

Empleo de ChatBots educativos como recurso complementario en las prácticas docentes

José Federico Medrano, Mario Alberto Tejerina, César Alejandro Castillo

jfmedrano@fi.unju.edu.ar, mariotejerina@gmail.com, ce_al_castillo@yahoo.com.ar,

VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería

Universidad Nacional de Jujuy - Ítalo Palanca 10, +54 (388) 4221587

RESUMEN

Los entornos virtuales se ofrecen como claros complementos a la labor docente, con los avances tecnológicos, la inclusión de estos dentro del aula ha propiciado un cambio y una revolución digital por más de una década. Estos avances no sólo han cambiado el modo de enseñar de los docentes sino también el modo de aprender de los alumnos. El aula ya no se encuentra definida por el espacio físico, la virtualidad expande los límites hacia nuevos horizontes. Con estos cambios tecnológicos los alumnos reclaman la presencia online del docente. En este sentido se podría pensar que contar con instrumentos automatizados e interactivos como es el caso de los programas robots conversacionales o *chatbots*. La inclusión de este tipo de tecnologías como soporte al proceso de enseñanza-aprendizaje resultará claramente beneficioso para el alumnado y profesorado.

Este proyecto plantea el diseño y construcción de *chatbots* educativos como un elemento de ayuda y soporte a la labor docente del profesorado universitario de la Facultad de Ingeniería de la Universidad Nacional de Jujuy. La motivación viene dada por la necesidad de contar con un mecanismo automatizado que permita satisfacer rápidamente los requerimientos de los alumnos relacionados con

resolver dudas y consultas de una asignatura específica en el momento que lo precise.

Palabras clave: *ChatBot; Entorno virtual; Comunicación virtual; PLN; Interacción*

CONTEXTO

La línea de investigación de I+D+i aquí presentada nace como iniciativa del grupo de investigación VRAIn (Visualización y Recuperación Avanzada de Información) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy (UNJu), cuyo eje principal es la mejora del proceso de enseñanza de asignaturas teóricas en carreras tecnológicas. Puesto que el equipo de trabajo forma parte del plantel docente de las tres carreras informáticas de la Facultad de Ingeniería: Ingeniería Informática, Licenciatura en Sistemas y Analista Programador Universitario; somos conscientes del beneficio que representará para los alumnos contar con estos avances.

Desde el punto de vista de los recursos humanos, este proyecto contribuirá a mejorar la formación en temas de Inteligencia Artificial, Aprendizaje Automático y Procesamiento del Lenguaje Natural.

1. INTRODUCCIÓN

La incursión de las Tecnologías de la Información y Comunicación (TIC) en la educación modificó las prácticas de enseñanza tradicionales, dando apertura a modalidades educativas no convencionales, a distancia, online, híbridas y a mecanismos de interacción tanto síncronos y asíncronos donde el docente y estudiante mantienen encuentros en entornos virtuales distintos al aula (Islas Torres, 2015). Este cambio posibilitó la aparición de conceptos como *b-learning* (*blended learning*, aprendizaje combinado), una modalidad de enseñanza-aprendizaje prometedora puesto que el híbrido de presencialidad y virtualidad incrementa las opciones para la construcción de conocimiento. En ese mismo sentido, como una aplicación de este nuevo concepto, se planteó un cambio de lugar entre las dos principales actividades del modelo tradicional: “los deberes en clase y las lecciones en casa”, que pasó a denominarse Aula Invertida (Inverted Classroom, Classroom Flip, Flipped Classroom, Flipped Learning) (Baker, 2000; Lage, Platt, & Treglia, 2000; Lacleta, Luisa, Fidalgo Blanco, & García-Peñalvo, 2015).

El aspecto común de estas nuevas denominaciones es que las clases se sustituyen por videos, lecturas o cualquier otra actividad que el alumnado pueda hacer de forma independiente fuera del horario de clase y el aula se transforma en un punto de encuentro donde se realizan actividades prácticas y de forma cooperativa. Si bien no se puede afirmar que todas las clases de las carreras impartidas en la UNJu adoptan este nuevo concepto de forma total (aula invertida), es necesario destacar la importancia y relevancia del uso del Aula Virtual que al menos la Facultad de Ingeniería y

sobre todo las carreras tecnológicas, han implementado desde hace años (desde el año 2011 con UnjuDigital¹ y desde el año 2015 con VirtualUnju², actualmente UnjuDigital está migrando a VirtualUnju como única plataforma oficial de educación a distancia).

El empleo del Aula Virtual funciona brindando soporte a la labor realizada por los docentes, principalmente como un repositorio de material educativo y como un entorno virtual que posibilita la comunicación y la interacción docente-alumno. El Aula Virtual aporta ciertas ventajas frente al modelo de enseñanza tradicional, por ejemplo, elimina las barreras espacio-temporales, ofrece agilidad y rapidez para solventar dudas, permite acceder a los tutores sin necesidad de desplazamientos, presenta un medio de trabajo colaborativo entre los alumnos y claramente actúa de facilitador para acceder a la información.

Si bien el Aula Virtual ofrece espacios para interactuar (foros de conversación por ejemplo), el alumno prefiere expresar su duda o consulta a través del mensaje privado, se ha detectado que al alumno le cuesta iniciar el debate/consulta/pregunta, principalmente por la timidez y por el miedo a quedar expuesto.

Por otro lado, a pesar de contar con estos entornos virtuales, el alumno reclama la presencia online del docente más allá del correo electrónico, los foros o mensajería instantánea. En este sentido se puede apelar a mecanismos más interactivos y eficientes que pueden operar con cierta autonomía con una amplia disponibilidad de acceso, como los programas robots conversacionales, o *chatbots* (Batista, 2017).

¹ <http://www.unjudigital.unju.edu.ar/>

² <https://virtual.unju.edu.ar/>

Los asistentes conversacionales o mejor conocidos como *bot* o *chatbot* o *chatterbot*, al que se puede definir como un asistente virtual, es un conjunto de programas informáticos que posee la habilidad de mantener una conversación con un ser humano mediante el lenguaje natural (Abdul-Kader & Woods, 2015). La inclusión de este tipo de tecnologías como soporte al proceso de enseñanza-aprendizaje resultará claramente beneficioso para el alumno.

El uso de *chatbots* dentro el ámbito escolar no es nuevo, han habido algunos acercamientos con resultados muy satisfactorios (Jia, 2009; Kerly, Ellis, & Bull, 2008; Kerly, Hall, & Bull, 2007; Hiremath, Hajare, Bhosale, Nanaware, & Wagh, 2018; Clarizia, Colace, Lombardi, Pascale, & Santaniello, 2018), y en cada uno de ellos se destaca el propósito de cada *chatbot*, es decir, cada *chatbot* es específico de un área de conocimiento o el menos si está dedicado a una sola área, los resultados mejoran notablemente, esto responde al hecho de que las técnicas empleadas para proporcionar el dialogo y sobre todo la respuesta a una pregunta dada, mejoran cuando se trata de un ámbito específico en contraposición a un *chatbot* de propósito genérico.

En una etapa inicial se planea realizar la construcción de un *chatbot* para una asignatura de la carrera Analista Programador Universitario, la asignatura en cuestión cuenta con alrededor de 80 alumnos y posee abundante material teórico. En este sentido se plantea automatizar las respuestas a dudas o consultas que el estudiante pueda tener en cualquier momento, puesto que estas herramientas se encuentran operativas las 24hs del día. La asignatura a evaluar posee mucho contenido

teórico y conceptual, por tanto es una candidata ideal para este tipo de experimentos en donde el Procesamiento del Lenguaje Natural (PLN) se realiza para hallar relaciones semánticas entre la pregunta realizada y la respuesta que eventualmente brindaría la herramienta.

La herramienta que se propone desarrollar se encuadra en el campo de la Inteligencia Artificial, específicamente lo que corresponde a PLN (Manning, 1999), puesto que un *chatbot* interactúa mediante el lenguaje natural, este tipo de técnicas resultan adecuadas para lograr una correcta aproximación.

La construcción del primer prototipo se realizará íntegramente en Python³, empleando la librería *Gensim* (Rehurek & Sojka, 2010) que posee implementaciones de los algoritmos más comúnmente utilizados para PLN. En este caso se seleccionaron *Term Frequency - Inverse Document Frequency* (TF-IDF) (Aizawa, 2003), *Latent Semantic Indexing/Latent Semantic Analysis* (LSI/LSA) (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997) y *Word Mover's Distance* (WMD) (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013; Kusner, Sun, Kolkin, & Weinberger, 2015)

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo tiene como principal línea de investigación el diseño, desarrollo, implementación y puesta en marcha de *chatbots* para el ámbito académico. Claramente es un trabajo interdisciplinario pues requiere de conocimiento en programación, desarrollo web, bases de datos, manejo de técnicas de

³ <https://www.python.org/>

aprendizaje automático, procesamiento del lenguaje natural, entre otras.

Los continuos avances tecnológicos, la generación de información y la abundancia de dispositivos hiperconectados requieren la inclusión de herramientas e instrumentos como los que se proponen con este proyecto. Por esta razón, la inclusión de un *chatbot* dentro de una asignatura, representará un avance importante en el proceso de enseñanza-aprendizaje.

3. RESULTADOS OBTENIDOS/ESPERADOS

Tras el desarrollo del presente proyecto se espera:

- Poder evaluar el compromiso de los alumnos con el uso responsable de la herramienta.
- Analizar el comportamiento de los alumnos identificando temáticas, dudas o preguntas similares que permitan establecer ciertos patrones, a partir de la información almacenada al momento de utilizar la herramienta.
- Definir los lineamientos tecnológicos para la correcta implementación y despliegue de un *chatbot* para una asignatura.
- Establecer el formato de la información a ser utilizada y procesada por los modelos de predicción diseñados.
- Establecer el protocolo de reentrenamiento de los modelos de predicción diseñados.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática presentada, todos los miembros del proyecto imparten docencia en carreras informáticas, con lo cual, el desarrollo de este proyecto brinda un marco para que tanto los docentes como los estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico.

En el transcurso del proyecto se tiene como objetivo consolidar la formación en investigación de los integrantes con menos antecedentes en proyectos y también está contemplado que los alumnos del último año de las carreras informáticas de grado, pueden optar por realizar su trabajo fin de grado con temáticas afines.

5. BIBLIOGRAFÍA

- Abdul-Kader, S. A., & Woods, J. C. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications*, 6(7).
- Aizawa, A. (2003). An information-theoretic perspective of TF-IDF measures. *Information Processing & Management*, 39(1), 45-63.
- Baker, J. W. (2000). The 'Classroom Flip': Using Web Course Management Tools to Become the Guide by the Side en J. A. Chambers (Ed.). *11th International Conference on College Teaching and Learning*, (págs. 9-17). Jacksonville, FL.
- Batista, A. (2017). Uso de chatbots como apoyo para la comunicación en el Aula. Un asistente virtual 24x7x365 colaborando con el curso. *4º Jornadas de TIC e Innovación en el Aula*. La Plata.
- Clarizia, F., Colace, F., Lombardi, M., Pascale, F., & Santaniello, D. (2018). Chatbot: An Education Support System for Student. *Cyberspace Safety and Security*, (págs. 291-302).

- Deerwester, S., Dumais, S., Furnas, G., Landauer, T., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6), 391.
- Hiremath, G., Hajare, A., Bhosale, P., Nanaware, R., & Wagh, K. S. (2018). Chatbot for education system. *International Journal of Advance Research, Ideas and Innovations in Technology*, 4(3), 37-43.
- Islas Torres, C. (2015). La interacción en el b-learning como posibilitadora de ambientes de aprendizaje constructivistas: perspectiva de estudiantes. *Pixel-Bit. Revista de Medios y Educación*(47), 7-22.
- Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems*, 22(4), 249-255.
- Kerly, A., Ellis, R., & Bull, S. (2008). CALMsystem: a conversational agent for learner modelling. *Applications and Innovations in Intelligent Systems XV* (págs. 89-102). Springer.
- Kerly, A., Hall, P., & Bull, S. (2007). Bringing chatbots into education: Towards natural language negotiation of open learner models. *Applications and Innovations in Intelligent Systems XIV*, (págs. 179-192).
- Kusner, M., Sun, Y., Kolkin, N., & Weinberger, K. (2015). From word embeddings to document distances. *International Conference on Machine Learning*, (págs. 957-966).
- Lacleta, S. E., Luisa, M., Fidalgo Blanco, Á., & García-Peñalvo, F. J. (2015). Metodología de enseñanza inversa apoyada en b-learning y gestión del conocimiento. *III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad*, (págs. 464-468).
- Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment. *The Journal of Economic Education*, 31(1), 30-43.
- Landauer, T., & Dumais, S. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, 104(2), 211.
- Manning, C. D. (1999). *Foundations of statistical natural language processing*. MIT press.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Neural Information Processing Systems*, 3111-3119.
- Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*.