



Análisis Predictivo con R

Presentado por :
MCI. Carlos Morales Castro



Temas de la presentación

1

Introducción

Breve descripción de la importancia del análisis predictivo

2

Conceptos y definiciones

Se detallan diferentes conceptos necesarios para entender los modelos de Machine Learning

3

Ejemplo

Se implementa y estudia el modelo llamado "Estadísticas por año, materia y despacho "

4

Implementación

Se implementa y estudia el modelo llamado "Monto susceptible a no ejecutarse"

5

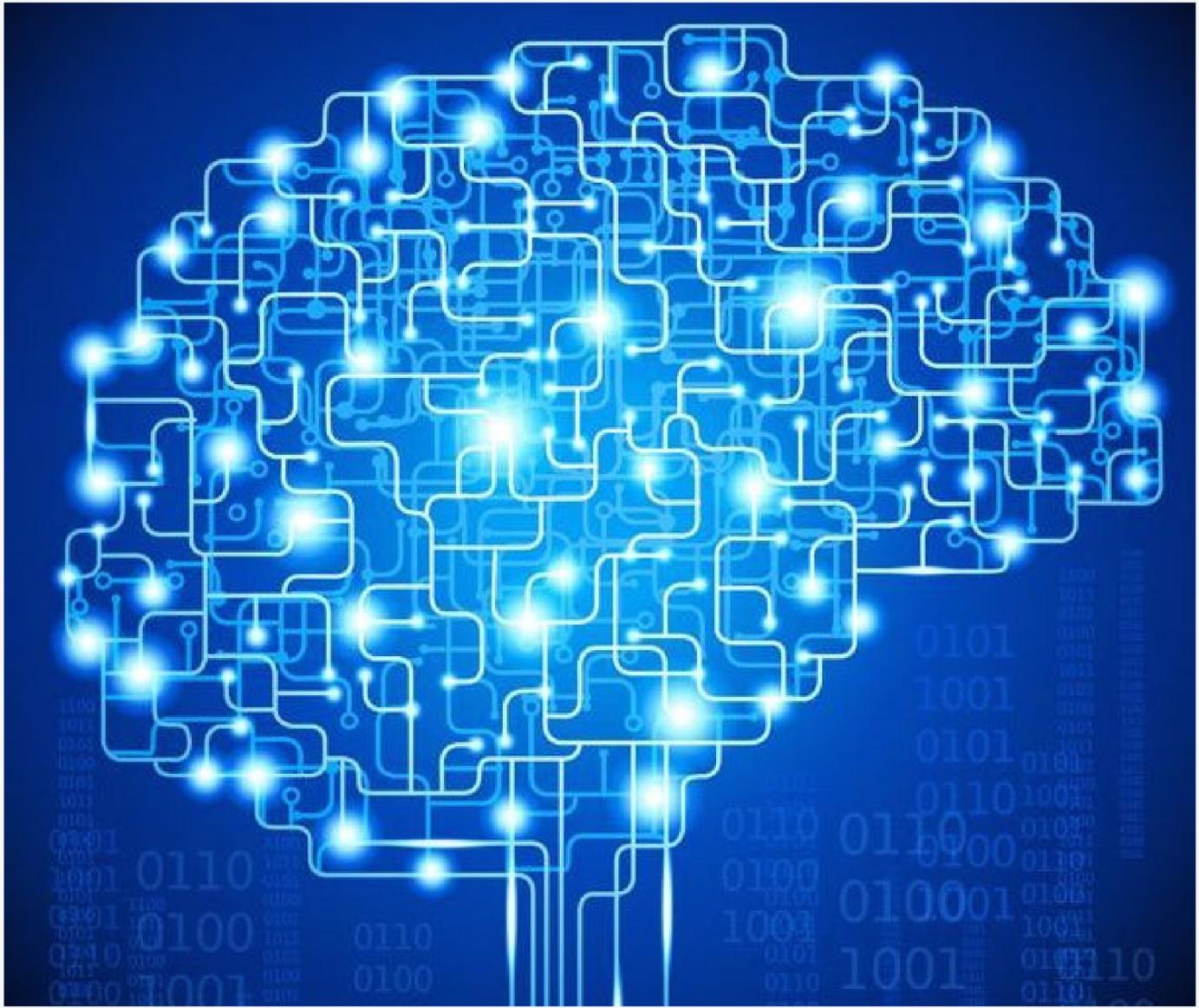
Consideraciones y costo de Azure ML Studio

Consejos y mejores practicas para Azure ML Studio. Detalle de costos.

6

Conclusiones

Comentarios finales de la presentación



Introducción

Breve descripción de la importancia del análisis predictivo

Importancia de análisis predictivo

Nos permite identificar debilidades en los procesos.

Nos permite simular cambios para valorar impactos o situaciones.

Nos permite identificar y analizar tendencias.

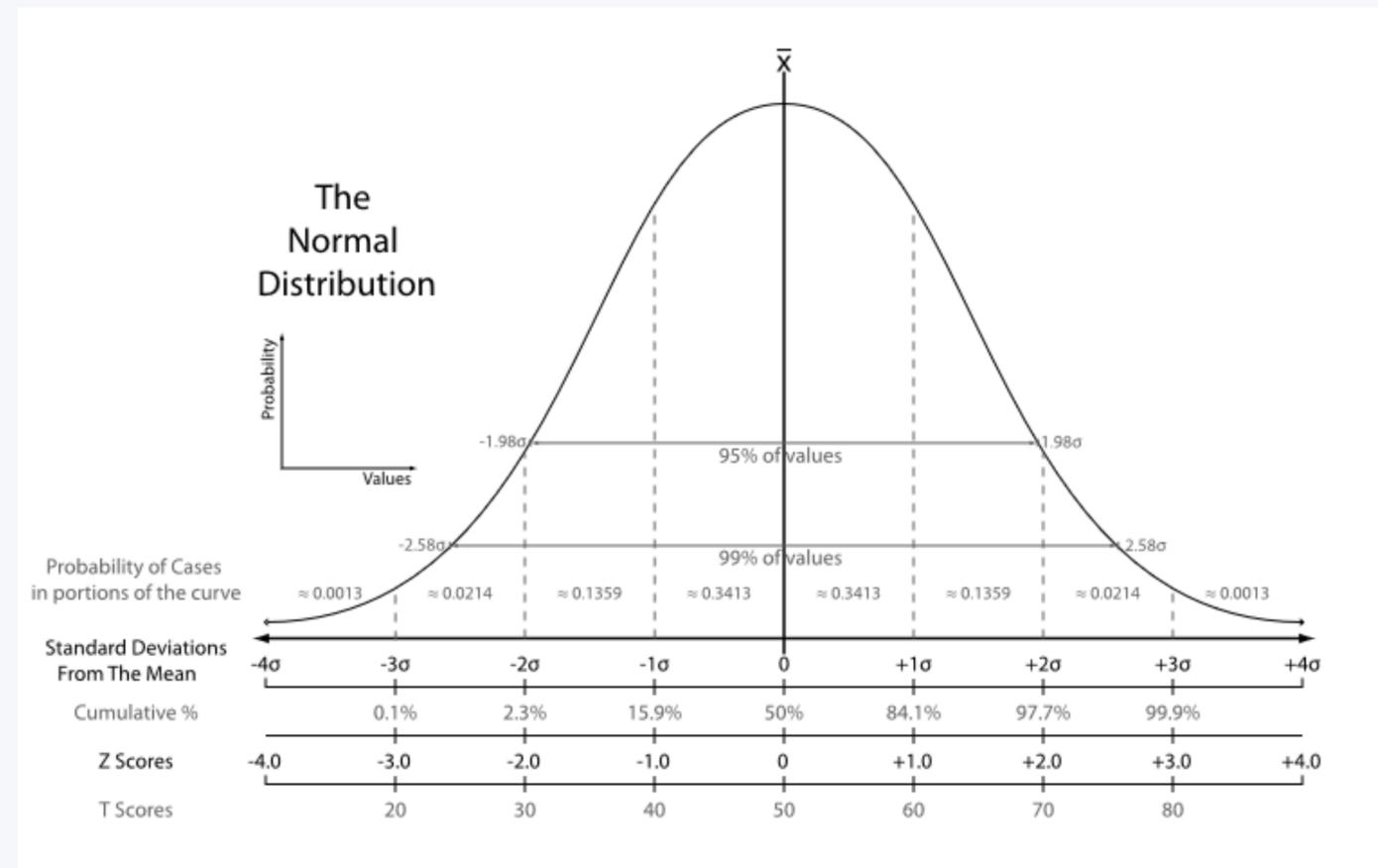
Nos permite desarrollar procesos proactivos para evitar eventos no deseados.



Conceptos y definiciones

Se detallan diferentes conceptos necesarios para entender los modelos de Azure Machine Learning Studio

“ Usos y análisis provenientes de una muestra representativa de datos ”



Probabilidad

“ Se refiere a la certeza con la que se la a producir un evento ”



Ejemplo práctico

Elige un número.....

1 2 3 4

Datos

El 75% elige el número 3

El 20% eligen el número 2 o el 4

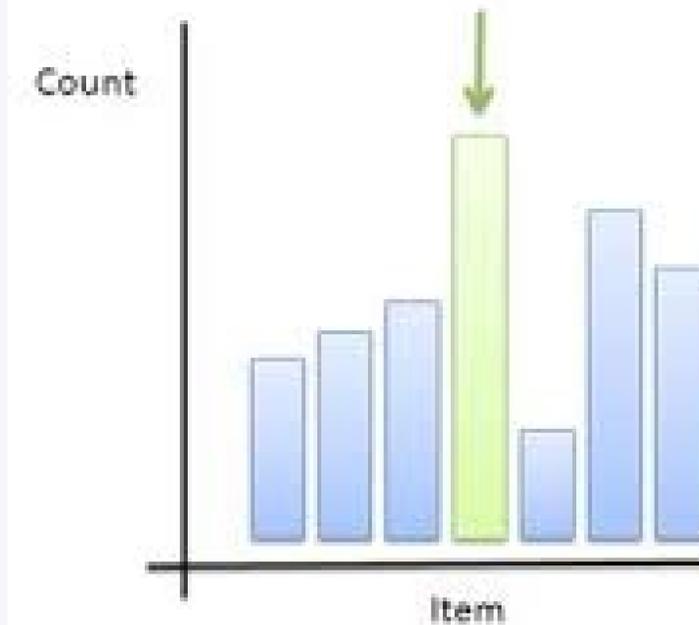
El 5% eligen el número 1

Media, mediana, modo

$$\bar{x} = \frac{\sum x}{n}$$

$$M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n+1}{2}}}{2}$$

Mode (Most Popular)



Media

Se refiere al promedio

Mediana

Valor que se encuentra en el medio de todos los valores cuando se ordenan

Modo

Se refiere al valor que más se repite

Desviación estándar

“ Representa la distancia entre el valor y el promedio ”

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Correlación

Representa la relación sistemática entre dos variables cuantitativas

Fecha	Lugar	Limonada	Naranja	Temperatura	Folletos	Precio	Ventas	Ingresos
01/07/2016	Parque	97	67	70	90	0,25	164	41
02/07/2016	Parque	98	67	72	90	0,25	165	41,25
03/07/2016	Parque	110	77	71	104	0,25	187	46,75
04/07/2016	Playa	134	99	76	98	0,25	233	58,25
05/07/2016	Playa	159	118	78	135	0,25	277	69,25
06/07/2016	Playa	103	69	82	90	0,25	172	43
07/07/2016	Playa	143	101	81	135	0,25	244	61
08/07/2016	Playa	123	86	82	113	0,25	209	52,25
Total de ingr								412,75

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \left[n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}}$$

Tipos de Variables

- Cuantitativas
- Ordinales
- Categóricas

	Limonada	Naranja	Temperatura	Folletos	Precio	Ventas	Ingresos
Limonada	1						
Naranja	0,99671434	1					
Temperatura	0,47734477	0,45311565	1				
Folletos	0,85751196	0,82633261	0,28720915	1			
Precio	-0,27053036	-0,31808337	-0,03357457	0,03204048	1		
Ventas	0,99930904	0,99903613	0,46661642	0,8439048	-0,29257238	1	
Ingresos	0,46927552	0,42695527	0,3394465	0,60059337	0,70225907	0,45023895	1

Machine Learning

Es un campo de la Inteligencia Artificial que le da la habilidad de aprender y mejorar a los sistemas computacionales



Ejemplo

Se implementa y estudia el modelo llamado
"Estadísticas por año, materia y despacho"

Implementación del modelo "Estadísticas por año, materia y despacho"

A	B	C	D	E	F	G	H	I	J	K	L
Anno	Mes	NombreM	CodMater	NombreMater	CodCircuito	NombreCircuito	CodDespacho	NombreDespacho	CodTipoDespacho	NombreTipoDespacho	CirculanteInicial
2017	1	Enero	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	9160
2017	2	Febreo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	9328
2017	3	Marzo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	9431
2017	4	Abril	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	9607

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Anno	Mes	NombreM	CodMater	NombreMater	CodCircuito	NombreCircuito	CodDespacho	NombreDespacho	CodTipoDespacho	NombreTipoDespacho	CirculanteInicial	Circulante
161	2015	1	Enero	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	7525	
695	2015	2	Febreo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	7723	
984	2015	3	Marzo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	7925	
1544	2015	4	Abril	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	7790	
1830	2015	5	Mayo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	7587	
2382	2015	6	Junio	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	7776	
2703	2015	7	Julio	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	7947	
5384	2016	1	Enero	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	8389	
5790	2016	2	Febreo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	8651	
5462	2016	3	Marzo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	8864	
5786	2016	4	Abril	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	8575	
7586	2016	5	Mayo	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	8182	
7943	2016	6	Junio	PA	PENSIONES AL	17	III CIRCUITO JUDIC	256	JUZGADO DE PENS	JA	JUZGADO DE PEN	8138	

Anno	Mes	NombreM	CodMater	NombreMater	CodCircuito	NombreCircuito	CodDespacho	NombreDespacho	CodTipoDespacho	NombreTipoDespacho	CirculanteInicialPre
2017	5	Mayo	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9149
2017	6	Junio	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9234
2017	7	Julio	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9149
2017	8	Agosto	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9255
2017	9	Septiembre	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9342
2017	10	Octubre	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9083
2017	11	Noviembre	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9211
2017	12	Diciembre	PA	PENSIONES A	17	III CIRCUITO J	256	JUZGADO DE JA		JUZGADO DE PENSI	9333

Implementación

Se implementa y estudia el modelo llamado
"Monto susceptible a no ejecutarse"

Predicciones

Monto susceptible a no ejecutarse

Pedido

- Sin disponible presupuestario
- En estado Visado o Caducado
- Fecha de afectación de la transferencia
- Fecha del Ultimo caduco

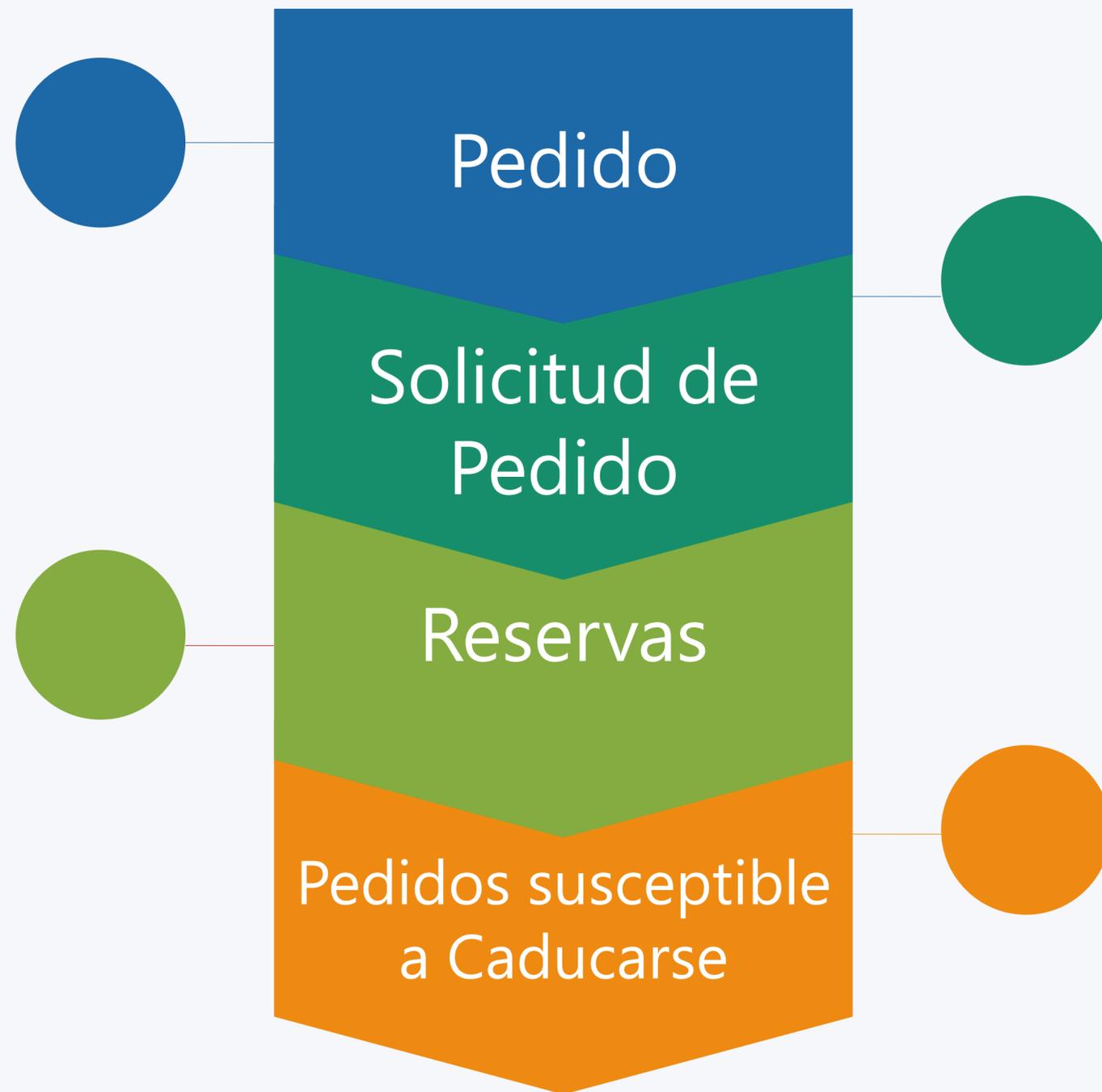
Reservas

- Sin disponible presupuestario
- En estado Visado o Caducado
- Fecha de creación de pedido
- Fecha del ultimo caduco

Solicitudes del pedido

- Sin disponible presupuestario
- En estado Visado o Caducado
- Fecha de creación de pedido
- Fecha del ultimo caduco

- Se consideran pedidos cuyo monto a comprar es igual a la cantidad a comprar
- Sin facturas en tramites (cancelado para evitar malos entendidos en recalculo)
- Con disponible mayor a cero



* Fecha al 01 de julio del 2018 (Prueba de concepto)

Predicciones (consideraciones)

Monto susceptible a no ejecutarse

Pedido (16.974)

- Periodo
- Indicador de consultoría
- Indicador de provienen de expediente
- Indicador de CND
- Oficina
- Subpartida
- Área tramitadora
- Tipo de Pedido
- Rubro
- Tipo de Gasto
- Tipo de moneda
- Fecha de creación

Solicitudes del pedido (17.893)

- Periodo
- Indicador de consultoría
- Indicador de provienen de expediente
- Indicador de CND
- Oficina
- Subpartida
- Área tramitadora
- Tipo de Pedido
- Rubro
- Tipo de Gasto
- Tipo de moneda
- Fecha de creación

Reservas (18.946)

- Periodo
- Indicador de CND
- Subpartida
- Tipo de reserva
- Tipo de Gasto
- Fecha de creación

- Reservas 2013
- Solicitudes de Pedido 2014
- Pedidos 2015
- CND desde el 2015



* Fecha al 01 de julio del 2018 (Prueba de concepto)

Consideraciones

Consejos y mejores practicas.



Pasos para desarrollar un modelo de Azure ML Studio

- 1 Identificar caso de uso
- 2 Identificar datos
- 3 Limpieza de datos
- 4 Elegir algoritmo
- 5 Optimización del modelo

- 6 Evaluar resultados
- 7 Publicar modelo

- 8 Entrenar nuevamente



Conclusiones

Comentarios finales de la presentación



Conclusiones



Limites???



Futuros pasos...



Metas....

01

Se pueden llegar a predecir muchas variables.

02

Aplicar un modelo completo no es sencillo se debe analizar la necesidad los datos y evaluar los modelos.

03

La muestra de datos para entrenar debe ser lo más completa y limpia posible.

04

La implementación de modelos requieren conocimiento de la lógica del negocio, conocimiento técnico y experiencia para tener el mejor resultado.