

Sentiments in Social Context of Student Modelling

Regina Motz

Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay
rmotz@fing.edu.uy

Ofelia Cervantes

Dept. Computación, Electrónica y Mecatrónica
Universidad de las Américas Puebla
San Andrés de Cholula, México
ofelia.cervantes@udlap.mx

Paula Echenique

División de Informática
Consejo de Formación en Educación
Montevideo, Uruguay
paula.echenique@cfe.edu.uy

Resumen—Social learning analytics is an emerging discipline that offers new methods to explore data from online educational devices in order to obtain a better understanding of student behavior. As learning takes place in heterogeneous and complex online environments, the incorporation of contextual information about the student has attracted major interest. For some time, most methods of student modelling have considered interactions as a key dimension of the student’s social context, but only recently, automatic extraction software agents begin to tackle interactions in a non-exclusively quantitative way. In this paper, we propose the discovery of sentiments in online social interactions as an additional property for the modeling of students in order to produce a contextualized diagnosis when performing learning analytics. We propose answers to the questions of “Which are the sentiments in students context modeling?”, “Why are they important for social learning analytics?”, “How can we visualize them?”.

Index Terms—social learning analytics, context student modeling, sentiment analysis

I. INTRODUCTION

El amplio uso que se hace actualmente de la tecnología en la educación genera un gran volumen de interacciones en línea como las que ocurren en foros o chats de los espacios de los LMS (Learning Management Systems) como por ejemplo Moodle, Schoology o Blackboard, o incluso en grupos de redes sociales informales como por ejemplo Facebook, Google+ o Twitter. En particular los datos de las interacciones generadas en estos ambientes son una excelente fuente para la analítica del aprendizaje social que busca evidenciar que las nuevas ideas y las nuevas habilidades no se producen de manera individual, sino que se estimulan y se desarrollan a través de la interacción y la colaboración entre los miembros de una comunidad [10], [17].

Según George Siemens [23], quien acuñó el término *conectivismo*, el aprendizaje en la era digital ya no es una actividad individualista. El conocimiento se distribuye a través de las redes. En nuestra sociedad digital, las conexiones y las conectividades dentro de las redes conducen al aprendizaje. Stephen Downes [9], enfatiza la naturaleza interconectada del aprendizaje. Sin entrar en la discusión si estas son nuevas teorías del aprendizaje, se observa que las redes de aprendizaje, se construyen a través del flujo de las interacciones grupales. Como señala Caldeiro *et al.* [6], estas interacciones, a diferencia de otras variantes de interacción grupal, por su naturaleza digital se vuelven visibles a través de los rastros que las interacciones dejan a lo largo del proceso. Es entonces

que apoyándonos en este rastro podemos realizar análisis para detectar cuáles son las interacciones que impactan en el aprendizaje con el objetivo de brindar al profesor alertas sobre el proceso de sus estudiantes.

Trabajos del análisis de redes sociales proporcionan los medios para abordar el análisis de la estructura de la red social, mientras que el análisis de contenido nos permite centrarnos en la naturaleza del vínculo [13]. Nuestro enfoque es estudiar cómo estas dos líneas (la estructura de la red social y la naturaleza del vínculo) se impactan mutuamente resultando en contextos diferentes para alcanzar el logro en los aprendizajes. Centrados en las necesidades del docente de conocer oportunamente las situaciones de riesgo de sus estudiantes, es nuestro objetivo diseñar una herramienta que le ofrezca oportunidades de fortalecimiento de su estrategia pedagógica basado en las interacciones en las redes sociales de sus estudiantes. Para lograr este objetivo estamos trabajando en el proyecto DIIA ¹ de forma interdisciplinaria con docentes, psicólogos e informáticos. Considerando al estudiante como el centro de los procesos educativos, una etapa inicial significativa es realizar el modelado del estudiante, o sea realizar la selección de las características apropiadas de los estudiantes que deben ser consideradas y representadas. En este trabajo presentamos la especificación de las propiedades del estudiante que resultan de interés para una aplicación de análisis de las interacciones según los sentimientos que en ellas se expresan y una propuesta inicial de captura de estos datos. Algunos trabajos relacionados a esta propuesta son el de Wen, Yang, and Rosé [24] que realizan un estudio de análisis de sentimientos en discusiones en línea de estudiantes de un MOOC y encuentran una fuerte correlación entre la expresión de sentimientos negativos con la probabilidad de desvinculación del estudiante del curso. Resultados similares son presentados por Wen *et al.* [25] quienes muestran que el uso de palabras con connotación de sentimientos positivos junto con el uso de pronombres en primera persona pueden servir como medida del nivel de motivación del estudiante en el curso. Nuestra propuesta extiende estos trabajos agregando la característica de sentimientos en el contexto social en línea del estudiante para estudiar de qué forma los sentimientos en las interacciones de los estudiantes son influenciadas por los

¹DIIA: Descubrimiento de Interacciones que impactan en el Aprendizaje, ANII FSED_2_2016_1_130712

sentimientos prevalentes en el grupo.

En la Sección II presentamos en qué consiste el modelado del estudiante y algunos trabajos relacionados con particular atención a las dimensiones de interacciones y de contexto del estudiante. Presentamos respuesta a la pregunta de cuáles son las características que entendemos relevantes para el contexto social en línea del estudiante considerando los sentimientos en sus interacciones. En la Sección III presentamos el modelado de las interacciones en el grafo social, el cálculo de las métricas sociales así como el análisis de sentimientos. En la Sección IV desarrollamos nuestra propuesta de análisis de sentimientos dentro del proyecto DIIA y la interpretación de su impacto en el aprendizaje a través de un componente de visualización. Finalmente, el estado actual y perspectivas del crecimiento de la plataforma se presentan en la Sección V.

II. MODELADO DEL ESTUDIANTE

El estudiante, como centro de los procesos educativos, es la pieza fundamental sobre la cual es necesario definir cuáles son sus propiedades relevantes, definir en qué forma se obtendrán esas propiedades y definir la estructura de almacenamiento y administración de los datos que representan esas propiedades. A este proceso se le llama *modelado del estudiante*. El Modelo de Estudiante debe ser el más adecuado para alcanzar el objetivo del sistema a realizar (por ej. recomendar, adaptar, personalizar, comparar, predecir, etc.), a partir del cual se puedan aplicar las técnicas de análisis de datos y construir la dimensión buscada en el perfil del estudiante.

De acuerdo a Devedžić [1] los desafíos del modelado se organizan en tres etapas: (i) Definir qué características del usuario se consideran importantes para ser incluidas en el modelo; (ii) cómo representarlas; y (iii) cómo extraerlas y mantenerlas a través del tiempo. En el trabajo de Anter *et al.* [4] se identifican de forma genérica seis dimensiones que agrupan las características que se pueden obtener al realizar modelado de usuario de un sistema web, ellas son: Personal (datos demográficos), Intereses (datos sobre su dominio de conocimiento), Expectativas (objetivos a alcanzar, niveles de calidad de datos), Preferencias (volumen de datos, forma de presentación), Seguridad (niveles de seguridad que utiliza) y el Histórico de las Interacciones (comportamiento del usuario, trazas de navegación). En el ámbito educativo se han desarrollado muchos trabajos que focalizan el modelado del usuario estudiante, en el trabajo de Chrysafiadi & Virvou [7] se presenta un estado del arte en el modelado del estudiante con el objetivo de personalización. Al igual que [4] identifican la dimensión Personal, la de Preferencias situada para el estilo de aprendizaje (si prefiere estudiar en grupo o sólo, si prefiere material visual o textual, etc) y la dimensión Cognición (que coincide con la de Intereses de [4] que refleja los datos de su conocimiento pero distinguiendo entre su conocimiento correcto y el erróneo). Chrysafiadi & Virvou [7] agregan las dimensiones de Meta-cognición (importante para la administración del aprendizaje, esta dimensión incluye habilidades de reflexión, monitoreo, autoexplicación, autoevaluación y

autogestión) y la del Estado Emocional (donde listan las emociones de: feliz, triste, enojado, interesado, frustrado, aburrido, distraído, enfocado, confundido y argumentan que algunas de estas emociones, como el aburrimiento o frustración, son proclives a llevar a los estudiantes a un comportamiento fuera de las tareas del curso).

Algunos autores han marcado también la necesidad de considerar el contexto del estudiante para evaluar su actividad en los sistemas de educación en línea, entre ellos [3], [12], [18], [19].

Dependiendo de la aplicación, algunas dimensiones se vuelven más importantes que otras y necesitan una caracterización más detallada. Considerando la dimensión Interacción en el trabajo de Paiva *et al.* [21] se modelan las interacciones considerando cuatro categorías: (1) Colaborativas (aquellas que ayudan a otros estudiantes, como compartir materiales, etc.), (2) de Gamificación (las que sirven para incorporar elementos de un juego (puntos, rankings, etc.), (3) Individuales (interacción en solitario con el material, como ver un video, realizar un ejercicio) y (4) Sociales (las que promueven la participación social, como chats, foros, etc.). El trabajo de Crossley *et al.* [8] analiza las interacciones usando cohesión en la red social para predecir qué estudiantes van a terminar con éxito cursos en un ambiente de MOOC.

Inspirados de todas estas propuestas de modelado, en el proyecto DIIA se busca alcanzar un balance entre la simplicidad en la interpretación y la riqueza semántica de los modelos complejos ofreciendo a los docentes, un modelo eficiente, con el nivel de abstracción adecuado para representar, analizar e interpretar las interacciones de los estudiantes en las redes sociales.

El proyecto DIIA (Descubrimiento de Patrones de Interacción que Impactan al Aprendizaje) ofrece a los docentes una plataforma DIIA orientada a aportarle el conocimiento (información estratégica) sobre el desempeño de sus estudiantes y el impacto de sus interacciones en las redes sociales formales e informales. De esta manera, el docente puede detectar situaciones de riesgo y tomar decisiones oportunas que favorezcan su aprendizaje.

La Figura 1 muestra de forma gráfica las características del modelo del estudiante que son relevantes para los objetivos de DIIA. En el centro se encuentra el estudiante con sus características Personales demográficas, nos interesa especialmente modelar las Interacciones que ocurren con otras personas del sistema (estudiantes, docentes) y también con los recursos disponibles del sistema, las Preferencias que el estudiante tiene respecto al aprendizaje (si prefiere trabajar solo o en grupo, si prefiere videos o textos, etc.) y el rendimiento Académico alcanzado (sus conocimientos y sus calificaciones en evaluaciones).

Todas estas dimensiones se encuentran influenciadas por la dimensión **Contexto** que rodea al estudiante. El contexto influencia al estudiante y alcanza todas sus dimensiones, así las **Preferencias** pueden variar por ejemplo según la característica **tiempo** del contexto (prefiere leer en la mañana, pero ver videos en la noche) o la característica **lugar** del contexto (no

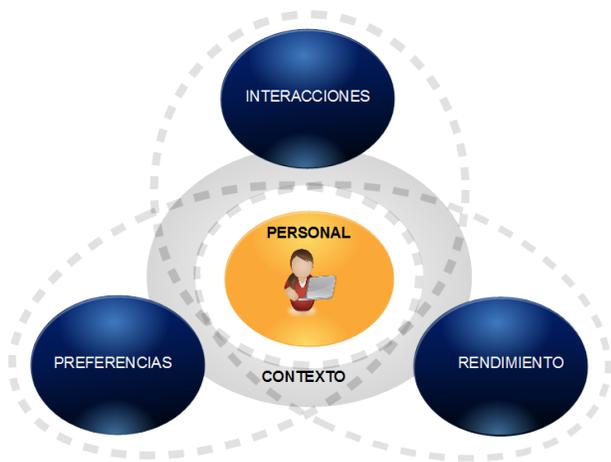


Figura 1. Modelo estudiante.

lee en un transporte pero sí atiende audios mientras se desplaza por la ciudad), o la dimensión **Rendimiento Académico** puede variar según las características de los dispositivos de su contexto (conectividad, hardware, software). Para las **Interacciones** personales observamos que estas se dan en un contexto social donde una característica importante es la polaridad del sentimiento (positivo, neutro o negativo) del conjunto de las interacciones del grupo social. Nuestro enfoque se centra en el análisis de las interacciones que integran el contexto social en línea del estudiante, ya que el peso de los sentimientos negativos o positivos en las interacciones estudiante-estudiante deben ser considerado en relación al sentimiento predominante en el grupo que constituye el contexto, donde éstas interacciones ocurren. Todas estas dimensiones tienen a la vez, puntos de contacto en sus influencias recíprocas y múltiples, lo que convierte al Modelo del Estudiante en una representación con gran riqueza semántica. Nuestro interés es proponer una visualización en la que el profesor pueda interpretar de manera sencilla las influencias positivas o negativas descubiertas en el análisis de las interacciones sociales.

En la próxima sección describimos en detalle las interacciones representadas en el grafo social de la red de estudiantes de un curso.

III. ANÁLISIS DE LA RED SOCIAL

El elemento central de análisis de la red social son las interacciones que en ella ocurren. El análisis de las interacciones puede tener enfoques puramente cuantitativos, o enfoques de análisis del contenido que arrojan información cualitativa, o enfoques híbridos donde se combinan los dos anteriores. F. Henri [17] presenta una investigación enfocada en el análisis de los contenidos de los intercambios en un entorno de aprendizaje colaborativo en línea. Parte del hecho que un educador que pueda leer entre líneas en los contenidos que se generan en un entorno de este tipo encontrará información no disponible en cualquier otra situación de aprendizaje. El análisis que realiza el educador debe ir más allá de la comprensión de

lo que se dice sobre un tema en sí, debe comprender los procesos sociales y cognitivos de los estudiantes. En este sentido, plantea que el análisis de contenido, cuando se realiza con el objetivo de comprender el proceso de aprendizaje, proporciona información sobre los alumnos como aprendices y sobre sus formas de tratar un tema determinado. Con esta información, el educador estará en condiciones de cumplir su función principal, que es ofrecer apoyo inmediato al individuo y al proceso de aprendizaje colectivo. Podrá analizar la razón de la eficiencia de los entornos colaborativos de aprendizaje y proponer estrategias que a partir de los patrones detectados, estimulen el aprendizaje de sus estudiantes.

La estructura interactiva del grupo lleva a los usuarios a tratar el aprendizaje en un modo cooperativo, que le da valor al conocimiento colectivo. El contenido a ser aprendido se construye colectivamente, con el intercambio e interacciones. Se plantea entonces, la necesidad de desarrollar una metodología específica para el análisis de mensajes que sea útil para los educadores que desean obtener una mejor comprensión de los intercambios mediados por computadora. La solución que propone es establecer un método analítico apropiado y desarrollar herramientas para comprender mejor los mensajes e identificar los elementos útiles de su significado. Un método analítico apropiado identificaría los procesos y estrategias de aprendizaje seleccionados o desarrollados por los estudiantes. Los resultados de este análisis constituyen una base para el desarrollo de un marco para orientar las intervenciones y apoyar el proceso de aprendizaje.

En la propuesta de Henri [17], luego implementado por Gunawardena *et al.* [15] se desarrolla un modelo denominado Modelo de Análisis de Interacción (IAM por sus siglas en inglés). El modelo se basa en la teoría constructivista social de Vygotsky. Describe cinco fases de la construcción del conocimiento: compartir y comparar constituyen la Fase I; la disonancia la Fase II; la negociación y la co-construcción la Fase III, las construcciones tentativas de prueba la Fase IV, y las declaraciones y la aplicación del conocimiento recién construido la Fase V. Incluyen numerosas operaciones para cada una de sus cinco fases. El modelo en sí sirve como un marco que define la construcción social del conocimiento como una función de la interacción.

En el trabajo de Flor *et al.* [14] se retoma la definición de IAM, planteando que hasta el momento, es el modelo más usado para el análisis de contenidos de interacciones. La propuesta incorpora al modelo herramientas de analítica de aprendizaje y análisis de redes sociales. Se plantea que al combinar el análisis de interacción con el análisis de aprendizaje y el análisis de redes sociales han logrado conceptualizar el proceso mediante el cual tiene lugar la construcción del conocimiento en plataformas en línea.

El modelo IAM propone un marco de cinco dimensiones para el análisis de contenido donde además de las dimensiones cognitiva y metacognitiva, destacan las dimensiones participativa, interactiva y social. A continuación presentamos estas últimas tres dimensiones por ser el foco de nuestro proyecto de estudio y relacionamos su interpretación con el análisis de

las métricas en redes sociales.

La **dimensión participativa** se refiere a la cantidad de mensajes o publicaciones de un estudiante. Esta dimensión es muy útil en combinación con herramientas del análisis de redes sociales y sus métricas asociadas. Las métricas de centralidad pueden proporcionar información sobre el rol de un alumno en un curso, sobre su integración, incluso puede poner en evidencia situaciones de riesgo con respecto al curso.

La **dimensión interactiva** utiliza una definición de interacción en tres pasos: Paso 1: comunicación de información. Paso 2: una primera respuesta a esta información. Paso 3: una segunda respuesta relacionada con la primera. Esta es una forma descriptiva de entender la interacción y que facilita su detección en la totalidad de los intercambios. Gunawardena *et al.* [15] plantean que la interacción debe verse y analizarse en la totalidad, la definen como el proceso a través del cual se produce la negociación del significado y la co-creación del conocimiento. De estas definiciones se desprende que no necesariamente un conjunto de mensajes implica interacción. La interacción puede ser explícita si se menciona explícitamente otro mensaje, idea o persona o implícita si se refiere obviamente a uno o más mensajes o ideas, pero no específicamente menciona la conexión. Esta dimensión puede analizarse desde múltiples perspectivas y aportar información tanto individual como grupal. Mediante aplicación de análisis de redes sociales y sus métricas a los datos de interacciones, se puede obtener información sobre un conjunto de características de los estudiantes y de los grupos de estudiantes que la mera cuantificación de intervenciones no proporcionaría. Para valorar esta dimensión es necesario incorporar análisis de contenidos. Aplicando técnicas de análisis de sentimientos se podría deducir si una intervención en una interacción expresa acuerdo o desacuerdo. En Gunawardena *et al.* [15] basándose en las propuestas de análisis de Henri [17] proponen diversas formas de utilizar la dimensión interactiva: la identificación de interacciones permite mostrar la interrelación de los mensajes en forma gráfica, también proponen realizar análisis de términos más usados en interacciones. Se pueden identificar hilos de mensajes en una conversación, identificar mensajes que son particularmente “influyentes” por la producción de muchas respuestas o secuencias largas de respuestas anidadas. Este último caso podría identificarse como mensaje “polémico”.

La **dimensión social** se refiere a la dinámica social de los intercambios. Resalta la importancia de este aspecto de la comunicación para la participación, la cohesión social dentro del grupo y el sentimiento de pertenencia. Se puede detectar del análisis de contenidos ya sea por presencia de palabras clave o asumiendo que la presencia de la interacción social está en cualquier mensaje que no esté relacionada con el contenido formal de la materia. Un alto nivel de mensajes con contenido social puede indicar falta de motivación, dispersión, falta de concentración, tanto en referencia a un alumno a al grupo.

Representación como grafo Las métricas de análisis de grafo usadas para el análisis de interacciones representan a la red social como un grafo donde los actores de la red (estudiantes, docentes, recursos educativos) son los nodos del

grafo y las aristas representan la relación entre los actores que puede ser fuerte o débil dependiendo de su frecuencia, importancia o tipo. El enfoque centrado en el estudiante identifica a las relaciones o vínculos que inciden o salen del estudiante, mientras que el análisis de la red global nos proporciona una idea de los intereses y las prácticas de las personas, identificando las características que mantienen integrada la red social. Este método también tiene el potencial de ayudar a identificar grupos dentro de la red, que pueden apoyar procesos de aprendizaje, como comunidades y grupos de afinidad.

Formalmente el grafo se representa por $G = (V, E, \alpha)$ donde:

$V = \{v_i \mid i = 1, 2, \dots, n\}$ es el conjunto finito de vértices del grafo que consiste de entidades en un curso: *estudiantes, docente, recurso educativo*

$E \subseteq V \times V$ es el conjunto finito de aristas que representan una de las siguientes interacciones: *publicación, reacción, comentario*

$\alpha : \rightarrow L$ es una función que asigna una etiqueta L a un par de vértices asociados por medio de una arista.

Para medir la importancia de un nodo en el grafo se le asocian medidas de centralidad.

III-A. Métricas de Centralidad

Las métricas de centralidad de grafos [11] se refieren a una familia de medidas estructurales calculadas para encontrar la posición o importancia de un nodo o vértice dentro de un grafo dado. Estas métricas han sido usadas de manera exitosa para descubrir la relevancia de vértices en distintos tópicos como el Análisis de Redes Sociales [22], los Sistemas de Recuperación de Información [5] o las herramientas de Procesamiento del Lenguaje Natural [2] entre otros. A continuación, se describen las métricas de centralidad empleadas en el proyecto DIIA:

Centralidad de grado (degree centrality): Es definida como el número de aristas incidentes a un vértice dentro de un grafo dado. Existen dos subversiones de este tipo de centralidad, la centralidad de grado de entrada (in-degree), la cual se refiere al número de aristas que apuntan a un vértice dado, y la centralidad de grado de salida (out-degree), en la cual se mide el número de aristas que apuntan a otros vértices en el grafo desde un vértice dado.

Centralidad de cercanía (closeness centrality): Está definida como el promedio de la suma de los caminos más cortos desde un vértice dado a todos los demás en el grafo.

Centralidad de intermediación (betweenness centrality): Es una medida del grado en que un vértice está en el camino más corto entre otros dos vértices en el grafo.

Centralidad de vector propio (eigenvector centrality): Esta es una medida de la importancia o influencia de cada vértice en el grafo. La suposición de esta métrica es que cada medida de centralidad sobre un vértice es la suma de las medidas de centralidad de los vértices que están conectados a éste. Para calcular esta métrica se debe obtener antes una matriz clásica de adyacencia al grafo.

Métrica (alto valor)	Tipo de interacción	Interpretación
Grado de entrada	Estudiante-Estudiante	Estudiante con mucha importancia en la red y con soporte entre sus pares.
	Estudiante-Docente	Estudiante o docente prestigioso, tiene mucha importancia dentro de la red.
Grado de salida	Estudiante-Estudiante	Estudiante con mucha influencia entre sus pares.
	Estudiante-Docente	Estudiante o docente muy sociable, altamente comunicativo, influyente.
Proximidad	Estudiante-Estudiante	Estudiante con capacidad de comunicar rápidamente a sus pares. Puede ser influenciado o influenciar a los demás.
	Estudiante-Docente	Estudiante o docente con mucho alcance en la difusión de mensajes.
Excentricidad	Estudiante-Estudiante	Estudiante con un amplio rol dentro de la red. Es fácilmente accesible por todos, puede ser influenciado o influenciar a los demás.
	Estudiante-Docente	Estudiante o docente con una amplia cobertura de la red. Es fácilmente accesible por todos los integrantes del grupo. Idealmente el docente debería tener este valor.
Vector propio	Estudiante-Estudiante	Estudiante estratégico para difundir información entre sus pares.
	Estudiante-Docente	Estudiante o docente influyente, importante para difundir información.
Interconexión	Estudiante-Estudiante	Estudiante estratégico para comunicar entre subgrupos de la red de pares.
	Estudiante-Docente	Estudiante o docente estratégico para comunicar entre subgrupos.

Figura 2. Interpretación de las métricas de las interacciones.

Excentricidad (eccentricity): La excentricidad no es una medida de centralidad, pero sirve para detectar la distancia máxima desde un vértice hacia cualquier otro vértice en un grafo siguiendo los caminos más cortos presentes en la estructura.

Las fórmulas que utilizamos para calcular estas métricas pueden encontrarse en [26]. La interpretación de las métricas se describe en la Figura 2. Coincidimos con las interpretaciones de [8] donde un grado de entrada con valor alto en la interacción estudiante-estudiante es un indicador de la frecuencia con la que se elige un alumno como compañero de aprendizaje y de su popularidad en la clase. También para [8] un valor alto en el grado de salida de la interacción estudiante-estudiante es un indicador de la frecuencia con la que los estudiantes solicitan ayuda a sus compañeros de aprendizaje social y de la apertura del alumno.

Como ejemplo del grafo anteriormente descrito, en la Figura 3 puede observarse un ejemplo del modelado de un curso, tomando en cuenta los distintos tipos de vértices y aristas propuestos. Considerando las propiedades del grafo propuesto, se puede apreciar que existen distintos tipos de vértices que representan a las instancias más comunes en un curso. Dependiendo del valor de un vértice con respecto a una medida de centralidad, el tamaño de éste cambiará (entre mayor la centralidad, más grande el vértice). Las aristas presentes en el grafo pueden identificar las distintas interacciones, como un mensaje entre estudiantes o una reacción de un estudiante a un recurso educativo propuesto para el curso. El grosor de estas aristas estará determinado por el número de interacciones del mismo tipo, existentes entre los nodos. La dirección de las aristas en el grafo representa el orden de la relación entre las entidades, por ejemplo, la arista: *Lupita* → *Ejercicio* representa el acceso de la estudiante *Lupita* al contenido del recurso educativo *Ejercicio*. La etiqueta asociada a cada arista describe el tipo de dicha interacción entre dos entidades en el grafo (lectura, comentario, reacción, publicación, etc).

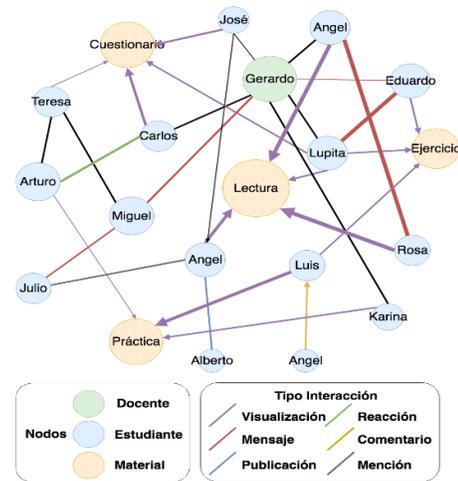


Figura 3. Ejemplo de grafo de la red social.

III-B. Análisis de Sentimientos

Además de estudiar los patrones topológicos presentes en las distintas interacciones entre los principales actores de un curso, otra importante herramienta para entender el comportamiento social de estudiantes y docentes es el análisis de sentimientos. Según Nasukawa & Yi [20] el análisis de sentimientos, también conocido como análisis de opiniones, se refiere al uso de técnicas de minería de texto y procesamiento de lenguaje natural para determinar la actitud de una persona sobre un tema, o determinar la polaridad (positiva, negativa o neutra) de un documento, como pueden ser de las publicaciones o comentarios compartidos en los cursos para determinar el estado de ánimo de los distintos actores.

Existen varias técnicas para realizar análisis de sentimientos, dentro del proyecto DIIA se propuso la creación de un clasificador de textos supervisado [16] que por medio de la extracción de características léxico-sintácticas como los trigramas de palabras (ventana de 3 elementos contiguos en el texto) con alta frecuencia de aparición, pueda diferenciar un sentimiento de otro. Es importante remarcar que dicho clasificador se basa en el entrenamiento de un modelo que busca las características presentes en un conjunto de datos de entrenamiento, para posteriormente buscar la existencia de los patrones en nuevos contenidos asociados a un curso. La Figura 4 ilustra un ejemplo de grafo construido para reflejar el sentimiento acumulado en las interacciones con contenido textual. El grosor de la arista muestra la cantidad de interacciones y el tono del color refleja el promedio de las polaridades del sentimiento contenido en los textos analizados.

IV. EXPERIENCIA DE MODELADO Y ANÁLISIS DEL CONTEXTO SOCIAL

La aplicación de métricas de análisis de datos en grafos para dar explicaciones ha aportado conocimiento útil para apoyar los procesos de enseñanza-aprendizaje. En este trabajo, se analizan e interpretan estas interacciones para los estudiantes y

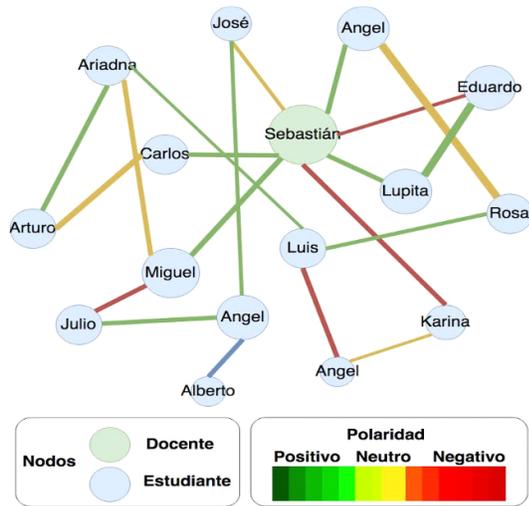


Figura 4. Ejemplo de grafo de la red social con sentimientos.

docentes que forman parte de un grupo de Facebook ilustrando la manera en que aportan a los docentes, conocimiento clave para la toma de decisiones pedagógicas que beneficien al estudiante. Al análisis de métricas se le incorpora análisis de sentimientos para obtener datos de contexto de la interacción social.

En el marco del proyecto DIIA se analizaron las interacciones sociales que ocurren dentro de un grupo de Facebook, donde se generaron datos de prueba, creados para ilustrar el conocimiento útil que el docente puede visualizar y usar para tomar decisiones que impacten favorablemente el aprendizaje de los estudiantes. Los datos generados permiten proponer la detección de posibles situaciones de riesgo o de uso positivo de la influencia de los diferentes miembros de la red social integrada por las personas que interactúan simultáneamente en un salón de clases y en un espacio virtual como los grupos de Facebook.

Las métricas de centralidad utilizadas son las presentadas en la Sección III. Para el análisis de sentimientos se utiliza la herramienta Meaning Cloud ². La presentación de resultados se realiza mediante la herramienta de análisis y visualización de grafos Gephi ³.

Las interacciones se modelan como un grafo en el que los nodos representan a los alumnos y al docente. Las aristas tienen origen en el nodo que representa a la persona que inicia la interacción, y destino en la persona destino de la interacción. El peso que se le asigna a cada arista corresponde con la cantidad de interacciones de origen a destino. En la Figura 5 se muestra el grafo de interacciones donde el grosor de la arista depende de su peso y el tamaño del nodo está configurado según la métrica de centralidad de vector propio.

En la tabla de la Figura 2 se presenta una propuesta de interpretación de las métricas visualizadas en la Figura 4 de acuerdo a la definición utilizada de la Sección II. La

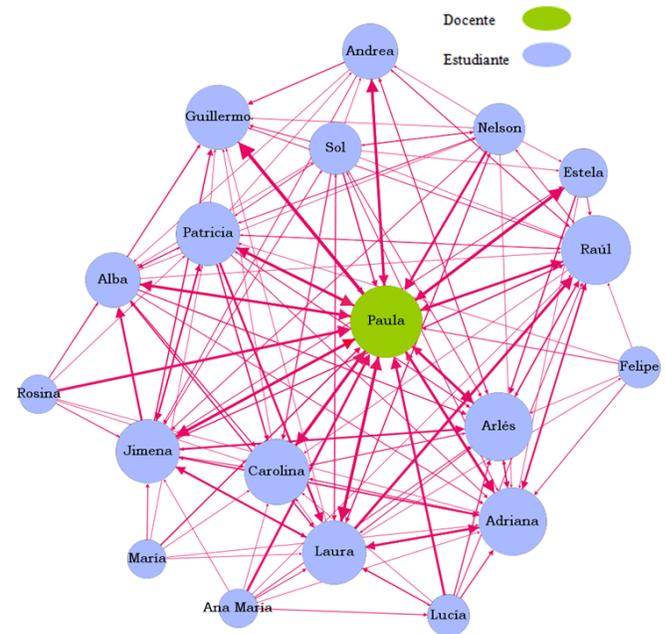


Figura 5. Ejemplo de grafo de interacciones usando Gephi.

propuesta pretende proporcionar información cualitativa a las mediciones. A cada nivel de la interpretación presentada en la tabla de la Figura 2 se le puede proveer de contenido semántico. Por ejemplo, si un estudiante presenta un nivel bajo del grado de entrada en las interacciones estudiante – docente se podrían establecer diferentes conclusiones: puede tratarse de un estudiante independiente y con muy buena escolaridad, o puede ser un estudiante desinteresado por el curso con riesgo de abandonar. Si él tiene valores bajos en el grado de entrada, pero en las interacciones estudiante – estudiante puede tratarse de un estudiante con muy buena escolaridad, pero sin interés en colaborar con sus pares o puede deberse a un caso de baja popularidad y aislamiento, o ambas. El análisis de sentimiento con el que se complementa este estudio puede aportar información a este tipo de situaciones.

Para cada una de las interacciones que involucran envío de texto se clasifica el texto según su polaridad. En la Figura 6 se muestra el grafo de sentimientos. Los nodos correspondientes a estudiantes se representan con un color y el docente con otro. Las aristas se colorean según el peso, de forma que a la de menor peso se le hace corresponder el color de la polaridad N+ y se matiza hacia el color de P+ a medida que aumenta el peso. El tamaño de los nodos se corresponde con su grado de salida, de forma que el nodo de quien ha interactuado (como origen) con más personas, tiene mayor tamaño y a medida que la cantidad de personas con la que ha interactuado un individuo decrece, el tamaño del nodo correspondiente decrece.

V. CONCLUSIONES Y TRABAJOS FUTUROS

En este trabajo hemos presentado el enfoque seguido en el proyecto DIIA para modelar el comportamiento social del estudiante y producir un diagnóstico contextualizado al aplicar

²MeaningCloud LLC : <https://www.meaningcloud.com/es/>

³Gephi : <https://gephi.org/>

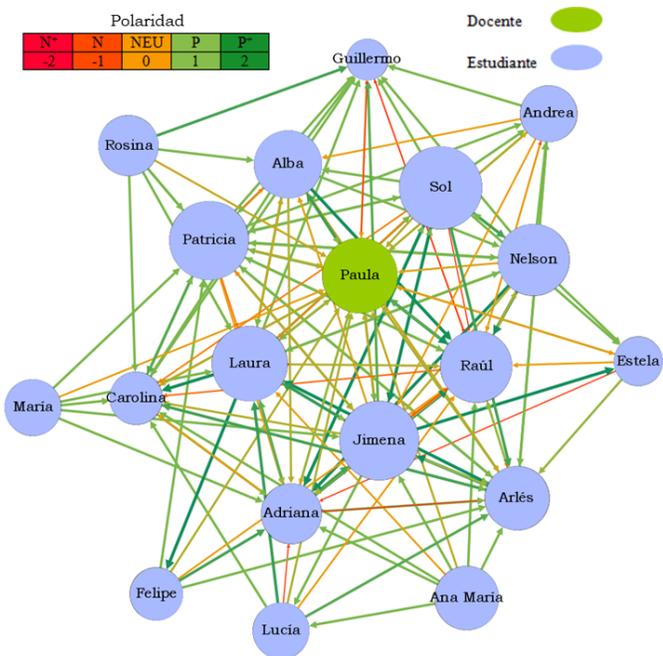


Figura 6. Ejemplo de grafo de sentimientos usando Gephi.

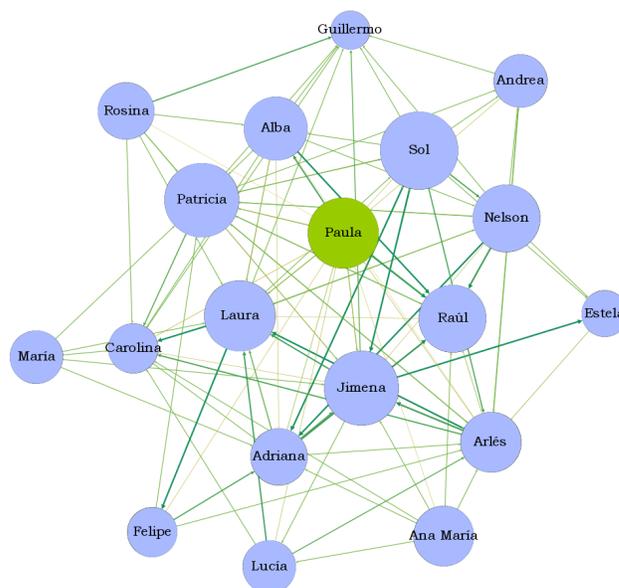


Figura 8. Ejemplo de grafo de Sentimientos Solo Positivos usando Gephi.

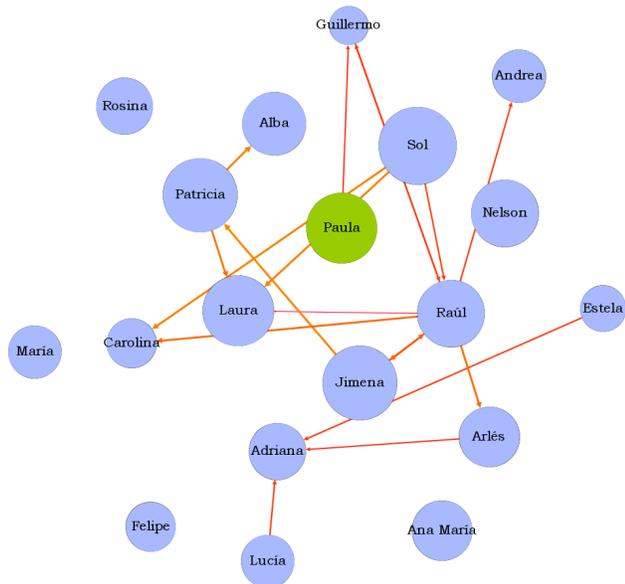


Figura 7. Ejemplo de grafo de Sentimientos Solo Negativos usando Gephi.

métricas sociales sobre los grafos construidos a partir de las interacciones existentes entre estudiante-estudiante, estudiante-docente e incluso estudiantes y los materiales propuestos por el docente para favorecer su aprendizaje.

Las diferentes métricas sociales así como el análisis de la polaridad de los sentimientos expresados en las interacciones de la red social permiten aportar al docente un conocimiento valioso sobre el perfil social del estudiante. Nuestra propuesta de visualización usando grafos que ilustran el impacto de las interacciones que tienen lugar en las redes sociales y explotan la polaridad del sentimiento, permite detectar patrones del contexto social que pueden impactar favorable o desfavorablemente el desempeño académico del estudiante.

Los ejemplos de visualización realizados en Gephi, sobre un conjunto de datos de una red social experimental, muestran el impacto del análisis de los sentimientos expresados en las interacciones, abriendo nuevas vías para trabajos futuros en los que se podrá explorar la detección automática de subgrupos cuya dominante de polaridad de sentimiento podría ser positiva, muy positiva, negativa, muy negativa o neutra, reflejando con ello el clima general del ánimo del grupo. La visualización propuesta es sencilla de interpretar por el docente aportándole un primer acercamiento de lo que sucede con sus estudiantes. El ambiente de análisis del aprendizaje provisto por la plataforma de DIIA está diseñado para ofrecerle herramientas complementarias que le permitirán que una vez detectada una situación de alerta y/o riesgo, pueda profundizar y conocer el detalle de las interacciones que la produjeron y detectar a los estudiantes a quienes tiene que acercarse para brindarles el apoyo pedagógico que necesitan.

VI. AGRADECIMIENTOS

La investigación que da origen a los resultados presentados en la presente publicación recibió fondos de la Agencia Nacional de Investigación e Innovación bajo el código FSED-2-2016-1-130712.

REFERENCIAS

- [1] V. A proposed paradigm for smart learning environment based on semantic web. *Semantic web and education*, volume 12. Springer Science & Business Media, 2006.
- [2] G. Algaphari, F. M. Ba-Alwi, and A. Moharram. Text summarization using centrality concept. *International Journal of Computer Applications*, 79(1), 2013.
- [3] M. A. Alimam, H. Seghioer, and Y. Elyusufi. Building profiles based on ontology for career recommendation in e-learning context. In *multimedia computing and systems (ICMCS), 2014 international conference on*, pages 558–562. IEEE, 2014.
- [4] S. Anter, M. H. El Yazidi, A. Zellou, and A. Idri. Towards a generic model of a user profile. In *Intelligent Systems: Theories and Applications (SITA), 2016 11th International Conference on*, pages 1–6. IEEE, 2016.
- [5] M. R. Bouadjenek, H. Hacid, and M. Bouzeghoub. Social networks and information retrieval, how are they converging? a survey, a taxonomy and an analysis of social information retrieval approaches and platforms. *Information Systems*, 56:1–18, 2016.
- [6] G. P. Caldeiro and C. Martinez. El aprendizaje en red y el trabajo colaborativo en entornos mediados por tecnología. *Virtualidad, Educación y Ciencia*, 5(9):102–103, 2013.
- [7] K. Chrysaftadi and M. Virvou. Student modeling for personalized education: A review of the literature. In *Advances in Personalized Web-Based Education*, pages 1–24. Springer, 2015.
- [8] S. Crossley, M. Dascalu, D. S. McNamara, R. Baker, and S. Trausan-Matu. Predicting success in massive open online courses (moocs) using cohesion network analysis. Philadelphia, PA: International Society of the Learning Sciences., 2017.
- [9] S. Downes. Connectivism and connective knowledge. *Essays on meaning and learning networks*, pages 493–557, 2012.
- [10] R. Ferguson and S. B. Shum. Social learning analytics: five approaches. In *Proceedings of the 2nd international conference on learning analytics and knowledge*, pages 23–33. ACM, 2012.
- [11] L. C. Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1978.
- [12] S. Graf, K. B. Kinshuk, F. A. Khan, P. Maguire, A. Mahmoud, T. Rambharose, and Q. Zhang. Adaptivity and personalization in learning systems based on students’ characteristics and context. In *The 1st international symposium on smart learning environment*, pages 33–36, 2012.
- [13] A. Gruzd, D. Paulin, and C. Haythornthwaite. Analyzing social media and learning through content and social network analysis: A faceted methodological approach. *Journal of Learning Analytics*, 3(3):46–71, 2016.
- [14] C. N. Gunawardena, N. V. Flor, D. Gómez, and D. Sánchez. Analyzing social construction of knowledge online by employing interaction analysis, learning analytics, and social network analysis. *Quarterly Review of Distance Education*, 17(3):35, 2016.
- [15] C. N. Gunawardena, C. A. Lowe, and T. Anderson. Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of educational computing research*, 17(4):397–431, 1997.
- [16] P. Harrington. *Machine learning in action*, volume 5. Manning Greenwich, CT, 2012.
- [17] F. Henri. Computer conferencing and content analysis. In *Collaborative learning through computer conferencing*, pages 117–136. Springer, 1992.
- [18] J. Jovanović, D. Gašević, C. Knight, and G. Richards. Ontologies for effective use of context in e-learning settings. *Journal of Educational Technology & Society*, 10(3), 2007.
- [19] V. Luna, R. Quintero, M. Torres, M. Moreno-Ibarra, G. Guzmán, and I. Escamilla. An ontology-based approach for representing the interaction process between user profile and its context for collaborative learning environments. *Computers in Human Behavior*, 51:1387–1394, 2015.
- [20] T. Nasukawa and J. Yi. Sentiment analysis: Capturing favorability using natural language processing, 01 2003.
- [21] R. Paiva, I. I. Bittencourt, T. Tenório, P. Jaques, and S. Isotani. What do students do on-line? modeling students’ interactions to improve their learning experience. *Computers in Human Behavior*, 64:769 – 781, 2016.
- [22] F. Riquelme and P. González-Cantergiani. Measuring user influence on twitter: A survey. *Information Processing & Management*, 52(5):949–975, 2016.
- [23] G. Siemens. Connectivism: A learning theory for the digital age. *International journal of instructional technology and distance learning*, 2(1):3–10, 2005.
- [24] M. Wen, D. Yang, and C. Rose. Sentiment analysis in mooc discussion forums: What does it tell us? In *Educational data mining 2014*. Citeseer, 2014.
- [25] M. Wen, D. Yang, and C. P. Rosé. Linguistic reflections of student engagement in massive open online courses. In *ICWSM*, 2014.
- [26] x. xxx. Use of social metrics to discover interaction patterns that impact learning. *cubierto para mantener proceso de evaluación anónima*, 2(1):3–10, 2000.