

# Proyecto Semestral

El siguiente trabajo tiene por finalidad acercarlo a usted y a su equipo a un problema real en una empresa u organización (a continuación, se usará solamente la palabra empresa), con el objetivo de dar solución a este utilizando las herramientas de minería de datos aprendidas en el curso.

Para ello, usted deberá cumplir con diferentes hitos y entregables a lo largo del semestre. Para llevar a cabo esta tarea, usted debe conformar un equipo de trabajo de **4 o 5 alumnos/as**<sup>1</sup> y proseguir de la siguiente manera:

- **Selección de una empresa:** Deben buscar una empresa que pueda facilitarles una base de datos para trabajar resolviendo un problema desde la perspectiva de minería de datos, de lo contrario no servirá. El equipo de trabajo debe conseguir los datos y estudiar cómo generar valor a partir de dicha información. En caso de ser requerido por la empresa, existe un formato de carta donde se explica que los motivos del trabajo son sólo académicos.
- **Inscripción del Equipo:** El representante del grupo debe inscribir su grupo en un post de U-Cursos que se habilitará para tales fines, el plazo para dicha inscripción es hasta el día **Domingo 21 de Agosto a las 23:59 hrs.**

El trabajo semestral cuenta con dos informes, **un video**<sup>2</sup> y un informe final.

Los informes y el video deben ser **AUTOCONTENIDOS**; además es obligatorio adjuntar el código utilizado en las entregas que así corresponda.

---

<sup>1</sup>Por composición del curso, puede que los grupos sean de menos integrantes, dependiendo del número final de inscritos.

<sup>2</sup>Se decidió utilizar el formato de video reemplazando una presentación final para entregarles mayor flexibilidad en la entrega.

# Entrega 1

En esta entrega se espera mayor foco sobre una descripción del problema que se abordará en el semestre. Para ello será necesario que mantengan reuniones con su contraparte en la empresa para entender el problema que se abordará. Para el informe se recomienda la siguiente estructura:

1. Descripción de la empresa y datos disponibles.
2. Descripción del problema u oportunidad detectada.
3. Validación del problema basada en datos<sup>3</sup>.
4. Propuesta de mejora (descripción cualitativa de lo que pretenden hacer).

**Extensión máxima: 2 páginas sin contar portada o anexos**

**Fecha de entrega: Viernes 23 de Septiembre a las 23:59**

# Entrega 2

Para la segunda entrega se espera que reporten un avance sobre la propuesta de mejora planteada en la entrega 1, es decir, análisis que se hayan realizado, iteraciones que haya tenido el proyecto y hallazgos relevantes de su trabajo. La estructura mínima para el informe es la siguiente:

1. Breve descripción de la empresa y problema que se aborda.
2. Imputaciones y tratamiento de datos (análisis descriptivo y outliers).
3. Transformación y escalamiento de atributos.
4. Extracción de atributos.
5. Justificación del modelo que se pretende utilizar.

**Extensión máxima: 4 páginas sin contar portada y anexos<sup>4</sup>**

**Fecha de entrega: Viernes 21 de Octubre a las 23:59**

# Entrega 3 (vídeo)

El objetivo es presentar el proyecto desarrollado, por lo que se incluirán los puntos de las entregas anteriores. La idea central de esta presentación es que todos puedan ver cuáles fueron los principales problemas y desafíos del proyecto, y cómo se resolvieron. Como el punto (3) es evaluado en la presentación anterior, no es necesario una explicación extensa del proceso.

La estructura recomendada para la presentación es la siguiente:

1. Breve descripción de la empresa (foco en el área del proyecto) y problema que se aborda.
2. Descripción de la base de datos que les fue entregada.
3. Acciones ejecutadas con la base de datos<sup>5</sup>.
4. Evaluación y validación de los modelos desarrollados.
5. Comparación de los resultados obtenidos con la situación actual de la empresa<sup>6</sup>
6. Evaluación económica de los modelos desarrollados en base a los resultados obtenidos<sup>7</sup>.

**Extensión máxima 7 minutos**

**Fecha de entrega: Viernes 26 de Noviembre a las 23:59**

---

<sup>3</sup> De forma sencilla: ratios, indicadores o visualizaciones.

<sup>4</sup> Deben adjuntar en un archivo aparte el código utilizado.

<sup>5</sup> Imputaciones, tratamiento de outliers, escalamiento y selección de atributos.

<sup>6</sup> Es decir, ¿Su solución resuelve el problema de la empresa o es mejor que la solución actual?

<sup>7</sup> No se busca una evaluación exhaustiva de los modelos, sino una estimación económica del impacto.

# Informe Final

Se espera que los alumnos entreguen un informe final en donde se detalle lo realizado, justificando las decisiones tomadas en cada una de las etapas, así como también las conclusiones alcanzadas por el grupo con respecto al trabajo realizado y cómo éste agrega valor para la empresa. La estructura es libre, pero debe contener las siguientes formalidades:

1. Resumen Ejecutivo/Abstract
2. Introducción
3. Desarrollo
4. Conclusiones

El desarrollo/cuerpo del informe debe contener:

1. Descripción de la institución con la que se trabajó.
2. Descripción de la base de datos que les fue entregada<sup>8</sup>
3. Acciones ejecutadas con la base de datos<sup>9</sup>
4. Evaluación y validación de los modelos desarrollados<sup>10</sup>
5. Relevancia de su trabajo: en base a los resultados, señalar las implicancias para la empresa.

Se recomienda escribir el informe en Latex usando el template ACL de dos columnas: <https://www.overleaf.com/latex/templates/acl-2020-proceedings-template/zsrkcwjtpcd>

**Extensión máxima: 6 páginas sin contar portada o anexos<sup>11</sup>.**

**Fecha de entrega: Viernes 26 de Noviembre a las 23:59**

## Cálculo de nota

La nota grupal del proyecto se calculará de la siguiente manera:

$$Nota\ Grupal\ (NG) = 0.1 * Entrega_1 + 0.1 * Entrega_2 + 0.2 * Entrega_3 + 0.6 * Informe\ Final$$

Además, se considera una coevaluación grupal. Para dicho proceso se habilitará un documento en el momento adecuado, a continuación se detalla el funcionamiento de la coevaluación:

- Cada estudiante debe asignar 100 puntos entre las y los integrantes del grupo (incluido quién evalúa). Una mayor asignación de puntos se relaciona con un mayor aporte al proyecto.
- Al final del proceso de coevaluación, cada estudiante tendrá un puntaje total de aporte al desarrollo del trabajo(CT), correspondiente a la suma de los puntajes que les fueron asignados por el resto del grupo.
- La nota del proyecto se les calculará de la siguiente manera:

$$Nota\ Personal\ (NP) = 1 + CT * ((NG - 1)/100)$$

Finalmente, la nota del proyecto viene dada por:

$$Nota\ Final\ Proyecto\ (NFP) = 0.25 * NG + 0.75 * NP$$

Un ejemplo de autoevaluación es el que se indica:

---

<sup>8</sup>Atributos, número de registros, temporalidad de los datos, etc.

<sup>9</sup>Imputaciones, tratamiento de outliers, escalamiento y selección de atributos.

<sup>10</sup>Especificación de los modelos y métricas de desempeño.

<sup>11</sup>Deben adjuntar en un archivo aparte el código utilizado.

| <i>Integrante</i> | <i>CT</i> | <i>NP</i> |
|-------------------|-----------|-----------|
| 1                 | 100       | 5.5       |
| 2                 | 125       | 6.9       |
| 3                 | 55        | 3.1       |
| 4                 | 120       | 6.6       |

Se tiene un grupo de 4 integrantes, con Nota Grupal (*NG*) igual a 5.5..La tabla muestra el puntaje total recibido por estudiante(*CT*) y la nota personal (*NP*). Se observa que la suma de los *CT* es igual a 400. Esto debe ser así, pues cada evaluador/a debe asignar en total 100 puntos entre los 4 integrantes del grupo. Finalmente, si el grupo decide que cada integrante mantendrá la nota grupal, cada evaluador debe asignar la misma cantidad de puntos a cada integrante del grupo. Es decir, cada evaluador deberá asignar 25 puntos a cada integrante del grupo.

### **Consideraciones**

Recuerde que la autoevaluación grupal mide el grado de contribución personal al trabajo, y no el resultado de éste. Un integrante no podrá asignarse a sí mismo más del 50% de los puntos. Por ejemplo, si el grupo tiene 4 integrantes hay 100 puntos para repartir, el o la estudiante no podrá asignarse más de 50 puntos.

Situaciones excepcionales serán analizadas con el cuerpo docente. Finalmente, es importante indicar que la autoevaluación de cada estudiante es tratada de forma confidencial, y serán revisadas sólo por el profesor y profesor auxiliar del curso. De esta manera, la información final que poseerá cada estudiante corresponderá a la nota final obtenida en esta evaluación, que estará relacionado al total de puntaje recibido.

### **Formulario de autoevaluación**

Para realizar la autoevaluación se dispondrá de un formulario especial que deberá completar cada estudiante. Este formulario estará disponible en la sección "Tareas" de u-cursos en el módulo de entrega de la autoevaluación. Recuerden que el total de puntos que deben distribuir es 100, por lo que no deben superar esta cantidad.