

¿Es posible detectar el fraude energético mediante técnicas de minería de datos?

PROBLEMA

Las pérdidas eléctricas presentes en las redes de distribución están vinculadas especialmente al consumo de la electricidad que no es parcial o totalmente facturada al usuario debido a la ejecución de prácticas fraudulentas. Esta problemática produce grandes pérdidas económicas tanto para la empresa eléctrica como para el país. Así mismo origina un rápido deterioro de las redes e instalaciones.

OBJETIVO GENERAL

Evaluar técnicas de minería de datos aplicado a registros de medidores inteligentes para detectar fuentes de pérdidas eléctricas debido al fraude energético en redes de distribución.

PROPUESTA

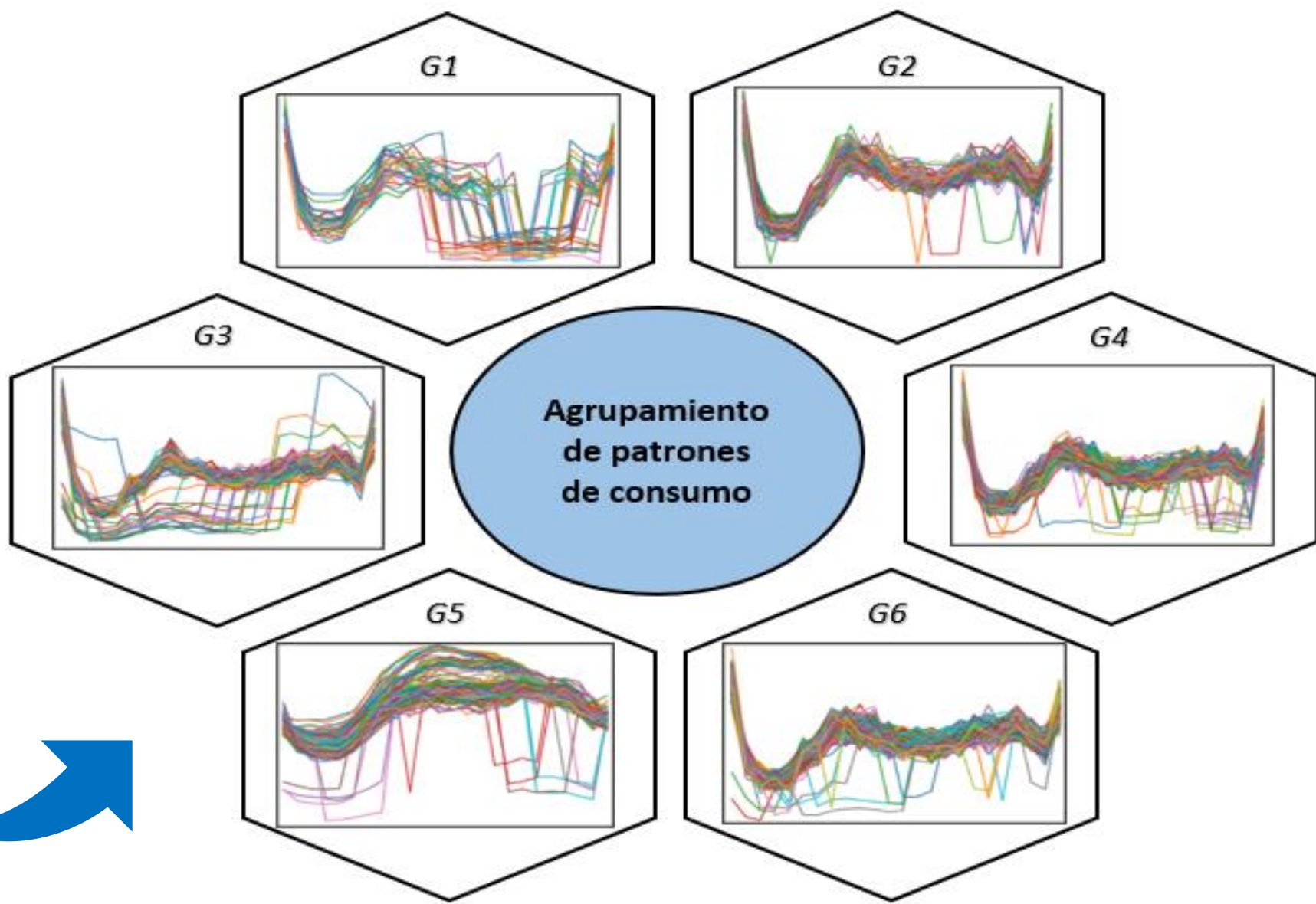
Se implementó el proceso de minería de datos a un conjunto de datos de la empresa eléctrica de New Jersey PSEG, recolectados desde medidores inteligentes los cuales representan el consumo eléctrico diario de 2555 usuarios; para identificar si es posible, anomalías, irregularidades o alteraciones en los patrones de consumo de energía. Se aplicó herramientas y algoritmos de aprendizaje de maquina no supervisado para agrupar (Clustering) y determinar a los consumidores que con mayor duración e intensidad estarían produciendo pérdidas a la red.

Adquisición de datos desde medidores inteligentes



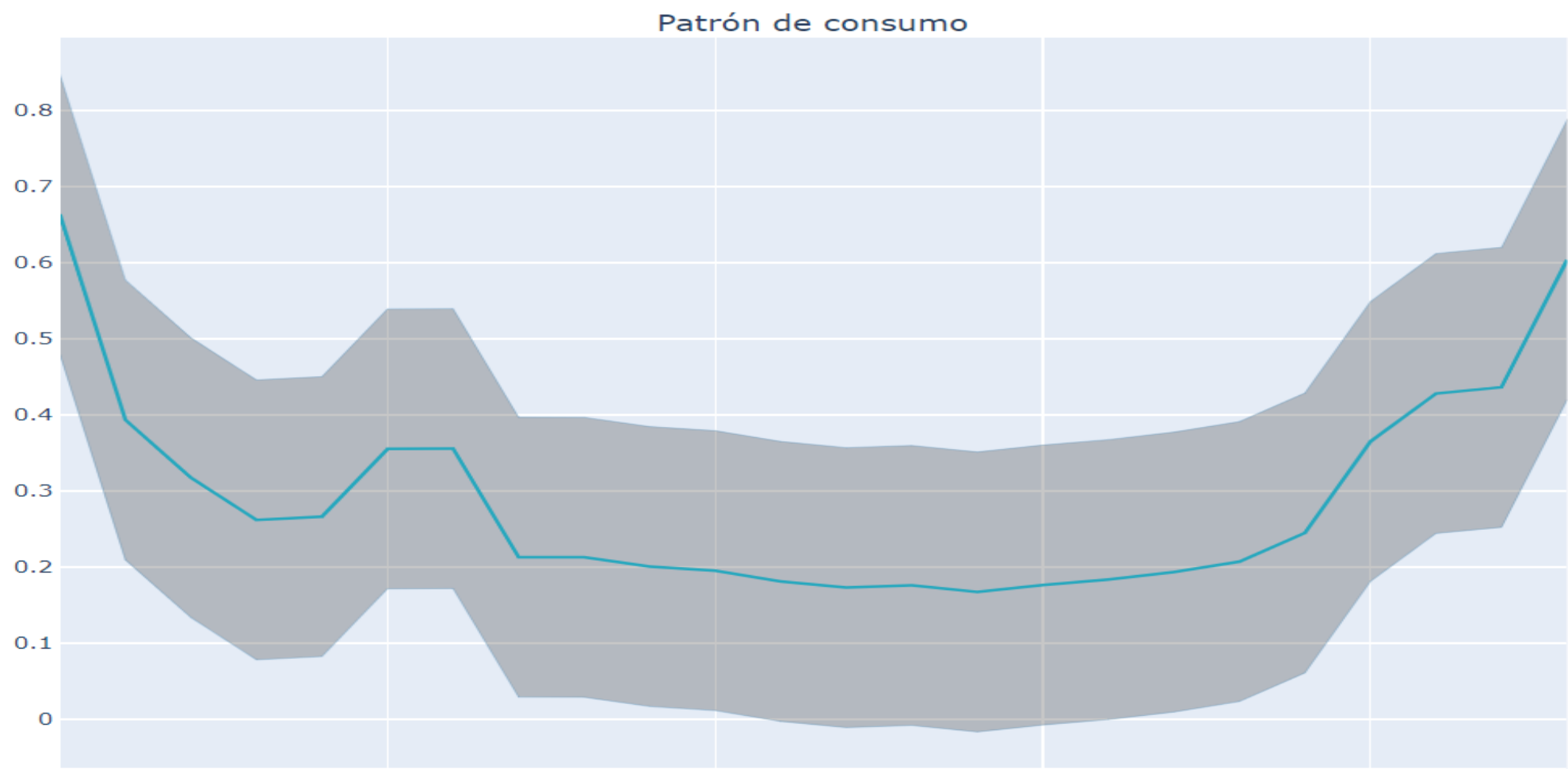
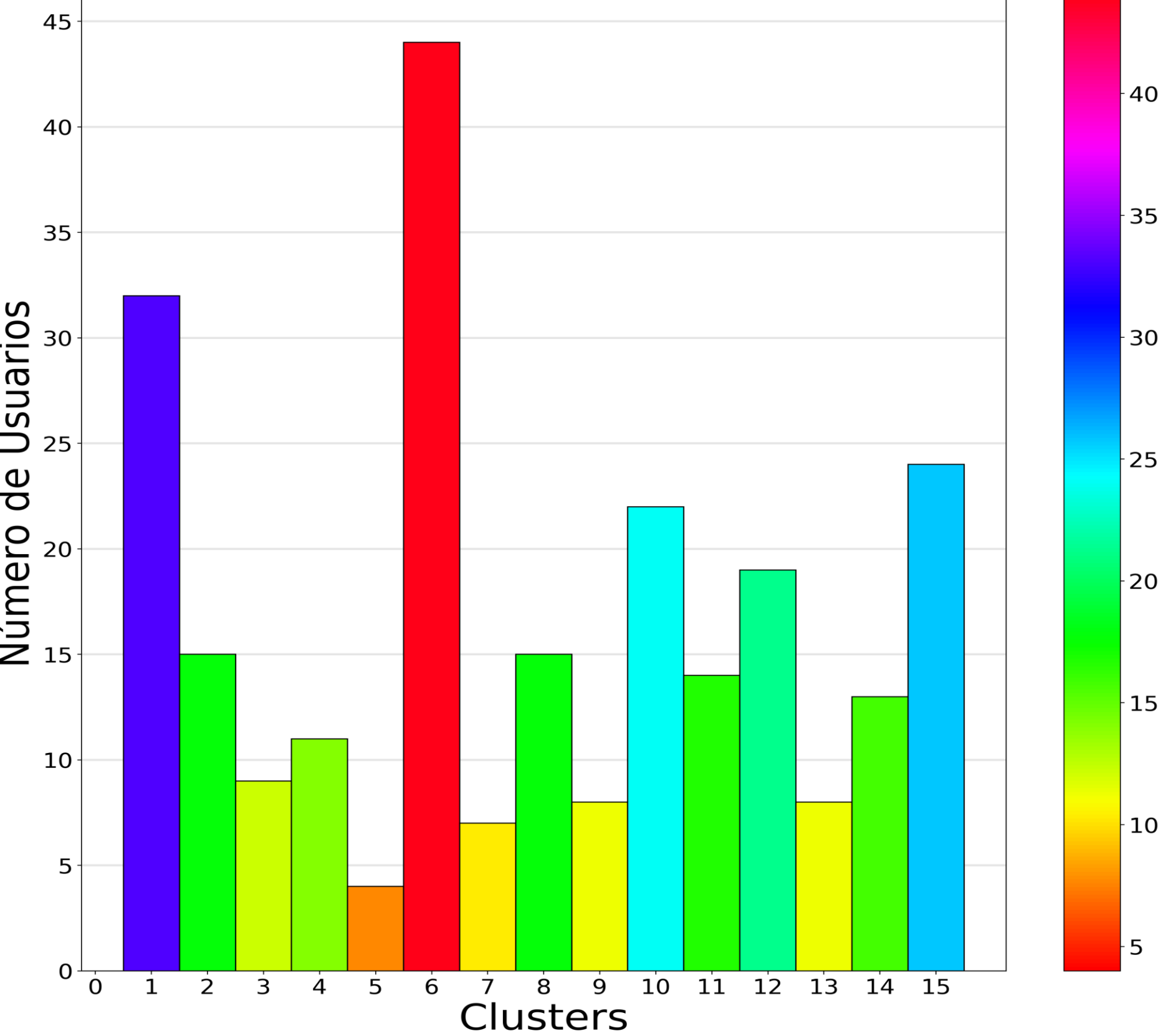
DATA MINING

Selección, Preprocesamiento, Transformación, Minería e Interpretación de datos.

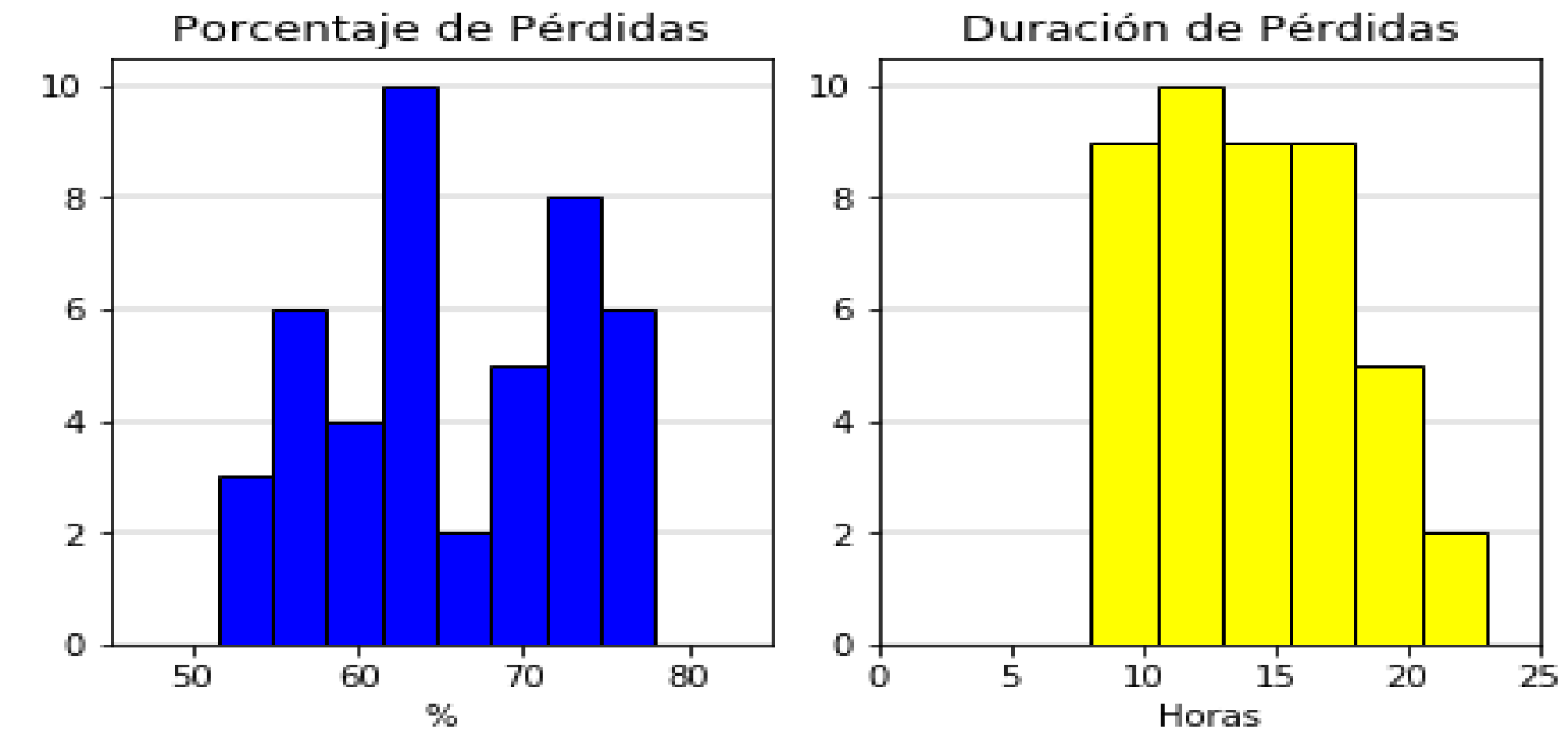


RESULTADOS

Usuarios Fraudulentos por Cluster



Cluster 6



- Aproximadamente el 10 % de los usuarios presentaron anomalías o irregularidades en su patrón de consumo.
- Se realizó el agrupamiento de usuarios teniendo como resultado 15 grupos o clusters en total, y en cada grupo se presenta al menos 4 posibles usuarios fraudulentos.
- Se identifica al cluster 6 como el grupo con mayor cantidad de posibles usuarios fraudulentos por cluster, con 44 usuarios en total; seguido del cluster 1 y 15, con 32 y 24 usuarios respectivamente.
- Precisamente el cluster 6 manifiesta el mayor porcentaje de pérdidas por cluster. Así mismo, durante más horas en el transcurso del día. Teniendo pérdidas durante 10 a 12 horas en promedio.

CONCLUSIONES

- Los resultados obtenidos muestran que es posible identificar y detectar a usuarios con defraudación mediante el análisis descriptivo de los patrones de consumo eléctrico.
- Teniendo en cuenta los resultados obtenidos, es de suma importancia contar con metodologías alternas que permitan la solución de problemas, para el caso en el que se tiene como objetivo general, la detección del fraude de energía eléctrica en sistemas de distribución.
- Se desarrolló un modelo de pérdidas para simular el comportamiento de usuarios fraudulentos debido a la carencia de este tipo de datos.
- Es de gran utilidad los métodos aplicados para modelos que necesiten entrenamiento para la predicción de usuarios fraudulentos mediante algoritmos de aprendizaje de máquina supervisados .