



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



**TFG del Grado en Ingeniería
Informática**

KairósWatch

**Entorno de compra y venta de
relojes por subasta.**



Presentado por Rodrigo Pérez Ubierna
en Universidad de Burgos — 20 de abril
de 2024

Tutor: Dra. Sandra Rodríguez Arribas
y D. Jose Antonio Barbero Aparicio



UNIVERSIDAD DE BURGOS
ESCUELA POLITÉCNICA SUPERIOR
Grado en Ingeniería Informática



D. nombre tutor, profesor del departamento de nombre departamento, área de nombre área.

Expone:

Que el alumno D. Rodrigo Pérez Ubierna, con DNI dni, ha realizado el Trabajo final de Grado en Ingeniería Informática titulado título de TFG.

Y que dicho trabajo ha sido realizado por el alumno bajo la dirección del que suscribe, en virtud de lo cual se autoriza su presentación y defensa.

En Burgos, 20 de abril de 2024

Vº. Bº. del Tutor:

Vº. Bº. del co-tutor:

D. nombre tutor

D. nombre co-tutor

Resumen

En este primer apartado se hace una **breve** presentación del tema que se aborda en el proyecto.

Descriptores

Palabras separadas por comas que identifiquen el contenido del proyecto Ej: servidor web, buscador de vuelos, android ...

Abstract

A **brief** presentation of the topic addressed in the project.

Keywords

keywords separated by commas.

Índice general

Índice general	iii
Índice de figuras	iv
Índice de tablas	v
1. Introducción	1
2. Objetivos del proyecto	3
3. Conceptos teóricos	5
3.1. Parte teórica Machine Learning	5
4. Técnicas y herramientas	9
5. Aspectos relevantes del desarrollo del proyecto	11
6. Trabajos relacionados	13
7. Conclusiones y Líneas de trabajo futuras	15

Índice de figuras

Índice de tablas

1. Introducción

Kairós no es un trabajo cualquiera. Kairós nace de un hoobie que arrastro desde hace unos años, concretamente desde la herencia de mi primer reloj. Yo no solía fijarme mucho en ellos, pues siempre creí que no era más que una simple herramienta para saber qué hora era en ese momento. Sin embargo, poco a poco decidí informarme sobre este mundo y pase a verlo como lo que era: arte. Un reloj no es un simple accesorio. Un reloj define la identidad de la persona. Muchas personas afirman que un conjunto bonito pierde todo *glamour* si no es acompañado de un reloj en la muñeca. Por esta razón, nunca salgo sin que haya algo que pese en mi brazo.

Actualmente, el coleccionismo de relojes atrae a numerosas personas, especialmente hombres. Podríamos hablar de que civilización se atribuye la invención del reloj, aunque no sería de apoyo al trabajo. Sin embargo, una gran curiosidad es quién portaba los primeros relojes de pulsera en la historia: las mujeres. Los hombres siempre portaban relojes de bolsillo agarrados con una cadena. ¿Curioso verdad? Pues no es hasta la Primera Guerra Mundial cuando los hombres deciden utilizar los relojes de pulsera debido a la comodidad detrás de la trinchera.

¿Y por qué este apartado anterior? La respuesta es sencilla: las principales marcas de relojes hacen colecciones para el género masculino, pues son ellos más propensos a portarlos de manera diaria. De ahí que la gran mayoría de relojes que veremos a lo largo del trabajo lleven la característica del género.

Tras un tiempo informandome, hablando con otros amantes de este mundo... ví que hay un gran obstáculo en esta sociedad cuando hablamos de relojes: la compra y venta de estos productos. No existen apenas sitios web o aplicaciones que se dediquen exclusivamente a la adquisición, bien por subasta o bien por compra directa, de relojes. Por esto, surge Kairós.

Kairós es una aplicación multiplataforma donde cualquier usuario podrá comprar y/o vender su reloj de una manera sencilla. La adquisición o venta de estos se realizará a través de varios tipos de subastas a elegir por el usuario, así como la venta directa del reloj si el usuario lo deseara.

Y ahora una de las preguntas que surgen dentro de este entorno: pero, ¿cuánto vale mi reloj? Esta pregunta es la más formulada a la hora de vender esta pieza de arte. Desde Kairós queremos eliminar cualquier duda del usuario y, por ello, la aplicación cuenta con un sistema de predicción del precio del reloj según diversas características propias del accesorio.

Con todo situado, podemos marcar cuáles van a ser los principales apartados de la memoria del trabajo y la dinámica que se va a seguir para su realización:

1. Objetivos del proyecto
2. Conceptos teóricos
3. Técnicas y herramientas
4. Aspectos relevantes del desarrollo del proyecto
5. Trabajos relacionados
6. Conclusiones y líneas de trabajo futuras

Cada apartado expondrá distintas cosas que se han ido realizando durante el proceso, aunque todos compartiran una misma estructura. En todo ellos se explicarán tanto la parte de la aplicación como la parte del Machine Learning, de forma que todo quede lo más limpio y ordenado posible.

A modo de ser diferente al resto, aprovecho esta parte del informe para desearles que disfruten del trabajo tanto como yo voy a disfrutar contándoselo.

2. Objetivos del proyecto

Este apartado explica de forma precisa y concisa cuales son los objetivos que se persiguen con la realización del proyecto. Se puede distinguir entre los objetivos marcados por los requisitos del software a construir y los objetivos de carácter técnico que plantea a la hora de llevar a la práctica el proyecto.

3. Conceptos teóricos

3.1. Parte teórica Machine Learning

Tras varios momentos pensando como esta aplicación podría situarse como una clara competidora en el mercado, me di cuenta de que muchas aplicaciones no se centran en lo que de verdad importa: el cliente. Siempre se ha dicho que el cliente tiene la razón, que el cliente debe encontrarse cómodo... pero a la hora de la verdad, las aplicaciones no destinan el esfuerzo suficiente a dejar todo lo más sencillo, amigable y accesible que el cliente merece. Por ello, lo primero que hice fue ponerme en su piel y pensar que me gustaría encontrarme en mi aplicación, que problemas podría a llegar a tener... y di con ello: desconozco cuánto vale mi reloj. Sé que puede sonar raro, pero es así. Son tantos los precios que se estipulan a un mismo reloj que sé, a ciencia cierta, que sería incapaz de marcar un precio de venta directa donde consiguiese el mayor valor de venta posible. De aquí nace la idea de investigar acerca del aprendizaje automático o Machine Learning. Mi objetivo principal es crear un modelo capaz de predecir cuál es el precio más recomendable para la venta de mi reloj, apoyándome en un dataset de más de 280000 relojes de diferentes marcas, estados, géneros, precios... pero no nos adelantemos. Veamos que es todo esto.

El aprendizaje automático o Machine Learning es una parte de la informática muy ligada al concepto de AI o Inteligencia Artificial. Definir el objetivo de esta es simple: conseguir que una máquina se comporte como un ser humano. Conseguir el objetivo de esto es algo más complicado. Principalmente se busca entrenar a la máquina con el uso de algoritmos y datos de forma que esta aprenda de manera gradual. No hay que decir que cuánto mayor sea el número de datos y cuánto mayor sea la calidad de estos, mayor precisión se conseguirá.

El término Machine Learning podemos atribuírselo a Arthur Samuel quien, en 1952, creó un software capaz de aprender y jugar a las damas. Aun así, este campo ha evolucionado a pasos agigantados, siendo actualmente uno de los puntos más relevantes dentro del campo de la informática. Entre los hitos más significativos, destacan principalmente unos estudiantes de la Universidad de Stanford. Estos alumnos consiguieron desarrollar en 1979 un software capaz de pilotar un carro de manera autónoma sin que este chocara con ningún obstáculo. No podemos olvidarnos de uno de los momentos más significativos de nuestra época cuando el gran Garri Kaspárov fue derrotado por Deeper Blue en una partida de ajedrez en 1997, siendo la primera vez que el ajedrecista perdía contra una máquina.

Como se ha marcado en esta introducción, esta rama de la informática se centra en desarrollar algoritmos y modelos autosuficientes, es decir, que sean capaces de aprender de manera automática a través de datos y la experiencia. Dentro de esta rama, encontramos infinidad de técnicas y algoritmos, aunque me voy a quedar con los más comunes.

El aprendizaje supervisado es una rama del aprendizaje automático donde se entrena a un modelo con una serie de datos que contienen cuál es su entrada y cómo debe ser su salida. Un ejemplo de ello es la venta de un local: el modelo recibe como variables de entrada el número de baños, los metros cuadrados, el año de edificación... y el modelo procesa todo dando como salida el precio del local. Dentro de este tipo de aprendizaje pueden abordarse dos tipos de problemas: regresión si lo que devuelve la máquina es un valor justo (siguiendo con el ejemplo anterior, el precio del local), o clasificación si lo que devuelve es una categoría (clasificar imágenes de vehículos entre coches o camiones). Hay que añadir que son muchos los algoritmos de los que consta, aunque destacan principalmente:

Regresión lineal: se modela la relación de una variable dependiente con una o varias variables independientes.

Regresión logística: se modela para solventar problemas de clasificación binaria.

Regresión logística: se modela para solventar problemas de clasificación binaria.

SVM: se utiliza tanto en regresión como en clasificación. El objetivo es separar las clases en el espacio de características a través de la búsqueda del hiperplano.

Árboles de decisión: el objetivo es dividir el espacio de características en regiones y asignar una etiqueta a cada una de ellas.

Random forest: conjunto de árboles de decisión que trabajan de manera combinada para alcanzar una mayor precisión y así evitar el sobreajuste.

Por otro lado estaría el aprendizaje no supervisado donde los datos no tienen ni etiquetas ni salidas asociadas. Su objetivo es que la máquina aprenda a base de identificar patrones en los datos y extraer información útil de ellos. Para explicar mejor esto, pongamos un ejemplo: teniendo un conjunto de transacciones bancarias, determinar grupos de personas según las similitudes de compras. Dentro de ello podemos distinguir tres tipos de problemas:

Clustering: el objetivo es agrupar los datos en conjuntos o clusters según existan similitudes entre ellos. Por indicar alguno de los algoritmos dentro de este tipo, destacan K-Means, Clustering Jerárquico o DBSCAN.

Reducción de dimensionalidad: el objetivo es reducir el número de variables en un conjunto de datos sin perder la mayor parte de la información. Ejemplos de algoritmos son Análisis de Componentes Principales (PCA) y T-Distributed Stochastic Neighbor Embedding (t-SNE)

Asociación: el objetivo de la máquina es que aprenda a base de relacionar variables en un conjunto de datos determinado. Un ejemplo dominante dentro de este grupo son la sección de recomendación de webs de compras. Ejemplos de algoritmos dentro de ello son Eclat y FP-Growth.

Si combinamos las ideas fundamentales de estos dos aprendizajes anteriores, damos lugar al aprendizaje semisupervisado. Simplemente permite que la máquina aprenda a base de datos sin etiqueta y con ella.

Otro tipo de Machine Learning es el conocido como aprendizaje por refuerzo. La idea fundamental de este tipo de aprendizaje es que la respuesta del modelo mejore a partir de una retroalimentación en forma de recompensas. En otras palabras, el algoritmo va a aprender basándose en lo que le rodea y siguiendo la filosofía “ensayo-error”. Se utiliza en infinidad de sectores y los algoritmos por excelencia dentro de este aprendizaje son Q-Learning, SARSA, Algoritmo de Policy Gradient...

Por último, otro aprendizaje muy presente en diversos sectores es el Deep Learning. El objetivo de este campo es emular el comportamiento de un cerebro humano a partir de algoritmos de redes neuronales artificiales formadas por múltiples capas de procesamiento capaces de representar datos de alto nivel. Ejemplos muy característicos son el reconocimiento de imágenes, detección de objetos, detección de fraudes, traducción automática...

4. Técnicas y herramientas

4.1. Técnicas y herramientas en Machine Learning

Habiendo introducido los aspectos clave del Machine Learning, podemos afirmar que nuestro caso corresponde a un modelo de regresión. Voy a realizar mi tratamiento del dataset de relojes con el fin de encontrar cual es el precio de venta más recomendable para cualquier reloj. Para ello, las dos herramientas clave que utilizaré son: Python (aprovechando la biblioteca de data science Panda) y mi dataset en formato .csv.

Las tres fases fundamentales dentro de esta labor práctica son:

- Puesta a punto de los datos
- Entrenamiento de nuestro modelo
- Predicciones

El dataset que voy a utilizar ha sido importado de Kaggle, cuyo link dejo marcado a continuación: <https://www.kaggle.com/datasets/philmorekoun911/luxury-watch-listings> . Siguiendo el orden que marcan las columnas del archivo, podemos crear una leyenda tal que:

unnamed: número de línea predefinido

name: nombre del reloj

price: precio del reloj

brand: marca del reloj

model: modelo del reloj

ref: número de referencia del reloj

mvmt: tipo de movimiento del reloj

casem: material de la caja

bracem: material del brazalete

yop: año de producción

cond: estado

sex: género

size: tamaño

condition: estado

5. Aspectos relevantes del desarrollo del proyecto

De manera que el trabajo quede estructurado de la mejor manera posible, vamos a dividir este apartado en dos grandes subgrupos donde Como marcamos en anteriores apartados, el trabajo se divide en dos grandes partes:

1. **Travis:** realizaba una compilación del proyecto, ejecutaba los test unitarios, ejecutaba *Lint*, ponía en marcha un emulador de Android, y ejecutaba los Android test sobre este. Al finalizar, enviaba los resultados a Codecov y SonarQube.
2. **Codecov:** realizaba un análisis sobre la cobertura de los test unitarios.
3. **CodeClimate:** ejecutaba cuatro motores de chequeo (*checkstyle*, *fixme*, *pmd* y *markdownlint*) sobre el código para detectar posibles problemas o vulnerabilidades en él.
4. **SonarQube:** analizaba código duplicado, violaciones de estándares, cobertura de tests unitarios, *bugs* potenciales, etc.
5. **VersionEye:** chequeaba todas las dependencias utilizadas en la aplicación y comprobaba si estaban actualizadas, si tenían algún problema de seguridad conocido, o si violaban la licencia del proyecto.

6. Trabajos relacionados

Este apartado sería parecido a un estado del arte de una tesis o tesina. En un trabajo final grado no parece obligada su presencia, aunque se puede dejar a juicio del tutor el incluir un pequeño resumen comentado de los trabajos y proyectos ya realizados en el campo del proyecto en curso.

7. Conclusiones y Líneas de trabajo futuras

Todo proyecto debe incluir las conclusiones que se derivan de su desarrollo. Éstas pueden ser de diferente índole, dependiendo de la tipología del proyecto, pero normalmente van a estar presentes un conjunto de conclusiones relacionadas con los resultados del proyecto y un conjunto de conclusiones técnicas. Además, resulta muy útil realizar un informe crítico indicando cómo se puede mejorar el proyecto, o cómo se puede continuar trabajando en la línea del proyecto realizado.