

SWAN como herramienta para la docencia de física de altas energías

SWAN as tool for teaching High Energy Physics

Víctor Rodríguez Bouza (vrbouza@uniovi.es),*
José Enrique Palencia Cortezón (enrique.palencia@uniovi.es)*

Palabras clave: CERN, SWAN, nube, docencia, física de partículas, física de altas energías, HEP, openUp2U, CERNBox, Jupyter.

Relevancia del artículo para con los objetivos de la Revista: puesto que en este texto se expone una novedosa herramienta que permite realizar labores docentes de una forma diferente y con la que se alcanzaron buenos resultados, siendo el fin del mismo el dar a conocer la experiencia a otro profesorado que lo pueda encontrar útil, entendemos que conectamos plenamente con los objetivos de la revista, en particular el objetivo (iii).

*Universidad de Oviedo - Instituto de Ciencias y Tecnologías Espaciales de Asturias (ICTEA). C/ Federico García Lorca 18, 33007 Oviedo/Uviéu, Asturias (España).

Resumen

In this article we present the results of a teaching innovation project based on the use of the tool SWAN (Service for Web-based ANalysis), developed by CERN in the teaching of experimental high energy physics. It has been applied upon a fourth-year subject of the Degree on Physics, in which part of the usual reports to do have been replaced by practical exercises using SWAN. The main task was based on an analysis of the process in which a Higgs boson decays into two tau leptons using open data from the CMS Collaboration, to which we have added a short introduction to machine learning techniques. Despite of the current situation due to the COVID-19 pandemic, and the low number of students enrolled this academic year in the subject, the results show that the pupils appreciate learning these topics and positively evaluate the use of SWAN, showing it as useful to ease the learning.

1. Introducción

Hoy en día muchas herramientas informáticas empleadas tanto en ciencia y tecnología, como en casi cualquier área profesional, están «deslocalizándose» de nuestros ordenadores personales para ubicarse en la nube. Este movimiento presenta una serie de ventajas incuestionables, entre las que podemos contar la eliminación de la necesidad de poseer la herramienta (*software* y *hardware*) individualmente, a cambio de tener una conexión a Internet. Esto presenta oportunidades no solo para el día a día, en el uso productivo de estos instrumentos, sino también para, por ejemplo, la docencia y la divulgación científica.

Una de las herramientas que se han hecho populares en los últimos tiempos han sido los cuadernos de Jupyter [1], que permiten programar en varios lenguajes (aunque sobre todo se emplea con Python) a través de casillas de código (celdas), de un modo similar al de otro *software* como wxMaxima o Mathematica. Además, se accede al programa a través de un navegador de Internet, estando el editor embebido en una página web que se puede ejecutar en el mismo ordenador desde el que se accede, o se puede correr en un servidor dedicado. Y aquí radica una de sus mayores potencialidades: el poder «externalizar» la ejecución del código y la necesidad de, tan solo, un navegador web para programar. La gran popularidad de este tipo de cuadernos está bastante correlacionada con el auge de la ciencia de datos, el *big data*, aprendizaje automático o inteligencia artificial en los últimos años, siendo casi omnipresente en estas áreas el uso de cuadernos Jupyter, y muy especialmente en el aprendizaje de las mismas.

En este contexto, el CERN [2], el centro europeo de investigación en física de altas energías, ha desarrollado una nueva herramienta: **SWAN** (*Service for Web-based ANalysis*) [3]. Un programa que, basándose en los cuadernos Jupyter, permite crear un servidor de los mismos modificado y mejorado, que está internamente conectado con su *software* de almacenamiento en la nube, CERN-Box [4], y que emplea EOS [5] como gestor interno de datos: el mismo programa que se usa para almacenar y mover todos los datos de colisiones del *Large Hadron Collider*: el mayor colisionador de partículas del planeta ubicado en el CERN. Además de la interrelación con CERNBox y EOS, SWAN incluye mejoras en el editor de los cuadernos Jupyter normales, nuevos kernel que permiten correr por ejemplo el intérprete de C++ de ROOT [6] (la librería informática por excelencia para física experimental de partículas), así como acceso a todo el *software* necesario en este campo a través de CernVM-FS [7].

Siguiendo la filosofía respecto al *software* del CERN, SWAN es *software* gratuito y abierto. Junto a la gran cantidad de datos de acceso público que el CERN, dentro de su portal de ciencia abierta [8], ha sacado a la luz, esto muestra una gran potencialidad para el uso de estos datos y de la herramienta para fines divulgativos y docentes. Todo ello unido a otras características de la misma: facilidad de uso (solo se necesita un navegador y conexión a Internet), herramienta popular en la industria y la academia (sobre todo en el campo de ciencia de datos) o que permite una mayor interactividad a través de la ejecución por celdas con el código o lo que se quiera transmitir con un cuaderno.

Debido a estas características, nos propusimos el pasado curso 2019-2020 desarrollar un proyecto piloto para explorar el uso de herramientas como esta en la docencia de física de altas energías.

2. Aplicación a la docencia y desarrollo del plan

El proyecto se implementó en la asignatura de forma voluntaria, de modo que los alumnos interesados no tendrían que llevar a cabo las entregas de los trabajos habituales. Los estudiantes

tenían que realizar varios ejercicios empleando las herramientas, que les fueron explicadas en clase expositiva brevemente. Para hacer uso de ellas, tan solo debían conectarse a través de su navegador web a la plataforma en línea OpenUp2U [9], que fue el medio escogido para realizar el proyecto. En esa plataforma, haciéndose una cuenta de usuario, tenían acceso a un servidor de SWAN para poder hacer las tareas.

Los ejercicios que el estudiantado tuvo que hacer están basados en un análisis realizado con datos en abierto del experimento CMS [10] de búsqueda de desintegraciones de bosones de Higgs a dos leptones tau [11]. Aprovechando el trabajo realizado, creamos cuadernos de Jupyter en los que los estudiantes tenían parcialmente desarrollados ejemplos de código con los que trabajar y entender la física de los procesos que aparecían, los observables físicos, ejemplos de uso de modelos de aprendizaje automático, etc.. Los cuadernos venían con ejemplos preconfigurados y explicaciones intercaladas en **Markdown** con vínculos a más referencias para entender mejor el código y las herramientas. Ejemplos de estos cuadernos, tal y como se observan en el navegador, se muestran en la figura 1.

Los ejercicios eran principalmente dos: en el primero, se hacían distintas preguntas en referencia al análisis que se estudiaba (la desintegración de bosones de Higgs a dos leptones tau), pidiéndoles entender distintos observables físicos y que intentaran entender cómo afectaba la física de cada proceso físico de fondo en cada uno de ellos. En el segundo ejercicio, se les pedía que intentaran aplicar un modelo de aprendizaje automático a estas simulaciones, para intentar clasificar los sucesos recogidos que eran en efecto desintegraciones de bosones de Higgs a leptones tau (la señal) de los procesos de fondo.

Al final del curso, se pidió que hicieran una breve presentación con lo que habían realizado y aprendido. Para hacer las exposiciones más interesantes para el alumnado, se hicieron variantes de cada ejercicio (cambiando e.g. el modelo de aprendizaje automático a entrenar, o los observables físicos a estudiar), de forma que lo que otros estudiantes resultara más interesante para ellos a la hora de presentarlos. Finalmente, solicitamos que rellenaran una encuesta para recoger su experiencia con el proyecto.

3. Resultados

Antes de comentar las valoraciones recibidas, cabe destacar que el curso en el que se hizo estuvo claramente caracterizado por la pandemia por COVID-19, que afectó a las clases y vidas tanto de alumnado como profesorado. En nuestro caso, pudimos continuar a distancia con la asignatura, pero con un ritmo más lento. Tampoco se pueden eliminar los efectos indirectos (físicos o psíquicos) que sobre las personas mismas pudo tener la pandemia.

En cuanto a la propia realización de los ejercicios, los estudiantes en general consiguieron terminarlos y las puntuaciones globales de los mismos arrojan una media de 8 sobre 10. Cabe destacar que debido a cuestiones técnicas, y también indirectamente por la pandemia, el alumnado dispuso de menos tiempo del que en principio se quería para hacer los ejercicios, pero para compensar por ello y por el contexto global, se redujeron los requisitos de los mismos y se amplió en el tiempo todo lo posible la fecha de entrega.

Respecto a las presentaciones, cabe mencionar positivamente la sesión de defensa (por videollamada) y la diversificación de los ejercicios con el uso de distintas técnicas y el estudio de diferentes observables, pues al final apreciamos cierto interés por todo el alumnado en comprender lo que sus compañeros/as habían visto.

Las respuestas de la encuesta propuesta como parte del proyecto (todo el alumnado contestó a la misma) muestran valoraciones globalmente positivas. A la pregunta «en una escala del uno al cinco, (siendo uno lo peor y cinco lo mejor), ¿cuán útiles te parecen, globalmente, las siguientes

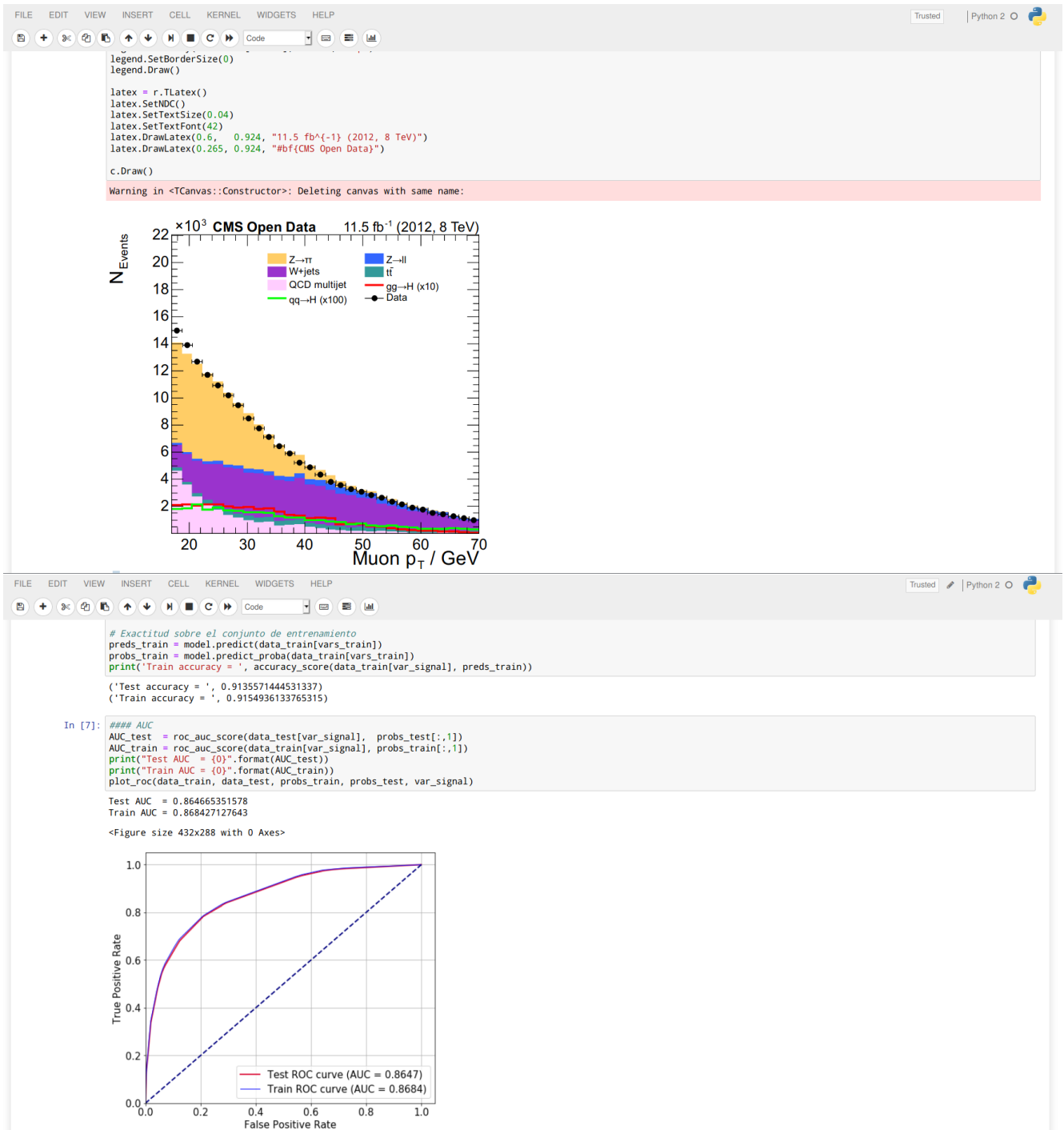


Figura 1: Visualización de los cuadernos de Jupyter dados a los estudiantes para realizar el proyecto.

herramientas para hacer prácticas de clase?», de media respondieron con un 3.8 en el caso de SWAN y en el caso de los cuadernos Jupyter con un 4.0. Resulta significativo también que a la pregunta «¿cuán positivo o útil crees que enseñemos herramientas de aprendizaje automático o inteligencia artificial...» de media se respondiera con un 4.2 en el caso de «...en esta asignatura?» y 4.6 en el caso de «...en alguna asignatura optativa / de último curso del Grado en Física?», lo que es muestra del interés por el alumnado por los contenidos de aprendizaje automático. En la figura 2 se pueden observar resumidos los resultados.

Finalmente, la encuesta incluía un recuadro para incluir sugerencias o comentarios de forma optativa. Un estudiante hace referencia a la falta de tiempo para hacer estos ejercicios, diciendo en

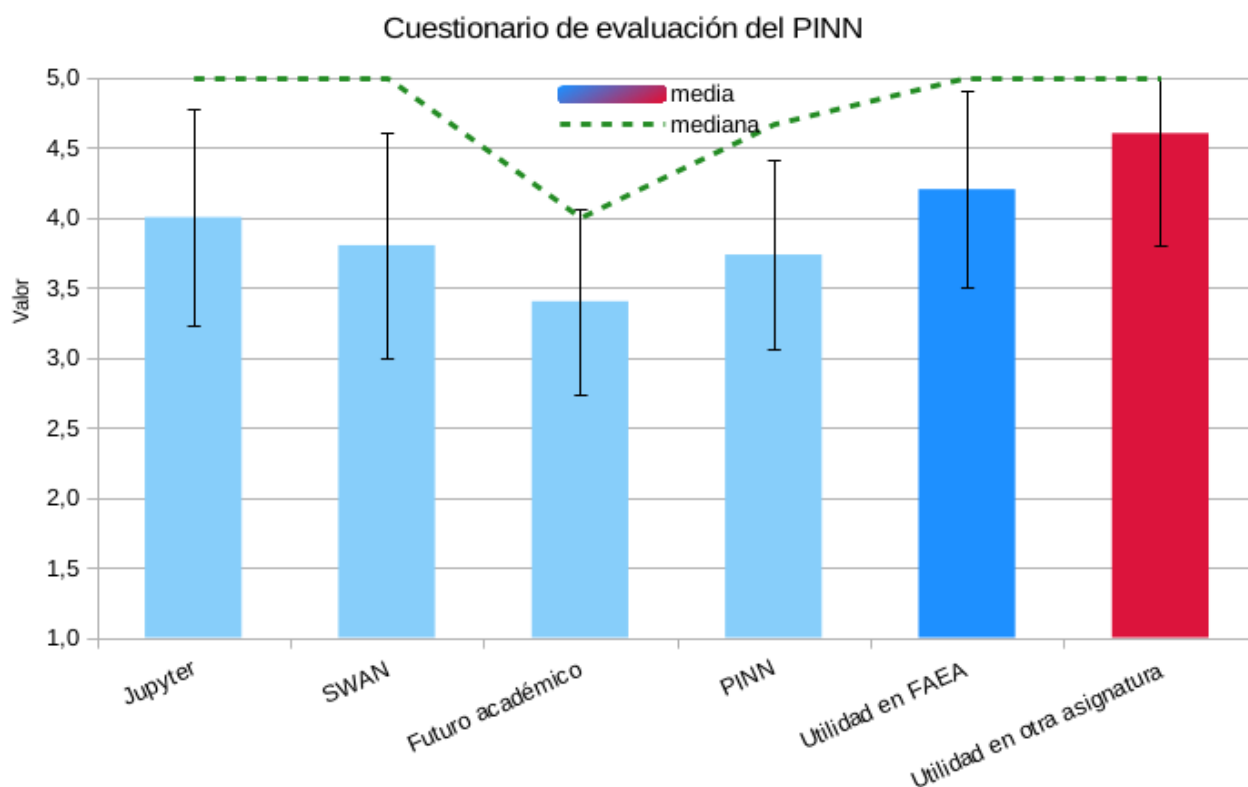


Figura 2: Gráfico de barras que muestra la media y la mediana de las respuestas a la valoración del uno al cinco de distintos aspectos del proyecto: los cuadernos Jupyter, SWAN, si cree que será útil para su futuro académico, el proyecto en general, la utilidad de contenidos de aprendizaje automático en la asignatura (Física de Altas Energías y Aceleradores) y la utilidad en otra asignatura.

un caso que «le dedicamos una parte muy pequeña del semestre cuando realmente no estamos familiarizados con ello». Otro destaca que esta herramienta es «muy útil para pensar un poco menos en la programación y mucho más en la física detrás de los procesos con los que se trabaja». Otro se congratula del cambio de método pedagógico, frente al tradicional informe de práctica o trabajo («me parece muy "refrescante" que [...] se haya propuesto sustituir el enésimo informe que haríamos por algo distinto y más liviano») y menciona que le haya motivado para aprender autónomamente sobre los contenidos de los ejercicios («he acabado dedicando bastante tiempo a leer sobre ello solo movido por la propia curiosidad, cosa que llevaba mucho tiempo sin ocurrirme, y he conseguido comprender conceptos fundamentales a la par que importantes simplemente por el afán de saber»).

Los comentarios recogidos vienen a corroborar en cierta forma algunas apreciaciones que ya hicimos en párrafos anteriores de esta sección, como el poco tiempo que finalmente dispusieron para hacer los ejercicios, frente al inicialmente contemplado. También el interés que suscitaron los contenidos en el alumnado y el hecho de que a través de estos cuadernos Jupyter ciertos aspectos de la física experimental de altas energías podrían ser más accesibles a quien los lee (y ejecuta).

4. Conclusiones

La retroalimentación recibida y explicada en la sección anterior, así como el desarrollo del propio proyecto nos permite decir que lo hemos culminado de forma exitosa, a pesar de los impedimentos directos e indirectos de la pandemia por COVID-19.

SWAN y los cuadernos de Jupyter no constituyen una «revolución» para enseñar física experimental de altas energías ni mucho menos, ni tampoco constituyen una guía sobre la que hacer docencia de este área científica. Sin embargo, hemos constatado que sí pueden ser buenas herramientas para ayudar en esta labor. Las buenas valoraciones recibidas nos animan a intentar explorar más utilidades de estas herramientas en cursos próximos.

Referencias

- [1] Página web oficial del proyecto Jupyter: <https://jupyter.org> .
- [2] Página web oficial del CERN: <https://cern.ch> .
- [3] Página web oficial de SWAN: <https://swan.web.cern.ch/> .
- [4] Página web oficial de CERNBox: <https://cernbox.web.cern.ch> .
- [5] Página web oficial de EOS: <https://eos-web.web.cern.ch> .
- [6] Página web oficial de ROOT: <https://root.cern.ch/> .
- [7] Página web oficial de CernVM-FS: <https://cernvm.cern.ch/portal/filesystem> .
- [8] Portal oficial de ciencia abierta del CERN: opendata.cern.ch .
- [9] Portal oficial de OpenUp2U: <https://up2university.eu/open/> .
- [10] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, [doi:10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- [11] *Analysis of Higgs boson decays to two tau leptons using data and simulation of events at the CMS detector from 2012* [http://opendata.cern.ch/record/12350](https://opendata.cern.ch/record/12350) .