



Evaluación Final del Módulo

Caso Estudio Precios Propiedades

A continuación, se define el instrumento de evaluación final del módulo:

Componentes de la evaluación	
Evidencias	Las evidencias de la evaluación serán <ul style="list-style-type: none"> • De producto, consistente en: <ul style="list-style-type: none"> ○ Código Python con algoritmo • Conocimiento fundamental, que consiste en: <ul style="list-style-type: none"> ○ Hoja de respuesta
Situación Evaluativa	Se evaluará el desempeño mediante: <ul style="list-style-type: none"> • Un encargo
Tipo de trabajo	Individual
Tiempo	24 horas
Lugar	Trabajo remoto

Enunciado del Problema:

Se está haciendo un estudio de propiedades en la octava región del país y a Usted le han encargado analizar un set de datos provisto por el ministerio de vivienda y urbanismo. El data set, tiene información respecto a los precios de las propiedades y la caracterización de los siguientes atributos:

- CRIM: Tasa de criminalidad per cápita del sector
- ZN: Proporción de suelo residencial para lotes sobre 25,000 sq ft
- INDUS: Proporción de negocios no-retail por sector
- CHAS: 1 si el tramo limita con el Río Biobío, caso contrario, 0
- NOX: Concentración óxido nítrico (part per 10 million)
- RM: Promedio número de dormitorios por vivienda
- AGE: Proporción de viviendas ocupadas por sus propietarios construidas previo a 1940
- DIS: Distancias ponderadas a cinco centros de empleo de Boston
- RAD: Índice de accesibilidad a carreteras radiales
- TAX: Tasa de impuesto a la propiedad de valor total por \$ 10,000
- PTRATIO: Proporción alumno-profesor por ciudad
- B: $1000 (B_k - 0,63)^2$, donde B_k es la proporción de (personas de otras etnias) por sector
- LSTAT: Porcentaje de menor estatus de la población

Se le solicita, que en su análisis de datos, incorpore como mínimo lo siguiente:

1. Realice un modelamiento regresivo multivariable con selección de variables hacia adelante (Forward Stepwise Regression) ¿Qué nivel de ajuste tiene el modelo? ¿Cuáles son las variables que más inciden en el precio de una propiedad?.
2. ¿Cambiano el método de selección de variables fue posible superar el mejor R^2 ajustado de 0.735? (valor reportado en el notebook guía calculado con el método de eliminación hacia atrás, Backward Stepwise Regression). ¿Se encontró un modelo diferente o llegó al mismo modelo?, explique posibles causas.

Recuerde:

Forward Stepwise Regression. Las variables se introducen secuencialmente en el modelo. La primera variable que se introduce es la de mayor correlación (+ o -) con la variable dependiente. Dicha variable se introducirá en la ecuación solo si cumple el criterio de entrada. A continuación se considerará la variable independiente cuya correlación parcial sea la mayor y que no esté en la ecuación. El procedimiento termina cuando ya no quedan variables que cumplan el criterio de entrada. Seleccionar el modelo con el mejor R^2 ajustado y comparar.

Entregable:

El entregable es un notebook, ordenado, documentado y reproducible que contiene su análisis y conclusiones, utilizando las buenas prácticas aprendidas a lo largo del módulo.

Si bien ChatGPT es una tecnología impresionante, también tiene sus limitaciones. Puede haber ocasiones en las que la respuesta generada no sea completamente precisa o no se ajuste al contexto. Por eso, es importante usarlo como una guía y complementar la información con otras fuentes confiables.

Criterios de evaluación:

Aspectos a evaluar	Descripción
Entrada y Salida de Datos	Utiliza técnicas de extracción y almacenamiento de datos adecuadas al problema y de forma eficiente
Exploración inicial	Se realiza un análisis exploratorio inicial para conocer la consistencia de la información
Valores nulos	identificación, análisis y tratamiento de valores nulos
Utilización de conceptos de estadística descriptiva	Utiliza adecuadamente los conceptos generales de estadística descriptiva, tales como medidas de tendencia central, dispersión y de posición
Análisis <u>univariado</u>	Realiza un análisis a la variable principal para caracterizarla, utilizando estadística descriptiva y análisis visual
Análisis de Correlación	Realiza análisis de correlación en donde explica aquellas variables con una mayor correlación
Modelo de regresión	Realiza un modelo regresivo múltiple aplicando selección de variables y distinguiendo aquellas con mayor preponderancia en la explicación de la variable objetivo
Análisis visual	Denota dominio en la utilización de herramientas de análisis visual
Verificación de supuestos de regresión	Realiza verificación de supuestos de la regresión, tal como análisis de residuales
Conclusiones y hallazgos	Se presentan conclusiones claras y significativas basadas en los hallazgos del análisis exploratorio.
Documentación y presentación	El análisis exploratorio está debidamente documentado y presenta una estructura clara.