

I CONGRESO IBEROAMERICANO DE DOCENTES

CONGRESO VIRTUAL DEL 26 NOVIEMBRE AL 08 DICIEMBRE DE 2018

ALGECIRAS (CÁDIZ) DEL 06 AL 08 DICIEMBRE DE 2018

Actas del Congreso Iberoamericano de Docentes

Aprendizaje orientado a proyectos en
Reconocimiento Automático de Escritura

Verónica Romero Gómez

Emilio Granell Romero

Emilio Granell

Alejandro H. Toselli

Moisés Pastor

ISBN: 978-84-948417-0-5

Edita **Asociación Formación IB.**

Coordinación editorial: **Joaquín Asenjo Pérez, Óscar Macías Álvarez, Patricia Ávalo Ortega y Yoel Yucra Beisaga**

Año de edición: **2018**

Presidente del Comité Científico: **César Bernal.**

El I Congreso Iberoamericano de Docentes se ha celebrado organizado conjuntamente por la Universidad de Cádiz y la Asociación Formación IB con el apoyo del Ayuntamiento de Algeciras y la Asociación Diverciencia entre otras instituciones.

<http://congreso.formacionib.org>



red
iberoamericana
de docentes



formaciónib))

Aprendizaje orientado a proyectos en Reconocimiento Automático de Escritura

Verónica Romero, Emilio Granell, Alejandro H. Toselli, Moisés Pastor
Centro PRHLT, Universitat Politècnica de València, España
{vromero,egrnell,ahector,mpastor}@prhlt.upv.es

Resumen

En este artículo se presenta un método de enseñanza innovador, basado en el método de aprendizaje orientado a proyectos, en el que los estudiantes deben aplicar las habilidades y conocimientos adquiridos en clase en el desarrollo de un sistema real. Este método se ha aplicado en la asignatura “Reconocimiento de Escritura” (RES) que se imparte en el máster universitario “Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital” de la Universitat Politècnica de València con buenos resultados. En los sistemas de reconocimiento de texto manuscritos (RTM) hay varias tecnologías involucradas que colaboran para obtener resultados aceptables. Entre estas tecnologías destacan el preproceso y la limpieza de las imágenes de texto, la segmentación automática de dichas imágenes en líneas, el modelado morfológico de los caracteres, el modelado del lenguaje y finalmente el reconocimiento. En RES, se desea que el alumno no solo conozca dichas tecnologías, si no que sea capaz de utilizarlas e integrarlas para construir un sistema de reconocimiento de texto del estado del arte. Para ello, durante el curso, a medida que se les presentaron a los alumnos las diferentes tecnologías involucradas, se les proporcionaron las herramientas básicas para construir un sistema RTM. Finalmente, con el fin de evaluar la asignatura, se les pidió a los alumnos que llevaran a cabo el reconocimiento automático de un documento con el sistema desarrollado. Para medir el nivel de aceptación de este nuevo planteamiento por parte de los alumnos, se les dio la opción de realizar un test en lugar de llevar a cabo el desarrollo y la evaluación de un sistema. Solo un 5% prefirió la evaluación mediante un test, el 95% restante desarrolló su propio sistema de RTM obteniendo resultados similares a los obtenidos en el estado del arte.

1. Introducción

El reconocimiento automático de documentos de texto manuscrito es un tema candente de investigación y desarrollo en el área del reconocimiento de formas. Esto se debe especialmente a la ingente cantidad de manuscritos históricos que se conservan en archivos y bibliotecas de todo el mundo, muchos de los cuales se han estado digitalizando a lo largo de las últimas décadas para dejarlos al alcance del público en general. Para que las imágenes de texto manuscrito sean realmente útiles deben ser anotadas con información acerca de su contenido. Dado el inmenso volumen de documentos de interés involucrados, dicha transcripción no puede obtenerse de manera manual, por lo que el proceso pasa necesariamente por el uso de métodos automatizados [1,2].

En los últimos años ha habido notables avances en el campo de “Reconocimiento de Texto Manuscrito” (RTM). Estos avances se han llevado a cabo no solo en el desarrollo de sistemas de transcripción totalmente automática, si no que también se ha avanzado mucho en el desarrollo de sistemas de transcripción asistida [3], donde el usuario y la máquina colaboran para obtener la transcripción perfecta, y en sistemas de búsqueda de palabras clave.

En la asignatura “Reconocimiento de Escritura” (RES), que se imparte en el máster universitario “Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital” (MIARFID) de la Universitat Politècnica de València (UPV), se estudia como crear un sistema automático de reconocimiento de escritura manuscrita de última generación, así como conceptos avanzados de inteligencia artificial, reconocimiento de formas, imagen digital y tecnologías del lenguaje. También se introducen los sistemas interactivos de asistencia a la transcripción [3] y los sistemas de búsqueda de palabras clave [4].

A diferencia de como se venía haciendo hasta este último curso, donde la asignatura se presentaba a los alumnos como una secuencia de lecciones magistrales, en este último curso la asignatura se ha planteado con un enfoque muy práctico, basada en el método de aprendizaje orientado a proyectos. En este enfoque los estudiantes deben aplicar las habilidades y conocimientos adquiridos en clase en el desarrollo de un sistema real. En este caso, la asignatura se ha articulado alrededor de un proyecto de implementación de un sistema interactivo de reconocimiento de texto manuscrito e indexación para la transcripción de un libro manuscrito antiguo: “Historia de España del arzobispo Don Rodrigo” del siglo XVI [5]. En estos sistemas hay varias tecnologías involucradas que colaboran para obtener resultados aceptables:

- Adquisición y mejora de imágenes: binarización y eliminación de ruido [6].
- Análisis de la estructura de páginas: detección automática de bloques de texto y líneas [7].
- Normalizaciones geométricas: rotación, inclinación y tamaño [8].
- Extracción de características [8].
- Modelado óptico, léxico y de lenguaje [8].
- Decodificación del texto manuscrito [8].
- Interacción hombre-máquina [3].
- Indexación [4]

En RES, se desea que el alumno sea capaz de utilizar e integrarlas estas tecnologías para construir un sistema de reconocimiento del estado del arte. Para ello, durante el desarrollo del curso, a medida que se les han presentado a los alumnos las diferentes tecnologías, se les han proporcionado las herramientas necesarias para su utilización e implementación. De esta forma, el proyecto global en el que se centra la asignatura se puede dividir en etapas o subproyectos que, tras su implementación e integración, darán lugar al sistema deseado.

1.1 Resultados de aprendizaje

Los resultados de aprendizaje se pueden definir como las habilidades y los conocimientos que se espera que el alumno adquiera, conozca y/o comprenda al final de un recorrido formativo. Una correcta definición de los mismos es muy importante pues informa a los alumnos sobre lo que se espera de ellos y, por otro lado, ayuda al profesor en la organización de las clases y de las pruebas objetivas a utilizar en la evaluación. Los resultados de aprendizaje en la asignatura RES son:

- Implementar técnicas de limpieza de imágenes de texto.
- Utilizar técnicas de análisis de la estructura de página: detección de bloques y líneas.
- Clasificar las diferentes técnicas de normalizaciones geométricas existentes en la literatura.
- Aplicar las técnicas de normalizaciones geométricas adecuadas a los documentos estudiados.

- Enumerar las diferentes técnicas de extracción de características disponibles en la literatura.
- Describir los diferentes tipos de modelado óptico, léxico y de lenguaje
- Desarrollar sistemas de procesamiento del lenguaje natural y su aplicación al reconocimiento de escritura.
- Integrar tecnologías y sistemas propios de la inteligencia artificial, el reconocimiento de formas, la imagen digital y las tecnologías del lenguaje.
- Diseñar sistemas de interacción hombre-máquina.
- Evaluar sistemas de reconocimiento de escritura, de interacción hombre-máquina y de búsqueda de palabras clave.

1.2 Tareas del profesorado

La innovación docente presentada en este artículo requiere una serie de tareas por parte del profesorado para el logro de los resultados de aprendizaje propuestos con la nueva metodología de aprendizaje elegida:

- Preparar nuevo material docente.
- Adaptar la transmisión del conocimiento.
- Diseñar las actividades de aprendizaje presenciales y no presenciales.
- Establecer los sistemas de seguimiento y de evaluación.

2. Antecedentes

Desde su inicio, en el curso 2007-2008, la asignatura “Reconocimiento de Escritura” (RES) se ha venido impartiendo siguiendo una metodología de aprendizaje basada en lecciones magistrales. Para cada una de las etapas involucradas en el reconocimiento de escritura un experto del tema presentaba a los alumnos la tecnología existente en el estado del arte y su utilización. Dicha presentación solía llevarse a cabo mediante una lección o clase magistral en ocasiones acompañada por algunos ejercicios prácticos. Finalmente, la evaluación se realizaba mediante una prueba objetiva de respuesta abierta.

3. Aprendizaje orientado a proyectos

El aprendizaje de los humanos se basa en la construcción activa del conocimiento, y no por la transmisión del mismo. Por lo tanto, para mejorar el aprendizaje de los alumnos, es necesario utilizar métodos didácticos que activen a los estudiantes. El aprendizaje orientado a proyectos cumple con esta necesidad, pues consiste en la aplicación de habilidades y conocimientos adquiridos por parte del alumnado con la intención de solucionar una situación problemática concreta, mediante la planificación, diseño y realización de una serie de actividades.

Las etapas más características del aprendizaje orientado a proyectos son:

1. Definir el proyecto.
2. Planificar todos los detalles del proyecto. Distribuirlo en tareas.
3. Seleccionar el material necesario. Obtener y estructurar la información.
4. Realizar un seguimiento del proyecto.
5. Llevar a cabo el proyecto.
6. Analizar y evaluar el trabajo realizado y las aportaciones individuales.

Dadas las características de la asignatura RES, se optó por esta metodología de aprendizaje. Las diferentes tecnologías involucradas en el reconocimiento de escritura se les presentaron a los alumnos en 8 sesiones participativas donde, además de su descripción detallada, se les proporcionaron herramientas básicas para su utilización e integración. A continuación se detallan las etapas características de proyecto que debían llevar a cabo los alumnos:

-Definición del proyecto

Desarrollo de un sistema interactivo de reconocimiento de texto manuscrito e indexación.

-Planificar todos los detalles del proyecto. Distribuirlo en tareas.

El proyecto se divide en una tarea por cada una de las etapas que forman el sistema más una tarea final de integración:

1. Desarrollo de un sistema de limpieza de imágenes.
2. Implementación de un sistema de análisis de la estructura de páginas
3. Desarrollo de herramientas para la normalización de las imágenes: skew, slope, slant, warping y tamaño.
4. Entrenamiento de los modelos ópticos.
5. Entrenamiento del modelado de lenguaje.
6. Decodificación de las muestras de test y obtención de los grafos de palabras
7. Integración del sistema interactivo de reconocimiento (CATTI)
8. Obtención de los índices necesarios para llevar a cabo la indexación. (KWS)

En la Ilustración 1 se puede ver un diagrama de un sistema interactivo de reconocimiento de texto manuscrito e indexación.

-Seleccionar el material necesario. Obtener y estructurar la información.

El desarrollo del sistema se realizó para la transcripción de un libro manuscrito antiguo llamado: "Rodrigo" del siglo XVI disponible en: <https://www.prhlt.upv.es/wp/es/resource/the-rodrigo-corpus>.

La evaluación del sistema se llevó a cabo con el documento publico conocido como "Washington" y disponible en: <http://www.fki.inf.unibe.ch/databases/iam-historical-document-database/washington-database>

Por otro lado, las herramientas necesarias estaban disponibles gratuitamente para investigación en:

- Laia. Entrenamiento de los modelos ópticos basados en redes neuronales: <https://github.com/jpuigcerver/Laia>
 - SRILM. Entrenamiento de los modelos de lenguaje: <http://www.speech.sri.com/projects/srilm/>
 - EESSEN. Decodificación basada en transductores de estados finitos: <https://github.com/srvk/eesen>
 - CATTI. Sistema de transcripción asistida: <https://github.com/PRHLT/CATTI>
 - KWS. Sistema de búsqueda de palabras clave: <https://github.com/PRHLT/KwsEvalTool>
- y <https://github.com/PRHLT/WordGraph2Index>

- BaseLinePage. Sistema para segmentación de líneas:
<https://github.com/PRHLT/BaseLinePage>

-Realización del seguimiento del proyecto

A medida que iban avanzando en el curso los alumnos tenían retroalimentación por parte del profesor sobre su avance en las diferentes etapas del desarrollo del sistema.

-Llevar a cabo el proyecto

Los alumnos llevaron a cabo el proyecto de forma individual y continuada a lo largo del curso.

-Analizar y evaluar el trabajo realizado y las aportaciones individuales

Para su evaluación, además del proceso llevado a cabo durante el curso, en el que el alumno contaba con la tutorización constante del profesor, utilizando el documento “Rodrigo”, el alumno debía aplicar el sistema de reconocimiento implementado en un nuevo documento antiguo conocido como “George Washington”.

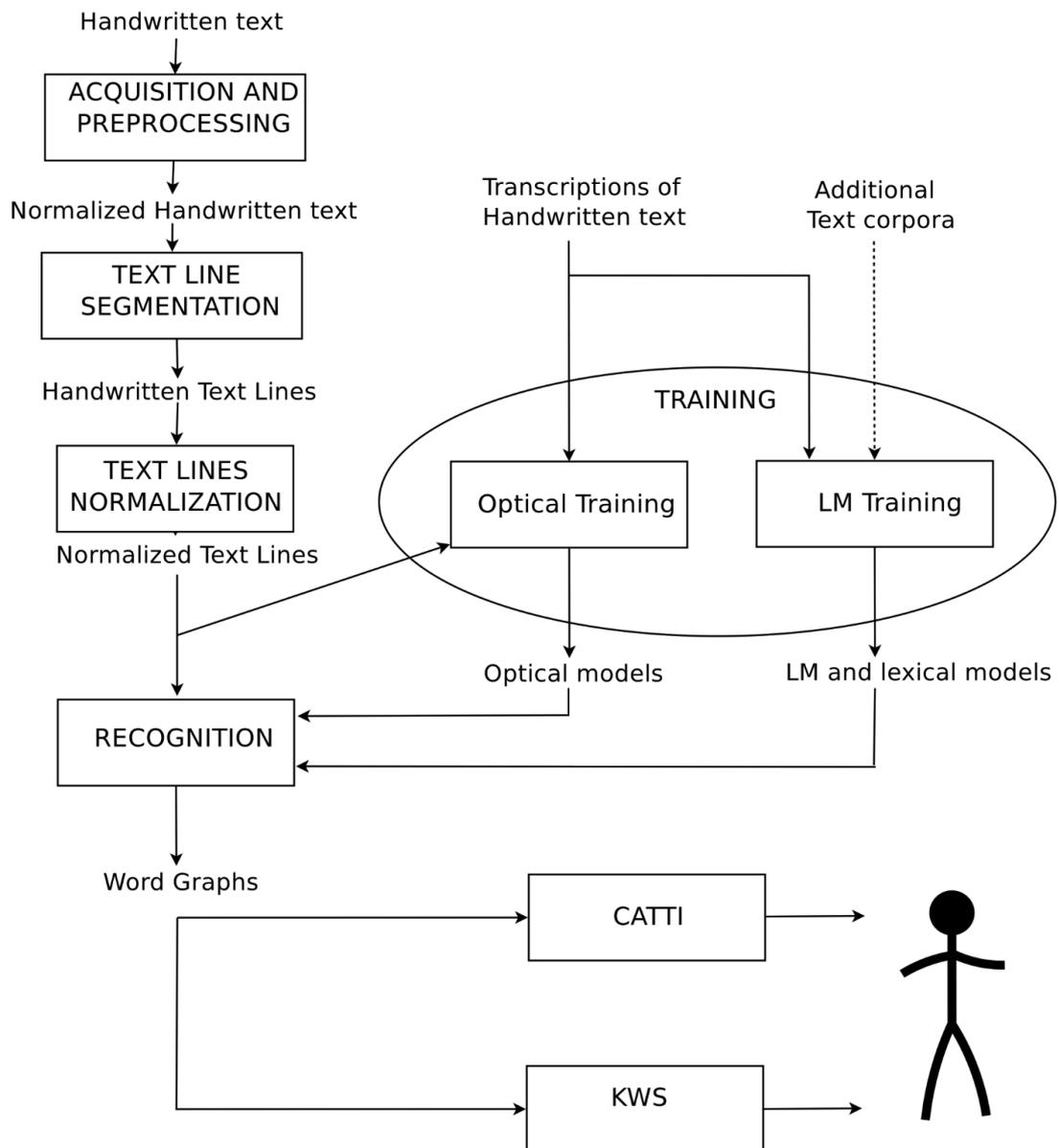


Ilustración 1: Diagrama de un sistema interactivo de reconocimiento de texto manuscrito e indexación

4. Evaluación

Con el fin de evaluar el nivel de aceptación de la innovación docente introducida en la asignatura por parte de los alumnos, se les dió la opción de llevar a cabo la evaluación de la asignatura de dos formas diferentes. La primera de ellas consistía en un test objetivo de respuesta abierta, tal y como se venía haciendo en años anteriores. La otra opción ofertada a los alumnos consistía en llevar a cabo el desarrollo y evaluación de un sistema de reconocimiento de texto manuscrito tal y como se ha comentado en el apartado anterior.

Solo un 5% de los alumnos prefirió la evaluación mediante un test, mientras que el 95% restante desarrolló el sistema obteniendo resultados similares a los obtenidos en el estado del arte.

5. Conclusiones

En este artículo se ha presentado un método de enseñanza basado en el aprendizaje orientado a proyectos en la asignatura “Reconocimiento de Escritura” del máster MIARFID de la UPV. Dadas las características de la asignatura y tras los resultados obtenidos, podemos concluir que dicha metodología de aprendizaje se adapta mejor a la asignatura y permite a los alumnos adquirir los conocimientos y habilidades mejor y mas eficientemente.

6. Referencias Bibliográficas

- [1] Bazzi, I., Schwartz, R., & Makhoul, J. (1999). An omnifont open-vocabulary OCR system for English and Arabic. *IEEE TPAMI*, 21(6), 495-504.
- [2] Bluche, T. (2015). *Deep neural networks for large vocabulary handwritten text recognition* (Doctoral dissertation, Université Paris Sud-Paris XI).
- [3] Romero, V., Toselli, A. H., & Vidal, E. (2012). *Multimodal interactive handwritten text transcription* (Vol. 80). World Scientific.
- [4] Toselli, A. H., Vidal, E., Romero, V., & Frinken, V. (2016). HMM word graph based keyword spotting in handwritten document images. *Information Sciences*, 370, 497-518.
- [5] Serrano, N., Castro, F., & Juan-Ciscar, A. (2010). *The RODRIGO Database*. LREC.
- [6] Villegas, M., Romero, V., & Sánchez, J. A. (2015, June). On the Modification of Binarization Algorithms to Retain Grayscale Information for Handwritten Text Recognition. In *IbPRIA 2015* (pp. 208-215). Springer, Cham.
- [7] Fawzi, A., Pastor, M., & Martínez-Hinarejos, C. D. (2017, August). Baseline Detection on Arabic Handwritten Documents. In *Proceedings of the 2017 ACM Symposium on Document Engineering* (pp. 193-196). ACM.
- [8] Puigcerver, J. (2017, November). Are Multidimensional Recurrent Layers Really Necessary for Handwritten Text Recognition?. In *ICDAR, 2017 14th IAPR* (Vol. 1, pp. 67-72). IEEE.

Aprendizaje orientado a proyectos en Reconocimiento Automático de Escritura

Verónica Romero, Emilio Granell, Alejandro H. Toselli, Moisés Pastor
 Pattern Recognition and Human Language Technology Research Center,
 Universitat Politècnica de València, Camí de Vera s/n, 46022, Valencia



- ✘ El reconocimiento automático de texto manuscrito es un tema candente de investigación y desarrollo en el área del reconocimiento de formas.
- ✘ En la asignatura Reconocimiento de Escritura (RES) se estudian los sistemas de reconocimiento de escritura manuscrita de última generación.
- ✘ La asignatura se ha planteado con un enfoque práctico, basada en el método de aprendizaje orientado a proyectos.

- ✘ La asignatura RES se ha venido impartiendo siguiendo una metodología de aprendizaje basada en lecciones magistrales y ejercicios prácticos.
- ✘ La evaluación se llevaba a cabo mediante una prueba objetiva de respuesta abierta.
- ✘ Debilidad de la sesión magistral: fomenta la pasividad y la falta de participación del estudiante.
- ✘ Solución propuesta: metodología orientada a proyectos.

Definición del proyecto

Desarrollo de un sistema interactivo de reconocimiento de texto manuscrito e indexación.

Distribución del proyecto en tareas

Una tarea por cada una de las tecnologías involucradas en el reconocimiento más una tarea de integración:

1. Desarrollo de un sistema de limpieza de imágenes
2. Implementación de un sistema de análisis de la estructura de páginas
3. Desarrollo de herramientas para la normalización de las imágenes
4. Entrenamiento de los modelos ópticos
5. Entrenamiento del modelado de lenguaje
6. Decodificación de las muestras de test y obtención de grafos de palabras
7. Integración de un sistema interactivo de transcripción (CATTI)
8. Obtención de los índices para búsqueda de palabras clave (KWS)

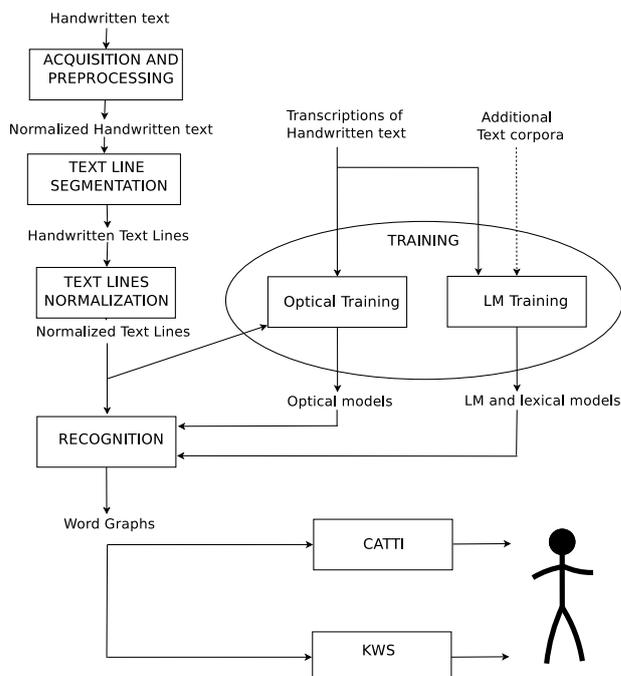
Material necesario

Herramientas utilizadas:

- ✘ Laia: Entrenamiento de los modelos ópticos basados en redes neuronales. <https://github.com/jpuigcerver/Laia>
- ✘ SRILM: Entrenamiento de los modelos de lenguaje. <http://www.speech.sri.com/projects/srilm/>
- ✘ EESEN: Decodificación basada en transductores de estados finitos. <https://github.com/srvk/eesen>
- ✘ CATTI: Sistema de transcripción asistida. <https://github.com/PRHLT/CATTI>
- ✘ KWS: Sistema de búsqueda de palabras clave. <https://github.com/PRHLT/WordGraph2Index>
- ✘ BaseLinePage: Sistema para segmentación de líneas. <https://github.com/PRHLT/BaseLinePage>

Conjunto de datos:

- ✘ Rodrigo. Corpus medieval del siglo XVI en español.
- ✘ Washington. Corpus en inglés del siglo XVIII.



Desarrollo del proyecto

El proyecto se llevó a cabo de forma individual a lo largo de todo el curso. En todo momento tuvieron retroalimentación por parte del profesorado.

Evaluación del trabajo realizado

- ✘ Proceso llevado a cabo durante el curso con Rodrigo
- ✘ Aplicación del sistema implementado en Washington

✘ Los alumnos podían elegir entre dos tipos de evaluación:

- ⊖ Un test objetivo de respuesta abierta.
- ⊖ Desarrollo y evaluación de un sistema interactivo de reconocimiento de texto manuscrito e indexación

✘ El 95% de los alumnos desarrolló el sistema de reconocimiento con resultados similares a los del estado del arte.

✘ Se ha presentado un método de enseñanza basado en el aprendizaje orientado a proyectos en la asignatura RES.

✘ Los resultados obtenidos muestran que dicha metodología se adapta mejor a la asignatura.

✘ Los alumnos adquieren los conocimientos y habilidades de forma amena y eficiente.