



UNIVERSIDAD NACIONAL
DE MOQUEGUA

ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS E INFORMATICA

SEMESTRE 2021 - II

INTELIGENCIA ARTIFICIAL
PROYECTO UNIDAD II

VIII
Ciclo



2021
Ilo

ESTUDIANTE
Jhonny Enrique Andy
Vilca Condori

DOCENTE
MORALES GONZALES, RUSO ALEXANDER

IMPLEMENTACION DE UN PROGRAMA PARA LA PREDICCION DE PRECIOS DE LAS CASAS PARA LA EMPRESA GOLOD GROUP - 2021

Una breve descripción

Una predicción precisa del precio de la vivienda es importante para los posibles propietarios, promotores, inversores, tasadores, asesores fiscales y otros participantes del mercado inmobiliario, como, prestamistas hipotecarios y aseguradoras. La predicción tradicional del precio de la vivienda es basada en la comparación de costos y precios de venta sin un estándar aceptado y una certificación proceso. Por lo tanto, la disponibilidad de un modelo de predicción del precio de la vivienda ayuda a completar una importante brecha de información y mejorar la eficiencia del mercado inmobiliario.

A medida que prospera el mercado de las viviendas, el precio de la vivienda se convierte en un factor crucial para los buscadores de casa.

En general, las investigaciones han mejorado el poder predictivo de los modelos al enfatizar los atributos del valor de la propiedad, como el lugar de la vivienda, la calidad de la vivienda, la ubicación geográfica y el ambiente. Mientras que a su vez otros estudios más recientes se centran en la Ubicación, costos, factores internos, entre otros.

Con este trabajo académico se busca la implementación de un software o aplicación para la empresa GOLOD GROUP que, basándose en ciertos factores o características de las casas, se puede predecir el precio de las viviendas o inmuebles que se podrán a disposición o a la venta del público.

Problemática:

El principal problema por el cual se enfrentan diversos propietarios o inversionistas inmobiliarios como GOLOD GROUP es poder dar un precio real dada las condiciones del mercado, muchas veces por falta de información, así, como no estar al tanto de las tendencias o evolución del mercado. Sin esta información o guía, muchos inversionistas o personas en este rubro inmobiliario no pueden responder de manera adecuada a la aparición de nuevas oportunidades tanto de compra como venta, lo cual a largo plazo generara pérdidas significativas.

Dada la problemática , la realidad a la que se enfrentan los inversionistas es buscar soluciones ante esto, como por ejemplo realizar una plantilla con las variables que intervienen o llamar a los dueños de los inmuebles aledaños y hacer cálculos en base a esta información o la otra opción que suelen acudir es consultar a un asesor personalizado para que evalúan las viviendas y les den el precio , sin embargo estas opciones requieren de dos factores críticos como el tiempo y el dinero , dado que un asesor cobra comisiones por este tipo de trabajos.

Por este motivo se presenta este proyecto para ahorrar tiempo y esfuerzo que normalmente se emplea a la hora de buscar los precios óptimos para invertir o vender propiedades.

Objetivos

Objetivo Principal

- Crear un programa que permita la predicción de los precios de las viviendas mediante la utilización de un modelo de regresión lineal múltiple utilizando información referida al valor de las propiedades o inmuebles.

Objetivos Específicos:

- Generar un modelo de predicción (regresión lineal múltiple) que permita la predicción de las viviendas basándose en su perfil o características.
- Consolidar información que se relaciona con el precio de las viviendas.
- Realizar un análisis de los principales factores que intervienen en el valor de una vivienda.

Resultados Esperados

- El modelo fue implementado y se demostró que está en la capacidad de predecir el precio aproximado al cual debe venderse una vivienda, utilizando el modelo de regresión múltiple con un coeficiente de R2 de **0.66**, el cual es aceptable para poder tener una guía de que precio pagar o cobrar por una vivienda dados los factores expuestos.

	name	coef
1	intercept	6.48591e+06
2	bedrooms	-38793
3	bathrooms	45925.2
4	sqft_living	170.511
5	sqft_lot	-0.257118
6	floors	23823.3
7	waterfront=0	-286904
8	waterfront=1	286904
9	view	45335
10	condition	18866.8
11	grade	124262
12	yr_built	-3573.35
13	yr_renovated	8.58039

Modelo de regression multiple:

$$\begin{aligned} \text{PrecioCasa} = & 45925.2 * \text{bathrooms} - (\text{bedrooms} * 38793) + (\text{sqft_living} * 170.511) \\ & - (\text{sqft_lot} * 0.257118) + (\text{floors} * 23823.3) - (\text{waterfront} * 286904) \\ & + (\text{view} * 45335) + (\text{condition} * 18866.8) + (\text{grade} * 124262) \\ & - (\text{yr_built} * 3573.35) + (\text{yr_renovated} * 8.58039) \end{aligned}$$

- Además, el modelo de regresión múltiple fue generado con éxito basándonos en valores o propiedades de las casas los cuales contribuían en gran medida para la efectividad y precisión del modelo.
- Por otro lado, se buscó la información relacionada al valor de las propiedades, dando a conocer así las características como el número de dormitorios, numero de baños, los pies cuadrados del espacio habitable al interior de las viviendas, la medida de pies cuadrados del espacio del terreno, el número de pisos , la vista al mar, nivel de vista de la propiedad, estado de la vivienda, nivel de construcción y diseño, el año de

construcción y el año de renovación ; los cuales tiene un gran valor o impacto en el modelo propuesto.

- En base a los factores que intervienen en el valor final de la vivienda, se puede afirmar que una casa es muy dependiente de los valores como numero de dormitorios, baños, medida de pies cuadrados del terreno y la condición o estado de la construcción, y que aún puede mejorarse el modelo mediante la utilización o consideración de factores como la zona o espacio geográfico de ubicación de las casas.

ANEXOS

DATASET:

El dataset utilizado para el trabajo propuesto consta de la siguiente estructura y 21 640 registros con sus respectivos precios finales y atributos considerados, los cuales fueron necesarios para la elaboración del modelo de regresión múltiple.

A	B	C	D	E	F	G	H	I	J	K	L
price	bedrooms	bathrooms	floors	waterfront	view	condition	grade	sqft_living	sqft_lot	yr_built	yr_renovated
221900	3	1	1	0	0	3	7	1180	5650	1955	0
538000	3	2.25	2	0	0	3	7	2570	7242	1951	1991
180000	2	1	1	0	0	3	6	770	10000	1933	0
604000	4	3	1	0	0	5	7	1960	5000	1965	0
510000	3	2	1	0	0	3	8	1680	8080	1987	0
1.23E+06	4	4.5	1	0	0	3	11	5420	101930	2001	0
257500	3	2.25	2	0	0	3	7	1715	6819	1995	0
291850	3	1.5	1	0	0	3	7	1060	9711	1963	0
229500	3	1	1	0	0	3	7	1780	7470	1960	0
323000	3	2.5	2	0	0	3	7	1890	6560	2003	0
662500	3	2.5	1	0	0	3	8	3560	9796	1965	0
468000	2	1	1	0	0	4	7	1160	6000	1942	0
310000	3	1	1.5	0	0	4	7	1430	19901	1927	0
400000	3	1.75	1	0	0	4	7	1370	9680	1977	0
530000	5	2	1.5	0	0	3	7	1810	4850	1900	0
650000	4	3	2	0	3	3	9	2950	5000	1979	0
395000	3	2	2	0	0	3	7	1890	14040	1994	0
485000	4	1	1.5	0	0	4	7	1600	4300	1916	0
189000	2	1	1	0	0	4	7	1200	9850	1921	0
230000	3	1	1	0	0	4	7	1250	9774	1969	0
385000	4	1.75	1	0	0	4	7	1620	4980	1947	0
2000000	3	2.75	1	0	4	3	9	3050	44867	1968	0
285000	5	2.5	2	0	0	3	8	2270	6300	1995	0
252700	2	1.5	1	0	0	3	7	1070	9643	1985	0
329000	3	2.25	2	0	0	4	8	2450	6500	1985	0
233000	3	2	1.5	0	0	5	6	1710	4697	1941	0
937000	3	1.75	2	0	0	3	8	2450	2691	1915	0
667000	3	1	1.5	0	0	5	8	1400	1581	1909	0
438000	3	1.75	1	0	0	3	7	1520	6380	1948	0
719000	4	2.5	2	0	0	3	8	2570	7173	2005	0
580500	3	2.5	2	0	0	3	8	2320	3980	2003	0
280000	2	1.5	3	0	0	3	7	1190	1265	2005	0
687500	4	1.75	1.5	0	0	4	7	2330	5000	1929	0
535000	3	1	1.5	0	0	4	8	1090	3000	1929	0
322500	4	2.75	1	0	0	3	7	2060	6659	1981	0
696000	3	2.5	1.5	0	0	3	8	2300	3060	1930	2002
550000	4	1	1	0	0	1	5	1660	34848	1933	0

Imagen1: Dataset utilizado

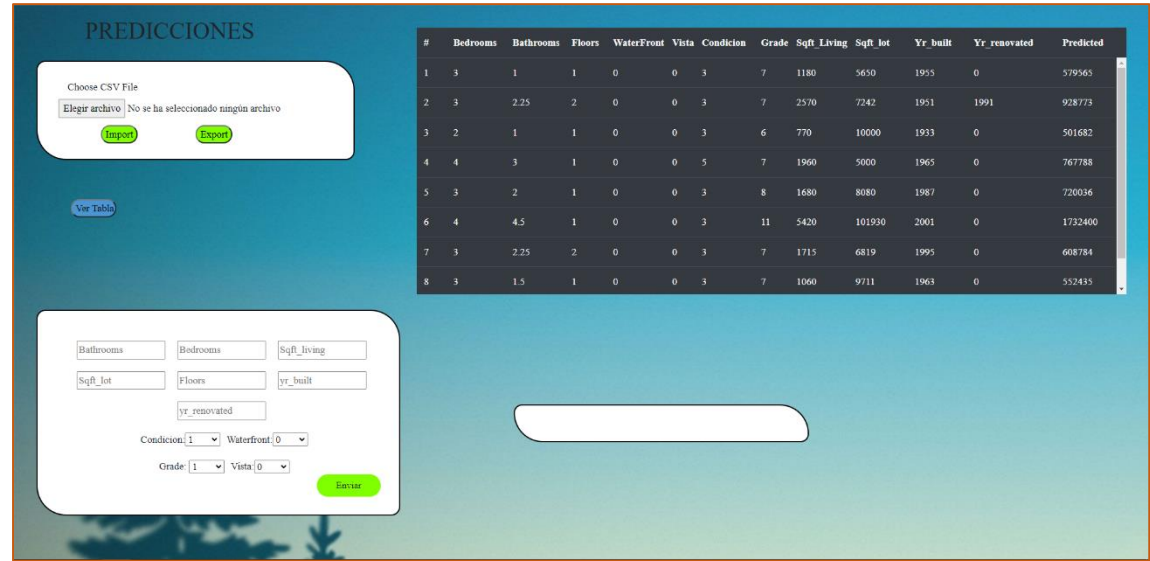
Los atributos considerados para la elaboración del modelo final de regresión múltiple fueron los siguientes:

```
house_prices house
# bedrooms : int(11)
# bathrooms : float
# floors : float
# waterfront : int(11)
# vista : int(11)
# condicion : int(11)
# grade : int(11)
# sqft_living : int(11)
# sqft_lot : int(11)
# yr_built : int(11)
# yr_renovated : int(11)
# price : float
```

Imagen2: Atributos utilizados

SOFTWARE:

El software consta de 2 funcionalidades, una de ellas posibilita al usuario la suba de archivos CSV, una vez importados el usuario puede descargarlos en el mismo formato y ya agregado el valor de la predicción o precio final de la vivienda basándose en los factores o atributos de entrada del usuario.



Asimismo, también tiene la opción de consultar manualmente el valor de una casa en específico mediante la utilización del formulario que se le ofrece para llenar los datos.

PREDICCIONES

Choose CSV File

Elegir archivo

No se ha seleccionado ningún archivo

Import

Export

Ver Tabla

Bathrooms

Bedrooms

Sqft_living

Sqft_lot

Floors

yr_built

yr_renovated

Condicion1

Waterfront0

Grade1

Vista0

Enviar

Segun Regresion Multiple, el precio de la casa seria :
\$79564.9133