

De "discusión táctica" en juego de colaboración a "conductas": enfoque de clasificación en etapas.

Franco D. Berdun, Francisco Serrano, Marcelo G. Armentano
ISISTAN Research Institute (CONICET/UNICEN), Campus Universitario - Paraje Arroyo Seco - Tandil, Argentina
{franco.berdun, marcelo.armentano}@isistan.unicen.edu.ar; fserrano@exa.unicen.edu.ar

Abstract The analysis of group dynamics is extremely useful for understanding and predicting the performance of teamwork's, since in this context, collaboration problems can naturally arise. Artificial intelligence, and specially machine learning techniques, enables automating the observation process and the analysis of groups of users who use an online collaborative platform. Among the online collaborative platforms available, games are an attractive alternative for all audiences that enable capturing the players' behavior by observing their social interactions, while engaging them in a pleasant activity. In this paper, we present experimental results of classifying observed conversations in an online game to collaborative behaviors, guided by the Interaction Process Analysis, a theory for categorizing social interactions. The proposed automation of the classification process can be used to assist teachers or team leaders to detect alterations in the balance of group reactions and to improve their performance by indicating actions to improve the balance.

Resumen El análisis de la dinámica de grupos es extremadamente útil para entender y predecir el desempeño de equipos de trabajos, puesto que en este contexto, pueden surgir naturalmente problemas de colaboración. La inteligencia artificial, y especialmente las técnicas de aprendizaje de máquina, permite automatizar el proceso de observación y análisis de grupos de usuarios que utilizan una plataforma colaborativa en línea. Entre las diversas plataformas colaborativas en línea, los juegos son una alternativa apta para todo público y permiten capturar el comportamiento de los jugadores mediante la observación de sus interacciones sociales, al mismo tiempo que los involucra en una actividad agradable. En este trabajo se presentan resultados experimentales de clasificación de conversaciones observadas en un juego en línea a conductas de colaboración, guiados por una teoría de categorización de interacciones sociales específica. La automatización propuesta se puede utilizar para asistir a profesores o líderes de equipo a detectar alteraciones en el balance de las reacciones grupales y permitirá mejorar el desempeño al indicar acciones para mejorar el equilibrio.

Keywords: automatic classification, group dynamic, gamification, user modeling.

Palabras clave: clasificación automática, dinámica de grupo, gamificación, modelado de usuario.

1 Introducción

La dinámica de grupo se define como el proceso de interacción en un grupo para resolver una determinada tarea [1]. Un enfoque para estudiar la dinámica de grupo se basa en el análisis de las interacciones entre los miembros de un equipo para extraer información útil sobre el comportamiento de los mismos. Por ejemplo, "SYstem Multi-Level Observation Group" (SYMLOG) [2]. Otra teoría ampliamente utilizada es el "Interaction Process Analysis" (IPA) [3], la cual propone una categorización de las interacciones manifestadas en discusiones colaborativas. Al mismo tiempo, el éxito de un trabajo en equipo requiere de un cierto equilibrio de diferentes características por parte de los miembros del grupo, de lo contrario, el desbalance de estas características tiene un impacto negativo en el grupo y dificultan el buen desempeño y la obtención de logros.

Este equilibrio es más difícil de conseguir en trabajos colaborativos mediados por computadora. Hoy en día, en muchos entornos (como empresas, escuelas, universidades y gobiernos) las personas tienen que colaborar con pares ubicados en diferentes espacios físicos utilizando plataformas en línea. Por ejemplo, empresas de software donde existen equipos de empleados ubicados en diferentes regiones. En este escenario cualquier conocimiento previo del perfil de los empleados podría permitir una mejor organización de los grupos de desarrollo. Adicionalmente, la motivación, que es el motor de las organizaciones y uno de los principales obstáculos, puede ser atacado con la aplicación de la gamificación. El concepto de gamificación hace referencia al uso de los elementos de sistemas de juegos con objetivos y en contextos diferentes al del entretenimiento [4]. La gamificación mejora el desempeño de las personas al emplear la diversión como motivadora intrínseca que las estimula a tomar de manera dinámica y proactiva acciones que generalmente requieren un esfuerzo de la voluntad [5].

Muchas de las plataformas colaborativas disponibles proporcionan la información necesaria para capturar el comportamiento de los usuarios. Alternativamente, las plataformas de videojuegos colaborativos proponen un escenario más atractivo para el usuario por sus características recreativas. Las prestaciones provistas por estas plataformas posibilitan registrar grandes volúmenes de información referente a las interacciones entre los usuarios. Esto permite, la ejecución de un análisis tanto para detectar y caracterizar desequilibrios en la dinámica de los equipos, como para corregir y mejorar el proceso de trabajo, por medio de nuevos instrumentos de asistencia o mejorando los existentes. Adicionalmente, se han desarrollado asistentes inteligentes que, en el marco de una plataforma colaborativa y en base a una estrategia de trabajo [6], o a las interacciones del grupo [7], llevan a cabo la detección de conflictos en la dinámica grupal y alerta al supervisor de equipo para una rápida intervención, sugiriendo un plan de acciones correctivas. Por otro lado, el empleo de plataformas de videojuegos en las áreas de aprendizaje y análisis de características cuenta con numerosos precedentes [8], [9], [10], [11], [12], [13], [14].

Partiendo de la adaptación digital del juego “El Señor de los Anillos”¹ y un modelo de categorización de interacciones, se presentan resultados experimentales de clasificación en etapas de texto libre y datos del contexto, observados durante las sesiones de juego, a los patrones definidos por Bales [2], [3]. Los datos de las interacciones sociales, las conversaciones, serán capturados de la participación de un grupo de usuarios que usarán el juego como plataforma de colaboración, los datos del contexto serán inferidos en base a las variables del estado del juego (Fig. 1). En la adaptación digital del juego de mesa, los usuarios se deben comunicar por medio de un chat integrado, el cual registra los mensajes de las interacciones sociales que surgen en la dinámica del juego. Los resultados obtenidos servirán, en trabajos futuros, para la materialización de un instrumento de modelado de usuarios y desarrollo de una asistencia inteligente personalizada. Esto último, posibilitará la mejora mediante sugerencias de acciones correctivas al equipo en caso de la existencia de alteraciones en la dinámica grupal.

2 Marco teórico y trabajos relacionados

2.1 Gamificación

Los juegos son ambientes estructurados con reglas claramente definidas, donde los jugadores tienen objetivos y desafíos claros, generalmente con la victoria como meta final. Los jugadores están motivados a participar, mejorando su desempeño, y pueden involucrarse en una experiencia simulada sin enfrentar riesgos [15]. El potencial que tiene el uso de juegos es ampliamente reconocido y su empleo en ambientes no lúdicos han derivado en lo que actualmente se conoce como gamificación. Es decir, el uso de elementos de los sistemas de juegos con objetivos y en contextos diferentes al del entretenimiento [3], por ejemplo: salud, finanzas, gobierno, educación [16], [17], [18].

¹ <https://goo.gl/tDwWl2>

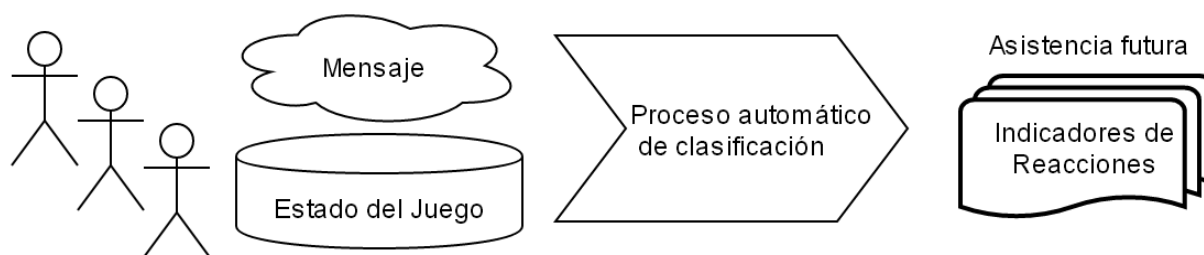


Figura 1. Esquema general

Particularmente, existen varios estudios sobre el uso de los videojuegos en el contexto de habilidades sociales. Linehan et al. [19], por ejemplo, proponen una serie de fundamentos para la mecánica de un juego para el entrenamiento de "soft skills", cuyas características lo hacen óptimo para estimular la participación colaborativa y la comunicación entre usuarios. También describen qué partes específicas de tales mecánicas desempeñan un papel especialmente importante, por ejemplo, la introducción de "fases de retroalimentación" durante los juegos, que permiten a un tutor detectar conflictos colaborativos y ayudar a resolverlos. Esta mecánica se utilizó en el desarrollo de DREADED, un juego que propone un método de enseñanza que permite a los usuarios entrenar sus habilidades sociales en un entorno virtual [20].

Los trabajos previamente mencionados basaron su investigación en enfoques ad-hoc para modelar usuarios, sin una teoría de base para estudiar la dinámica de los grupos. En nuestro trabajo, basamos nuestra investigación en un modelo teórico bien conocido y ampliamente utilizado para la dinámica de grupos.

Un enfoque similar en el sentido de que utiliza juegos para la construcción de perfiles de usuario, es la investigación de Feldman et al. [9]. Dichos autores implementan un conjunto de juegos serios que permiten la construcción automática de perfiles preceptuales de usuarios. A partir de los perfiles aprendidos por el sistema, Feldman et al. [9] explican cómo los cursos pueden adaptarse agrupando estudiantes con un tipo de percepción similar en busca de un mejor desempeño. Mediante el diseño de una red Bayesiana, Feldman et al. [9] interpretan la información sobre el desempeño de un usuario aislado observada en diferentes sesiones de juego. De manera diferente, nuestro trabajo busca detectar los patrones de reacciones de las interacciones de usuarios que interactúan en un grupo que participa en un juego colaborativo y no a un usuario actuando de manera aislada.

En cuanto a juegos colaborativos, Zagal et al. [21] analiza un juego de mesa de colaboración e identifica lecciones y dificultades en la dinámica para la creación de nuevos juegos. En el juego elegido, todos los jugadores deben cooperar activamente para ganar, mediante el logro de un objetivo común. No hay ningún usuario individual que gane o pierda en este juego: el grupo entero gana o pierde según el funcionamiento del equipo. Dado que un buen rendimiento es extremadamente dependiente de una buena comunicación y cooperación entre los jugadores, creemos que este juego de mesa es una opción apropiada para el estudio de la dinámica de equipo en una plataforma digital de colaboración, el cual se ha adaptado a una versión digital [22].

2.2 Análisis de interacciones

Como se mencionó previamente, el buen desempeño en un juego colaborativo es extremadamente dependiente de la buena comunicación, y hace al contenido verbal uno de los factores más importantes en el análisis. Entonces, para clasificar las interacciones sociales, se requiere un método que pueda describir varios patrones de comportamiento. Para este propósito, Bales desarrolló dos sistemas: SYMLOG [2] e IPA [3]. El primero distingue tres dimensiones estructurales en interacciones grupales: estado, atracción y orientación de objetivos. Es decir, analiza la actitud dominante (U) o sumisa (D) de quienes interactúan, la tendencia positiva (P) o negativa (N) y finalmente estudia la cuestión de si las personas están involucradas con la tarea (F) o con comportamientos socio-emocionales (B). Cada dimensión también contempla una posible conducta neutral, logrando de esta forma 27 combinaciones (Tabla 1). Particularmente, la última dimensión (orientación de objetivos), fue presentada en el trabajo previo de Bales: IPA. Es importante destacar que, a pesar de que SYMLOG fue desarrollado como una extensión del modelo IPA, estas dos teorías se complementan. En la Tabla 2 se muestra la relación entre ambas teorías: en la primera columna, las tres dimensiones SYMLOG; en la segunda columna, los atributos extremos para cada dimensión; y en la tercera y cuarta columnas, asociadas a los atributos extremos de la tercera dimensión de SYMLOG, se mencionan las diferentes conductas establecidas por Bales para clasificar las interacciones sociales basadas en mensajes en un grupo en colaboración. Bales divide las interacciones sociales en cuatro categorías principales: reacciones positivas, reacciones de respuesta, reacciones de requerimiento (o pregunta) y reacciones

negativas. De esta manera, las interacciones sociales pueden entenderse desde dos perspectivas: orientadas hacia la tarea y orientadas hacia lo socio-emocional, independientemente de los contenidos detallados de los mensajes. Aunque hay tres subcategorías en cada una de las cuatro categorías, se cree que las categorías del segundo nivel de IPA han proporcionado una categorización suficientemente clara para clasificar las interacciones sociales. Adicionalmente, analizar el primer nivel de categorías reducirá considerablemente la posibilidad de clasificar erróneamente las interacciones sociales.

El modelo IPA contribuye a la cuantificación de los distintos tipos de interacciones en una serie de etapas sucesivas típicas, que identificó Bales, por las que pasa cualquier grupo que desarrolla una tarea colaborativa. De esta forma, el desequilibrio generado por la manifestación inapropiada de los distintos tipos de interacciones en cada etapa genera alteraciones en la correcta dinámica del grupo. Para esto, Bales definió los umbrales entre los cuales una cantidad de manifestaciones de cada tipo de interacción pueden ser consideradas apropiadas.

Partiendo, entonces, de esta clasificación y de las prestaciones que la versión digitalizada adaptada del juego colaborativo provee para analizar la participación de cada usuario, puede llevarse a cabo el mapeo de estas interacciones a las categorías IPA para detectar el tipo de reacción de cada contribución. Aquí surge un nuevo desafío: partiendo de un conjunto de datos sobre la participación (interacciones sociales mediante la conversación textual), hay que vincularlos a las reacciones del modelo IPA. De ser un análisis “manual” (es decir, llevado a cabo por una persona), indudablemente esta tarea conlleva una carga de trabajo importante para los analistas y plantea importantes desafíos si se desea hacer de manera automática.

Tabla 1: Cuestionario de autopercepción SYMLOG.

Ítem		Consigna
1	U	Buscas éxito individual, prominencia personal y poder.
2	UP	Buscas popularidad y éxito social, ser querido y admirado.
3	UPF	Prefieres el trabajo en equipo activo dirigido a metas comunes, integridad organizacional.
4	UF	Buscas la eficiencia, administración imparcial estricta.
5	UNF	Buscas el refuerzo activo a la autoridad, repaso de reglas y regulaciones.
6	UN	Mantienes firmeza, obstinación y constancia para cumplir un objetivo, asertividad auto-orientada.
7	UNB	Eres vigoroso, individualista auto-orientado, pones resistencia a la autoridad.
8	UB	Buscas pasarla bien, renunciar a la tensión, un control relajante.
9	UPB	Buscas proteger a los miembros menos capaces, proveer ayuda cuando se necesita.
10	P	Buscas igualdad, participación democrática en la toma de decisión.
11	PF	Tienes un idealismo responsable, trabajo colaborativo.
12	F	Eres conservador, establecido, prefieres las formas “correctas” de hacer las cosas.
13	NF	Reprimes tus deseos individuales por las metas organizacionales.
14	N	Buscas la auto-protección, primero el interés personal, ser auto-suficiente.
15	NB	Rechazas los procedimientos establecidos, rechazas la conformidad.
16	B	Cambias a nuevos procedimientos, diferentes valores, creatividad.
17	PB	Buscas amistad, placer mutuo, recreación.
18	DP	Buscas confianza en la bondad de otros.
19	DPF	Trabajas con dedicación, confianza, eres leal al equipo
20	DF	Obedeces a la cadena de mando, eres complaciente con la autoridad.
21	DNF	Haces auto-sacrificio si es necesario para alcanzar las metas organizacionales.
22	DN	Tienes un rechazo pasivo a la popularidad.
23	DNB	Admites un error, dejas de esforzarte en defender trabajo mal hecho
24	DB	Tienes actitudes de no-cooperación pasiva con la autoridad.
25	DPB	Buscas satisfacción, te lo tomas con calma.
26	D	Renuncias a tus necesidades personales y deseos, pasividad.

Tabla 2: Modelo propuesto por Bales.

Dimensiones SYMLOG	Atributos extremos	IPA (2° nivel)	IPA (3° nivel)
Estado	Dominante (U)	-	-
	Sumiso (D)	-	-
Atracción	Amistoso (P)	-	-
	No amistoso (N)	-	-
Orientación de objetivo (IPA →)	Tarea (F) (IPA -1° nivel)	Requiere (o Pregunta)	Pide información
			Pide opinión
			Pide sugerencias
		Responde	Da sugerencia
			Da opiniones
			Da información
	Socio-emocional (B) (IPA -1° nivel)	Positiva	Muestra solidaridad
			Muestra relajamiento
		Negativa	Muestra acuerdo
			Muestra desacuerdo
		Muestra tensión	
		Muestra antagonismo	

3 Material y metodología

En esta Sección se explica el enfoque propuesto para la clasificación de la discusión táctica de un grupo de participantes en un juego colaborativo en línea. Primero, en la Sección 3.1, se mencionan las características más relevantes de la versión digital del juego empleado. Luego, en la Sección 3.2, se detalla el procedimiento de detección automatizada de las interacciones observadas.

3.1 Juego de colaboración

El desarrollo de la versión digital del juego de mesa “El Señor de los Anillos” respeta todos los requerimientos para trabajos de cooperación sugeridos por Johnson y Johnson [23]:

- interdependencia positiva, el juego debe requerir que los jugadores participen en equipo, es decir deben comprender rápidamente que una participación individualista impedirá el logro de los objetivos;
- responsabilidad individual, se debe incorporar un sistema de puntos para evaluar el desempeño personal de cada participante. Los puntajes deben poder ser visibles para todo el grupo;
- capacidades de interacción entre participantes, es decir, poder ayudarse e informarse por medio de un canal de comunicación. De esta forma los participantes que decidan colaborar incrementarán sus posibilidades de éxito;
- comprometer las habilidades sociales, este tipo de capacidades deben poder entrenarse y mejorarse si los participantes son observados y guiados por un tutor desde el inicio del uso del entorno;
- consenso grupal, es decir, el juego debe facilitar instrumentos para que los participantes estén motivados a discutir tácticamente su progreso, y mecanismos de consenso.

Adicionalmente, se conservaron todos los aspectos resaltados en las recomendaciones de Zagal [21]:

- Cada jugador cuenta con un marcador individual y gana puntos ayudando al grupo. Por ejemplo, los jugadores obtendrán puntos si colaboran en la decisión táctica para resolver una tarea, sin embargo, pueden optar por ser egoístas y solo recolectar recursos;
- Los jugadores pueden actuar y moverse libremente en el transcurso del juego, en la medida que se le permita. Ningún jugador es obligado a participar cumpliendo un determinado papel, sin embargo algunas acciones solo son posibles mediante el consenso de todos los jugadores;
- Los resultados de las decisiones que se toman son visibles para todos los jugadores. Por ejemplo, en el escenario concreto en que un jugador que puede impedir la ejecución de un determinado evento, o puede emplear el turno para recolectar recursos. Si este decide no evitar evento, el grupo puede verse obstaculizado para cumplir con el objetivo. Sin embargo el jugador egoísta recibirá recursos extras;
- Los jugadores son provistos con recursos heterogéneos con el fin de distribuir diferentes poderes y responsabilidades (las cartas especiales por ejemplo);
- La dinámica rotativa del juego evita que todas las responsabilidades recaigan sobre un solo jugador;

- Como todos los jugadores ganan o pierden como grupo, estos son motivados a jugar en forma colaborativa y participar en todas las decisiones tácticas.

La plataforma requiere que los jugadores se registren con un nombre de usuario y una contraseña, lo cual permite individualizar a los jugadores en el proceso experimental. Una vez que los usuarios ingresan, son dirigidos a una segunda pantalla donde pueden crear sesiones de juego o unirse a una sesión ya creada. Cuando un grupo de mínimamente dos jugadores están de acuerdo para comenzar una sesión de juego, se redirige a una nueva pantalla donde se despliega el tablero y los elementos del juego tal como se muestran en la Fig. 2. Cuando una sesión de juego finaliza, a causa de una victoria o una derrota, los jugadores regresan a la segunda pantalla, en la cual pueden unirse a una sesión de juego o crear una nueva.

En referencia a las características del juego, los jugadores usan representaciones de los elementos originales del juego de mesa: dados especiales, cartas, fichas, etc. Los jugadores tienen recursos asociados, por ejemplo un conjunto de cartas que solo el jugador mismo puede manipular, o atributos especiales, como un indicador de salud que llegado a cierto valor puede significar la eliminación del jugador en la sesión de juego.

El juego puede ser jugado desde 2 a 5 jugadores. El objetivo final es transportar un objeto especial hasta el último escenario. Para lograrlo, los jugadores deben avanzar a través de diferentes escenarios impidiendo obstáculos. Un enemigo se ubica en el lado opuesto desde donde comienzan los jugadores, y durante el juego avanza en dirección opuesta a estos. Para impedir cruzarse con el enemigo y/o que avance, los jugadores deben tomar decisiones en forma individual y grupal en cada turno. Este tipo de decisiones, permiten el cumplimiento de los objetivos grupales a costa de sacrificios individuales. Por ejemplo, al cruzar de un escenario a otro, para prevenir que el enemigo avance posiciones, los jugadores deben haber coleccionado diferentes elementos en el transcurso del escenario. Por otro lado, en cada escenario, el objeto especial es reasignado, de acuerdo al nivel de salud de los jugadores. Si el enemigo llega a la misma posición en la que se encuentra el portador del objeto, la sesión de juego finaliza en derrota. En caso de que el enemigo llegue a la misma posición de un jugador que no es portador del objeto especial, entonces ese jugador es eliminado de la sesión de juego, pero el juego continúa.



Figura 2. Adaptación digital del juego de mesa

La mayor parte del progreso en el juego es compartido entre todos los jugadores: el grupo debe completar una serie de escenarios para ganar el juego. Y el progreso es común para todos los jugadores. Los jugadores deben resolver determinadas tareas en forma individual, por ejemplo elegir entre tirar el dado o retroceder algunas posiciones para elevar el indicador de salud, y otras en forma grupal, por ejemplo, seleccionar al jugador que deberá descartar una carta especial. Durante el juego se deben balancear las acciones entre beneficiarse a sí mismo y beneficiar a los demás con el fin de obtener la victoria final.

En este contexto, la plataforma registra el estado de la sesión del juego, para el grupo de jugadores, considerando el balance entre:

- el total de los recursos recolectados por los jugadores en relación a la cuota que cada escenario requiere para pasar al siguiente nivel del juego;
- la distancia promedio entre todos los jugadores al enemigo en contraste con la distancia al fin del escenario;
- y la cantidad de cartas especiales.

De esta forma, para cada interacción se registra el estado de la sesión de juego discretizado en uno de tres posibles valores: “Bueno”, “Neutro” o “Malo”.

3.2 Enfoque propuesto

El enfoque que se propone aplicar divide el proceso de clasificación en dos etapas (Fig. 3). La primera etapa de clasificación estará dirigida a la detección de la orientación del mensaje. La clasificación resultante será utilizada para enriquecer los datos de entrada del clasificador de la segunda etapa, que estará dirigido a la detección de la reacción del mensaje.

Este procesamiento de las interacciones sociales llevadas a cabo en las sesiones de juego contribuye al cálculo de un indicador de contribuciones individual y grupal. El indicador considera las reacciones detectadas y determina la cantidad de contribuciones para cada una de las categorías y el porcentaje asociado respecto a la cantidad indiscriminada de interacciones manifestadas. Este mismo indicador se computa para cada usuario y para cada equipo de jugadores (Fig. 1). Como se mencionó en la Sección 3.1, en conjunto con estos indicadores, para cada interacción también se observa el estado de la sesión de juego. Esto permite evaluar el rendimiento de cada usuario en contraste con el equipo.

El proceso de las interacciones requiere ejecutar la clasificación de cada interacción como muestra de una determinada reacción y consecuentemente la distinción del tipo socio-emocional u orientada a tarea. Al finalizar el procesamiento de una base de registros, se busca reconocer la existencia de perturbaciones en la dinámica grupal del equipo. Para esto, se verifican, para cada grupo, que el porcentaje de cada reacción no exceda los umbrales altos y bajos que definen el rango de manifestaciones recomendadas: entre 34% y 3% para reacciones positivas; entre 70% y 18% para reacciones de respuesta; y para reacciones negativas y de requerimiento (o pregunta) entre un 20% y un 1%. De esta manera, se pueden iniciar acciones correctivas específicas para cada caso particular.

En cuanto a la clasificación de interacciones, detección automática de reacciones y análisis a partir de la conversación, en la literatura se han estudiado diversas cuestiones al respecto. En particular, el trabajo realizado por Zhang, C., & Zhang, C. [24] para analizar el contenido de los mensajes en grupos de noticias (medio asincrónico) los clasifican a las categorías IPA. En este estudio emplean como algoritmo de clasificación, específicamente para categorización de texto, SVM [25]. Previa categorización, los mensajes fueron pre procesados aplicando filtro de “stemming”, “stop-words” y construyéndose un registro de frecuencia de apariencia por término o característica. En cuanto a la selección de características, los símbolos especiales y emoticones en este estudio fueron considerados; sólo se consideran aquellas palabras que aparecen en al menos dos registros diferentes; y consideran alguna metacaracterística de la fuente del dataset. La información generada, finalmente, es normalizada. Zhang, C., & Zhang, C. [24] logran una precisión media con respecto a las reacciones IPA de 84.1% para un dataset y 87.2% con otro dataset. Un trabajo similar a Zhang, C., & Zhang, C. [24], pero en un entorno de trabajo colaborativo en línea, fue llevado a cabo por Cincunegui et al. [26] en el contexto de aprendizaje colaborativo. Los registros empleados fueron obtenidos de las conversaciones (medio sincrónico) que mantuvieron los usuarios involucrados en las experiencias. Estos registros fueron pre procesado mediante filtrado de “stop-words” y “stemming”, con los que luego se generó un dataset para entrenar clasificadores con diferentes implementaciones de las técnicas “SVM”, “Decision Tree” y “Naive Bayes”, a las cuatro categorías de reacción y a las doce categorías de colaboración IPA. Obteniendo resultados poco alentadores (SVM accuracy: 54.89%; Decision Tree accuracy: 54.34%; Naive Bayes accuracy: 55.42%), que luego serían levemente mejorados al aplicar un filtro PoSTagging en el pre procesamiento (Naive Bayes accuracy: 54,07%; Decision Tree accuracy: 58,62%; SVM accuracy: 60,67%) [27].

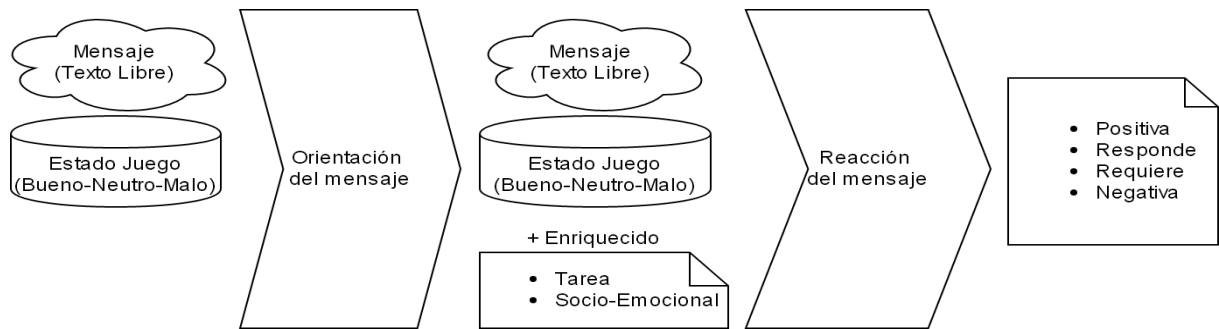


Figure 3. Enfoque de clasificación dividido en fases.

Partiendo de esta serie de trabajos, las diferencias en nuestra propuesta radican en la utilización de una plataforma de juego y la observación de variables externas a las interacciones registradas. Adicionalmente, se pretende incursionar en el empleo de recursos complementarios que ofrecen otros modelos, como es el caso de SYMLOG para características no verbales. A continuación se clasificará el dataset resultante a las reacciones en dos etapas: 1) Inicialmente se agruparán las conductas por la orientación del objetivo: “hacia lo socio-emocional” (“Positivo”, “Negativo”) y “hacia la tarea” (“Pregunta”, “Responde”) de la interacción; 2) Finalmente, se busca detectar directamente las reacciones de las conductas (“Pregunta”, “Respuesta”, “Positivo”, “Negativo”). En la siguiente sección describiremos el proceso experimental que se llevó a cabo para la clasificación automática de las interacciones.

4 Experimentación

Esta sección se encuentra organizada de la siguiente manera. En la Sección 4.1, se detalla el conjunto de datos utilizados para realizar la evaluación experimental. En la Sección 4.2, se detalla el procedimiento para efectuar el experimento. Finalmente, en la Sección 4.3, se muestran los resultados obtenidos y un análisis de los resultados y sus implicancias.

4.1 Conjunto de datos

Para realizar los experimentos se recolectó un conjunto de datos correspondiente a la participación voluntaria realizada por alumnos de la carrera Ingeniería de Sistemas de la Universidad Nacional del Centro de la Provincia de Bs. As., Argentina, durante una materia curricular de 3er año. Participaron 35 alumnos que fueron divididos en 8 grupos de 4 y 5 integrantes cada uno y debían utilizar el juego intentando lograr 2 victorias. Los datos fueron obtenidos mediante el registro de las interacciones de los alumnos al utilizar el chat provisto en el juego. Una vez concluida la etapa de juego, se analizaron los chats de 163 sesiones de juego (Fig. 4) y se estableció de forma manual la conducta (ver sección 2.2) más asociada a cada interacción y el contexto, tanto en el juego (“Bueno”, “Neutro”, “Malo”) como en el flujo de la conversación donde se emite. Sobre el dataset resultante, con un total de 2135 interacciones, se efectuó un pre-procesamiento aplicando stemming, eliminación de stopwords, clasificación Part-of-Speech Tagging (PoST) usando FreeLing², se construyó un registro de frecuencia de aparición por término, considerando los términos que aparecen en al menos 5 registros del juego y se enriqueció el dataset con una variable referida a contexto de la interacción respecto al estado del juego. Finalmente, para la segunda etapa de clasificación, se aumentó cada registro con la categoría de orientación de la interacción a “Tarea” o “Socio-emocional” clasificada en la primera etapa.

² <http://nlp.lsi.upc.edu/freeling/>

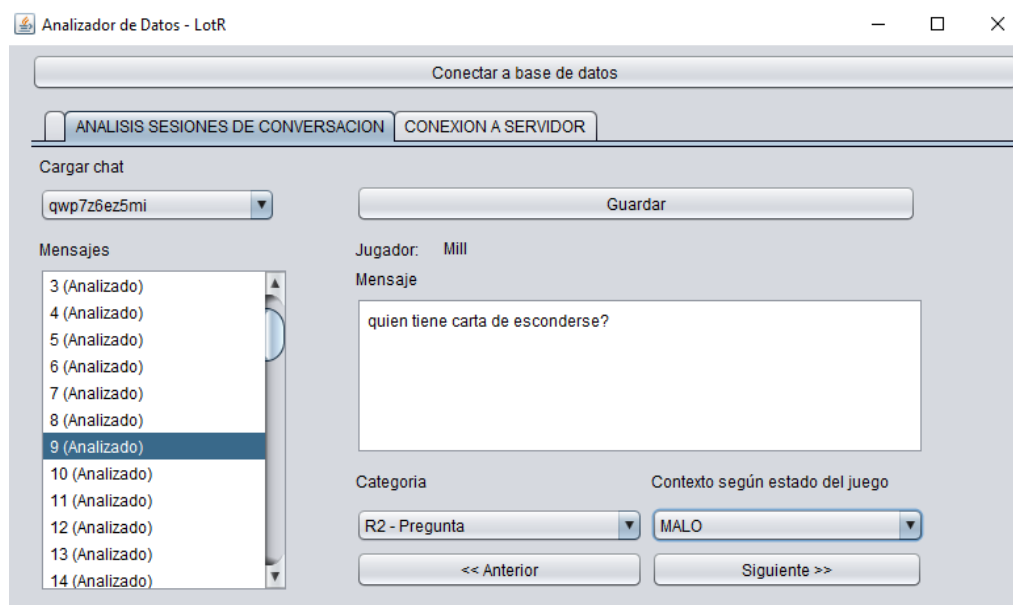


Figura 4. Herramienta de inspección de conversaciones de sesiones de juego

4.2 Proceso

El objetivo de este experimento es encontrar un modelo que permita categorizar en forma automática las interacciones de los alumnos para acelerar los procesos de análisis y caracterización de perfiles de conducta en juegos de colaboración, disminuyendo el alto consumo de recursos humano-temporal que requiere la categorización de interacciones por parte de personas especializadas en el tema. Para lograr dicho objetivo, se plantearon las siguientes preguntas de investigación: (1) ¿Qué algoritmo de clasificación permitirá obtener mejores resultados de clasificación? (2) ¿Qué mejora introduce una división en etapas de la clasificación con respecto al enfoque directo? (3) ¿Es posible lograr una automatización de la detección de reacciones en un juego con el enfoque de clasificación en etapas? Para poder responder, se ejecuta una iteración sobre el dataset utilizando diferentes algoritmos de clasificación. Se busca obtener los resultados más eficientes para su posterior utilización en una asistencia inteligente para los usuarios jugadores, sistemas multi-agentes o personas que trabajen con el método IPA para reconocer patrones en la dinámica social de grupos de juego o trabajo.

4.3 Resultados

En primer lugar, se evaluó la influencia de dividir el proceso de clasificación en dos etapas. Se probaron diferentes algoritmos de clasificación, utilizando 10-fold cross validation sobre el conjunto de datos de entrenamiento. La Tabla 3 muestra la precisión obtenida, mencionando en cada fila el algoritmo utilizado y en cada columna el enfoque empleado. La Tabla 3 contrasta la precisión lograda con el enfoque de clasificación de texto libre directamente a las categorías de reacción y con el enfoque de clasificación en etapas. Se puede observar que emplear una clasificación en partes mejora los resultados de los clasificadores con respecto a la clasificación directa de texto libre a las categorías de orientación del mensaje. Es importante recordar que en el enfoque de clasificación en partes, el resultado de clasificación que se obtiene en la primera etapa es utilizado para enriquecer los datos de entrada de la segunda etapa, aumentando las posibilidades de una mejor clasificación en la etapa final. El mejor clasificador se logra con la implementación J48 (del algoritmo árbol de decisión), obteniendo una precisión de 79,59 % de instancias correctamente clasificadas. La técnica SMO también logra una precisión cercana, con un 78,10% de instancias correctamente clasificadas. Sin embargo, Naive Bayes solo mejora en 5 puntos la precisión (66,34%) y obtiene el porcentaje más bajo de instancias clasificadas correctamente de entre las tres técnicas empleadas para realizar los experimentos.

En la Tabla 4 se muestra la matriz de confusión obtenida para el clasificador J48, indicando en las filas la categoría de los registros y en las columnas como fueron clasificados. Como se puede ver, la pre-clasificación de la "Orientación" (hacia la tarea/hacia lo socio-emocional) de la interacción, produce que en la segunda etapa se acote el margen de error, es decir que las instancias mal clasificadas se mantienen por lo general en la misma categoría de "Orientación". Esto es consecuencia directa de la incorporación del resultado del primer clasificador como entrada del segundo clasificador, sin limitar la categorización de conducta resultante.

Tabla 3: Precisión de clasificadores.

Técnica de clasificación	Directa	Por etapas
Árbol de decisión (J48)	61,69	79,59
Naive Bayes (NB)	61,08	66,34
SVM (SMO)	62,16	78,10

Tabla 4: Matriz de confusión del clasificador J48

Clasifica como →	a	b	c	d
Positivo (a)	583	49	0	0
Negativo (b)	101	51	0	0
Pregunta (c)	0	0	116	168
Responde (d)	0	0	113	954

La matriz de confusión para el clasificador Bayesiano (Tabla 5) muestra el bajo desempeño logrado para este conjunto de datos, indicando en las filas la categoría de los registros y en las columnas como fueron clasificados. De forma similar, la primera etapa de clasificación reduce el margen de error a las categorías dentro de la misma “Orientación”, aunque se observan algunas excepciones.

En la Tabla 6 se muestra la matriz de confusión obtenida para el clasificador SMO, indicando en las filas la categoría de los registros y en las columnas como fueron clasificados. En cuanto a esta última técnica (SMO), muestra resultados similares a los obtenidos con el clasificador J48 sin lograr mejorar la precisión de este último.

Estos experimentos sugieren, entonces que se obtendrán mejores resultados para este dominio con una ejecución en partes mediante la agrupación de las conductas de reacción según la orientación de la contribución. Podemos entonces responder las preguntas planteadas al principio de esta sección:

1. La técnica J48 en el enfoque de clasificación por partes logra el clasificador más eficiente, con una precisión de 79,59% de instancias correctamente clasificadas.
2. La clasificación en partes de las interacciones textuales, clasificando primero la orientación de la contribución (hacia la tarea/ hacia lo socio-emocional) y luego emplear esta primer categorización para identificar la reacción, impacta en forma positiva en la generación de clasificadores. En el mejor de los casos, con el algoritmo J48, aumentó en un 17.9% la precisión. Mientras que el peor de los casos, con la técnica Naive Bayes, la precisión mejoró en un 4.54%.
3. La automatización de la detección de reacciones demostradas vía comunicación textual en el marco de un juego con el enfoque de clasificación por etapas es posible. Aunque los valores pueden seguir siendo mejorados, los clasificadores resultantes son aptos para sugerir la reacción más probable, reduciendo de esta manera la carga de la persona a cargo de la supervisión de los equipos en las sesiones de juego y la materialización de un asistente que sugiera acciones correctivas.

5 Aplicación del clasificador resultante

El objetivo principal de automatizar el proceso de clasificar las interacciones manifestadas en reacciones es asistir a profesores o líderes de equipo en el análisis de la dinámica de los grupos. Para ello, se implementó una herramienta que utiliza el modelo de clasificación entrenado en el primer experimento para detectar alteraciones en la dinámica expresada durante las sesiones de juego. Mientras los jugadores participan de una nueva sesión de juego, la herramienta observa las interacciones textuales y las clasifica en reacciones usando el modelo de clasificación entrenado en el primer experimento. De acuerdo con esta clasificación, se verifica el balance de reacciones manifestadas durante la partida. Si el número de interacciones excede los umbrales definidos (Sección 3.2), se destaca una posible alteración al profesor o líder del equipo. La Fig. 5 muestra una captura de pantalla de la herramienta visualizando los resultados, destacando en naranja las alteraciones detectadas para alertar al profesor o líder del equipo.

Tabla 5: Matriz de confusión del clasificador NB.

Clasifica como →	a	b	c	d
Positivo (a)	556	55	5	16
Negativo (b)	94	48	0	10
Pregunta (c)	0	0	210	74
Responde (d)	5	2	451	609

Tabla 6: Matriz de confusión del clasificador SMO.

Clasifica como →	a	b	c	d
Positivo (a)	601	31	0	0
Negativo (b)	100	52	0	0
Pregunta (c)	0	0	77	207
Responde (d)	0	0	125	942

Se realizó un segundo experimento con el objetivo de comparar el número de alteraciones detectadas por un análisis asistido con nuestro enfoque con respecto al número de alteraciones detectadas por un análisis manual de las interacciones de las sesiones de juego. En este segundo experimento, 35 estudiantes fueron divididos en 7 grupos de 5 miembros cada uno para realizar una primera asignación. El mismo grupo de estudiantes fue asignado a 7 grupos diferentes (con 5 miembros) para realizar una segunda asignación. En total, recogimos las interacciones de 14 grupos participando en dos sesiones de juego. La Tabla 7 muestra el número de alteraciones de cada tipo detectados por nuestro enfoque de clasificación automática y por un análisis manual de las interacciones del grupo. Podemos ver que, aunque el enfoque propuesto alcanza una precisión de clasificación del 79,59%, el número de alteraciones detectadas por nuestro enfoque es una buena aproximación al análisis manual. Nuestro enfoque fue capaz de detectar el 72.7% de las alteraciones existentes, con una tasa de falsos positivos del 24,2%. La tasa de falsos positivos implica que, a partir de las alteraciones detectadas automáticamente por nuestro enfoque, sólo el 24,2% son falsas alarmas y el 75,8% son alteraciones reales.

Un análisis posterior de cada tipo de alteración detectada mostró que la mayoría de los grupos manifestaron una cantidad elevada de interacciones para intercambiar información. De acuerdo a este tipo de reacción, podemos concluir que los estudiantes tendieron a mostrar un comportamiento “enfocado a la tarea” mientras colaboraron para completar victoriosamente el juego. Adicionalmente, los resultados permitieron generar indicadores individuales para cada jugador y aprender el perfil que cada uno demostró dentro del juego. Estos “perfiles de juego”, inferidos a partir de las reacciones identificadas en las discusiones dentro de la plataforma en línea ofrecen una mejor interpretación de los datos y consecuentemente una asistencia de mayor calidad para el profesor o el líder de equipo a cargo.

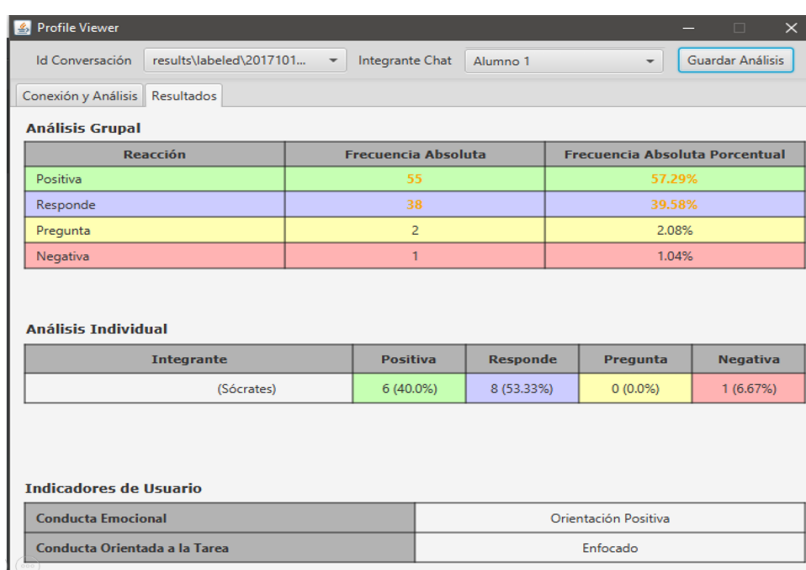


Figura 5 Herramienta desarrollada para mostrar resultados del análisis

Tabla 7: Contraste entre análisis manual y análisis asistido.

Reacción	Análisis manual	Análisis asistido	Falsos positivos
Positivo	5	5	2
Responde	13	7	0
Pregunta	8	7	4
Negativo	7	5	2
Total	33	24	8

6 Conclusiones

En este trabajo se presentaron resultados experimentales de clasificación de texto libre obtenido de las conversaciones surgidas durante sesiones de juego a reacciones de usuarios. Para el dominio de detección de perfiles de usuario mediante la observación de sus conductas en juegos se abre una nueva puerta en este trabajo al obtener clasificadores aptos para la sugerencia de reacciones. El reconocimiento automático de las reacciones de un grupo de usuarios que participan de un juego colaborativo puede a su vez ser mejorado. Los hallazgos de nuestro estudio podrán ser utilizados como evidencia en trabajos futuros de la necesidad de trabajar complementando las interacciones con un análisis de las acciones y contexto más profundo. Los valores resultantes de los clasificadores han permitido determinar que dividir el proceso de clasificación en dos etapas y aumentando la entrada con los resultados de la primera etapa, mejora las predicciones.

Como consecuencia de que la literatura no dispone de muchos estudios sobre la lengua española, creemos que este trabajo efectúa una contribución importante al área de análisis de interacciones. Adicionalmente, pocos estudios trabajan con grupos de usuarios que participan en un juego adaptado a una plataforma colaborativa en línea.

Particularmente, en contraste con el trabajo realizado por Zhang, C., & Zhang, C. [24] cuyo clasificador supera los resultados logrados en el presente trabajo, la diferencia sustancial se debe al canal de comunicación donde se contempla la interacción de los usuarios. Es decir, en un foro (canal de comunicación asincrónico), los usuarios tienen la posibilidad (y deben) elaborar los mensajes de una forma que el mensaje no sea mal interpretado. Por otro lado, en un canal de comunicación sincrónico y en un escenario particular los usuarios pueden emitir mensajes espontáneos, sin previa elaboración y ambiguos. De esta manera se introduce ruido al dataset con el que se trabajará y consecuentemente pérdida de precisión en el desempeño de los clasificadores.

Como trabajo futuro, se estudiará identificar de acuerdo al estado del juego una clasificación más precisa e incursionar en el modelado del comportamiento de usuarios para la construcción de perfiles más robustos. Por otro lado se planea incorporar otros factores al análisis que puedan afectar positivamente a los resultados, como el enriquecimiento del dataset con la incorporación de las acciones que son tomadas en el juego por los usuarios. Alternativamente, se experimentará con instrumentos correspondientes a teorías complementarias sobre la dinámica de grupo. Finalmente, se recolectarán nuevos conjuntos de datos, con grupos más heterogéneos, que permitan replicar el estudio y corroborar los resultados de esta experiencia.

Referencias

- [1] Romero, R. R., & Saune, S. T. (1995). Grupo: objeto y teoría.
- [2] Bales, R. F., Cohen, S. P., & Williamson, S. A. (1979). SYMLOG: A system for the multiple level observation of groups. Free Pr.
- [3] Bales, R. F. Interaction process analysis; a method for the study of small groups (1950).
- [4] Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments (pp. 9–15). ACM. Doi:10.1145/2181037.2181040
- [5] Herger, M. (2014). Gamification in Human Resources. Enterprise Gamification, 3.
- [6] Casamayor, A., Amandi, A., & Campo, M. Intelligent assistance for teachers in collaborative e-learning environments. Computers & Education, 53(4), 1147-1154 (2009). Doi:10.1016/j.compedu.2009.05.025
- [7] Costaguta, R., Garcia, P., & Amandi, A. Using Agents for Training Students Collaborative Skills. Latin America Transactions, IEEE (Revista IEEE America Latina), 9(7), 1118-1124 (2011). Doi: 10.1109/TLA.2011.6129712

- [8] Feldman, J., Monteserin, A., & Amandi, A. (2016). Can digital games help us identify our skills to manage abstractions?. *Applied Intelligence*, 45(4), 1103-1118.
- [9] Feldman, J., Monteserin, A., & Amandi, A. (2014). Detecting students' perception style by using games. *Computers & Education*, 71, 14-22. Doi:10.1016/j.compedu.2013.09.007
- [10] Jin, C.-H. (2014). The role of users' motivations in generating social capital building and subjective well-being: The case of social network games. *Computers in Human Behavior*, 39, 29-38. Doi:10.1016/j.chb.2014.06.022
- [11] Kosterman, S., & Gierasimczuk, N. (2015). Collective Learning in Games Through Social Networks. In *Proceedings of the 1st International Conference on Social Influence Analysis - Volume 1398* (pp. 35-41). Aachen, Germany, Germany: CEUR-WS.org.
- [12] Popescu, M., Romero, M., & Usart, M. (2012). Using serious games in adult education serious business for serious people-the MetaVals game case study. In *ICVL 2012-7th International Conference on Virtual Learning* (pp. 125-134).
- [13] Romero, M., Usart, M., & Almirall, E. (2011). Serious games in a finance course promoting the knowledge group awareness. In *EDULEARN11 Proceedings* (pp. 3490-3492). IATED.
- [14] Wendel, V., Gutjahr, M., Göbel, S., & Steinmetz, R. (2013). Designing collaborative multiplayer serious games. *Education and Information Technologies*, 18(2), 287-308.
- [15] Kirriemuir, J., & McFarlane, A. (2004). Literature review in games and learning. Retrieved from <https://telearn.archives-ouvertes.fr/hal-00190453/>
- [16] Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving Participation and Learning with Gamification. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications* (pp. 10-17). New York, NY, USA: ACM.
- [17] Zichermann, G. (2011). The purpose of gamification. A look at gamification's applications and limitations. *Radar*, April, 26.
- [18] Zackariasson, P., & Wilson, T. L. (2012). *The Video Game Industry: Formation, Present State, and Future*. Routledge.
- [19] Linehan, C., Lawson, S., & Doughty, M. (2009, March). Tabletop Prototyping of Serious Games for 'Soft Skills' Training. In *Games and Virtual Worlds for Serious Applications, 2009. VS-GAMES'09. Conference in* (pp. 182-185). IEEE. Doi: 10.1109/VS-GAMES.2009.9
- [20] Haferkamp, N., Kraemer, N. C., Linehan, C., & Schembri, M. (2011). Training disaster communication by means of serious games in virtual environments. *Entertainment Computing*, 2(2), 81-88. Doi:10.1016/j.entcom.2010.12.009
- [21] Zagal, J. P., Rick, J., & Hsi, I. (2006). Collaborative games: Lessons learned from board games. *Simulation & Gaming*, 37(1), 24-40. Doi: 10.1177/1046878105282279
- [22] Berdun, F. (2014). Identification of collaborative skills with serious games. In *XLIII Jornadas Argentinas de Informática e Investigación Operativa (43JAIIO)-Doctoral Consortium (IJCAI)(Buenos Aires, 2014)*.
- [23] Johnson D, Johnson R (1994) *Learning Together and Alone, Cooperative, Competitive, and Individualistic Learning*. Needham Heights, MA: Prentice-Hall.
- [24] Zhang, C., & Zhang, C. (2005). Discovering Users' Participant Roles in Virtual Communities with the Help of Social Interaction Theories. *PACIS 2005 Proceedings*, 65.
- [25] Joachims, T. (1998). Making large-scale SVM learning practical (No. 1998, 28). Technical Report, SFB 475: Komplexitätsreduktion in Multivariaten Datenstrukturen, Universität Dortmund.
- [26] Cincunegui, M., Berdun, F., Armentano, M. G., & Amandi, A. (2015). Clasificación de conductas colaborativas a partir de interacciones textuales. In *Argentine Symposium on Artificial Intelligence (ASAI 2015)-JAIIO 44* (Rosario, 2015).
- [27] Berdun, F. D., Armentano, M. G., Berdun, L., & Mineo, M. (2017). Classification of collaborative behavior from free text interactions. *Computers & Electrical Engineering*. Doi:10.1016/j.compeleceng.2017.07.015