



ETSI SISTEMAS
INFORMÁTICOS



POLITÉCNICA

HACKATHON

gmv[®]
INNOVATING SOLUTIONS

BASES

- 1** Hackaton GMV es un concurso de resolución de desafíos propuestos por GMV y profesores de la UPM que están inspirados en la operativa de la compañía. El concurso está dirigido a estudiantes universitarios y su objetivo es promover la resolución de problemas reales.
- 2** Los participantes deben ser estudiantes de la Universidad Politécnica de Madrid.
- 3** Se propondrán tres desafíos durante el concurso. Cada participante puede resolver tantos desafíos como quiera. Sin embargo, solo podrá resultar premiada una de sus soluciones.
- 4** Se concederá un premio a la mejor solución de cada uno de los desafíos, y una opción a una beca dentro de la Cátedra GMV universidad / empresa en el laboratorio de GMV en la Escuela Técnica Superior de Ingeniería de Sistemas Informáticos y una Raspberry.
- 5** Se valorará la utilidad y carácter innovador de las soluciones a los desafíos.
- 6** Para participar, será necesario inscribirse en la web del evento. <https://forms.gle/H7QKhZEGtNCCdNuWA>
- 7** La entrega de premios se realizará en un acto público en la Escuela Técnica Superior de Ingeniería de Sistemas Informáticos durante el mes de septiembre. Los ganadores de cada desafío presentarán su solución durante el acto.
- 8** El jurado estará formado por especialistas de la empresa GMV y por profesores de la Universidad Politécnica de Madrid. Su composición se anunciará oportunamente y su veredicto será inapelable.
- 9** La mera participación en el concurso supone la aceptación de las presentes bases.
- 10** Fecha fin de presentación de solicitudes: 30 de julio. Fecha fin de presentación de los resultados de los retos: 15 de septiembre.

PROBLEMA DE OPTIMIZACIÓN

Un camión tiene que repartir un conjunto de piezas entre 100 hangares de una planta de fabricación de aviones. En la hora de reparto se nos indica el número de piezas de cada tipo y a qué hangar va cada pieza. Por otro lado tenemos las coordenadas de cada hangar, y suponemos que tenemos una tabla de distancias para llegar desde cada hangar a cada otro.

■ Problema 01

Se pide encontrar una ruta que permita visitar los 100 hangares de modo que el trayecto recorrido sea lo más eficiente posible, partiendo del hangar-0 y terminando de nuevo en él. Para calcular la eficiencia supondremos que el camión de reparto consume de media 25 litros a los 100 kilómetros, y el coste del diesel será de 2€ por litro. El objetivo es minimizar el coste del diesel consumido.

■ Problema 02

Suponiendo que cada pieza tiene un coste de penalización por cada turno en que no se ha entregado (ya que los obreros están parados esperándola), es decir, si se entrega en primer lugar no tiene penalización; si se entrega en segundo lugar tiene penalización $1 p n$, siendo p el coste de penalización de ese tipo de pieza y n el número de piezas de ese tipo a entregar, etc.

Dado este sistema de penalización, encontrar la ruta con menor penalización sin importar el combustible consumido.

■ Problema 03

Si queremos optimizar conjuntamente los dos objetivos, encontrar posibles soluciones de modo que la ruta sea corta pero también que la penalización sea baja.

Para ello utilizaremos el coste económico de ambas partes, es decir, el coste del combustible consumido más el coste de las penalizaciones.

■ Funciones auxiliares

Se recomienda implementar las siguientes funciones auxiliares:

```
def distancia_total (ch):
    """dado un camino ch, devuelve la distancia total recorrida al seguir ese camino"""
    pass

distancia_total([0, 1]) # debe dar 1660.786
distancia_total(range(100)) # debe dar 106348.437

def penalizacion (ch):
    """Dado un camino ch, devuelve el coste de la penalización en que incurre ese camino"""
    pass

penalizacion([0,1]) # debe dar 10.4
penalizacion([1,0]) # debe dar 8.4
penalizacion(range(100)) # debe dar 51476.44

def optimizar ():
    """Lanza el proceso de optimización"""
    pass
```

■ Entregable

Se deberá entregar un notebook de Jupyter en Google Colab con el código en python que resuelva el problema. Está permitido el uso de cualquier librería de código externa, indicando el proceso de instalación.

Se valorará la generación de gráficas con matplotlib que ayuden a visualizar las soluciones obtenidas.

El código entregado debe permitir la utilización de otras tablas de distancias y de piezas para las pruebas por el jurado.

■ Material

Podrás acceder a todo el material que necesitas pinchando en el siguiente link: <https://drive.upm.es/s/CczQBXeOu18z1sW>

PROBLEMA DE DATA ANALYTICS: CAZA DE EXOPLANETAS EN EL ESPACIO PROFUNDO

Los datos describen el cambio de flujo (intensidad luminosa) de varios miles de estrellas. Cada estrella tiene una etiqueta binaria de 2 o 1. 2 indica que se ha confirmado que la estrella tiene al menos un exoplaneta en órbita; algunas observaciones son de hecho sistemas multiplanetarios.

Como se puede imaginar, los planetas en sí no emiten luz, pero las estrellas que orbitan sí. Si se observa dicha estrella a lo largo de varios meses o años, es posible que se produzca una “atenuación” regular del flujo (la intensidad de la luz). Esto es una prueba de que puede haber un cuerpo en órbita alrededor de la estrella; dicha estrella podría ser un sistema “candidato”. Un estudio más detallado de nuestro sistema candidato, por ejemplo, mediante un satélite que capte la luz en una longitud de onda diferente, podría consolidar la creencia de que el candidato puede, de hecho, ser “confirmado”.

■ Descripción

Conjunto de entrenamiento:

- 5087 filas u observaciones.
- 3198 columnas o características.
- La columna 1 es el vector de etiquetas. Las columnas 2 - 3198 son los valores de flujo en el tiempo.
- 37 estrellas exoplanetas confirmadas y 5050 estrellas no exoplanetas.

Conjunto de pruebas:

- 570 filas u observaciones.
- 3198 columnas o características.
- La columna 1 es el vector de etiquetas. Las columnas 2 - 3198 son los valores de flujo en el tiempo.
- 5 exoplanetas-estrellas confirmados y 565 no exoplanetas-estrellas.

■ Objetivo

El objetivo es entrenar un clasificador que sea capaz de determinar si las observaciones en el flujo de luz corresponden a un exoplaneta o no. Los modelos de clasificación se validarán y ordenarán en función de la métrica F1-score obtenida en la predicción del conjunto de pruebas por parte del clasificador. Los participantes solo tendrán acceso al conjunto de datos de entrenamiento.

■ Material

Podrás acceder a todo el material que necesitas pinchando en el siguiente link: <https://drive.upm.es/s/yvVDTDFoUpoq0qH>

RETO FORENSE

■ El asunto (Contexto)

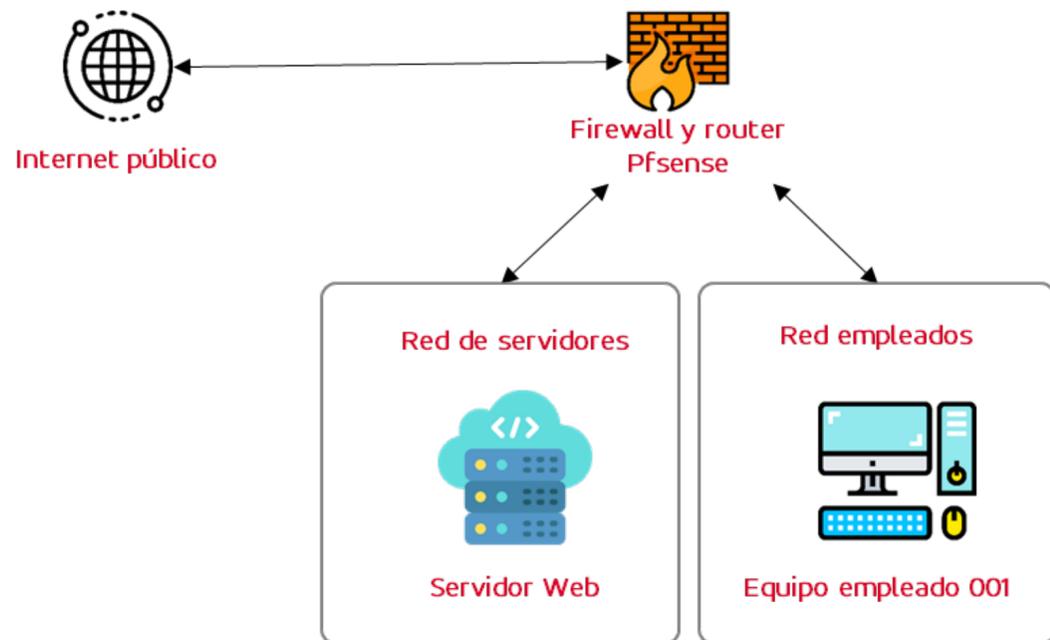
La empresa “Ay madre que me la han colao S.A.” celebra en pocos días su Junta general de accionistas donde se presentarán entre otros, los datos de sus resultados y por ende los beneficios a sus socios.

Sin embargo, se descubre que los documentos confidenciales de la empresa han sido vulnerados por un atacante que ha logrado acceder a la carpeta compartida por el consejo de administración y los datos de uno de sus documentos esenciales se han modificado, saboteando por tanto dicha junta.

■ La intervención Forense

El equipo forense Tras su Intervención, adquiere las evidencias de los equipos que han podido intervenir. Éstas son:

a) Esquema de la red atacada.



- b) Lista de Archivos de Logs Obtenidos.
 - a. Los relacionados con el sistema operativo Linux
 - b. Los relacionados con el sistema operativo Windows
 - c. Los relacionados con las reglas de un firewall.

Estos archivos de Logs (Evidencias), contienen la actividad de las diferentes máquinas que se han visto implicadas en el incidente.

■ La prueba

El objetivo es dar respuesta a las siguientes preguntas:

- a. ¿Qué ha sucedido?
- b. ¿Cuándo ha sucedido?
- c. ¿Cómo ha sucedido?
- d. ¿Quién o que ha causado el incidente?

■ Material

Podrás acceder a todo el material que necesitas a partir del 25 de Julio pinchando en el siguiente link:

<https://drive.google.com/drive/folders/1bzgDN6PD7xIFt6xYcgrDwlR4Vn0qZ-MK?usp=sharing>