

Problem Set 2: Making Money with ML?

Curso: BIG DATA Y MACHINE LEARNING PARA ECONOMÍA APLICADA-UNIFICADO (MECA4107_1-2)

Criterios	Insuficiente	Regular	Suficiente	Sobresaliente	Puntuación del criterio
introducción (5%)	1.25 puntos No plantea la introducción	2.5 puntos Presenta la introducción, aunque los antecedentes no son los adecuados. No contiene un resumen claro de los resultados y las principales conclusiones.	3.75 puntos Se plantea la introducción, aunque existen algunas imprecisiones que no afectan la comprensión del problema. Contiene un resumen de los resultados y las principales conclusiones.	5 puntos La introducción enuncia con claridad y precisión el problema, además explica los antecedentes, datos y pertinencia para responder a dicho problema. Contiene un resumen de los resultados y las principales conclusiones.	/5

Criterios	Insuficiente	Regular	Suficiente	Sobresaliente	Puntuación del criterio
Descripción de los Datos (30%)	7.5 puntos No describe los datos	15 puntos Realizan una descripción de la base de datos, aunque no es muy claro el propósito de esta y la utilidad que puede representar para abordar el problema planteado en el taller. No es claro cómo se obtiene la muestra final de los datos. El análisis descriptivo no se encuentra articulado consistentemente con las tablas y/o gráficos presentados. No es claro el uso de conocimiento profesional que agregue valor a la sección.	22.5 puntos Existe una descripción de la base de datos, es clara, aunque falta mayor precisión de su utilidad para abordar el problema planteado en el taller. Hay imprecisiones en la descripción de cómo se obtiene la muestra final. Aunque realiza el análisis descriptivo de los datos, algunas de las tablas o gráficos presentan inconsistencias o no permiten comprender plenamente la variación de los datos y la elección de las variables.	30 puntos Realizan una descripción de la base de datos y su utilidad en la solución del problema planteado en el taller. Se describe cómo se obtiene la muestra final de los datos, y se realiza el análisis descriptivo de los datos, donde se presentan tablas y/o gráficos que ayudan al lector a entender la variación de los datos y la elección de las variables. Se presentan al menos 2 mapas. Se usa conocimiento profesional para agregar valor a la sección.	/30

Criterios	Insuficiente	Regular	Suficiente	Sobresaliente	Puntuación del criterio
Modelo y resultados (40%)	10 puntos No se presenta la sección	20 puntos La descripción del mejor modelo entregado en Kaggle, no es precisa, no se incluyen los detalles suficientes de cómo fue estimado o de los hiper-parámetros relevantes. No se hace la comparación con otros modelos entregados en Kaggle.	30 puntos Se describe el mejor modelo entregado en Kaggle, aunque existen imprecisiones sobre cómo fue estimado, o en los hiper-parámetros relevantes. No se comparó con al menos otras 4 especificaciones entregadas en Kaggle.	40 puntos Se describe el mejor modelo entregado en Kaggle. Se incluyen los detalles de la estimación del modelo, incluyendo la selección de los hiper-parámetros relevantes. Se compara con al menos otras 4 especificaciones entregadas en Kaggle.	/40
Conclusión (5%)	1.25 puntos No presenta la conclusión.	2.5 puntos Presenta la conclusión, pero no se resumen los resultados adecuadamente, ni las recomendaciones que surgen sobre la utilidad de los diferentes modelos predictivos	3.75 puntos El resumen del trabajo no es claro, existen imprecisiones sobre la utilidad de los diferentes modelos predictivos.	5 puntos Se resumen claramente los resultados del trabajo y las recomendaciones que surgen sobre la utilidad de los diferentes modelos predictivos	/5

Criterios	Insuficiente	Regular	Suficiente	Sobresaliente	Puntuación del criterio
“Entrega” en Kaggle (10%)	2.5 puntos No se hizo entrega en Kaggle.	5 puntos Hay menos de 2 entregas en Kaggle.	7.5 puntos Hay mas de 2 y menos de 4 entregas en Kaggle.	10 puntos Hay al menos 4 entregas en Kaggle.	/10

Criterios	Insuficiente	Regular	Suficiente	Sobresaliente	Puntuación del criterio
Documento y Repositorio de GitHub (10%)	2.5 puntos No contiene un repositorio	5 puntos El documento es autocontenido, pero excede las 8 páginas o las 8 tablas o figuras. Existe el repositorio y contiene un README. El repositorio no sigue el template, o el README no permite navegar el repositorio o reproducir el trabajo. No hay contribuciones significativas por parte de los miembros. El código no es fácilmente legible o reproducible.	7.5 puntos El documento es no es completamente autocontenido, no excede las 8 páginas y las 8 tablas o figuras. Existe el repositorio siguiendo el template provisto. El README tiene algunas imprecisiones que no afectan la comprensión del lector para navegar y replicar el trabajo. La rama del repositorio principal muestra más de (2) y menos de cinco (5) contribuciones que aportan significativamente al taller, por parte de cada miembro del equipo. El código en el repositorio es reproducible, pero su	10 puntos El documento es autocontenido, no excede las 8 páginas y las 8 tablas o figuras. Existe el repositorio siguiendo el template provisto y contiene un README que ayuda al lector a navegar el repositorio e incluye instrucciones breves para replicar completamente el trabajo. La rama del repositorio principal muestra al menos cinco (5) contribuciones que aportan significativamente al taller, por parte de cada miembro del equipo. El código en el repositorio es legible y reproducible	/10

Criterios	Insuficiente	Regular	Suficiente	Sobresaliente	Puntuación del criterio
			legibilidad no es buena.		

Total	/100
-------	------

Puntuación general

Nivel 1

0 puntos mínimos

Nivel 2

5 puntos mínimos

Nivel 3

8 puntos mínimos

Nivel 4

11 puntos mínimos