

Tecnológico de Monterrey Santa Fe

Del mundo físico al mundo digital con ciencia de datos (Semana i 2019)
Profesor: Saul Juárez Ordoñez



Lineamientos de evaluación

Los criterios de evaluación para esta semana son los siguientes:

Criterios de evaluación

Actividades	Porcentaje
Jupyter Notebook del caso resuelto	60%
Presentación oral final	35%
Auto y co-evaluación	5%

Para finalizar la semana i satisfactoriamente, deben cumplir con lo siguiente:

Jupyter notebook del caso de estudio: Fighting Fraud.

- Un miembro del equipo subirá el *jupyter notebook* del caso resuelto a nuestra clase [Semana i 2019](#) en classroom.google.com a más tardar el viernes a las 10 am.
 - Las explicaciones en el *notebook* deben ser claras y la información debe de apreciarse en sus gráficas y tablas. Se anexa la rúbrica correspondiente.

Presentación oral final

Es muy importante que puedan leer las gráficas y tablas que generaron en su notebook adecuadamente y puedan comunicar efectivamente los resultados del análisis exploratorio de datos y del estudio de *machine learning*. La presentación debe cumplir con los cuatro criterios de calidad: clara, atractiva, en el tiempo adecuado y con el uso correcto de los apoyos visuales. Los miembros del equipo deben estar organizados y no interrumpirse entre ellos.

Cada equipo presentará oralmente un resumen del caso de estudio. Las presentaciones empiezan a las 10 am con una duración máxima de 20 minutos. El orden de presentación es aleatorio. La presentación debe abordar los siguientes puntos:

- Para entrar en contexto: ¿Cómo funciona la detección de fraude crediticio y cómo ha evolucionado a través del tiempo? ¿por qué es relevante?
- ¿Por qué el problema de detección de fraude crediticio se puede ver como un problema de clasificación de *machine learning*?
- Interpretación adecuada de las gráficas y tablas que generaron en su *notebook*.
- ¿Cuál es el problema de lidiar con una base de datos desbalanceada? ¿Por qué es necesario un submuestreo (*undersampling*)? ¿Hay otras técnicas de balanceo?
- Interpretación de la comparación numérica y visual de los algoritmos de clasificación. ¿Qué algoritmo es más eficiente para la detección de fraude crediticio?

Se harán preguntas individuales, así que todos deben estar preparados.

Auto y co-evaluación

Evaluarás de manera honesta y objetiva tu desempeño y el de cada integrante de tu equipo en la realización de este reto. El instrumento de evaluación es un cuestionario *google form* que estará disponible el viernes a las 10 am.