

CONVERSATIONAL AGENT TO ASSIST IN SOLVING MATHEMATICS PROBLEMS

AGENTE CONVERSACIONAL PARA AUXILIAR EN LA RESOLUCIÓN DE PROBLEMAS DE MATEMÁTICA

José F. F. Cairello ¹

¹ Utec, Rivera / República Oriental del Uruguay

jffontoura@gmail.com

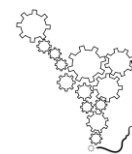
Recibido: 24 Septiembre 2023 / Aceptado: 20 Febrero 2024 / Publicado: 31 Marzo 2024

Abstract. Given a growing interest in identifying potential uses of tools based on Artificial Intelligence (AI) in the educational field, the need arises to propose solutions focused on improving teaching and learning processes. The use of conversational agents is then presented as a possible answer. The present study pursues the objective of developing an AI-based conversational agent (CA) whose function is to assist student teachers in solving a mathematical problem. As a methodology, an experiment was designed to test the ability of AC to assist students during the problem resolution stage and the results of two groups of students are contrasted: with and without access to AC. Some encouraging indicators are identified in relation to the potential of CAs as virtual assistants when solving math problems.

Keywords: Education, Mathematics Teaching, Artificial Intelligence, ChatBot, Natural Language Processing, Conversational Agent, Problem Solving.

RESUMEN. Ante un creciente interés en identificar usos potenciales de herramientas basadas en Inteligencia Artificial (IA) en el ámbito educativo surge la necesidad de proponer soluciones enfocadas a mejorar los procesos de enseñanza y aprendizaje. La utilización de agentes conversacionales se presenta entonces como posible respuesta. El presente estudio persigue el objetivo de desarrollar un agente conversacional (AC) basado en IA cuya función es asistir a estudiantes de magisterio en la resolución de un problema de matemática. Como metodología, se diseñó un experimento para poner a prueba la capacidad del AC de asistir a los estudiantes durante la etapa de resolución del problema y se contrastan los resultados de dos grupos de alumnos: con y sin acceso al AC. Se identifican algunos indicadores alentadores en relación al potencial de los AC como asistentes virtuales al momento de resolver problemas de matemática.

Palabras Clave: Educación, Enseñanza de Matemática, Inteligencia Artificial, ChatBot, Procesamiento del Lenguaje Natural, Agente Conversacional, Resolución de problemas.



1 INTRODUCCIÓN

1.1 JUSTIFICACIÓN

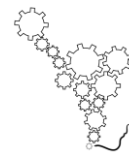
La vida en las sociedades del siglo XXI experimenta cambios de forma acelerada, en buena medida impulsados por los avances científicos y el desarrollo de nuevas tecnologías. Actualmente, en el mapa del desarrollo científico se destaca el campo de la Inteligencia Artificial (IA). Al decir de Russell y Norvag, desde hace muchísimos años la humanidad ha intentado entender cómo pensamos. Actualmente, con el desarrollo de la inteligencia artificial, va más lejos aún, esforzándose no sólo en comprender, sino en construir entidades inteligentes (Russell y Norvag, 2004).

La IA comprende así “el desarrollo de máquinas que tienen cierto nivel de inteligencia, con la capacidad de realizar funciones similares a las humanas, incluidas las cognitivas, el aprendizaje, la toma de decisiones y la adaptación al entorno” (Chen, L. et al., 2020). Su presencia recurrente en los titulares de los medios masivos de comunicación; con noticias relacionadas al desarrollo de vehículos autónomos, máquinas capaces de realizar precisos diagnósticos de enfermedades, el reconocimiento facial, traducciones en tiempo real y un largo etcétera de soluciones a problemas en variados contextos o respuestas más eficientes a viejos problemas, dan cuenta de su relevancia actual y en un futuro cercano.

En ese escenario es de esperar que el desarrollo de la IA, impacte también en el campo de la educación. A decir de Sandoval, la relación entre aprendizaje, revolución digital e IA demandará a los educadores respuestas a nuevas necesidades, intereses y habilidades emergentes que presenta el alumnado (Sandoval, E., 2018).

Por esa razón resulta necesario indagar acerca de las oportunidades que la IA puede brindarnos en el campo de la enseñanza y aprendizaje; y cuestionar, desde el rol de educador cómo puede ponerse al servicio de la enseñanza y los aprendizajes y colaborar con la tarea profesional docente.

En el terreno de la Educación existe un amplio consenso de aprovechar las herramientas y técnicas de la IA para desarrollar el potencial de los sistemas educativos con el objetivo de formar ciudadanos críticos y capaces de solucionar problemas del siglo XXI.



El denominado Consenso de Beijing sobre IA y Educación, coorganizado por la Unesco y con la participación de más de 100 naciones en mayo de 2019; incluye entre sus recomendaciones “planificar y aplicar estrategias coherentes a nivel de todo el sistema para el uso de la inteligencia artificial en la educación que estén armonizadas e integradas con las políticas educativas” así como “tener presentes las necesidades de inversión para la implantación de la inteligencia artificial en las políticas y los programas educativos” (UNESCO, 2019).

Una de las áreas en que es posible avanzar en la introducción de la IA en el mundo de la educación es la introducción de agentes conversacionales o chatbots que asuman tareas con menor demanda de habilidades intelectuales y emocionales para permitir mientras tanto a los profesionales del área implicarse de lleno en otros menesteres y de ese modo avanzar en una educación personalizada; capaz de atender las necesidades individuales de cada estudiante con el apoyo de estas tecnologías.

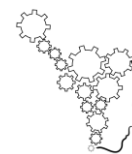
Las herramientas virtuales basadas en IA, son potencialmente capaces de realizar las tareas simples que se les encomienda de forma muy eficiente pues están permanentemente disponibles, cualquier día y a cualquier horario y no son susceptibles ante consultas que se reiteran o al desgaste natural que suele provocar en las personas la atención al público.

En particular, son cada vez más, capaces de comprender el lenguaje natural y desarrollar hasta cierto punto el diálogo con las personas sin que se requiera para ello ajustarse a los estrictos cánones de los lenguajes formales. Eso nos lleva a pensar en este tipo de herramientas y en su utilización con el objetivo de colaborar en las tareas propias de la docencia.

1.2 CARACTERIZACIÓN DEL PROBLEMA

En momentos en que la nueva sociedad del conocimiento pone en cuestión los roles al interior de la educación (Paur, A. et al., 2006), surge un creciente interés en identificar usos potenciales de las herramientas basadas en IA en el ámbito educativo.

El presente trabajo pretende avanzar en dicha indagación y se propone, desde el terreno de la Matemática, obtener respuestas a la siguiente pregunta de investigación: ¿de qué manera un tutor virtual, basado en IA, puede emular el andamiaje que realiza el docente durante actividades de resolución de problemas abiertos de matemática?



La investigación tiene como tema los AC basados en IA en el contexto de la enseñanza y aprendizaje de matemáticas. Para delimitar el alcance del presente estudio, la hipótesis de trabajo que estaremos sometiendo a prueba es que los estudiantes con acceso a un AC logran mejores resultados ante un problema de matemática que los obtenidos por estudiantes que no tienen acceso al mismo.

1.3 OBJETIVOS

1.3.1 Objetivo general

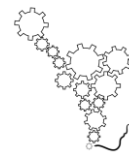
Desarrollar un agente conversacional basado en IA para asistir a los estudiantes en la resolución de un problema abierto de matemática.

1.3.2 Objetivos secundarios

- Diseñar e implementar un agente conversacional a través de Dialogflow, que asista a los estudiantes mientras resuelven el problema titulado Supercuadrados.
- Entrenar al agente y obtener información acerca de las interacciones que realizan los estudiantes cuando resuelven un problema específico.
- Evaluar el desempeño del agente e indagar acerca de la pertinencia del uso de agentes virtuales en la enseñanza de la matemática.
- Obtener retroalimentación de la forma en que piensan los estudiantes cuando se enfrentan a un problema.
- Establecer una posible hoja de ruta para el desarrollo de nuevos AC.

2. MÉTODO

La presente investigación es de naturaleza mixta (cuali-cuantitativa) del tipo experimental y aplicada con relación a los objetivos propuestos, de acuerdo con Gil (2002). El público implicado directamente en la investigación es el de los estudiantes de primer año de magisterio en el Instituto de Formación Docente del departamento de Rivera, Uruguay. Los resultados de la investigación estarán al servicio de docentes y estudiantes de matemática en cursos de graduación.



La investigación se divide en tres etapas:

- Desarrollo del AC
- Investigación Experimental
- Análisis del experimento

2.1 DESARROLLO DEL AGENTE CONVERSACIONAL

En la presente investigación se implementa un agente conversacional utilizando la plataforma de comprensión de lenguaje natural *Dialogflow ES*¹ que es una herramienta presente en el ecosistema de *Google Cloud Platform*. La elección de *Dialogflow ES* responde a los siguientes motivos:

- a) Es posible su utilización sin costo dentro de los fines de esta investigación.
- b) Se integra fácilmente con los sistemas de gestión de aprendizajes LMS² que utilizan docentes y estudiantes en sus cursos en Uruguay.
- c) Dispone de buena cantidad de material de referencia (guías y tutoriales) sobre el funcionamiento de la herramienta en comparación con otras plataformas similares.

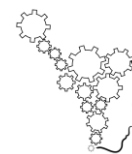
Debe tenerse en cuenta que este trabajo pretende no solamente avanzar en la exploración acerca del potencial de una herramienta de este tipo en la educación sino también mostrar una posible hoja de ruta para que otros docentes puedan eventualmente desarrollar sus propios AC y utilizarlos en sus prácticas pedagógicas. Por eso la opción por una herramienta accesible, sólida al momento de interpretar y procesar el lenguaje natural y con una curva de aprendizaje no tan pronunciada.

Para la definición de los *Intents* y para el entrenamiento previo del AC se requirió la colaboración de una docena de experimentados docentes de matemática a los que se les pidió resolver el problema de los “supercuadrados” y realizar un análisis *a priori* de las posibles estrategias de resolución y dificultades que podrían surgir en la resolución del problema. En el grupo de colaboradores se logró incluir a tres inspectores³ de matemática en enseñanza

¹ <https://dialogflow.cloud.google.com>

² Sistema de Gestión de Aprendizajes (Learning Management System)

³ Inspectores de Matemática de Educación Secundaria y Técnico Profesional. Son la principal referencia en la enseñanza de la asignatura en el organigrama del sistema de educación de Uruguay.



media, tres articuladores⁴ de la asignatura de diferentes departamentos de Uruguay, cuatro referentes del Departamento de Matemática de Ceibal⁵ y dos formadores de formadores⁶. Se le proporcionó a cada uno de los docentes una plantilla (documento) en el que debían redactar:

(A) las preguntas que a su juicio podrían llegar a formular los estudiantes en relación al problema planteado;

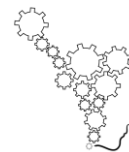
(B) las respuestas sugeridas a esas preguntas bajo la consigna de desbloquear al estudiante sin cerrar el problema;

Seguidamente el investigador realizó manualmente una clasificación de las preguntas sugeridas por los docentes agrupándolas de acuerdo a la intención de cada pregunta. El siguiente cuadro muestra el resultado de dicha clasificación con el conjunto de diecinueve intenciones definidas para el agente:

⁴ Los PAD (Profesores Articuladores Docentes) colaboran con los inspectores en la orientación a los docentes en su jurisdicción.

⁵ Ceibal es un proyecto socioeducativo que promueve la inclusión digital.

⁶ Docentes de educación terciaria: formación de maestros (Ed. Inicial y Primaria) y profesores (Ed. Media).



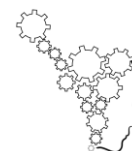
CUADRO 1 - *INTENTS*

Saber cómo se resuelve	Informar que no ha recibido respuesta adecuada
Ver ejemplos de cuadrados perfectos	Saber qué es un cuadrado perfecto
Saber qué es un supercuadrado	Saber cuántas cifras posee un supercuadrado
Saber si la respuesta es correcta	Saber qué es un número natural
Ver ejemplo de SC	Saber qué son cifras consecutivas
Saber cómo empezar	Saber la respuesta
Default Fallback Intent	Agradecer
Informar el nombre	Default Welcome Intent
Ver otro ejemplo de SC	Pedir ayuda - pista
Saber si son infinitos	

FUENTE: AUTOR (2021).

El paso siguiente fue la redacción, por parte del investigador, de un conjunto de frases de entrenamiento para cada *intent*. La inteligencia artificial presente en *Dialogflow* permite que a partir de unas cuantas expresiones etiquetadas bajo un mismo *intent*, se puedan reconocer otras frases que persiguen la misma intención de usuario. En el entrenamiento básico de nuestro agente utilizamos más de 10 frases de entrenamiento para cada intención incluyendo naturalmente las sugerencias proporcionadas por el grupo de expertos y otras provenientes de la experiencia del propio investigador.

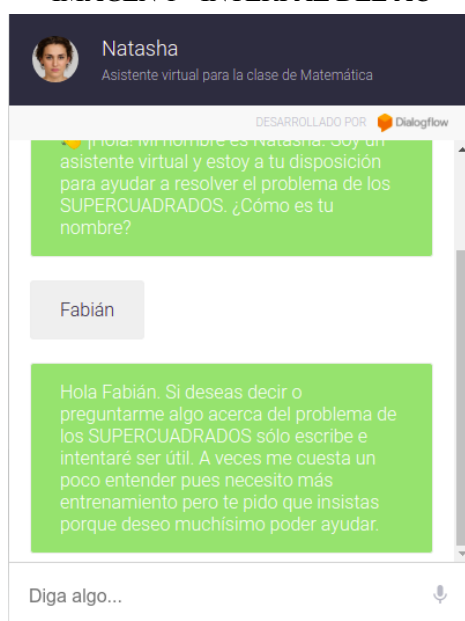
En la etapa siguiente del desarrollo del AC se procedió a la identificación y extracción dentro de los mensajes de usuario final de determinados parámetros pertenecientes a ciertas categorías llamadas *entities* (entidades). Las entidades son parámetros, “datos estructurados que se pueden usar fácilmente para realizar alguna lógica o generar respuestas” (Google Cloud, 2021). Por ejemplo, en determinado entorno de ejecución de nuestro AC se obtiene el nombre del estudiante, parámetro que pertenece a una entidad de sistema (predefinida) llamada *sys.given-name*. Pero para otros parámetros fue necesario definir nuevas entidades como la de los “supercuadrados”. Otros ítemes que se debieron definir durante el desarrollo con *Dialogflow ES* fueron los contextos, de entrada y salida, del AC. Los contextos de entrada y salida dirigen los flujos de la comunicación permitiendo la activación de determinados *intents* y manteniendo activa determinada



información por el tiempo que dura ese contexto, medido en cantidad de interacciones. En el anexo 1, se muestra un cuadro con ejemplos de preguntas y respuestas para cada *intent* del AC, así como los contextos de entrada y salida y parámetros correspondientes.

Completadas las etapas anteriores se ejecutó el primer entrenamiento del AC y se realizaron pruebas y ajustes a su funcionamiento. Finalmente, para permitir el acceso de los estudiantes al AC se realizó su integración a una sección de la plataforma del curso (CREA-Schoology) a través de código HTML. Completadas las etapas anteriores se ejecutó el primer entrenamiento del AC y se realizaron pruebas y ajustes a su funcionamiento. Finalmente, para permitir el acceso de los estudiantes al AC se realizó su integración a una sección de la plataforma del curso (CREA-Schoology) a través de código HTML. En la imagen 1 se muestra un detalle de la interfaz del AC del modo en que ven la misma los estudiantes⁷.

IMAGEN 1 - INTERFAZ DEL AC



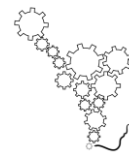
CAPTURA DE PANTALLA DE LA INTERFAZ DEL AC BAUTIZADO COMO “NATASHA”. LA IMAGEN DE AVATAR DE NATASHA FUE CREADA ARTIFICIALMENTE EN *GENERATED PHOTOS*⁸. FUENTE: AUTOR (2021).

2.2 INVESTIGACIÓN EXPERIMENTAL

⁷ Es posible interactuar con el AC de este proyecto en:

<https://sites.google.com/view/supercuadrados/asistente-virtual?authuser=0>

⁸ <https://generated.photos/>



En la etapa de investigación experimental se adoptó un enfoque que se caracteriza por el diseño de un experimento, ejecución y medición de sus resultados (Somekh, L., 2015). El estudio se presenta a partir del interrogante: ¿de qué manera un tutor virtual basado en IA puede emular el andamiaje que realiza el docente durante actividades de resolución de problemas abiertos de matemática?

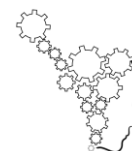
Para intentar responder a esa pregunta diseñamos el experimento que consiste en proponer un mismo problema a resolver a los 220 estudiantes que se encuentran cursando primer año de la carrera de magisterio en el Instituto de Formación Docente (IFD) de Rivera (Uruguay). A la mitad de los estudiantes, elegidos al azar, se les proporciona acceso a un agente conversacional pre entrenado para asistirlos en la resolución del problema. A los restantes solamente la consigna:

“A un número natural N , mayor que 10, lo llamaremos supercuadrado si el número formado por cada dos cifras consecutivas de N es siempre un cuadrado perfecto. Por ejemplo, 8164 es “supercuadrado” porque los números 81, 16 y 64 son cuadrados perfectos. Otros ejemplos de supercuadrados son 25 y 649. a) ¿cuántos supercuadrados hay? b) ¿cuál es el mayor supercuadrado?” (Banco de Preguntas OBMEP, 2013).

El diseño del experimento corresponde al de un *posttest* (problema propuesto) donde un grupo recibe el tratamiento (acceso al AC) y un grupo de control no lo recibe y sirve de referencia para identificar posibles variaciones en el primer grupo. Debido a las restricciones a la presencialidad a raíz de la pandemia por Covid-19, se optó por fijar una fecha (19 de junio de 2021) y solicitar a los estudiantes que participen de la investigación en el transcurso de ese día a través de la plataforma (LMS⁹) que habitualmente utilizan en el dictado del curso de matemática (CREA-Schoology). En la plataforma todos los estudiantes matriculados y con alguna actividad en los 30 días previos pueden disponer, tal cual se muestra en la Imagen 2, de los siguientes elementos:

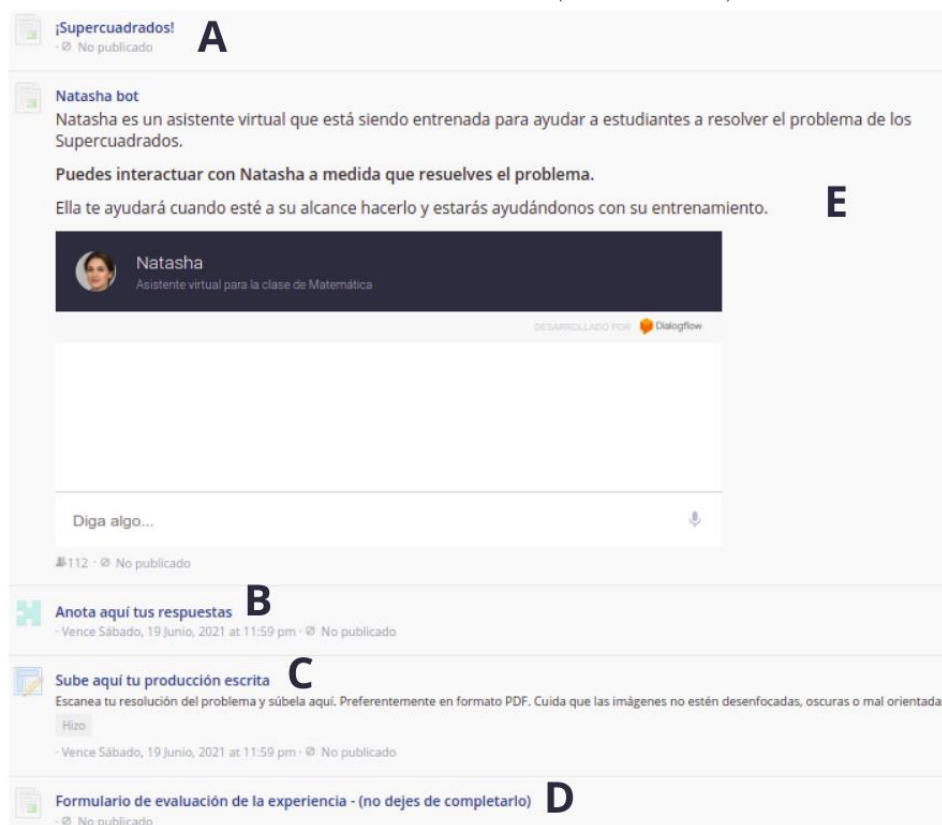
- A. consigna del problema,
- B. cuestionario autocorregible para anotar las respuestas,
- C. espacio para subir archivo con el detalle de la producción escrita,

⁹ Learning Manage System (Sistema de gestión de aprendizajes)



- D. encuesta acerca de la experiencia
- E. acceso al AC (sólo los estudiantes habilitados).

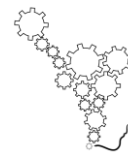
IMAGEN 2 - CAPTURA DE PANTALLA DE LA ACTIVIDAD SUPERCUADRADOS EN PLATAFORMA CREA (SCHOOLGY)



FUENTE: AUTOR (2021)

La selección de los estudiantes con y sin acceso al AC se hace al interior de cada grupo de alumnos (7 grupos en total) ordenando alfabéticamente a los estudiantes y asignando el acceso en forma alternada. Ningún estudiante sabía en forma previa si tendría o no acceso al AC. Se tiene especial cuidado en que ambos grupos sean estadísticamente equivalentes y no ocurran sesgos por características de origen en algunos de los grupos. Por ejemplo, una de las clases está formada únicamente por estudiantes que recursan primer año, lo que llevó a la decisión de sortear al interior de cada clase y no clases enteras con y sin acceso al AC.

Una de las características distintivas del enfoque experimental es la manipulación de la variable independiente para observar su efecto en las variables dependientes que deben ser medidas o cuantificadas de algún modo. La variable independiente en esta investigación



es el acceso de los estudiantes al agente conversacional. Las variables dependientes a observar en cada grupo, con y sin acceso al AC, se presentan en el siguiente cuadro en el que se incluyen además los instrumentos o técnicas utilizados para la recogida de los datos.

CUADRO 2 - RELACIÓN DE ANÁLISIS ENTRE VARIABLES E INSTRUMENTOS.

<i>variable dependiente a observar al interior de cada grupo experimental</i>	<i>instrumentos para la obtención de los datos</i>
Respuestas correctas e incorrectas al problema	<ul style="list-style-type: none">● cuestionario/evaluación CREA● producción escrita de los estudiantes
Tiempo empleado por los estudiantes que resolvieron correctamente el problema	<ul style="list-style-type: none">● cuestionario/evaluación CREA
Valoración de la experiencia de los estudiantes con el AC	<ul style="list-style-type: none">● encuesta (formulario de Google)

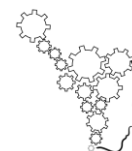
FUENTE: AUTOR (2021)

2.3 ANÁLISIS DEL EXPERIMENTO

A partir de los resultados obtenidos por cada grupo (con y sin acceso al AC) en relación al porcentaje de respuestas correctas al problema, tiempo empleado para resolver el problema y valoración de la experiencia con el AC, se realizó un análisis utilizando parámetros estadísticos de tendencia central y dispersión de los datos así como técnicas de prueba de hipótesis (test Z para comparación de medias y de proporciones) para comparar los resultados de ambos grupos e inferir posibles conclusiones acerca del desempeño del AC en cuestión.

Cabe aún señalar que en la etapa de validación del AC se realizó también un análisis del funcionamiento del AC a partir del historial de interacciones (registro que proporciona *Dialogflow*) de los estudiantes con acceso al agente, verificando manualmente cada interacción, según el criterio asumido por el investigador, procediendo a etiquetarla de acuerdo a una de las siguientes categorías:

- *bien clasificado*: el AC interpreta correctamente la intención del usuario final



- *mal clasificado*: el AC ha clasifica el mensaje con una intención diferente a la que corresponde de acuerdo al criterio del investigador (error atribuible al AC)
- *inclasificable*: el investigador no consigue decodificar el mensaje

Se pretende así contar con otros indicadores acerca del funcionamiento del AC para evaluar el nivel de acierto en las intenciones de usuario (*intents*), la pertinencia del entrenamiento inicial realizado a partir de la información proporcionada por el grupo de expertos docentes y la necesidad eventual de incluir nuevas respuestas a intenciones aún no especificadas en el chatbot para futuros entrenamientos y eventuales reutilizaciones del mismo.

3 RESULTADOS

3.1 VALIDACIÓN DEL AC CON LOS ESTUDIANTES

Se realizó la evaluación del AC a partir de los siguientes indicadores:

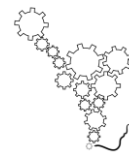
I. Comparación de las proporciones de respuestas correctas e incorrectas

Fueron habilitados todos los estudiantes matriculados en los cursos que habían ingresado al menos una vez a la plataforma Crea en los últimos 30 días, lo que da un total de 223 estudiantes. Participaron en la experiencia 81 estudiantes, lo que representa un 36,3% del total de habilitados.

CUADRO 3 -PARTICIPACIÓN EN LA INVESTIGACIÓN

AC	Habilitados	Participaron
Si	113	40
No	110	41
<i>Total</i>	223	81

FUENTE: AUTOR (2021).



En relación a la cantidad de estudiantes que resolvieron correctamente el problema los resultados en ambos grupos fueron los siguientes: 60,0% y 51,2% de respuestas correctas en los grupos con el AC y sin el AC respectivamente.

CUADRO 4 - CANTIDAD DE RESPUESTAS CORRECTAS E INCORRECTAS

AC	Correcto	Incorrecto/No hizo	Total
Si	24	16	40
No	21	20	41
Total	45	36	81

Fuente: Autor (2021).

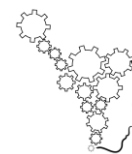
Se realiza una prueba de hipótesis (Test Z - diferencia de proporciones) para examinar si hay evidencia suficiente que permita afirmar que los resultados son mejores cuando se tiene acceso al AC. Eso equivale a decir que la proporción de aciertos en la muestra con el AC es significativamente superior a la proporción de aciertos en la muestra sin el acceso al AC. Verificadas las condiciones de aleatoriedad, normalidad e independencia de los datos; establecemos un nivel de significación $\alpha = 5\%$ y procedemos con el tratamiento estadístico correspondiente.

Establecemos la hipótesis nula H_0 : la proporción P_{AC} de aciertos en la muestra con acceso al AC es menor o igual a la proporción $P_{no AC}$ de aciertos en la muestra sin acceso al AC.

$$H_0: P_{AC} \leq P_{no AC}$$

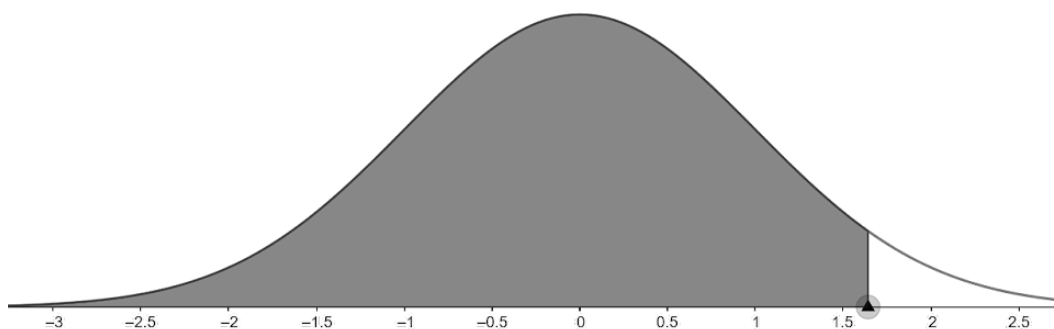
Establecemos la hipótesis alternativa H_1 :

$$H_1: P_{AC} > P_{no AC}$$



Fijamos la zona de rechazo de H_0 para el nivel de significación $\alpha = 5\%$ cuyo valor crítico es igual a 1,642.

GRÁFICO 1 - ZONAS DE ACEPTACIÓN (GRIS) Y RECHAZO DE LAS HIPÓTESIS A CONTRASTAR



FUENTE: AUTOR (2021).

Calculamos el estadístico de contraste:

$$z = \frac{P_{AC} - P_{no AC}}{\sigma_{P_{AC} - P_{no AC}}} \approx \frac{\frac{24}{40} - \frac{21}{41}}{\sqrt{\frac{\frac{45}{81} \cdot \frac{36}{81}}{40} + \frac{\frac{45}{81} \cdot \frac{36}{81}}{41}}} \approx 0,7951$$

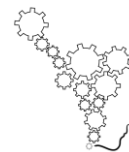
Como z se encuentra en la zona de aceptación ($z < 1,642$) no existe evidencia suficiente para rechazar H_0 y afirmar, con 95% de confianza, que estudiantes con acceso al AC obtuvieron mejores resultados que los estudiantes que no tenían acceso al AC.

II. Comparación del tiempo empleado para resolver el problema

En relación al tiempo utilizado para resolver el problema, la mediana del tiempo utilizado por los que contaban con el AC fue de 11 minutos y la del grupo que no contaba con el AC fue de 14 minutos. Eso equivale a decir que la mitad de los estudiantes con acceso al AC resolvieron el problema en menos de 11 minutos mientras que la mitad de los estudiantes sin acceso al AC lo hicieron en menos de 14 minutos.

CUADRO 5 - TIEMPOS PROMEDIO EN MINUTOS DE CULMINACIÓN DEL PROBLEMA DE ACUERDO AL ACCESO AL AC

AC	Si	No
----	----	----

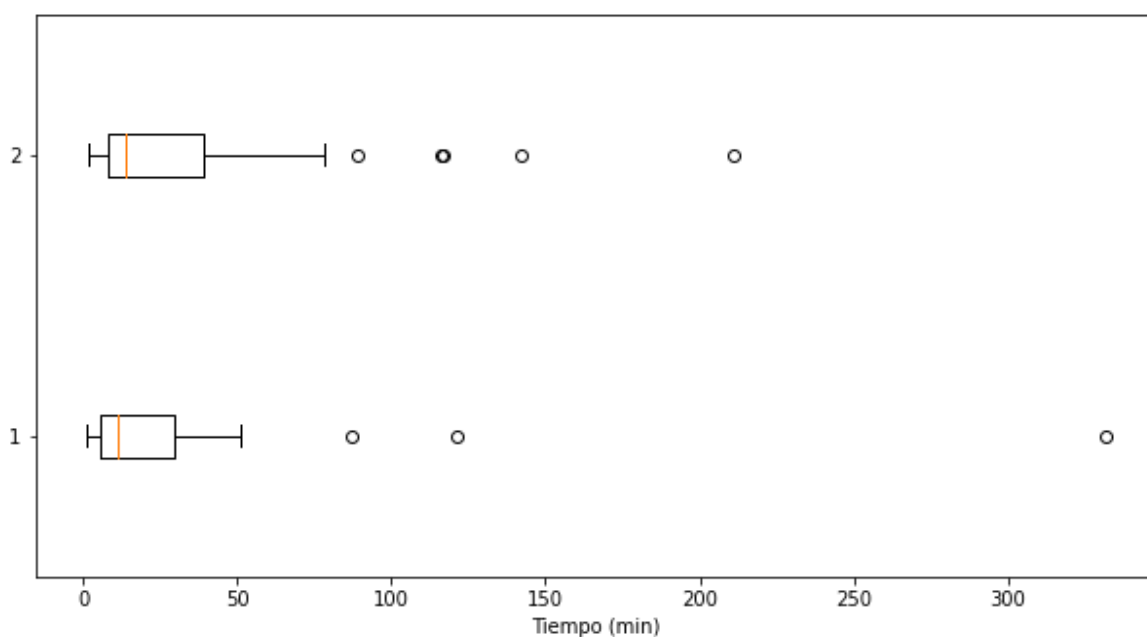


Promedio (min)	28,6	33,6
Desviación Estándar	54,8	45,2
Mediana	11	14
N	40	41

Fuente: Autor (2021).

El test Z - diferencia de medias aplicado a ambas distribuciones tampoco permite afirmar con 95% de confianza que el promedio de tiempo empleado para resolver el problema por parte de los estudiantes con acceso al AC es inferior al promedio de los que no tienen acceso al AC.

GRÁFICO 2 - DIAGRAMAS DE CAJA CORRESPONDIENTES A LOS TIEMPOS UTILIZADOS POR LOS ESTUDIANTES CON EL AC (1) Y SIN EL AC (2)

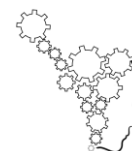


FUENTE: AUTOR (2021).

III. Valoración de los estudiantes

En cuanto a la valoración de los estudiantes acerca de la experiencia con el AC se realizó una encuesta que contestaron 32 de los 40 estudiantes con acceso al agente.

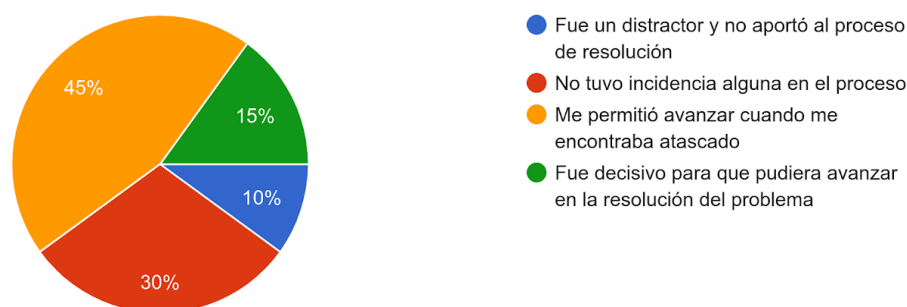
Uno de esos estudiantes manifestó no haber sentido necesidad de usar el AC y 11 de ellos (27,5%) indicaron que no supieron interactuar con el agente.



Entre los 20 restantes que llegaron a usar el AC, el 60% valoró positivamente la experiencia indicando que el agente le permitió avanzar cuando estaba atascado o que resultó decisivo para que pudiera resolver el problema (gráfico 3).

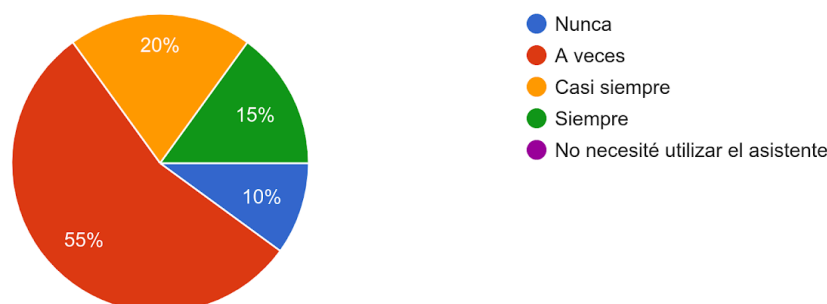
Sólo a veces las respuestas del asistente estaban adecuadas a lo preguntado según la mayoría (55%) de los encuestados. Para el 35% las respuestas siempre o casi siempre eran adecuadas mientras que para un 10% las respuestas nunca estaban adecuadas a lo preguntado (gráfico 4).

GRÁFICO 3 - CUÁL DE LAS SIGUIENTES AFIRMACIONES COINCIDE MÁS CON TU OPINIÓN ACERCA DE LA EXPERIENCIA RECIENTE CON EL ASISTENTE VIRTUAL



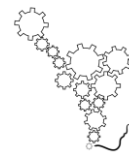
Fuente: Autor (2021).

GRÁFICO 4 - EN RELACIÓN A LAS RESPUESTAS OFRECIDAS POR EL ASISTENTE VIRTUAL: ¿ESTABAN ADECUADAS A LO PREGUNTADO?



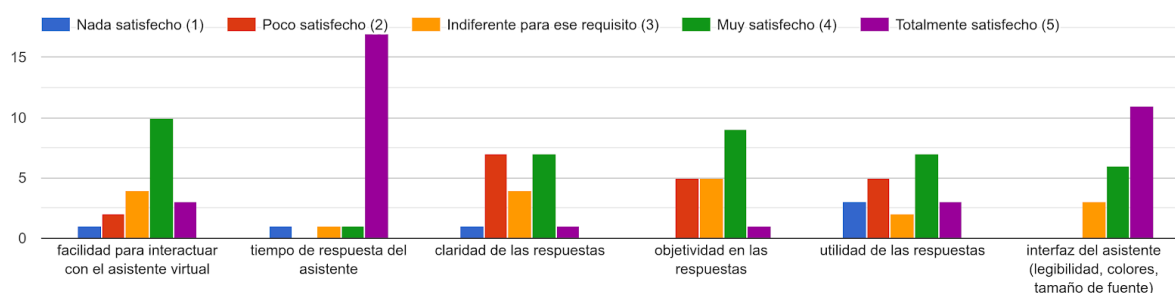
FUENTE: AUTOR (2021)

Al momento de calificar la experiencia con el asistente en relación a ciertas características particulares del mismo hubo una valoración positiva (muy satisfecho y totalmente satisfecho) superior a la valoración negativa (poco satisfecho y nada satisfecho).



Especialmente en relación al tiempo de respuesta, la interfaz, la facilidad de interacción y la objetividad en las respuestas. Los resultados se muestran más equilibrados en relación a la utilidad de las respuestas y particularmente en la claridad de las mismas que recoge la misma cantidad de valoraciones positivas y negativas (gráfico 5).

GRÁFICO 5 - SI PUDIERAS CALIFICAR TU EXPERIENCIA EN EL PRESENTE USO DEL ASISTENTE VIRTUAL CON UN PUNTAJE ENTRE 1 (NADA SATISFECHO) Y 5 (TOTALMENTE SATISFECHO). CONSIDERA LOS SIGUIENTES ASPECTOS Y EVALÚA DE MANERA SEPARADA:

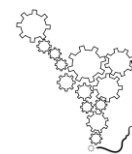


FUENTE: AUTOR (2021).

IV. *Análisis del historial de interacciones*

Del análisis del historial de interacciones de los estudiantes con el agente para evaluar el nivel de acierto en las intenciones de usuario (intents), la pertinencia del entrenamiento previo con los docentes y la necesidad de incluir valoración de los estudiantes acerca de la experiencia con el AC nuevas respuestas a intenciones no especificadas en el chatbot surge que hubo un total de 173 interacciones con el AC de las cuales el 76% fueron clasificadas por el agente como una de las intenciones con las que fue entrenado.

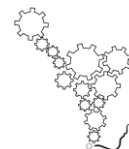
Los principales requerimientos de los estudiantes, según la clasificación del agente, fueron para saber qué es un supercuadrado (25%), saber si la respuesta era correcta o no (11%), saber qué es un cuadrado perfecto (10%) y saber la respuesta a las preguntas del problema (8%). Un 24% de las solicitudes no pudieron ser clasificadas por el agente.



CUADRO 6 - CANTIDAD DE SESIONES, INTERACCIONES Y PORCENTAJE DE SALIDAS (FIN DE SESIÓN) SEGÚN INTENTS

Intenciones	Sesiones	Interacciones	Salida %
Saber cómo se resuelve	1	1	2%
Ver ejemplos de cuadrados perfectos	10	13	4%
Saber qué es un supercuadrado	30	43	31%
Saber si la respuesta es correcta	16	19	11%
Ver ejemplo de SC	2	2	0%
Saber cómo empezar	1	1	0%
Default Fallback Intent	26	43	24%
informar el nombre	1	1	0%
Ver otro ejemplo de SC	1	1	0%
Saber si son infinitos	2	2	0%
Default Welcome Intent	2	2	0%
Pedir ayuda - pista	1	1	0%
Informar que no ha recibido respuesta adecuada	1	1	2%
Saber qué es un cuadrado perfecto	16	18	9%
Saber cuántas cifras posee un SC	4	7	4%
Saber qué es un número natural	2	2	2%
Saber qué son cifras consecutivas	1	1	2%
Saber la respuesta	13	14	11%
Agradecer	1	1	0%
Total	131	173	100%

FUENTE: DIALOGFLOW ANALYTICS (2021).



Para comprobar si el AC había realizado correctamente la clasificación de las intenciones de los estudiantes, realizamos una verificación manual de las 173 interacciones a partir del historial registrado en *Dialogflow*. De ese modo cada interacción fue etiquetada con uno de los rótulos de las siguientes categorías:

1) Bien clasificada (el agente identificó y clasificó correctamente la intención del estudiante),

2) Mal clasificada (el agente atribuyó una intención diferente a la intención real del estudiante, incluye aquellos casos en que tal intención no fue prevista en el desarrollo del AC) y

3) Inclasificable (el texto ingresado por el usuario no es inteligible o no guarda relación con la temática). Se obtuvieron los siguientes resultados:

CUADRO 7 - CLASIFICACIÓN DE LAS INTERACCIONES DE ACUERDO A LAS
DIFERENTES CATEGORÍAS

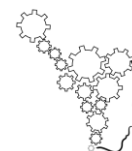
CATEGORÍA	NRO. INTERACCIONES	PORCENTAJE
inclasificable	29	16,76%
bien clasificado	108	62,43%
mal clasificado	36	20,81%
Total	173	100,00%

FUENTE: AUTOR (2021)

4 CONCLUSIONES Y TRABAJOS FUTUROS

4.1 CONCLUSIONES

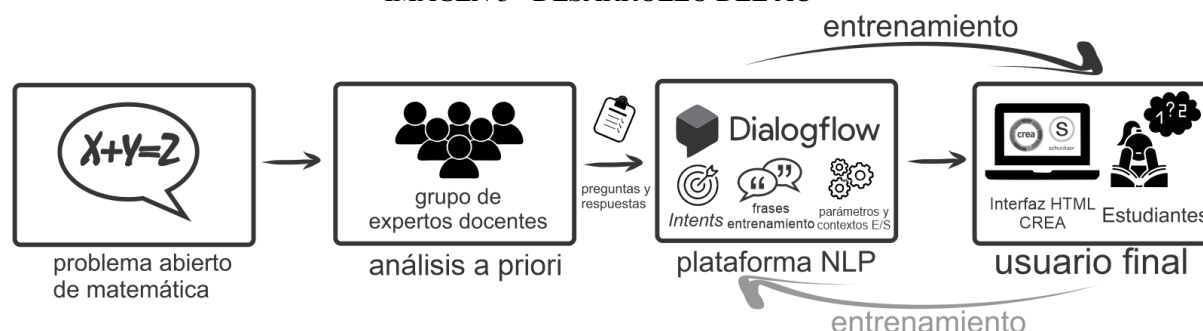
Cuando se analiza la proporción de respuestas correctas y el tiempo empleado por los estudiantes durante la resolución de un problema, a pesar que los resultados de los estudiantes que han utilizado el AC se presentan ligeramente mejores que los que no lo han hecho, no es posible como ya hemos mencionado, descartar que esos resultados sean producto meramente del azar.



A pesar de ello, hay varios indicadores, subjetivos y objetivos, que se presentan como alentadores en relación al potencial de los AC como asistentes virtuales al momento de resolver problemas de matemática. Empezamos destacando que seis de cada diez estudiantes que utilizaron el AC plantean que les permitió avanzar en momentos en que estaban atascados o fue decisivo para que pudieran resolver el problema.

Como mostramos antes, más del 62% de las frases de usuario final fueron bien clasificadas y resueltas por el AC. Cabe destacar que a partir de la puesta en práctica directamente con los estudiantes es razonable esperar que el agente continúe aprendiendo de las interacciones y mejore aún más su desempeño. Una revisión de las interacciones luego de una o más sesiones permitirá en el futuro clasificar mejor las intenciones de los estudiantes y eventualmente añadir nuevas intenciones en un proceso de aprendizaje supervisado.

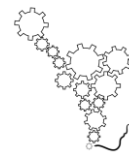
IMAGEN 3 - DESARROLLO DEL AC



FUENTE: AUTOR (2021)

De todos modos, la hoja de ruta (ver Imagen 3) presentada para el desarrollo y entrenamiento inicial de un AC cuya finalidad es asistir a estudiantes que resuelven un problema de matemática se ha mostrado efectiva. Básicamente, el camino consiste en el análisis a priori del problema, la redacción de preguntas y respuestas modelo, el agrupamiento de las preguntas y la definición de *intents*, la identificación de entidades y parámetros y la implementación del agente.

Un aspecto a considerar es que una parte de las preguntas realizadas por los estudiantes (17%) era inclasificable, imposible o de muy difícil decodificación por parte de un ser humano y probablemente también por los modelos de PLN. Eso nos lleva a la cuestión



acerca de que muchos estudiantes tienen dificultades no sólo en la formulación de preguntas adecuadas sino para expresarlas correctamente.

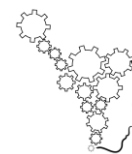
4.2 TRABAJOS FUTUROS Y ALGUNAS CONSIDERACIONES FINALES

Una diferencia importante, entre la tarea del asistente virtual y el trabajo que realiza un docente, es que éste no actúa solamente a demanda, a partir de las preguntas de los estudiantes, sino que al identificar a través de lenguaje gestual a un estudiante que aparenta estar bloqueado, realiza algún tipo de intervención para ayudarlo a continuar. Esto no resulta necesariamente positivo, pues algunas intervenciones a destiempo o el propio lenguaje corporal del docente pueden contribuir a cerrar el problema o privar a los estudiantes de transitar por determinados caminos de solución que en perspectiva podrían resultar enriquecedores. Sin embargo, puede pensarse en el futuro en un AC que no sólo actúa a demanda sino a partir de otro tipo de señales que sirvan de disparador.

Un aspecto que debió considerar esta investigación y no lo hizo fue la falta de habilidades de parte considerable de los estudiantes para interactuar con el AC a través de un chat ya que el 27,5% de los alumnos indicó no haber sabido utilizarlo.

Quizás el aspecto más importante a tener en cuenta en futuras investigaciones relacionadas sea el de poder contar con un ambiente más controlado. En esta oportunidad y debido a la pandemia por Covid-19 no se pudo trabajar presencialmente con los estudiantes. Aunque se trató de minimizar los riesgos a la investigación, el trabajo presencial hubiera garantizado una mayor participación de los estudiantes y acotado cualquier posibilidad de contacto de éstos entre sí o con otras personas durante el experimento, aspecto que en teoría pudo haber incidido en los resultados obtenidos.

Debe recordarse que esta investigación se centra en el desempeño de un AC entrenado, a partir de la opinión de un grupo de expertos para un problema específico: el de los supercuadrados, por lo que otro camino de investigación es el contraste con otros problemas de diversas áreas.

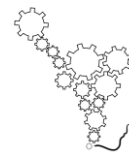


REFERENCIAS

- ASIMOV, Isaac. Isaac Asimov Isaac Asimov. **Popular science**, v. 7, p. 1.
- BOONSTRA, Lee. Dialogflow essentials concepts. In: **The definitive guide to conversational AI with dialogflow and google cloud: build advanced enterprise chatbots, voice, and telephony agents on google cloud**. Berkeley, CA: Apress, 2021. p. 59-91.
- CASTRILLÓN, Omar D.; SARACHE, William; RUIZ-HERRERA, Santiago. Predicción del rendimiento académico por medio de técnicas de inteligencia artificial. **Formación universitaria**, v. 13, n. 1, p. 93-102, 2020.
- CRUZ, Miguel. La enseñanza de la Matemática a través de la Resolución de Problemas. 2006.
- FONTOURA CAIRELLO, José Fabián. Agente conversacional para auxiliar en la resolución de problemas de Matemática. 2020.
- GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. Editora Atlas SA, 2002.
- GONZÁLEZ, Héctor M.; DUQUE, Nestor D.; OVALLE, Demetrio A. Modelo del estudiante para sistemas adaptativos de educación virtual. **Revista Avances en sistemas e informática**, v. 5, n. 1, p. 199-206, 2008.
- HELMS, Fernanda Krumreich et al. Modelando um sistema tutor multiagentes para auxiliar na aprendizagem da matemática. **Scientia Plena**, v. 11, n. 8, 2015.
- CHEN, Lijia; CHEN, Pingping; LIN, Zhijian. Artificial intelligence in education: A review. **Ieee Access**, v. 8, p. 75264-75278, 2020.
- LOPES, Arcanjo; NETTO, José Francisco; DE LIMA, Dhanielly PR. O uso de agentes conversacionais no apoio do ensino de resolução de problemas matemáticos: Uma revisão sistemática da literatura. In: **Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)**. 2018. p. 1403.
- LUAN, Hui et al. Challenges and future directions of big data and artificial intelligence in education. **Frontiers in psychology**, v. 11, p. 580820, 2020.
- MONLEÓN-GETINO, Antonio. El impacto del Big-data en la Sociedad de la Información. Significado y utilidad. **Historia y comunicación social**, v. 20, n. 2, p. 427-445, 2015.
- MORENO, Luis; WALDEGG, Guillermina. Constructivismo y educación matemática. **Educación matemática**, v. 4, n. 02, p. 7-15, 1992.
- NIETO, Lorenzo J. Blanco. Una clasificación de problemas matemáticos. **Epsilon: Revista de la Sociedad Andaluza de Educación Matemática "Thales"**, n. 25, p. 49-60, 1993.
- OBANDO, Eduardo Sandoval. Aprendizaje e inteligencia artificial en la era digital: implicancias socio-pedagógicas reales o futuras?. **Revista boletín REDIFE**, v. 7, n. 11, p. 155-171, 2018.



www.relainep.ufpr.br



- PAUR, Alicia Beatriz; ROSANIGO, Zulema Beatriz; BRAMATI, Pedro. La educación en la sociedad del conocimiento. In: **I Congreso de Tecnología en Educación y Educación en Tecnología**. 2006.
- POLYA, George et al. **Cómo plantear y resolver problemas**. México: Trillas, 1965.
- PUGLIANO, John. **The Robots are Coming: A human's survival guide to profiting in the age of automation**. Simon and Schuster, 2017.
- RUSSELL, Stuart J.; NORVIG, Peter. Inteligencia Artificial: un enfoque moderno. 2004.
- SAMPEDRO, Sergio Ballester. Estrategias para la resolución de problemas en matemáticas. **Revista digital, Innovación y experiencias educativas**, v. 19, p. 1-8, 2009.
- SERRANO GONZÁLEZ-TEJERO, José Manuel; PONS PARRA, Rosa María. El constructivismo hoy: enfoques constructivistas en educación. **Revista electrónica de investigación educativa**, v. 13, n. 1, p. 1-27, 2011.
- VÁSQUEZ, Augusto Cortez et al. Procesamiento de lenguaje natural. **Revista de investigación de Sistemas e Informática**, v. 6, n. 2, p. 45-54, 2009.