Técnicas de Minería de Datos

02-10-2020

Lindsey Zugey Alejandro Castillo 1676950
Equipo 03
Grupo 003
Minería de Datos

Detección de Outliers

La detección de outliers (del inglés traducido como “valores atípicos”) pertenece a la categoría descriptiva de las tareas de la minería de datos. El objetivo es encontrar patrones que den un resumen de las relaciones ocultas dentro de los datos. De esta forma, la detección de valores atípicos estudia el comportamiento de valores extremos que difieren del patrón general de una muestra.

Pero ¿qué son los valores atípicos? son observaciones cuyos valores son muy diferentes a las otras observaciones del mismo grupo de datos. Los datos atípicos son ocasionados por errores de entrada de datos y procedimiento, acontecimientos extraordinarios, valores extremos y/o faltantes, causas no conocidas. Los datos atípicos distorsionan los resultados de los análisis, y por esta razón hay que identificarlos y tratarlos de manera adecuada.

Y ¿cuándo un valor es atípico?, ¿cómo identificarlo?. Visualmente podemos apreciar en gráficas de dispersión, de líneas, diagramas de caja, distribuciones (normal, por ejemplo), de barras, etc., datos que son en extremo diferentes al promedio, datos que se ubican en lugares apartados de los demás, datos que son muy muy grandes y/o datos que son muy muy pequeños. Los valores observados en estas zonas de las gráficas corresponden a los outliers.

Existen distintos tipos de técnicas para detectarlos y se pueden dividir en dos categorías principales: métodos univariantes y métodos multivariantes. Una vez detectados los outliers se pueden eliminar o sustituir si se corrobora que se deben a un error de captura o en la medición de la variable.

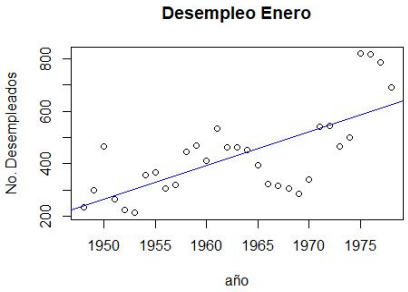
Podemos analizar distintos significados de outliers: error, límites, punto de interés.

Ejercicio

Se tiene una base con datos de desempleo por cada mes de 1948 hasta 1978, encuentre en que meses hubo datos atípicos.

Año	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiem	Octubr	Noviem	Diciem
1948	235.1	280.7	284.6	240.7	201.4	240.8	241.1	223.8	206.1	174.7	203.3	220.5
1949	239.5	347.4	338.3	327.7	351.6	396.6	438.6	395.6	363.5	376.6	357	369
1950	464.6	479.1	431.3	366.5	326.3	355.1	331.6	261.3	249	205.5	235.6	240.9
1951	264.5	253.6	232.3	193.6	177	213.2	207.2	180.6	186.6	175.4	199	179.6
1952	225.6	234	200.2	163.6	178.2	203.2	208.5	191.6	172.6	148	159.4	154.5
1953	213.2	196.4	162.6	178.4	153.6	173.2	171	151.2	161.9	157.2	201.7	236.4
1954	358.1	398.3	403.7	394.6	365.8	398.1	367.9	347	343.3	292.9	311.5	300.9
1955	368.5	358.9	329.7	316.2	269	289.3	266.2	253.6	233.6	228.4	253.6	280.1
1956	306.6	309.2	309.5	271	279.9	317.9	298.4	246.7	227.3	209.1	259.9	266
1957	320.6	308.5	262.2	262.7	263.5	310.1	294.3	252.6	250.3	246.5	312.7	333.2
1958	446.4	511.6	515.5	506.4	463.2	522.3	503.6	460.7	405.6	375	378.5	406.6
1959	467.8	469.8	429.8	355.8	332.7	378	360.5	334.7	319.5	323.1	363.6	352.1
1960	411.5	388.6	416.4	360.7	336	417.2	398.4	371.1	331.5	353.7	396.7	447
1961	533.5	565.4	542.3	488.7	467.1	531.3	496.1	444	403.4	396.3	394.1	404.1
1962	462.1	448.1	432.3	398.3	395.2	421.9	362.9	394.2	345.5	323.4	372.6	376
1963	462.7	497	444.2	399.3	394.9	459.4	414	375.5	347	339.4	385.8	378.8
1964	451.8	446.1	422.5	383.1	352.8	445.3	367.5	355.1	326.2	319.8	331.8	340.9
1965	394.1	417.2	369.5	349.2	321.4	405.7	342.9	316.5	294.2	270.9	288.8	278.9
1966	324.4	310.9	299	273	279.3	359.2	305	282.1	250.3	246.5	257.9	266.5
1967	315.5	318.4	295.4	269.4	245.6	362.6	324.5	294.2	268.5	295.2	296.3	272
1968	307.4	328.7	292.9	249.1	238.4	361.5	321.7	277.2	260.7	251	257.6	241.9
1969	287.5	292.3	274.7	254.2	230	339	318.2	287	295.6	284	271	262.7
1970	340.6	379.4	373.3	355.2	338.4	466.9	451	422	429.2	425.9	460.7	463.6
1971	541.4	544.2	517.5	469.4	439.4	549	533	506.1	484	457	481.5	469.5
1972	544.7	541.2	521.5	469.7	434.4	542.6	517.3	485.7	465.6	447	426.6	415.6
1973	467.5	484.5	451.2	417.4	379.9	494.7	455	420.8	416.5	376.3	405.6	405.8
1974	500.8	514	475.5	430.1	414.4	536	526	468.5	520.2	504.4	568.5	610.6
1975	616	639.9	635.5	782	762.3	856.9	828.9	769.6	752.2	724.4	723.1	719.5
1976	817.4	803.3	752.5	689	630.4	765.5	757.7	732.2	702.6	683.3	709.5	702.2
1977	784.6	810.9	755.6	656.6	676.1	745.3	694.1	675.7	643.7	622.1	634.6	586
1978	683.7	673.9	647.9	568.8	545.7	632.6	643.8	593.1	579.7	546	562.9	572.5

Podemos observar los datos atípicos graficando la recta de regresión lineal sobre el gráfico de dispersión, dichos datos son los que se encuentran más alejados de la línea de regresión.



Dato Sospechoso: 856.4 (Junio)

Prueba de Grubbs:

Estadístico T=|(Xo-Xm)|/S, (Xm= Media, Xo=Dato sospechoso, S=desviación std.) Según la prueba de Grubbs si T>Gtab entonces dicho dato es atípico, (Gtab: valor crítico en tablas para prueba de Grubbs)

S=161.26

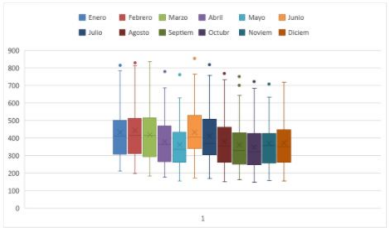
Xo=856.4

T=2.632

Gtab=2.76 (n=31, a= 0.05)

NO ES ATÍPICO

Resumen con Diagramas de caja



Xm= 433.92

Reglas de asociación

Búsqueda de patrones frecuentes, asociaciones, correlaciones o estructuras causales entre conjuntos de elementos u objetos en bases de datos de transacciones, bases de datos relacionales y otros repositorios de información disponibles.

Se pueden aplicar en análisis de datos de la banca, cross-marketing (poner la crema batida junto a las fresas), diseño de catálogos.

Entre los conceptos clave que se utilizan en las reglas de asociación están: soporte (fracción de transacciones que contiene un itemset), conjunto de elementos frecuentes (un conjunto de elementos cuyo soporte es mayor o igual que un umbral de mínimo, conjunto de elementos (una colección de uno o más artículos. k-itemset, un conjunto de elementos que contienen k elementos. Recuento de soporte (frecuencia de ocurrencia de un itemset), confianza (mide qué tan frecuente items en Y aparecen en transacción que contienen X.

El objetivo de la minería de reglas de asociación, dado un conjunto de transacciones T, es encontrar todas las reglas teniendo: umbral mínimo de soporte y umbral mínimo de confianza.

En las reglas de asociación de monería (enfoque en dos pasos) se encuentra la generación de elementos frecuentes donde se generan todos los conjuntos de elementos cuyo soporte $\geq$ min sup. Y la generación de reglas donde se generan reglas de alta confianza a partir de un conjunto de elementos frecuentes.

En las reglas de asociación principio “apriori” el principio si un conjunto de de elementos es frecuente, entonces todos sus subconjuntos también deben ser frecuentes. El soporte de un conjunto de elementos nunca excede el conjunto de sus elementos. Esto se conoce como la propiedad anti-monótona de soporte.

Ejercicio

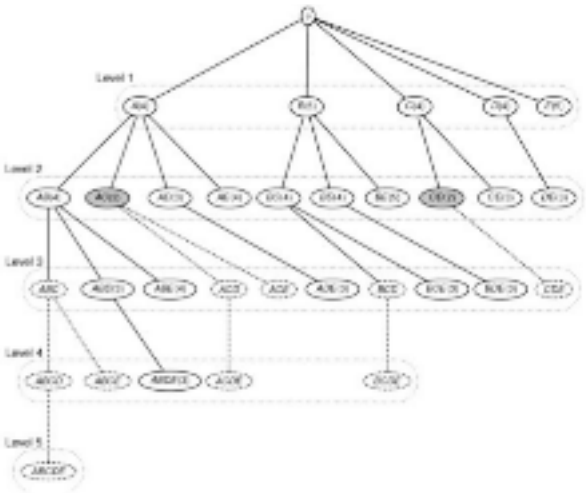
Dada la siguiente base de datos y un soporte mínimo de 3, genere todos los conjuntos de elementos frecuentes.

D	A	B	C	D	E
1	1	1	0	1	1
2	0	1	1	0	1
3	1	1	0	1	1
4	1	1	1	0	1
5	1	1	1	1	1
6	0	1	1	1	0

Base de Datos Binaria

t	i(t)
1	*ABDE*
2	*BCE*
3	*ABDE*
4	*ABCE*
5	*ABCDE*
6	*BCD*

Base de Datos Transaccionaría



Regresión

Una regresión lineal es un modelo matemático para determinar el grado de dependencia entre una o más variables, es decir conocer si existe relación entre ellas. Existen dos tipos de regresión: regresión lineal (cuando una variable independiente ejerce influencia sobre otra variable dependiente) y regresión lineal múltiple (cuando dos o más variables independientes influyen sobre una variable dependiente.

La regresión en minería de datos se encuentra en la categoría predictivo, donde el objetivo es analizar los datos de un conjunto y en base a eso, predecir lo que puede ocurrir con ese conjunto de datos en un futuro.

El análisis de regresión permite examinar la relación entre dos o más variables e identificar cuáles son las que tienen mayor impacto en un tema de interés. Donde las variables dependientes son el factor más importante, el cual se está tratando de entender o predecir. Y las variables independientes son el factor que tú crees que puede impactar en tu variable dependiente.

El análisis de regresión nos permite explicar un fenómeno y predecir cosas acerca del futuro, por lo que nos será de ayuda para tomar decisiones y obtener los mejores resultados.

Ejercicio

Temperatura	Uso/1000
x	y
21	185.79
24	214.47
32	288.03
47	424.84
50	454.68
59	539.03
45	320.05

Se cree que la cantidad de libras de vapor usadas en la planta por mes está relacionada con la temperatura ambiente promedio. A continuación se presentan los consumos y las temperaturas del último año. Indique la ecuación estimada para la regresión (modelo lineal), su ajuste R^2 ajd, y su gráfica correspondiente.

```
import numpy as np

x = [21,24,32,47,50,59,45]
y = [185.79,214.47,288.03,
424.84,454.68,539.03,320.05
n = len(x)
n = len(y)
x = np.array(x)
y = np.array(y)
sumx = sum(x)
sumy = sum(y)
sumx2 = sum(x**2)
sumy2 = sum(y**2)
sumxy = sum(x*y)
promx = sumx/n
promy = sumy/n
m = (sumx*sumy - n*sumxy) /
(sumx**2 - n*sumx2)
b = promy - m*promx

sigmax = np.sqrt(sumx2/n -
promx**2)
sigmay = np.sqrt(sumy2/n -
promy**2)
sigmaxy = sumxy/n -
promx*promy
R2 = (sigmaxy /
(sigmax*sigmay))**2
porc = R2*100
print("\nR2 = " + str(porc) + " %")
```

Clustering

Traducido del inglés como “agrupamiento”, el clustering es una de las técnicas de minería de datos, su proceso consiste en la división de los datos en grupos de objetos similares. Las técnicas de clustering son las que, utilizando algoritmos matemáticos, se encargan de agrupar objetos. Usando la información que brindan las variables que pertenecen a cada objeto, se mide la similitud entre los mismos, y una vez hecho esto se colocan en clases que son muy similares internamente y a la vez diferentes entre los miembros de las diferentes clases.

Para la introducción a ésta técnica se deben conocer los conceptos de: cluster (colección de objetos de datos similares entre sí dentro del mismo grupo y disimilar a los objetos en otros grupos), análisis de cluster (dado un conjunto de puntos de datos tratar de entender su estructura, encuentra similitudes entre los datos de acuerdo con las características encontradas en los datos, es un aprendizaje no supervisado ya que no hay clases preferidas).

Entre las aplicaciones que tiene el clustering se encuentran los estudios de terremotos, aseguradoras, uso del suelo, marketing y planificación de la ciudad. Por otro lado están los métodos de agrupamiento: asignación jerárquica frente a punto, datos numéricos y/o simbólicos, determinística vs probabilística, exclusivo vs superpuesto, jerárquico vs plano, de arriba abajo y de abajo a arriba.

Algoritmos de clustering: simple k-means (debe definir el número de clusters que se desean obtener, así se convierte en un algoritmo voraz para particionar) y x-means (selecciona a priori el número de clusters que se deseen obtener, se le define un límite inferior k-min y k-max, obteniendo así, en ese rango, el número óptimo de clusters, usa parámetros cobweb y EM).

Ejercicio

```
from sklearn.cluster import KMeans
from sklearn import metrics
import numpy as np
```

```
#X = [[3,5],[1,4],[1,6],[2,6],[1,5],[6,8],[6,6],[6,7],[5,6],[6,7],[7,1],[8,2],[9,1],[8,2],[9,3],[9,2],[8,3], ...]

v1=[0, 3, 1, 1, 2, 1, 6, 6, 5, 6, 7, 5, 6, 7, 8, 10, 9, 8, 9, 9, 9, 11, 13, 9, 15, 14, 13, 12, 14, 12]
v2=[5, 5, 4, 6, 6, 5, 8, 6, 7, 7, 6, 7, 1, 2, 0, 1, 2, 3, 2, 3, 12, 11, 10, 13, 12, 13, 10, 10, 11]

x1 = np.array(v1)
x2 = np.array(v2)
```

```
X = np.array(list(zip(x1, x2))).reshape(len(x1), 2)
print(X)

import matplotlib.pyplot as plt
plt.plot(v1, v2, 'ro')
plt.axis([-1, 16, -1, 15]) #Eje x: de -1 a 16; Eje Y: de -1 a 15
plt.show()
```

```
#Algoritmo K-Means
K = 4
kmeans_model = KMeans(n_clusters=K).fit(X)
```

```
for i, l in enumerate(kmeans_model.labels_):
    print("(x1,x2) -> Clase")
    print("({0},{1}) ->:{2}".format(x1[i], x2[i], l))
```

```
predicciones = kmeans_model.predict(X)
print(predicciones)
```

```
test=[[10,15]]
prediccion = kmeans_model.predict(test)
print("Prediccion {0}->{1}".format(test,prediccion))
```

```
x1 = np.array([2, 5, 8, 12])
x2 = np.array([5, 5, 4, 14])

X = np.array(list(zip(x1, x2))).reshape(len(x1), 2)
prediccion = kmeans_model.fit_predict(X)
print(prediccion)
```

```
print(kmeans_model.cluster_centers_)
```

```
[[12. 14.]
 [ 2.  5.]
 [ 8.  4.]
 [ 5.  5.]]
```

Predicción

De la categoría predictiva de las tareas de la minería de datos, la predicción es una técnica que se utiliza para proyectar los tipos de datos que se verán en el futuro o predecir el resultado de un evento. Existen cuestiones relativas a la relación temporal de las variables de entrada o predictores de la variable objetivo. Los valores son generalmente continuos. Las predicciones son a menudo sobre el futuro.

Cualquiera de las técnicas utilizadas para la clasificación y estimación puede ser adaptada para su uso en la predicción mediante el uso de ejemplos de entretenimiento donde el valor de la variable que se predijo que ya es conocido, junto con los datos históricos de esos ejemplos. Los datos históricos se utilizan para construir un modelo que explica el comportamiento observado en los datos. Cuando este modelo se aplica a nuevas entradas de datos, el resultado es una predicción del comportamiento futuro de los mismos.

Entre algunas de sus aplicaciones están: revisar los historiales crediticios de los consumidores y las compras pasadas para predecir si serán un riesgo crediticio en el futuro, predecir si va a llover en función de la humedad actual, predecir el precio de venta de una propiedad y predecir la puntuación de cualquier equipo durante un partido de fútbol.

La mayoría de las técnicas de predicción se basan en modelos matemáticos: modelos estadísticos simples como regresión, estadísticas no lineales como series de potencias, redes neuronales, etc. Todo esto basado en ajustar una curva a través de los datos, es decir, encontrar una relación entre los predictores y los pronosticados.

Ejercicio

Pronóstico de ventas futuras.

```
ultimosDias = df['2018-11-16':'2018-11-30']
ultimosDias
```

```
fecha
2018-11-16 152
2018-11-17 111
2018-11-19 207
2018-11-20 206
2018-11-21 183
2018-11-22 200
2018-11-23 187
2018-11-24 189
2018-11-25 76
2018-11-26 276
2018-11-27 220
2018-11-28 183
2018-11-29 251
2018-11-30 189
Name: unidades, dtype: int64
```

```
values = ultimosDias.values
values = values.astype('float32')
# normalize features
values=values.reshape(-1, 1) # esto lo hacemos porque tenemos 1 sola dimension
scaled = scaler.fit_transform(values)
reframed = series_to_supervised(scaled, PASOS, 1)
reframed.drop(reframed.columns[[7]], axis=1, inplace=True)
reframed.head(7)
```

	var1(t-7)	var1(t-6)	var1(t-5)	var1(t-4)	var1(t-3)	var1(t-2)	var1(t-1)
7	-0.24	-0.65	0.31	0.30	0.07	0.24	0.11
8	-0.65	0.31	0.30	0.07	0.24	0.11	0.13
9	0.31	0.30	0.07	0.24	0.11	0.13	-1.00
10	0.30	0.07	0.24	0.11	0.13	-1.00	1.00
11	0.07	0.24	0.11	0.13	-1.00	1.00	0.44
12	0.24	0.11	0.13	-1.00	1.00	0.44	0.07
13	0.11	0.13	-1.00	1.00	0.44	0.07	0.75

```
values = reframed.values
x_test = values[6:, :]
x_test = x_test.reshape((x_test.shape[0], 1, x_test.shape[1]))
x_test
```

```
array([[[ 0.11000001, 0.13 , -1. , 1. ,
0.44000006, 0.06999993, 0.75 ]]], dtype=float32)
```

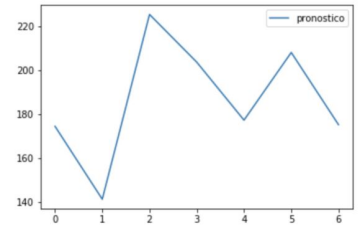
```
def agregarNuevoValor(x_test,nuevoValor):
    for i in range(x_test.shape[2]-1):
        x_test[0][0][i] = x_test[0][0][i+1]
    x_test[0][0][x_test.shape[2]-1]=nuevoValor
    return x_test

results=[]
for i in range(7):
    parcial=model.predict(x_test)
    results.append(parcial[0])
    print(x_test)
    x_test=agregarNuevoValor(x_test,parcial[0])
```

```
adimen = [x for x in results]
inverted = scaler.inverse_transform(adimen)
inverted
```

```
array([[174.48904094],
[141.26934129],
[225.49292353],
[203.73262324],
[177.30941712],
[208.1552254 ],
[175.23698644]])
```

```
prediccion1SemanaDiciembre = pd.DataFrame(inverted)
prediccion1SemanaDiciembre.columns = ['pronostico']
prediccion1SemanaDiciembre.plot()
prediccion1SemanaDiciembre.to_csv('pronostico.csv')
```



Patrones secuenciales

Dos importantes conceptos conforman esta técnica de la minería de datos perteneciente a la categoría predictiva. El primero corresponde a minería de datos secuenciales (extracción de patrones frecuentes relacionados con el tiempo u otro tipo de secuencia, son eventos que se enlazan con el paso del tiempo) y reglas de asociación secuencial (expresión de patrones secuenciales, es decir, que se dan en instantes distintos en el tiempo).

Las características de los patrones secuenciales son: el orden importa, el objetivo es encontrar patrones secuenciales, el tamaño de una secuencia es su cantidad de elementos, la longitud de la secuencia es su cantidad de ítems, el soporte de una secuencia es el porcentaje de secuencias que la contienen en un conjunto de secuencias S, las secuencias frecuentes son las subsecuencias de una secuencia que tiene soporte mínimo.

Predecir si un compuesto químico causa cáncer, el comportamiento de compras y el reconocimiento de spam de un correo electrónico son ejemplos de las aplicaciones de patrones secuenciales. Donde los primeros dos corresponden al agrupamiento de patrones secuenciales y el último a clasificación con datos secuenciales.

En el proceso de los patrones secuenciales se encuentran las secuencias donde |s| es el número de elementos de una secuencia, una k-secuencia es una secuencia con k eventos. Así mismo están las subsecuencias que son secuencias que están dentro de otras pero cumpliendo ciertas normas. El ítem del evento i de la subsecuencia, tiene que estar dentro del evento i de la secuencia.

El análisis de secuencias va desde la base de datos hasta el elemento (ítem) pasando por la secuencia y el elemento (transacción).

Ejemplo

Candidatos

$w_1 = \langle \{1\}\ \{2\ 3\}\ \{4\} \rangle\ w_2 = \langle \{2\ 3\}\ \{4\ 5\} \rangle$

Producen $\langle \{1\}\ \{2\ 3\}\ \{4\ 5\} \rangle$

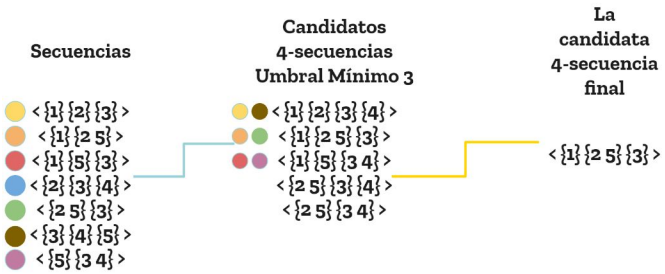
$w_1 = \langle \{1\}\ \{2\ 3\}\ \{4\} \rangle\ w_2 = \langle \{2\ 3\}\ \{4\}\ \{5\} \rangle$

Producen $\langle \{1\}\ \{2\ 3\}\ \{4\}\ \{5\} \rangle$

$w_1 = \langle \{1\}\ \{2\ 6\}\ \{4\} \rangle\ w_2 = \langle \{1\}\ \{2\}\ \{4\ 5\} \rangle$

Producen $\langle \{1\}\ \{2\ 6\}\ \{4\ 5\} \rangle$

Generalized Sequential Pattern



Visualización

Perteneciente a la categoría descriptiva de las tareas de la minería de datos, la visualización de datos nos sirve para representar gráficamente los elementos más importantes de nuestra base de datos, esto es, la presentación de información en formato gráfico o ilustrado. Al utilizar elementos visuales como cuadros, gráficos o mapas, nos proporcionan una manera accesible de ver y comprender tendencias, valores atípicos y/o patrones en los datos.

Entre los tipos de visualización más comunes están gráficos (circulares, líneas, columnas, barras aisladas o agrupadas, burbujas, áreas, diagramas de dispersión, mapas de tipo árbol), mapas, infografías, cuadros de mando (dashboards).

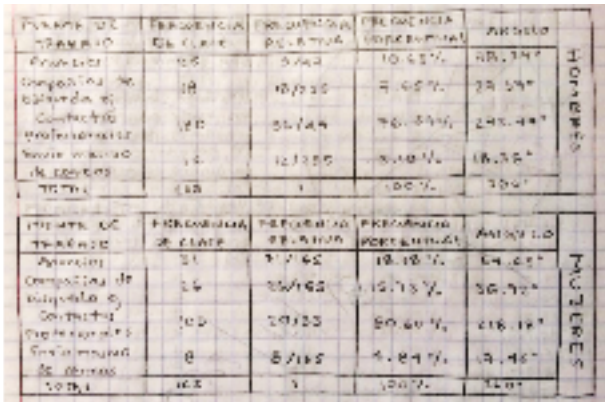
La mayoría de los analistas de datos utilizan software avanzado para explorar y visualizar datos, Y las herramientas de software van desde hojas de cálculo, sencillas con Excel o Google sheets a software de analítica más sofisticado, como R.

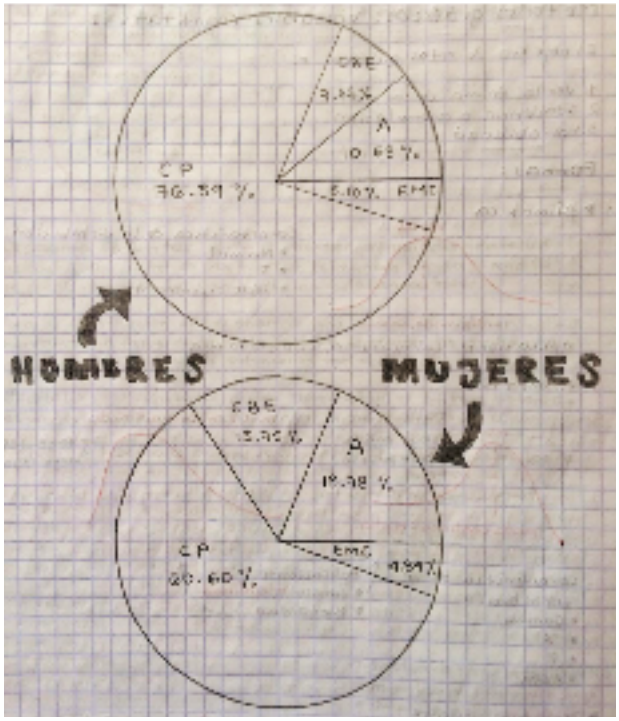
El objetivo de aplicar esta técnica es comprender la información con rapidez, identificar relaciones y patrones, identificar tendencias emergentes y comunicar la historia a otras personas. Una buena visualización cuenta una historia, eliminando el ruido de los datos y resaltando la información útil.

Ejemplo

Se lleva a cabo un estudio para determinar cómo obtenía trabajo la gente. La tabla siguiente lista los resultados de una muestra de sujetos en edad laboral. Los datos están basados en resultados del centro nacional de estrategias profesionales. Utilizando un método apropiado, ¿cuál parece ser el enfoque más apropiado para conseguir trabajo?

Fuente de trabajo	Hombre	Mujer
Anuncios	25	31
Compañías de búsqueda ejecutivas	18	26
Contactos profesionales	180	100
Envíos masivos de correos	12	8





Clasificación

La clasificación es una técnica de minería de datos de la categoría predictiva, es el ordenamiento o disposición por clases tomando en cuenta las características de los elementos que contiene. Uno de los objetivos de la minería de datos es la clasificación, la cual tiene como fin clasificar una variable dentro de una de las categorías de una clase.

La clasificación en minería de datos es una técnica supervisada, donde generalmente se tiene un atributo llamado clase y se busca determinar si los atributos pertenecen o no a un determinado concepto.

La clasificación, es la habilidad para adquirir una función que mapee (clasifique) un elemento de dato a una de entre varias clases predefinidas. Un objeto se describe a través de un conjunto de características (variables o atributos) $X \rightarrow \{X1, X2, \ldots, Xn\}$. El objetivo de la tarea de clasificación es clasificar el objeto dentro de una de las categorías de la clase $C = \{C1, \ldots, Ck\}$ $f: X1 \times X2 \times \ldots \times Xn \rightarrow C$.

Para clasificar existen métodos, entre los cuales se encuentran: análisis discriminante (método utilizado para encontrar una combinación lineal de rasgos que separan clases de objetos o eventos), reglas de clasificación (buscan términos no clasificados de forma periódica, si se encuentra una coincidencia se agrega a los datos de clasificación), árboles de decisión (método analítico que a través de una representación esquemática facilita la toma de decisiones) y redes neuronales artificiales (también conocido como sistema conexionista, es un modelo de unidades conectadas para transmitir señales).

Aunque diferentes todos ellos comparten las mismas características: precisión en la predicción, eficiencia, robustez, escalabilidad, interpretabilidad.

Ejercicio

En EUA los maestros clasifican a los estudiantes en A, B, C, D o F según sus notas. Utilizando simples límites (60, 70, 80, 90) las siguientes clasificaciones son posibles.

nota ≥ 90	A
$80 \leq$ nota < 90	B
$70 \leq$ nota < 80	C
$60 \leq$ nota < 70	D
nota < 60	F

Se obtuvieron las siguientes clasificaciones de un grupo al azar y se les pide que se les clasifique por resultados obtenidos en los grupos mostrados anteriormente.

Jesús = 81	B	Daniel = 94	A
María = 74	C	Tania = 34	F
Sebastián = 99	A	Armando = 63	D
Manuel = 50	F	Paulina = 44	F
Karla = 82	B	Iván = 96	A