UNIVERSIDAD DE EXTREMADURA

Escuela Politécnica

Grado en Ingeniería Informática en Ingeniería de Software

Trabajo Fin de Grado

Modelos predictivos aplicados a IoT

<Nombre y Apellidos del Autor>

<Convocatoria, Año>

UNIVERSIDAD DE EXTREMADURA

Escuela Politécnica

Grado en Ingeniería Informática en Ingeniería de Software

Trabajo Fin de Grado

Modelos predictivos aplicados a IoT

Autor: José Luis Pérez García

Tutor: Andrés Caro Lindo

Co-Tutor/es: José Carlos Sancho Núñez

**Tribunal Calificador**

Presidente: <Nombre y Apellidos>

Secretario: <Nombre y Apellidos>

Vocal: <Nombre y Apellidos>

**RESUMEN**

**ÍNDICE GENERAL DE CONTENIDOS**

[1 INTRODUCCIÓN 5](#_Toc105238215)

[2 OBJETIVOS 6](#_Toc105238216)

[3 ESTADO DEL ARTE 7](#_Toc105238217)

[**3.1** **MACHINE LEARNING** 7](#_Toc105238218)

[**3.2** **ESTADO ACTUAL DEL MACHINE LEARNING** 11](#_Toc105238219)

[**3.1** **DISPOSITIVOS IoT** 12](#_Toc105238220)

[3.2 **ESTADO ACTUAL DE LOS DISPOSITIVOS IoT** 14](#_Toc105238221)

[**3.3** **MACHINE LEARNING APLICADO A IOT** 15](#_Toc105238222)

[4 METODOLOGÍA 16](#_Toc105238223)

[5 IMPLEMENTACIÓN Y DESARROLLO 17](#_Toc105238224)

[5.1 XGBoost 17](#_Toc105238225)

[5.2 Linear Regression Múltiple 18](#_Toc105238226)

[5.3 Linear Regression Múltiple Stats Model 19](#_Toc105238227)

[5.4 Ramdom Forest 20](#_Toc105238228)

[5.5 Support Vector Regression 21](#_Toc105238229)

[5.6 Logistic Regression 22](#_Toc105238230)

[5.7 Stochastic Gradient Descent 23](#_Toc105238231)

[5.8 LASSO Regression 24](#_Toc105238232)

[5.9 Gausian Naïve Bayes 25](#_Toc105238233)

[5.10 Decision Tree Classifier 26](#_Toc105238234)

[5.11 Robust Regression RANSAC 27](#_Toc105238235)

[6 RESULTADOS Y DISCUSIÓN 28](#_Toc105238236)

[7 CONCLUSIONES 29](#_Toc105238237)

[8 REFERENCIAS BIBLIOGRÁFICAS 30](#_Toc105238238)

[Anexo 1: Instalación del entorno 31](#_Toc105238239)

[Anexo 2: Librerías usadas 34](#_Toc105238240)

**ÍNDICE DE FIGURAS**

[Figura 1: Beneficios del Machine Learning en el ámbito empresarial. Iberdrola 8](#_Toc104916815)

# INTRODUCCIÓN

# OBJETIVOS

# ESTADO DEL ARTE

Para proceder a realizar un análisis del proyecto realizado, antes se va a introducir a grandes rasgos en que consiste el Machine Learning y en que consiste IOT y en qué estado se encuentran actualmente cada tecnología. Posteriormente se indicará como se aplica una tecnología a otra.

## **MACHINE LEARNING**

A día de hoy podemos encontrar muchos ejemplos en Machine Learning en nuestro día a día, pero desconocemos que lo es, por ejemplo, coches que conducen solos (son ya muy famosos los coches de la marca Tesla), asistentes que traducen instantáneamente de un idioma a otro (la cámara en el traductor de Google) o sugerencias de compra personalizadas (por ejemplo el mail que nos llega de Amazon con el mensaje “Te podría interesar”).

**¿Qué es el Machine Learning?**

El Machine Learning o Aprendizaje Automático se conoce por ser una disciplina del campo de la Inteligencia Artificial que mediante el uso de algoritmos, proporciona a los ordenadores la capacidad de poder identificar patrones en grandes conjuntos de datos y con ello poder elaborar predicciones. A esto también se le conoce como análisis predictivo. Este aprendizaje dota a los ordenadores de la posibilidad de poder realizar tareas específicas de forma autónoma. Así, de esta manera no sería necesario que un usuario los programe.

El término Machine Learning fue usado por primera vez en el 1959. No obstante hasta en estos últimos años no se ha empezado a escuchar de verdad. Esto ha sido debido al gran aumento en la capacidad de computación y al uso de los conjuntos masivos de datos. De hecho, en el mundo del Big Data el Machine Learning es una parte fundamental.

**Algoritmos de Machine Learning**

Los algoritmos de Machine Learning se dividen en tres categorías, siendo las dos primeras las más usadas:

Aprendizaje supervisado: este tipo de algoritmos necesitan un aprendizaje previo basado que está basado en un sistema de etiquetas asociadas a datos que permiten poder tomar decisiones o hacer predicciones. Por ejemplo, un detector de spam que se usa para etiquetar un mail como spam hace uso patrones que ha aprendido del histórico de correos (remitente, relación entre texto e imágenes, palabras clave en el asunto, etc.).

Aprendizaje no supervisado: estos algoritmos no requieren de un conocimiento previo. Analizan los datos (incluso desordenados) con el objetivo de encontrar patrones que permitan organizarlos de alguna forma. Por ejemplo, en marketing es muy usado para extraer patrones de grandes conjuntos de datos que provienen de las redes sociales y así poder crear campañas de publicidad altamente segmentadas.

Aprendizaje por refuerzo: el objetivo de este tercer tipo de algoritmos es que este aprenda de la experiencia. De forma que sea capaz de tomar la mejor decisión ante diferentes situaciones mediante un proceso de prueba y error en el que se recompensan las decisiones correctas. Por ejemplo en la actualidad lo podemos encontrar en reconocedores faciales, clasificación de secuencias de ADN o diagnósticos médicos.

**Beneficios del Machine Learning en el ámbito empresarial**

En el siguiente gráfico se pueden ver los beneficios que las empresas obtienen las al usar Machine Learning, además después de dicho gráfico se puede ver una explicación de cada cosa.

Figura 1: Beneficios del Machine Learning en el ámbito empresarial. [Iberdrola](https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico)

Impulsa la innovación: El aprendizaje automático se sirve de los datos masivos, los analiza y busca nuevas soluciones.

Reduce costes: Dado que con el aprendizaje automático se pueden automatizar tareas, esto hace que se las empresas puedan ahorrar capital humano. Además les permite optimizar tiendas gracias a los datos de navegación y flujo de clientes.

Mejora la relación con el cliente: Tener un Chatbot online las 24 horas, los 7 días de la semana respondiendo permite recoger datos para profundizar en el conocimiento del consumidor

Mejora la segmentación de anuncios: Con los algoritmos de Machine Learning se pueden predecir los contenidos que son más eficaces para cada objetivo.

Mejora la segmentación del público objetivo: Gracias a los algoritmos de aprendizaje no supervisado se pueden encuentrar algunos patrones en los datos que recompilan las empresas sobre los consumidores.

Predice tendencias: Mediante un análisis de los hábitos de compra, los algoritmos de Machine Learning pueden predecir los productos que tendrán más demanda y en que momento se pueden subir o bajar precios.

**Aplicaciones prácticas del Machine Learning**

El Machine Learning es uno de los fundamentos básicos de la transformación digital. Actualmente se está utilizando para el desarrollo de nuevas soluciones en diferentes campos, entre los que podemos encontrar:

Recomendaciones: le permite por ejemplo poder hacer sugerencias personalizadas de compra en algunas plataformas online. La forma más sencilla de explicar cómo funciona esto sería la siguiente: se encarga de analizar nuestro historial de compras y lo compara con las tendencias de compra de otros usuarios con gustos iguales o parecidos.

Vehículos inteligentes: actualmente como ya se señalaba anteriormente en el documento, ya podemos encontrar algunos prototipos de la marca Tesla con funciones de conducción autónoma. Además también, gracias al aprendizaje automático, estos vehículos podrán ajustar las configuraciones internas como son por ejemplo la temperatura, música, inclinación del respaldo, etc. Esto de acuerdo a las preferencias del conductor e, incluso, mover el volante solos para reaccionar a cualquier situación extrema que surja del entorno.

Redes sociales: Twitter, por ejemplo, se sirven de algoritmos de Machine Learning para reducir de forma drástica el spam que se publica por los usuarios de esta red. Mientras Facebook, lo usa para poder detectar noticias falsas o contenidos no permitidos en retransmisiones de videos en directo, los cuales son bloqueados en el acto de forma automática.

Procesamiento de Lenguaje Natural: mediante la comprensión de la voz humana, los asistentes virtuales como Alexa, Google o Siri pueden traducir de forma instantánea un idioma a otro. Además, pueden reconocer la voz del usuario e incluso analizar sus sentimientos.

Búsquedas: los buscadores como por ejemplo Google o Bing usan el aprendizaje automático para mejorar sus resultados en función de su eficacia, midiendo esta, a través de los clics de los usuarios.

Medicina: se está utilizando el Machine Learning para detectar en la mayor brevedad y con la mayor antelación el cáncer de mama, ya que su detección temprana aumenta las probabilidades de curación También se utiliza con una alta eficacia para detectar neumonía y algunas enfermedades de la retina que pueden provocar ceguera en el futuro.

Ciberseguridad: los nuevos ya usan el Machine Learning para mejorar el escaneado, de esta forma se acelera la detección y se mejorara la habilidad de reconocer anomalías.

## **ESTADO ACTUAL DEL MACHINE LEARNING**

## **DISPOSITIVOS IoT**

IoT, Internet de las cosas o Internet of Things está muy presente en la actualidad, y cada vez hay más fabricantes que crean dispositivos de este tipo.

**¿Qué es IoT o Internet de las cosas?**

IOT es una red de **interconexión digital entre dispositivos**, personas e Internet que hace posible compartir datos, lo que permite que se pueda capturar información sobre el uso y el rendimiento de los dispositivos para así poder detectar patrones y hacer por ejemplo recomendaciones con las que mejorar la experiencia del usuario.

Por ejemplo, IOT es la conexión entre tu smartphone y los dispositivos Smart que tienes en casa que pueden ser: un aire acondicionado o una Raspberry Pi (que controla la programación de tu televisor) o también puede ser un sensor de temperatura (que puede almacenar datos históricos de la temperatura de un lugar).

El termino IOT se empieza a usar cuando los dispositivos (no solo ordenadores) empiezan acceder a la red, para obtener información que necesitan y así poder dar sus servicios.

**Cómo funciona el IoT**

Los dispositivos IoT se conectan entre si con un proceso llamado Machine to Machine (máquina a máquina). En este proceso, los dos dispositivos se comunican entre sí utilizando cualquier tipo de conectividad como por ejemplo Wifi o Bluetooth. De esta manera se consigue que puedan realizar un trabajo sin necesidad de que haya un usuario de por medio.

Una vez se han conectado los dispositivos, estos generan gran cantidad de datos que se almacenan en una plataforma IoT, la cual recolecta y procesa y posteriormente analiza los datos. Gracias a estos datos el usuario puede sacar conclusiones de hábitos y preferencias de él mismo.

**Influencia de los dispositivos IoT en nuestra vida actual**

Durante nuestro día a día nos podemos encontrar con una enorme cantidad de dispositivos que forman parte del Internet de las cosas. Se van a describir a continuación algunos ejemplos:

Vehículos autónomos: como ya se habló anteriormente en la sección de Machine Learning, cada vez se tiene más presente el futuro de los coches con conducción autónoma (lo cual, no es otra cosa que coches que conducen solos). Conforme pasa el tiempo, estos tienen más tecnología, y se basan en el uso de sensores para poder realizar dicha conducción autónoma.

Robots aspiradora: estos tienen gran cantidad de sensores que les permiten realizar una limpieza del sin chocarse con ningún obstáculo.

Smart home: los dispositivos smart pueden ser desde sensores de presencia que encienden las luces automáticamente hasta sensores de temperatura o humedad (son en estos en los que se ha basado el estudio) que recogen constantemente datos que luego pueden ser analizados por el usuario.

## **ESTADO ACTUAL DE LOS DISPOSITIVOS IoT**

## **MACHINE LEARNING APLICADO A IOT**

# METODOLOGÍA

# IMPLEMENTACIÓN Y DESARROLLO

En esta sección se van a describir los algoritmos usados para hacer las predicciones.

## XGBoost

¿Qué es XGBoost?

Conocido como eXtreme Gradient Boosting, se trata de una librería o algoritmo

## Linear Regression Múltiple

<https://realpython.com/linear-regression-in-python/>

## Linear Regression Múltiple Stats Model

https://realpython.com/linear-regression-in-python/

## Ramdom Forest

penalized linear regression

## Support Vector Regression

M5p -> Predicion tree -> <https://www.analyticsvidhya.com/blog/2016/04/tree-based-algorithms-complete-tutorial-scratch-in-python/>

Gradient Boosting -> xgboost -> revisar proyecto telecoming

random forest -> https://towardsdatascience.com/random-forest-in-python-24d0893d51c0

## Logistic Regression

SVM -> <https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python>

Para regresión: <https://medium.com/pursuitnotes/support-vector-regression-in-6-steps-with-python-c4569acd062d>

## Stochastic Gradient Descent

Stochastic Gradient Descent -> <https://scikit-learn.org/stable/modules/sgd.html>

## LASSO Regression

eARTH/bagEARTH (py-earth) -> http://mehdidc.github.io/pyearth-doc/

## Gausian Naïve Bayes

Naïve Bayes -> <https://scikit-learn.org/stable/modules/naive_bayes.html>

## Decision Tree Classifier

https://scikit-learn.org/stable/modules/generated/sklearn.isotonic.IsotonicRegression.html

## Robust Regression RANSAC

<https://machinelearningmastery.com/make-predictions-scikit-learn/>

# RESULTADOS Y DISCUSIÓN

# CONCLUSIONES

# REFERENCIAS BIBLIOGRÁFICAS

**Libros**

Burkov, Andriy, 2019: *The hundred-page machine learning book*. ISBN 9781999579517

<https://realpython.com/linear-regression-in-python/>

<http://webapplayers.com/>

<https://wpbeaches.com/installing-configuring-apache-on-macos-using-homebrew/>

<https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico>

# Anexo 1: Instalación del entorno

Para poder ejecutar nuestro proyecto, vamos a necesitar instalar apache como servidor. Pero, ¿qué es apache? Apache es un software de servidor web gratuito y de código abierto para plataformas Unix(Linux, Mac por ejemplo) con el cual se ejecutan el 46% de los sitios web de todo el mundo. Es mantenido y desarrollado por la Apache Software Foundation.

Le permite a los propietarios de sitios web o a un particular en su propio servidor o pc servir contenido en la web o de forma local, de ahí el nombre de servidor web. Es uno de los servidores web más antiguos y confiables. Su primera versión fue lanzada hace más de 20 años, en 1995.

Cuando alguien quiere visitar un sitio web, ingresa un nombre de dominio en la barra de direcciones de su navegador. Luego, el servidor web envía los archivos solicitados actuando como un repartidor virtual.

Sabiendo que es Apache, proseguimos con los pasos para realizar la instalación, en esta documentación el proyecto se ha realizado sobre un Mac, con lo cual los pasos para la instalación van a ser en lenguaje de consola de Mac OS

El comando de instalación que vamos a usar es:

**brew install httpd**

Una vez instalado vamos a indicar en el archivo httpd.conf cuál va a ser la ruta que debe abrir.

Primero abrimos el archivo para editarlo con el siguiente comando:

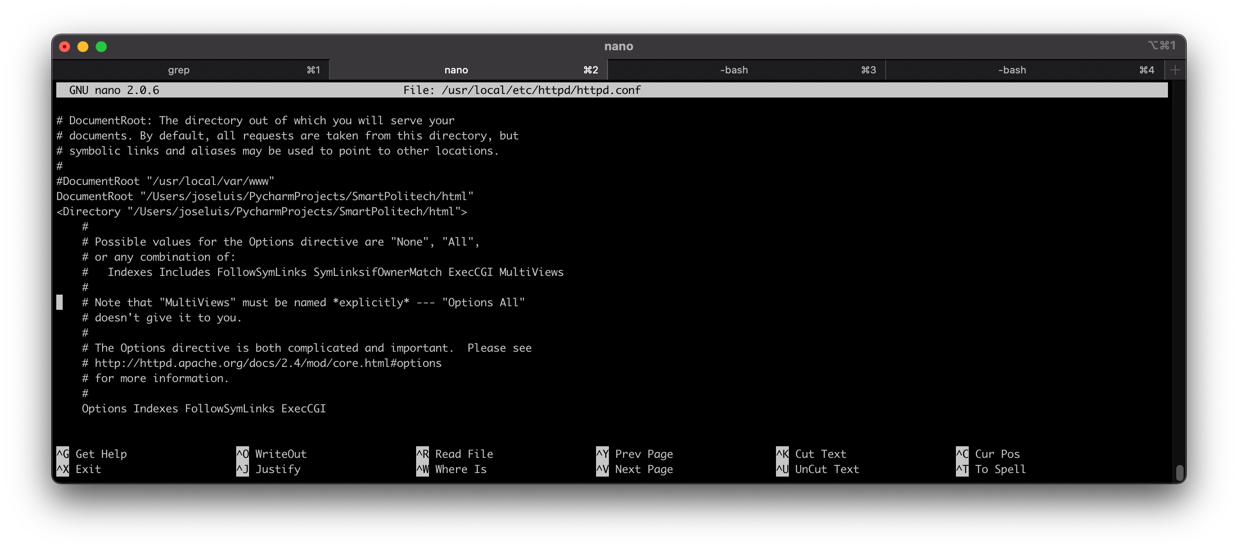
**nano /usr/local/etc/httpd/httpd.conf**

Bajamos hasta la linea donde pone:

# DocumentRoot: The directory out of which you will serve your

# documents. By default, all requests are taken from this directory, but

# symbolic links and aliases may be used to point to other locations.



Justo debajo de esta linea indicaremos donde se encuentra nuestra carpeta con el main principal. Por defecto aparecerá #DocumentRoot “/usr/local/var/www”, con lo cual, tendríamos que cambiar “/usr/local/var/www” por nuestra ruta. En mi caso será “/Users/joseluis/PycharmProjects/SmartPolitech/html”, quedando la linea así:

**#DocumentRoot “/Users/joseluis/PycharmProjects/SmartPolitech/html”**.

Una vez hecho esto, debajo de esta linea hay otra que también debemos modificar:

<Directory “/usr/local/var/www”> y haremos lo mismo, sustituiremos “/usr/local/var/www” por nuestra ruta, en mi caso quedando así:

**<Directory “/Users/joseluis/PycharmProjects/SmartPolitech/html”>**

Hecho esto, bajamos un poco hasta donde encontraremos el fragmento de código siguiente:

<IfModule dir\_module>

DirectoryIndex index.html

</IfModule>

Si nuestro main, se llama index.html, lo dejaremos como está, sin por el contrario se llama de otra forma, deberemos cambiarlo. En mi caso se llama main.html, por lo que mi fragmento de código quedaría así:

<IfModule dir\_module>

DirectoryIndex main.html

</IfModule>

Una vez hemos terminado, pulsamos control + X para guardar los cambios y salir del archivo. Ya tenemos configurado el archivo de configuración para poder ejecutar nuestro servidor local con el proyecto.

Ahora tenemos que levantar el servidor, para ello usaremos los siguientes comandos:

Inicio de httpd: **brew services start httpd**

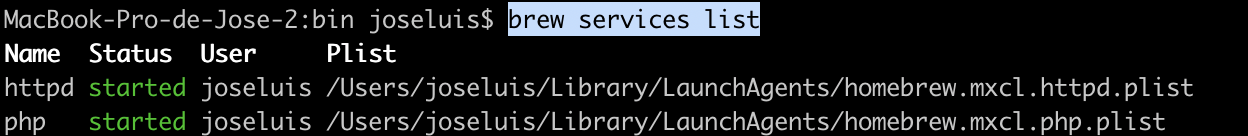
Interrupción de httpd: **brew services stop httpd**

Reinicio de httpd: **brew services restart httpd**

Para comprobar que se ha iniciado el servidor podemos usar el siguiente comando

**brew services list**

Este comando mostrará todos los servicios iniciados en el sistema y su estado.

****

Cómo podemos ver en la captura anterior tenemos servicio httpd en estado started.

# Anexo 2: Librerías usadas

Durante el Proyecto, se han usado varias librerías en los diferentes lenguajes con lo que se ha programado:

Librería XGBoost -> se instala con el comando pip3 install xgboost

Librería Sklearn -> se instala con el comando pip3 install sklearn

Librería Statsmodels -> se instala con el comando pip3 install statsmodels

Librería Bartpy -> se instala con el comando pip3 install bartpy

# 