Proyección de demanda en Industrias Metalmecánicas San Judas

Diana Barboza[1], Juan Caro[2], Iván Salazar[3]
[1−2−3] Estudiantes de la facultad de ingeniería.

Resumen—

En la empresa Industrias Metalmecanicas San Judas Ltda, se fabrican, reparan, mantienen y montan repuestos en los diferentes metales y/o materiales industriales y navales. Por ende se quiere proyectar la demanda de las tapas de alcantarillado fabricadas por esta misma, utilizando métodos de regresiones (Lineal, polinomial, entre otras. . .), con el fin de saber cuál de todas es la más conveniente.

Palabras claves: Demanda, regresiones, alcantarillado.

1. Introducción

Las distintas regresiones, por lo general, se utilizan para realizar una predicción de los datos a partir de otros datos. Por esa razón, se pondrá en práctica diferentes tipos de regresiones y comparar los resultados de cada uno de ellos, con el fin de determinar que tan relacionadas se encuentran nuestras variables(datos) y saber cuál de todas es una mejor opción para la empresa y para su sistema económico.

2. Objetivos

2.1. General

Comparar diferentes metodologías para la proyección de la demanda de las tapas de alcantarillado fabricadas por la empresa Industrias Metalmecánicas San Judas Ltda, basadas en datos históricos, con el fin de determinar cuál es la más adecuada para la toma de decisiones de la organización.

2.2. Especifico

- Calcular la relación entre variables para cada tipo de regresión
- Determinar el modelo mas adecuado para la predicción

3. Regresión

Regresión es una palabra un tanto rara. La utilizan los biólogos, los médicos, los psicólogos... y suena como ïr hacia atrás", "volver al pasado", y realmente este es verdadero significado del vocablo.

Fue un biólogo y estadístico inglés, SIR FRANCIS GALTON*, quien introdujo en 1889 el término regresión en Estadística. Empleó este concepto para indicar la relación que existía entre la estatura de los niños de una muestra y la estatura de su padre. Observó, que si los padres son altos, los hijos generalmente también lo son, y si los padres son bajos los hijos son también de menor estatura. Pero ocurría un hecho curioso: cuando el padre es muy alto o muy bajo, aparece una perceptible regresión"hacia la estatura media de la población, de modo que sus hijos retroceden hacia la media de la que sus padres, por cierto, están muy alejados.

Hoy día, el término no se utiliza en ese sentido. En muchas ocasiones, se desea conocer algo acerca de la relación o dependencia entre dos características cuantitativas, o más de una, consideradas sobre la misma población objeto de estudio (por ejemplo la talla y el peso).

Hay muchos casos en los que ya de antemano se ”sospecha”que puede existir algún tipo de relación, y por consiguiente, se pretende saber por ejemplo, en el caso de que tengamos únicamente dos variables:

- Si ambas variables están realmente relacionadas entre sí o si, por el contrario, pueden considerarse independientes.
- Si existe dependencia, es necesario conocer el ”grado de relación”, así como el ”tipo”de relación entre ambas.
- Si puede predecirse la variable que es considerada como dependiente a partir de los valores de la otra, que es considerada independiente, y si es así, con qué precisión.

4. Proyección en una empresa

Para una empresa tener ingresos debe de ser una de las cosas más importantes para ella; utilizando maniobras que le convenga a la empresa. Por ejemplo, producir por cantidad de productos y obtener el doble o el triple de ganancia estipulada es el objetivo principal de toda empresa, al igual que en el caso que le estamos presentando. Por eso tratamos de tener un análisis de manera exploratorio donde se busquen sistemas de solución para el trabajo de la empresa.

Problemática

Industrias Metalmecánicas San Judas Ltda. es una empresa de carácter mecánico, metalúrgico, metalmecánico y naval con más de 40 años de experiencia, al servicio de la industria, dedicada a la venta de bienes y servicios. Comprende la fabricación, reparación, mantenimiento y montaje de repuestos en los diferentes metales y/o materiales, industriales y navales. La organización trabaja bajo un modelo Job Shop o Taller, es decir, fabrica bajo pedido de sus clientes artículos específicos.

Las tapas de alcantarillado a diferencia de los demás productos de la empresa, se encuentran estandarizados y se producen durante todo el año. Por lo anterior se requiere analizar el comportamiento de ésta demanda a fin de predecir los niveles de producción que deben mantenerse para cumplir con los pedidos de los clientes. Actualmente, la empresa inicia el proceso de fabricación con el pedido (sistema de producción PULL), pero, se está explorando la idea de mantener inventarios de estos productos (sistema de producción PUSH). En este punto es que el estudio toma relevancia: permitiría tener una base cuantitativa para la toma de decisiones de la organización.

5. Metodología

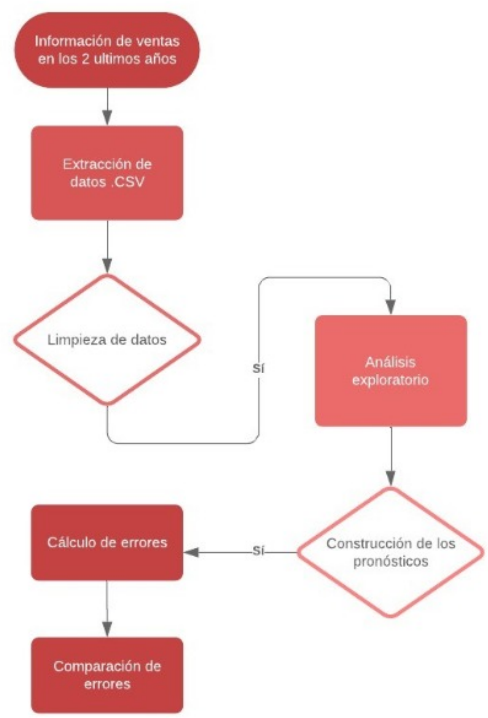

Figura 1: Diagrama de flujo de la metodología usada para diseñar el algoritmo.

6. Solución y análisis de datos

En primera instancia se recopilaron los datos de ventas de tapas de alcantarilla de esta empresa en los últimos dos(2) años en la siguiente tabla:

Mes	Ventas
1	40
2	42
3	44
4	40
5	49
6	52
7	58
8	45
9	46
10	58
11	54
12	41
13	42
14	43
15	45
16	41
17	50
18	50
19	54
20	43
21	47
22	60
23	56
24	40

Tabla 1: Ventas efectuadas por la empresa en los últimos 24 meses

Se guardaron en un archivo de formato csv para hacer la posterior lectura de ellos en el lenguaje de programación a usar, en este caso usamos Python para realizar nuestros modelos.
Como medida de seguridad se le realiza una limpieza a los datos, para evitar problemas con casillas vacías, datos faltantes, puntos, comas, etc, usando las siguientes lineas de código:

```
#Proyecto - Análisis de pronósticos de ventas
#Diana Barboza, Juan Caro e Ivan Salazar
#Extracción de datos de ventas desde un archivo .csv
import pandas as pd
import numpy as np
from scipy.interpolate import *
import matplotlib.pyplot as plt
#Mostrar solo dos decimales
def formato(dato):
    return "{0:.2f}".format(dato)
#Quitar puntos al User name random.randint(20,100)
def ClearDots(text):
    forbidden = {"?", "¿", "¡", "!", " ", ",", ".", ";", ":"}
    return "".join(c for c in text.lower() if c not in forbidden)
#Pone un numero aleatorio en los datos faltantes
def CleanData(cell):
    cell=ClearDots(cell)
    #Crear lista con el randint y concatenar los valores generados
    if(cell=="n.a." or cell=="-1" or cell=="not available" or cell=="NaN"):
        return random.randint(40,58)
    return cell
Ventas=pd.read_csv('Ventas.csv',index_col=0,usecols=['Ventas','Periodo'],sep=";",encoding='UTF-8',converters={'Ventas':CleanData
print("ANÁLISIS DE MÉTODOS DE PRONÓSTICOS DE VENTAS")
print("Aplicado a Industrias Metalmecánicas San Judas Ltda.\n")
print("DATOS CARGADOS DESDE EL ARCHIVO")
print(Ventas)
y=np.array(Ventas['Ventas'], dtype=int)
ListaAux=[]
```

Figura 2: Código para limpieza y lectura del archivo csv

```
ANÁLISIS DE MÉTODOS DE PRONÓSTICOS DE VENTAS
Aplicado a Industrias Metalmecánicas San Judas Ltda.

DATOS CARGADOS DESDE EL ARCHIVO
         Ventas
Periodo
1           40
2           42
3           44
4           40
5           49
6           52
7           58
8           45
9           46
10          58
11          54
12          41
13          42
14          43
15          45
16          41
17          50
18          50
19          54
20          43
21          47
22          60
23          56
24          40
```

Figura 3: Datos cargados desde el archivo

Luego de la lectura y limpieza de los datos, procedemos a guardarlos en un Array para poder usar la librería Numpy, la cual ayudará a realizar las distintas regresiones con los datos proporcionados en la Tabla 1.

```
y=np.array(Ventas['Ventas'], dtype=int)
ListaAux=[]
for i in range (0,len(y)):
    ListaAux.append(i+1)
x=np.array(ListaAux)
```

Figura 4: Guardado en un Array

Primero se realizó el modelo matemático de las ventas a través de regresión lineal simple usando la librería anteriormente mencionada, de la siguiente forma:

```
#Regresión Lineal
Reg1=np.polyfit(x,y,1)
print("\n Regresión lineal")
print ("Y = ",formato(Reg1[1])," + ",formato(Reg1[0]),"x")
```

Figura 5: Código para generar el resultado de la regresión lineal simple

Como resultado se generó la siguiente función lineal:

```
Regresión lineal
Y =  44.90  +  0.21 x
```

Figura 6: Modelo de ventas lineal

Se realizó otro proceso para generar un modelo a parir de una regresión polinomial de segundo grado usando las siguientes lineas de codigo:

```python
#Regresión Polinomial de grado 2
Reg2=np.polyfit(x,y,2)
print("\n Regresión polinomial de segundo grado")
print ("Y = ",formato(Reg2[2])," + ",formato(Reg2[1]),
      "x"," + ",formato(Reg2[0]),"x^2")
```

Figura 7: Código para generar el resultado de la regresión polinomial de segundo grado

Generando la siguiente expresión cuadrática:

```
Regresión polinomial de segundo grado
Y =  42.99  +  0.65 x  +  -0.02 x^2
```

Figura 8: Modelo de ventas cuadrático

Como ultimo proceso de modelado, se usó la regresión polinomial de tercer grado, el cual se ideó de la siguiente forma usando los datos anteriormente tabulados:

```python
#Regresión Polinomial de grado 3
Reg3=np.polyfit(x,y,3)
print("\n Regresión polinomial de tercer grado")
print ("Y = ",formato(Reg3[3])," + ",formato(Reg3[2]),
"x"," + ",formato(Reg3[1]),"x^2"," + ",formato(Reg3[0]),"x^3")
```

Figura 9: Código para generar el resultado de la regresión polinomial de tercer grado

El resultado fue la siguiente función cúbica:

```
Regresión polinomial de tercer grado
Y =  37.31  +  3.13 x  +  -0.26 x^2  +  0.01 x^3
```

Figura 10: Modelo de ventas cubico

Procedemos a calcular los valores de ventas(y) para un tiempo en meses(x), para posteriormente imprimir en forma de tabla, utilizando las siguientes lineas de codigo:

```python
for i in range (len(x)):
    y1.append(Reg1[1]+Reg1[0]*x[i])
    y2.append(Reg2[2]+Reg2[1]*x[i]+Reg2[0]*x[i]*x[i])
    y3.append(Reg3[3]+Reg3[2]*x[i]+Reg3[1]*x[i]*x[i]+Reg3[0]*x[i]*x[i]*x[i])

Ventas['Reg. LINEAL']= pd.DataFrame(y1)
Ventas['Reg. Pol. GRADO 2']= pd.DataFrame(y2)
Ventas['Reg. Pol. GRADO 3']=  pd.DataFrame(y3)
print (Ventas)
```

Figura 11: Código para generar los pronósticos de venta para cada modelo

La tabla generada fue:

RESULTADOS DE LOS PRONÓSTICOS

Periodo	Ventas	Reg. LINEAL	Reg. Pol. GRADO 2	Reg. Pol. GRADO 3
1	40	45.317826	44.214816	42.570185
2	42	45.525652	44.776549	44.518301
3	44	45.733478	45.302891	46.062102
4	40	45.941304	45.793843	47.240420
5	49	46.149130	46.249404	48.092090
6	52	46.356957	46.669574	48.655947
7	58	46.564783	47.054354	48.970825
8	45	46.772609	47.403743	49.075558
9	46	46.980435	47.717741	49.008981
10	58	47.188261	47.996348	48.809926
11	54	47.396087	48.239565	48.517230
12	41	47.603913	48.447391	48.169726
13	42	47.811739	48.619827	47.806249
14	43	48.019565	48.756871	47.465632
15	45	48.227391	48.858525	47.186710
16	41	48.435217	48.924789	47.008317
17	50	48.643043	48.955661	46.969288
18	50	48.850870	48.951143	47.108457
19	54	49.058696	48.911234	47.464658
20	43	49.266522	48.835935	48.076725
21	47	49.474348	48.725245	48.983493
22	60	49.682174	48.579164	50.223795
23	56	49.890000	48.397692	51.836467
24	40	NaN	NaN	NaN

Figura 12: Resultados de los pronósticos

Luego de tener los pronósticos para cada una de las regresiones con sus respectivos modelos, procedemos a calcular el error para cada una de las formas de regresión, esto se realiza apartir del calculo del coeficiente r^2:

```
#ERRORES
print ("\nANÁLISIS DE LOS PRONÓSTICOS - (R^2)\n")
#Calculo de errores
def calculo_R2(yreal,ypron):
    sumayreal=0
    for i in range (len(ypron)):
        sumayreal+=yreal[i]
    yrpromedio=sumayreal/len(yreal)
    res=0
    res2=0
    for i in range (len(ypron)):
        res+=(ypron[i]-yrpromedio)**2
        res2+=(yreal[i]-yrpromedio)**2
    return res/res2
```

Figura 13: Código para calcular los errores en cada tipo de regresión

Por ultimo se comparan los errores, y se realiza una linea de código que determiné cual de estos modelos es el mas adecuado a partir del análisis del coeficiente r^2:

```
#Comparación de errores
metodos=["Reg. Lineal","Reg. Pol. grado 2","Reg. Pol. grado 3"]
r2=[calculo_R2(y,y1),calculo_R2(y,y2),calculo_R2(y,y3)]
for i in range (len(r2)):
    print ("R^2 para el método ["+str(metodos[i])+"] = "+formato(r2[i]))
def posmayor (array):
    aux=0
    for i in range (len(array)):
        if(array[i]==max(array)):
            aux=i
            break
    return aux
#CONCLUSIONES
#Selección del método de prónostico
print("\nRESULTADO FINAL\n")
print("Debe seleccionarse el método ["+str(metodos[posmayor(r2)])+"]
```

Figura 14: Código para comparar los errores y determinar el modelo adecuado para la predicción

El resultado para los datos de la Tabla 1 fueron los siguientes:

```
ANÁLISIS DE LOS PRONÓSTICOS - (R^2)

R^2 para el método [Reg. Lineal] = 0.05
R^2 para el método [Reg. Pol. grado 2] = 0.07
R^2 para el método [Reg. Pol. grado 3] = 0.14

RESULTADO FINAL

Debe seleccionarse el método [Reg. Pol. grado 3]
```

Figura 15: Resultados de los errores para cada modelo y presentación del mas adecuado

7. Conclusiones

El mejor modelo para pronosticar las ventas de tapas de alcantarillado de la empresa Industrias Metalmecánicas San Judas es el modelo obtenido a partir de la regresión polinómica de tercer grado, ya que presentó un coeficiente de correlación menor al de los demás, esto se debe a que los datos están muy dispersos y no siguen un comportamiento lineal o cuadrático, generando errores muy grandes en estos dos modelos, situación que se atenúa, en gran parte, al usar expresiones cubicas

Referencias

1. Universidad de Salamanca. (s.f.). Regresión y Correlación. Recuperado 21 mayo, 2019, de http://biplot.usal.es/problemas/libro/6 %20 %20Regresion.pdf

2. S. Chapra, R. Canale. “Ajuste de Curvas” en Métodos numéricos para ingenieros. McGraw-Hill. 2007.

3. Universidad EAFIT. (s.f.). Teorías y modelos en los pronósticos de ventas. Recuperado 21 mayo, 2019, de http://publicaciones.eafit.edu.co/index.php/revista-universidad-eafit/article/download/1632/1619/

4. R. Chase, R. Aquilano y N. Jacobs. “Administración y Pronóstico de la demanda” en Administración de Operaciones: Producción y Cadena de Suministros. México: McGraw-Hill. 2009.