**1. Autores del trabajo, planificación y entrega**

**1.1 Autores**

En este apartado se debe indicar el número de grupo y los nombres de los autores, poniendo en primer lugar al coordinador del grupo.

**1.2 Planificación**

En este apartado se debe incluir un enlace (URL) compartido a la planificación del trabajo utilizando una herramienta online de diagramación Gantt (por ejemplo, GanttPro, versión gratuita).

En este enlace hay un ejemplo de una posible planificación del trabajo, que se puede usar como referencia, pero indicando los nombres reales de las tecnologías y de los participantes.

[Ejemplo planificación](https://app.ganttpro.com/shared/token/b0b82a0da290d4dcc93d8813795ad00093b8c583b346f796b38148ef71895eb1).

Hay que tener en cuenta que, como puede verse en el ejemplo, cada participante del grupo debe tener asignadas tareas que sumen al menos 15 horas. El peso de este trabajo en la calificación total de la asignatura es de un 10%, por tanto requiere de una dedicación de 15 horas del total de 150 horas de la asignatura.

**1.3 Entrega**

En este apartado debe incluirse un enlace (URL) a un repositorio en GitHub o en BitBucket creado para el trabajo.

En dicho repositorio debe encontrarse, al menos los siguientes archivos en la rama máster:

* Trabajo terminado: del trabajo terminado con el nombre TG1\_final.ocx
* Presentación del trabajo: TG1\_final.pptx

Dichos archivos serán los que se tendrán en cuenta para la calificación del trabajo.

**2. Descripción del tipo de tecnología**

En este apartado se debe indicar el tipo de tecnología en general y las tecnologías específicas sobre las que trata el trabajo.

El objetivo del trabajo es ponerse en la situación de una persona ya titulada en el Grado en Sistemas de Información, y que desea actualiza sus conocimientos sobre dichas tecnologías.

En este documento se debe recoger toda la información que se ha recopilado para poder aprender la teoría y práctica de dichas tecnologías, así como las ayudas que existen para poder financiar su estudio o su implementación en empresas u otras organizaciones.

**3. Fuentes de información (documentos)**

En los sub-apartados de este apartado se deben indicar documentos de interés para aprender sobre el tipo de tecnología en general, y sobre cada una de las tecnologías elegidas.

Sobre cada documento se debe

**3.1 Fuentes sobre el tipo de tecnología en general**

3.1.1 Fuente de información 1 sobre el tipo de tecnología en general

3.1.2 Fuente de información 2 sobre el tipo de tecnología en general

3.1.n Fuente de información n sobre el tipo de tecnología en general

**3.2 Fuentes sobre la tecnología específica A**

3.2.1 Fuente de información 1 sobre la tecnología específica A

3.2.2 Fuente de información 2 sobre la tecnología específica A

3.2.n Fuente de información n sobre la tecnología específica A

**3.3 Fuentes sobre la tecnología específica B**

3.3.1 Fuente de información 1 sobre la tecnología específica B

3.3.2 Fuente de información 2 sobre la tecnología específica B

3.3.n Fuente de información n sobre la tecnología específica B

**4. Fuentes de información (cursos no gratuitos)**

**4.1 Cursos no gratuitos sobre el tipo de tecnología en general**

4.1.1 Curso no gratuito 1 sobre el tipo de tecnología en general

4.1.2 Curso no gratuito 2 sobre el tipo de tecnología en general

4.1.n Curso no gratuito n sobre el tipo de tecnología en general

**4.2 Cursos no gratuitos sobre la tecnología específica A**

4.2.1 Curso no gratuito 1 sobre la tecnología específica A

4.2.2 Curso no gratuito 2 sobre la tecnología específica A

4.2.n Curso no gratuito n sobre la tecnología específica A

**4.3 Cursos no gratuitos sobre Spark**

**4.3.1 Cloudera Developer Training for Apache Spark:**



Este curso tiene una duración total de 21 horas y su precio asciende a 1360€. Va dirigido a desarrolladores y diseñadores de software con experiencia en lenguajes de programación Python y Scala.

Los contenidos de los que consta este curso son los siguientes:

Introducción y conceptos básicos de Spark, trabajos y operaciones RDD, Spark en un Cluster, programación en paralelo con Spark, trabajo con aplicaciones Spark, algoritmos comunes y mejoras de rendimiento.

Enlace: <https://www.pue.es/cursos/curso-cloudera-developer-training-for-apache-spark>

**4.3.2 Módulo de especialización: Tecnología Spark**



Curso diseñado por U-tad orientado a profesionales de ingeniería (en especial a profesionales TIC) para que se mantengan actualizados con las nuevas tendencias y herramientas del sector.

El módulo tiene una duración de 20 horas distribuidas en dos fines de semana, con un coste total de 375€.

El curso consta de los siguientes módulos:

- Introducción a Scala y lenguajes funcionales orientados a la visión general de Big Data y Spark.

- Desplegar un entorno de desarrollo Spark y desarrollar aplicaciones.

- Creación de cluster de Spark en sistemas de producción en máquinas virtuales y streaming mediante la realización de ejercicios práctico.

- Tecnologías Spark.

Enlace: <https://www.u-tad.com/estudios/modulo-de-especializacion-tecnologia-spark/>

**4.3.3 Curso Apache Spark Madrid:**



Curso presencial de 20 horas destinado a desarrolladores interesados en las tecnologías de procesamientos de datos. El curso junto a la certificación está valorado en 399€.

El temario del curso se muestra a continuación:

* Características de Spark Resilients Distributed Datasets.
* Desarrollo de aplicaciones con Spark.
* Procesamiento de datos a tiempo real con Spark Streaming.
* Spark SQL.

Enlace: <http://formacionhadoop.com/curso-apache-spark-madrid/>

**5. Fuentes de información (cursos gratuitos)**

**5.1 Cursos gratuitos sobre Big Data**

**5.1.1 Introducción a la ciencia de datos: Coursera**

Curso gratuito impartido por la universidad de Washington. No consta de un plazo fijo, puede completarse entre 8 y 10 semanas, dependiendo del número de horas semanales que dedique el alumno (80-100 horas en total).

Este curso cubre técnicas y tecnologías clave, así como el uso de las bases de datos relacionales y diseño de experimentos con visualización y modelado estadístico.

**5.1.2 Curso gratuito sobre ciencia de datos de Harvard: Data Science**

Curso impartido por la universidad de Harvard, completamente gratuito. Este curso se centra en:

- Data Wrangling

- Gestión, exploración y análisis de datos.

- Predicción.

- Visualización de resultados.

**5.1.3 Curso sobre el diseño de algoritmos en Stanford: diseño de Algoritmos**

Este curso gratuito impartido por la universidad de Stanford, cubre temas fundamentales en Big Data cómo el diseño de algoritmos y paradigmas, gráficos y estructuras de datos y diseño de algoritmos aleatorios y de probabilidad.

**5.1.4 Curso sobre Machine Learning: Aprendiendo de los datos**

Curso gratuito impartido por el instituto tecnológico de California. Se centra en el campo con mayor crecimiento del Big Data, Machine Elearning. Se trata del diseño de algoritmos autodidactas, que se adaptan automáticamente en función de los datos recibidos sin intervención humana.

**5.1.5 Visualización de datos: visualización de datos**

Curso online gratuito impartido por la universidad de Berkeley que se centra en la parte más creativa y complicada de Big Data.

Abarca las representaciones gráficas, infogramas y otros tipos de representaciones de los datos, representaciones claves para entender y asimilar los resultados obtenidos de las cantidades ingentes de datos.

**5.1.6 La caja de herramientas para el científico de datos: Universidad Johns Hopkings**

Este curso gratuito es perfecto para situarnos en el contexto de las herramientas, hipótesis y datos con los que trabajan los científicos de datos.

Introduce a los conceptos de las ideas que hay detrás de los datos para transformarlo en conocimiento aplicado e introduce a la práctica de herramientas y técnicas utilizadas.

**5.2 Cursos gratuitos sobre Hadoop**

**5.2.1 Big Data University:**



Este curso enseña los fundamentos de Hadoop y el concepto de grandes volúmenes de datos. Los materiales y herramientas software utilizadas en este curso son completamente gratis y se incluyen ejercicios prácticos.

Este curso está diseñado para proporcionar una comprensión básica de Hadoop, comenzando por una descripción de lo que es Big Data y la utilidad de Hadoop para procesar datos de una forma muy ágil.

Enlace: <https://www.bigdatauniversity.com.br/courses/BDUPT/BD001PT/2015/about>

**5.2.2 Code Jobs:**



Este curso gratuito ofrece una serie de tutoriales y documentación oficial proporcionada por los desarrolladores de Hadoop para su instalación y configuración. Además, ofrece una serie de documentación online para una primera toma de contacto de dicho software.

Enlace: <https://www.codejobs.biz/es/blog/2014/10/16/que-es-hadoop>

**5.3 Cursos gratuitos sobre Spark**

**5.3.1 Big Data University:**



Curso online gratuito puesto en marcha en junio de 2015 y que se puede iniciar en cualquier momento que desee el usuario. Además, no consta de tiempo límite para la realización de este curso.

Abarca la instalación, primer contacto y utilización completa de dicha herramienta para manejar una gran variedad de situaciones de procesamiento de datos.

Enlace: <https://www.bigdatauniversity.com.br/courses/BDUPT/BD095PT/2015/about>

**5.3.2 Página oficial de Spark:**



Spark ofrece una serie de documentación gratuita muy completa que enseña al usuario todo lo necesario para el dominio de esta herramienta (funcionalidades, herramientas, etc).

Toda la documentación proporcionada se encuentra totalmente gratis en el enlace inferior. No consta de tiempo de formación, ya que al tratarse de la documentación permanente proporcionada por Spark.

Además, Spark ofrece un canal de contacto para posibles problemas que puedan surgir y también cuentan con una serie de recursos externos completamente gratuitos (incluyendo vídeos y ejemplos) donde el usuario puede aprender más.

Enlace: <http://spark.apache.org/docs/latest/>

**6. Ayudas económicas para estudiar las tecnologías**

**- Becas Fundación EOI:**



Master online Big Data que cuenta con una beca del 25% de descuento de los horarios de matrícula en el programa, apoyado por el Ministerio de Industria, Energía y Turismo.

Está destinado al impulso de la formación y capacitación de profesionales en el sector de la economía digital y su objetivo es atraer el talento y contribuir a la formación y desarrollo de jóvenes profesionales que deseen mejorar su formación en el área de Big Data como ventaja competitiva en la búsqueda de empleo.

La beca está dirigida a titulados universitarios o de formación profesional con escasos recursos económicos y que acrediten una gran proyección profesional a juicio del comité de becas.

Enlace: <https://www.eoi.es/es/becas/2178/becas-fundacion-eoi-redes-master-online-big-data>

**- Beca del Banco Santander para el master en Big Data y Business Analytics:**

CIFF Business School, con la colaboración del Banco Santander, ofrece 2 becas destinadas a la realización de este curso. El valor de la matrícula es de 8900€ y esta beca cubre entorno al 55% de la cuantía.

La duración del programa es de 9 meses y engloba el siguiente programa académico:

1. Introducción: Big Data, Data Science Toolkit, medición para el negocio y aspectos legales y éticos del Big Data.
2. Gestión de proyectos en Big Data.
3. Herramientas de análisis: entornos de data science y programación estadística.
4. Técnicas de análisis: aprendizaje automático, aprendizaje estadístico, NLP y minería de textos, análisis de redes sociales y visualización de datos.
5. Gestión de datos: BBDD no convencionales.
6. Adquisición de datos: la web de los datos y extractores de datos.
7. Paralelización de datos: modelos y gestión de datos paralelos.

Enlace: <http://gestion.fundacioncarolina.es/programas/4243>

**- Beca CUNEF en el master de Data Science para finanzas:**

Ofrece tres ayudas económicas de 12000€ para la realización de este master, cuyo valor de matrícula es de 16000€. La duración del programa será de 10 a 11 meses y abarca los siguientes campos desde el punto de vista del Big Data: investigación, análisis de datos, herramientas de programación, ETL, técnicas de análisis estadístico y de clasificación, entorno legal y tecnológico, y análisis del mercado.

Enlace: <http://gestion.fundacioncarolina.es/programas/4246>

**7. Recursos para implementar las tecnologías**

**7.1 Recursos para implementar la tecnología A**

7.1.1 Recursos gratuitos para implementar la tecnología A

7.1.2 Recursos no gratuitos para implementar la tecnología A

**7.2 Recursos para implementar la tecnología B**

7.2.1 Recursos gratuitos para implementar la tecnología B

7.2.2 Recursos no gratuitos para implementar la tecnología B

**8. Conclusiones**