

Applying Data Mining to Discover Successful Collaborative Groups Styles

Rosanna Costaguta
Univ. Nac. de Santiago del Estero
Santiago del Estero, Argentina
rosanna@unse.edu.ar

María de los Ángeles Menini
Univ. Nac. de Santiago del Estero
Santiago del Estero, Argentina
marameni@unse.edu.ar

Germán Lescano
CONICET
Univ. Nac. de Santiago del Estero
Santiago del Estero, Argentina
gelescano@unse.edu.ar

Abstract—In a computer supported collaborative learning environment, students are organized into groups and work together to achieve a common goal, using the computer as a communication, collaborate and coordination tool. Under the premise that every student in a group behaves in a certain way to collaborate, thus demonstrating a learning style, this research demonstrates the existence of certain relationships between learning styles of students in a collaborative group and group academic performance. To discover these combinations of learning style, collaborative experiences were carried out with groups of degree students and data mining techniques was applied.

Keywords—computer supported collaborative learning; learning style; group performance; knowledge discovery; association rules.

I. INTRODUCCIÓN

En la actualidad, las ventajas del ACSC son ampliamente reconocidas y cada día son más las instituciones del ámbito educativo que lo adoptan. Bajo esta modalidad los estudiantes están organizados en grupos y trabajan juntos para la consecución de un objetivo común, utilizando la computadora como medio de comunicación, de colaboración y de coordinación. Un grupo de aprendizaje es una estructura formada por personas que interactúan para lograr aprender a través de su participación [1].

Los estudiantes construyen su aprendizaje colaborativamente manifestando ciertas actitudes, por ejemplo: explicar y justificar sus opiniones, articular su razonamiento, elaborar y reflexionar sobre sus conocimientos o incentivar a sus compañeros a responder. La colaboración conduce a resultados positivos, sólo, cuando los integrantes del grupo dan y reciben ayuda, comparten recursos, negocian con sus pares, construyen conocimientos, adquieren competencias, resuelven conflictos y se regulan mutuamente manifestando habilidades interpersonales. Cada estudiante miembro de un grupo es responsable de su accionar individual pero también del rendimiento grupal.

Crear grupos y lograr que sus miembros resuelvan una consigna de manera colaborativa, no garantiza que el comportamiento y el rendimiento de esos grupos sean adecuados, ni que la experiencia de enseñanza y aprendizaje

sea exitosa. Un integrante tiene un comportamiento particular cuando realiza actividades dentro de su grupo. Este comportamiento no sólo pone de manifiesto, una forma de colaborar, sino un estilo de aprendizaje específico. Es por ello que la conformación de los grupos, en ACSC, resulta ser una tarea que requiere particular atención.

De acuerdo con el estudio realizado por Costaguta [2], se sabe que los grupos de ACSC se conforman de diferentes maneras: en algunos casos es el profesor quien selecciona aleatoriamente a los miembros de cada grupo, o establece algún criterio para dicha selección; en otros, son los mismos alumnos quienes se agrupan por amistad, afinidad en la forma de trabajar, etc. En ninguno de los casos analizados, la selección de los integrantes de un grupo se lleva a cabo considerando el estilo de aprendizaje de los estudiantes involucrados, tendiendo a promover un mejor desempeño de los mismos.

Esta investigación buscó dar respuesta a los siguientes interrogantes: ¿Existe alguna relación entre los estilos de aprendizaje de los estudiantes que constituyen un grupo colaborativo y el rendimiento del mismo? ¿Es posible identificar, mediante técnicas de minería de datos, combinaciones de estilos de aprendizaje de estudiantes que se vinculen a grupos con buen rendimiento? A priori, las respuestas a estas preguntas parecieron ser afirmativas, sin embargo, fue necesario responderlas rigurosa y científicamente.

El presente artículo se organiza de la siguiente manera: la sección 2 se refiere al modelo de estilos de aprendizaje que fue usado en esta investigación, en la sección 3 se describe el Proceso de descubrimiento de conocimiento (KDD), la sección 4 consiste en la aplicación de KDD, detallando los pasos que se llevaron a cabo y los resultados obtenidos en cada paso. Finalmente, en la sección 5, se enuncian algunas conclusiones.

II. ESTILOS DE APRENDIZAJE

Identificar el estilo de aprendizaje de una persona implica reconocer el método o estrategia utilizada por esa persona para aprender. Aunque las estrategias varían según lo que se quiera aprender, cada individuo tiende a desarrollar ciertas preferencias o tendencias globales que definen un estilo de aprendizaje. Según Felder [3], existen diferencias en los estudiantes con respecto a la manera en que éstos procesan la

información. Asimismo, un estudiante puede presentar varias características, por lo que el docente debería ser capaz de adaptar su estilo de enseñanza a los estilos de aprendizaje, de tal forma de no afectar negativamente el rendimiento académico del alumno o su actitud frente a los contenidos propuestos [4].

Un modelo de estilo de aprendizaje clasifica a los estudiantes de acuerdo con la forma en cómo reciben y procesan la información. En esta investigación se utilizó el modelo propuesto por Felder y Silverman [5]. Los autores consideran cuatro dimensiones como variables de análisis y dos estilos de aprendizaje por cada una de esas dimensiones. La dimensión Procesamiento se vincula con cómo el estudiante prefiere adquirir la información, la dimensión Percepción los hace con el tipo de información que el estudiante prefiere recibir, la dimensión Entrada se relaciona con la vía sensorial que el estudiante prefiere utilizar para captar la información, y finalmente, la dimensión Comprensión se asocia con el modo de presentación que facilita el entendimiento de contenidos. Los tipos de aprendizajes resultantes, considerando las dimensiones propuestas, se describen brevemente en Tabla 1.

TABLA 1. DIMENSIONES Y ESTILOS DE APRENDIZAJE

Dimensión	Estilos	Descripción
Procesamiento	<ul style="list-style-type: none"> • Activo • Reflexivo 	Un estudiante activo se siente más cómodo con la experimentación que con la observación reflexiva, al revés de un estudiante reflexivo. Un estudiante activo no aprende en situaciones de pasividad y trabaja bien en grupo. Un estudiante reflexivo no aprende en situaciones que no proporcionan la oportunidad de pensar sobre la información que se le presenta, trabaja mejor solo y tiende a ser teórico.
Percepción	<ul style="list-style-type: none"> • Sensitivo • Intuitivo 	Un estudiante sensitivo gusta de hechos, datos, y experimentaciones; resuelve problemas comprendiendo los métodos, no le agradan las sorpresas ni las complicaciones; es paciente con los detalles; es bueno memorizando hechos; es cuidadoso pero lento. Un estudiante intuitivo prefiere principios y teorías; gusta de innovaciones y complicaciones; y no de repetición; se aburre con detalles; es bueno para asimilar nuevos conceptos.
Entrada	<ul style="list-style-type: none"> • Visual • Verbal 	Un estudiante visual recuerda mejor lo que ve (figuras, diagramas, cuadros, demostraciones, etc.). Un estudiante verbal recuerda más lo que escucha, aprende a partir de la discusión, y prefiere las explicaciones verbales a las demostraciones visuales.
Comprensión	<ul style="list-style-type: none"> • Secuencial • Global 	Un estudiante secuencial aprende en un orden de progresión lógica, regido por el tiempo y el calendario. Un estudiante global no se rige por el tiempo ni el calendario, puede pasar días ocupado en resolver un simple problema o demostrando poca comprensión hasta que de repente logra una rápida comprensión del todo.

A partir de las combinaciones de los estilos, Felder y Silverman [5] establecieron la existencia de 16 estilos de aprendizaje. Así, por ejemplo, es posible encontrar estudiantes que responden a un estilo de aprendizaje sensitivo, visual, activo, secuencial; mientras que otros lo hacen a un estilo intuitivo, visual, reflexivo, global; o a cualquier otra combinación posible. Respondiendo a este modelo, Felder y Soloman [6] crearon el Test de estilos de aprendizaje, [7], un instrumento que permite determinar el estilo de aprendizaje de un estudiante, mediante cuarenta y cuatro preguntas, con dos opciones de respuesta cada una, en las que el estudiante debe elegir, necesariamente, sólo una de ellas.

III. EL PROCESO DE DESCUBRIMIENTO DE CONOCIMIENTO

Para El Proceso de Descubrimiento de Conocimiento en Bases de Datos (Knowledge Discovery in Database - KDD), consiste en identificar en grandes volúmenes de datos, patrones significativos que resulten válidos, novedosos, potencialmente útiles y comprensibles para un usuario. El proceso KDD consta de tres fases: preprocesamiento, minería de datos y postprocesamiento [8], [9] y [10].

La fase de preprocesamiento tiene por objetivo preparar los datos para la fase siguiente del proceso KDD. Dentro de las técnicas posibles de aplicar durante el preprocesamiento cabe mencionar [8]: la limpieza de datos, a fin de remover ruido e inconsistencias; la integración de datos, para generar un único almacén coherente en aquellos casos donde los datos provienen de diferentes fuentes; las transformaciones de datos, para normalizarlos y, la reducción de datos, a fin de eliminar características redundantes. La importancia del preprocesamiento se debe a que la calidad de los datos sobre los que se aplica el proceso KDD impacta de manera directa en la calidad del conocimiento que se descubre [8].

La fase de minería de datos puede definirse sobre la base de primitivas, diseñadas especialmente para facilitar el descubrimiento de conocimiento eficiente y fructífero. Tales primitivas incluyen [8]: la especificación de las porciones de la base de datos o del conjunto de datos en los que se quiere trabajar; la clase de conocimiento a ser descubierto; los conocimientos existentes que podrían resultar útiles para guiar el proceso de KDD; las métricas de interés para llevar a cabo la evaluación de patrones en los datos analizados y, finalmente, las formas en que el conocimiento descubierto podría visualizarse. Esta fase se materializa mediante la realización de tareas del tipo predictivas o del tipo descriptivas. Dentro de las tareas predictivas se pueden mencionar: la clasificación y la regresión, mientras que, entre las tareas descriptivas figuran el clustering, las correlaciones y las reglas de asociación.

La fase de postprocesamiento implica la realización de alguna reformulación de los resultados obtenidos a fin de hacer que los conocimientos descubiertos sean más fáciles de entender y utilizar por el usuario a quien finalmente están destinados.

IV. APLICACIÓN DE KDD SOBRE EL ESCENARIO PROPUESTO

Las hipótesis que guiaron esta investigación fueron las siguientes:

- Existen determinadas relaciones entre los estilos de aprendizaje de los estudiantes que constituyen un grupo colaborativo y el rendimiento académico del mismo.
- Existen combinaciones de estilos de aprendizaje entre los estudiantes integrantes de grupos colaborativos que se vinculan con un buen rendimiento académico grupal.

Para validar estas hipótesis se realizaron experiencias de aprendizaje colaborativo con estudiantes universitarios de entre 20 y 24 años en diversas asignaturas. En estas experiencias se utilizó como soporte educativo la plataforma Moodle, dado que el Centro Universitario Virtual (CUV) dependiente de la FCEyT, utiliza dicha plataforma para brindar apoyo a las actividades de docencia, investigación y extensión. Además, la misma dispone de código abierto, lo que permitió realizar modificaciones acordes con los objetivos de la investigación. El resultado de estas experiencias colaborativas fue uno de los insumos necesarios para aplicar el proceso KDD.

En primer lugar se trabajó sobre datos históricos (logs), recopilados mediante el registro de las sesiones colaborativas efectuadas. Específicamente se almacenaron, en un archivo de base de datos, los datos personales de los estudiantes y sus estilos de aprendizaje (esto último fue el resultado de la aplicación del Test de estilos de aprendizaje que respondieron todos los alumnos al comienzo de la cursada). La Figura 1 muestra el resultado del Test de estilos de aprendizaje obtenido por el estudiante a. En los extremos de cada fila de dicha figura se visualizan los estilos posibles para cada dimensión del modelo propuesto por Felder y Silverman [5], según se describió en Tabla 1. La fila 1 corresponde a la dimensión Procesamiento, la fila 2 a la dimensión Percepción, la fila 3 a la dimensión Entrada, y la fila 4 a la dimensión Comprensión. Cada una de las filas cuenta con una escala numérica que va desde -11 a 11, y un carácter x que identifica el valor asignado para la dimensión. Si x está posicionada en el intervalo comprendido entre -3 y 3, se deduce un valor neutro para la dimensión en cuestión, mientras que si señala un valor comprendido entre 5 y 11, o entre -5 y -11, está indicando el estilo correspondiente al extremo derecho o izquierdo de la fila respectivamente. Por ejemplo, el resultado obtenido por el alumno a, conforme lo muestra la Figura 1, evidencia un estilo reflexivo para la dimensión Procesamiento (x sobre el número 7 de la derecha en la primera fila); un estilo sensitivo para la dimensión Percepción (x sobre el número -7 de la izquierda en la segunda fila); un estilo visual para la dimensión Entrada (x sobre el número -9 de la izquierda en la fila 3), y finalmente, un estilo secuencial para la dimensión Comprensión (x sobre el número -5 de la izquierda en la fila 4). Por lo tanto el estilo de aprendizaje del estudiante a es: reflexivo, sensitivo, visual y secuencial.

Una vez obtenida la base de datos con los datos personales, los datos de las interacciones y los estilos de aprendizaje de cada uno de los estudiantes de los diferentes grupos, se llevaron a cabo ciertas tareas de preprocesamiento. Así, se eliminaron

los datos que para esta investigación resultaban irrelevantes y también se depuraron algunos registros incompletos. La base de datos quedó constituida por los atributos: el estilo de aprendizaje del estudiante, el identificador de grupo del cual formó el alumno y la evaluación del rendimiento académico grupal asignada por el profesor. Dado que algunos profesores evaluaron el rendimiento académico en forma cuantitativa y otros de manera cualitativa, también fue necesario realizar algunas transformaciones sobre este atributo para disponer únicamente de datos de tipo cualitativo (Tabla 2).

NC STATE UNIVERSITY	
Learning Styles Results	
Results for:	
ACT 11 9 7 5 3 1 3 5 7 9 11 REFL	x
SEN 11 9 7 5 3 1 3 5 7 9 11 INT	x
VIS 11 9 7 5 3 1 3 5 7 9 11 VRB	x
SEC 11 9 7 5 3 1 3 5 7 9 11 GLO	x
<p>If your score on a scale is 1-3, you are fairly well balanced on the two dimensions of that scale.</p> <p>If your score on a scale is 5-7, you have a moderate preference for one dimension of the scale and will learn more easily in a teaching environment which favors that dimension.</p> <p>If your score on a scale is 9-11, you have a very strong preference for one dimension of the scale.</p>	

Fig. 1. Ejemplo de resultado del Test de estilos de aprendizaje

Posteriormente, se agruparon las tuplas de la base de datos con igual identificador de grupo para transformarlas en una única tupla que contenga los datos correspondientes a un grupo. Cada nueva tupla incluye así el identificador del grupo, el estilo de aprendizaje de cada uno de los estudiantes integrantes del mismo expresado mediante cuatro valores (uno para cada dimensión), y la evaluación del rendimiento académico grupal asignada por el profesor.

TABLA 2. RELACIÓN ENTRE LA ESCALA CUANTITATIVA Y LA VALORACIÓN CUALITATIVA

Escala cuantitativa	Escala cualitativa
0-3	Malo
4-5	Regular
6-8	Bueno
9-10	Muy Bueno

A continuación se realizó una última transformación que consistió en definir clases en función del porcentaje de aparición de los diferentes estilos de aprendizaje en cada tupla, conforme se muestra en la Tabla 3. Esta transformación permitió reemplazar los estilos de aprendizaje de los estudiantes por nuevos atributos que indicaron la distribución de los diferentes estilos para cada una de las cuatro dimensiones pero tomando al grupo como un todo.

La equivalencia planteada en Tabla 3 se explica a continuación. Suponga la existencia del grupo con identificador

JJJ conformado por tres estudiantes con los estilos de aprendizaje (reflexivo, sensitivo, visual, secuencial), (activo, intuitivo, verbal, secuencial) y (activo, sensitivo, visual, secuencial). Puede observarse que para la dimensión Procesamiento el estilo reflexivo aparece sólo una vez en el grupo, representando el 33,33% de los estudiantes del grupo. Según las equivalencias indicadas en Tabla 3, ese porcentaje equivale a pocos estudiantes con el estilo reflexivo. Continuando con el análisis del ejemplo, los tres estudiantes evidencian para la dimensión Comprensión un estilo secuencial. Esto representa el 100% de los integrantes del grupo, por lo que según las equivalencias de la Tabla ya mencionada, se asocia el valor todos con el estilo secuencial. Del mismo modo, pueden realizarse las transformaciones correspondientes con los estilos restantes.

TABLA 3. EQUIVALENCIA ENTRE PORCENTAJES Y CLASES

Porcentaje	Clases
= 0	NINGUNO
>= 1 y <= 40	POCOS
>= 41 y <= 69	EQUILIBRADO
>= 70 y <= 99	DEMASIADO
= 100	TODOS

Luego de esta transformación se eliminó el atributo identificador de grupo por considerárselo innecesario, puesto que cada grupo origina una tupla distinta. Para finalizar el preprocesamiento el archivo de base de datos se transformó en uno con extensión .arff [7]. Esto fue necesario porque en la siguiente fase del proceso KDD se usará el software WEKA [11]. En la Figura 2 se muestra una porción del archivo resultante. La primera línea corresponde a la sección de encabezado (@RELATION) que incluye el nombre de la relación, seguidamente se especifican los atributos contenidos en cada tupla (@ATTRIBUTE) especificando en cada caso los valores que el atributo puede tomar, y por último, aparece la sección de datos (@DATA) con todas las tuplas o instancias contenidas.

```

@RELATION prueba
@ATTRIBUTE intuitivo {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE sensitivo {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE neutralis {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE activo {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE reflexivo {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE neutralar {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE global {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE secuencial {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE neutralgs {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE verbal {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE visual {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE neutralvv {ninguno,pocos,equilibrado,demasiado,todos}
@ATTRIBUTE rendimiento {malo,regular,bueno,excelente}

@DATA
ninguno,pocos,demasiado,ninguno,ninguno,todos,pocos,pocos,pocos,pocos,pocos,excelente
ninguno,equilibrado,pocos,ninguno,ninguno,todos,ninguno,ninguno,todos,pocos,ninguno,equilib
ninguno,equilibrado,equilibrado,pocos,pocos,equilibrado,pocos,pocos,equilibrado,ninguno,der
ninguno,demasiado,pocos,ninguno,equilibrado,equilibrado,ninguno,ninguno,todos,ninguno,todos
pocos,ninguno,equilibrado,pocos,ninguno,equilibrado,pocos,ninguno,equilibrado,ninguno,pocos
ninguno,ninguno,todos,pocos,pocos,equilibrado,ninguno,ninguno,todos,pocos,pocos,equilibrad
ninguno,equilibrado,equilibrado,pocos,ninguno,demasiado,ninguno,demasiado,pocos,ninguno,poc
equilibrado,equilibrado,todos,ninguno,ninguno,ninguno,ninguno,demasiado,ninguno,pocos,equi
ninguno,demasiado,pocos,ninguno,ninguno,todos,ninguno,ninguno,todos,pocos,pocos,equilibrad
equilibrado,equilibrado,ninguno,ninguno,demasiado,pocos,equilibrado,pocos,pocos,equilibrad
ninguno,pocos,equilibrado,todos,ninguno,ninguno,ninguno,pocos,equilibrado,ninguno,equilibr
ninguno,equilibrado,pocos,pocos,ninguno,equilibrado,pocos,ninguno,equilibrado,ninguno,todos
ninguno,ninguno,todos,equilibrado,pocos,ninguno,ninguno,pocos,equilibrado,ninguno,equilibr
ninguno,ninguno,todos,ninguno,ninguno,todos,ninguno,ninguno,todos,ninguno,ninguno,todos,bue
ninguno,todos,ninguno,equilibrado,ninguno,equilibrado,equilibrado,ninguno,equilibrado,ningu
ninguno,equilibrado,pocos,ninguno,ninguno,todos,ninguno,pocos,equilibrado,ninguno,equilibr
ninguno,pocos,equilibrado,todos,ninguno,ninguno,ninguno,pocos,equilibrado,ninguno,equilibr

```

Fig. 2. Porción del archivo con extensión .arff

La fase de minería de datos se estructuró a través de las siguientes preguntas planteadas por [8]:

a. ¿Cuál es el conjunto de datos que se consideró relevante para realizar el descubrimiento de conocimiento?

Se decidió trabajar con la base de datos resultante del preprocesamiento explicado en los párrafos previos de esta sección. Específicamente se dispusieron de 160 tuplas o instancias expresadas con el formato que muestra la sección @DATA de la Figura 2.

b. ¿Qué clase de conocimiento se quiere descubrir?

Dada la necesidad de identificar relaciones no explícitas entre los estilos de aprendizaje presentes en un grupo y el rendimiento académico grupal, se decidió utilizar una técnica de minería de datos del tipo descriptivo. Se optó por las Reglas de Asociación porque permiten descubrir hechos que ocurren en común, dentro de un conjunto de datos. Para nuestro caso, descubrir las combinaciones de estilos de aprendizaje que se asocian con un determinado rendimiento académico.

Las reglas de asociación tiene la forma $X \Rightarrow Y$, indicando que el antecedente “X” determina al consecuente “Y”. Para nuestro caso, “X” representó la combinación de los estilos de aprendizaje de los estudiantes miembros de un grupo y “Y” el rendimiento académico grupal.

El procesamiento se realizó usando el software Weka, en su opción Reglas de Asociación, y con el algoritmo “A priori”. Asimismo, se establecieron valores de soporte y de confianza en la aplicación de las Reglas de asociación. Entendiendo por soporte, el porcentaje mínimo de instancias que cubre una regla con respecto al total de instancias y, por confianza la medida que indica la frecuencia con que el consecuente “Y” es cierto, cuando también lo es, el antecedente “X”. Así, se establecieron como valores iniciales para soporte (MinMetric = 0,6) y para confianza (LowerBoundMinSupport = 0.6), siendo el objetivo de la tarea encontrar todas las reglas que superaran esos valores. También se estableció en 200 el número máximo de reglas a generar.

c. ¿Qué conocimiento de background resultará de utilidad?

Como background, resultó de utilidad recordar que cada tupla sintetiza los estilos de aprendizaje presentes en un grupo, y que esa síntesis es el resultado de una serie de acciones de preprocesamiento aplicadas sobre tuplas que inicialmente incluían la identificación del grupo, el estilo de aprendizaje del estudiante expresado en base a las cuatro dimensiones del modelo de estilos de aprendizaje de Felder y Silverman Felder y Silverman [5], y la calificación del rendimiento académico grupal otorgada por el profesor.

d. ¿Qué medidas pueden utilizarse para estimar patrones de interés?

Para este trabajo sólo resultaron de interés aquellas reglas de asociación cuyo consecuente fuera el atributo rendimiento académico grupal. Efectuado el procesamiento Weka devolvió exactamente 100 reglas (la última con una confianza de 0.62). Como primer paso se procedió a analizar las mismas a fin de descartar aquellas reglas que no tuvieran al atributo rendimiento como consecuente. Para ellos se aplicó un primer filtro de reglas mediante plantillas [12], luego del cual reglas como las que se muestran en la Figura 3 fueron eliminadas.

1. neutralgs=equilibrado 90 ==> intuitivo=ninguno 89 conf:(0.99)
2. neutralgs=equilibrado verbal=ninguno 85 ==> intuitivo=ninguno 84 conf:(0.99)
3. neutralgs=equilibrado rendimiento=bueno 85 ==> intuitivo=ninguno 84 conf:(0.99)
4. neutralgs=equilibrado verbal=ninguno rendimiento=bueno 80 ==> intuitivo=ninguno 79 conf:(0.99)
5. reflexivo=ninguno rendimiento=bueno 109 ==> verbal=ninguno 107 conf:(0.98)

Fig. 3. Ejemplos de reglas descartadas

Aplicado el filtro por plantilla las 100 reglas quedaron reducidas a 16. La Figura 4 muestra una porción de las mismas. Estas reglas pueden interpretarse fácilmente, por ejemplo, la regla de asociación 17 (primera regla contenida en la Figura 4), predice que en un grupo donde no hay integrantes intuitivos, ni reflexivos, ni verbales, el rendimiento académico grupal será bueno con un valor de confianza igual al 94%.

17. intuitivo=ninguno reflexivo=ninguno verbal=ninguno 109 ==> rendimiento=bueno 103 conf:(0.94)
20. neutralgs=equilibrado 90 ==> rendimiento=bueno 85 conf:(0.94)
23. intuitivo=ninguno neutralgs=equilibrado 89 ==> rendimiento=bueno 84 conf:(0.94)
26. neutralgs=equilibrado verbal=ninguno 85 ==> rendimiento=bueno 80 conf:(0.94)
28. intuitivo=ninguno neutralgs=equilibrado verbal=ninguno 84 ==> rendimiento=bueno 79 conf:(0.94)

Fig. 4. Ejemplos de reglas que cumplen el primer filtro

Profundizando el análisis de las reglas obtenidas, se observó la aparición de reglas redundantes por lo que se aplicó otro filtrado a fin de eliminarlas. Este filtrado se basó en el trabajo de [13], que establece que dadas dos reglas R1: $X \Rightarrow Y$, R2: $X, Z \Rightarrow Y$, con valores de confianza similares, puede considerarse que R1 es redundante respecto de R2, por cuanto no aporta nueva información y por tanto debe eliminarse. Por ejemplo, si consideramos las reglas 17, 31, 39, 41, 45, 48 y 52 (Figura 5) puede observarse que todas tienen igual consecuente (rendimiento bueno) con un valor de confianza similar, esto es, comprendido entre 91% y 94%. Sin embargo, el antecedente de la regla 17 incluye el antecedente de las restantes, dado que incorpora más estilos de aprendizaje. Por lo expuesto las reglas 31, 39, 41, 45, 48 y 52 fueron eliminadas por considerárselas redundantes.

Con idéntico criterio se eliminaron otras reglas que presentaban redundancia con las reglas 28 y 46. De esta forma las reglas resultantes del proceso KDD fueron cuatro (Figura 6).

17. intuitivo=ninguno reflexivo=ninguno verbal=ninguno 109 ==> rendimiento=bueno 103 conf:(0.94)
31. reflexivo=ninguno verbal=ninguno 114 ==> rendimiento=bueno 107 conf:(0.94)
39. intuitivo=ninguno reflexivo=ninguno 113 ==> rendimiento=bueno 105 conf:(0.93)
41. reflexivo=ninguno 118 ==> rendimiento=bueno 109 conf:(0.92)
45. intuitivo=ninguno verbal=ninguno 135 ==> rendimiento=bueno 124 conf:(0.92)
48. verbal=ninguno 145 ==> rendimiento=bueno 132 conf:(0.91)
52. intuitivo=ninguno 144 ==> rendimiento=bueno 130 conf:(0.91)

Fig. 5. Ejemplos de reglas redundantes

e. ¿Cómo se visualizarán los patrones descubiertos?

Los patrones descubiertos se visualizarán mediante cuatro sentencias IF – THEN, tal como muestra la Figura 7.

- 17: intuitivo=ninguno reflexivo=ninguno verbal=ninguno 109 ==> rendimiento=bueno 103 conf:(0.94)
- 28: intuitivo=ninguno neutralgs=equilibrado verbal=ninguno 84 ==> rendimiento=bueno 79 conf:(0.94)
- 46: intuitivo=ninguno global=ninguno verbal=ninguno 84 ==> rendimiento=bueno 77 conf:(0.92)
- 69: secuencial=ninguno 87 ==> rendimiento=bueno 73 conf:(0.84)

Fig. 6. Reglas resultantes del proceso KDD

- 1: IF estudiantes intuitivos = 0 AND estudiantes reflexivos = 0 AND estudiantes verbales = 0 THEN rendimiento = bueno
- 2: IF estudiantes intuitivos = 0 AND estudiantes neutral global-secuencial = cantidad total de integrantes del grupo +/- 1 AND estudiantes verbales = 0 THEN rendimiento = bueno
- 3: IF estudiantes intuitivos = 0 AND estudiantes globales = 0 AND estudiantes verbales = 0 THEN rendimiento = bueno
- 4: IF estudiantes secuenciales = 0 THEN rendimiento = bueno

Fig. 7. Patrones descubiertos expresados como IF-THEN

V. CONCLUSIONES

La originalidad de esta investigación consiste, no sólo, en la aplicación de KDD para determinar la relación existente entre los estilos de aprendizaje de los estudiantes que constituyen un grupo colaborativo y el rendimiento académico del mismo, sino además, en el uso de técnicas de minería de datos para identificar combinaciones de estilos de aprendizaje de estudiantes que se vinculen con un buen rendimiento académico grupal.

Los resultados obtenidos son un claro indicio de que existen ciertos vínculos entre las combinaciones de estilos de aprendizaje de los estudiantes que conforman un grupo y el rendimiento académico logrado por el mismo. Sin embargo, dado que sólo se lograron descubrir cuatro patrones, todos ellos vinculados con el rendimiento grupal bueno, se juzga necesario recolectar datos provenientes de nuevas sesiones colaborativas que permitan incrementar de forma significativa las 160 tuplas iniciales, para luego replicar el proceso KDD y eventualmente corroborar o acrecentar la cantidad de patrones descubiertos.

En el futuro también se pretende descubrir si existen vinculaciones del rendimiento académico grupal con otros factores además del ya considerado estilo de aprendizaje. Para esto se está evaluando la posibilidad de incorporar información tal como: características de personalidad, nivel de conocimiento y estados emocionales.

Por último, cabe resaltar que el conocimiento obtenido a partir de esta investigación fue usado para definir la función de evaluación de un algoritmo genético que realiza la conformación de grupos colaborativos. Las pruebas realizadas sobre el algoritmo genético, en la definición de grupos integrados por estudiantes que responden a una combinación de estilos de aprendizaje asociada con un buen rendimiento académico grupal, fueron exitosas [14].

AGRADECIMIENTOS

Este trabajo se realizó financiado parcialmente por los proyectos PICTO UNSE 2012-0016 y SECYT UNSE 23-C089.

REFERENCIAS

[1] Souto, M. 1990. Didáctica de lo grupal. Ministerio de Educación y Justicia, INPAD.

[2] Costaguta, R. 2015. Algorithms and machine learning techniques in collaborative group formation. En: *Advances in Artificial Intelligence and Its Applications, Lecture Notes in Computer Science (LNCS)*, Springer, Vol. 9414, pp. 249-258.

[3] Felder, R. 1996. Matters of Style. En: *ASEEE Prism*, vol 6(4), pp. 18-23.

[4] Felder, R. 2004. ¿Cómo estructurar la currícula en Ingeniería?. IV Congreso Argentino de Enseñanza de la Ingeniería (CAEDI), Argentina.

[5] Felder, R., y Silverman, L. 1988. Learning and Teaching Styles in Engineering Education Application. *Journal of Engineering Education*, vol. 78(7).

[6] Felder, R., y Soloman, V. 1984. Index of Learning Styles. Disp. en: <http://www.ncsu.edu/felder-public/ILSpage.html>. Último acceso: 15/11/2015.

[7] Test para descubrir estilos de aprendizaje ideado por Felder R. y Soloman, V. Disp. en: http://www.ua.es/dpto/dqino/RTM/Invest_docente/ilsweb_es.html. Último acceso: 15/11/2015.

[8] Han, J., y Kamber, M. 2001. *Data Mining: Concepts and Techniques*. USA, Academic Press.

[9] Hernández Orallo, J., Ramírez Quintana, M., y Ferri, C. 2006. *Introducción a la minería de datos*. España, Prentice Hall – Addison Wesley.

[10] Witten, I., y Frank, E. 1999. *Data Mining: Practical machine, learning tools and techniques with Java implementations*. USA, Morgan Kauffmann Publishers.

[11] Softwre WEKA. Disp. en: <http://www.cs.waikato.ac.nz/ml/weka>. Último acceso: 15/11/2015.

[12] Klemettinen, M., Mannila, H., Ronkainen, P., Toivonen, H., y Verkamo, A. 1994. Finding interesting rules from large sets of discovered association rules. En: *ACM Proc. 3rd. International Conference on Information and Knowledge Management CIKM'94*, pp. 401–407, New York, NY, USA.

[13] Liu, B., Hsu, W., y Ma, Y. 1999. Pruning and summarizing the discovered associations. En: *ACM Proc. 5th. International Conference on Knowledge Discovery and Data Mining SIGKDD'99*, pp. 125–134, New York, USA.

[14] Lescano, G., Costaguta, R., y Amandi, A. 2016. Genetic Algorithm for Automatic Group Formation Considering Student's Learning Styles. En: *ACM Proc. EATIS, Colombia*.