

Taller 2: Profundizando el análisis de *CheMarket Inc.*

Contexto

Después de su primer informe, la mesa directiva de *CheMarket Inc.* quedó interesada en seguir profundizando el análisis. En particular, quiere saber si los resultados obtenidos hasta ahora se mantienen cuando se usan métodos más flexibles que permitan descubrir relaciones no lineales y efectos heterogéneos.

Su encargo en este Taller 2 es elaborar un **informe de extensión del Taller 1**, donde comparen directamente los hallazgos anteriores con nuevos resultados basados en modelos de árboles.

Entrega final

- Un **informe escrito** de máximo **6 páginas** (sin contar anexos). Deben contrastar los nuevos resultados con los hallazgos previos y discutir similitudes y diferencias.
- Dos sets de **slides para presentación**:
 - Slides Parte A (máx. 4)
 - Slides Parte B (máx. 4)
- La presentación debe durar 15 minutos, enfatizando la comparación con el Taller 1 y las implicaciones estratégicas.

Formato de entrega:

- `informe_equipo_XX.pdf` (reemplazar XX por el número del equipo con dos dígitos)
- `slides_parteA_equipo_XX.pdf`
- `slides_parteB_equipo_XX.pdf`

Parte A: Árboles de regresión y predicción

La dirección quiere saber si un modelo basados en arboles de regresión puede mejorar la predicción del gasto en comparación con los modelos lineales usados en el Taller 1.

Aspectos a incluir en el informe:

- Construyan un árbol de regresión para predecir **Revenue**.

- Describan la estructura del árbol: ¿qué variables aparecen como nodos de decisión?, ¿qué reglas de segmentación emergen?
- Comparen el desempeño predictivo (ej. MSE en el conjunto de prueba) con los resultados obtenidos anteriormente.
- Discutan qué nuevas intuiciones ofrece el árbol frente a los modelos lineales, y qué limitaciones tiene.

Preguntas que podrían anticipar o responder en el informe:

- ¿Qué variables aparecen como más relevantes en el árbol de regresión?
- ¿El árbol confirma o contradice las relaciones lineales encontradas en el Taller 1?
- ¿Qué tan distinto es el poder predictivo (MSE) del árbol frente a los modelos lineales?
- ¿El árbol ofrece intuiciones adicionales sobre subgrupos de usuarios con distinto nivel de gasto?
- ¿Vale la pena preferir este método para la predicción en *CheMarket Inc.*?

Parte B: Árboles causales y heterogeneidad de efectos

El interés ahora es entender no solo el efecto promedio de facilitar el registro, sino si este efecto varía entre diferentes tipos de usuarios.

Aspectos a incluir en el informe:

- Apliquen un árbol causal para estimar los efectos heterogéneos del tratamiento `sign_up`.
- Identifiquen subgrupos en los que el efecto sea más alto o más bajo.
- Comparen estos hallazgos con los resultados promedio obtenidos en el Taller 1: ¿el árbol confirma o matiza lo que se había concluido antes?
- Discutan la plausibilidad de los patrones encontrados y qué implicaciones tienen para la estrategia de la empresa.

Preguntas que podrían anticipar o responder en el informe:

- ¿El árbol causal confirma el efecto promedio del tratamiento estimado en el Taller 1?
- ¿Existen grupos de usuarios en los que el efecto de facilitar el registro sea más alto o más bajo?

- ¿Qué características de los usuarios marcan la diferencia en el efecto del tratamiento?
- ¿Son plausibles estos patrones desde el punto de vista del negocio?
- ¿Qué tan robustos parecen los hallazgos? ¿Se deben tomar como exploratorios o como evidencia sólida?
- ¿Qué recomendaciones estratégicas se derivan para la mesa directiva? ¿Sugeriría aplicar la intervención a todos o focalizarla en ciertos segmentos?