimientos del mercado de trabajo en la economía del conocimiento. Estudio sobre los egresados del CUCEA. Revista de la Educación Superior. ANUIES, 41(161), 9-30.

Durán Ramos, T. (2002). Globalización y formación profesional: ¿Nuevas reglas? *Revista Educación*, 26(2), 107-116.

Garré, M., Cuadrado, J. J., & Sicilia, M. Á. (s.f.). Universidad del País Vasco. Facultad de Informática. Documento: Comparación de diferentes algoritmos de clustering en la estimación de coste en el desarrollo de software. Recuperado el marzo de 2016, de <http://www.sc.ehu.es/jiwdocoj/remis/docs/GarreAdis05.pdf>

Hernández, Fernández y Baptista. (2003). *Métodología de la Investigación*. 3era. Edición. México. Mc Graw Hill.

Kisilevsky, M., & Roca, E. (2010). *Evaluación. Indicadores, Meta y Políticas Educativas*. Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura. OEI.

Martínez, S. (2015). Índice y modalidades de titulación en comunicación: Caso Universidad Autónoma de Chiapas". *Razón y Palabra*. ITESM-CEM, sep-nov(91).

Narro Robles, J., Marthuscelli Quintana, J., & Barzana García, E. (2012). *Plan de diez años para desarrollar el Sistema Educativo Nacional*. (Dirección General de Publicaciones y Fomento Editorial, UNAM) Recuperado el marzo de 2016, de www.planeducativonacional.unam.mx

Periódico Oficial del Estado de Nayarit. (s.f.). Unidad General de Asuntos Jurídicos. Secretaría de Gobernación. Recuperado el marzo de 2016, de <http://www.ordenjuridico.gob.mx/Publicaciones/CDs2008/CDEducacion/pdf/NAYLEY05.pdf>

Sánchez, C. (21 de octubre de 2009). Expansión con alianza con CNN. Recuperado el marzo de 2016, de www.cnnexpansion.com/mi-carrera/2009/10/20/titularse-o-no-titularse-es-la-cuestion

Scheerens, J. (2004). Perspectives on Education Quality, Education Indicators and Benchmarking. *European Educational Research Journal*, 3(1), 115-138.

Toscano de la Torre, B. A., Ponce Gallegos, J. C., Díaz Mendoza, J. C., & Carrillo Ortíz, M. G. (2015). Análisis de la Eficiencia Terminal Basado en Técnicas de Minería de Datos para Identificar los Factores que Determinan su Logro. En *Tendencias de las Tecnologías de la Información y Comunicaciones* (págs. 309-322). Puerto Vallarta, Jalisco, México: Alfa Omega.

Universidad Autónoma de Nayarit. (2002). Acuerdo que Establece las Bases Generales para la Titulación de Licenciatura por Área del Conocimiento. Tepic, Nayarit: Gaceta UAN.

Vargas Leyva, M., & Jiménez Hernández, M. (2). *Programas Acreditados y Estrategias de Titulación*. *Revista Electrónica ANFEI*, julio-diciembre 2015(3).

Vargas Pureko, A., & Rivera Michelena, N. (2006). El examen de titulación en odontología como indicador de calidad profesional en la Universidad Michoacana de San Nicolás de Hidalgo, México. *Educación Médica Superior*, 20(3).