

¿Puede un videojuego ayudarnos a predecir los resultados de aprendizaje?

Ana Illanas, Faraón Llorens, Rafael Molina, Francisco Gallego, Patricia Compañ,
Rosana Satorre, Carlos Villagrà

Grupo Informática Industrial e Inteligencia Artificial. Equipo ByteRealms. Universidad de
Alicante

{aillanas, faraon.llorens, rmolina, fran.gallego, company,
rosana.satorre, villagra}@ua.es

Resumen. Los videojuegos educativos tienen características (inmediatez, motivación, diversión, interacción...), que permiten mejorar el aprendizaje. Además, la acción del juego genera una gran cantidad de datos que pueden ayudarnos a conocer el progreso en el aprendizaje y a predecir sus resultados. Presentamos CoIn, un videojuego educativo en el que el jugador se convierte en un intérprete que debe completar las traducciones al español de varios discursos en inglés. CoIn, además, recoge multitud de datos durante el desarrollo del juego, que luego se utilizan en un sistema de predicción basado en técnicas de *data mining*. En particular, se propone una red neuronal cuya entrada es este conjunto de datos y cuya salida es un rango de notas finales en las actividades docentes de la asignatura. A pesar de la limitada cantidad de datos disponible, los resultados son prometedores y nos permiten aventurar que es posible medir el progreso y predecir el aprendizaje.

Palabras clave. Videojuegos educativos, educational data mining, aprendizaje

1 Introducción

El uso de los videojuegos con fines educativos es un campo en auge en los últimos años. Se han realizado estudios sobre sus beneficios en la enseñanza, no sólo por su carácter atractivo y motivador para el estudiante, sino también por la gran variedad de campos y materias que pueden abarcar así como las habilidades que pueden entrenar. Considerando estos beneficios como premisa ¿es posible extraer alguna conclusión interesante del análisis de la gran cantidad de datos que se generan en un juego como resultado de la interacción y de la propia acción de jugar? Es más, en un juego educativo ¿es posible utilizar estos datos para conocer cómo está siendo el aprendizaje e incluso tratar de predecir sus resultados?

En este trabajo presentamos CoIn (*The Conference Interpreter*), un videojuego educativo en el que el jugador se convierte en un intérprete que debe completar las traducciones en castellano de varios conferenciantes que realizan sus discursos en inglés. CoIn es, además, una plataforma de experimentación que permite recoger

multitud de datos durante el desarrollo del juego, que luego se utilizan en un sistema de predicción basado en técnicas de *data mining*. De esta forma, se consigue convertir un videojuego educativo en una plataforma de enseñanza-aprendizaje, aprovechando su carácter lúdico y motivador para que el estudiante avance en el conocimiento.

Los objetivos principales son dos: proporcionar herramientas de medición del progreso en el aprendizaje y de predicción de la evolución del estudiante, y experimentar y avanzar en las técnicas de *data mining*.

En el apartado 2, se presentan los antecedentes de esta investigación, realizando un repaso del uso de los videojuegos en la enseñanza y de algunos trabajos interesantes sobre *data mining*. En el apartado 3, se presenta el videojuego CoIn, explicando su funcionamiento y los datos que almacena, y experimentando con esos datos como entrada de un sistema predictivo de *data mining*. A continuación, el apartado 4 recoge los resultados de la experimentación. Por último, se presentan las conclusiones que extraemos de este trabajo y se esbozan las líneas futuras de investigación.

2 Antecedentes

El uso de la tecnología en general y de los videojuegos en particular en nuestros momentos de ocio es una realidad y la facilidad con la que los nativos digitales los utilizan puede ser aprovechada por los docentes para el aprendizaje [1]. Los videojuegos nos atraen por diversos motivos: nos incitan a participar, nos motivan introduciendo pequeños objetivos alcanzables, nos premian y nos castigan de forma inmediata, favoreciendo nuestro proceso de aprendizaje, nos permiten ganar a oponentes, los niveles de dificultad están adaptados en función de nuestras destrezas, edades, conocimiento del juego, etcétera [2].

Son numerosos los estudios que indican que la diversión favorece el aprendizaje y la motivación y que se reduce el estrés [3-5]. Además, en los juegos el miedo a cometer errores se reduce y, por tanto, el proceso de aprendizaje se ve favorecido [6].

Los videojuegos tratan de conseguir una completa inmersión del jugador en la tarea que está realizando, posibilitando que éste decida sobre qué hacer en cada momento, además de favorecer la competitividad y el trabajo colaborativo. Junto a esto, el hecho de que durante el juego el jugador obtenga una retroalimentación inmediata tras cada acción, le permite el aprendizaje mediante prueba y error, factores deseables en un entorno de aprendizaje constructivista y centrado en el estudiante [7, 8].

Otras investigaciones han mostrado la efectividad de los videojuegos como herramienta para el aprendizaje en diversas áreas tales como álgebra, ingeniería, negocios y gestión empresarial o medicina [9-12]. Son numerosos los autores que han analizado juegos de rol online multijugador masivos (MMORPG), como Ever Quest II o World of Warcraft, y otros de realidad virtual, como The Sims o Second Life, para evidenciar las posibilidades educativas que ofrecen estas herramientas, partiendo del desarrollo de identidades individuales o de grupo [8, 13-15].

Pero no sólo está extendido el uso de los videojuegos en los ámbitos citados anteriormente, algunos autores destacan las ventajas de la utilización de entornos de aprendizaje. Al contrario de lo que se cree, el docente también tiene posibilidad de

intervención tal y como ocurre en la enseñanza tradicional, puesto que en las plataformas *e-learning* los participantes dejan rastros que proporcionan al docente pistas sobre el desarrollo de la actividad de forma muy precisa [16].

Por otro lado, los videojuegos ofrecen un contexto sociocultural que favorece el aprendizaje de segundas lenguas, particularmente la comprensión auditiva, el aprendizaje de vocabulario y la adquisición de competencias comunicativas [17-20].

En este trabajo proponemos utilizar técnicas de *data mining*, por lo que se hace necesario revisar esta disciplina. *Data mining* es el proceso computacional de descubrir correspondencias y patrones útiles entre grandes cantidades de datos [21, 22]. Cuando los datos explorados provienen del campo educativo, hablamos de *educational data mining*. Algunos objetivos, en este caso, son mejorar el entendimiento de los estudiantes y los métodos de enseñanza [23] o guiar a los estudiantes en el aprendizaje [24], utilizando *logs* e interacciones de los estudiantes como fuente de datos [25].

Algunos de los principales campos en los que se han aplicado los modelos de *educational data mining* son el aprendizaje individual para software educativo, el apoyo automatizado para aprendizaje colaborativo, las pruebas adaptativas automatizadas, la detección de factores asociados al fracaso escolar, la mejora de los modelos de los estudiantes (características y estados de los mismos) y el descubrimiento o mejora de los modelos de estructuras de conocimiento [26]. Por otro lado, se han utilizado técnicas de *data mining* para adaptar la respuesta del software de manera más personalizada [27]. También hay experiencias que tratan de predecir los factores que causan el fracaso de los estudiantes [28]. Además, se han utilizado estas técnicas para el estudio de los cursos on-line y en *learning management systems*, mostrando las ventajas de utilizar un entorno de aprendizaje en lugar de una simple web como fuente de datos para un sistema de *data mining* [29, 30].

3 CoIn. The Conference Interpreter

3.1 CoIn como juego

CoIn es un videojuego sobre interpretación simultánea donde los jugadores deben realizar traducciones simultáneas de varias conferencias. El jugador debe superar varios niveles, tratando de conseguir la mejor puntuación. Es una herramienta educativa para complementar la enseñanza en estudios de lenguas extranjeras. De hecho, ya ha sido utilizado en una experimentación orientada a la docencia en una asignatura de inglés con estudiantes universitarios y se ha observado una mejora en las notas finales con respecto a los estudiantes que no utilizaron el videojuego [31]. En este trabajo vamos un paso más allá, y utilizamos el videojuego como sistema de predicción.

En CoIn la componente lúdica es fundamental. Por eso se basa en el desarrollo de las habilidades prácticas de los estudiantes y no incluye explicaciones teóricas durante el juego. Este videojuego puede ser considerado como un minijuego conceptual: está centrado en mejorar una habilidad particular del jugador mediante varios niveles con unos ejercicios prácticos sencillos [32].

En la versión que nos ocupa, la pantalla principal simula una sala donde un conferenciante está hablando en inglés (figura 1). Sin embargo, el videojuego es fácilmente

adaptable a cualquier idioma, sin más que reemplazar el audio y su transcripción. En cualquier caso, el jugador, como intérprete, debe escuchar la conferencia y completar la traducción de la misma a otro idioma, en nuestro caso al español. En la parte inferior de la pantalla la traducción aparece al mismo tiempo que el conferenciante está hablando. Sin embargo, esta traducción no está completa. Algunas palabras han sido reemplazadas por huecos vacíos y la tarea del jugador es elegir una de las cuatro opciones que el juego ofrece como respuesta para cada palabra oculta.



Fig. 1. Pantalla principal de CoIn

Las cuatro opciones que CoIn ofrece no son siempre las mismas para cada ejecución del juego. Las tres opciones incorrectas son escogidas de una lista de términos similares de forma aleatoria. Además, la conferencia y el juego continúan independientemente de si el jugador contesta a la palabra de forma correcta o incorrecta, o incluso si no contesta.

El jugador debe mantener un nivel de vida durante el juego. Existe una barra de energía en la parte superior izquierda de la pantalla, que empieza con un 30% al principio de cada nivel. Si se elige una respuesta correcta se incrementa en un 3% pero, si por el contrario se elige una respuesta incorrecta, entonces esta barra decrece un 10%. Si la barra de vida se vacía por completo, el juego termina, aunque podrá volver a jugar de nuevo el nivel desde el principio.

El tiempo de duración de la conferencia y la progresión de la misma se muestran en una segunda barra en la parte superior de la pantalla. Cuando el nivel finaliza, habiendo el jugador ganado, perdido o abandonado, aparece la pantalla de resultados (figura 2). Como ayuda, el juego incorpora tres comodines que pueden utilizarse en cada conferencia: *Thing* (permite ignorar el hueco en blanco sin ser penalizado), *Potion* (incrementa el nivel de vida un 25%) y *50%* (reduce las cuatro posibles respuestas a dos).



Fig. 2. Pantalla de resultados de CoIn

CoIn está dividido en varias historias. Cada historia contiene un mapa (figura 3) con un itinerario de conferencias que el jugador debe superar. Cada punto en el mapa representa un nivel con una conferencia diferente para ser interpretada. El jugador tiene la libertad de jugar cualquier nivel una o varias veces. Pero algunos niveles están bloqueados y el jugador deberá superar otros para desbloquearlos.



Fig. 3. Pantalla de mapa en CoIn

3.2 CoIn como recolector de datos

CoIn es fácilmente adaptable a cualquier texto y a cualquier idioma, puesto que las transcripciones de los textos de cada nivel están organizadas en ficheros XML. Para crear nuevos niveles o adaptar el juego a otros idiomas basta con crear estos archivos XML y sincronizar el audio y el texto de la transcripción.

En estos ficheros XML, las frases se sincronizan con el audio utilizando atributos que indican el momento en el que la frase debe mostrarse en la pantalla y durante cuánto tiempo. Un ejemplo de archivo XML se muestra en la tabla 1.

```

<text lang="es">
  <s begin="0'00.00" end="0'00.00">
    Hoy <note id="presentamos">presentamos</note>
  </s>
  <s begin="0'02.84" end="0'03.72">
    el <note id="iPad2">iPad2</note>,
  </s>
  <s begin="0'04.53" end="0'06.10">
    el <note id="iPad">iPad</note>
    de <note id="segunda">segunda</note> generación.
  </s>
  ...
</text>

```

Tabla 1. Ejemplo de archivo XML conteniendo la transcripción y sincronización de un texto

Los términos que tienen que ser reemplazados por un hueco en blanco durante el juego se indican con la etiqueta `note` que admite otro atributo `cat` usado para clasificar gramaticalmente cada término: categoría gramatical (nombre, adjetivo, verbo...), género y número. Esta característica hace posible identificar el grupo de términos que resultan más complicados para los estudiantes y centrar en ellos la enseñanza.

Las diferentes respuestas que pueden aparecer en el juego como posibles soluciones para cada término se especifican también en el archivo XML, como puede observarse en el ejemplo de la tabla 2.

```

<notes>
  <note id=""archivar"" cat="vpl">
    <sub>
      <true>"archivar"</true>
      <false>"menos";"atrás";"adelante";"pausa";</false>
    </sub>
  </note>
  <note id=""buscar"" cat="vpl">
    <sub>
      <true>"buscar"</true>
      <false>"salir";"menos";"pausa";"cancelar";</false>
    </sub>
  </note>
  ...
</notes>

```

Tabla 2. Ejemplo de archivo con las respuestas alternativas

El juego solo muestra cuatro términos como solución (uno correcto y tres incorrectos) pero puede haber más de tres términos incorrectos en el archivo XML. El juego escogerá tres de estos términos incorrectos de forma aleatoria para hacer el juego diferente cada vez que se juega.

CoIn recoge una gran cantidad de datos durante el juego; además la puntuación del jugador, también almacena cada una de las acciones que el jugador realiza mientras juega. Esto resulta muy útil a la hora de realizar análisis exhaustivos del proceso de aprendizaje de cada estudiante.

CoIn genera dos archivos de datos diferentes durante el juego, con diferente información en cada uno de ellos: *game_stats* y *word_stats*.

Los archivos *game_stats* contienen información general de cada nivel. En la tabla 3 encontramos un ejemplo de este tipo de archivos.

id	archivo	fecha	hora	dur	result	punt	vida	tot	Aciertos	fallos	ignore	tpausa	tpoción	t50%	tthing
1	app1.xml	29/11	19:25	39	abandon	0	10	9	0	2	0	30	0	0	0
2	app1.xml	29/11	19:27	24	abandon	0	20	9	0	1	0	20	0	0	0
3	app1.xml	29/11	19:29	68	complet	89	44	9	8	1	0	60	0	0	0

Tabla 3. Información en un fichero *game_stats*

Sin ser exhaustivos, en los archivos *game_stats* se almacena información como el identificador de la partida que se está jugando, el archivo XML, la fecha y hora de inicio, la duración de la partida, el resultado del juego, la puntuación obtenida, el nivel de vida, el número total de términos y de ellos, cuántos se han acertado, fallado e ignorado, los tiempos de espera, los comodines utilizados, etc.

Los archivos *word_stats* contienen información detallada sobre cada acción del jugador para cada hueco vacío. La tabla 4 presenta un ejemplo de estos archivos.

note_id	id	archivo	t_apar	t_resp	estado	método	respuesta	comodín
iPad	1	app1.xml	97	17.385	Falla	raton	iPod	nil
aprendido	1	app1.xml	19.063	37.483	Falla	raton	retocado	nil
iPad	2	app1.xml	95	11.003	Falla	raton	iPod	nil

Tabla 4. Información en un archivo *word_stats*

Entre la información almacenada, se encuentra el identificador del término, el identificador de la partida, el nombre del archivo XML con el nivel, los tiempos de aparición y respuesta, el resultado de la respuesta, el método de interacción utilizado, el texto seleccionado o si se utilizó un comodín.

Los datos recogidos se usan en el sistema de predicción presentado a continuación.

3.3 CoIn como sistema de predicción

El objetivo final de esta propuesta es utilizar el videojuego para predecir las notas finales que obtendrán los estudiantes en una determinada asignatura. Nos apoyamos en técnicas de *data mining*, en particular en el uso de una red neuronal [33, 34].

En el experimento participaron un conjunto de 55 estudiantes de una asignatura oficial de inglés. Mientras los estudiantes jugaban a CoIn durante dos horas se fueron

registrando sus datos. También se obtuvieron las notas finales de los estudiantes en la asignatura, a través de la evaluación tradicional que realiza el profesor. En concreto, se obtuvieron 5 notas para cada estudiante: nota media, nota de *quiz* (ejercicio tipo test), nota de la prueba oral, nota de redacción (*writing*), participación en un blog. El objetivo es predecir cada una de estas notas, utilizando los datos de los 55 estudiantes.

Puesto que CoIn genera una gran cantidad de información mientras el jugador está jugando, es necesario hacer una selección de los datos que se consideren más adecuados. En este caso, para cada jugador, los datos seleccionados han sido siete:

- Número total de palabras a resolver en todo el juego.
- Número total de palabras acertadas por el jugador.
- Número total de palabras falladas por el jugador.
- Número total de palabras ignoradas por el jugador.
- Tiempo medio de respuesta (global, independiente de si falla o acierta).
- Tiempo medio que el jugador empleó en resolver las palabras acertadas.
- Tiempo medio que el jugador empleó en resolver las palabras que falló.

Planteamos una red neuronal independiente para predecir cada una de las 5 notas consideradas: media, *quiz*, oral, *writing* y blog. Según los datos seleccionados en la fase de preprocesado, cada una de estas redes neuronales tendrá un esquema como muestra la figura 4.

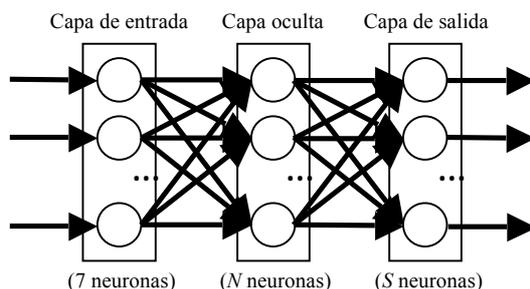


Fig. 4. Esquema de cada una de las 5 redes neuronales

Estas redes disponen de 7 neuronas en la capa de entrada (una por cada dato de entrada del juego), N neuronas en la capa oculta y S neuronas en la capa de salida de la red, parámetros que se ajustarán junto con otros durante la experimentación.

Cada neurona de la capa de salida produce un valor entre 0 y 1, que representa la probabilidad de que la nota del estudiante esté comprendida en un rango predefinido. La salida de la red neuronal se obtiene escogiendo la salida con probabilidad más alta. El número de neuronas en la capa de salida determina, por tanto, el número de rangos en los que se dividirá el espacio de posibles notas de los estudiantes. Los rangos se construyen tomando como notas menor y mayor, respectivamente, la nota mínima y máxima dentro de los valores reales de la muestra (figura 5).

El número de neuronas en la capa de salida S es configurable y los rangos que representan cada una de estas neuronas de salida se ajustan de forma automática

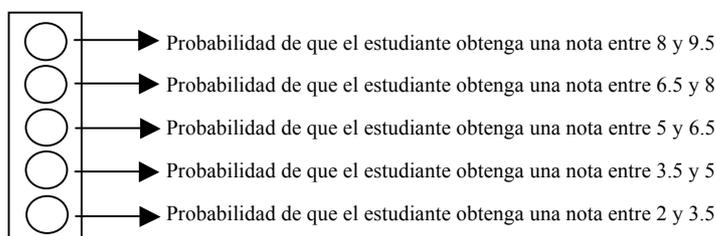


Fig. 5. Ejemplo de los rangos de salida para una capa de salida de cinco neuronas, siendo la nota mínima 2 y la máxima 9.5

Para el entrenamiento de la red neuronal de hace necesario dividir los 55 estudiantes participantes en dos conjuntos: el *training set*, conjunto de estudiantes para entrenar la red neuronal, y *test set*, conjunto de estudiantes para probar el funcionamiento de la red una vez ya está ajustada (la red no debe conocer los datos de estos estudiantes hasta que no haya sido entrenada completamente para que no la contaminen). La proporción de muestras seleccionadas ha sido del 84% para el *training set* (46 estudiantes) y del 16% para el *test set* (9 estudiantes), seleccionados al azar.

Para el entrenamiento se ha elegido el algoritmo de *backpropagation* [35], que permite ajustar iterativamente los pesos de las neuronas para minimizar el error de predicción. Durante el entrenamiento se han fijado también los valores de S (número de neuronas en la capa de salida), de N (número de neuronas en la capa oculta) y de λ (parámetro de regularización del algoritmo de *backpropagation*, encargado de ayudar a la red a generalizar mejor). Para S , se ha comprobado que el valor más adecuado es 5, obteniéndose por lo tanto 5 rangos de notas. En cuanto a los valores de N y λ , los mejores valores dependen de cada experimento y se presentan en la experimentación.

4 Experimentación

Una vez entrenadas las redes neuronales con el conjunto de los 7 datos provenientes del juego a la entrada y con las 5 notas de las que se dispone para cada estudiante a la salida, se les presentan los datos correspondientes a los 9 estudiantes que las redes neuronales no conocen ya que no han sido utilizados para su aprendizaje.

Como se ha dicho anteriormente, para la salida se han utilizado 5 neuronas donde cada una representa un rango de notas que el estudiante tiene una cierta probabilidad de obtener. Estos rangos serán diferentes en cada una de las redes neuronales, pues provienen de las notas reales de los estudiantes, así como los valores de N y λ que provienen de la fase de entrenamiento. En la tabla 5 se presentan las predicciones realizadas por las 5 redes neuronales para cada nota, comparándolas con las notas reales obtenidas por cada estudiante. Sombreadas aparecen aquellas predicciones que difieren de la nota real. Se presenta también el valor de los parámetros N y λ que se han obtenido durante el entrenamiento.

	Media		Blog		Oral		Quiz		Writing	
	$\lambda = 3, N = 5$		$\lambda = 1, N = 15$		$\lambda = 3, N = 55$		$\lambda = 3, N = 15$		$\lambda = 1, N = 65$	
Estud	Real	Pred	Real	Pred	Real	Pred	Real	Pred	Real	Pred
1	1	1	2	2	1	1	2	2	1	1
2	1	1	2	2	3	3	4	3	5	4
3	4	4	4	4	3	3	4	4	2	1
4	2	2	3	2	3	3	2	2	2	2
5	4	4	5	5	4	4	4	4	3	3
6	4	4	4	4	4	4	4	4	3	3
7	3	3	4	4	2	2	4	3	1	1
8	2	1	4	4	2	3	3	3	2	1
9	5	4	4	4	4	4	1	1	4	4

Tabla 5. Resultados de la experimentación

Como puede observarse, las predicciones se ajustan razonablemente bien a los resultados reales. Además, en los casos en los que las predicciones son erróneas, se les adjudica a los estudiantes un intervalo contiguo al real.

5 Conclusiones y trabajos futuros

Las sinergias entre los videojuegos y las técnicas de *educational data mining* están aún por explorar. Existen muy pocas experiencias previas del uso de videojuegos como fuente de datos para un sistema de *data mining*, por lo que los resultados obtenidos son un buen comienzo para continuar en este campo y confirman que es posible predecir con cierto grado de exactitud los resultados de un estudiante a partir de su actividad en un videojuego, dos actividades con escasa relación aparentemente.

Siendo conscientes de la insuficiencia de datos con los que se ha contado para el experimento, el sistema de predicción que aquí se ha presentado representa un paso importante para la elaboración de herramientas de medición del progreso y de predicción que ayuden tanto a profesores como a estudiantes a orientar la enseñanza de un modo cada vez más personalizado.

El trabajo que se ha presentado cumple el objetivo de ser un punto de partida hacia un sistema de predicción basado en videojuegos. Los resultados obtenidos son prometedores a pesar de contar con una muestra tan reducida. Se han diseñado, implementado y probado cinco redes neuronales, una por cada nota que se quería predecir y se ha conseguido un porcentaje de acierto del 88.9% y del 66,7% en el mejor y el peor de los casos respectivamente. El ampliar y mejorar estas redes neuronales es un trabajo a realizar en el futuro así como investigar con otras técnicas de minería de datos. Otro aspecto muy interesante es utilizar estas técnicas para identificar las variables de entrada que resultan más significativas para la predicción.

El mayor problema que se ha presentado es la reducida cantidad de datos con los que se contaba como entrada al sistema de *data mining*. Disponer de una muestra tan

pequeña ha limitado enormemente la calidad y precisión de las redes neuronales que se han utilizado, por lo que en un futuro, se tendrá que poner empeño en conseguir un número de datos muy superior para poder mejorar las investigaciones en este campo.

Además, CoIn almacena muchos más datos de los que se han utilizado en este trabajo. Sería posible obtener otras conclusiones utilizando otra combinación de datos como entrada al sistema de *data mining*. Estudiaremos este aspecto en un futuro.

Con respecto al estudio de lenguas, un posible experimento para el futuro sería el determinar si los resultados obtenidos son independientes del idioma. Para ello se adaptaría CoIn a otra lengua y se recogerían los datos que genera mientras los estudiantes juegan. Estos nuevos datos se utilizarían en el sistema de predicción y sería posible estudiar si el idioma influye de algún modo en los resultados.

Finalmente, otra posible ampliación de este trabajo sería expandirlo a otros campos, desarrollando otros videojuegos orientados a otras disciplinas y sometiendo los datos que generan a sistemas de predicción similares al que aquí se presenta.

6 Referencias

1. Vassileva, J (2008). "Toward Social Learning Environments". IEEE Transactions on Learning Technologies, vol. 1 no. 4.)199-214
2. Prensky M. 2001. *Digital Game-Based Learning*. McGraw-Hill.
3. Bisson C., Luckner J. 1996. "Fun in Learning: The Pedagogical Role of Fun in Adventure Education", *Journal of Experiential Education*, 19(2):107-112.
4. Kirriemuir J., McFarlane A. 2004. "Literature Review in Games and Learning", *Futurelab*. Bristol.
5. Cordova D.I., Lepper M.R., 1996. "Intrinsic Motivation and the Process of Learning: Beneficial Effects of Contextualization, Personalization, and Choice", *Journal of Educational Psychology*, 88(4): 715-730.
6. Nemitcheva, N. (1995). The psychologist and games in the intensive foreign language game-based course. In K. Arai & D. Crookall (Eds.), *Simulation and gaming across disciplines and cultures* (pp. 70-4). Thousand Oaks, CA: Sage Publications, Inc.
7. Clinton, K. A. (2004). Embodiment in digital worlds: What being a videogame player has to teach us about learning. In Annual Meeting of the American Educational Research Association. San Diego.
8. de Freitas, S. (2006). Learning in immersive worlds: A review of game-based learning. London: Joint Information Systems Committee.
9. Mayo, M. J. (2007). Games for science and engineering education. *Communications of the ACM*, 50(7), 31-35.
10. Collier, B. D., & Scott, M. J. (2009). Effectiveness of using a video game to teach a course in mechanical engineering. *Computers & Education*, 53(3), 900-912.
11. Blunt, R. (2009). Does game-based learning work? Results from three recent studies. *ELearn Magazine*. Education and Technology in Perspective, 2009(12).
12. Göbel, S., Hardy, S., Wendel, V., Mehm, F., & Steinmetz, R. (2010). Serious games for health: Personalized exergames. In Proceedings of the international conference on multimedia (pp. 1663-6). New York: ACM.
13. de Freitas, S., & Neumann, T. (2009). The use of 'exploratory learning' for supporting immersive learning in virtual environments. *Computers & Education*, 52(2), 343-352.

14. Melchor-Couto, S. (2011). Second life en el aula de ELE: Un tándem virtual. Suplementos MarcoELE. Revista De Didáctica Español Lengua Extranjera, (13), 63-72.
15. Thorne, S. L., Fischer, I., & Lu, X. (2012). The semiotic ecology and linguistic complexity of an online game world. *ReCALL*, 24(3), 279-301.
16. Marty, J.C. & Carron, T. (2011). Observation of Collaborative Activities in a Game-Based Learning Platform. *IEEE Transactions on Learning Technologies* 4(1), 98-110
17. Hu, M. M. & Chang, B. (2008). Massively Multiplayer Online Game Supported Foreign Language Listening Ability Training. First IEEE International Conference on Digital Games and Intelligent Toys Based Education (pp. 176-8). Washington, D.C.: IEEE Computer Society.
18. Huyen, N. T. T. & Nga, K. T. T. (2003). Learning vocabulary through games. *Asian EFL Journal*, 5(4).
19. Schwienhorst, K (2002). Why virtual, why environments? Implementing virtual reality concepts in computer-assisted language learning. *Simulation & Gaming*, 33(2), 196-209.
20. Nesselhauf, N. & Tschichold, C. (2002). Collocations in CALL: An investigation of vocabulary-building software for EFL. *Computer Assisted Language Learning*, 15(3), 251-279.
21. Witten, I.H. and Frank, E. 1999. *Data mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, San Francisco, CA
22. Klossgen, W., Zytkow, J. 2002. *Handbook of data mining and knowledge discovery*. New York: Oxford University Press.
23. Educational Datamining Society (<http://www.educationaldatamining.org>)
24. Romero, C., Ventura, S., Bra, P.D. 2004. Knowledge discovery with genetic programming for providing feedback to courseware author. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, 14(5), 425-464.
25. Pahl, C., Donnellan, C. 2003. Data mining technology for the evaluation of web-based teaching and learning systems. In *Proceedings of the congress e-learning*, Montreal.
26. Baker R.S.J.D, Yacef, K. 2010. The State of Educational Data Mining in 2009: A review and Future Visions. *Journal of Educational Data Mining*. Vol. 1, issue1, Pages 3-17.
27. Corbett, A.T. 2001. Cognitive Computer Tutors: Solving the Two-Sigma Problem. In *Proceedings of the International Conference on User Modeling*, 137-147.
28. Romero, C., Ventura, S. Espejo, P.G. & Hervas, C. 2008. Data Mining Algorithms to Classify Students. *Proceedings 1st Int. Conf. on Educational Data Mining*, 8-17.
29. Zaïne, O. 2001. Web usage mining for a better web-based learning environment. In *Proceedings of conference on advanced technology for education*, 60-64.
30. Romero, C., Ventura, S., García, E. 2008. Data mining in course management systems: Moodle case study and tutorial. *Journal Computers & Education*. vol. 51, 368-384.
31. Calvo Ferrer, J.R. (2012). Videojuegos y aprendizaje de segundas lenguas: Análisis del videojuego *The Conference Interpreter* para la mejora de la competencia terminológica. *Tesis doctoral, departamento de Filología Inglesa, Universidad de Alicante, 2012*
32. Illanas Vila, A. I., Gallego Durán, F. J., Satorre Cuerda, R. y Llorens Largo, F. (2008). Conceptual mini-games for learning. In *INTED2008 Proceedings*. <http://rua.ua.es/dspace/bitstream/10045/8495/1/illanas08conceptual.pdf>
33. Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice Hall
34. Anderson, J. A. (1995). *Introduction to Neural Networks*. Cambridge, MA: MIT Press
35. Rumelhart, D.E.; Hinton, G.E., Williams, R.J. (1986). "Learning representations by back-propagating errors". *Nature* 323 (6088): 533-536