

Machine Learning para cada Sector- Industria

BIG DATA – BUSINESS INTELLIGENCE – MACHINE LEARNING



ÍNDICE

ÍNDICE	2
1. INTRODUCCIÓN.....	5
2. RETAIL	7
A. PREDICCIÓN DE VENTA.....	7
I. MODELOS ARIMA.....	7
II. HOLT WINTER.....	8
III. COMPARACIÓN.....	9
B. ESTRATEGIA DE PRECIO.....	10
I. REDES NEURONALES.....	10
II. CONJOINT	12
C. LA GESTIÓN DEL INVENTARIO.....	12
D. CLASIFICACIÓN DE CLIENTES.....	12
3. E-COMMERCE	14
A. SISTEMAS DE RECOMENDACIÓN.....	14
I. KVECINOS.....	14
B. GESTIÓN DEL INVENTARIO	15
C. OPTIMIZACIÓN DE PRECIOS	15
D. PROTECCIÓN CONTRA EL FRAUDE.....	15
I. ALGORITMOS.....	15
II. ÁRBOL DE DECISIÓN.....	15
III. RANDOM FOREST.....	16
IV. GRADIENT-BOOSTED TREES.....	17
E. RANKING DE BÚSQUEDA.....	17
F. ATENCIÓN AL CLIENTE Y AUTOSERVICIO.....	17
4. FINANCIERO	19
A. PORTFOLIO MANAGEMENT.....	19
I. SUPPORT VECTOR MACHINES (SVM).....	19
B. ALGORITHMIC TRADING	20
C. DETECCIÓN DEL FRAUDE.....	20
D. SUSCRIPCIÓN DE PRÉSTAMOS O SEGUROS	20
E. SERVICIO AL CLIENTE.....	20
F. SEGURIDAD 2.0.....	21
G. ANÁLISIS DE SENTIMIENTOS Y NOTICIAS	21
H. VENTAS O RECOMENDACIONES DE PRODUCTOS FINANCIEROS.....	22
5. INDUSTRIA.....	23

A.	MANTENIMIENTO PREDICTIVO	23
B.	MEJORA DE LA CAPACIDAD DE PRODUCCIÓN	24
C.	OPTIMIZACIÓN DE PRECIOS	25
D.	GESTIÓN DE INVENTARIO	25
E.	AUTOMATIZACIÓN	25
I.	REGRESIÓN DE MÍNIMOS CUADRADOS ORDINARIOS	25
II.	REGRESIÓN DE RIDGE	26
III.	REGRESIÓN DE LASSO (LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR)	26
IV.	AYUDA SOBRE NORMAS	27
F.	PERSONALIZACIÓN	27
I.	CHATBOTS	27
II.	USO DE DATOS EN TIEMPO REAL	27
6.	SANIDAD	29
A.	REALIZACIÓN DE PREDICCIONES	29
B.	PROCESADO DE IMAGEN	29
C.	PÚBLICA	30
D.	PRIVADA	30
I.	ASISTENTES VIRTUALES COMO AGENTE DE SALUD	30
II.	SEGUROS DE SALUD	30
7.	ADMINISTRACIÓN PÚBLICA	31
A.	LOS BOTS Y LOS CHATBOTS	31
B.	DIGITALIZACIÓN DE LA ADMINISTRACIÓN PÚBLICA	31
C.	FRAUDE FISCAL	31
I.	SUPPORT VECTOR MACHINES (SVM)	32
D.	FRAUDE EN APUESTAS EN INTERNET	32
8.	EDUCACIÓN	34
A.	CUALIFICACIÓN CIEGA	34
B.	PREDICCIÓN	34
C.	APRENDIZAJE COLABORATIVO	34
D.	REDUCIR LAS TASAS DE ABANDONO	34
9.	TURISMO	36
A.	PREDICCIÓN DE LA NACIONALIDAD DE LOS TURISTAS	36
B.	PREDICCIÓN DE CANCELACIONES	36
C.	PREDICCIÓN DE OCUPACIÓN	36
D.	PREDICCIÓN DE RETRASOS EN VUELOS	36
E.	MAXIMIZAR EFICACIA DE CAMPAÑAS PROMOCIONALES	37
10.	INTERNET DE LAS COSAS	38
A.	TELECOMUNICACIONES	38

B.	SANIDAD.....	39
C.	AGRICULTURA Y GANADERÍA	39
D.	INDUSTRIA.....	39
E.	HOGARES INTELIGENTES.....	40
F.	WEARABLES	40
G.	TRANSPORTE.....	41
H.	COMERCIO.....	41
11.	SMART CITY	42
A.	ALUMBRADO INTELIGENTE	42
B.	SMART PARKING.....	42
C.	EADMINISTRACIÓN.....	43
D.	MEDIOAMBIENTAL	43
E.	SMART BUILDINGS.....	43
F.	ALIMENTACIÓN Y OCIO.....	44
G.	SISTEMAS DE RECICLADO AUTOMÁTICO.....	44
H.	TURISMO	44
12.	CONCLUSIÓN.....	45
13.	RESUMEN.....	46
14.	BIBLIOGRAFÍA.....	50
15.	EJEMPLOS PRACTICOS.....	51
A.	NÚMERO DE TRABAJADORES.....	51
B.	PROVEEDORES.....	51
C.	ANALIZAR LA EFICACIA DE UNA CAMPAÑA PUBLICITARIA.....	51
D.	DETECCIÓN DE FRAUDE EN COMPRAR CON TARJETAS	51
E.	SEGMENTACIÓN DE AUDIENCIA.....	52
16.	SOBRE STRATEBI.....	53
17.	TECNOLOGÍAS.....	54
18.	DESCRIPCIÓN DE LAS SOLUCIONES.....	55
19.	INFORMACIÓN SOBRE STRATEBI.....	56
20.	REFERENCIAS STRATEBI.....	58
21.	EJEMPLOS DE DESARROLLOS ANALYTICS	59
22.	DEMOS ONLINE, VIDEOTUTORIALES Y MANUALES.....	62

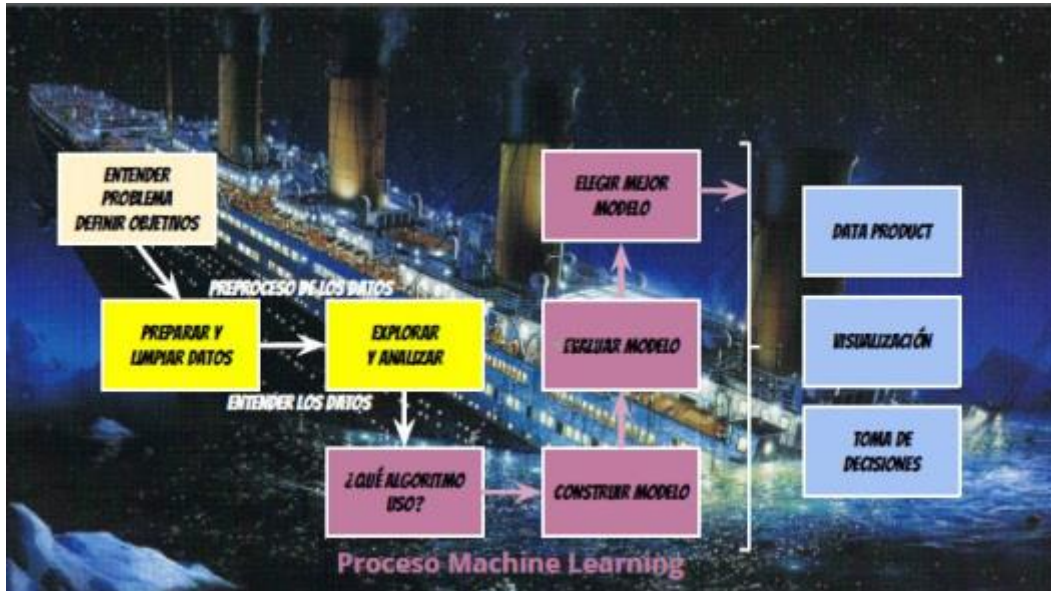
1. INTRODUCCIÓN

En este informe se van a exponer que modelos son los más adecuados, con sus ventajas e inconvenientes para resolver problemas de Retail, E-Commerce, Financieros, y otros campos, con los principales algoritmos que incluimos en todos nuestros proyectos



La idea de este documento es la facilitación para explicar a cada organización que tipo de problema pueden tener y cómo podemos ayudarles con nuestros conocimientos y aplicación de algoritmos predefinidos

Se incluye una imagen resumen bastante explícita de lo que hay que hacer para resolver un problema con machine learning



2. RETAIL

En esta sección trataremos problemas que se suelen dar en Retail, explicando el problema y analizando que algoritmos se deben usar.

RETAIL BUSINESS SOLUTIONS



a. Predicción de venta

Este problema es un clásico, aplicable a muchísimos casos, considerando el histórico de las ventas se pueden predecir las mismas y usar los datos obtenidos para ver qué cantidad de materia prima obtener, cuantos beneficios se van a obtener, etc. . . .

Voy a explicar los distintos algoritmos que se pueden utilizar en este problema y cómo funcionan.

i. Modelos ARIMA

Antes de definir lo que es un modelo ARIMA tenemos que definir una serie de conceptos que serán importantes para entender mejor que representa.

Un **proceso estocástico** es una familia $\{X_t, t \in T\}$ de variables aleatorias, cada una de ellas definida sobre el mismo espacio de probabilidad (Ω, F, P) . No voy a extenderme en la teoría, vamos a las definiciones básicas para poder entender mejor en lo que consiste un modelo Arima.

Una **realización** de un proceso estocástico es una particular función $X_t(\omega)$ de t ($\omega \in \Omega$ *fijo*).

Una **serie temporal** es un proceso estocástico donde T es un conjunto de tiempos en que se estudia el proceso, como sería en nuestros casos prácticos. Por ejemplo, $T = \{1, 2, 3, \dots\}$, $[0, \infty)$.

Diremos que una serie $\{X_t\}$ sigue un proceso **ARIMA(p,d,q)** si $\phi_p(B)\nabla^d X_t = c + \theta_q(B)Y_t$, donde el coeficiente $\phi_p(B)$ es el polinomio del operador de retardo del proceso autorregresivo de orden p, $\nabla^d = (1 - B)^d$, c es una constante, $\theta_q(B)$ es el polinomio del operador de retardo del proceso de media móvil, $BX_t = X_{t-1}$ conocido B como el operador de retardo y $\{Y_t\} \sim WN(0, \sigma^2)$.

Aún se extiende más toda esta teoría introduciendo los procesos ARIMA estacionales multiplicativos que es una combinación de un proceso ARIMA con otro ARIMA estacional en el que se añaden las diferencias estacionales.

Los modelos ARIMA analizan cómo se comportan los datos con el paso del tiempo y se ajusta a un modelo matemático que intenta asemejar la manera de actuar de los datos.

Los principales paquetes de R para el análisis de series temporales son:

- **FontS:** Funciones y datos.
- **forecast:** Es una de las librerías más potentes de R. Métodos y herramientas para visualizar y realizar pronósticos de series de tiempo univariadas usando técnicas de suavizado exponencial basado en espacio de estados y la selección automática de modelos Arima.
- **stats:** Paquete base con funciones estadísticas básicas.
- **tseries:** Funciones y datos para el análisis de series temporales y finanzas computacionales.
- **uroot:** Contrastes de raíces unitarias para series temporales con componentes periódicas.

ii. Holt Winter

Este método también es conocido como triple suavizado exponencial. Se puede decir que es un caso particular de un modelo Arima.

Se entiende también como una manera de filtrar series temporales proporcionando un filtrado. El triple suavizado exponencial se usa para series con una componente de tendencia y una componente de estacionalidad, también están el doble suavizado exponencial usado para series con tendencia, pero no estacionarias y el suavizado exponencial simple usado en series sin tendencia ni estacionalidad.

El triple suavizado exponencial es un método que aparte de retener información acerca de la tendencia con la señal suavizada y la tendencia suavizada una tercera cantidad se utiliza para

describir la estacionalidad (puede ser aditiva o multiplicativa). A continuación, voy a expresar las fórmulas y voy a señalar que representa cada parámetro:

Estacionalidad aditiva:

$$s_i = \alpha(x_i - p_{i-k}) + (1 - \alpha)(s_{i-1} + t_{i-1})$$

$$t_i = \beta(s_i - s_{i-1}) + (1 - \beta)t_{i-1}$$

$$p_i = \gamma(x_i - s_i) + (1 - \gamma)p_{i-k}$$

$$x_{i+h} = s_i + ht_i + p_{i-k+h}$$

Estacionalidad multiplicativa:

$$s_i = \alpha \frac{x_i}{p_{i-k}} + (1 - \alpha)(s_{i-1} + t_{i-1})$$

$$t_i = \beta(s_i - s_{i-1}) + (1 - \beta)t_{i-1}$$

$$p_i = \gamma \frac{x_i}{s_i} + (1 - \gamma)p_{i-k}$$

$$x_{i+h} = (s_i + ht_i)p_{i-k+h}$$

Dónde s_i es el dato suavizado, t_i es la componente de tendencia suavizada, p_i es la componente periódica suavizada, α , β y γ son parámetros constantes entre 0 y 1 que suavizan los datos, la tendencia y la periodicidad respectivamente, y x_{i+h} es el valor de nuestra predicción siendo h el número de predicción de nuestra variable.

Este método se utiliza cuando existe una clara componente de tendencia y de estacionalidad a diferencia de los modelos Arima que pueden ajustar cualquier conjunto de datos temporales. Y aquí está la diferencia, el método Holt Winter es más específico para este tipo de series temporales (estacionales y con tendencia) y el método Arima es más general.

iii. Comparación

Una de las diferencias es que Holt Winter tiene tres parámetros por lo que es más simple que un modelo Arima que tiene más parámetros (puede llegar a tener infinitos, es decir, muchos más que HW). He estado buscando y visitando distintos foros y páginas y todos llegan a lo mismo, no existe

una diferencia clara para conocer la mejor manera de predecir. Como en la mayoría de los casos (por no decir todos) se deben comparar los dos modelos y ver cuál de ellos es el mejor para cada caso.

b. Estrategia de precio

El objetivo de este problema es muy simple, averiguar qué precio poner para el producto o servicio a tratar para que maximice las ganancias, las ventas y que mejore a la empresa en la guerra de precios. Otro problema dentro de este es el de cómo puede afectar el cambio de precio de un producto a nuestros clientes.

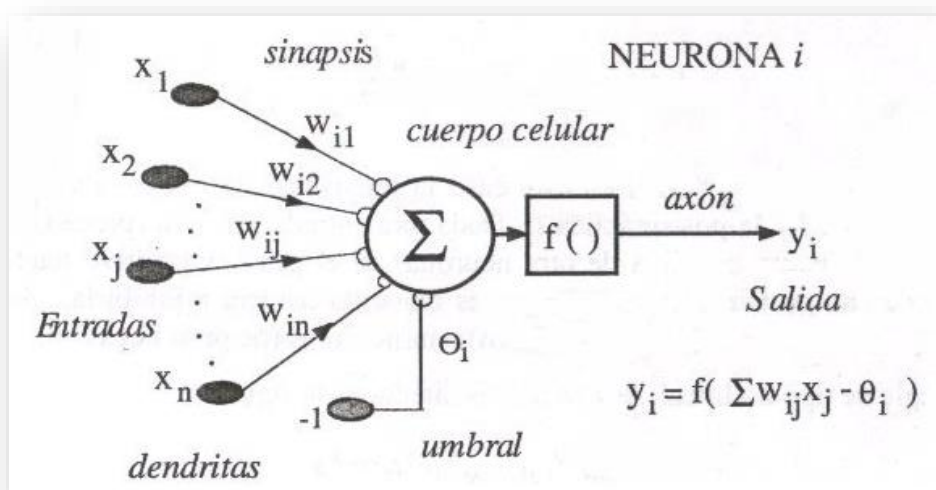
Para ello se usan distintos algoritmos:

i. Redes neuronales

Nos centraremos en las que trabajamos, las redes neuronales artificiales, aunque el nombre viene de las biológicas ya que se asemejan en comportamiento.

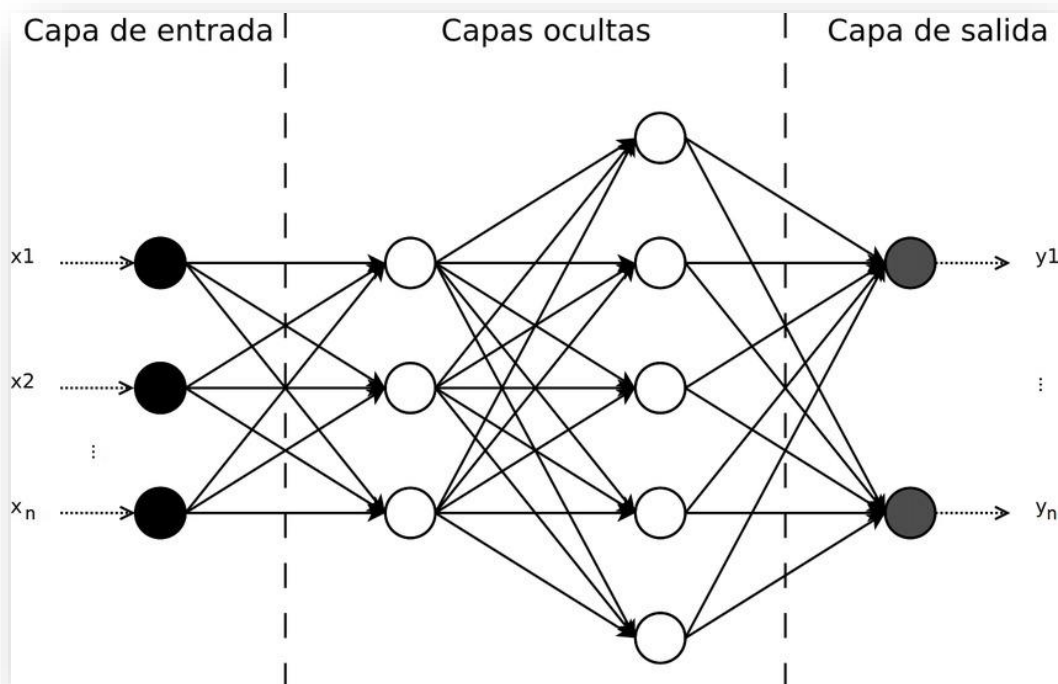
Antes de definir que es una red neuronal que como su nombre indica está formada por neuronas, vamos a definir éstas.

Existen distintos tipos de neuronas artificiales, la más común es la de tipo McCulloch-Pitts, aquí tenemos una representación gráfica:



Una neurona artificial es algo así como un procesador, se introduce un vector $x = (x_1, x_2, \dots, x_n)$ de entradas y tras procesarlo devuelve una respuesta o salida única y_i . Los principales elementos de una neurona se pueden ver en la figura y son estos:

- Las entradas que reciben los datos de otras neuronas.
- Los pesos sinápticos ($\omega_{i1}, \omega_{i2}, \dots, \omega_{in}$). A las entradas que vienen de otras neuronas se les asigna un peso, un factor de importancia. Este número se modifica en el entrenamiento de la red neuronal, y es aquí por tanto dónde se almacena la información que hará que la red sirva para un propósito u otro.
- Una regla de propagación. Con esas entradas y los pesos sinápticos, se suele hacer algún tipo de operación para obtener el valor del potencial postsináptico. Una de las operaciones más comunes es sumar las entradas, pero teniendo en cuenta la importancia de cada una. Es lo que se llama suma ponderada, aunque otras operaciones también son posibles.
- Una función de activación. El valor obtenido con la regla de propagación, se filtra a través de una función conocida como función de activación y es la que nos da la salida de la neurona. Según para lo que se desee entrenar la red neuronal, se suele escoger una función de activación u otra en ciertas neuronas de la red.



Por lo que respecta a la terminología habitual en redes neuronales artificiales los nodos del grafo representan a las neuronas y las conexiones a las sinapsis. Se denominan neuronas de entrada a aquellas neuronas sin sinapsis entrantes. A las neuronas sin sinapsis salientes se les denomina neuronas de salida, y finalmente a aquellas neuronas que no son ni de entrada ni de salida se les

llama neuronas ocultas. Y un conjunto de cada una de ellas forman las distintas capas de una red neuronal.

Y esto es básicamente en lo que consiste una red neuronal. Existen distintos tipos de ellas dependiendo si son unidireccionales o bidireccionales, si tienen una o más capas o dependiendo el tipo de aprendizaje. Pero no nos adentraremos en esto ya que es simplemente para tener una idea de cómo se comporta una red neuronal artificial.

ii. Conjoint

Es un proceso muy útil que se encarga de averiguar qué niveles de los atributos de un producto son más valorados por los clientes y que peso tienen en la decisión de compra. Consiste en realizar una encuesta a distintos clientes y analizar los resultados. A continuación dejo un enlace que ayuda mucho a entender este método: <https://www.netquest.com/blog/es/que-es-un-estudio-conjoint-y-como-utilizarlo>.

Aquí voy a nombrar unas cuantas librerías de R que trabajan con redes neuronales, que son: nnet, neuralnet y devtools. También nos permiten hacer un dibujo de las mismas.

c. La gestión del inventario

Una de las áreas de actuación en las que la Inteligencia Artificial tendrá un protagonismo cada vez mayor es en el control de inventarios a través de Machine Learning. Con una adecuada optimización del inventario se resolverán una gran cantidad de problemas logísticos que padecen las empresas con modelos más tradicionales. Además, no se debe olvidar que el foco de este modo de actuación está en la utilización de datos. Es decir, si se utilizan los datos que actualmente se poseen a nivel de compra, por ejemplo, se podrán realizar predicciones de compra en tiempo real. Al igual que conocer los productos que hay en stock en todo momento.

Como en otros modelos, se podría usar Modelos Arima o Holt Winter para la predicción.

d. Clasificación de clientes

En este y en muchos campos es interesante conocer que tipo de cliente tenemos y para ello se pueden usar diversos algoritmos de clasificación, como por ejemplo el algoritmo de los k vecinos (explicado más adelante) entre otros. Se utilizará Python con Scikit-Learn.

3. E-COMMERCE

E-COMMERCE BUSINESS SOLUTIONS



a. Sistemas de recomendación

Un sistema de recomendación es clave en cualquier sector, Retail incluido, consisten en proponer sugerencias a los clientes usando distintas fuentes de información para que sigan visitando el establecimiento o web y aumente la probabilidad de que efectúen compras.

He encontrado un trabajo de fin de grado titulado: Sistemas de Recomendación en Apache Spark, dejo el link para más información [TFG Sistemas de Recomendación en Apache Spark](#).

i. K vecinos

Es un método de clasificación que consiste en introducir cada dato nuevo en un determinado grupo según los k elementos que estén más cerca de él. Se puede entender en el método de sistemas de recomendación como agrupar los clientes según lo que hayan comprado. Y si un nuevo cliente quiere comprar determinado producto x, enseñarle otros que otro cliente haya comprado junto con el producto x.

También se podría usar otros algoritmos de clasificación, ya que la idea es la misma. Agrupar en base a los datos de clientes que ya hayan realizado alguna compra para mostrar productos a nuevos clientes que les pueda interesar. Más adelante se explican otros métodos de clasificación.

El software que se puede usar aquí es Scikit-Learn que trabaja con Python. El cuál es capaz de implementar distintos algoritmos de clasificación, en particular el que hemos mencionado.

b. Gestión del inventario

Al igual que en Retail, se podría gestionar un sistema de automatización del inventario.

c. Optimización de precios

Llegar a poder cambiar el precio teniendo en cuenta muchos factores a la vez: los precios, la demanda, la hora del día, los competidores, etc.

d. Protección contra el fraude

Las empresas de comercio electrónico son susceptibles de fraude. En algunos casos, una reputación dañada puede manchar permanentemente la reputación de una compañía.

Detectar y prevenir esto a escala es casi imposible sin la ayuda del aprendizaje automático. El Machine Learning puede procesar los datos repetitivos y tediosos rápidamente y evitar transacciones fraudulentas antes de que sucedan.

i. Algoritmos

Scikit-Learn dispone de una variedad de algoritmos de clasificación como el de los k vecinos ya nombrado o árboles de decisión. A continuación, voy a explicar más detalladamente como funciona un árbol de decisión.

ii. Árbol de decisión

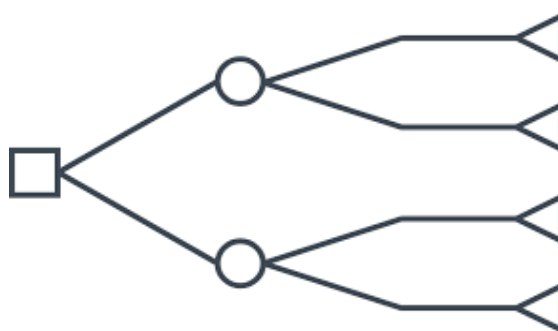
Este algoritmo realiza una partición binaria recursiva del espacio de características. La división elegida en cada nodo del árbol se elige del conjunto dónde está la ganancia de información cuando se aplica una división a un conjunto de datos. Uno de los puntos fuertes es que este método no requiere previa normalización de los datos.

Los árboles de decisión se componen de distintos elementos:

- Los **nodos** que se pueden definir como el momento en el que se ha de tomar una decisión. Existen dos tipos de nodos:
 - Los puntos de decisión, representados en la imagen por cuadrados.
 - Los nodos aleatorios que indican los posibles estados de ese punto, representados por círculos.
- Los **vectores**, son los resultados esperados, representados por triángulos en la figura.

- Las **flechas** o **ramas** que son el punto de selección entre diferentes alternativas.

Se puede observar una figura resumen de la representación de un árbol de decisión.



iii. Random Forest

Ya que hemos introducido los árboles de decisión aprovecho para indicar un algoritmo usado en regresión y clasificación que son los Random Forests (Bosques aleatorios).

Un Random Forest es un conjunto de árboles de decisión que haciendo una combinación de los mismos se reduce el riesgo de sobreajuste. Su algoritmo básico consiste en entrenar varios árboles de decisión por separado, con lo cual lo realiza en paralelo. Se incluye aleatoriedad en el proceso de entrenamiento. La combinación de las predicciones de cada árbol reduce la varianza de las predicciones, mejorando el rendimiento en los datos de entrenamiento. La aleatoriedad nombrada incluye un submuestreo del conjunto de datos para cada iteración (consiguiendo un entrenamiento diferente) y también se consideran diferentes subconjuntos aleatorios de características.

Cada árbol predice un valor real, entonces la predicción en este método de regresión es el promedio de las predicciones de los árboles. El método Random Forest es sencillo de sintonizar debido a sus dos parámetros, el número de árboles y el número de características que se seleccionarán en cada nodo y es más difícil de sobreajustar.

En cambio, el principal problema del Random Forest es que una elevada cantidad de árboles puede hacer que sea lento para predicciones en tiempo real. Para datos que incluyen variables categóricas con diferente número de niveles, los Random Forest están sesgados a favor de aquellos atributos con más niveles. Por lo tanto, los puntajes de importancia del Random Forest no son confiables para este tipo de datos. Y por último, si los datos contienen grupos de características correlacionadas de relevancia similar para el resultado, entonces los pequeños grupos se ven favorecidos en grupos más grandes.

iv. Gradient-Boosted Trees

Los árboles potenciados por gradiente (GBT) son otro método de un conjunto de árboles de decisión.

Este método entrena iterativamente una secuencia de árboles de decisión. En cada iteración, el algoritmo usa el conjunto actual para predecir la etiqueta de cada instancia de entrenamiento y luego compara la predicción con la etiqueta verdadera. El conjunto de datos se vuelve a etiquetar para poner más énfasis en las instancias de entrenamiento con predicciones pobres. Por lo tanto, en la siguiente iteración, el árbol de decisiones ayudará a corregir errores previos.

Dado que los árboles impulsados se derivan mediante la optimización de una función objetivo, este modelo se puede utilizar para resolver casi todas las funciones objetivas que podamos escribir con gradiente de salida ∇f .

Los GBT son más sensibles al sobreajuste de datos ruidosos, son más lentos debido al hecho de que los árboles se entrenan secuencialmente y son más difíciles de sintonizar debido a los parámetros número de árboles, profundidad de los mismos y la velocidad de aprendizaje.

e. Ranking de búsqueda

Este apartado es muy amplio y es conocido por el mundillo como SEO (Search Engine Optimization) que se define como el posicionamiento en buscadores u optimización de motores de búsqueda que consisten en el proceso de mejora para la visibilidad de un sitio web en los resultados de distintos buscadores.

Existen muchos factores que afectan a un motor de búsqueda, pero se podría decir que hay dos factores básicos:

- La autoridad que resumidamente es la popularidad de la web. Cuanto más se comparta un contenido significa que a más usuarios les ha sido útil.
- La relevancia es la relación que tiene una página frente a una búsqueda dada.

Este problema es muy amplio y existe una gran cantidad de información al respecto debido a que es uno de los principales objetivos de toda empresa que disponga de una página web.

f. Atención al cliente y autoservicio

Proporcionar un servicio al cliente de calidad en el comercio online es una tarea difícil. Hacerlo a escala aún es peor. Usando Machine Learning en chatbots para poder establecer una comunicación con el cliente para identificar y poder llegar a resolver un problema.

4. FINANCIERO

FINANCE BUSINESS SOLUTIONS



a. Portfolio Management

El Portfolio Management es la ciencia y el arte de tomar decisiones sobre la combinación de inversiones y las políticas, hacer coincidir las inversiones con los objetivos, la asignación de activos para individuos e instituciones y equilibrar el riesgo con el rendimiento. El Portfolio Management trata de determinar las fortalezas, debilidades, oportunidades y amenazas en la elección de la deuda frente a la equidad, nacional o internacional, crecimiento o seguridad, y muchas otras compensaciones encontradas en el intento de maximizar el rendimiento a un determinado apetito por riesgo.

Aquí adjunto un artículo con un ejemplo de cómo se puede usar aquí Machine Learning, se usa en método de Python, es un conjunto de métodos de aprendizaje supervisado usado en regresión, clasificación o valores atípicos, conocido como Support Vector Machines (SVM).

i. Support Vector Machines (SVM)

Como ventajas este método tiene efectividad en espacios de alta dimensión, sigue siendo eficaz cuando el número de dimensiones es mayor que el de muestras, es eficiente en su forma de entrenar y es versátil.

Para conocer más en profundidad este método y otros que se usan en Python, recomiendo visitar la página oficial de Scikit Learn: <http://scikit-learn.org/stable/index.html>.

b. Algorithmic Trading

La negociación algorítmica es una modalidad de operación en mercados financieros donde se usan procedimientos automáticos utilizando distintos algoritmos.

No es que exista un algoritmo en particular para este problema, sino que he encontrado que se usa Python con los potentes pandas de la biblioteca de análisis de datos y más paquetes extra de Python.

c. Detección del fraude

La tecnología de la información avanza a pasos gigantes, pero por desgracia los delincuentes también, pero con la implementación del Machine Learning en el análisis de datos para poder reconocer patrones sospechosos y poder actuar notificando a los clientes de los posibles incidentes se podría solucionar este problema.

Se puede usar cualquier algoritmo de clasificación de los que usa Python. Consisten, como su nombre bien indica, en clasificar una variable dependiendo de otras. En este caso, se clasificarían los clientes en fraude o no fraude buscando patrones de ataque.

d. Suscripción de préstamos o seguros

Muchas empresas importantes del sector están investigando el uso de algoritmos de Machine Learning para determinar si se concede o no un crédito o un seguro al cliente.

Para poder realizar esto se usarán algoritmos de clasificación por ejemplo con Python y Scikit-Learn, ya que de momento desconozco otro software que los utilice.

Antes de continuar, estos primeros cuatro problemas del campo financiero son en donde se está utilizando actualmente Machine Learning y demás, los cuatro siguientes son el posible futuro del uso de estos métodos.

e. Servicio al cliente

Este es un problema que ya no sólo afecta a este campo, sino a todo tipo de empresa. Actualmente se sigue investigando en este problema para consolidar servicios de chat digitales que tengan la capacidad de comprender, aprender y poder dar al cliente un buen servicio de venta y postventa.

f. Seguridad 2.0

Con el gran aumento de ciberataques y la enorme cantidad de malware, el usar la inteligencia artificial para luchar contra todos estos fenómenos era casi una necesidad. La automatización del malware fue probablemente el factor decisivo que forzó el empleo del aprendizaje automático en la clasificación, prevención y detección del malware, y en las acciones para solucionarlo. Y la industria de la seguridad informática no tardó en investigar, emplear y promocionar el uso del Machine Learning como respuesta a los retos actuales y seguramente futuros.

Los algoritmos de clasificación son cruciales en este problema ya que nos permiten agrupar los elementos analizados tales como ficheros, procesos, código o técnicas de ataque en benignos y malignos.

g. Análisis de sentimientos y noticias

En muchas ocasiones, cuando hablamos de reputación online, aparece el concepto de “análisis de sentimiento” pero, ¿sabemos realmente qué significa? El análisis de sentimiento se refiere a los diferentes métodos de lingüística computacional que ayudan a identificar y extraer información subjetiva del contenido existente en el mundo digital (redes sociales, foros, webs, etc.). Gracias al análisis del sentimiento, podemos ser capaces de extraer un valor tangible y directo, como puede ser determinar si un texto extraído de la red Internet contiene connotaciones positivas o negativas.

El análisis de sentimientos, también conocido como minería de opinión, se trata de una tarea de clasificación masiva de documentos de manera automática, que se centra en catalogar los documentos en función de la connotación positiva o negativa del lenguaje ocupado en el mismo.

Con las redes sociales, los usuarios tienen hoy en día todo tipo de facilidades para mostrar sus opiniones sobre cualquier tema que deseen. Tener constancia sobre las opiniones referentes a una marca o producto y medir su impacto es actualmente de vital importancia para todas las empresas, ya que es tu imagen lo que está en juego.

A toda la información que se recopila de esta forma se le denomina minería de opinión (opinión mining) y gracias a ella, las empresas tienen una inmediata disponibilidad de la información deseada. Además, la minería de opinión no solo permite responder “qué opinan los internautas sobre su propia marca o producto” sino que facilita, mediante los medios adecuados, obtener ventajas competitivas en diferentes ámbitos.

Gracias al análisis de sentimiento o minería de opinión podemos recopilar información suficiente para conocer qué piensa o qué opinan los usuarios (o target) en la red Internet.

h. Ventas o recomendaciones de productos financieros

Como en otros apartados de recomendaciones de productos dependiendo del cliente se podrían usar algoritmos de clasificación para organizar a los clientes según ciertas características para poder recomendarle productos que clientes con sus mismas características hayan adquirido.

5. INDUSTRIA

TELECOMMUNICATIONS BUSINESS SOLUTIONS



La combinación de inteligencia artificial y Machine Learning va a suponer, en poco tiempo, un cambio más grande que el que significó la llegada de Internet, a nivel de usuario. En la industria 4.0 se trata de digitalizar los procesos productivos. Precisamente en esta transformación lo que pretende es generar información de los datos digitales.

a. Mantenimiento predictivo

La industria manufacturera puede usar inteligencia artificial y Machine Learning para descubrir patrones significativos en los datos de fábrica. Las prácticas de mantenimiento correctivo y preventivo son costosas e ineficientes mientras que el mantenimiento predictivo minimiza el riesgo de fallas inesperadas y reduce la cantidad de actividades innecesarias de mantenimiento preventivo.

Para el mantenimiento predictivo, se puede construir una arquitectura Machine Learning que consta de datos históricos del dispositivo, entorno de análisis flexible, herramienta de visualización del flujo de trabajo y ciclo de retroalimentación de operaciones. Se supone que el activo tiene un patrón de degradación progresivo. Este patrón se refleja en la medición del sensor del activo. Para predecir fallas futuras, el algoritmo Machine Learning aprende la relación entre el valor del sensor y los cambios en los valores del sensor a fallos históricos. [Ejemplo de mantenimiento predictivo con la plataforma Azure ML.](#)

El éxito de los modelos de mantenimiento predictivo depende de tres componentes principales: disponer de los datos correctos, enmarcar el problema adecuadamente y evaluar las predicciones correctamente.

Al recopilar los datos es importante tener en cuenta lo siguiente:

- ¿Cuáles son los posibles fallos?
- ¿Cuáles trataremos de predecir?

- ¿Cómo es el fallo? ¿Es un proceso lento o repentino?
- ¿Qué partes de la máquina o sistema podrían estar relacionadas con cada fallo?
- ¿Cómo se puede medir su estado?
- ¿Qué frecuencia y qué precisión se deben realizar estas mediciones?

Aunque, lo que suele suceder es que los datos se han recopilado antes de que llegue el científico de datos. Ahora, dependiendo de las características del sistema y de los datos disponibles, es esencial un encuadro adecuado del modelo que se quiere construir, ¿qué pregunta queremos que responda el modelo y es posible con los datos que tenemos? A la hora de enmarcar un modelo predictivo, es importante tener en cuenta:

- ¿Qué tipo de salida debería dar el modelo?
- ¿Hay suficientes datos históricos disponibles o solo datos estáticos?
- ¿Está etiquetado cada evento registrado correctamente? O por lo menos, ¿se sabe cuándo cada máquina falló (si es que lo hizo)?
- Cuando los eventos están bien etiquetados, ¿cuál es la proporción de la cantidad de eventos fallidos y los eventos que funcionan bien?
- ¿Cuánto tiempo de anticipación debería el modelo indicar que ocurrirá el fallo?
- ¿Cuáles son los objetivos de rendimiento para los que se debe optimizar el modelo?
- ¿Cuál es la consecuencia de no predecir un fallo o predecir un fallo que no sucederá?

Con toda esta información a mano, ahora podemos decidir qué estrategia de modelado se ajusta mejor a los datos disponibles y al resultado deseado, o al menos cuál es el mejor candidato para empezar. Existen múltiples estrategias de modelado para el mantenimiento predictivo y en el siguiente enlace se explican brevemente: [Machine Learning para el mantenimiento predictivo](#).

b. Mejora de la capacidad de producción

El Machine Learning en este caso consiste en proporcionar a las máquinas de inteligencia cognitiva. Se les enseña a través de la introducción de datos históricos para predecir comportamientos futuros, de respuestas ante diferentes eventualidades... Se busca la forma de optimizar la producción a partir de la información de los datos históricos.

Por ejemplo, si tenemos informatizado el perfil de nuestros clientes (tamaño, pedidos, frecuencia, materiales, precios...), el sistema es capaz de predecir demanda, alertar de posibles bajas, fijar precios, detectar fraudes, prevenir la morosidad o identificar nuevos patrones de consumo. En este sentido, en un mundo en el que la capacidad de las empresas para personalizar su producción es un

valor en alza, el Machine Learning demuestra ser una herramienta idónea para segmentar perfiles y ofrecer la flexibilidad que exigen los pedidos a la carta.

c. Optimización de precios

Al igual que en otros muchos campos, el gran problema de optimización de precios es muy amplio y se pueden usar otros conocimientos a parte del Machine Learning como pueden ser estrategias de marketing, por ejemplo.

d. Gestión de inventario

Esta innovación no sólo impacta a nivel de producción, también a la hora de organizar el mantenimiento de las máquinas de una manera eficiente, en función de la demanda del momento, así como en la gestión del inventario. El Machine Learning es clave para optimizar el almacenamiento, y no tener stocks innecesarios y/o desabastecimiento. Para ello, es crucial tener un seguimiento automatizado, que en tiempo real nos pueda predecir falta de alguna pieza o componente, o anticipar una partida deteriorada, para así salvar rápidamente el problema y cumplir con los plazos de entrega del pedido.

e. Automatización

Las máquinas van aprendiendo a partir de las incidencias del mundo real (paradas imprevistas, pedidos urgentes, falta de personal...). A partir de ahí, muchas de las respuestas que ofrecen se pueden automatizar, con lo que no es necesaria la presencia de un operario en la planta, pudiendo éste dedicarse a otras actividades de creación de valor. Asimismo, se puede enseñar a las máquinas a identificar patrones de no calidad, reduciéndose, así, los retrabajos. Como vemos, la automatización ofrece una gran agilidad y rapidez en la producción, mientras las personas pueden poner el foco en otras áreas. La rama de automatización se centra en predecir patrones de comportamiento y para ello se podrían usar algoritmos de regresión entre otros, como, por ejemplo:

i. Regresión de mínimos cuadrados ordinarios

He recopilado información sobre estos algoritmos de la [página oficial de Scikit Learn](#). Esta regresión lineal trata el siguiente problema:

$$\min_w \|Xw - y\|_2^2$$

Dónde tenemos que:

- $X = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix}$ siendo p el número de predicciones y n el número de observaciones.
- $w = \begin{pmatrix} w_1 \\ \vdots \\ w_p \end{pmatrix}$ es el vector a minimizar
- $y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$ es el vector de los valores predichos

Este método penaliza que los términos estén correlacionados (lo cual es un gran problema pues la mayoría de los términos tienen relación por lo que podemos decir que este es el peor método) y las columnas de X tienen una dependencia lineal entonces esta estimación se vuelve sensible a error aleatorios en la respuesta observada, que se representa con una gran varianza.

ii. Regresión de Ridge

Esta regresión soluciona un problema de la regresión de mínimos cuadrados porque impone una penalización al tamaño de los coeficientes con el término $\alpha \|w\|_2^2$. A mayor valor de α los coeficientes se vuelven más robustos a la colinealidad.

$$\min_w \|Xw - y\|_2^2 + \alpha \|w\|_2^2$$

Los componentes son los mismos a la regresión de mínimos cuadrados.

iii. Regresión de Lasso (Least Absolute Shrinkage and Selection Operator)

Este es un modelo que estima coeficientes dispersos. Es útil en algunos contextos debido a su tendencia a preferir soluciones con menos valores de parámetros, reduciendo efectivamente el número de variables de las que depende la solución dada. Por esta razón, el alzo y sus variantes son fundamentales en el campo de la detección comprimida. Tiene como característica una selección de variables implícita. Es decir, que a las variables que no aportan información les pone un coeficiente de cero y en cambio el modelo Ridge asigna coeficientes a todas las variables.

$$\min_w \frac{1}{2n_{samples}} \|Xw - y\|_2^2 + \alpha \|w\|_1$$

iv. Ayuda sobre normas

Para que se entienda las distintas normas, básicamente existen distintos tipos, de hecho, se denomina **norma** a una función $\|\cdot\|: \mathbb{R}^n \rightarrow \mathbb{R}$ (se le introduce un vector de varios elementos y devuelve un número) que cumple las siguientes características:

- i. $\|x\| \geq 0 \forall x \in \mathbb{R}^n$ y, además $\|x\| = 0 \Leftrightarrow x = 0$
- ii. $\|\alpha x\| = |\alpha| \|x\| \forall \alpha \in \mathbb{R} \text{ y } \forall x \in \mathbb{R}^n$
- iii. $\|x + y\| \leq \|x\| + \|y\| \forall x, y \in \mathbb{R}^n$

Nosotros conocemos la norma más utilizada, conocida como la norma euclídea $\|x\|_2 = \sqrt{x_1^2 + \dots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2}$. Pero existen muchas más, como por ejemplo la norma l_1 conocida también como distancia de Manhattan o geometría del taxista, dada por $\|x\|_1 = |x_1| + \dots + |x_n| = \sum_{i=1}^n |x_i|$, que es la que se utiliza en la segunda parte de la regresión de Lasso.

f. Personalización

La fabricación en masa ha pasado a un segundo plano. Se imponen los pedidos a la carta y con el Machine Learning ya no será, de nuevo, necesario que una persona esté pendiente de la customización. Las máquinas serán capaces de hacerlo, con la máxima precisión y con una producción rápida. Esto supone, además, un nuevo modelo de negocio para la industria, más allá de la fabricación y venta de un producto. Es decir, la posibilidad de ofrecer servicios alrededor del artículo, como es la personalización del mismo. Todo esto se basa en mejorar la experiencia del usuario.

i. Chatbots

A la hora de personalizar un producto o servicio se podría automatizar todo este proceso con chatbots. Uno de los algoritmos que usan son las redes neuronales (explicadas en 2.2.1). Gracias a esto se podría automatizar, por ejemplo, el grabado de un texto en un cierto artículo que lo permita, añadir ciertas características específicas, en definitiva personalizar el producto o el servicio del cliente.

ii. Uso de datos en tiempo real

Esta parte es más Big Data que Machine Learning, pero como habremos llegado a la conclusión estás dos tecnologías van de la mano. Procesar datos en tiempo real y usarlos en algoritmos de regresión para predicciones para ayudar a la empresa.

6. SANIDAD

HEALTH BUSINESS SOLUTIONS



El Machine Learning es una herramienta muy importante cuando decimos que es capaz de salvar vidas, y así es, con una buena implementación de herramientas que usen aprendizaje automático los médicos tendrían un diagnóstico muy preciso de máquinas que podrían servirles de mucha ayuda. Incluso sería una herramienta para ofrecer diagnósticos en países en vías de desarrollo o zonas remotas donde no hay personal médico tan cualificado.

a. Realización de predicciones

Este tipo de técnicas, aplicadas en el ámbito de la salud, pueden utilizarse para realizar predicciones que permitan prevenir enfermedades, realizar prescripciones personalizadas, alertar y proporcionar recomendaciones de tratamiento, diagnóstico precoz a través de imagen médica, análisis de riesgos, etc.

Desde predicción de número de pacientes en función de distintos factores como pueden ser el calendario, el clima, una enfermedad que se está propagando, etc. Hasta la predicción de enfermedades por el estado actual del paciente. Algoritmos de regresión como los nombrados en 5.5 sería una opción de uso, redes neuronales 2.2.1 también tienen esta capacidad.

b. Procesado de imagen

La potencia que tiene las redes neuronales para el análisis de imágenes hace que el Machine Learning sea una herramienta crucial en la nueva era de la medicina con casos prácticos como identificar dos enfermedades graves de la retina que pueden provocar ceguera desarrollado por un equipo de científicos de EEUU. También se pueden usar estas herramientas para por ejemplo el diagnóstico de la neumonía infantil, la principal causa de muerte en niños menores de cinco años. Y, además también se puede aplicar a cáncer y muchos otros tipos de enfermedades.

Todo este problema es muy amplio es son muchas las especializaciones que hay dependiendo la enfermedad u otros casos (fetos, por ejemplo) que se va a tratar. El software Tensor Flow que se especializa en el uso de redes neuronales es muy potente para este problema por la gran capacidad de este algoritmo (explicado en 2.2.1).

c. Pública

El Machine Learning en este campo puede jugar con la ética ya que se podrían utilizar algoritmos para automatizar decisiones sanitarias y es necesario que sus resoluciones no solo sean correctas, sino también justas. Por ejemplo, un orden de preferencia de enfermedades para priorizar.

Adjunto el enlace a un trabajo fin de master relacionado con Machine Learning y la sanidad pública.
[TFM-Análisis y predicción de datos de entrada en urgencias relativos a problemas respiratorios en la ciudad de Valencia](#)

d. Privada

i. Asistentes virtuales como agente de salud

Siri o Cortana son unos ejemplos que aplicados a la salud ofrecerían la posibilidad de activar soluciones con cierta facilidad. Podría ser público o privado, pero la idea de gestionar de manera privada este servicio sería una buena idea de negocio para empresas privadas de la salud

ii. Seguros de salud

La capacidad del Machine Learning de analizar la capacidad de un cliente de contraer una enfermedad la cuál conlleve un gasto elevado de tratamiento y establecer un precio u otro dependiendo del cliente. También se juega con la ética de si aceptar o no a un cliente en una aseguradora con un alto porcentaje de contraer este tipo de enfermedades. Desde algoritmos de regresión, hasta redes neuronales para el procesamiento de imágenes y la detección de anomalías siendo éstas últimas (2.2.1) las más potentes por su gran capacidad y extensión.

7. ADMINISTRACIÓN PÚBLICA

PUBLIC ADMIN BUSINESS SOLUTIONS



a. Los bots y los chatbots

El sistema de la administración pública es bastante peculiar, pero una de las aplicaciones de Machine Learning que creo que sería más adecuada sería la implementación de bots y chatbots para procesos en este campo ya que ayudaría a automatizar y agilizar multitud de dudas y procesos.

b. Digitalización de la administración pública

Adjunto un enlace con información al respecto <https://www.revistabyte.es/actualidad-byte/iecisa-sap-retos-administracion-publica/>.

No he encontrado que tipo de problemas dentro de la administración podríamos usar machine Learning ya que uno de los mayores problemas es la digitalización de la misma. En España se aprobó dicha digitalización siendo el año 2020 el fijado para que la Administración española sea totalmente digital y entonces podremos plantearnos el posible uso de Machine Learning dentro de este campo.

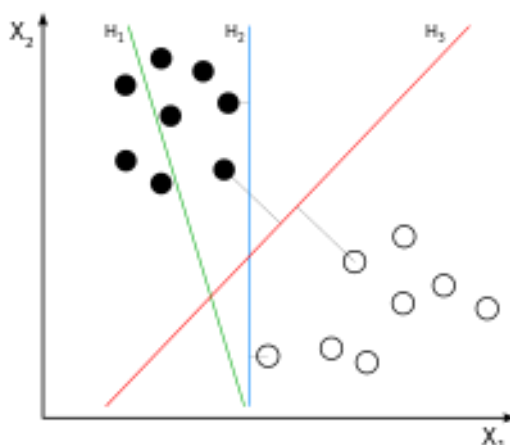
c. Fraude fiscal

Este problema consiste en detectar cualquier infracción administrativa como la evasión de impuestos y poder analizar y clasificar si esa acción administrativa es correcta o no lo es.

En este problema se usará algoritmos de clasificación como pueden ser algunos modelos de regresión ya explicados, otros a tener en cuenta son las redes neuronales que, en este caso, deberían ser más precisos que estos últimos. A continuación, voy a explicar un algoritmo que se puede usar en clasificación y en regresión, conocido como máquina de vectores de soporte (SVM), ya que los otros algoritmos nombrados ya están explicados a lo largo de este documento.

i. Support Vector Machines (SVM)

Una máquina de vectores de soporte es un modelo de machine Learning supervisado que usa un clasificador lineal binario no probabilístico para agrupar registros en un conjunto de datos. Esto significa que el algoritmo divide los datos en dos categorías con una línea clara. La línea de división se define haciendo varios hiperplanos en el espacio multidimensional (entiéndase como hiperplano como una extensión de la idea de un plano, si un plano es dibujar algo de 2 dimensiones en 3 dimensiones, un hiperplano es dibujar algo de $n-1$ dimensiones en n dimensiones). El algoritmo se encarga de seleccionar el hiperplano que separa los registros de la manera más óptima.



Las ventajas que aporta SVM es que son buenas para trabajar con sistemas multidimensionales complejos. Comparándolo con los bosques aleatorios, permite evitar el sobreajuste. Por el contrario, los modelos SVM requieren de mucho esfuerzo para afinar el algoritmo y lograr una alta precisión. Además, son muy lentas y computacionalmente pesadas y por ello requieren una fuerte arquitectura informática.

d. Fraude en apuestas en internet

Según he encontrado se estima que el 3.5% de los pagos en línea de los juegos de apuestas son fraudulentos y los ataques de bots también van en aumento por lo tanto se tiene que usar machine learning para detectar estos bots y poder cerrar esas cuentas. Los algoritmos tienen en cuenta la ubicación del jugador, su comportamiento navegando en la web, los detalles de sus transacciones, la información del dispositivo, la información histórica del pedido, pudiendo ser más extensas estas variables. Básicamente este problema consiste en analizar si una cuenta es verdadera y tomar medidas cuando dicha cuenta no lo sea, problema de clasificación.

Algoritmos de clasificación como pueden ser SVM, redes neuronales, o algunos de regresión (vistos previamente en este documento) que también sirven para clasificar serían lo que se usarían en este problema.

8. EDUCACIÓN

EDUCATION BUSINESS SOLUTIONS



En el campo educativo, especialmente en el aprendizaje en la Universidad que debe ser más competitivo e innovador, permite que los estudiantes accedan a una experiencia educativa más adaptada a sus necesidades y más flexible, según las circunstancias de cada estudiante.

a. Cualificación ciega

Si es una máquina la encargada de evaluar y calificar al alumno, se evitan tratos de favor y se garantiza que se apliquen los mismos baremos, sin excepciones. Se tendría que estudiar una forma de digitalizar y ver que algoritmos podríamos usar, en principio alguno de clasificación.

b. Predicción

Las máquinas inteligentes destacan por su capacidad analítica y predictiva, lo que permite tener una idea más fiable del rendimiento que obtendrá cada alumno y de lo que necesita para conseguir las competencias fijadas.

c. Aprendizaje colaborativo

Mediante algoritmos de clasificación se permite agrupar mejor a los estudiantes y docentes, según intereses y necesidades. Por tanto, se puede mejorar el trabajo en equipo y la cooperación a la hora de avanzar en el aprendizaje.

d. Reducir las tasas de abandono

El continuo análisis de datos y resultados ayuda a detectar a los estudiantes que están en riesgo de exclusión por su bajo rendimiento y que corren peligro de abandonar los estudios. El Machine

Learning es más hábil en la detección de estos casos, de manera que la Universidad pueda intervenir.

A continuación, dejo el enlace a un artículo muy interesante que habla sobre este tema.

<https://www.u-planner.com/es/blog/machine-learning-y-retencion-estudiantil>

9. TURISMO

TOURISM BUSINESS SOLUTIONS



a. Predicción de la nacionalidad de los turistas

Este es un caso muy real descrito en el siguiente enlace donde se muestra un vídeo explicando este caso. <https://cleverdata.io/datos-predicciones-machine-learning-real/>

b. Predicción de cancelaciones

Ver que patrones de los clientes son los que afectan a probabilidad de que un turista cancele su reserva para poder tomar decisiones a tiempo y evitar la pérdida de dinero. Existen varias opciones, árboles de decisión o, en general, algoritmos de clasificación.

c. Predicción de ocupación

Dados los datos de un hotel, se podría ver y calcular la predicción de ocupación del mismo hotel para poder estimar el número de trabajadores, etc.

d. Predicción de retrasos en vuelos

El Machine Learning puede ser capaz de predecir cuándo un vuelo se va a retrasar antes del anuncio del vuelo retrasado por la propia compañía.

Algoritmos de regresión, series temporales o redes neuronales se pueden en cualquiera de estos apartados de predicción.

e. Maximizar eficacia de campañas promocionales

La predicción de segmentos de público objetivo para cada tipo de campaña, el medio o la hora de la comunicación aumentan la eficiencia de las campañas.

10. INTERNET DE LAS COSAS

TELECOMMUNICATIONS BUSINESS SOLUTIONS



Internet of things, conocido por sus siglas, IoT es la interconexión digital de objetos cotidianos con internet permitiendo el intercambio automático de información sin intervención humana. Es la capacidad de poder automatizar distintos procesos, o aplicar la mínima intervención humana en los mismos.

IoT tiene unos retos que debe superar o tiene que superar diariamente como son la **privacidad** de esa transmisión de datos que se produce a través de sensores que está relacionado con la **seguridad** que deben de tener esos datos almacenados. Otro de estos retos es el **precio**, ya que un mismo objeto llega a valer mucho menos que si ese mismo objeto sea inteligente. Como por ejemplo una lavadora inteligente que sea capaz de conectarte a través de tu teléfono móvil a ella vale mucho más que una lavadora normal y es un inconveniente para muchas personas hoy en día. Y uno de los retos más importantes a superar es la **parte social**, concienciar que la tecnología es una solución y no es parte del problema.

Está muy relacionado con la tecnología Big Data pues la cantidad y variedad de los datos que se generan sumados a la velocidad necesaria para la lectura y el uso de dichos datos hacen de esta tecnología la idónea para ello.

a. Telecomunicaciones

Desde análisis predictivo para evitar interrupciones en la red cuando surgen datos durante eventos deportivos a monitorizar y planificar el mantenimiento de manera más efectiva y demás. Son muchas las aplicaciones de IoT, Big Data y Machine Learning en este campo. Dejo un enlace con un artículo que profundiza algo más este apartado. <https://www.digitalistmag.com/iot/2017/08/31/iot-telecommunications-industry-05323468>

b. Sanidad

Además, uno de los beneficios clave de los dispositivos IoT es que se pueden usar para habilitar sistemas remotos de monitoreo de estado y notificación de emergencias muy importante para ciertos sectores de nuestra sociedad como por ejemplo las personas mayores.

Existen proyectos enfocados a este sector como por ejemplo el servicio de dispensador de medicamentos que ayuda a los pacientes a tomar sus medicinas con capacidad de alertar automáticamente cuando es la hora de tomar medicamentos al igual que avisar a familiares o al personal médico cuando el paciente no toma su dosis medicinal.

Dentro de la Sanidad, se podría utilizar IoT de muchas maneras, con personal sanitario cualificado trabajando con personas que conozcan estas tecnologías se podrían monitorizar todo tipo de enfermedades para poder controlar la evolución de una enfermedad. Como por ejemplo en el sector de la urología existe una solución IoT de administración de fluidos que ayuda a evitar infecciones y detectar de forma muy rápida diabetes, cáncer de próstata y demás.

Cuando nombro algún tipo de predicción se pueden usar muchos tipos de algoritmos, de regresión como pueden ser los árboles de decisión, random forest, GBT pasando por redes neuronales y llegando a modelos ARIMA de series temporales.

c. Agricultura y ganadería

Básicamente a todo lo que se le pueda implementar un sensor o una forma de medir algo de manera que se pueda transformar en información digital es lo que se conoce como IoT y se le añade el adjetivo de inteligente.

Las aplicaciones en el sector de la agricultura van desde medir la humedad en el ambiente dónde se produce el cultivo hasta la implementación de los sensores que estamos nombrando en drones. Predecir si va a llover o no para la gestión automática del riego el control de cultivos, la idea de monitorizar todo el proceso de cultivo en un invernadero inteligente, el reconocimiento de plagas en cultivos para su rápida intervención para poder erradicarla

En el sector de la ganadería se puede utilizar las tecnologías de IoT para poder recolectar datos sobre el bienestar, la salud o incluso la ubicación de su ganado. Con todo esto se podría averiguar que animal porta una enfermedad para poder separarlo de la manada y así evitar la propagación de la misma

d. Industria

El objetivo del IoT en este campo es crucial para perfeccionar la producción, gestión, reparto y/o otras actividades de una empresa dentro del sector industrial. Gracias a los sensores en máquinas de

producción, vehículos de transporte y demás que permiten ver en tiempo real en que estado se encuentran o en el caso de los vehículos por dónde están circulando para optimizar el trayecto en función del tráfico y otros factores (hablaremos en el apartado 10.7 Transporte, más sobre este tema.

Gracias a todos estos dispositivos se podría mejorar el mantenimiento de las máquinas con la mejora del mantenimiento predictivo. Gracias a los sensores que permiten transmitir en tiempo real la información del estado de la máquina se podría utilizar para usar algoritmos de predicción y detectar cuál es el mejor momento para cambiar un componente de dicha máquina o realizar un saneamiento de la misma.

Otros proyectos del IoT en este campo es el de optimizar y mejorar la productividad y los procesos gracias sobre todo a la visualización de las personas a cargo del estado en tiempo real de la industria. También se pueden usar algoritmos de clasificación y clustering para poder etiquetar o agrupar respectivamente.

e. Hogares inteligentes

Este apartado también se conoce como Domótica. Como he nombrado en otros apartados la capacidad que tiene un sensor de poner transformar datos reales en datos informáticos para que se puedan analizar y trabajar con ellos hace de todos estos apartados de IoT un campo tan amplio que es cuestión de echarle imaginación.

Algunos de los ejemplos más usados es la gestión energética de la casa desde evitar el derroche de energía (por ejemplo, una app en el móvil que te indique que ventanas o puertas están abiertas para no perder temperatura en la casa por la calefacción o el aire acondicionado) hasta detectar si se tiene contratado un plan de ahorro de energía eléctrica y algunos electrodomésticos como el lavavajillas o la lavadora se enciendan en esas horas de bajo consumo. Aunque una de las aplicaciones más relevantes en mi opinión son los sistemas de seguridad que permiten conectar tu casa a tu teléfono móvil y tener control de lo que sucede.

f. Wearables

Se conoce como Wearable al conjunto de aparatos electrónicos que podemos llevar en nuestro cuerpo. Dentro de este campo están las zapatillas con GPS, los smartwatches, las pulseras que miden nuestro estado de salud, gafas inteligentes, etc.

Son muchos los algoritmos de ML que se pueden usar aquí desde cualquier tipo de algoritmos de predicción para algo que queramos saber hasta clasificación de rutas (zapatillas con gps y

smarwatches) o de tipo de enfermedades según las características medidas con las pulseras que controlan nuestro estado de salud.

g. Transporte

Instalando sensores en los vehículos de transporte y la capacidad que existe de transmitir esa información de todos los coches conectados utilizando la tecnología que viene con IoT permite optimizar la mejor ruta incluyendo variables como el tráfico, la temperatura, el clima, en fin, variables nuevas que los antiguos sistemas de navegación no tenían en cuenta.

También se puede usar esta tecnología en el transporte público para que el usuario pueda ver como va el transporte, las rutas que hace actualmente y también en el sector logístico ver de una manera más clara e incluyendo nuevas variables cuál de las rutas posibles es la que maximiza los beneficios.

Incluso se podría crear una aplicación con distintos idiomas para gestionar el transporte público a nivel usuario para que turistas o trabajadores de otros países puedan de forma sencilla llegar al destino que quieran.

Realmente si buscamos podemos encontrar un uso de cada tipo de algoritmo de Machine Learning en cada campo, realmente es ver que problema queremos resolver y pensar si se puede usar un algoritmo. En este caso desde predicción para el mantenimiento de vehículos a clasificación de rutas en buena ruta o mala ruta en función del tráfico.

h. Comercio

Muchos centros comerciales tienen implantados sistemas de renacimiento. Conectando simplemente el móvil a la red WiFi gratuita del centro comercial que te pide una serie de datos para acceder a la red como por ejemplo identificarse en Facebook, registrarse con un correo electrónico, etc. Con todo esto le permite al centro comercial enlazar esa cuenta de Facebook al teléfono móvil y saber cuándo entras, cuando sales y capacidad de analizar todos estos datos.

Existe un dicho que es digno de mencionar, "Cuando no te están vendiendo algo el producto eres tú" y también se puede entender que eres tú y todos los datos que generas. Con toda esa información se puede ver cuánto dinero se va a gastar el cliente en el centro comercial, calcular beneficios totales ese día pudiendo ser más específicos por horas, ver qué horas son las más fuertes.

Como en casi todos los problemas de los distintos campos se pueden usar la mayoría de los algoritmos, si buscamos una posible aplicación de los mismos se acabaría encontrando en todo problema.

11.SMART CITY

PUBLIC ADMIN BUSINESS SOLUTIONS



La traducción literaria de este término es ciudad inteligente, aunque realmente el que usa incluso en países de habla no inglesa es "Smart City". El objetivo principal que se proponen las Smart Cities es el de ganar calidad de vida para sus ciudadanos y visitantes de la misma.

Muchas son las aplicaciones y a continuación voy a indicar algunas que he encontrado y otras que están en desarrollo o se me ocurran para ver si así entre todos planteamos alguna buena idea para este proyecto tan actual como son estas peculiares ciudades.

a. Alumbrado inteligente

Un sistema de control de iluminación exterior inteligente e inalámbrico. Con características como la iluminación adaptativa, ayuda a ahorrar electricidad al poder detectar a personas y vehículos y llegar a regular la intensidad de la luz de forma automática. También se podría revelar el estado del sistema de alumbrado para el mantenimiento del mismo.

Ya no solo en el alumbrado dentro de la ciudad sino en polígonos industriales donde es menos habitual que haya tanto tráfico a ciertas horas y incluso en alumbrado en eventos deportivos como el que usa el Club Portugalete, de la ciudad con su mismo nombre en la provincia de Vizcaya ([Iluminación inteligente en el Club Portugalete](#)). Sistemas como este ayudarían a las ciudades y entidades a ahorro de energía y económicos significativos.

b. Smart Parking

Todos hemos visto en algún centro comercial un sistema con luces rojas y verdes dependiendo si el estacionamiento está ocupado o libre respectivamente. Pues la idea es llevar este concepto a una ciudad y a aplicaciones móvil que permitan saber dónde hay un estacionamiento libre para ahorrar a

los ciudadanos el tiempo que se invierte en encontrar aparcamiento. Incluso conocer el estado del tráfico para optimizar nuestro desplazamiento por dicha ciudad.

El apartado de transporte nombrado anteriormente se podría añadir perfectamente a las Smart Cities.

c. eAdministración

El poder automatizar todos los procesos administrativos y que todo se pueda hacer a través de internet es una idea muy innovadora y difícil de hacer, pero cada día estamos más cerca de todo esto gracias en parte al IoT. Desde notificar algún problema del estado de la vía pública, parques hasta el trato con empresas que tienen los ayuntamientos. Las opciones son muy amplias pero los medios que se necesitan para ello están creciendo ahora. Convencido estoy de que llegará ese momento.

Otra opción muy clara que veo aquí es el de una plataforma de gestión de incidencias en tiempo real que permita notificar con material fotográfico cualquier desperfecto que se produzca en la ciudad creando así una plataforma de relación más directa ciudadano-ayuntamiento.

d. Medioambiental

Todas las soluciones que se proponen en Smart City tienen como objetivo final el ahorro de energías y el comienzo de implementación de energías renovables tales como la energía solar o la eólica y que ayuden al futuro de nuestro planeta.

En las Smart Cities también se reducen las emisiones de CO2 por lo tanto bajaría el impacto de las ciudades en el medio ambiente. A través de sensores se puede medir la calidad del aire para llegar a puntos extremos en ciudades llegar a poner una solución a ello.

Otra opción es la de implementar un sistema inteligente que orqueste el riego automáticamente, dependiendo de variables como la humedad, el clima, el tiempo y demás para evaluar si es necesario y cuando el inicio y fin del riego. Como objetivo el ahorro de agua para que no llegue a desperdiciarse estando los jardines municipales con un buen estado.

e. Smart Buildings

Podemos conocerlos como edificios inteligentes que sean capaces de producir un ahorro energético y generar soluciones para poder autoabastecerse energéticamente. Esto está muy enlazado con la domótica, hablado anteriormente.

Otra opción para los edificios de las ciudades es crear un sistema de comunicación donde los edificios a parte de interactuar con sus inquilinos o visitantes, puedan recoger y mandar información a todo

tipo de dispositivo que se encuentre en la red gracias a IoT con el objetivo final de gestionar eficientemente los consumos y gastos de ellos mismos y los servicios públicos.

f. Alimentación y ocio

Una aplicación enlazada a la ciudad que te recomiende sitios dónde comprar, ir a comer, hacer ocio según tus gustos. Ver si está abierto, si existen promociones u ofertas incluso realizar una reserva en tiempo real. Quizá ya existan aplicaciones de este estilo, pero el llevarlas un paso más adelante como por ejemplo ver si hay una zona de parking cerca con hueco o ver las rutas de transporte público a seguir para llegar a la zona en sí. El poder enlazar todas las tecnologías del IoT en este campo.

g. Sistemas de reciclado automático

Esta es una idea y no sé si realmente sería posible. Instalar en cada zona de recogida de basuras un sistema que permita organizar las basuras en los distintos tipos: plástico, vidrio, orgánico y papel. Analizar las basuras y clasificarlas según distintas variables con algoritmos de clasificación. Y aquí es dónde más dudas tengo, el cómo poder analizar la basura para poder clasificarla.

De aquí me viene la siguiente idea, el instalar un clasificador en casa que sea capaz de decirte a que apartado de basura pertenece lo que quieres tirar y que tienes que hacer con el ya sea un mueble, un televisor y demás y que sea capaz de hacerlo todo automáticamente como por ejemplo llamar a una empresa de recogida de estos aparatos o indicarte dónde están situados los contenedores de reciclaje más cercanos a ti.

h. Turismo

Se podría nombrar como Smart Tourism o Turismo Inteligente. Implementar información a través de una aplicación de los monumentos o ciudades que se visitan, sitios interesantes que visitar que no sean conocidos, dónde probar platos típicos y poder reservar. Todo dentro de una misma tecnología. Aunque centrándose en la idea de sitios que visitar, se podrían agregar otras funciones como por ejemplo calendario de eventos o un chat o foro interactivo con las personas que están visitando el mismo lugar.

Ya existe una nueva tecnología de marcadores, una especie de código parecido al famoso código de barras o a los códigos QR, pero la gran diferencia entre estos nuevos marcadores llamados DDtags es la distancia a la que se puede leer, llegando hasta los 20 metros. Por lo tanto, colocando estos marcadores en museos, estatuas o monumentos de las ciudades o sitios conocidos se podría conocer en tiempo real a través de un Smartphone toda la información requerida al respecto.

12.CONCLUSIÓN

Al fin y al cabo, existen muchos algoritmos dentro de Machine Learning, y concretamente no existe un algoritmo que sea mejor en cada problema, sino que como siempre todo depende de los datos obtenidos, puede ser que para un mismo problema sea interesante usar un algoritmo para unos datos y otro distinto para otros. De ahí la generalización cuando pongo por ejemplo clasificación en el resumen ya que se podría usar un algoritmo de ese tipo.

Se puede usar distintos algoritmos de Machine Learning para cada problema, lo que nos tenemos que centrar previamente es en contestar a la pregunta: ¿cuál es nuestro objetivo?

Toda la información es un modo resumen. Dicha información explicada en el mismo se puede extender ya que existen grandes cantidades de páginas web y documentación al respecto, aunque la que hemos encontrado no explica cómo se usan los algoritmos realmente, sino que te dicen que gracias al Machine Learning todo será mejor, básicamente. A excepción de las páginas oficiales del software de uso.

13.RESUMEN

Aquí se presenta una tabla resumen con los problemas que hemos visto según sector, algoritmos o métodos y software utilizado con sus respectivos complementos del mismo. Los elementos que aparezcan vacíos no significan que no existan métodos que se puedan usar, sino que no he encontrado ejemplos o aplicaciones de esos casos o que por el contrario son distintos algoritmos los que se pueden usar en ese caso incluso varios a la vez. Con el paso del tiempo este resumen se irá mejorando.

Sector	Problema	Algoritmo	Software
Retail	Predicción de ventas	Arima	R. Librerías: FonTS, forecast, stats, tseries, uroot
		Holt Winter	R. Librerías: FonTS, forecast, stats, tseries, uroot
		Regresión	Python: Scikit-Learn y R
	Estrategia de precio	Redes neuronales	R. Librerías: nnet, neuralnet, devtools
			Tensor Flow
	Análisis conjoint		R. Librerías: conjoint
	Gestión del inventario	Arima	R. Librerías: FonTS, forecast, stats, tseries, uroot
		Holt Winter	R. Librerías: FonTS, forecast, stats, tseries, uroot
E-Commerce	Clasificación de clientes	Clustering	Python: Scikit-Learn y R
		Clasificación	Python: Scikit-Learn y R
	Sistemas de recomendación	K Neighbors (Clustering)	Python: Scikit-Learn y R
		Árbol de decisión	Python: Scikit-Learn y R
		Random Forest	Python: Scikit-Learn y R
		GBT	Python: Scikit-Learn y R
	Gestión del inventario	Arima	R. Librerías: FonTS, forecast, stats, tseries, uroot

		Holt Winter	R. Librerías: FonTS, forecast, stats, tseries, uroot
	Optimización de precios		
	Protección contra el fraude	K Neighbors	Python: Scikit-Learn y R
		Árbol de decisión	Python: Scikit-Learn y R
		Random Forest	Python: Scikit-Learn y R
		GBT	Python: Scikit-Learn y R
	Ranking de búsqueda		
	Atención al cliente y autoservicio	Redes Neuronales	R. Librerías: nnet, neuralnet, devtools Tensor Flow
Financiero	Portfolio Management	SVM	Python: Scikit-Learn y R
	Algorithmic Trading		Python
	Detección del fraude	Clasificación	Python: Scikit-Learn y R
	Suscripción de préstamos o seguros	Clasificación	Python: Scikit-Learn y R
	Servicio al cliente	Clasificación	Python: Scikit-Learn y R
	Seguridad 2.0	Clasificación	Python: Scikit-Learn y R
	Análisis de sentimientos y noticias	Clasificación	Python: Scikit-Learn y R
	Ventas o recomendaciones de productos financieros	Clasificación	Python: Scikit-Learn y R
Industria	Mantenimiento predictivo	Regresión	Python: Scikit-Learn y R
	Mejora de la capacidad de producción	Regresión y clasificación	Python: Scikit-Learn y R
	Gestión de inventario	Regresión	Python: Scikit-Learn y R
	Automatización	Regresión y clasificación	Python: Scikit-Learn y R
	Personalización		
Sanidad	Predicciones	Regresión	Python: Scikit-Learn y R
		ARIMA o Holt Winters	R. Librerías: FonTS, forecast, stats, tseries, uroot
	Procesado de imágenes	Redes Neuronales	R. Librerías: nnet, neuralnet, devtools Tensor Flow

	Sanidad Pública	Redes Neuronales	R. Librerías: nnet, neuralnet, devtools
			Tensor Flow
		Clasificación o clustering	Python: Scikit-Learn y R
	Sanidad Privada	Regresión	Python: Scikit-Learn y R
		Redes Neuronales	R. Librerías: nnet, neuralnet, devtools
			Tensor Flow
		Regresión	RStudio, Python: Scikit-Learn
		Clasificación	RStudio, Python: Scikit-Learn
Administración pública	Chatbots	Redes Neuronales	R. Librerías: nnet, neuralnet, devtools y Tensor Flow
	Digitalización		
	Fraude fiscal	SVM (Clasificación)	RStudio, Python: Scikit-Learn
	Fraude en apuestas online	Redes Neuronales	R y Tensor Flow
		Clasificación	R con RStudio, Python: Scikit-Learn
Educación	Cualificación ciega	Clasificación	RStudio, Python: Scikit-Learn
	Predicción	Regresión	RStudio, Python: Scikit-Learn
	Aprendizaje colaborativo	Clasificación	RStudio, Python: Scikit-Learn
	Reducir las tasas de abandono	Regresión	RStudio, Python: Scikit-Learn
Turismo	Predicción de la nacionalidad de los turistas	Regresión y clasificación	RStudio, Python: Scikit Learn
	Predicción de cancelación	Regresión y clasificación	RStudio, Python: Scikit Learn
	Predicción de ocupación	Regresión	RStudio, Python: Scikit Learn
	Predicción de retraso en vuelos	Regresión y clasificación	RStudio, Python: Scikit Learn
	Maximizar eficacia de campañas promocionales	Regresión	R y Python: Scikit-Learn

El apartado de IoT junto con Smart City más que centrarse en los algoritmos de Machine Learning que usan o que se pueden usar se indica que gracias a sensores de todo tipo y a tecnología Big Data y la

implementación de hacer “inteligente” todo se pueden llegar a usar algoritmos de ML que usen esos datos para predecir, clasificar y demás.

14.BIBLIOGRAFÍA

A continuación, voy a dejar enlaces a páginas web interesantes sobre Machine Learning, dónde se explican algoritmos y se comparan y demás:

- [Fraud Detection: How Machine Learning Systems Help Reveal Scams in Fintech, Healthcare, and eCommerce.](#)
- <http://scikit-learn.org/stable/>
- Durante la redacción del documento se ha dejado enlaces a páginas web.
- (Ampliable)

15.EJEMPLOS PRACTICOS

A continuación, se indican algunas ideas para poder realizar una demo a nivel formativo y ser capaz de demostrar de lo que el Machine Learning es capaz.

a. Número de trabajadores

La primera de las ideas que me se me han ocurrido es la respuesta a la pregunta: ¿Cuántos empleados debo de tener trabajando un cierto día? En una empresa de servicio o ventas (un hotel, un restaurante, un supermercado, ...), tendrán temporadas de clientes y un buen planteamiento sería el número óptimo de trabajadores en ese momento para poder realizar el trabajo de ese día, contestando esa pregunta se optimizarían los beneficios pues no tendríamos trabajadores de más.

Se necesitarían las variables: número de clientes, alguna que mida cuantos clientes puede cubrir un trabajador (o cada trabajador), fecha (se podría calcular por semanas o meses en vez de por día). Se podría dividir por sectores, por puestos distintos de trabajo

b. Proveedores

Otra de las ideas es la de, cuál proveedor me vende el mejor producto a un buen precio (viene siendo la respuesta a calidad/precio. El precio de los productos va variando y se podría realizar una demo de proveedores que disponen del producto X, viendo su precio, el tiempo de entrega, calidad, etc. Que ayuden al gerente de la pequeña empresa a decidir dónde comprar su producto X.

c. Analizar la eficacia de una campaña publicitaria

Tras haber tomado datos históricos y haber realizado distintas campañas publicitarias, estudiar el número de clientes tras esa campaña y poder llegar a la conclusión de que campaña es la más eficaz.

d. Detección de fraude en comprar con tarjetas

Partiendo de los datos de un cierto cliente sabiendo las que son buenas (transacciones realizadas por él mismo), meter esos datos en un algoritmo de predicción para futuras intervenciones (para saber si son o no aceptadas por el algoritmo) o si lo que queremos es analizar las transacciones realizadas se podría usar un algoritmo de clasificación.

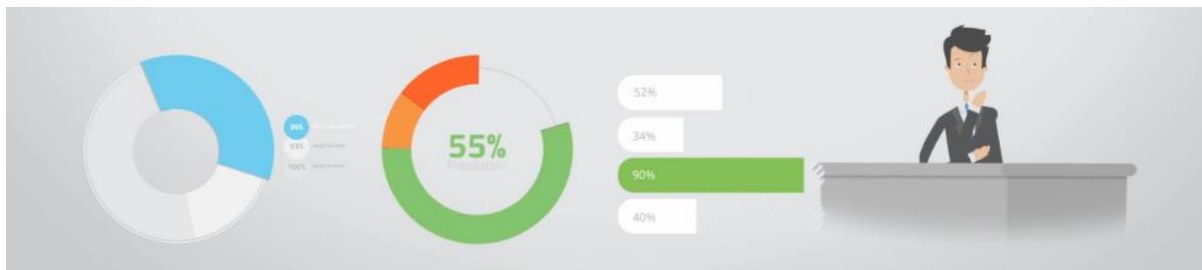
e. Segmentación de audiencia

Para podemos aplicar un algoritmo sin etiquetas (lo que se conoce como aprendizaje no supervisado) y poder observar que patrones de comportamiento tienen los clientes de una página web, una aplicación, un establecimiento y demás. Es una especie de clasificación, pero sin decidir nosotros a que grupos pertenecen. Y sería interesante aplicarla en el campo del E-Commerce ya que podríamos establecer pack de productos si a un grupo les interesan varios o simplemente ver qué características tienen relación y poder llegar a la conclusión de toma de decisiones para mejorar los beneficios que son el objetivo primordial de cada comercio. **Resumen:** agrupar datos de los cuales desconocemos sus características en común y queremos descubrirlas. **Algoritmos de clustering.**

16. SOBRE STRATEBI

En Stratebi ofrecemos **gran cantidad de soluciones analíticas** por una compañía de **rápido crecimiento**, innovando en las áreas tecnológicas de mayor desarrollo en la actualidad: **Business Intelligence, Big Data y Machine Learning**, muchas de ellas, basadas en soluciones **Open Source**.

Además, somos **Partners Certificados en Microsoft PowerBI, Talend y Vertica**, con gran número de proyectos con ambas tecnologías y los principales especialistas en **Pentaho** y muchas otras soluciones Open Source



Desarrollamos nuevas soluciones analíticas basadas en Open Source, para la generación de Cuadros de Mando en tiempo real, con tecnologías IoT para SmartCities, machine learning, etc...

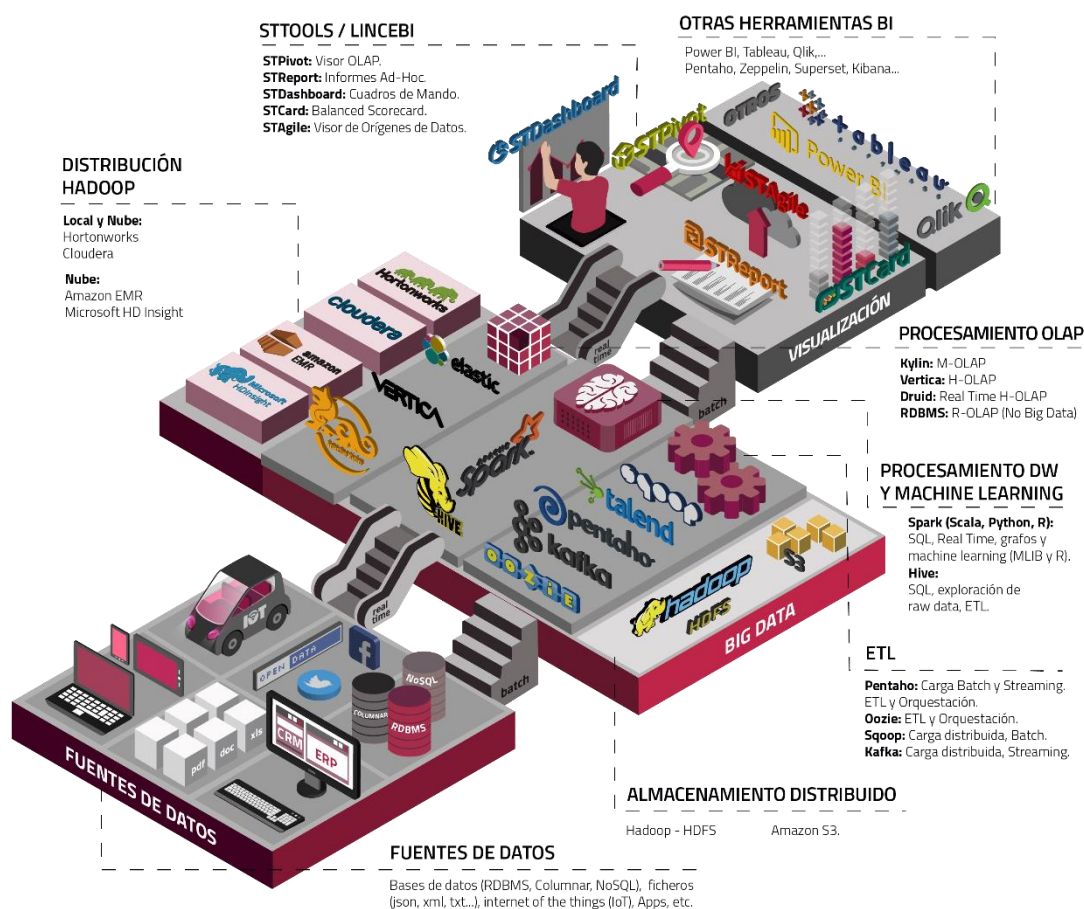


17. TECNOLOGÍAS

Recientemente, hemos sido nombrados Partners Certificados de Vertica, Talend, Hortonworks y Microsoft PowerBI



18. DESCRIPCIÓN DE LAS SOLUCIONES



Stratebi son los creadores de la solución LinceBI.com (Big Data Analytics.)



19. INFORMACIÓN SOBRE STRATEBI



[Stratebi](#) es una empresa española, con sede en Madrid y oficinas en Barcelona, Alicante y Sevilla, creada por un grupo de profesionales con amplia experiencia en sistemas de información, soluciones tecnológicas y procesos relacionados con soluciones de Open Source y de inteligencia de Negocio.

Esta experiencia, adquirida durante la participación en proyectos estratégicos en compañías de reconocido prestigio a nivel internacional, se ha puesto a disposición de nuestros clientes a través de Stratebi.

Stratebi es la única empresa española que ha estado presente en los 11 Pentaho Developers celebrados en Europa.

En Stratebi nos planteamos como **objetivo** dotar a las compañías e instituciones, de herramientas escalables y adaptadas a sus necesidades, que conformen una estrategia Business Intelligence capaz de rentabilizar la información disponible. Para ello, nos basamos en el desarrollo de soluciones de Inteligencia de Negocio, mediante tecnología Open Source.

Stratebi son **profesores y responsables de proyectos** del Master en Business Intelligence de la Universidad UOC.

Los profesionales de Stratebi son los creadores y autores del primer weblog en español sobre el mundo del Business Intelligence, Data Warehouse, CRM, Dashboards, Scorecard y Open Source.

Stratebi es partner de las principales soluciones BI Open Source: Pentaho, Talend, Jedox, SQL Power...

Todo Bi, se ha convertido en una referencia para el conocimiento y divulgación del Business Intelligence en español.



Stratebi ha sido elegida como **Caso Éxito del Observatorio de Fuentes Abiertas de Cenatic**.



Asociaciones empresariales de Software Libre empresarial en las que participamos.



20. REFERENCIAS STRATEBI

Trabajamos en los principales sectores y con algunas de las compañías y organizaciones más importantes de España y multinacionales

SECTOR PRIVADO

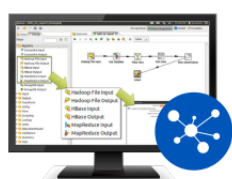


SECTOR PÚBLICO



21. EJEMPLOS DE DESARROLLOS ANALYTICS

A continuación se presentan **ejemplos de algunos screenshots** de cuadros de mando diseñados por Stratebi, con el fin de dar a conocer lo que se puede llegar a obtener, así como Demos Online en la web de Stratebi:



Data Ingestion
Manipulation
Integration



Enterprise and
Ad Hoc Reporting



Data Discovery
Visualization



Predictive
Analytics

Pentaho Analytics Platform

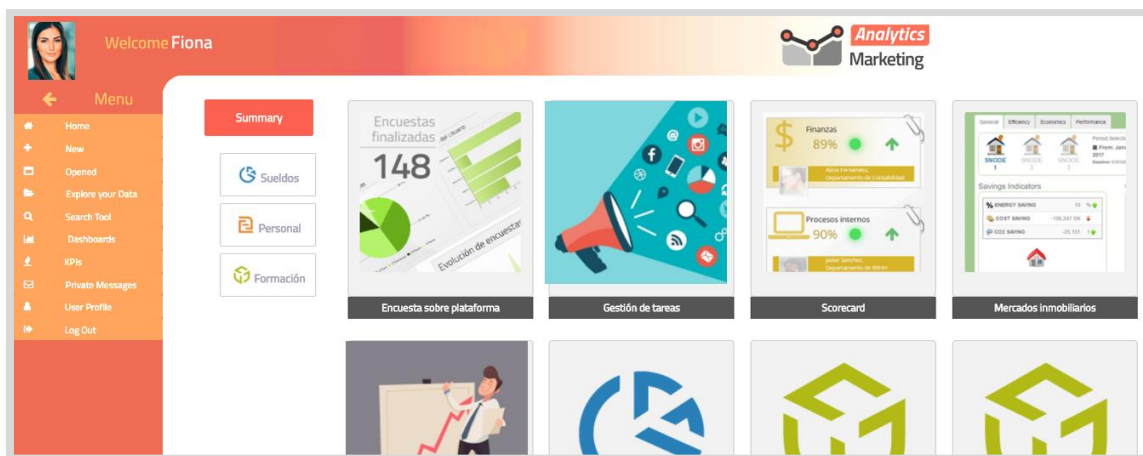
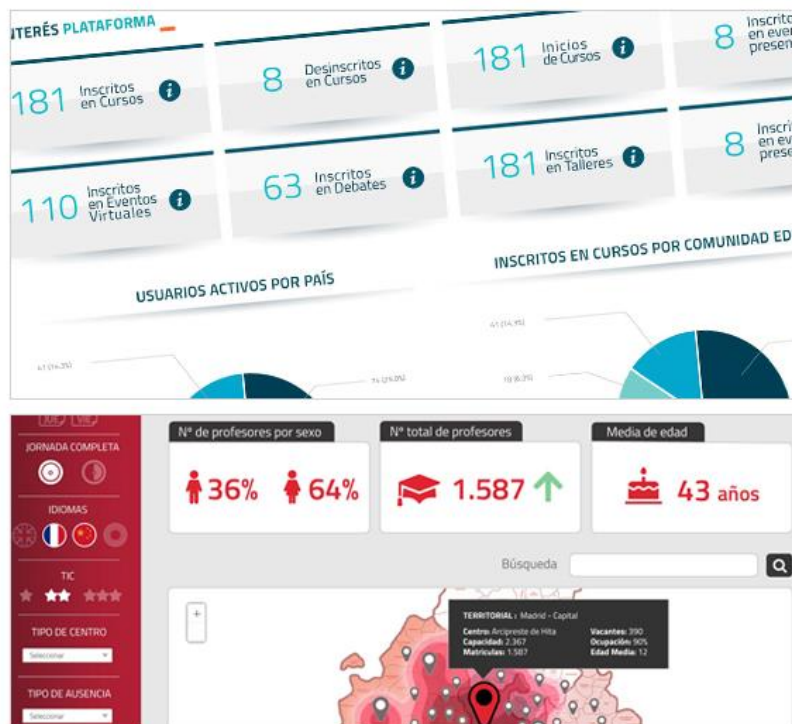
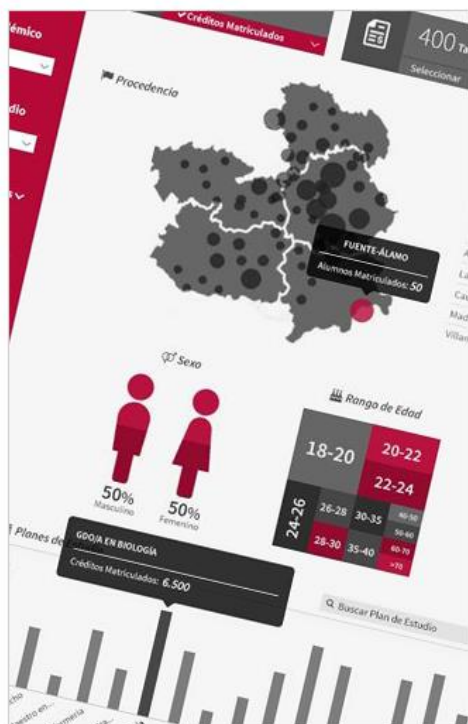
Hadoop

NoSQL

Analytic
Databases

Relational





22. DEMOS ONLINE, VIDEOTUTORIALES Y MANUALES

En este apartado incluimos una serie de recursos de utilidad para mostrar las funcionalidades y capacidades de aplicación del Machine Learning en diferentes Industrias

- Plataforma Online de Soluciones Big Data y Machine Learning ([Demos y Aplicaciones](#))
- Aplicación práctica de Machine Learning con Bases de Datos Analytics. [Descargar Paper](#)
- [VideoTutorial](#) Machine Learning para Predicción de Ventas
- VideoTutorial [PowerBI con R](#)
- Aplicación de técnicas de Data Mining para [Sentiment Analysis](#). [Whitepaper descargable](#)
- Otros VideoTutoriales:
 - o [Jupyter Notebooks con Anaconda y Python 3 Scikit Learn, Pandas, Numpy](#)
 - o [Curso de Introducción a Pentaho](#)
 - o [Introducción a Data Mining](#)
- Otras Presentaciones:
 - o [Benchmark Big Data OLAP Analytics](#)
 - o [Machine Learning con Anaconda, Jupyter y Python](#)
 - o [Introducción al Machine Learning](#)
 - o [53 Claves para conocer Machine Learning](#)
 - o [Introducción al Data Mining](#)

Artículos interesantes:

[DIFERENCIAS ENTRE BUSINESS INTELLIGENCE Y DATA SCIENCE \(BIEN EXPLICADO\)](#)

[CURSO DE MACHINE LEARNING EN MADRID\(\)](#)

[GLOSARIO DE INTELIGENCIA ARTIFICIAL \(AI\)](#)

[MACHINE LEARNING DEMO TUTORIAL](#)

[COMO HACER DEEP LEARNING CON PENTAHO](#)

[COMO PENTAHO AYUDA A LOS CIENTIFICOS DE DATOS](#)

[LOS MEJORES POSTS DE TODOBI SOBRE MACHINE LEARNING Y BIG DATA DURANTE EL VERANO](#)

[VIDEOTUTORIAL: USANDO R PARA MACHINE LEARNING CON POWERBI](#)

[CUAL ES EL ORDEN CORRECTO PARA HACER UN PROYECTO DE MACHINE LEARNING?](#)

[EJERCICIO PRACTICO DE MACHINE LEARNING CON JUPYTER NOTEBOOKS, ANACONDA Y PYTHON 3](#)

[AI, MACHINE AND DEEP LEARNING TIMELINE](#)

[FORMACION DATA ENGINEER: MACHINE LEARNING CON SPARK](#)

[APRENDE MACHINE LEARNING CON FACEBOOK](#)

[MACHINE LEARNING CON PENTAHO](#)

[THE TIMELINE OF STATISTICS \(MACHINE LEARNING\)](#)

[LA TABLA PERIODICA INTERACTIVA DEL MACHINE LEARNING](#)

[LAS MEJORES APIS PARA MACHINE LEARNING](#)

[LAS 20 MEJORES LIBRERIAS PARA DATA SCIENTISTS EN R](#)

[CURSO EN MADRID DE MACHINE LEARNING](#)

[RSTUDIO PAPERS: LISTOS PARA DESCARGAR](#)