Actividad grupal: Detección de anomalías y técnicas de agrupamiento

**Objetivos**

Mediante este trabajo se pretende que pongas en práctica la aplicación de los algoritmos de detección de anomalías u *outliers* y las técnicas de agrupamiento. El objetivo es que comprendas de forma práctica con un problema determinado los pasos que hay que realizar para detección automática de valores inusuales y por otro lado obtener grupos similares en los datos e interpretar los mismos.

**Descripción** **de la actividad**

Importa los datos del Abstract data set for Credit card fraud detection:

<https://www.kaggle.com/shubhamjoshi2130of/abstract-data-set-for-credit-card-fraud-detection>

Contiene la información necesaria sobre los datos

Realiza los siguientes pasos:

* Análisis descriptivo de los datos:
  + De las variables numéricas hallar el valor mínimo, el máximo, la mediana y la media.
  + De las variables categóricas, listar las diferentes categorías y hallar la frecuencia de cada una de ellas.
  + Hallar todas las correlaciones existentes entre las variables numéricas del conjunto de datos.
* Determinar el conjunto de modelización y el de validación.
* Aplicar una técnica de detección de anomalías. Se sugiere utilizar la de Isolation Forest. Si el estudiante está familiarizado o desea aplicar otra técnica es libre de hacerlo.
* Aplicar una técnica de agrupamiento: k-means o cluster jerárquico son las sugeridas, pero al igual que el punto anterior, el estudiante es libre de elegir otra.
* Calcular una métrica de evaluación para el ejercicio de detección de anomalías. Podría utilizar validación cruzada.
* Comentar los resultados obtenidos:
  + Analizar el rendimiento del modelo de detección de anomalías. ¿Son realmente útiles los modelos creados para el conjunto de datos propuesto?
  + Para el ejercicio de técnicas de agrupamiento analizar si los clúster creados tienen sentido o no.
  + ¿Qué ventajas y desventajas se encuentran en cada uno de los modelos trabajados?
* Otros comentarios que parezcan adecuados.
* Se puede usar R o Python.
* Se deben comentar los resultados obtenidos y el código.

**Extensión** **máxima:** 5 páginas en formato Word o PDF con el código adjunto aparte.

**Rúbrica**

|  |  |  |  |
| --- | --- | --- | --- |
| Actividad | Descripción | Puntuación máxima  (puntos) | Peso  % |
| Criterio 1 | Metodología | 6 | 60% |
| Criterio 2 | Resultados | 2 | 20% |
| Criterio 3 | Informe | 2 | 20% |
|  |  | **10** | **100 %** |

**Organización y gestión de equipos**

En el foro «Pregúntale al profesor» de la asignatura encontrarás un nuevo tema específico para la organización de equipos donde el profesor explicará todos los detalles.

Una vez cerrado el equipo de trabajo os podéis poner en contacto a través de vuestras cuentas @comunidadunir.net y comenzar a trabajar. Puedes ampliar la información sobre el trabajo en equipo, consultando los [Tutoriales de trabajo en grupo](http://tv.unir.net/secciones/3967/4883/0/0/0/0/).

**IMPORTANTE:** Aquellos **estudiantes que no comiencen su trabajo dentro de los 7 primeros días**, contados a partir del día de inicio de la actividad, **quedarán excluidos** de la actividad, no pudiendo tomar parte en ella. Se trata de una actividad colaborativa, por lo que unos estudiantes no pueden beneficiarse del trabajo que hayan realizado sus compañeros.

**Entrega de la actividad grupal**

Al finalizar la actividad grupal, todos los miembros del equipo entregarán la misma actividad a través del apartado «Envío de actividades» del aula virtual. El documento a entregar debe ir nombrado así:

APELLIDO1\_APELLIDO2\_NOMBRE\_Titulo\_actividad (sin tildes ni apóstrofes ni ningún otro carácter que pudiera resultar conflictivo).

Todos los miembros del equipo deben hacer la entrega en el aula virtual y deben adjuntar el mismo documento.

Indica en la actividad el nombre de todos los componentes del equipo y cumplimenta la siguiente tabla de valoración individual:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Sí | No | A veces |
| Todos los miembros se han integrado al trabajo del grupo |  |  |  |
| Todos los miembros participan activamente |  |  |  |
| Todos los miembros respetan otras ideas aportadas |  |  |  |
| Todos los miembros participan en la elaboración del informe |  |  |  |
| Me he preocupado por realizar un trabajo cooperativo con mis compañeros |  |  |  |
| Señala si consideras que algún aspecto del trabajo en grupo no ha sido adecuado |  |  |  |