



ESCUELA TÉCNICA SUPERIOR DE INGENIEROS
INFORMÁTICOS

UNIVERSIDAD POLITÉCNICA DE MADRID

Aprendizaje automático para flujos de datos

TRABAJO FIN DE MÁSTER
MÁSTER UNIVERSITARIO EN INTELIGENCIA ARTIFICIAL

AUTOR: Javier Ramos Fernández
TUTOR/ES: María Concepción Bielza Lozoya y
Pedro María Larrañaga Múgica

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Metodología	3
1.4. Organización de la memoria	4
2. Aprendizaje automático	5
2.1. Notación	6
2.2. Algoritmos de aprendizaje supervisado	6
2.2.1. Clasificadores bayesianos	6
2.2.2. Inducción de reglas	6
2.2.3. Árboles de decisión	7
2.2.4. Redes neuronales	7
2.2.5. k -Vecinos más cercanos	7
2.2.6. Máquinas de vector soporte	8
2.2.7. Regresión logística	8
2.2.8. Métodos combinados de aprendizaje	9
2.3. Algoritmos de aprendizaje no supervisado	10
2.3.1. Agrupamiento	10
2.4. Series temporales	11
2.5. Redes bayesianas para el descubrimiento de conocimiento	11
3. Estado del arte: Aprendizaje automático para flujos de datos	15
3.1. Introducción	15
3.1.1. Conceptos	15
3.2. Algoritmos de aprendizaje supervisado	16
3.2.1. Clasificadores bayesianos	16
3.2.2. Inducción de reglas	16
3.2.3. Árboles de decisión	16
3.2.4. Redes neuronales	17
3.2.5. k -Vecinos más cercanos	17
3.2.6. Máquinas de vector soporte	18
3.2.7. Regresión logística	18
3.2.8. Métodos combinados de aprendizaje	19
3.3. Algoritmos de aprendizaje no supervisado	19

3.3.1. Agrupamiento	20
3.4. Series temporales	20
3.5. Redes bayesianas para el descubrimiento de conocimiento	21
4. Conclusiones y líneas futuras de trabajo	23
5. Sobre las referencias	25
A. Anexos	29
Bibliografía	31

Índice de figuras

2.1. SVM Unagar and Unagar (2017)	9
---	---

Índice de tablas

2.1. Notación utilizada en el trabajo	6
---	---

Índice de algoritmos

Capítulo 1

Introducción

1.1. Motivación

Actualmente vivimos en la era de la información, una época en la que se está generando una cantidad ingente de información. Esta abundancia de datos se debe principalmente a la aparición de ordenadores y de otros dispositivos que son capaces de recoger información de lo que nos rodea, procesarla y transmitirla, como son los teléfonos móviles. Además, su capacidad de conexión a Internet hace que haya una generación de tal cantidad de información que es imposible tener un control absoluto de todo lo que circula por esta red informática de nivel mundial.

Este aumento exponencial del volumen y variedad de información ha generado la necesidad de llevar a cabo un almacenamiento masivo de los datos, y con ello el interés por analizar, interpretar y extraer información útil de los mismos con el objetivo de obtener conocimiento. Para manejar toda esta información, los sistemas tradicionales de almacenamiento de datos no son convenientes puesto que no tienen la capacidad necesaria para su correcto procesamiento. El volumen, la variedad y la velocidad de los grandes datos causan inconvenientes de rendimiento cuando se utilizan técnicas tradicionales de procesamiento de datos Mohanty et al. (2015). Por ello surge lo que se denomina **Big Data**, un concepto relativamente nuevo que se refiere a conjuntos de datos cuyo tamaño va más allá de la capacidad de las herramientas típicas de software de bases de datos para almacenar, gestionar y analizar datos Oguntimilehin and Ademola (2014).

Los datos son la materia prima para conseguir información provechosa, que se puede utilizar para llevar a cabo una toma de decisiones y la realización de conclusiones, por lo que se han desarrollado nuevas herramientas que sobrepasan las herramientas disponibles anteriormente para tratar este tipo de volúmenes de datos. De esta manera, surge el concepto de **minería de datos**, que se define como el proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos. Es decir, la tarea primordial de la minería de datos es encontrar modelos inteligibles a partir de los datos que posibiliten el hallazgo de aspectos previamente desconocidos de los mismos.

Para ejecutar el proceso de extracción del conocimiento, una de las posibilidades más populares es la aplicación de una rama de la inteligencia artificial denominada **aprendizaje automático**, la ciencia (y el arte) de programar computadoras para que puedan aprender de los datos Gron (2017). En este caso, el objetivo es que los ordenadores aprendan automáticamente sin intervención humana. Este proceso de aprendizaje se realiza proporcionándoles a los algoritmos pertinentes una serie de datos sobre los que se entrenan con el objetivo de buscar patrones en los mismos y llevar a cabo mejores decisiones en el futuro.

En general, los algoritmos de aprendizaje automático asumen que los datos están disponibles a la hora de llevar a cabo el entrenamiento y que se generan a partir de una distribución estática. No obstante, la información presente hoy en día circula en un entorno que cambia de manera *continua y rápida*, de manera que la distribución que genera los datos puede sufrir transformaciones. Todo esto ha propiciado la aparición de lo que se denominan **flujos de datos** (*data streams*), que son secuencias continuas y ordenadas de datos en tiempo real.

A la hora de tratar con flujos de datos, los algoritmos de aprendizaje automático tradicionales no son capaces de funcionar correctamente puesto que el volumen de los mismos puede llegar a ser *infinito*. Para que puedan realizar el entrenamiento, el conjunto de datos debe estar almacenado en memoria, y solo pueden llevar a cabo tareas de predicción cuando la fase de entrenamiento haya finalizado. Sin embargo, el almacenamiento de datos generados continuamente se hace inviable. Además, para que los algoritmos puedan manejar flujos de datos cuya distribución que los subyace puede cambiar es necesario un *procesamiento en tiempo real*, característica que los algoritmos tradicionales no poseen.

En este trabajo pretendemos realizar una revisión de literatura sobre métodos propuestos para aplicar aprendizaje automático en flujos de datos con el objetivo de tener una visión global de las diferentes posibilidades existentes para resolver el problema planteado anteriormente. Para ello, vamos a tener en cuenta diferentes algoritmos de aprendizaje automático existentes en el mercado actual.

1.2. Objetivos

El objetivo principal de este trabajo es realizar una **comparativa amplia de las diferentes aproximaciones propuestas para resolver el problema de aprendizaje automático para flujos de datos, principalmente tanto supervisado como no supervisado**. Para abordar esta meta, nos centraremos en los algoritmos de aprendizaje automático más populares. En el caso de aprendizaje supervisado, nos enfocaremos en propuestas relacionadas con *árboles de decisión*, *redes bayesianas*, *redes neuronales*, *inducción de reglas*, *vecinos más cercanos*, *máquinas de soporte vectorial* y *métodos combinados de aprendizaje* (ensemble). Con respecto al aprendizaje no supervisado, consideraremos primordialmente los algoritmos de *agrupamiento* (clustering), uno de los principales de este tipo de aprendizaje.

También se hará hincapié en algunas propuestas relacionadas con aprendizaje

semisupervisado

Para alcanzar este objetivo, se han contemplado una serie de metas específicas dentro del proyecto:

- *Búsqueda de artículos relacionados con cada uno de los algoritmos de aprendizaje automático contemplados.* Con el objetivo de llevar a cabo una comparativa de los diferentes métodos propuestos de aprendizaje automático para flujos de datos, es de vital importancia realizar una búsqueda de diferentes artículos propuestos para cada uno de los algoritmos de aprendizaje automático mencionados anteriormente. Para ello hemos utilizado diferentes revistas que aborden esta rama de la inteligencia artificial, así como otros recursos como *Google Académico*.
- *Anotación de aspectos claves de los diferentes artículos encontrados.* Con ello se pretende tener disponible información resumida de las diferentes propuestas que se encuentran a nuestra disposición y utilizarla para compararlas con otros artículos con la finalidad de dar una perspectiva general de la utilidad de las diferentes aproximaciones existentes.
- *Estructuración de la revisión de la literatura de algoritmos de aprendizaje automático para flujos de datos en función de los algoritmos en los que se centren las diferentes propuestas encontradas.* Se persigue comparar los diferentes artículos hallados según el algoritmo de aprendizaje automático que aborden para que funcionen correctamente con flujos de datos. Para ello hemos realizado una división de los mismos en diferentes apartados.
- *Exploración de las diferentes revisiones halladas sobre algoritmos de aprendizaje automático para flujos de datos.* El objetivo de esto es conocer qué propuestas de las que hemos encontrado se encuentran referenciadas en esas revisiones para tener en cuenta qué artículos son novedosos con respecto a dichas revisiones. Además, se lleva a cabo esta exploración para tener conocimiento de que algoritmos de aprendizaje automático tratan con el fin de aportar nuevos algoritmos. Todo esto se realiza con la finalidad de establecer qué características nos diferencian de las revisiones encontradas.

1.3. Metodología

Para desarrollar este proyecto se ha seguido una metodología que permitiera sobrellevar las dificultades del mismo de la mejor forma posible. A continuación se enumeran los pasos ejecutados durante este proceso:

- En primer lugar, hemos procedido a realizar una búsqueda **exhaustiva de artículos relacionados con aprendizaje automático para flujos de datos**. Antes de realizar una comparación entre las diferentes propuestas, es necesario recabar la mayor cantidad de aproximaciones desarrolladas con el fin de

ampliar nuestra visión genérica del estado del arte del tema abordado. Para lograr esto, hemos indagado en numerosas revistas que incluyen dentro de su temática el aprendizaje automático para flujos de datos. Algunas que se han consultado y que son relevantes en el mundo académico son *Journal of Machine Learning Research*, *IEEE Transactions on Knowledge and Data Engineering* y *Machine Learning Journal*. También se han explorado diferentes editoriales como *Elsevier* y *Springer* en las que, aparte de artículos publicados en revistas, también aparecen propuestas de conferencias. De forma complementaria, hemos buscado artículos en *Google Académico*, una herramienta de búsqueda de Google que permite hallar literatura del mundo científico de diferentes recursos (bibliotecas, editoriales, etcétera).

- Tras aglomerar una cantidad aceptable de artículos, hemos iniciado la **lectura de los mismos**. Durante este proceso, hemos ido apuntando características relevantes de las propuestas con el fin de poder tener la información fundamental para realizar la comparativa entre diferentes artículos. Para guardar dicha información, hemos creado un documento en el que, por cada propuesta, se apunta el *título*, la *fecha de publicación* y *contenido de interés* de las mismas, así como *comparaciones con otras aproximaciones* que se encontraba dentro de los artículos para utilizarlas en el estado del arte. Asimismo, hemos añadido información adicional con respecto a la *presencia o no de los artículos de las diferentes revisiones* encontradas que abordan el tema de aprendizaje automático para flujos de datos con el objetivo de tener conocimiento sobre qué artículos de los que hemos encontrado aportan nueva información con respecto a dichas revisiones para establecer características que nos diferencian de las revisiones halladas. Estas propuestas se han ordenado por *fecha de publicación* para contemplar la evolución cronológica de las mismas.
- De forma paralela a la lectura de artículos, hemos procedido a **buscar más propuestas**, enfocándonos en encontrar aquellas que son más recientes para aportar más información que nos diferencie de las revisiones encontradas. También hemos aprovechado las referencias a otros papers que hemos hallado en los artículos revisados para añadirlos al estado del arte.
- Por otra parte, se ha creado un documento en el que se ha ido **almacenando información útil durante la lectura de los artículos para relacionar diferentes propuestas de la mejor manera posible**. Por ejemplo, hemos anotado diferentes categorías en las que se podrían clasificar las distintas propuestas.

1.4. Organización de la memoria

Pendiente

Capítulo 2

Aprendizaje automático

Arthur Samuel, uno de los pioneros del aprendizaje automático, estableció en 1959 una definición general de esta rama de la inteligencia artificial:

El aprendizaje automático es el campo de estudio que da a las computadoras la capacidad de aprender sin estar programadas explícitamente.

Tom Mitchell, otro investigador de aprendizaje automático reputado, propuso en 1997 una definición más precisa y más orientado a la ingeniería:

Se dice que un programa de ordenador aprende de la experiencia E con respecto a alguna tarea T y alguna medida de rendimiento P , si su rendimiento en T , medido por P , mejora con la experiencia E .

Por lo tanto, el aprendizaje automático se centra en aplicar sistemáticamente algoritmos para sintetizar las relaciones subyacentes de forma automática en un conjunto de datos proporcionados en forma de ejemplos a través de una fase de entrenamiento, de tal forma que en el futuro se utilice esta información para la ejecución de predicciones de eventos desconocidos y una mejor toma de decisiones. Los campos de aplicación que existen de esta rama de la inteligencia artificial son muy variados. Algunos de ellos son la predicción bursátil, predicción meteorológica, detección de correos spam, construcción de sistemas de recomendación y detección de fraude en el uso de tarjetas de crédito.

Según el propósito que persigan los algoritmos de aprendizaje automático, éstos se clasifican en dos categorías principales: **aprendizaje supervisado** y **aprendizaje no supervisado**. En el aprendizaje supervisado se engloban aquellos algoritmos que buscan aprender una *función de mapeo* entre una serie de características de entrada (variables predictoras) y una variable de salida (variable clase) mediante la ejecución de una fase de entrenamiento en la que se utilizan *datos de entrenamiento etiquetados* (valor de la variable clase conocida), de tal forma que se utiliza esta función para predecir el valor de la variable de salida a partir de los valores de las variables predictoras. En cambio, el aprendizaje no supervisado la finalidad es aprender una *función que modele la estructura o distribución subyacente en los datos*

a partir de datos de entrenamiento no etiquetados, de manera que se lleva a cabo una exploración de los datos sólo conociendo los valores de las variables predictoras.

Existe además otra categoría en la que se pueden incluir las técnicas de aprendizaje automático denominada **aprendizaje semisupervisado**. Los algoritmos que llevan a cabo este tipo de aprendizaje entrenan sobre un conjunto de datos parcialmente etiquetados, generalmente muchos datos sin etiquetar y una pequeña parte de datos etiquetados. Debido a que para la fase de entrenamiento se utilizan tanto datos etiquetados como no etiquetados, se considera que esta categoría se sitúa entre el aprendizaje supervisado y el no supervisado. En este trabajo se van a abordar algunas propuestas relacionadas con el aprendizaje semisupervisado para flujos de datos; no obstante, la contribución predominante de algoritmos de aprendizaje automático para flujos de datos que se va a realizar en este proyecto proviene de las categorías de aprendizaje supervisado y no supervisado.

2.1. Notación

Tabla 2.1: Notación utilizada en el trabajo

Símbolo	Explicación
X_i	Variable predictora i
x_i	Valor de la variable predictora i
C	Variable clase
N	Número de instancias de un conjunto de datos
x	Instancia

2.2. Algoritmos de aprendizaje supervisado

Improving Hoeffding Trees (2007)

2.2.1. Clasificadores bayesianos

Aprendizaje discriminativo de clasificadores Bayesianos Discrete Bayesian Network Classifiers: A Survey

2.2.2. Inducción de reglas

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

2.2.3. Árboles de decisión

El aprendizaje mediante **árboles de decisión** es un método de aproximación de una función objetivo de valores discretos en el cual la función objetivo es representada mediante un árbol de decisión. En estas estructuras de árbol, las *hojas* representan etiquetas de clase y las *ramas* representan las conjunciones de características que conducen a esas etiquetas de clase.

2.2.4. Redes neuronales

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

2.2.5. k -Vecinos más cercanos

El paradigma clasificatorio denominado **k -Vecinos más Cercanos** (*K-Nearest Neighbours*, K-NN) se fundamenta en la idea de identificar el grupo de k objetos en el conjunto de datos de entrenamiento que más cerca está de un nuevo objeto a clasificar, de manera que a este nuevo caso se le asigna la etiqueta de clase más frecuente en ese grupo de k objetos. Este método de clasificación supervisada se basa en tres componentes principales: un *conjunto de datos de entrenamiento etiquetados*, una *métrica de distancia* para calcular las distancias entre distintos objetos y el

número k de vecinos más cercanos. De esta forma, utilizando la métrica de distancia correspondiente, se calcula la distancia entre el nuevo caso a clasificar y los casos etiquetados para averiguar cuáles son las k instancias etiquetas que más próximas están al nuevo objeto y, una vez ejecutado este paso, se calcula la clase más repetida en el grupo de k casos más cercanos y se le asigna al nuevo objeto.

El algoritmo KNN lleva a cabo un **aprendizaje vago** (*lazy learning*) puesto que no aprende un modelo discriminativo de los datos de entrenamiento (*eager learning*), sino que memoriza el conjunto de datos de entrenamiento y la labor de predicción se realiza cuando llega un nuevo caso a clasificar. Por otra parte, es un algoritmo **no paramétrico** puesto que no hace ninguna suposición sobre la distribución de datos subyacente en el conjunto de datos de entrenamiento.

En la siguiente figura se presenta un pseudocódigo para el clasificador KNN básico:

COMIENZO

Entrada

Larranaga et al. (n.d.)

:

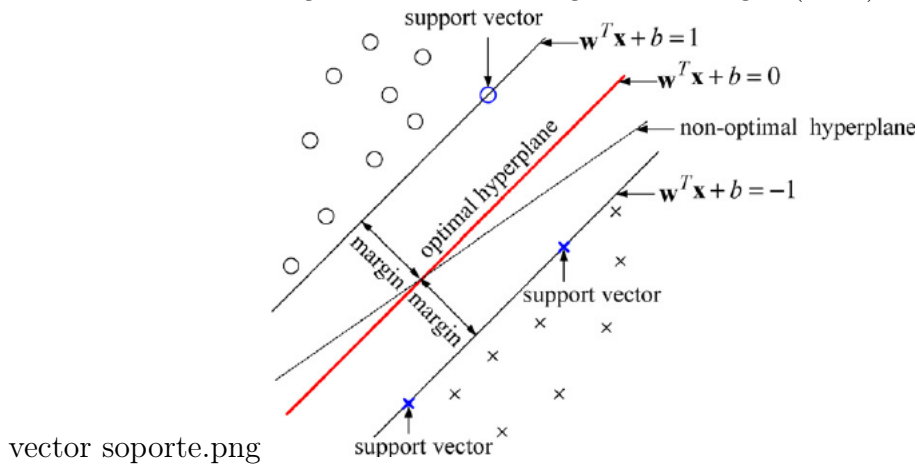
2.2.6. Máquinas de vector soporte

Las **máquinas de vector soporte** son un algoritmo de aprendizaje supervisado que puede utilizarse tanto para desafíos de clasificación como de regresión; no obstante, se utiliza principalmente en problemas de clasificación. En este algoritmo trazamos cada elemento de datos (cada muestra de entrenamiento) como un punto en el espacio n -dimensional (donde n es el número de características del problema a resolver), con el valor de cada característica mapeándose al valor de una determinada coordenada. A continuación, dado este conjunto de ejemplos de entrenamiento, cada uno perteneciente a una clase, entrenamos una SVM para construir un modelo que prediga la clase de una nueva muestra mediante la construcción de un **hiperplano** que separe las clases de los datos de entrenamiento y maximice el margen entre esas clases (maximice la distancia entre los puntos más cercanos de cada clase al hiperplano de separación óptimo, llamados vectores soporte) en el espacio n -dimensional:

2.2.7. Regresión logística

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Figura 2.1: SVM Unagar and Unagar (2017)



Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

2.2.8. Métodos combinados de aprendizaje

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

2.3. Algoritmos de aprendizaje no supervisado

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

2.3.1. Agrupamiento

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui.

Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

2.4. Series temporales

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

2.5. Redes bayesianas para el descubrimiento de conocimiento

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et

lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Durante la elaboración de la memoria del Trabajo Fin de Máster, el alumno deberá incluir toda la información que considere necesaria y útil para la descripción y justificación del trabajo desarrollado y de los resultados obtenidos. Además y con objeto de asegurar que ha cubierto todas las competencias transversales asociadas al mismo, es obligatorio que en los diferentes apartados de la memoria se desarrollen cuidadosamente los aspectos indicados a continuación.

Introducción y Objetivos: Justificar la necesidad de desarrollar la tesis en lugar de adquirir o aplicar directamente lo existente a lo largo de la gama de categorías de procesos, productos y servicios de la empresa o institución usuaria de la tesis.

Estado del arte: Demostrar que se ha comprendido como es el ámbito de negocio donde se enmarca la tesis, sus hábitos y necesidades de productos o servicios tecnológicos.

Evaluación de Riesgos: Demostrar que se han considerado diferentes soluciones tanto clásicas como novedosas o innovadoras al problema, y se han evaluado los riesgos y ventajas de cada una de ellas.

Resultados: Justificar que la tecnología resultante de la tesis satisface los deseos o necesidades del cliente (real, potencial o ficticio).

Conclusiones: Establecer las conclusiones del trabajo apoyándose fundamentalmente en los datos y observaciones obtenidas durante su desarrollo. Discutir que medios, cauces, etapas y tecnologías harían falta (si procede) para llevar a cabo una implantación real de los resultados.

Líneas futuras: Discutir los límites de las tecnologías actuales aplicadas al problema, planteando líneas de I+D+i realistas y capaces de superarlos.

Capítulo 3

Estado del arte: Aprendizaje automático para flujos de datos

3.1. Introducción

<http://i2pc.es/coss/Docencia/SignalProcessingReviews/Gaber2012.pdf> Cost-Sensitive Perceptron Decision Trees for Imbalanced Drifting Data Streams Incremental Decision Tree based on order statistics Learning Decision Trees from Data Streams with Concept Drift Fast Feature Selection for Naive Bayes Classification in Data Stream Mining Making Data Stream Classification Tree-based Ensembles Lighter (2018)

3.1.1. Conceptos

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2. Algoritmos de aprendizaje supervisado

A review of supervised machine learning algorithms

3.2.1. Clasificadores bayesianos

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2.2. Inducción de reglas

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2.3. Árboles de decisión

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac

habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2.4. Redes neuronales

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2.5. k -Vecinos más cercanos

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum

porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2.6. Máquinas de vector soporte

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2.7. Regresión logística

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.2.8. Métodos combinados de aprendizaje

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.3. Algoritmos de aprendizaje no supervisado

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque

tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

3.3.1. Agrupamiento

Quisque ullamcorper placerat ipsum. Cras nibh. Morbi vel justo vitae lacus tincidunt ultrices. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis. Nunc elementum fermentum wisi. Aenean placerat. Ut imperdiet, enim sed gravida sollicitudin, felis odio placerat quam, ac pulvinar elit purus eget enim. Nunc vitae tortor. Proin tempus nibh sit amet nisl. Vivamus quis tortor vitae risus porta vehicula.

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetur.

3.4. Series temporales

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellen-

tesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

3.5. Redes bayesianas para el descubrimiento de conocimiento

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Introducción y Objetivos: Justificar la necesidad de desarrollar la tesis en lugar de adquirir o aplicar directamente lo existente a lo largo de la gama de categorías de procesos, productos y servicios de la empresa o institución usuaria de la tesis.

Estado del arte: Demostrar que se ha comprendido como es el ámbito de negocio donde se enmarca la tesis, sus hábitos y necesidades de productos o servicios tecnológicos.

Evaluación de Riesgos: Demostrar que se han considerado diferentes soluciones tanto clásicas como novedosas o innovadoras al problema, y se han evaluado los riesgos y ventajas de cada una de ellas.

Resultados: Justificar que la tecnología resultante de la tesis satisface los deseos o necesidades del cliente (real, potencial o ficticio).

Conclusiones: Establecer las conclusiones del trabajo apoyándose fundamentalmente en los datos y observaciones obtenidas durante su desarrollo. Discutir que medios, cauces, etapas y tecnologías harían falta (si procede) para llevar a cabo una implantación real de los resultados.

Líneas futuras: Discutir los límites de las tecnologías actuales aplicadas al problema, planteando líneas de I+D+i realistas y capaces de superarlos.

Capítulo 4

Conclusiones y líneas futuras de trabajo

Establecer las conclusiones del trabajo apoyándose fundamentalmente en los datos y observaciones obtenidas durante su desarrollo. Discutir que medios, cauces, etapas y tecnologías harían falta (si procede) para llevar a cabo una implantación real de los resultados. Discutir los límites de las tecnologías actuales aplicadas al problema, planteando líneas de I+D+i realistas y capaces de superarlos.

Capítulo 5

Sobre las referencias

La bibliografía o referencias deben aparecer siempre al final de la tesis, incluso en aquellos casos donde se hayan utilizado notas finales. La bibliografía debe incluir los materiales utilizados, incluida la edición, para que la cita pueda ser fácilmente verificada.

Citar dentro del texto: Las fuentes consultadas se describen brevemente dentro del texto y estas citas cortas se amplían en una lista de referencias final, en la que se ofrece la información bibliográfica completa.

La cita dentro del texto es una referencia corta que permite identificar la publicación de dónde se ha extraído una frase o parafraseado una idea, e indica la localización precisa dentro de la publicación fuente. Esta cita informa del apellido del autor, la fecha de publicación y la página (o páginas) y se redacta de la forma que puede verse a través de los siguientes ejemplos:

Cuando se citan las palabras exactas del autor deben presentarse entre comillas e indicarse, tras el apellido del autor y, entre paréntesis, la fecha de publicación de la obra citada, seguida de la/s página/s.

Si lo que se reproduce es la idea de un autor (no sus palabras exactas) no se pondrán comillas y se indicará, entre paréntesis, el apellido del autor seguido de la fecha de publicación de la obra a la que se refiere.

No se puede eliminar una parte del texto citado sin señalarse; debe indicarse siempre con puntos suspensivos entre corchetes [...]

Ejemplos de como citar una referencia en el texto son los siguientes ? o ???.

Oguntimilehin and Ademola (2014)

Cómo ordenar las referencias:

1. Las referencias bibliográficas deben presentarse ordenadas alfabéticamente por el apellido del autor, o del primer autor en caso de que sean varios.
2. Si un autor tiene varias obras se ordenarán por orden de aparición.
3. Si de un mismo autor existen varias referencias de un mismo año se especificarán los años seguidos de una letra minúscula y se ordenarán alfabéticamente.

4. Si son trabajos de un autor en colaboración con otros autores, el orden vendrá indicado por el apellido del segundo autor, independientemente del año de publicación. Las publicaciones individuales se colocan antes de las obras en colaboración.

Cómo citar un artículo de revista: Un artículo de revista, siguiendo las normas de la APA, se cita de acuerdo con el siguiente esquema general: Apellido(s), Iniciales del nombre o nombres. (Año de publicación). Título del artículo. Título de la revista en cursiva, volumen de la revista (número del fascículo entre paréntesis), primera página- última página del artículo.

Cómo citar una monografía/libro: Las monografías, siguiendo las normas de la APA, se citan de acuerdo con el siguiente esquema general: Apellido(s), Iniciales del nombre. (Año de publicación). Título del libro en cursiva. Lugar de publicación: Editorial. Opcionalmente podremos poner la mención de edición, que irá entre paréntesis a continuación del título; y, si fuera el caso el volumen que irá en cursiva.

Cómo citar un capítulo de un libro: Los capítulos de los libros se citan de acuerdo con el siguiente esquema general: Apellido(s), Iniciales del nombre o nombres. (Año). Título del capítulo. En A. A. Apellido(s) Editor A, B. B. Apellido(s) Editor B, y C. Apellido(s) Editor C (Eds. o Comps. etc.), Título del libro en cursiva (pp. xxx-xxx). Lugar de publicación: Editorial.

Cómo citar un acta de un congreso: Apellido(s), Iniciales del nombre o nombres. (Año). Título del trabajo. En A. A. Apellido(s) Editor A, B. B. Apellido(s) Editor B, y C. Apellido(s) Editor C (Eds. o Comps. etc.), Nombre de los proceedings en cursiva (pp. xxx-xxx). Lugar de publicación: Editorial.

Cómo citar tesis doctorales, trabajos fin de máster o proyectos fin de carrera: Apellido(s), Nombre. (Año). Título de la obra en cursiva. (Tesis doctoral). Institución académica en la que se presenta. Lugar.

Cómo citar un recurso de Internet: Los recursos disponibles en Internet pueden presentar una tipología muy variada: revistas, monografías, portales, bases de datos... Por ello, es muy difícil dar una pauta general que sirva para cualquier tipo de recurso. Como mínimo una referencia de Internet debe tener los siguientes datos:

1. Título y autores del documento.
2. Fecha en que se consultó el documento.
3. Dirección (URL “uniform resource locator”)

Veamos, a través de distintos ejemplos, cómo se citan específicamente algunos tipos de recursos electrónicos.

Monografías: Se emplea la misma forma de cita que para las monografías en versión impresa. Debe agregar la URL y la fecha en que se consultó el documento

Artículos de revistas: Se emplea la misma forma de cita que para los artículos de revista en versión impresa. Debe agregar la URL y la fecha en que se consultó el documento.

Artículos de revistas electrónicas que se encuentran en una base de datos: Se emplea la misma forma de cita que para los artículos de revista en versión impresa, pero debe añadirse el nombre de la base datos, la fecha en que se consultó el documento.

Apéndice A

Anexos

Bibliografía

- Gron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, 1st edn, O'Reilly Media, Inc.
- Larranaga, P., Inza, I. and Moujahid, A. (n.d.). Tema 5. clasificadores k-nn.
- Mohanty, S., Nathrout, K., Barik, S., Das, S. and Prof, A. (2015). A study on evolution of data in traditional rdbms to big data analytics, *International Journal of Advanced Research in Computer and Communication Engineering* **4**: 230–232.
- Oguntimilehin, A. and Ademola, O. (2014). A review of big data management, benefits and challenges, *Journal of Emerging Trends in Computing and Information Sciences* **5**: 433–438.
- Unagar, A. and Unagar, A. (2017). Support vector machines. unwinded.
URL: <https://medium.com/data-science-group-iitr/support-vector-machinessvm-unraveled-e0e7e3ccd49b>