

# Devolución general TP1 Enfoque Estadístico del Aprendizaje 2023

## Comentarios sobre la corrección

- En muchos casos la resta de puntos se debió a que **no contestaron lo que se pedía en la consigna**. En la mayoría de estos casos fue por falta de respuesta a algo que se pedía explícitamente.
- En otros casos avanzaron en más de lo que se pide en las consignas, a veces cometiendo errores conceptuales. Sugerimos no hacerlo porque muchas veces se pierde el foco y se olvidan de contestar cosas que si se piden.
- Se hizo especial hincapié en aquellos temas desarrollados en extensión en el curso: interpretación de coeficientes, análisis de significatividad y evaluación y diagnóstico del modelo.
- Las **justificaciones o fundamentaciones** de decisiones o elecciones en algunos casos fueron muy buenas y en otros fueron escasas y/o confusas.
- Sugerimos que revisen antes de entregar las salidas de los notebooks en versión html. Queda desprolijo en algunos casos que al compilar no se ven bien los gráficos o lo que escribieron.
- Por fuera de los temas propios de la materia, se consideró un **error muy severo**: eliminar observaciones, reentrenar con el dataset de testing y/o incurrir en data leakage.

## Aspectos positivos

- Buenas interpretaciones de los coeficientes: de manera detallada acorde a lo visto en clase
- Fundamentación clara de las decisiones de selección de modelos.
- Creatividad y claridad en el diseño de nuevos modelos. Valoramos mucho el uso de gráficos o estadísticos descriptivos para mostrar las variables que pensaron incluir en los modelos.
- Ahondar en la comparación de las métricas de performance en el punto 6, listando posibles motivos de los resultados obtenidos.

## Aspectos negativos

### Análisis exploratorios

- **Eliminar registros** de la base. En las consignas no se pedía filtrado del dataset, y frente algunas preguntas en clase se había explicitado que no debían eliminar registros (error muy severo)
- **Mencionar sólo el número de la correlación** entre dos variables y no interpretarla: se esperaba interpretación de cómo es esa correlación, fuerte o moderada, positiva o negativa, etc.
- No realizar la apertura por género fue considerado un error moderado
- Incluir y analizar la correlación de variables de tipo identificador: aglomerado y código actividad

### Interpretación de coeficientes

Se requería interpretar en términos de **valor esperado**. Interpretar el **intercepto**, variables **continuas y categórica** (identificación de dummies y categorías basales), en términos de **modelo múltiple** (dadas las demás variables o manteniéndose constantes las demás variables).

#### *Fallas de interpretación:*

- No interpretación del intercepto.
- No interpretar en términos del valor esperado (en cualquiera de las formas vistas en la materia).
- No interpretar en términos del modelo de regresión múltiple
- Interpretación de manera independiente de los coeficientes de la experiencia potencial y el cuadrado de la experiencia potencial. Por ejemplo: “el  $\beta$  indica la variación esperada del salario horario ante un año más de experiencia potencial, dada la experiencia potencial al cuadrado constante”
- No interpretación de la categórica en relación a la categoría basal.
- Interpretación errónea de la interacción o no interpretar educación en relación a la basal
- Interpretación errónea de la semielasticidad
- Comparar la importancia de regresores en base a la magnitud de los estimadores con variables no estandarizadas.

La presencia de alguna de estas fallas fue considerado un error moderado o severo según el caso, la presencia de varias fallas fue considerado un error severo o muy severo según el caso

#### **Significatividad de coeficientes**

- No interpretar lo que implica que una variable sea estadísticamente significativa (significatividad individual, y global del modelo).
- No interpretar la significatividad individual de la variable dummy (como diferencia estadísticamente significativa de la categoría basal).
- Realizar análisis contradictorios en base a la información de los p-valores e intervalos de confianza. En algunos trabajos concluyen que rechazaban la hipótesis nula del test de significatividad individual y luego marcan que el intervalo de confianza contiene el cero
- Afirmar que una variable es “no significativa” solo por la magnitud del estimador.

#### **Evaluación de modelos**

- No transformar las predicciones de los modelos que usaban el logaritmo del salario horario para la evaluación y comparación de performance fue considerado un error severo
- En el punto 4 varias personas compararon los modelos usando el R-cuadrado simple de modelos con distinta cantidad de variables (a pesar de que no se solicitaba por este motivo). Se consideró un error moderado/severo
- En varios TPs se menciona que los modelos explican poca variabilidad. No es un error pero hay que recordar que es una afirmación que se puede hacer teniendo un conocimiento del dominio mucho más amplio.
- No interpretar los resultados obtenidos en RMSE y MAE. En varios casos hicieron una descripción literal de la tabla de métricas sin ahondar en posibles motivos de los resultados obtenidos.
- No interpretar los resultados del modelo robusto en base a que lo diferencia del resto de los modelos. Muchos dieron una explicación pobre, sin profundizar en el tratamiento del modelo robusto de los outliers y cómo esto influencia en los datasets y las métricas.

## Creación de modelos

- Incurrir en data leakage: inclusión del salario horario de manera directa fue considerado un error muy severo
- Se esperaba que justifiquen y expliquen sus decisiones: la inclusión de variables sin explicación o análisis exploratorios implicó la quita de puntos
- Reentrenar los modelos con el dataset de test fue considerado un error muy severo

## Diagnóstico

- En algunos TPs mencionaron que encontraron heterocedasticidad y otros homocedasticidad. Desde el curso entendemos que en el gráfico del **primer modelo** se observa **heterocedasticidad** de manera más clara, mientras que en el segundo modelo se observa **homocedasticidad**. Dado que es una herramienta gráfica se consideraron errores dependiendo de la justificación que brindaban
- En varios TPs señalan las observaciones identificadas con número en el plot como observaciones de alto leverage. Esto es incorrecto y se lo consideró un error moderado-severo según el caso.
- Varias personas no explicaron el criterio que utilizaron para interpretar el gráfico de leverage. También en varios casos interpretaron que había observaciones con alto leverage. Estos aspectos fueron considerados como errores leves-moderados según el caso.