



MEMORIA FINAL DEL PROYECTO

CONVOCATORIA PARA PROMOVER PROYECTOS Y ACCIONES DE INNOVACIÓN Y MEJORA EN LA UNIVERSIDAD DE MURCIA.

| TÍTULO DEL PROYECTO |
|---|
| USO DE UN SISTEMA INTELIGENTE AUTONOMO PARA EL APRENDIZAJE DE ANATOMÍA HUMANA |

| Relación completa de los Participantes | | | |
|---|-----------------------------|---------------------------------|--|
| | Apellidos y Nombre | Departamento | |
| 1 | López González, Laura | Anatomía Humana y Psicobiología | laura.lopezgonzalez@um.es |
| 2 | González Sequero, Ofelia | Anatomía Humana y Psicobiología | sequeros@um.es |
| 3 | Fernández Alemán, José Luis | Informática y sistemas | aleman@um.es |

1.- Resumen del proyecto y trabajo presentado

Con el fin de ayudar a mejorar las estrategias de aprendizaje de los estudiantes de Anatomía general y descriptiva del aparato locomotor del grado de medicina (Código 3498), realizamos un experimento con un sistema de aprendizaje basado en redes neuronales que constan de dos etapas:

1. Empleo de un sistema i-SIDRA basado en redes neuronales en un modo a distancia no presencial en 4 fechas distintas del primer cuatrimestre, que se correspondían con el final de los 4 bloques de teoría en los que se divide el experimento. Esta prueba está integrada en la aplicación SIDRA (www.docentis.inf.um.es/sidra). Las respuestas a estos 4 test facilitan la identificación de conceptos difíciles de cada una de las materias y proporcionar un feedback positivo inmediato a los alumnos mediante texto o imágenes.

Esta Red Neuronal ha sido entrenada primeramente para su correcto funcionamiento y posteriormente reentrenada añadiendo las respuestas de los alumnos del curso 12-13 para crear estados de conocimiento refinados, basados en las respuestas de los alumnos a las preguntas formuladas. Esto supondrá un perfeccionamiento de los feedback aportados por el sistema a los alumnos que utilicen de nuevo i-SIDRA, y por tanto una mejora en el aprendizaje.

2. Uso de un Sistema de Respuesta a Distancia Móvil (aplicación SIDRA) en modo presencial, similar a los sistemas de clickers (Educlik), que permita aportar a los alumnos las repuestas correctas y discutir las en clase proporcionando una mayor participación y atención en clases teóricas aumentando la interactividad entre profesor y alumno. Esta modalidad fue realizada en 4 ocasiones tras el empleo del sistema i-SIDRA.

El estudio se complementó con el uso del Aula Virtual, para el manejo de toda la información de los datos de los estudiantes y de comunicación (anuncios, mensajes privados, etc.).



Al final del trimestre se les proporcionó a los usuarios del sistema, una encuesta de satisfacción tipo Likert de 11 preguntas más una pregunta abierta libre para opinar, para conocer sus apreciaciones.

El análisis estadístico del estudio se realizó comparando un control que realizó el experimento utilizando sistema con la red neuronal entrenada (curso 2012-13) y un grupo experimental que empleó el mismo sistema y en las mismas condiciones con la red neuronal reentrenada (curso 2017-18).

2.- Nivel de logro de los objetivos

El nivel de logro o de consecución de objetivos es total. Se ha conseguido realizar el análisis de las siguientes variables:

1. Análisis del rendimiento académico con las calificaciones del examen final de los alumnos del grupo experimental (Estudiantes del curso 17-18 que utilizaron con i-SIDRA la red neuronal reentrenada) y grupo control (Estudiantes del curso 2012-13, que se usó i-SIDRA sin reentrenar).
2. Así mismo se realizó un análisis de los índices de éxito y rendimiento entre el grupo experimental y control.
3. Se ha realizado la encuesta de satisfacción a la que contestaron 38 estudiantes, menos de los previstos porque se hizo la última semana de clase antes de los exámenes y la asistencia era menor.

Se consiguieron alcanzar los objetivos propuestos y los resultados son descritos en detalle en las conclusiones y resultados finales en el punto 5.

3.- Descripción del proceso de trabajo.

José Luis Fernández Alemán se ocupó fundamentalmente de la preparación del sistema para la Red Neuronal, previamente a cada ensayo, incluyendo el entrenamiento o reentrenamiento de la red neuronal, en este caso con las respuestas a las preguntas de los cuestionarios de años anteriores.

Laura López González y Ofelia González Sequeros se ocuparon de la realización y evaluación de las distintas fases del ensayo a nivel docente, llevadas a cabo en la asignatura Anatomía general y descriptiva del aparato locomotor de 1º del Grado de Medicina. Y por tanto del análisis de los resultados obtenidos tras la obtención de las calificaciones medias obtenidas en el examen final de la asignatura del grupo experimental y grupos control. También se encargaron de la elaboración del feedback de los 4 cuestionarios.

Se hicieron 4 cuestionarios online cuya clave de acceso se enviaba a través del Aula Virtual, la experiencia constó de cuatro ensayos que realizaron los alumnos en los meses de octubre, Noviembre (se realizaron 2), y diciembre a modo de evaluación continua.

Cada cuestionario constaba de 10-15 preguntas de respuesta múltiple con una única respuesta correcta posible. Los alumnos contestaban en un tiempo dado, 60 minutos, todas las veces que querían para mejorar los resultados y después de leer el feedback efectuado por las profesoras. El feedback eran aclaraciones o indicaciones o incluso dibujos o imágenes que ayudaba al alumno a encontrar la respuesta correcta ya que el sistema i-SIDRA no la suministra. Este sistema intenta ayudarles con el aprendizaje de la manera más individualizada posible, ya que esta información de apoyo o feedback, es generada por el sistema en función de la respuesta específica de cada alumno. El sistema analiza las respuestas de cada alumno y lo introduce en un grupo o estado de conocimiento al que le suministra el feedback específico.



Los cuestionarios online se completaban con pruebas de Sistema Interactivo de Respuesta (SIDRA) en clase y en los días siguientes a los cuestionarios, para aclarar dudas y ver las respuestas correctas.

- Preparación del sistema:

Antes del inicio del curso, la parte técnica del equipo de trabajo procedió al reentrenamiento de la red neuronal introduciendo las respuestas de los alumnos de Anatomía General y Descriptiva del Aparato Locomotor de 1º de medicina del curso 2012-13, en la red neuronal primeramente entrenada.

Por tanto, con este reentrenamiento, el sistema i-SIDRA tenía nuevos datos para la realización de un nuevo experimento en el que se encajarían a los alumnos del curso 2017-18 en nuevos grupos o estados de conocimiento, generando nuevos feedback para los alumnos.

Además, durante la preparación y realización de las pruebas, se encargó de solventar problemas técnicos que iban surgiendo.

- La parte docente se encargó de preparar tanto los feedback como los test para que el Grupo experimental:

89 alumnos de Anatomía general y descriptiva del aparato locomotor del grado de medicina del curso 2017-18, hicieron los cuestionarios i-SIDRA, con el modo a distancia no presencial basado en redes neuronales. En la guía docente se les informaba que la participación en este sistema les proporcionaría un 5% adicional en la nota, como incentivo para su participación en la prueba.

Se realizaron 4 test, únicamente por recoger mejor los bloques temáticos de la asignatura: 1. Tronco, 2. cabeza y cuello, 3. miembro superior y 4. miembro inferior. Las fechas de estos test on line con i-SIDRA se realizaron en octubre (tronco), 2 en noviembre (cabeza y miembro superior) y el 3º y 4º en diciembre (miembro superior e inferior). Tras cada cuestionario online se efectuaba un SIDRA en el aula donde la participación y la interacción profesor-alumno es muy alta

- Análisis estadístico:

El análisis estadístico se realizó con los datos obtenidos del grupo experimental y del grupo control:

Grupo experimental:

Está constituido por 83 alumnos de Anatomía general y descriptiva del aparato locomotor de 1º del grado de medicina del curso 2017-18, que realizaron, al menos 3 de las 4 pruebas propuestas en el estudio (lo que supone al menos el 75 % de la materia) con el sistema basado en redes neuronales mediante i-SIDRA utilizando la red neuronal Reentrenada en una ocasión.

Esto supone un importante seguimiento de los alumnos (83 %) y por lo tanto la participación y el interés han sido muy altos.

Grupo control:

Está constituido por 90 alumnos de Anatomía general y descriptiva del aparato locomotor del 1º del grado de medicina del curso 2013-14, con idénticas características y sometidos a las mismas pruebas que las realizadas por el grupo experimental, pero con la red neuronal aplicada a i-SIDRA solamente entrenada.

Las Pruebas analíticas generales aplicadas fueron:

Para la comparación entre grupo control y experimental se realizaron el análisis de los datos empleando, mediante métodos estadísticos paramétricos y no paramétricos, según se cumplieron o no los criterios de aplicación de los paramétricos. Cuando se analizaron los



cambios habidos se utilizaron pruebas estadísticas para muestras no emparejadas.

Para las variables continuas se resumieron utilizando estadísticos descriptivos: media, desviación estándar (SD), mediana. Las variables categóricas se describieron utilizando frecuencias absolutas y relativas.

Para la comparación de las medias con una variable dicotómica se utilizó la prueba estadística t de Student para dos medias independientes, tras comprobar que se cumplían las condiciones de aplicación de dicha prueba: a) Normalidad o $n > 30$ en cada grupo, y b) Homogeneidad de varianzas. La normalidad se comprobó mediante la prueba de Kolmogorov-Smirnov.

Par variables era de tipo categórico, se realizó una tabulación cruzada (tablas de contingencia) empleando el estadístico chi-cuadrado de Pearson para contrastar la hipótesis de independencia (en tablas $r \times s$) o el test de Fisher (en tablas 2×2). Previamente se comprobaron las condiciones de aplicación de la chi-cuadrado: a) Que ambas variables fuesen cualitativas en escala nominal y, b) Que ninguno de los valores esperados fuese menor de 5 en tablas de 2×2 .

Todas las comparaciones estadísticas se realizaron utilizando una prueba bilateral con un nivel de significación $p = 0,05$.

4.- Resultados finales.

4.1 INDICADORES:

Los índices empleados en el estudio son:

- Índice de rendimiento= Alumnos matriculados/ Alumnos presentados a examen*100
- Índice de éxito= Alumnos presentados a examen/ Alumnos aprobados*100

4.1.1 ÍNDICE DE RENDIMIENTO:

De los 83 alumnos del grupo experimental (utilizaron RNr), se presentaron al examen final de la asignatura 75 de ellos, lo que constituye un índice de rendimiento del 90,36%, mientras de los 78 alumnos matriculados en el grupo control (Utilizaron RN), 67 de ellos realizaron el examen final, originando un índice de rendimiento del 85,9% final de la asignatura.

El cálculo de la odds ratio sobre los valores de los índices de rendimiento del grupo experimental respecto al grupo control indicó que:

El aumento del índice de rendimiento con el uso del sistema con la red neuronal reentrenada, aumentaría los odds de presentarse al examen final en 1,539 veces más que si se utiliza el sistema con la red neuronal sin reentrenar, si el resto de variables se mantuvieran constantes.

Ambos valores muestran una mayor predisposición a realizar el examen final de los alumnos que han utilizado el sistema con la red neuronal reentrenada.

Sin embargo, al aplicar el estadístico chi cuadrado de Pearson se obtuvo un valor de 0,77 con $p=0,380$, lo que indica que no existe diferencias estadísticamente significativas entre el reentrenamiento o no reentrenamiento de la red neuronal con respecto a la variable de presentación al examen final, aceptando por tanto la hipótesis nula (El uso del sistema basado en redes neuronales reentrenadas no mejora el índice de rendimiento de la signatura comparándolos con los alumnos que utilizan el mismo sistema basado en redes neuronales sin reentrenar)



4.1.2. ÍNDICE DE ÉXITO:

En cuanto al índice de éxito, en el que se representa el número de aprobados en cada grupo con respecto a los alumnos que realizan el examen final, se observó que:

En el grupo experimental (utilizaron RNr) 57 alumnos aprobaron el examen final de un total de 75 estudiantes presentados dando un índice de éxito del 76%. Mientras que en el grupo control (Utilizaron RN) 50 alumnos aprobaron de 67 alumnos presentados al examen final, con un índice de éxito de 74,63%.

Estos datos apuntarían hacia la idea de una mejor preparación de los alumnos pertenecientes al grupo experimental, al obtenerse un mayor porcentaje de estudiantes que superan la asignatura frente al grupo control.

El cálculo de la odds ratio sobre los valores de los índices de éxito del grupo experimental respecto al grupo control indicó que:

El aumento del índice de éxito al utilizar el sistema con la red neuronal reentrenada, aumentaría los odds de aprobar el examen final en 1,076 veces más que si se utiliza el sistema con la red neuronal sin reentrenar, si el resto de variables se mantuvieran constantes. Lo que indica prácticamente una igualdad en la probabilidad de aprobar al utilizar el sistema con redes neuronales reentrenadas o no.

Por último al aplicar el chi cuadrado de Pearson se obtuvo un valor de 0,04 con $p=0,841$, lo que indica que no existe diferencias estadísticamente significativas entre el reentrenamiento (grupo experimental) o no reentrenamiento (grupo control) de la red neuronal con respecto a la variable de presentación al examen final, aceptando por tanto la hipótesis nula (El uso del sistema basado en redes neuronales reentrenadas no mejora el índice de éxito de la signatura comparándolos con los alumnos que utilizan el mismo sistema basado en redes neuronales sin reentrenar)

4.2. COMPARACIÓN ESTADÍSTICA DE LAS CALIFICACIONES MEDIAS DEL EXAMEN FINAL DE LA ASIGNATURA:

Las calificaciones medias obtenidas en el examen final de la asignatura de Anatomía general y descriptiva fueron:

- De 6,33 en el grupo experimental que realizó cuatro pruebas, a lo largo del curso, con un sistema basadas en redes neuronales reentrenadas por primera vez. obtuvo una calificación media de 6,33 en el examen final de la asignatura.
- De 6,11 en el grupo control que realizó las mismas cuatro pruebas a lo largo del curso, con el mismo sistema basado en redes neuronales sin reentrenamiento.

Se realizó la comparación de las calificaciones medias obtenidas por el grupo experimental y grupo control mediante la prueba estadística U de Mann-Whitney (Prueba no paramétrica para 2 muestras independientes) una vez comprobado que las muestras no se comportan según la normalidad.

El valor obtenido de Z-Score es de 0.43732 con un valor de $P=0,65994$, lo que muestra que no existen diferencias estadísticamente significativas entre las calificaciones medias de ambos grupos. Aceptando por lo tanto la hipótesis nula (El uso del sistema basado en redes neuronales reentrenadas no incrementa los rendimientos académicos del examen final de la asignatura comparándolos con los alumnos que utilizan el mismo sistema basado en redes neuronales sin reentrenar)

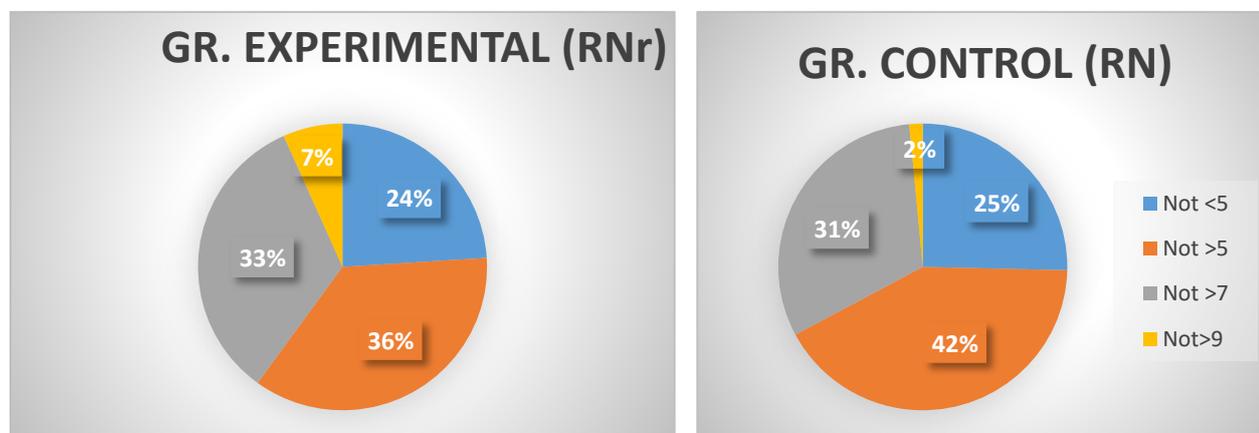
Por otro lado, desglosando los resultados de las notas obtenidas por ambos grupos (tabla 1, gráfico1), se observa que, a pesar de que el número de aprobados fue similar en ambos grupos, las proporciones en los diferentes rangos de notas son desiguales. De hecho, en el grupo control, 32 alumnos de los 50 aprobados (32,8%), se encuentran en el rango de notas superior a notable (entre 7 y 10), mientras que en el grupo experimental, muestra mayor proporción de alumnos situados en este rango (40%), lo que indicaría que los

alumnos del grupo experimental obtuvieron mejores resultados en el rango de notas superiores que el grupo control.

Tabla 1: Número y porcentaje de alumnos aprobados y suspensos y de alumnos aprobados en función de sus calificaciones, en el examen final de ambos grupos.

| CALIFICACIONES | % suspensos | % aprobados | 5-7 | 7-9 | 9-10 |
|--------------------------------------|--------------|--------------|--------------|--------------|-------------|
| GR. EXPERIMENTAL (n° alumnos) | 18 | 57 | 27 | 25 | 5 |
| (Total 75 alumnos) % | 24,0% | 76,0% | 36,0% | 33,3% | 6,7% |
| GR. CONTROL (n° alumnos) | 17 | 50 | 28 | 21 | 1 |
| (Total 67 alumnos)% | 25,4% | 74,6% | 41,8% | 31,3% | 1,5% |

Gráfico 1: Representación de los porcentajes de alumnos aprobados y suspensos y de alumnos aprobados en función de sus calificaciones, en el examen final de ambos grupos.



No obstante, al aplicar estadístico chi cuadrado de Pearson a las variables rangos de calificaciones estudiados en el área de aprobados con respecto al grupo experimental y grupo control el valor obtenido fue de 2,5858 con $p=0,27447$, lo que indica que no existe diferencias estadísticamente significativas entre ambos grupos. Aceptando por tanto la hipótesis nula (El uso del sistema basado en redes neuronales reentrenadas no mejora las calificaciones de los alumnos obtenidos en el examen final de la asignatura comparándolo con los alumnos que utilizan el mismo sistema basado en redes neuronales sin reentrenar).

4.3. ENCUESTA DE SATISFACIÓN.

Se realizó a los alumnos una encuesta de satisfacción constituida por 11 preguntas (escala Likert de 1-5) para conocer su opinión sobre varias cuestiones relacionadas con el uso del sistema como se muestra en la tabla. Los resultados obtenidos han resultado muy positivos.

En concreto, el 83,1 % de los encuestados está satisfecho con el uso del sistema basado en red neuronal y al 94.4% les resultó útil como autoevaluación de conceptos. Así mismo el 85.9% de los alumnos considera que los contenidos y la metodología empleada en esta asignatura les adecuada en su aprendizaje.

Por otro lado, 97.2% de los encuestados valora positivamente la implicación del profesor en metodologías innovadoras como los sistemas empleados en el estudio.

No obstante, hay que destacar que algunos alumnos no entienden el sentido de con ciertas partes del sistema, como las informaciones de feedback que no suministran la respuesta, ya que están acostumbrados al sistema tradicional donde reciben la respuesta correcta sin necesidad de reflexión. Lo que genera un aprendizaje de la respuestas pero no se aprenden o aclaran los conceptos básicos que se preguntan. Sin embargo el 53,6% creen que los feedback recibidos no han sido la información que necesitaban para aclarar sus conceptos y el 29.3% le resulta indiferente. Además el 47.9% no cree que los feedback recibidos le han sido de utilidad para mejorar sus conceptos y al 32.4 le resulta indiferente.

Además a la pregunta de si creen que el sistema les influirá positivamente en sus calificaciones, la respuesta ha sido moderada, ya que el 36.6% se muestran dudosos e indiferentes y el 54,9% piensan que es positivo,

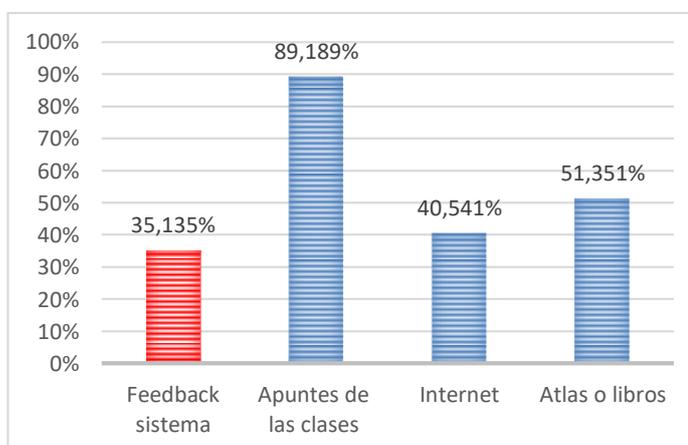
mientras que solo un 7% no lo creen.

A modo de apreciación y apuntando a los resultados obtenidos vemos que su percepción es inexacta, ya que la tasa de rendimiento, éxito y notas medias del examen final han sido mejores en el grupo experimental respecto al control.

Como comentario final queremos indicar que estos resultados nos animan a continuar en esa línea y que una propuesta de mejora debía ser la mejora y potenciación del feedback como medio para el autoperendizaje del alumno.

Por último, Se les realizó una pregunta adicional en la que les pedíamos a los alumnos que nos indicaran con sinceridad, ya que la encuesta era anónima, que indicaran de las fuentes de información propuestas (feedback del sistema, apuntes de las clases, internet o atlas y libros) utilizado para la realización de las pruebas a distancia.

Gráfico 2: Porcentaje del uso de las distintas fuentes de información.

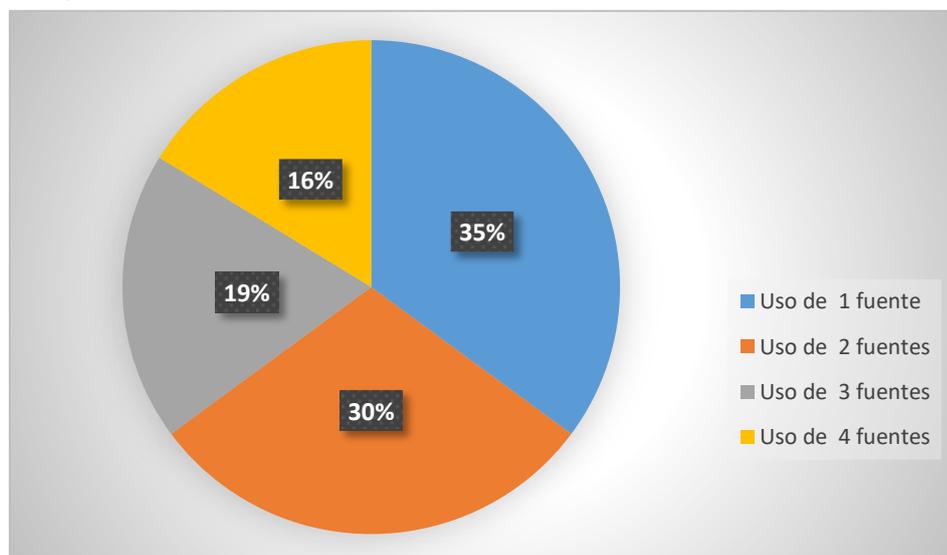


Los alumnos indicaron que la fuente de información menos empleada (35,1%) fueron los feedback que el sistema ofrecía, algo que no sorprende ya que los alumnos no valoran muy bien el feedback ofrecido por el sistema y creen que deben acudir a otras fuentes para hallar la respuesta (gráfico 2).

El sistema ofrece pistas sobre los errores cometidos en el test ofrecido, pero no proporciona directamente la respuesta correcta, con el objetivo de favorecer la reflexión del alumno sobre los temas tratados y reconocer el nivel de estudio que tiene, llegando a ser en parte un ejercicio de autoevaluación sin presión puesto que la puntuación obtenida en los ejercicios no se tiene en cuenta en las calificaciones finales.

En cuanto a cuantas fuentes fueron consultadas el 25% de ellos utilizaron más de 3 fuentes distintas para resolver sus dudas lo que parece mostrar que el estudio de la materia necesario en la preparación de la prueba era insuficiente. Además, el 35% de ellos utilizaron 1 sola fuente, pero fue principalmente los apuntes de clase y tan solo 1 alumno declaró que utilizó exclusivamente el feedback ofrecido por el sistema.

Gráfico 3: Porcentaje del uso de una o varias fuentes a la vez.



Además entre el 47% y 67% de los alumnos que han utilizado varias fuentes para resolver las pruebas, admiten no haber utilizado el feedback ofrecido por el sistema.

4.4. COMPLICACIONES OBSERVADAS

Varios han sido los motivos por los que se cree que no se han alcanzado los resultado esperados:

Es indispensable indicar que es la primera vez que se emplea el sistema i-SIDRA utilizando una red neuronal reentrenada, lo que teóricamente debería aumenta su rendimiento. Sin embargo, a lo largo del experimento han aparecido complicaciones prácticas que no han podido ser previstas y que creemos que han podido mermar el éxito de su uso.

Durante el experimento hemos observado que el reentrenamiento de la red neuronal generó grupos, con un feedback asociado, que no eran utilizados por i-SIDRA, lo que impedía que los alumnos recibieran la información más adecuada, como estaba previsto.

Esto se analizó posteriormente llegando a la conclusión de que, en el entrenamiento original de la red neuronal se utilizaron respuestas sin contestación o en blanco. Se ha barajado la idea de que el pequeño vacío de información original pudo ser asimilado por la red neuronal primeramente, sin embargo, al reentrenar de nuevo la red neurona este defecto fue mayor provocando ciertas alteraciones en la entrega de los feedback por parte de i-SIDRA.

No obstante, a pesar de los inconvenientes encontrados a lo largo del desarrollo del experimento, es destacable que todos los valores obtenidos por el grupo experimental, tanto para el rendimiento académico como para los índices de rendimiento y de éxito han resultado mayores que los del grupo control.

Esto, nos anima a continuar con nuestra línea de trabajo, ya que en otros estudios anteriores obtuvimos mejoras estadísticamente significativas con un grupo control de estudiantes con enseñanza tradicional y por tanto entendemos que seguimos obteniendo beneficio en el uso de este sistema.

Por último, también consideramos que es importante el perfeccionamiento del sistema a dos niveles:

1. Perfeccionamiento de la red neuronal mediante un entrenamiento y reentrenamiento más apropiados.
2. Mejora del sistema a nivel del feedback, ya que se ha observado en el estudio su uso incorrecto, puesto que los alumnos han admitido que emplean otras fuentes de información para conseguir la satisfacción del resultado.