

Aprendizaje experiencial en ciencia de datos: satisfacción de los estudiantes para tres modelos de enseñanza y aprendizaje

Experiential learning in data science: student satisfaction for three models of teaching and learning

Emilio Serrano, Martin Molina, Daniel Manrique, Luis Baumela, Damiano Zanardini
emilioserra@fi.upm.es, mmolina@fi.upm.es, dmanrique@fi.upm.es, lbaumela@fi.upm.es, damiano@fi.upm.es

Departamento de Inteligencia Artificial
Universidad Politécnica de Madrid
Madrid, España

Resumen- La Ciencia de Datos es una revolución que ya está cambiando la manera en la que nos ocupamos de negocios, sanidad, política, educación e innovación. Hay una gran variedad de cursos online, másteres, grados, y asignaturas que se enfocan a la enseñanza de este campo interdisciplinar, donde existe una demanda creciente de profesionales. Este artículo describe tres modelos de enseñanza y aprendizaje diferentes para Ciencia de Datos, inspirados en el paradigma del aprendizaje experiencial. Además, estos modelos han sido empleados en una asignatura de *Deep Learning*, dentro del contexto de un máster internacional de Ciencia de Datos. Finalmente, la satisfacción de los estudiantes es cuantificada por medio de una encuesta para: evaluar los tres modelos, decidir el enfoque favorito y considerar una metodología de aprendizaje experiencial pura.

Palabras clave: *Ciencia de Datos, Aprendizaje Profundo, Aprendizaje experiencial, Gamificación*

Abstract- Data science is a revolution that is already changing the way we do business, healthcare, politics, education and innovation. There is a great variety of online courses, masters, degrees, and modules that address the teaching of this interdisciplinary field, where is a growing demand of professionals. This paper describes three different teaching and learning models for Data Science inspired by the experiential learning paradigm. Moreover, these models have been employed in a Deep Learning course in the context of an international master of data science. Finally, the student satisfaction is quantified by a survey to: evaluate the three models, decide their favorite approach, and consider a pure experiential learning methodology.

Keywords: *Data Science, Deep Learning, Experiential Learning, Gamification*

1. INTRODUCCIÓN

La Ciencia de datos o *Data Science* (DS) es un campo interdisciplinar que se encarga de la extracción del conocimiento de los datos. Esta disciplina es particularmente compleja ante el Big Data: grandes volúmenes de datos que dificultan su almacenamiento, procesamiento y análisis con tecnologías estándar de las Ciencias de la computación. La Ciencia de los datos es una revolución que ya está cambiando la forma de hacer negocios, la sanidad, la política, la educación y la innovación (Mayer-Schonberger & Cukier, 2013).

La gran diversidad de aplicaciones y la creciente demanda de expertos en la materia, ha hecho que proliferen los cursos, libros y manuales en DS. El método pedagógico estándar que podemos apreciar en estos cursos consisten en: (1) la explicación de las distintas ramas de aprendizaje automático (supervisado, no supervisado, por refuerzo); (2) el detalle de algún paradigma de aprendizaje como árboles de decisión o redes de neuronas artificiales; y, (3) la ilustración mediante conjuntos de datos de juguete como Weather o Iris. Este método es empleado en algunos cursos de DS de gran aceptación en la comunidad como Machine Learning de la Universidad de Stanford¹, o la especialización en DS de la Universidad Johns Hopkins².

La existencia de distintos repositorios de datos sobre los que construir conocimiento ofrece un caldo de cultivo privilegiado para diseñar un curso de DS como una serie de experiencias en problemas del mundo real. Pocos campos permiten al estudiante ponerse en la piel de perfiles tan diversos e interesantes como DS: economistas, administradores de empresas, médicos, biólogos, administradores de sitios webs, etcétera. De la misma manera, pocas disciplinas pueden ofrecer recompensas tan atractivas para el aprendizaje experiencial como los tres millones de dólares que obtuvo el ganador del concurso para predecir los pacientes que eran admitidos en un hospital estadounidense en el siguiente año³; o el millón de dólares con el que la compañía Netflix⁴ premió al mejor predictor de valoraciones de películas.

En este trabajo, se presentan tres modelos de enseñanza y aprendizaje de DS que, inspirados en el *aprendizaje experiencial* (Jacobson & Ruddy, 2004), buscan que el estudiante aprenda mediante la reflexión sobre la experimentación, en lugar de ser recipientes pasivos de contenido indiscutible. Además, dichos modelos han sido utilizados en una asignatura centrada uno de los campos más punteros y complejos de DS, el Aprendizaje profundo o *Deep Learning* (DL) (Goodfellow, Bengio, & Courville, 2016).

¹ <https://www.coursera.org/learn/machine-learning>

² <https://www.coursera.org/specializations/jhu-data-science>

³ <https://www.kaggle.com/c/hhp>

⁴ <http://www.netflixprize.com/>

Finalmente, los estudiantes han sido encuestados para evaluar y elegir sus formas de enseñanza y aprendizaje favoritas.

2. CONTEXTO

Este artículo describe experiencias desarrolladas para el aprendizaje de DS en general, y DL en particular. En el DL, se parten de datos en bruto y, mediante arquitecturas compuestas de varias capas que realizan transformaciones no lineales de estos datos, se van obteniendo representaciones cada vez más abstractas. Por ejemplo, se pueden partir de los píxeles de una imagen como datos en bruto y, a través de distintas capas, aprender a distinguir: líneas, formas, y elementos más abstractos.

Las experiencias están diseñadas para estudiantes del “Master in Data Science (EIT Digital Master School)” impartido en la ETS de Ingenieros Informáticos de la Universidad Politécnica de Madrid (ETSINF). El perfil de estos estudiantes es eminentemente técnico, por lo que tienden a centrarse en aspectos de la programación de DS en detrimento de otras grandes disciplinas como la estadística o la reflexión sobre una aplicación concreta. Precisamente por ello, se considera que los modelos de enseñanza y aprendizaje aquí propuestos aportan una mejora significativa en el aprendizaje al centrarse en la solución de problemas concretos.

3. DESCRIPCIÓN

En esta sección se describen los tres modelos propuestos para la enseñanza y el aprendizaje en DS incluyendo las actividades y recursos utilizados.

A. Redes de neuronas artificiales (ANN).

En este modelo de enseñanza y aprendizaje, primero se introducen las arquitecturas neuronales, que son la piedra angular del campo de DL. Muchos de los grandes avances de la inteligencia artificial del siglo veintiuno se basan en la utilización de este paradigma de aprendizaje automático. Las ANN son modelos computacionales inspirados en la neurociencia capaces de predecir una salida a partir de datos etiquetados (*aprendizaje supervisado*), así como de encontrar estructuras subyacentes y ocultas en datos no etiquetados (*aprendizaje no supervisado*) (Witten, Frank, & Hall, 2011). En segundo lugar, se dan consejos prácticos en la solución de problemas con ANN y se presentan dos entornos de trabajo de DS: Weka⁵ y H2O.ai⁶. Finalmente, se propone una experiencia a los estudiantes: entrenar y evaluar arquitecturas neuronales para un problema concreto, proporcionando un conjunto de datos que permita inferir un modelo inteligente de predicción, como los que se encuentran disponibles en el repositorio UCI⁷.

El modelo de enseñanza y aprendizaje descrito sigue el flujo clásico de un curso en DS, aunque con una marcada componente práctica y experiencial, donde se pide abordar un problema concreto y realista de gran interés social.

B. Visión artificial (CV).

Este segundo modelo de enseñanza y aprendizaje aborda el problema de la visión artificial, es decir, la capacidad de los

⁵ <http://www.cs.waikato.ac.nz/ml/weka/>

⁶ <https://www.h2o.ai/>

⁷ [https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Original\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original))

computadores para obtener entendimiento a alto nivel a partir de imágenes o vídeos. Uno de los aspectos fundamentales de este modelo es que se busca que los estudiantes comprueben que las lecciones aprendidas en ANN no permiten resultados aceptables para la visión artificial. En consecuencia, tras una introducción al campo, se propone a los estudiantes predecir el objeto que contiene una imagen. Para ello, se facilita el conjunto de datos CIFAR-10⁸, que contiene 60K imágenes a color en tamaño 32x32 con objetos pertenecientes a diez categorías distintas (*clases* en terminologías de aprendizaje supervisado). Además, esta tarea se propone en la modalidad de concurso para incluir un esquema de *gamificación*: el ranking de las mejores predicciones se muestra en la plataforma virtual de la asignatura. Gracias a esto, los estudiantes pueden reflexionar sobre los problemas que llevan a las ANNs a tener bajos porcentajes de predicciones correctas. Ningún estudiante superó el 57% de aciertos.

Tras este concurso, se introducen las *redes de neuronas convolucionales (ConvNets)* y; en un enfoque práctico, se explica un entorno de trabajo para el diseño, entrenamiento, y uso de estas redes llamado Caffe⁹. La visión artificial requiere modelos específicos que sean capaces de, entre otros, aprovechar la información espacial de imágenes y vídeos: un pixel cercano a otro probablemente pertenecerá al mismo objeto. Las ConvNets han obtenido los mejores resultados en CV desde que en 2012 fuesen usadas en la competición *ImageNet* para reducir el error de predicción del 26% al 15%.

Después de introducir las ConvNets, se vuelve a plantear la misma experiencia en un nuevo concurso. Esta vez algunos estudiantes alcanzaron un 80% de predicciones correctas. Finalmente, se plantea una tercera experiencia en la que se reduce considerablemente los datos de entrenamiento, pero se permite utilizar ConvNets ya entrenadas para ajustarlas a los nuevos datos, i.e. *aprendizaje por transferencia* en terminología de DS.

La clave de este segundo modelo de enseñanza y aprendizaje para CV es que los estudiantes ven su conocimiento previamente adquirido desafiado ante un nuevo problema. Se tiene tiempo de intentar aplicar métodos conocidos a nuevas situaciones; reflexionar sobre los resultados; realizar la conceptualización del nuevo conocimiento; y comprobar experimentalmente las nuevas ideas en el siguiente concurso. De esta manera, se cubren las cuatro fases del ciclo de Kolb para el aprendizaje experiencial (Kolb, 1983). Asimismo, los concursos permiten un aprendizaje experiencial guiado por juegos como propone Shiralkar (Shiralkar, 2016) para el campo de las tecnologías de la información. Sin embargo, en este modelo, los métodos y soluciones utilizadas por los estudiantes están guiados y no tienen libertad para la propuesta de soluciones innovadoras o la investigación de métodos distintos a los tratados durante el curso.

C. Procesamiento del Lenguaje Natural (NLP).

Este modelo aborda otro problema de gran complejidad en DS: el procesamiento y entendimiento del lenguaje natural por las computadoras. El NLP, junto a la CV, tratada en el módulo anterior, son dos de los problemas más difíciles de la inteligencia artificial. Llamamos a estos problemas AI-

⁸ <https://www.cs.toronto.edu/~kriz/cifar.html>

⁹ <http://caffe.berkeleyvision.org/>

complete, o *AI-hard*, ya que buscan que las computadoras lleguen a ser tan inteligentes como las personas. Las raíces del NLP se encuentran en el trabajo “Computing Machinery and Intelligence” donde Alan Turing especulaba sobre los requisitos que debería cumplir una máquina para poder afirmar que es realmente inteligente. Desde 2014, los métodos de DL han liderado los resultados en NLP para una gran variedad de tareas como responder preguntas, traducir automáticamente a distintos idiomas, resumir grandes documentos o generar textos automáticamente.

Para este modelo de enseñanza y aprendizaje, se invierte el orden de los modelos anteriores: primero se da una experiencia realista a los estudiantes y luego se ofrecen algunas soluciones alternativas. Más concretamente, se trata el problema de la *clasificación de texto*: se busca un modelo predictivo que sea capaz de catalogar un texto en una clase específica. La clasificación de texto es una generalización de problemas como los filtros de SPAM, que existen en las cuentas de correo, el análisis de sentimientos de redes sociales que realizan compañías y gobiernos para detectar tendencias en la opinión pública, o la predicción de la temática de un documento. Este problema es muy desafiante con los conocimientos previos de la asignatura ya que, a diferencia de las ANN o la CV, la obtención de un vector de características de entrada de igual longitud para distintos textos no es trivial.

En esta experiencia, los estudiantes forman grupos de hasta cinco personas, a modo de grupos de investigación, en un laboratorio equipado con un puesto de trabajo por persona. Su tarea consiste en realizar una propuesta para una convocatoria competitiva de un proyecto de investigación en una gran compañía. El problema de investigación consiste en predecir la relevancia del título de un artículo respecto al cuerpo de este, basándose en un conjunto de datos con tres atributos: título, cuerpo, y clase (relevante/irrelevante)¹⁰. Además, para facilitar la exploración de diversas soluciones innovadoras, en lugar de pedir una implementación de código como en los modelos anteriores, se solicitan diagramas de flujo de trabajo para las propuestas así como los entornos de trabajo que se van a usar y referencias bibliográficas.

Los estudiantes, entonces, presentan sus propuestas que, si bien pueden incluir fallos en cuanto a las interfaces de las tecnologías utilizadas, fueron definitivamente muy creativas y hacían un buen uso de muchos de los conceptos ya vistos en DL y otras asignaturas del máster. En esta línea, algunos estudiantes plantearon el uso de ConvNets, estudiadas en CV, para el análisis de texto. Otras propuestas incluyeron tecnologías vanguardistas de NLP como el uso de redes neuronales recurrentes y formas densas de representación de palabras como Word2Vec y GloVe.

Posteriormente, los estudiantes deben revisar un extracto de lecturas seleccionadas (unas 5 páginas) con una introducción al campo del NLP, al problema de la clasificación de textos, y a plataformas de desarrollo de DL. Finalmente, se imparte una clase magistral donde se tratan diversas soluciones para el problema planteado. También se contextualizan estas soluciones con las distintas propuestas realizadas por los estudiantes.

¹⁰ El dataset era hipotético y no se proporcionó, a excepción de un ejemplo de caso.

Como en CV, este modelo de enseñanza y aprendizaje parte de un problema desafiante y que no se puede resolver con el conocimiento previo. El texto requiere un preproceso muy elaborado antes de poder alimentar a una red neuronal con él. Además, la petición de un diseño en lugar de una implementación es más complicada para los estudiantes del máster, dado su perfil técnico. También es clave en este modelo dar libertad y confianza a los estudiantes para que exploren distintas alternativas y soluciones innovadoras. Finalmente, el contenido de la clase magistral debe ser flexible y abordar los principales conceptos usados en las propuestas de los estudiantes, en lugar de limitarse a presentar una serie de soluciones no debatibles.

4. RESULTADOS

El objetivo del estudio es analizar la evolución de la satisfacción de los estudiantes para cada uno de los tres modelos de enseñanza y aprendizaje implementados. Como se describió en la sección anterior, los modelos se corresponden a las tres partes en que se divide la asignatura Deep Learning: redes de neuronas artificiales (ANN), visión artificial (CV), y procesamiento del lenguaje natural (NLP). Al final del curso, se suministró un cuestionario a los 24 estudiantes presentes de un total de 29 matriculados. Para cada uno de los modelos, se plantearon tres afirmaciones: A1 “El contenido del curso satisface mis necesidades de formación”, A2 “Lo que he aprendido será aplicable a mi trabajo”, A3 “La metodología aplicada, los recursos técnicos, y los materiales docentes eran apropiados”. A los participantes se les pidió que calificaran su acuerdo o desacuerdo en una escala Likert de 5 puntos. Estas preguntas son similares a las realizadas por Alonso et al. (Alonso, López, Font, & Manrique, 2010). Si bien, en este trabajo, se ha optado por una escala de 5 puntos en lugar de 6, dado que está más extendida y se consideró la posibilidad de dar una respuesta equidistante entre el máximo y el mínimo.

También se plantearon preguntas en base a un modelo de aprendizaje experiencial puro, donde no hay material suministrado ni ninguna guía a la hora de abordar una experiencia realista. “Considere un cuarto modelo en el cual sólo una experiencia realista es propuesta al principio de la asignatura, Deep Learning por ejemplo. En este caso, los estudiantes, todos juntos formando un grupo, serían libres de implementar e investigar los mejores métodos para completar esta tarea”. Este modelo es hipotético, en el sentido que no se siguió durante la asignatura, pero consideramos que la opinión de los estudiantes resulta valiosa para nuevas ediciones del máster. Para este modelo de aprendizaje, se plantearon otras tres afirmaciones: A4 “Me gustaría que al menos una asignatura siguiera este esquema”, A5 “Me gustaría que al menos una asignatura siguiera este esquema, pero con grupos más pequeños”, A6 “Me gustaría que todo el máster siguiera este esquema”. Una vez más, se pidió que calificaran su acuerdo o desacuerdo en una escala Likert de 5 puntos.

La Tabla 1 muestra la estadística descriptiva para la satisfacción de los estudiantes, siendo la primera columna el modelo de enseñanza y aprendizaje empleado (ANN, CV, NLP, o aprendizaje experiencial puro), y las declaraciones anteriormente mencionadas, numeradas de A1 a A6.

Tabla 1. Evaluación de modelos de enseñanza y aprendizaje

| | Mean | Std.Dev. |
|------------------|------|----------|
| ANN | | |
| A1 | 3,38 | 1,06 |
| A2 | 3,58 | 0,78 |
| A3 | 3,08 | 0,97 |
| CV | | |
| A1 | 4,50 | 0,59 |
| A2 | 4,46 | 0,51 |
| A3 | 4,00 | 0,88 |
| NLP | | |
| A1 | 4,13 | 0,68 |
| A2 | 4,08 | 0,97 |
| A3 | 4,21 | 0,93 |
| Exp.Puro. | | |
| A4 | 3,83 | 1,17 |
| A5 | 3,88 | 1,15 |
| A6 | 3,00 | 1,29 |

Respecto a los tres primeros modelos, se puede apreciar en todos los casos que la media de las respuestas dadas a las tres afirmaciones está por encima de la mitad de la escala (3). También destaca una desviación típica muy baja, inferior a un punto, excepto en un caso. Esto indica una escasa variación esperada respecto a la media aritmética y, sin la necesidad de un test ANOVA, que las medias dentro de los distintos modelos evaluados son similares. En cuanto a la comparación entre los distintos modelos, CV es el mejor valorado en las afirmaciones A1 y A2, mientras que NLP es la mejor valorada en A3, que incluye los aspectos metodológicos. En ese sentido, los estudiantes parecen haber apreciado las ventajas del aprendizaje experiencial y la libertad para investigar y aplicar soluciones no guiadas a situaciones realistas. El modelo de enseñanza y aprendizaje empleado en ANN, que es el más cercano a la enseñanza tradicional de DS, es el peor valorado, aunque superando la mitad de la escala. Hay casi un punto de diferencia entre ANN y el máximo alcanzado en la pregunta por los otros modelos (CV o NLP).

Se destaca también, a partir de los resultados obtenidos en la Tabla 1, que no hay un fuerte apoyo del modelo de aprendizaje experiencial puro, al menos en los términos planteados en las afirmaciones A4-A6. Los resultados son superiores a 3 para A4 y A5, y neutros para A6. La naturaleza profesional del máster podría justificar que el estudiante recelase de asignaturas donde no se proporciona material de apoyo aunque, como en el caso de NLP, este material se proporcione a posteriori, tras realizar una experiencia realista sin guías y ajustando el material a algunas de las soluciones propuestas por los estudiantes anteriormente.

Con el objetivo de profundizar más en la comparación entre los distintos modelos, se ha planteado a los estudiantes una votación donde para cada uno de los modelos empleados en la asignatura se proponen las siguientes afirmaciones en cuanto a la metodología empleada: A7 “Mi metodología favorita es la usada en”, A8 “La metodología que me permite aprender más es”, A9 “La metodología que (creo) que me permite un aprendizaje más duradero es”, A10 “La metodología que

considero más cercana al trabajo diario de un científico de datos es”. Las respuestas para cada una de estas declaraciones pueden ser ANN, CV, o NLP. Se decidió omitir la columna de “no sabe / no contesta” para forzar una respuesta no ambigua, considerándose que siempre debía haber un modelo favorito. Igualmente, se omitió la posibilidad de votar varias opciones simultáneamente. Sin embargo, algunos estudiantes decidieron realizar votaciones múltiples y estas respuestas se contabilizaron como votos extra para los modelos elegidos.

Tabla 2. Votación entre distintos modelos

| | ANN | CV | NLP | NºV |
|-----|-----|----|-----|-----|
| A7 | 1 | 17 | 7 | 25 |
| A8 | 0 | 19 | 6 | 25 |
| A9 | 1 | 19 | 4 | 24 |
| A10 | 3 | 15 | 11 | 29 |

La Tabla 2 muestra los resultados de la votación. Se incluye una columna NºV con los votos totales realizados para cada una de las afirmaciones. Como se puede ver, el modelo de visión artificial (CV) es superior en las cuatro afirmaciones, alcanzando la mayoría absoluta en todas las categorías, i.e. más de la mitad de los votos totales. En consecuencia, los estudiantes parecen haber apreciado los aspectos metodológicos innovadores de este modelo, como son: la experimentación con el conocimiento previo en una tarea nueva y desafiante; y la *gamificación* mediante concursos para enfrentar a los estudiantes en una competición con criterios de evaluación objetivos. Cabe destacar también que el modelo de NLP obtiene un número de votos muy cercano a CV en la afirmación A10. Este modelo NLP es cercano al aprendizaje experiencial puro, en el sentido de que se da libertad a los estudiantes para buscar soluciones propias en la resolución de la tarea en cuestión. Por consiguiente, los estudiantes parecen haber identificado este modelo experiencial, junto a CV, como algo muy cercano al día a día de un científico de datos.

5. CONCLUSIONES

En este artículo se han descrito tres modelos de enseñanza y aprendizaje que permiten a los estudiantes obtener nuevo conocimiento iterando sobre experiencias realistas, reflexión sobre ellas, conceptualización del nuevo conocimiento y experimentación. Los modelos están diseñados de manera que se va incrementando la responsabilidad de los estudiantes en la construcción de su propio conocimiento, dándose libertad total a la hora de proponer soluciones en la última propuesta.

Aunque los tres modelos han recibido evaluaciones superiores a la mitad de la escala empleada (3) por parte de los estudiantes, muestran un claro favorito: la metodología utilizada en visión artificial. En este modelo, se proponen tres experiencias realistas bajo la modalidad de concurso que se intercalan con clases magistrales. Tanto las encuestas anónimas realizadas como el material empleado en las experiencias propuestas en este trabajo están disponibles para el lector interesado solicitándolo a los autores.

Los modelos explicados están diseñados para un curso de Deep Learning pero son generalizables a otros cursos de Ciencia de Datos. Igualmente, se espera extrapolar estos modelos a cursos con estudiantes de perfiles distintos y menos técnicos que los del máster internacional de Data Science. Más

concretamente, se ha planificado su uso en el máster de Biología Computacional en la Universidad Politécnica de Madrid.

AGRADECIMIENTOS

Este trabajo de investigación está financiado por la Universidad Politécnica de Madrid, bajo el proyecto de innovación educativa: “Métodos, experiencias y herramientas para el aprendizaje experiencial de la Ciencia de Datos”; y por el Ministerio de Economía, Industria y Competitividad en el ámbito del proyecto “Datos 4.0: Retos y soluciones – UPM” (TIN2016-78011-C4-4-R, AEI/FEDER, UE).

REFERENCIAS

Alonso, F., López, G., Font, J. M., & Manrique, D. . (2010). Learner satisfaction when applying an instructional model in e-learning: an experimental study. *Proceedings of the 2nd International Conference on Computer Supported Education - Volume 1: CSEDU*, 141-146.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Jacobson, M., & Ruddy, M. (2004). *Open to Outcome: A Practical Guide for Facilitating & Teaching Experiential Reflection*. Wood 'N' Barnes.

Kolb, D. A. (1983). *Experiential Learning: Experience as the Source of Learning and Development*. Prentice Hall.

Mayer-Schonberger, V., & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt.

Shiralkar, S. (2016). *IT Through Experiential Learning: Learn, Deploy and Adopt IT through Gamification*. Apress.

Witten, I., Frank, E., & Hall, M. (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier Science.