

Estudio del impacto de un curso de nivelación en el desempeño de alumnos de ingeniería utilizando Minería de Datos Educativa

Beatriz A. González-Beltrán, Silvia B. González-Brambila,
Lourdes Sánchez-Guerrero, Irma Ardón-Pulido, Josué Figueroa-González

Universidad Autónoma Metropolitana, Unidad Azcapotzalco,
México
{bgonzalez,sgb,lsg,ifap,jfgo}@azc.uam.mx

Resumen. En un entorno educativo se genera gran cantidad de información que analizada adecuadamente puede ser de utilidad en la toma de decisiones. La Minería de Datos Educativa consiste en utilizar técnicas de Minería de Datos para analizar información académica, de tal manera que se pueda obtener conocimiento de diversos aspectos educativos, siendo uno de los más estudiados, el desempeño escolar. En la Universidad Autónoma Metropolitana, se crearon cursos de nivelación para que los alumnos de recién ingreso a un programa de ingeniería, puedan tener un nivel adecuado de matemáticas básicas y de física. Estos cursos pueden acreditarse a través de un examen diagnóstico o cursándolos de manera presencial. Actualmente no se ha analizado si la existencia de estos cursos, ya sea que se cursen o acrediten a través de un examen, tiene un efecto positivo en el desempeño de los alumnos en cursos posteriores. Este trabajo presenta un análisis, utilizando Reglas de Asociación, del impacto que tiene cursar o aprobar mediante un examen el curso de nivelación de Matemáticas básicas en el desempeño de los alumnos en dos cursos posteriores; de tal manera que se pueda decidir si es adecuado que este curso pueda acreditarse mediante un examen, deba cursarse o incluso, pudiera desaparecer de los planes de estudio.

Palabras clave: Análisis de planes de estudio, cursos de nivelación, desempeño escolar, minería de datos educativa, reglas de asociación.

Study of the Impact of a Leveling Course on the Performance of Engineering Students Using Educational Data Mining

Abstract. In an educational environment, a large amount of information is generated, when properly analyzed, can be useful in decision-making. Educational Data Mining uses Data Mining techniques for analyzing academic information, in such a way that knowledge of different educational aspects can be obtained, being one of the most studied, the academic performance. In the Universidad Autónoma Metropolitana, leveling courses were created so that students who have recently entered

to an engineering program can obtain an adequate level of mathematics and physics. These courses can be approved through a diagnostic exam or attending the course. Actually, it has not been analyzed if the existence of these courses, whether they are taken or approved through an exam, has a positive effect on the student's performance in subsequent courses. This paper presents an analysis, using Rules of Association, of the impact of attending a Basic Mathematics leveling course or passing it through an exam, over the performance of students in two subsequent courses; in such a way that it could be possible to decide if it is appropriate that this course can be accredited through an exam, it must be taken or even, disappear from the curricula.

Keywords: Analysis of curricula, leveling courses, school performance, educational data mining, association rules.

1. Introducción

Con el rápido crecimiento de la tecnología, la cantidad de datos que se producen ha permitido obtener información valiosa a través de su análisis. El proceso de obtener conocimiento a través de los datos es el objetivo de la Minería de Datos, la cual se aplica a muchos ámbitos de la sociedad, economía, medicina, etc. Cuando se analizan datos en un entorno educativo con técnicas de Minería de Datos, se habla de una rama denominada Minería de Datos Educativa [7,8], la cual ha tenido un gran crecimiento en los últimos años y ha permitido analizar diversos factores relacionados con la educación [12], siendo uno de los más estudiados el desempeño escolar de los estudiantes. Desde 2008 en la Universidad Autónoma Metropolitana (UAM), Unidad Azcapotzalco, los alumnos de nuevo ingreso de los diez programas de ingeniería que se imparten, presentan un examen diagnóstico que mide conocimientos básicos de matemáticas y de física; los que no aprueban, deben tomar los cursos de “Taller de Matemáticas” e “Introducción a la Física” [4] pertenecientes al Tronco de Nivelación Académica. El objetivo de estos cursos es que el alumno adquiera los conocimientos y habilidades necesarias para futuros cursos más avanzados; sin embargo, no se ha analizado si estos cursos de nivelación están ayudando a los estudiantes en los siguientes cursos. Es importante resaltar que el hecho de cursar estas dos materias implica un trimestre más en el tiempo de finalización de los estudios de los alumnos.

En este trabajo se analiza si el curso “Taller de Matemáticas” realmente está ayudando a que los alumnos tengan un mejor desempeño en dos cursos posteriores de matemáticas, “Complementos de Matemáticas” e “Introducción al Cálculo” [5]. Con el fin de identificar y analizar la relación entre el desempeño en “Taller de Matemáticas” con el desempeño en otros cursos, se utiliza la técnica de Reglas de Asociación analizando factores como la calificación obtenida en “Taller de Matemáticas”, la cantidad de veces que se cursó antes de aprobarlo y el tiempo transcurrido desde su aprobación hasta la inscripción de los siguientes cursos. El objetivo es determinar si la existencia de este curso y la forma en que

se puede aprobar está ayudando a los alumnos a estar mejor preparados para cursos más avanzados.

El contenido del documento es el siguiente: en la sección de Trabajos relacionados se presenta un análisis de los trabajos con un enfoque similar, en la sección de Metodología se presentan los pasos realizados para la obtención y análisis de las Reglas de Asociación y en la sección de Resultados y Análisis, se muestran las reglas obtenidas y su interpretación.

2. Trabajos relacionados

Existen diversos trabajos que a través de la Minería de Datos Educacional han analizado ciertas problemáticas en la educación superior. El trabajo de [3] propone descubrir Reglas de Asociación utilizando los algoritmos SLP-Growth (*Significant Least Pattern Growth*) y la medida CRS (*Critical Relative Support*). En [1,2], a partir de los datos de los alumnos de computación para el ingreso 2008/2009 de la University Malaysia Terengganu, se analiza un conjunto de datos que contiene los registros de los programas preferidos que son seleccionados por los estudiantes aceptados en dicha universidad. Los resultados muestran que se puede extraer un menor número de reglas de asociación interesantes comparado con las medidas tradicionales (análisis de correlación). Además, sugieren que sus resultados sean analizados por las autoridades universitarias para ofrecer programas más apropiados a los estudiantes prospectos en lugar de que sea de manera aleatoria.

En [11] se propone el uso del valor añadido (o equivalente a la medida *lift*) y del coseno como medidas adecuadas para datos educacionales. Esta propuesta se aplica en un curso titulado “Conceptos Básicos y Formales de Ciencias de la Computación” en la plataforma Moodle de la University of Applied Sciences TFH-Berlin durante el semestre de invierno 2007-2008. El curso de Moodle ofrece recursos adicionales a los ofrecidos en la clase presencial y los profesores están interesados en saber si el uso de estos recursos tienen un impacto positivo en sus calificaciones. Este trabajo propone el uso del coseno en primer lugar, y utilizar *lift* si la regla de asociación es considerada como no interesante con la medida del coseno. Si los resultados son contradictorios, entonces los profesores deben decidir si se toma o no en cuenta la regla de asociación.

El uso de árboles de decisión como técnica de clasificación, se propone en [6], esto aplicado sobre los datos de los alumnos del Departamento de Aplicaciones Computacionales en los años 2007-2010 de la VBS Purvanchal University, Jaunpur (Uttar Pradesh). El objetivo de este trabajo es predecir el desempeño de los alumnos al final del semestre. Los datos utilizados fueron: asistencia, pruebas en los cursos, seminarios y calificaciones de tareas. Los resultados podrían utilizarse para identificar a los estudiantes que necesitan atención especial para reducir la tasa de fracaso y tomar las acciones apropiadas para las siguientes pruebas semestrales.

En [9] se aplican algoritmos de clasificación basados en Reglas de Asociación y Árboles de Decisión sobre los datos de los alumnos de nuevo ingreso del

Programa II de la Unidad Académica Preparatoria de la Universidad Autónoma de Zacatecas del año académico 2009-2010 para predecir el estado académico de los estudiantes al final del primer semestre. Los resultados podrían utilizarse en la detección de alumnos con riesgo, cuya medida pudiera ser la asignación de un profesor-tutor para evitar la deserción del alumno.

En [10], se utilizan Reglas de Asociación basadas en el método Apriori para descubrir las conexiones entre las actividades de los estudiantes y sus calificaciones finales. La propuesta fue aplicada sobre la bitácora de datos de los alumnos del curso de “Programación 2” extraídos del sistema Moodle de la Universidad de Rijeka. Los resultados muestran que existe una influencia entre las actividades y el éxito del curso y que la creación de video-clases está justificada. La Tabla 1 presenta una comparativa de los trabajos analizados.

Tabla 1. Tabla comparativa de los trabajos relacionados.

Trabajo	Técnica de clasificación	Aplicación
[1,2] [11]	Reglas de asociación Reglas de asociación	Programas preferidos por los alumnos Relación del uso de los recursos en Moodle y la calificación obtenida en el curso
[6]	Árboles de decisión	Relación de notas obtenidas en los instrumentos de evaluación y nota final
[9]	Reglas de asociación y Árboles de decisión	Relación de datos de ingreso y nota al final del primer semestre
[10]	Reglas de asociación	Influencia entre las actividades y el éxito del curso

Como se presenta en los trabajos relacionados, el aplicar la Minería de Datos Educativa a información en un entorno educativo ha permitido encontrar aspectos que impactan en el desempeño escolar de los alumnos. La validez del curso de nivelación “Taller de Matemáticas” ha sido debatida en la UAM, por lo que es necesario realizar un estudio del impacto que pudiera tener. En este trabajo utilizamos la técnica de Reglas de Asociación para encontrar relaciones entre el desempeño en “Taller de Matemáticas” con dos cursos posteriores, estos primeros resultados serán validados en futuros trabajos con otro tipo de técnicas, por ejemplo, predictivas, las cuáles también han mostrado obtener resultados importantes en este tipo de análisis.

3. Metodología

Tomando como base la metodología CRISP-DM [13], se aplicaron los siguientes pasos para analizar la relación entre el curso de nivelación “Taller de Matemáticas” con los cursos “Complementos de Matemáticas” e “Introducción al Cálculo”.

3.1. Planteamiento de los objetivos

En esta fase se definen los objetivos del estudio. El curso de nivelación “Taller de Matemáticas” tiene como objetivo académico que los alumnos que lo cursan adquieran un mejor nivel en conceptos de matemáticas para cursos futuros. Con esta idea, se puede suponer que:

- Un alumno que no cursa “Taller de Matemáticas” por haber aprobado el examen diagnóstico, tendrá un buen desempeño en los cursos de “Complementos de Matemáticas” e “Introducción al Cálculo”.
- Un alumno que cursa y tiene un buen desempeño en “Taller de Matemáticas”, también lo tendrá en los cursos de “Complementos de Matemáticas” e “Introducción al Cálculo”.

Se busca obtener relaciones que permitan corroborar o refutar estas suposiciones.

3.2. Selección de los datos

Para este estudio se utilizó la siguiente información relacionada con los alumnos y su desempeño escolar:

- La modalidad en que se aprobó “Taller de Matemáticas”, ya sea por examen diagnóstico o por inscripción.
- La calificación obtenida en “Taller de Matemáticas”, ya sea que por inscripción o por examen diagnóstico. En caso de que lo haya inscrito, también se considera el número de oportunidades que utilizó el alumno para aprobar dicho curso.
- La calificación que obtuvo el alumno en su primera oportunidad al cursar “Complementos de Matemáticas”.
- La calificación que obtuvo el alumno en su primera oportunidad al cursar “Introducción al Cálculo”.

Los datos se obtuvieron de dos fuentes: el Archivo General de Alumnos (AGA) y el historial académico de los alumnos (llamado kardex en la UAM). Dado que los cursos “Taller de Matemáticas”, “Complementos de Matemáticas” e “Introducción al Cálculo” deben aprobarlos todos los estudiantes de ingeniería, se analizó el historial de los alumnos de las diez ingenierías que se imparten en la UAM. Se obtuvieron datos a partir del año 2008, fecha en la que surgió el curso de nivelación “Taller de Matemáticas”. De esta forma se analizaron 5,181 alumnos con las características mostradas en la Tabla 2.

Como primera aproximación del impacto de “Taller de Matemáticas” en “Complementos de Matemáticas” e “Introducción al Cálculo”, se obtuvo una estadística que muestra la cantidad de alumnos que obtuvieron una de las cuatro posibles calificaciones que se manejan en la UAM: MB (Muy bien), B (Bien), S (Suficiente), NA (No Acreditada), según su desempeño en “Taller de Matemáticas”. La Tabla 3 muestra los resultados para “Complementos de Matemáticas” y la Tabla 4 para “Introducción al Cálculo”.

Tabla 2. Desempeño de los alumnos en “Taller de Matemáticas”.

Calificación en “Taller de Matemáticas”	No. alumnos que aprobaron en examen diagnóstico	No. alumnos que cursaron y aprobaron		
		en primera oportunidad	en segunda oportunidad	en tercera o mas oportunidades
Muy Bien	277	771	113	9
Bien	463	1384	210	30
Suficiente	0	1439	394	97

Tabla 3. Desempeño de los alumnos en “Complementos de Matemáticas” de acuerdo a la calificación en “Taller de Matemáticas”.

Calificación en “Taller de Matemáticas”	Calificación en “Complementos de Matemáticas”			
	MB	B	S	N
MB en examen diagnóstico	108	73	35	51
B en examen diagnóstico	109	150	85	112
MB en primera oportunidad	197	211	157	194
B en primera oportunidad	180	321	306	514
S en primera oportunidad	133	259	308	739
MB en segunda oportunidad	15	32	28	36
B en segunda oportunidad	27	43	39	99
S en segunda oportunidad	27	72	90	205
MB en tercera oportunidad	0	0	4	5
B en tercera oportunidad	2	5	7	15
S en tercera oportunidad	6	25	13	53

Tabla 4. Desempeño de los alumnos en “Introducción al Cálculo” de acuerdo a la calificación en “Taller de Matemáticas”.

Calificación en “Taller de Matemáticas”	Calificación en “Introducción al Cálculo”			
	MB	B	S	N
MB en examen diagnóstico	86	52	21	53
B en examen diagnóstico	69	103	78	142
MB en primera oportunidad	139	150	131	210
B en primera oportunidad	105	185	232	548
S en primera oportunidad	77	152	181	771
MB en segunda oportunidad	13	23	15	46
B en segunda oportunidad	12	34	24	106
S en segunda oportunidad	28	50	59	205
MB en tercera oportunidad	3	1	1	4
B en tercera oportunidad	0	2	4	20
S en tercera oportunidad	6	12	17	58

3.3. Preparación de los datos

Se generaron el conjunto de antecedentes y consecuentes necesarios para analizar la información de las Tablas 3 y 4 con la técnica de Reglas de Asociación. Tomando en cuenta el trimestre de ingreso del alumno, si éste, en ese mismo trimestre había cursado y aprobado “Taller de Matemáticas” y cursado, mas no necesariamente aprobado, las materias de “Complementos de Matemáticas” e “Introducción al Cálculo”, significaba que había aprobado a través del examen diagnóstico la materia de “Taller de Matemáticas”. Se generaron un conjunto de antecedentes y consecuentes para la relación entre “Taller de Matemáticas” y “Complementos de Matemáticas” y otro conjunto para “Taller de Matemáticas” e “Introducción al Cálculo”.

Los antecedentes que se analizaron en ambos cursos fueron los siguientes:

Antecedente 1. Calificación del alumno en “Taller de Matemáticas” de acuerdo a la escala de calificaciones: MB, B y S.

Antecedente 2. La modalidad en que el alumno aprobó “Taller de Matemáticas”:

- APROBO. Acreditó el curso a través del examen diagnóstico
- PRIMER. Aprobó el curso en su primera oportunidad
- SEGUNDO. Aprobó el curso en su segunda oportunidad
- MAS_DE_DOS. Necesitó de más de dos oportunidades para aprobar (el número máximo de oportunidades es de cinco)

Antecedente 3. El número de trimestres que el alumno dejó pasar antes de inscribir “Complementos de Matemáticas” o “Introducción al Cálculo”:

- EXAMEN. Aprobó el examen diagnóstico y en el trimestre de ingreso cursó “Complementos de Matemáticas” e “Introducción al Cálculo”
- SIGUIENTE. Al siguiente trimestre que aprobó “Taller de Matemáticas”, cursó “Complementos de Matemáticas” o “Introducción al Cálculo”
- UNO. Dejó pasar un trimestre antes de inscribir “Complementos de Matemáticas” o “Introducción al Cálculo”
- DOS. Dejó pasar dos trimestres antes de inscribir “Complementos de Matemáticas” o “Introducción al Cálculo”
- MAS_DE_DOS. Dejó pasar tres o más trimestres antes de inscribir “Complementos de Matemáticas” o “Introducción al Cálculo”

El consecuente estaba relacionado con la calificación obtenida en “Complementos de Matemáticas” o “Introducción al Cálculo” la primera vez que se cursaron, pudiendo tenerse cualquiera de las siguientes calificaciones: MB, B, S o NA.

Un ejemplo del formato de los datos procesados que representaría que: El alumno aprobó en examen diagnóstico “Taller de Matemáticas” con MB y obtuvo B en “Complementos de Matemáticas” es:

MB,APROBO,EXAMEN,B

De manera similar, se tiene: El alumno cursó “Taller de Matemáticas”, lo aprobó al primer intento con B, dejó pasar un trimestre antes de inscribir “Introducción al Cálculo” y cuando la aprobó, obtuvo MB.

B,PRIMER,UNO,MB

En total se obtuvieron 5,095 conjuntos de antecedentes y consecuentes para la relación entre “Taller de Matemáticas” y “Complementos de Matemáticas” y 4231 para la relación “Taller de Matemáticas” e “Introducción al Cálculo”. La diferencia se debe a que no todos los alumnos habían cursado alguna de las materias.

3.4. Modelado

Una vez obtenidos los conjuntos de antecedentes y consecuentes, se obtuvieron las reglas de asociación utilizando el algoritmo Apriori. Los resultados se presentan en la sección de Resultados y análisis.

3.5. Evaluación

Una vez obtenidas las reglas de asociación, se seleccionaron las más representativas. Existen varios criterios para determinar la validez e importancia de las reglas, siendo la más conocida el **soporte**, que representa el porcentaje de elementos X en un conjunto de transacciones D y se define en (1):

$$\text{soporte}(X) = \frac{|X|}{|D|}. \quad (1)$$

Otro criterio es la **confianza**, que dada una asociación $X \Rightarrow Y$, indica la cantidad de reglas que tienen a X como antecedente y a Y como consecuente. Este valor se define en (2):

$$\text{confianza}(X \Rightarrow Y) = \frac{\text{soporte}(X \cup Y)}{\text{soporte}(X)}. \quad (2)$$

El criterio más utilizado es **lift**, que representa la frecuencia de un conjunto de elementos X en una determinada transacción que llevan a una consecuencia Y y está representado en (3):

$$\text{lift}(X \Rightarrow Y) = \frac{\text{soporte}(X \Rightarrow Y)}{\text{soporte}(X) * \text{soporte}(Y)}. \quad (3)$$

Cuando el valor de *lift* es igual a 1, significa que la relación entre antecedentes y consecuentes, puede deberse a factores aleatorios. Si el valor es menor a 1 significa que no hay tanta relación entre antecedentes y consecuentes y se dice que X y Y son sustitutos. Si el valor de *lift* es mayor a 1, se dice que hay una mayor relación entre X y Y y se denominan complementos.

Para la evaluación de las reglas, se consideraron aquellas reglas que tuvieran un valor de *lift* mayor o igual a 1.1 y un soporte mayor a 0.01. A partir de eso, se analizaron las reglas considerando un umbral de confianza de 0.3.

4. Resultados y análisis

Para facilitar el análisis, se asignaron los siguientes acrónimos a los antecedentes y consecuentes:

- CTA. Calificación en “Taller de Matemáticas”
- CUR_APR. Forma en que el alumno aprobó “Taller de Matemáticas” en examen
- TIEM. Tiempo antes de inscribir “Complementos de Matemáticas” o “Introducción al Cálculo”

Mientras que para los consecuentes:

- CCOM. Calificación en “Complementos de Matemáticas”
- CCAL. Calificación en “Introducción al Cálculo”

Tras el procesamiento, se obtuvieron un conjunto de 14 reglas para la relación entre “Taller de Matemáticas” y “Complementos de Matemáticas”, se realizó un filtrado para eliminar la redundancia obteniéndose finalmente 6, las cuales se muestran en la Tabla 5.

Tabla 5. Reglas de asociación para la relación entre “Taller de Matemáticas” y “Complementos de Matemáticas”.

No	Regla	Soporte	Confianza	Lift
1	{CTA=MB,CUR_APR=APROBO,TIEM=EXAMEN} ⇒ {CCOM=MB}	0.021	0.40	2.55
2	{CTA=B,CUR_APR=APROBO,TIEM=EXAMEN} ⇒ {CCOM=B}	0.029	0.32	1.40
3	{CTA=S,CUR_APR=PRIMER,TIEM=SIGUIENTE} ⇒ {CCOM=NA}	0.120	0.52	1.33
4	{CTA=S,CUR_APR=SEGUNDO,TIEM=SIGUIENTE} ⇒ {CCOM=NA}	0.021	0.51	1.29
5	{CTA=S,CUR_APR=SEGUNDO,TIEM=UNO} ⇒ {CCOM=NA}	0.011	0.47	1.20
6	{CTA=B,CUR_APR=SEGUNDO,TIEM=SIGUIENTE} ⇒ {CCOM=NA}	0.010	0.44	1.12

Para facilitar su interpretación, generamos una representación visual la cual se muestra en la Figura 1. Aquí, el tamaño del círculo representa el valor del *lift*, a mayor *lift*, mayor el tamaño del círculo, mientras que el tono del círculo representa la confianza, entre más oscuro, mayor confianza.

De acuerdo a la primera y segunda regla de la Tabla 5 y a la Figura 1, se puede observar que el aprobar el examen diagnóstico con MB o B sin tener que cursar “Taller de Matemáticas”, está relacionado con obtener MB o B en “Complementos de Matemáticas”.

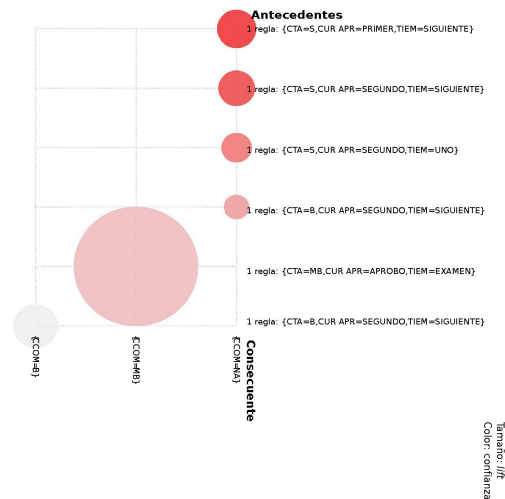


Fig. 1. Representación gráfica de las reglas entre “Taller de Matemáticas” y “Complementos de Matemáticas”.

Estas reglas tienen un valor de *lift* mayor que el resto y un nivel de confianza para el primer caso de 40 % y de 32 % para el segundo caso.

Los casos con mayor ocurrencia fueron aquellos donde el consecuente tiene una calificación de NA en “Complementos de Matemáticas” y un antecedente de S en “Taller de Matemáticas”, en la primera o segunda oportunidad. Esto se refleja en la tercera, cuarta y sexta regla, por lo que es posible determinar que el aprobar “Taller de Matemáticas” con S en una primera o segunda oportunidad u obtener B en la segunda oportunidad y tomar “Complementos de Matemáticas” al siguiente trimestre, tiene una mayor relación con una calificación de NA. Teniendo un nivel de confianza para el primer caso de 52 %, 51 % para el segundo y 44 % para el tercero.

De la quinta regla, se entiende que el aprobar “Taller de Matemáticas” con S en la segunda oportunidad y dejar pasar un trimestre para cursar “Complementos de Matemáticas” está relacionado con obtener una calificación de NA. Esto con un nivel de confianza de 47 %.

De manera similar, se obtuvieron 19 reglas para la relación entre “Taller de Matemáticas” e “Introducción al Cálculo”, las cuales tras ser filtradas para eliminar las redundantes, dieron como resultado las 7 reglas mostradas en la Tabla 6.

La representación gráfica de estas relaciones se muestra en la Figura 2. La relación entre tamaño y color con confianza y *lift* es la misma que para la Figura 1.

Analizando las reglas de la Tabla 6 y de la Figura 2, observamos que las reglas en donde la calificación en “Introducción al Cálculo” es MB tienen un *lift* mayor y son aquellas en las que se aprobó “Taller de Matemáticas” con MB en

Tabla 6. Reglas de asociación para la relación entre “Taller de Matemáticas” e “Introducción al Cálculo”.

No	Regla	Soporte	Confianza	Lift
1	{CTA=MB,CUR_APR=APROBO,TIEM=EXAMEN} ⇒ {CCAL=MB}	0.020	0.40	3.17
2	{CTA=S,CUR_APR=PRIMER,TIEM=SIGUIENTE} ⇒ {CCAL=NA}	0.129	0.66	1.29
3	{CTA=S,CUR_APR=SEGUNDO,TIEM=SIGUIENTE} ⇒ {CCAL=NA}	0.025	0.65	1.27
4	{CTA=S,CUR_APR=PRIMER,TIEM=MAS_DE_DOS} ⇒ {CCAL=NA}	0.025	0.64	1.26
5	{CTA=S,CUR_APR=PRIMER,TIEM=DOS} ⇒ {CCAL=NA}	0.013	0.64	1.26
6	{CTA=B,CUR_APR=SEGUNDO,TIEM=SIGUIENTE} ⇒ {CCAL=NA}	0.011	0.60	1.19
7	{CTA=S,CUR_APR=PRIMER,TIEM=UNO} ⇒ {CCAL=NA}	0.014	0.58	1.15

el examen diagnóstico. Esto se ve reflejado en la primera regla, con un nivel de confianza del 40 %.

Por otra parte, obtener NA en “Introducción al Cálculo” está relacionado con cursar y aprobar “Taller de Matemáticas” con S o B. Esto se observa en las reglas dos y tres, en dónde aprobar con S al primer o segundo intento “Taller de Matemáticas” y cursar en el siguiente trimestre “Introducción al Cálculo” está relacionado con una calificación de NA. Estas reglas tienen un nivel de confianza del 66 % para el primer caso y del 65 % para el segundo.

En el caso de la cuarta y quinta regla, el cursar “Taller de Matemáticas” y obtener una calificación de S en la primera oportunidad y dejar pasar dos o más trimestres para cursar “Introducción al Cálculo” también está relacionado con una calificación de NA, esto con un nivel de confianza del 64 % para ambas reglas.

La sexta regla indica que cursar y aprobar “Taller de Matemáticas” con una calificación de B en la segunda oportunidad y cursar al siguiente trimestre “Introducción al Cálculo”, está relacionada con una calificación de NA, esto con un nivel de confianza del 60 %.

Finalmente, la séptima regla señala que el cursar “Taller de Matemáticas”, obtener una calificación de S al primer intento y dejar pasar un trimestre para cursar “Introducción al Cálculo”, también se relaciona con una calificación de NA con un nivel de confianza del 58 %.

5. Conclusiones

En este trabajo se aplicó la Minería de Datos Educativa para determinar el impacto de un curso de nivelación en el desempeño escolar de los alumnos

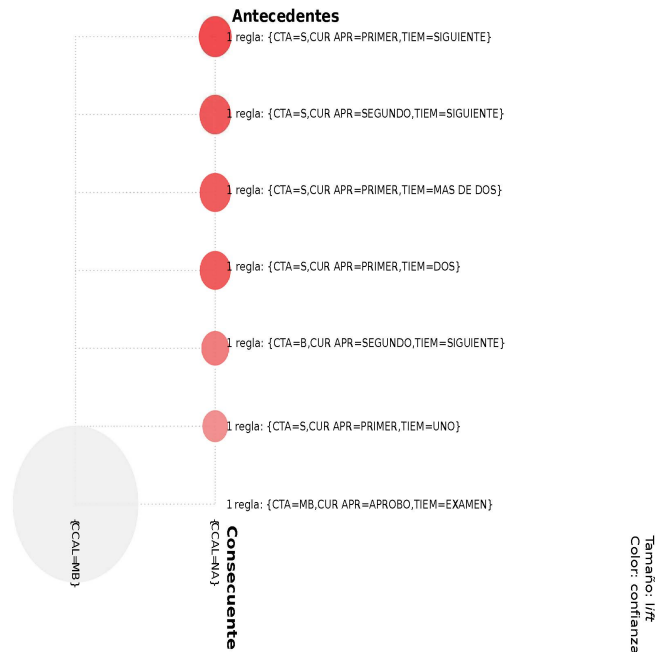


Fig. 2. Representación gráfica de las reglas entre “Taller de Matemáticas” e “Introducción al Cálculo”.

de ingeniería de la Universidad Autónoma Metropolitana de la Unidad Azcapotzalco. Para esto, se utilizó la metodología CRISP-DM con el objetivo de analizar la relación entre el curso de nivelación “Taller de Matemáticas” con los cursos de “Complementos de Matemáticas” e “Introducción al Cálculo”. Para el análisis de la información se utilizó la técnica de Reglas de Asociación empleando el algoritmo Apriori. Para esto, se establecieron varios supuestos a verificar o refutar con el análisis realizado.

El primer supuesto establecido fue que: «Un alumno que no cursa “Taller de Matemáticas” por haber aprobado el examen diagnóstico, tendrá un buen desempeño en los cursos de “Complementos de Matemáticas” e “Introducción al Cálculo”». Como resultado del análisis de este supuesto se cumple que los alumnos que obtuvieron una calificación de MB en el examen diagnóstico, obtuvieron un buen desempeño en los cursos posteriores anteriormente mencionados. Esto representa que aquellos alumnos que no cursan “Taller de Matemáticas” tienen un nivel aceptable de bases matemáticas, lo cual les permite tener un buen desempeño en sus siguientes cursos.

Sobre el segundo supuesto: «Un alumno que cursa y tiene buen desempeño en “Taller de Matemáticas”, también lo tendrá en los cursos de “Complementos de Matemáticas” e “Introducción al Cálculo”», no se encontraron reglas lo suficientemente representativas que permitieran sustentar este supuesto, por lo que no se puede decir que esto esté pasando, lo que implicaría que “Taller de

Matemáticas” no cumple con el objetivo para el que fue creado.

Otro resultado obtenido del análisis mostró que los alumnos que aprobaron el examen diagnóstico con una calificación de B, posteriormente acreditaron la materia de “Complementos de Matemáticas” pero no la de “Introducción al Cálculo”.

Los casos con mayor ocurrencia fueron aquellos alumnos que tuvieron que cursar “Taller de Matemáticas”, obteniendo una calificación de S en su primera o segunda oportunidad y que no aprobaron “Complementos de Matemáticas” ni “Introducción al Cálculo”. Otro factor presente en las reglas es que el tiempo que dejan pasar los alumnos después de aprobar la materia de “Taller de Matemáticas” e inscribir alguno de los otros cursos impacta en su desempeño de tal manera que el dejar pasar dos o más trimestres está relacionado fuertemente con no aprobar estos cursos.

Los resultados obtenidos sugieren que el contenido del curso “Taller de Matemáticas” debe de ser analizado para comprender porque no está ayudando en futuros cursos, en especial en “Introducción al Cálculo”.

Como trabajo futuro, tenemos contemplado incluir más variables en el análisis, como la edad de ingreso del alumno, desempeño en el examen de admisión y escuela de procedencia. Asimismo, plantea realizar un análisis similar en otras seriaciones del plan de estudio de diversas licenciaturas que se imparten en la Universidad Autónoma Metropolitana Unidad Azcapotzalco.

Agradecimientos. Agradecimientos a Sistemas Escolares, a la Dirección de Ciencias Básicas e Ingeniería y a Secretaría Académica de la Universidad Autónoma Metropolitana Azcapotzalco por proporcionar la información necesaria para la realización del presente trabajo.

Referencias

1. Abdullah, Z., Herawan, T., Deris, M.M.: Mining significant least association rules using fast slp-growth algorithm. In: Kim, T.h., Adeli, H. (eds.) *Advances in Computer Science and Information Technology*. pp. 324–336. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
2. Abdullah, Z., Herawan, T., Deris, M.M.: Tracing significant association rules using critical least association rules model. *International Journal of Innovative Computing and Applications* 5(1), 3–17 (2013)
3. Abdullah, Z., Herawan, T., Deris, M.M.: Discovering interesting association rules from student admission dataset. In: Herawan, T., Deris, M.M., Abawajy, J. (eds.) *Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)*. pp. 135–142. Springer Singapore, Singapore (2014)
4. Azcapotzalco, U.A.M.: División de ciencias básicas e ingeniería. http://cbi.azc.uam.mx/es/CBI/Tronco_de_Nivelacion_Academica, Última consulta 20 Abr 2018
5. Azcapotzalco, U.A.M.: División de ciencias básicas e ingeniería. http://cbi.azc.uam.mx/es/CBI/Planes_Programa_Estudio_Com, Última consulta 20 Abr 2018

6. Baradwaj, B., Pal, S.: Mining educational data to analyze students' performance 2, 63–69 (10 2011)
7. Castro, F., Vellido, A., Nebot, À., Mugica, F.: Applying data mining techniques to e-learning problems. In: Evolution of teaching and learning paradigms in intelligent environment, pp. 183–221. Springer (2007)
8. Kumar, J.: A comprehensive study of educational data mining. International Journal of Electrical Electronics & Computer Science Engineering Special Issue-TeLMISR pp. 2348–2273 (2015)
9. Marquez-Vera, C., Romero-Morales, C., Ventura-Soto, S.: Mining educational data to analyze students' performance 8, 63–69 (2 2013)
10. Matetic, M., Bakaric, M.B., Sisovic, S.: Association rule mining and visualization of introductory programming course activities. In: Proceedings of the 16th International Conference on Computer Systems and Technologies. pp. 374–381. CompSysTech '15, ACM, New York, NY, USA (2015)
11. Merceron, A., Yacef, K.: Interestingness measures for association rules in educational data. In: Educational Data Mining 2008 - 1st International Conference on Educational Data Mining, Proceedings. pp. 57–66 (01 2008)
12. Romero, C., Ventura, S.: Educational data mining: a review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 40(6), 601–618 (2010)
13. Wirth, R., Hipp, J.: Crisp-dm: Towards a standard process model for data mining. In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. pp. 29–39. Citeseer (2000)