
 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Semestre: 2018-10</p>	
--	---	---

## ENUNCIADO PRIMER PROYECTO

### OBJETIVO

Este trabajo tiene como objetivo

- Usar minería de datos para generar valor a una organización.

### DESCRIPCIÓN

El gobierno nacional ha decidido contratarlos con el fin de ver el potencial de los datos publicados en la página [www.datos.gov.co](http://www.datos.gov.co). En particular desea que utilice los datos disponibles en la parte de “descubre” que está orientada a la investigación, desarrollo de aplicaciones, creación de visualizaciones e historias. El objetivo es mostrar la utilidad de usar minería de datos sobre estos datos abiertos, para apoyar a las organizaciones en el logro de sus metas.

El contrato firmado para este proyecto, especifica que ustedes deben entregar un documento con las siguientes partes:

1. **(5%)** Selección de las fuentes de datos a utilizar en el proyecto, justificada en términos del problema de negocio que va a resolver, la cantidad y la calidad de datos disponible. La selección debe incluir por lo menos uno de los archivos disponibles en la página [www.datos.gov.co](http://www.datos.gov.co) y una fuente asociada al sector (e.g., *International Civil Aviation Organization* (ICAO) si se trata de un proyecto en transporte aéreo). Dependiendo del problema a resolver podrán complementar las fuentes con otros datos disponibles en páginas relacionadas con entidades regulatorias (e.g. aeronáutica civil –[aerocivil.gov.co](http://aerocivil.gov.co)- si se trata de un proyecto en transporte aéreo), que supervisan la organización para la que están diseñando el proyecto.
2. **(10%)** Descripción y preparación de los datos usados en los modelos. En este punto deben describir las fuentes utilizadas, para lo cual debe incluir el nombre de los campos y su tipo de datos, el número de registros y la frecuencia de generación. Adicionalmente, debe describir los datos que utilizará en el proyecto y el proceso que siguió para construirlos. Esto implica dar el nombre de las variables que usará, con su tipo de dato y realizar el perfilamiento de los datos. En el perfilamiento incluya para todas las variables el número de registros nulos. Además, para cada variable categórica, incluyan los valores que puede tomar y el número de registros por categoría, y para el caso de variables continuas indique los valores mínimo, máximo y promedio.
3. **(10%)** El resultado del análisis de calidad de los datos utilizados en los modelos propuestos, acompañado de un análisis del mismo.

4. **(5%)** Descripción de cómo el requerimiento de negocio es resuelto con el o los requerimientos de minería de datos propuestos, para lo cual debe diligenciar la tabla que se presenta a continuación.

Oportunidad/problema Negocio		
Descripción del requerimiento desde el punto de vista de minería de datos		
Detalles de la actividad de minería de datos		
Tarea	Técnica	Algoritmo y parámetros utilizados

5. **(30%)** El resultado de mínimo 2 modelos analíticos realizados, que de forma conjunta permitan resolver el requerimiento de negocio. Para cada modelo debe incluir la estrategia de selección de los atributos, la justificación de la selección de la técnica, el algoritmo utilizado en el modelo y sus parámetros (los más importantes), y el análisis de los resultados obtenidos, incluyendo los estadísticos arrojados por el modelo. Adicionalmente, si el modelo propuesto es predictivo, debe incluir el manejo dado a los datos para realizar las fases de entrenamiento y prueba.
6. **(20%)** Análisis del resultado de los modelos analíticos. En este punto se debe incluir una descripción de los resultados obtenidos que permita a la organización comprenderlos haciendo énfasis en el análisis de las medidas arrojadas por los modelos utilizados y en el aporte al negocio.
7. **(10%)** Las estrategias que la organización debe plantear relacionadas con los resultados obtenidos en los modelos y una justificación de porque esa información es útil para ellos.
8. **(10%)** Sustentación del proyecto.

## ENTREGA Y EVALUACIÓN

- El proyecto se realiza en grupos de 2 estudiantes.
- El documento a entregar tiene máximo 8 páginas (sin incluir portada, tabla de contenido, ni referencias), a una columna y con letra arial tamaño 12.
- Deben entregar los modelos analíticos desarrollados y todo lo relacionado con el proyecto para poder ejecutarlos en la sustentación.
- La fecha máxima de entrega es el **martes 6 de marzo a las 11:59 p.m.**