



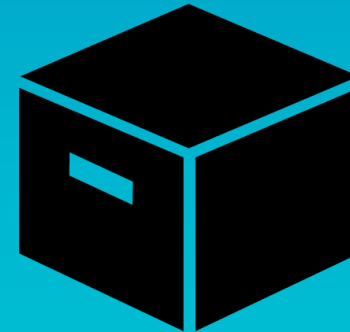
DMC ONLINE

#YoMeQuedoEnCasa

TALLER:

# Explicación e Interpretabilidad de Algoritmos de Machine Learning en la Industria

## Descifrando Cajas Negras



¡Hola! Soy,

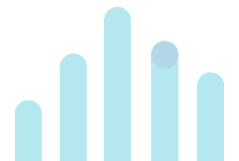
**André Chávez**

Lead Data Scientist Estadístico. MSc. Data Science



# AGENDA

- Conceptos básicos y palabras claves.
- ¿Qué es **Machine Learning** ?.
- Las 2 Culturas : Explicación vs. Interpretación.
- Evolución de los Algoritmos de Machine Learning.
- Aplicaciones de ML en la Industria.
- Interpretabilidad en Algoritmos de Machine Learning.
- Repaso desde Metodologías Estadísticas hasta Metodologías Computacionales.



# EXPLICACIÓN E INTERPRETABILIDAD MODELOS MACHINE LEARNING

## RELACIONES ENTRE TÉRMINOS

- ❑ **Conjunto de Datos (Data Set):** El total del conjunto de datos sobre los que queremos desarrollar un algoritmo de Machine Learning con el fin de obtener un modelo que lo represente lo mejor posible. Contendrá variables independientes y dependientes.
- ❑ **Variables Independientes (Features), (VI):** Aquellas columnas del Data Set que serán usadas por el algoritmo para generar un modelo que prediga lo mejor posible las variables dependientes.
- ❑ **Variables dependientes (Labels, Target), (VD):** Columna del data set que responde a una correlación de VI y que debe ser predicha por el futuro modelo
- ❑ **Conjunto de Datos de Entrenamiento (Training Set):** Subconjunto del Data Set que será utilizado para entrenar el modelo que se pretende generar.
- ❑ **Conjunto de Datos de Test (Test Set):** Subconjunto del data set que se le pasará al modelo una vez haya sido entrenado para comprobar, mediante el uso de diferentes métricas, sus indicadores más importantes de calidad.



# Explicación e Interpretabilidad Modelos Machine Learning

Relaciones entre términos

## ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



## MACHINE LEARNING

Machine learning begins to flourish.



## DEEP LEARNING

Deep learning breakthroughs drive AI boom.



# Machine Learning



“Es realmente difícil programar o hacer una tarea manual por lo general repetitiva pues resulta poco eficiente y tediosa. Tal vez un algoritmo o sistema pueda entender lo que necesito que haga. Tal vez pueda enseñarle cómo hacerlo, si tiene la estructura correcta y buenos ejemplos”.



# Proceso de Detección de Fraudes



# Explicación e Interpretabilidad Modelos Machine Learning

❑ Género.



❑ Rangos de Edad.



❑ Ingresos.



❑ Estado Civil.



$f_x$

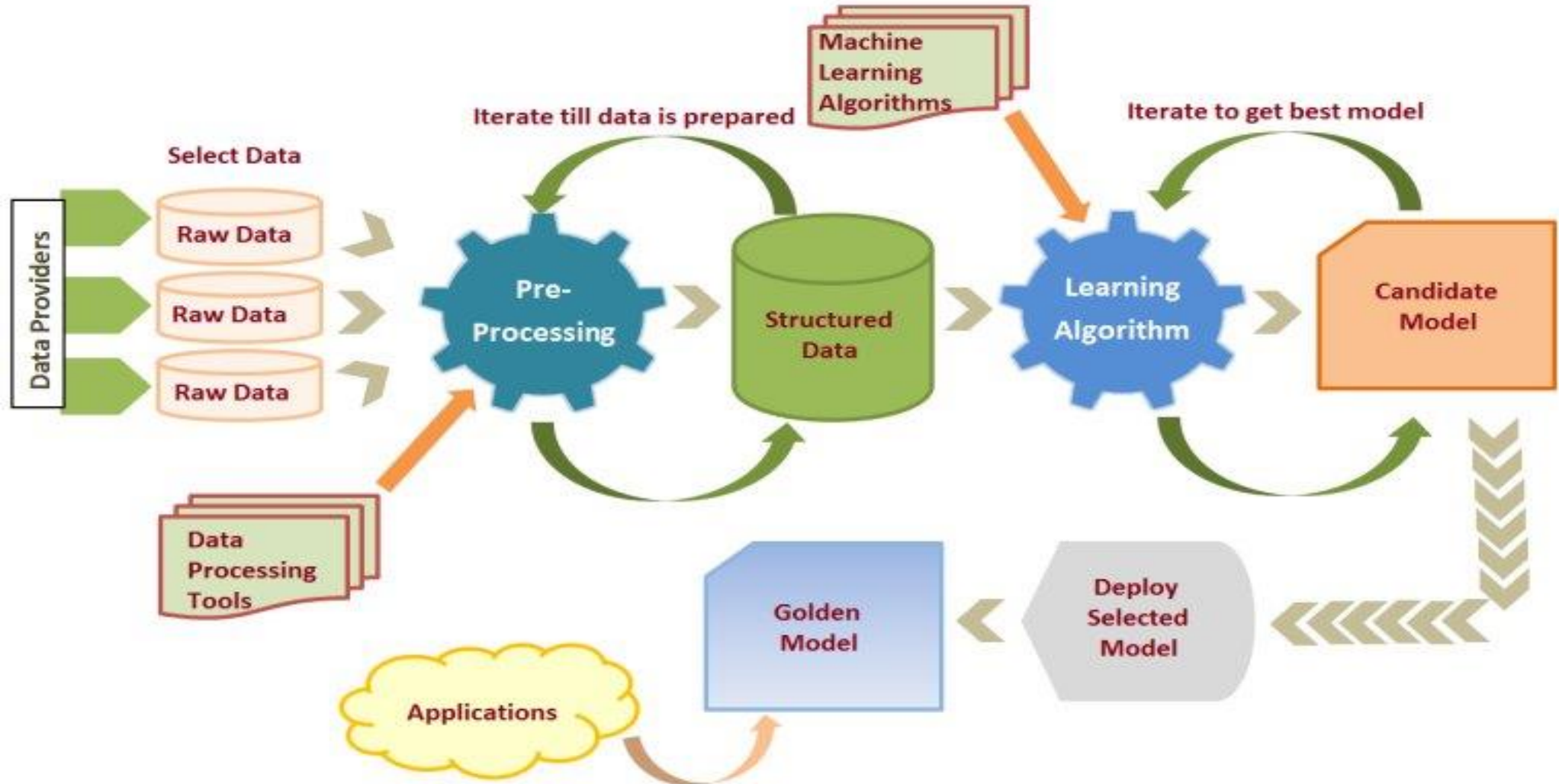
❑ Fraude

❑ No Fraude



# Explicación e Interpretabilidad Modelos Machine Learning

# Estructura de una solución de Machine Learning



## Applications

# Explicación e Interpretabilidad Modelos Machine Learning

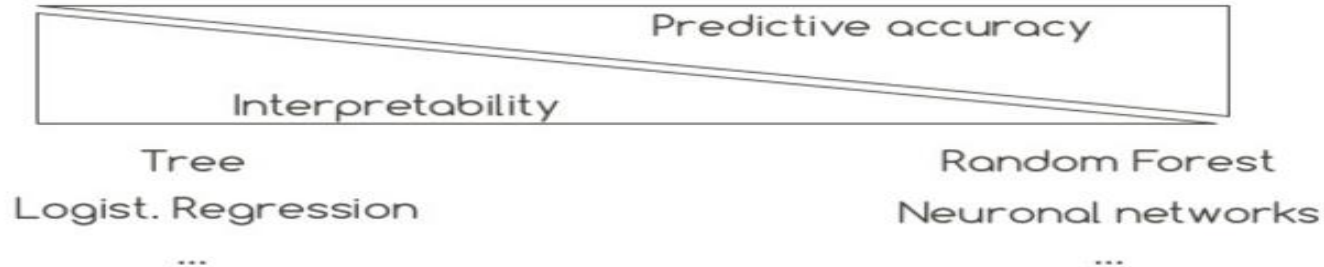
Modelamiento de Estadístico : Las dos culturas

*There are two cultures in the use of statistical modeling to reach conclusions from data. One assumes that the **data are generated by a given stochastic data model**. The other **uses algorithmic models and treats the data mechanism as unknown**. The statistical community has been committed to the almost exclusive use of data models... Algorithmic modeling, both in theory and practice, has developed rapidly in fields outside statistics. It can be used both on large complex data sets and as a more accurate and informative alternative to data modeling on smaller data sets. If our goal as a field is to use data to solve problems, then we need to move away from exclusive dependence on data models and adopt a more diverse set of tools.*

(Statistical Science, 2001, with discussants)

# Explicación e Interpretabilidad Modelos Machine Learning

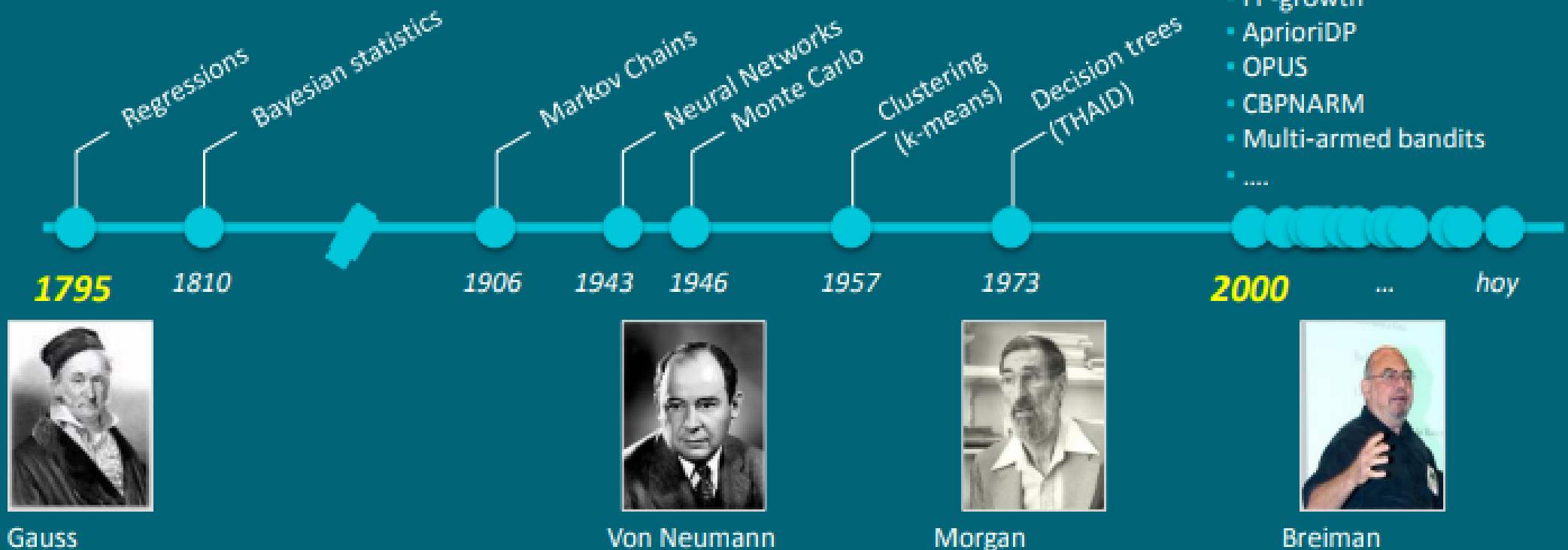
Modelamiento de Estadístico : Las dos culturas



## Evolución de algoritmos estadísticos

### Classical stats (pre Big Data era)

### Big Data native stats (Machine Learning)





# Explicación e Interpretabilidad Modelos Machine Learning

Aplicaciones de Machine Learning en la Industria.

1



## ENERGÍA

- Predecir fallas en refinerías
- Localizar nuevas fuentes de energía
- Analizar minerales

2



## SERVICIOS

- Fijar precios acorde a la demanda
- Alcanzar un ritmo de ventas óptimo

3



## GOBIERNO

- Elevar eficiencia y ahorros
- Minimizar el robo de identidad
- Prevenir la corrupción

4



## TRANSPORTE

- Identificar rutas más eficientes
- Predecir problemas

5



## MINORISTAS

- Mejorar campañas de mercadotecnia
- Personalizar la oferta
- Reducir la pérdida de clientes durante el proceso de compra
- Mejorar la experiencia de compra

6



## HOSPITALES

- Incrementar el éxito de una operación
- Predecir tiempos de espera en urgencias
- Prevenir infartos y convulsiones

7



## FINANCIERO

- Prevenir créditos incobrables
- Predecir reclamaciones
- Prevenir fraudes y lavado de dinero

# Explicación e Interpretabilidad Modelos Machine Learning

## Interpretación y Explicabilidad

- ❑ **Explique:** Aclara (una idea o situación) a alguien describiéndolo con más detalle o revelando hechos relevantes.
- ❑ **Interpretar:** Explicar el significado de (información o acciones).
- ❑ **Entender:** Percibir el significado deseado de (palabras, un idioma o un hablante).
- ❑ **Comprender:** Agarrar mentalmente; entender
- ❑ **Inteligible:** Capaz de ser entendido; comprensible.

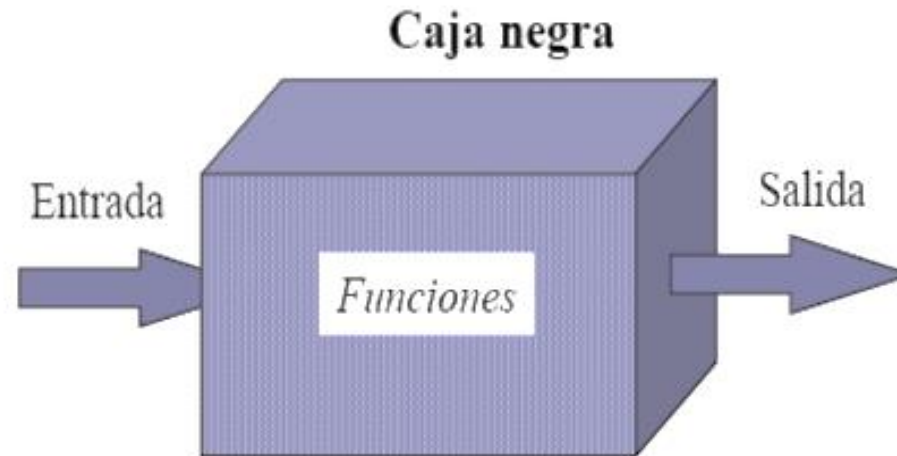
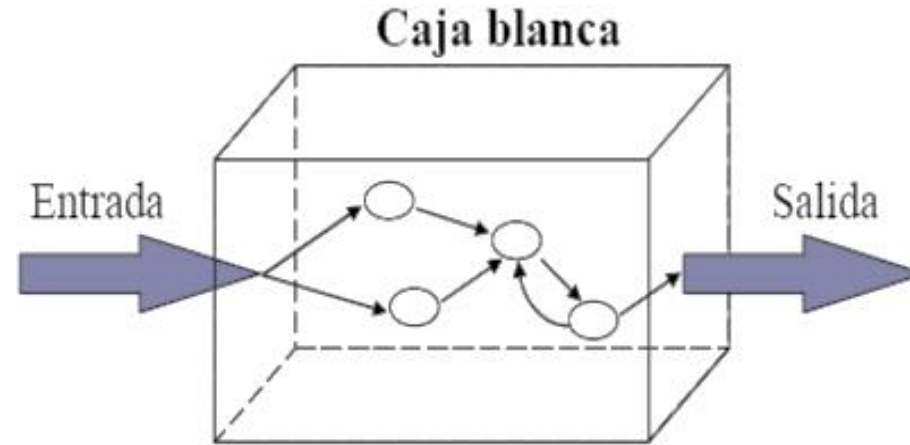
[Diccionario, Oxford English 2018]





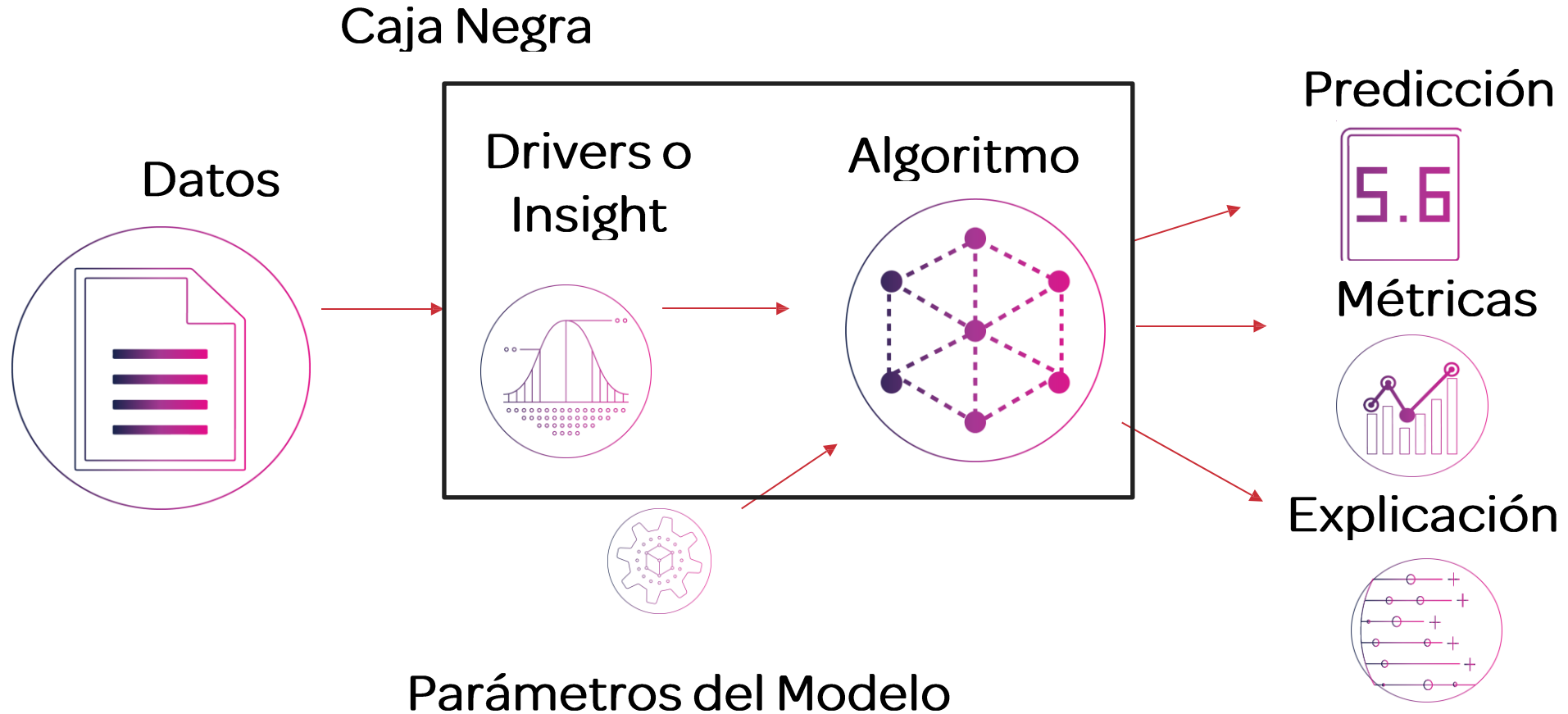
# Explicación e Interpretabilidad Modelos Machine Learning

Descifrando Algoritmos de Cajas Negras



# Explicación e Interpretabilidad Modelos Machine Learning

Descifrando Algoritmos de Cajas Negras

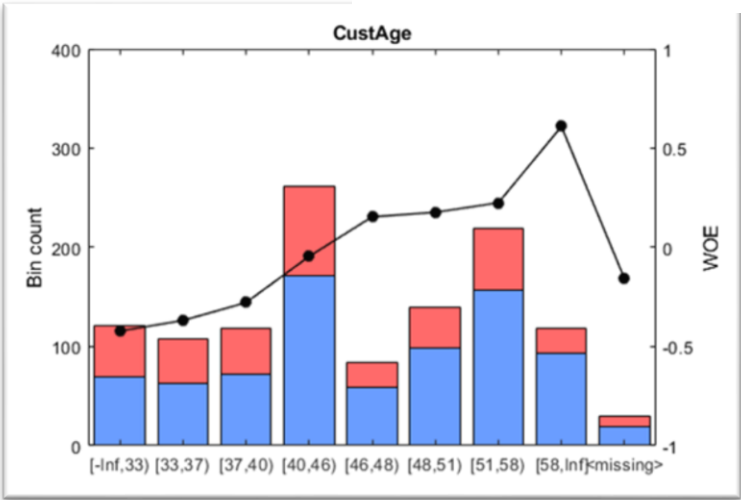


# Explicación e Interpretabilidad Modelos Machine Learning

Desde Metodologías Estadísticas hasta Computacionales.

Metodología Bivariadas  
WOESS

$$WOE = \ln \left( \frac{\text{Event\%}}{\text{Non Event\%}} \right)$$



Information Value (IV)	Predictive Power
< 0.02	useless for prediction
0.02 to 0.1	weak predictor
0.1 to 0.3	medium predictor
0.3 to 0.5	strong predictor
> 0.5	suspicious or too good to be true

# Explicación e Interpretabilidad Modelos Machine Learning

Desde Metodologías Estadísticas hasta Computacionales.

## Metodología Multivariadas ELI5

### Interpretación Global

Weight	Feature
0.4278	Sex=female
0.1949	Pclass=3
0.0665	Embarked=S
0.0510	Pclass=2
0.0420	SibSp
0.0417	Cabin=
0.0385	Embarked=C
0.0358	Ticket=1601
0.0331	Age
0.0323	Fare
0.0220	Pclass=1
0.0143	Parch
0	Name=Rothes, the Countess. of (Lucy Noel Martha Dyer-Edwards)
0	Name=Roebling, Mr. Washington Augustus II
0	Name=Rosblom, Mr. Viktor Richard
0	Name=Ross, Mr. John Hugo
0	Name=Rush, Mr. Alfred George John
0	Name=Rouse, Mr. Richard Henry
0	Name=Ryerson, Miss. Emily Borie
0	Name=Ryerson, Miss. Susan Parker "Suzette"
... 1972 more ...	

### Interpretación Local

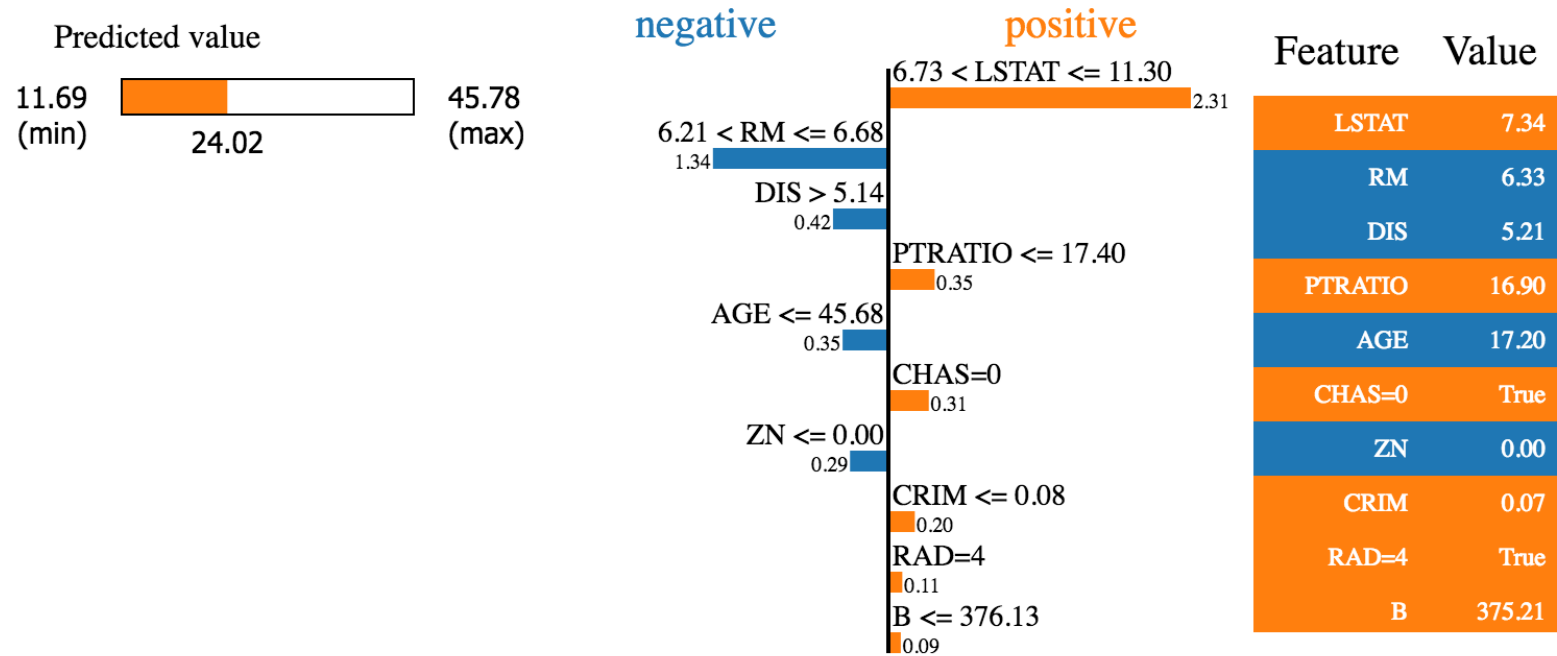
y=1 (probability 0.566, score 0.264) top features			
Contribution?	Feature	Value	
+1.673	Sex=female	1.000	
+0.479	Embarked=S	Missing	
+0.070	Fare	7.879	
-0.004	Cabin=	1.000	
-0.006	Parch	0.000	
-0.009	Pclass=2	Missing	
-0.009	Ticket=1601	Missing	
-0.012	Embarked=C	Missing	
-0.071	SibSp	0.000	
-0.073	Pclass=1	Missing	
-0.147	Age	19.000	
-0.528	<BIAS>	1.000	
-1.100	Pclass=3	1.000	

Un método similar se describe en Breiman, "Random Forests", Machine Learning, 45 (1), 5-32, 2001 (disponible en línea en <https://www.stat.berkeley.edu/%7Ebreiman/randomforest2001.pdf>).

# Explicación e Interpretabilidad Modelos Machine Learning

Desde Metodologías Estadísticas hasta Computacionales.

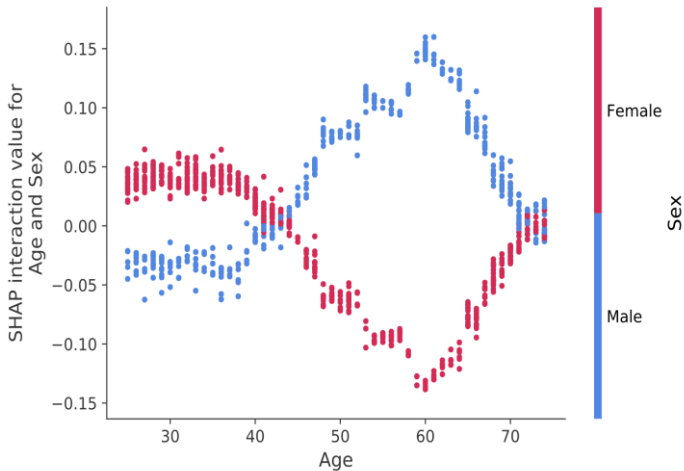
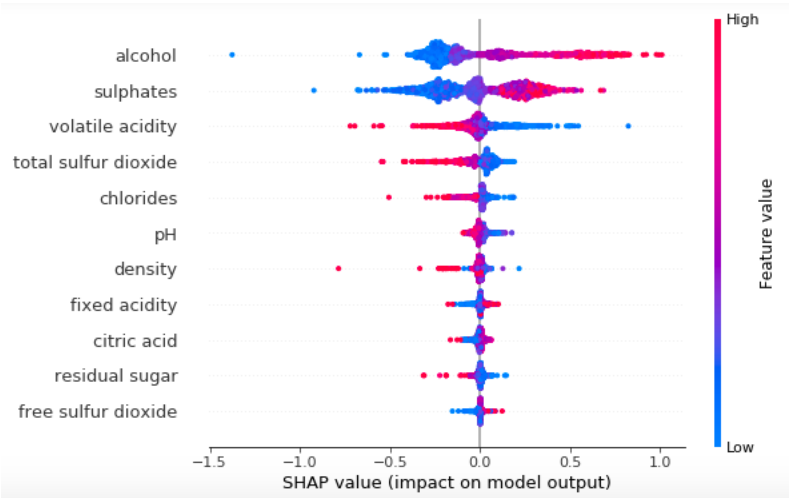
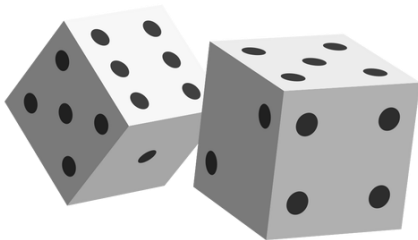
## Metodología Multivariadas LIME



# Explicación e Interpretabilidad Modelos Machine Learning

Desde Metodologías Estadísticas hasta Computacionales.

## Metodología Multivariadas SHAP

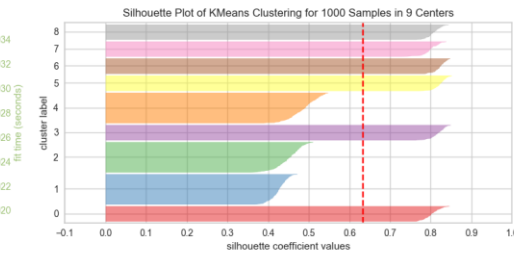
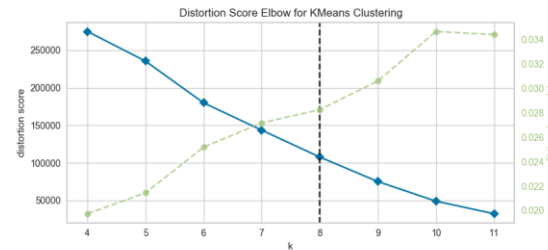
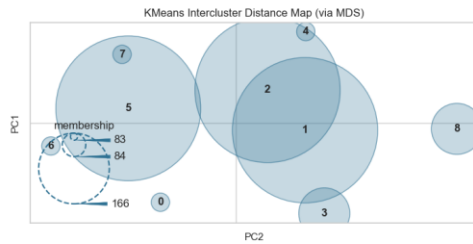
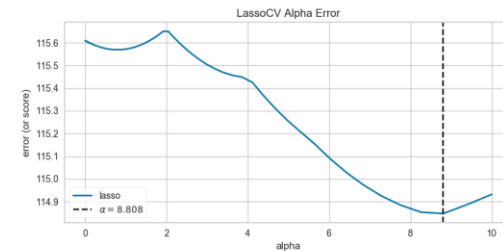
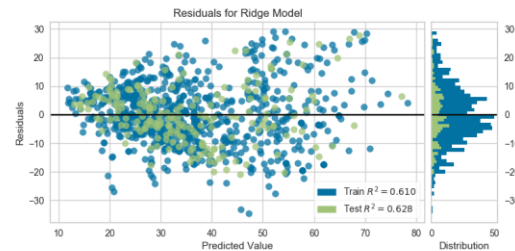
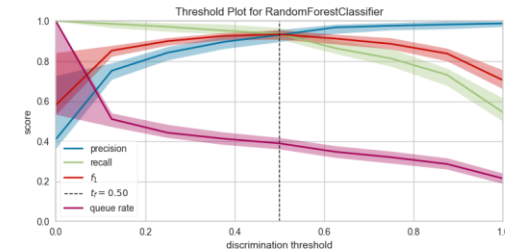
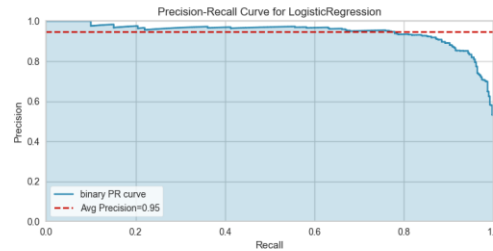
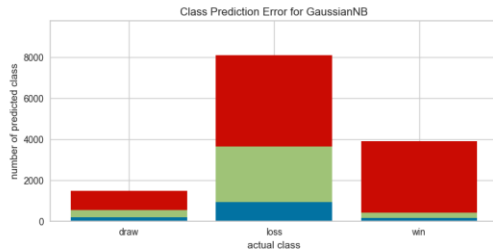




# Explicación e Interpretabilidad Modelos Machine Learning

Desde Metodologías Estadísticas hasta Computacionales.

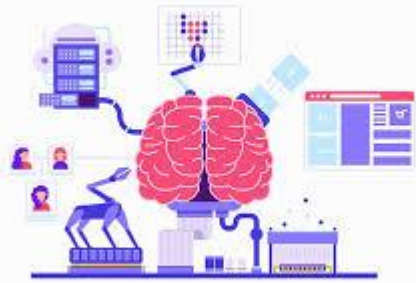
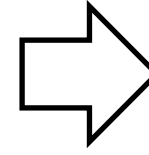
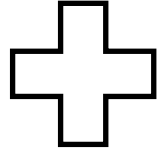
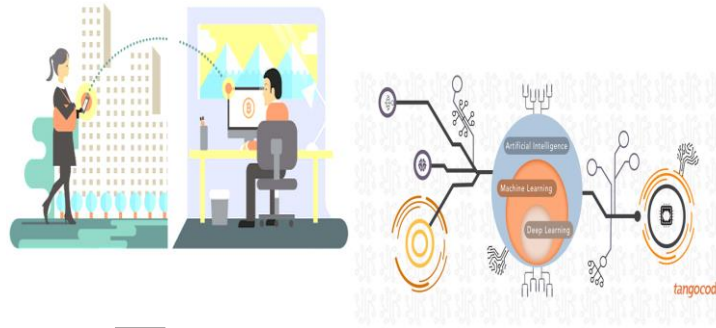
## Metodología Multivariadas YELLOWBRICK



# Explicación e Interpretabilidad Modelos Machine Learning

Conclusión

Explicación



Machine Learning



Experiencia de  
Negocio



Decisiones con  
Impacto en los  
Negocios

GRACIAS  
POR SU ATENCIÓN

