

**UNIVERSIDAD DE CASTILLA-LA MANCHA**

ESCUELA SUPERIOR DE INGENIERÍA INFORMÁTICA

**ANTEPROYECTO DEL TRABAJO FIN DE GRADO**

**GRADO EN INGENIERÍA INFORMÁTICA**

**TECNOLOGÍA ESPECÍFICA / INTENSIFICACIÓN / ITINERARIO DE COMPUTACIÓN**

**Predicción de precios de instancias “spot” en la nube de Amazon**

Autor: Luis Mendoza Montero

Directora: Carmen Carrión Espinosa

Directora: Mª Blanca Caminero Herráez

Mayo, 2018

**ÍNDICE DE CONTENIDOS**

1. INTRODUCCIÓN. 3

2. TECNOLOGÍA ESPECÍFICA CURSADA POR EL ALUMNO. 5

3. OBJETIVOS 6

4. MÉTODO Y FASES DE TRABAJO 7

5. MEDIOS QUE SE PRETENDEN UTILIZAR 8

5.1. Medios Hardware 8

5.2. Medios Software 9

6. REFERENCIAS 9

**ÍNDICE DE TABLAS**

[*Tabla 1. Tecnología Específica cursada por el alumno* 5](#_Toc513295072)

[*Tabla 2. Justificación de las competencias específicas abordadas en el TFG* 5](#_Toc513295073)

**ÍNDICE DE FIGURAS**

[*Ilustración 1. Resumen de la organización del trabajo* 5](#_Toc513295075)

[*Ilustración 2. Cronograma* 7](file:///C:\Users\Luis\Desktop\TFG\Anteproyecto\Anteproyecto%20Luis%20Mendoza.docx#_Toc513295076)

# 1. INTRODUCCIÓN.

*Amazon Web Services (AWS)* [1] es una plataforma que ofrece un conjunto de servicios fiables, escalables y económicos que se ofrecen en la nube pública [2] [3] con el fin de adquirir potencia de cómputo, almacenamiento de bases de datos, entrega de contenido y otras funcionalidades de gran ayuda para el mundo empresarial. Por tanto, ofrece un completo servicio de infraestructura en la red para las tecnologías de la información. Esta agrupación de servicios de Amazon ofrece inicialmente un acceso de prueba gratuito, aunque con limitaciones a algunos recursos.

En este Trabajo Fin de Grado nos centraremos en el servicio web de *Amazon Elastic Compute Cloud (Amazon EC2)* el cual proporciona capacidad informática en la nube para los desarrolladores. Los datos que necesitaremos para el desarrollo del trabajo serán los que Amazon pone a su disposición en la nube y que nosotros obtendremos gracias a los mecanismos que se irán desarrollando durante la fase de trabajo. Estos datos estarán distribuidos geográficamente en varios *data centers* organizados por regiones. Cada región representa un lugar determinado donde se encuentra disponible el tipo de instancia al que nos referimos [4]. Los datos que este servicio ofrece para el del usuario o desarrollador se estructuran por diferentes registros representando máquinas virtuales (herramientas software que simulan un computador virtual como si se tratase de uno real) con diferentes características respecto al sistema operativo (SUSE Linux, UNIX o Windows) [5] y con unas características determinadas en función de su familia. Las familias se estructuran en función de unas características generales (de uso general, optimizadas para cómputo, optimizadas para memoria, aceleración hardware y optimizadas para almacenamiento) y dentro de cada una de ellas en distintas categorías atendiendo al número de *cores*, a la memoria o al almacenamiento entre otras características en función de la categoría a la que pertenezcan [6]. Cabe mencionar que no todas las categorías están disponibles en todos los *data center* de los que habíamos hablado anteriormente.

Como pionera de la computación en la nube, EC2 utiliza tres opciones de compra diferentes para mejorar la administración del rendimiento: instancias reservadas, instancias bajo demanda e instancias de subasta o también denominadas puntuales o *Spot*.

* Con las instancias reservadas, los consumidores incurren en un coste fijo relativamente alto, pero tienen acceso garantizado al pedido, y las instancias no son anuladas una vez que son aprovisionadas.
* Con las instancias bajo demanda, los consumidores incurren en un alto coste fijo y no tienen acceso garantizado al pedido, pero las instancias no son anuladas una vez que son aprovisionadas.
* Con las instancias puntuales (*Spot*), los usuarios envían solicitudes de instancia puntual especificando el precio máximo por hora que están dispuestos a pagar por los recursos requeridos.

El tipo de instancias en las que se basará este TFG será con las instancias *Spot* y el proceso que se lleva a cabo frente a una solicitud será el siguiente:

* El usuario fija el precio máximo a pagar por la instancia o instancias (también llamado *bid*).
* Cada instancia puede pertenecer a una o a varias categorías cuyas características computacionales pueden establecer los requisitos de la CPU, la memoria, el almacenamiento, créditos por hora de la CPU, etc. Además, puede pertenecer a una o a varias de las regiones propuestas por la AWS, dentro de cada uno podemos hallar varias zonas de disponibilidad, pero no todos los tipos de instancia tiene por qué estar disponible en todas las regiones.
* Siempre y cuando el precio que determina el sistema para las instancias solicitadas por los usuarios sea inferior al precio especificado por el mismo, las instancias se desplegarán.
* En cuanto el precio de las instancias supere al ofrecido por el usuario, entonces las instancias se darán de baja con un preaviso de 2 minutos.

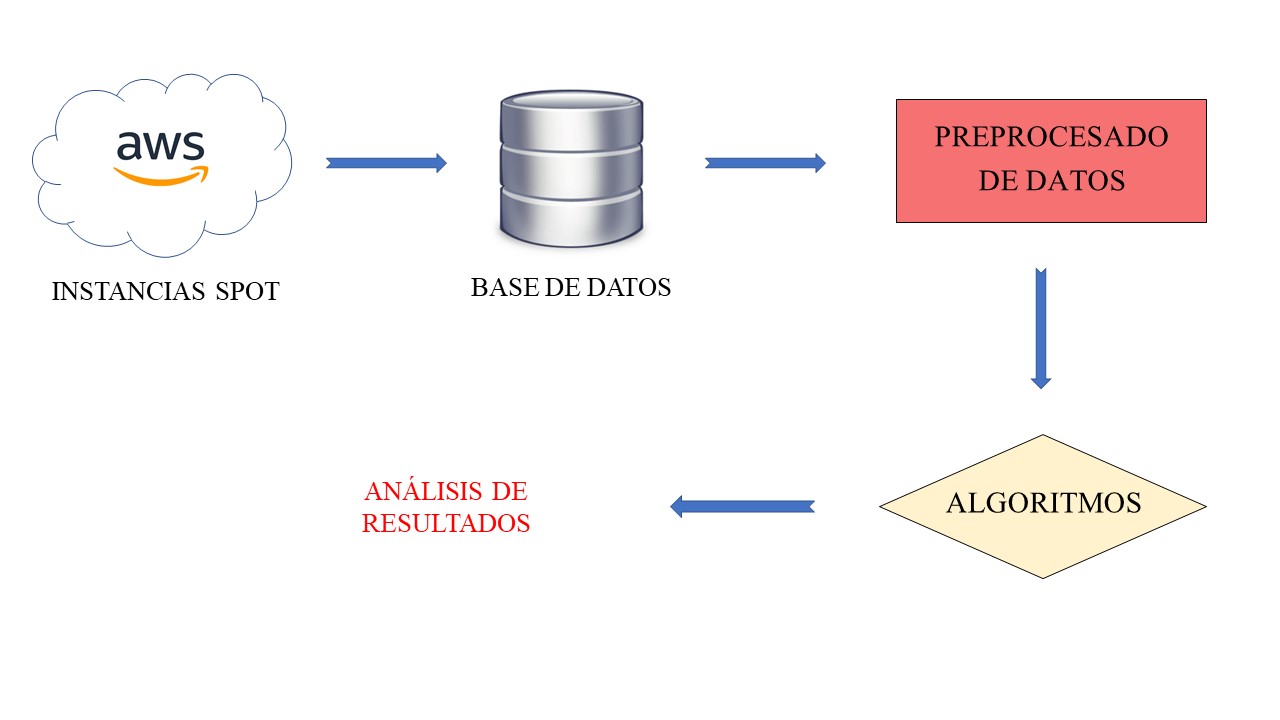
Estas instancias *Spot* ofrecen disponibilidad de capacidad de cálculo y/o almacenamiento a un precio inferior al de las instancias reservadas y al ser susceptibles de ser interrumpidas debido a la fluctuación de su precio, la aplicación que las utilice debe ser diseñada para poder funcionar con este tipo de interrupciones.

Analizar el historial de precios de *Spot* para una cantidad tan grande de mercados, entender la dinámica de precios dentro y entre ellos, y determinar el mejor tipo de instancia, zona de disponibilidad, intervalo de tiempo y precio de oferta no es sencillo. Por tanto, los postores potenciales necesitan ayuda debido a la gran cantidad de historial de precios *Spot* que está disponible para ellos. Pueden beneficiarse de un servicio que analiza e interpreta el historial de precios de *Spot* con el fin de cuantificar la volatilidad y variabilidad de cada mercado *Spot*, bien identificando patrones de precios y tendencias dentro y entre los mercados, o bien calificando el riesgo asociado con las pujas de las instancias *Spot*. Gracias a un enfoque estadístico basado en distintas métricas y algoritmos, podremos establecer una aproximación con respecto al precio de una determinada instancia puntual.

Con los datos ya recogidos en la base de datos y una vez preparados para ser ejecutados en los algoritmos, hay que especificar qué clase de métricas estadísticas emplearemos para que los resultados obtenidos después de la ejecución sean adecuados para la aproximación de nuestro objetivo. Algunos ejemplos de algoritmos que podemos utilizar son el coeficiente de Gini [7] que es muy útil ya que mide la desigualdad entre los valores de los precios [8]. También pueden ser de gran ayuda medidas como el Índice de Theil [9], métodos estadísticos de promedio móvil [10] o cualquier otro modelo de pronóstico [11].

Para una mejor comprensión del TFG propuesto se adjunta también en este documento un pequeño esquema con una visión global de todos los elementos que intervienen en el TFG (Ilustración 1):

*Ilustración 1. Resumen de la organización del trabajo*



# 2. TECNOLOGÍA ESPECÍFICA CURSADA POR EL ALUMNO.

A continuación, se mostrarán dos tablas donde se recogerá en la primera de ellas cuál es la intensificación estudiada de cara al desarrollo del trabajo (Tabla 1).

*Tabla 1. Tecnología Específica cursada por el alumno*

|  |
| --- |
| Tecnología cursada |
| Computación |

En la segunda tabla se expondrán las competencias previstas en relación con dicha especialidad que se pretenden alcanzar durante la evolución del TFG (Tabla 2).

*Tabla 2. Justificación de las competencias específicas abordadas en el TFG*

|  |  |
| --- | --- |
| Competencias | Justificación |
| C4: Capacidad para conocer los fundamentos, paradigmas y técnicas propias de los sistemas inteligentes y analizar, diseñar y construir sistemas, servicios y aplicaciones informáticas que utilicen dichas técnicas en cualquier ámbito de aplicación. | Se implementarán distintos tipos de algoritmos con el fin de conseguir los resultados requeridos en los objetivos del Trabajo Fin de Grado, así como la descarga de los datos y la preparación de los mismos antes de ser sometidos a los algoritmos. |
|  |  |
| C7: Capacidad para conocer y desarrollar técnicas de aprendizaje computacional y diseñar e implementar aplicaciones y sistemas que las utilicen, incluyendo las dedicadas a extracción automática de información y conocimiento a partir de grandes volúmenes de datos. | Para predecir el precio de nuestras instancias *Spot* aplicaremos algunas técnicas de preparación de datos, un script de descarga de estos desde la nube y aprenderemos gracias a los algoritmos con una gran cantidad de datos disponibles en la nube de Amazon. |

# 3. OBJETIVOS

El objetivo principal de este TFG consistirá esencialmente en obtener una predicción aproximada sobre el precio de una instancia *Spot* por medio de una serie de medidas estadísticas que serán implementadas y puestas en ejecución junto a los datos que obtendremos del historial de los precios. Se necesitará unas credenciales de acceso que nos permitirán acceder y utilizar los servicios de la AWS, descargar el conjunto de precios disponibles atendiendo a su categoría, familia y región a la que pertenecen, en este TFG nos centraremos en el estudio de la instancia 2xlarge.

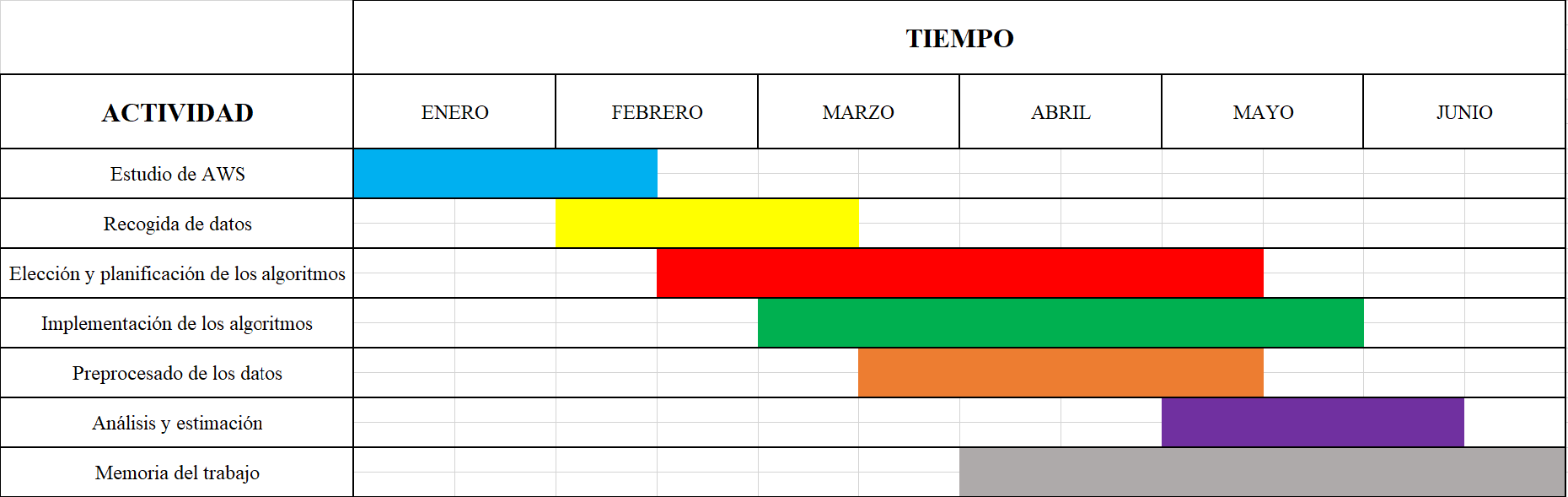
Para una mejor comprensión, el trabajo se subordinará en objetivos parciales que se detallarán a continuación:

* **Estudio de las características de la plataforma pública AWS:** antes de empezar a trabajar, analizaremos nuestro entorno de desarrollo, así conoceremos el dominio del problema. Se estudiarán previamente algunos conceptos relacionados con la computación en la nube, en concreto aquellos servicios que Amazon nos ofrece y la relación con los datos que necesitaremos durante todo el progreso.
* **Conocimiento y aplicación de medidas estadísticas:** será necesario inferir y hacer un exhaustivo estudio acerca de los algoritmos y métricas más eficientes para lograr unos resultados aceptables que nos permita predecir el patrón de coste de las instancias *Spot*.
* **Obtención, preprocesado y limpieza de los datos extraídos:** necesitaremos obtener los datos históricos por medio de la ejecución de un script implementado. Amazon nos permite obtener la información de las instancias disponibles desde los 90 últimos días anteriores desde que se ejecuta el script, una vez descargados, los registraremos en la BBDD. Los datos han de estar previamente preprocesados eliminando aquellos valores nulos que contengan para que los algoritmos sean ejecutados eficientemente y los resultados sean coherentes.
* **Desarrollo, implementación y ejecución:** una vez que se haya llevado a cabo la selección de algoritmos, estos serán implementados y ejecutados pasándoles los datos ya preprocesados y nos devolverán los resultados para la realización de un análisis estadístico predictivo.
* **Visualización de los datos:** por último, nos centraremos en hacer un análisis en vista de los resultados obtenidos que ayudarán a predecir el precio *Spot* de las instancias.

# 4. MÉTODO Y FASES DE TRABAJO

Los objetivos del TFG que se pretenden alcanzar constan de una serie de puntos que se especificarán detalladamente a continuación y que recogen resumidamente en forma de guion cada uno de los temas a tratar, que servirán de ayuda para un continuo seguimiento del trabajo. Esto es importante cumplimentarlo adecuadamente para que el funcionamiento sea el esperado y así prevenir algunos errores o fallos inesperados que pudieran llegar a aparecer. Para realizar un correcto desarrollo del trabajo se contemplará una de las metodologías ágiles más conocidas como es *SCRUM* [12] para conseguir una organización del trabajo adecuada, así como una planificación temporal del mismo (Ilustración 2):

*Ilustración 2. Cronograma*



* **Estudio de AWS:** Primero de todo se analizará y se estudiará todo el contenido del dominio del problema que nos encontramos como familiarizarnos con el entorno de los servicios que ofrece *Amazon Web Services* y en concreto con el EC2, conocer el concepto de *cloud computing* y cómo funciona el sistema de oferta y demanda de las instancias *Spot*, así como preparar todo nuestro entorno de trabajo tanto a nivel de hardware como de software.
* **Recolección de los datos:** Lo siguiente que se deberá realizar es implementar un script en Python para recoger toda la información acerca de los precios de las instancias de forma automática y conforme los obtengamos los vamos insertando en una base de datos que utilizaremos para almacenar y organizar todo el contenido necesario para posteriormente aplicar los distintos algoritmos de predicción.
* **Elección y planificación de los algoritmos:** Una vez conseguida toda la información necesaria y preparada la base de datos, es hora de investigar acerca de los algoritmos que puedan llegar a ser de utilidad para estimar el precio de una instancia atendiendo a que los resultados obtenidos sean de utilidad y coherentes a la definición del problema.
* **Implementación de los algoritmos:** Ya seleccionados los algoritmos que necesitemos, llevaremos a cabo la implementación de cada uno de ellos para que una vez que se ejecuten al pasarle los datos del precio de las instancias, nos revelen cuales han sido los resultados y que sean diseñados óptimamente desde un punto de vista computacional.
* **Preprocesado de los datos:** Es muy importante que antes de someter los datos a las ejecuciones de los algoritmos sean preprocesados de una forma eficaz evitando valores nulos que puedan entorpecer los resultados tras las ejecuciones por ejemplo y evitar problemas de sobreajuste y de sesgo. En cuanto estén los datos preparados podrán ser sometidos a la ejecución de los algoritmos correspondientes.
* **Análisis y estimación:** Por último, tras la ejecución de los algoritmos para explorar diversas propuestas de mejora en vista de los resultados esbozando unas gráficas que ayuden visualmente a analizar cuáles han sido estos resultados y si se ajustan a lo esperado pudiendo así sacar una serie de conclusiones que pondrán el broche de oro al proceso de trabajo del TFG.
* **Memoria de trabajo:** durante el proceso del trabajo se elaborará un informe que abarcará los contenidos del TFG con detalle.

# 5. MEDIOS QUE SE PRETENDEN UTILIZAR

A continuación, se especificarán tanto los medios hardware y software requeridos y recomendables para que la realización del TFG sea llevada a cabo con éxito.

# 5.1. Medios Hardware

Todo el TFG se llevará a cabo por medio de un ordenador portátil con conexión a Internet sin necesidad de ningún componente hardware especial, pero con una capacidad de almacenamiento suficiente para recoger todos los registros de las instancias *Spot* en la base de datos, ya que será una cantidad bastante voluminosa atendiendo al tipo de región o familia de la que se trate.

# 5.2. Medios Software

Las herramientas y aplicaciones software necesarias para descargar los datos desde la nube, insertarlos en la base de datos o la implementación de los algoritmos de predicción se llevará a cabo gracias a las:

* **Lenguaje de programación:** Python 3.6 [13] principalmente porque incorpora las librerías necesarias para el acceso a los datos de la EC2 de Amazon como la librería *boto* [14] u otras necesarias para esbozar gráficas entre otras.
* **Entorno de programación:** se utilizará la librería Jupyter Notebook [15] puesto que es una forma muy cómoda de trabajar y visiblemente aceptable gracias a la opción de poder añadir comentarios entre los diferentes fragmentos del código para la mejor comprensión de todo el proceso de implementación.
* **Base de datos:** Se usará una base de datos relacional que sirva para el almacenamiento de los datos, en este caso se utilizará SQLite ya que está disponible en una librería tras la instalación de los paquetes de Python.

# 6. REFERENCIAS

* [1] “Instancias *Spot* de Amazon EC2, 2018”, <https://aws.amazon.com/es/ec2/spot/>, [Acceso: Abril de 2018]
* [2] Alicia Rey, “Cloud computing: El servicio de almacenamiento en la nube”, Info-doc, Gestión de la información:

<http://www.huesca.es/_archivos/ficheros/bibliotecas_2651.pdf>

* [3] Instituto Nacional de Ciberseguridad, Cloud Computing, 2017 v.1, “Una guía de aproximación para el empresario”, Cloud Computing, 2017 v.1:

<https://www.incibe.es/sites/default/files/contenidos/guias/doc/guia-cloud-computing_0.pdf>

* [4] “Regions and Availability Zones, 2018”, <https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/using-regions-availability-zones.html>, [Acceso: Abril de 2018]
* [5] “Sistemas Operativos AWS, 2018”, <https://docs.aws.amazon.com/es_es/inspector/latest/userguide/inspector_supported_os_regions.html>, [Acceso: Abril de 2018]
* [6] “Tipos de instancias de Amazon EC2, 2018”, <https://aws.amazon.com/es/ec2/instance-types/> , [Acceso: Abril de 2018]
* [7] Jimena Sastre Álvarez, Laura P. Cogua Garzón, J.Manuel Cortés Rivera**, “**Transcripción del Coeficiente de Gini”, 28 de Mayo de 2015:

<http://www.icesi.edu.co/cienfi/images/stories/pdf/glosario/coeficiente-gini.pdf>

* [8] Autores desconocidos, “Medición de la desigualdad”, Páginas 3-32:

<http://decon.edu.uy/~mito/nip/desigualdad.pdf>

* [9] Carlos Gradín y Coral Del Río, “Desigualdad, Polarización y Pobreza en la Distribución de la renta en Galicia”, Capítulos 1 y 3, Instituto de Estudios Económicos de Galicia - Fundación P. Barrié de la Maza - nº 11, A Coruña, 2001:

<http://decon.edu.uy/~mito/nip/desigualdad.pdf>

* [10] D.R. Instituto Tecnológico y de Estudios Superiores de Monterrey, “Métodos estáticos para la estimación de ingresos”, México, 2006:

<http://www.cca.org.mx/funcionarios/biblioteca/html/finanzas_publicas/documentos/3/m3_metodos.pdf>

* [11] Dra. Fernanda Villarreal, “Introducción a los modelos de pronóstico”, Universidad Nacional del Sur- Departamento de Matemática, 2016:

<http://www.matematica.uns.edu.ar/uma2016/material/Introduccion_a_los_Modelos_de_Pronosticos.pdf>

* [12] “Qué es SCRUM”, <https://proyectosagiles.org/que-es-scrum/>, [Acceso: Abril de 2018]
* [13] “Python Official Page”, <https://www.python.org/>, [Acceso: Mayo de 2018]
* [14] “Librería Boto Python”, <https://pypi.org/project/boto/>, [Acceso: Mayo de 2018]
* [15] “Project Jupyter”, <http://jupyter.org/>, [Acceso: Mayo de 2018]