Reto KIO

cue-Bits

Problema

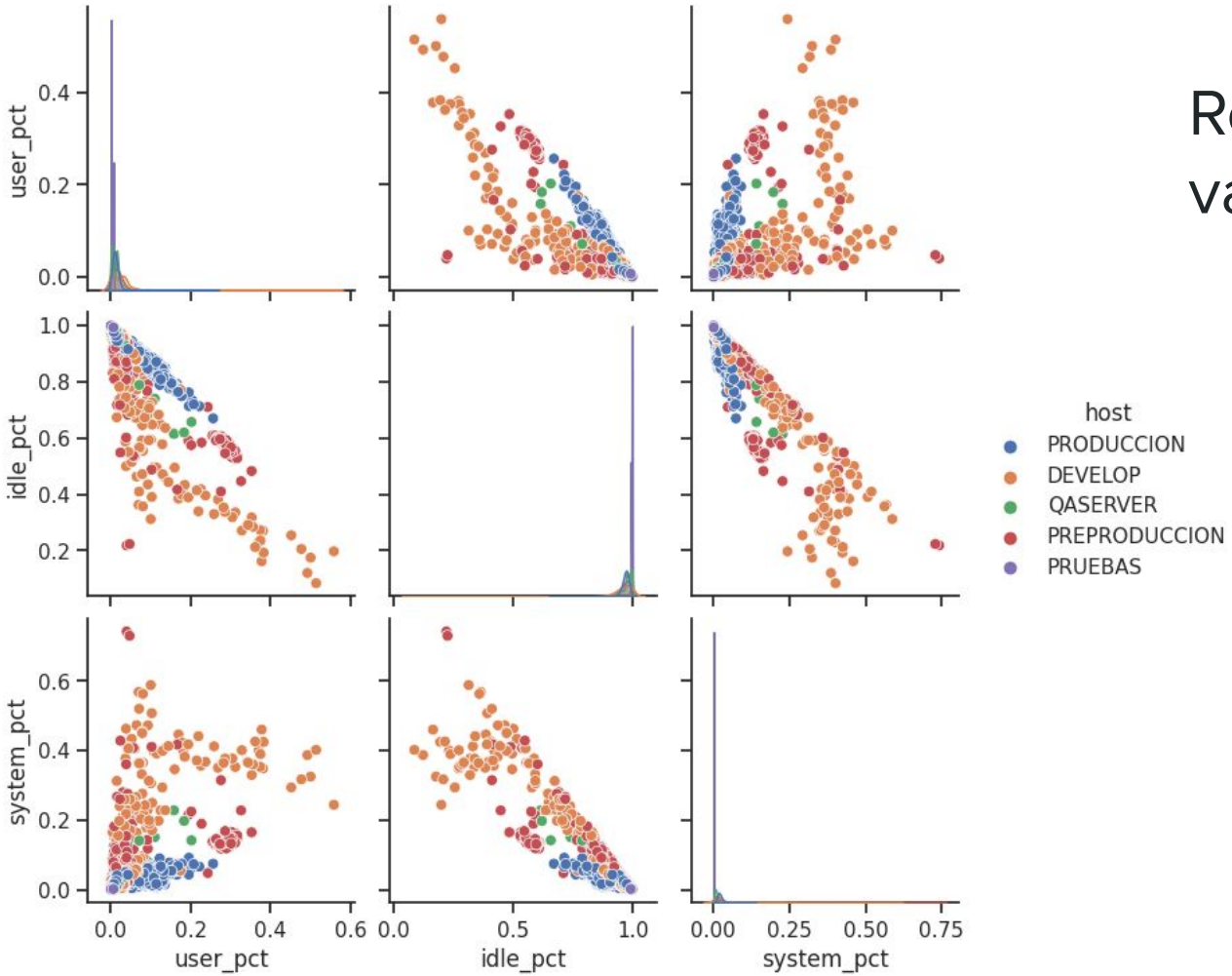
- ¿Cómo identificar un suceso anormal?
- ¿Cómo plantear escenarios probables?

Primeros pasos

- Acomodo de datos
- Muestreo aleatorio de datos (10% de los totales)
- Resampleo con intervalo de 10 minutos

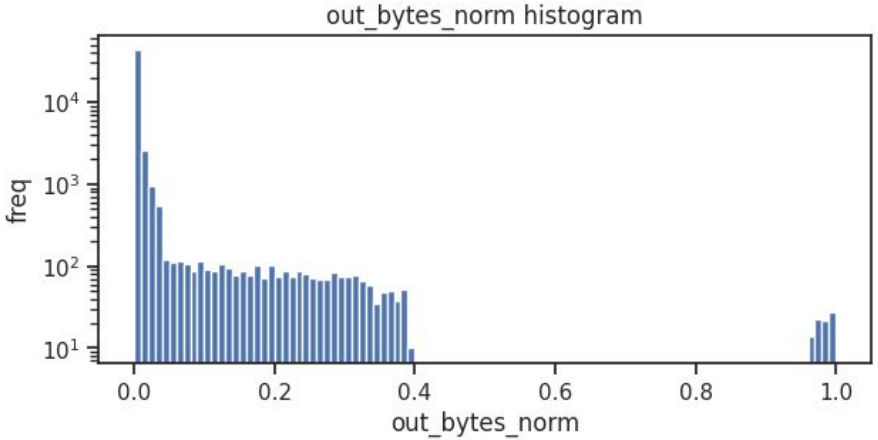
Fundamento matemático

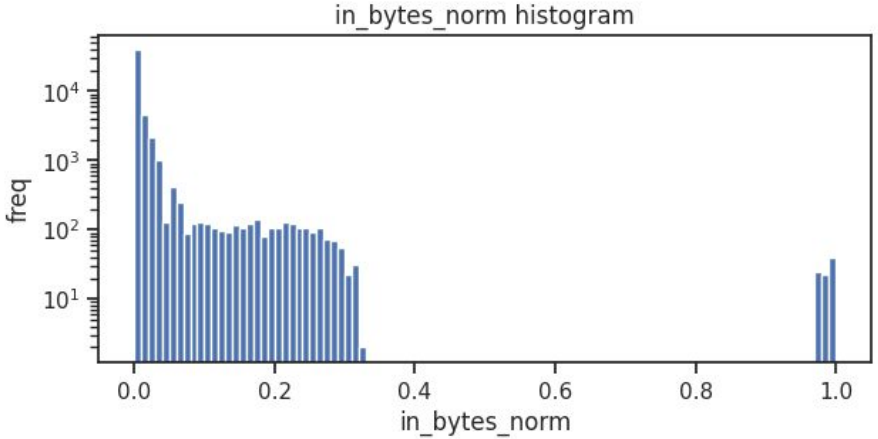
- Distribuciones con forma de función gamma
- Distribuciones sin forma conocida
  - Nos sugiere usar métodos no-paramétricos



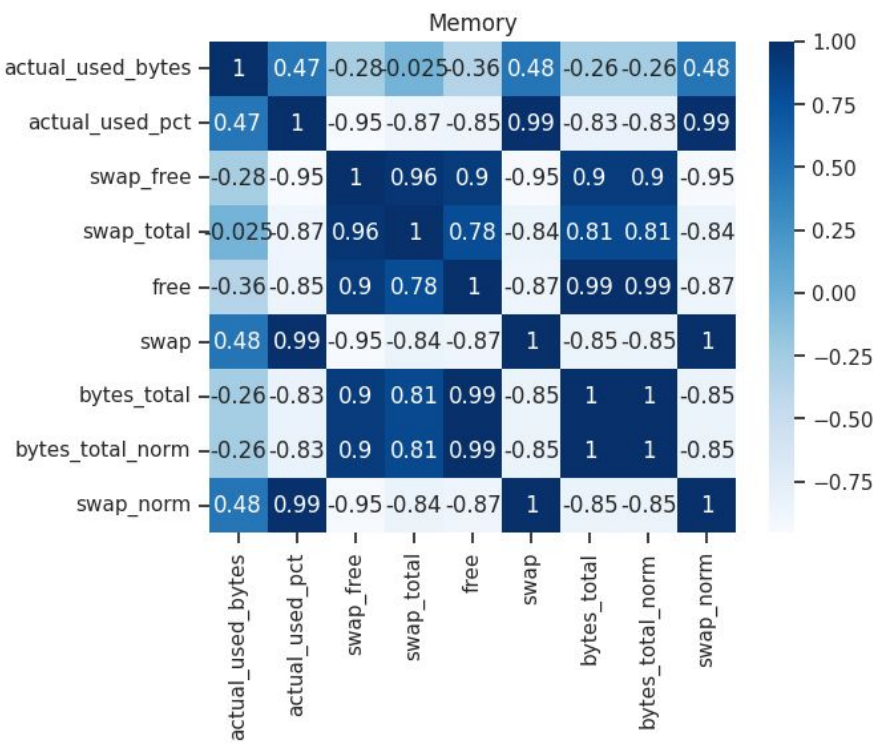
Relaciones entre variables

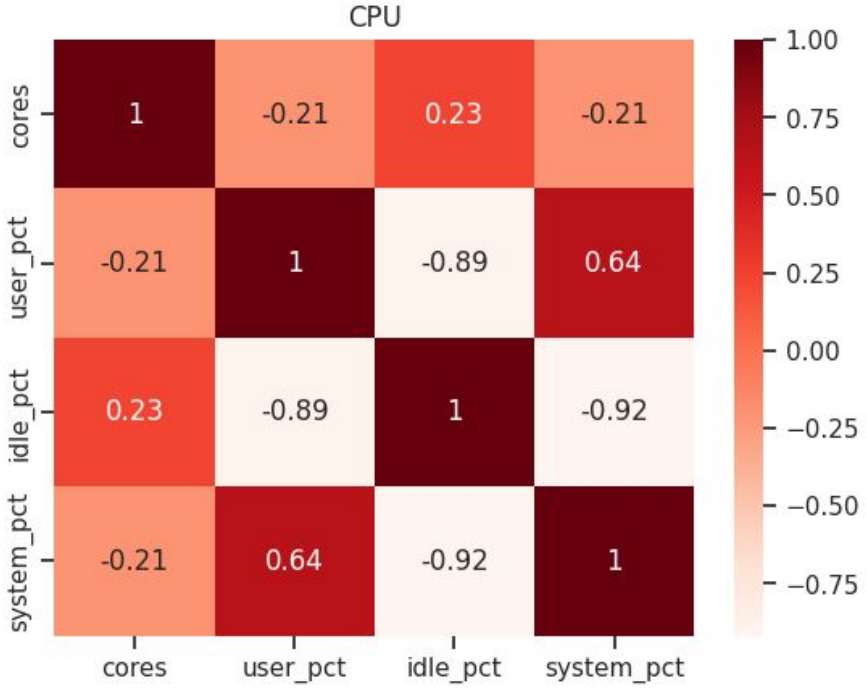
- Entre una misma categoría (cpu, mem, network)
- Mezclado (requiere resample)





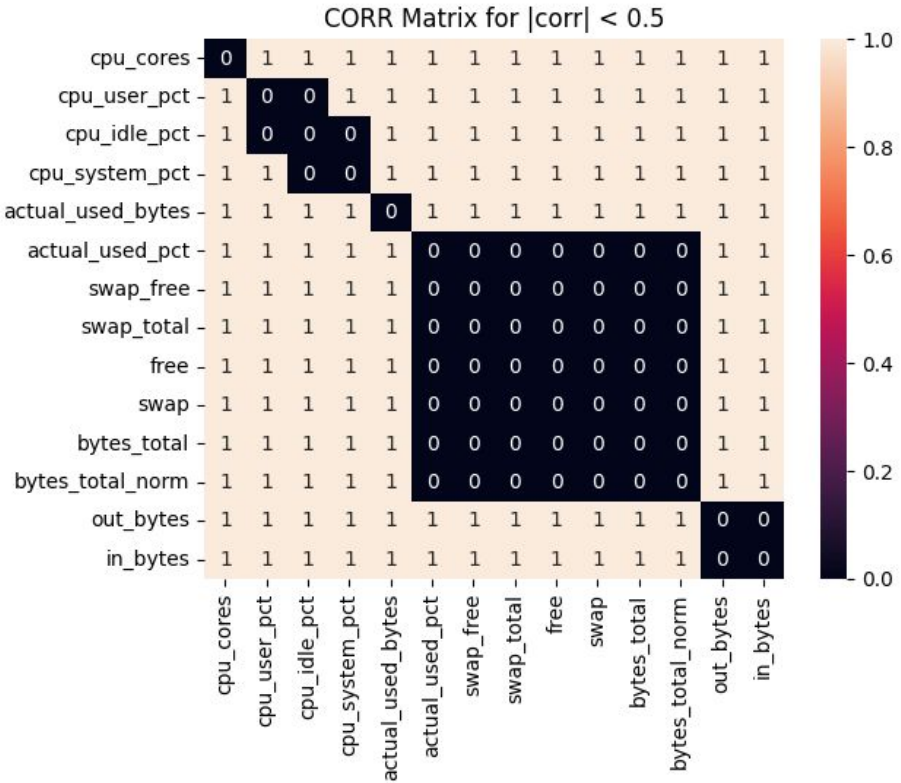
- Se observa con claridad que hay datos que se alejan de la norma.





Correlación ➜ indica qué variables podemos eliminar

Correlaciones por resample



- Encontramos relaciones entre indicadores

Indicadores propuestos

- Mem_libre_cpu: Memoria_libre / cpu_usada_pct

  Cómo se relaciona el cpu con la memoria libre

- Mem_usada_cpu: (memory_usuario_pct + memoria_sistema_pct) / cpu_usada_pct

  Cómo se relaciona el cpu con la memoria usada

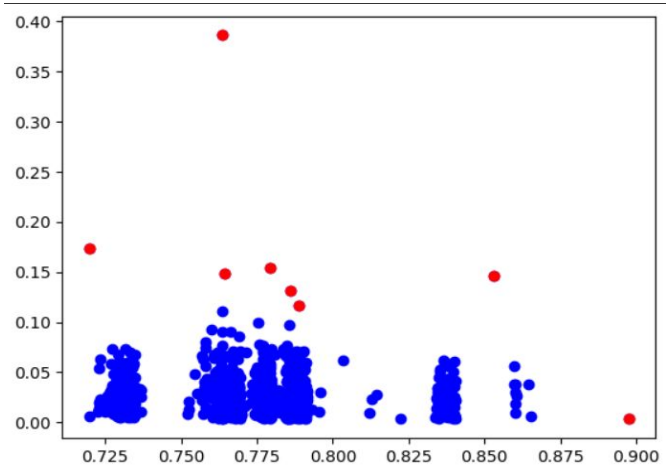
- Swap: Swap_libre / Swap_total
- usage: (user_pct + idle_pct) / system_pct

Modelos no supervisados

- KNN y métodos de densidad
- Probabilidad de ser declarado como anomalía
  - **Desviación del promedio de nearest neighbours**
  - Parámetros óptimos
  - Comparar con Isolation forest*

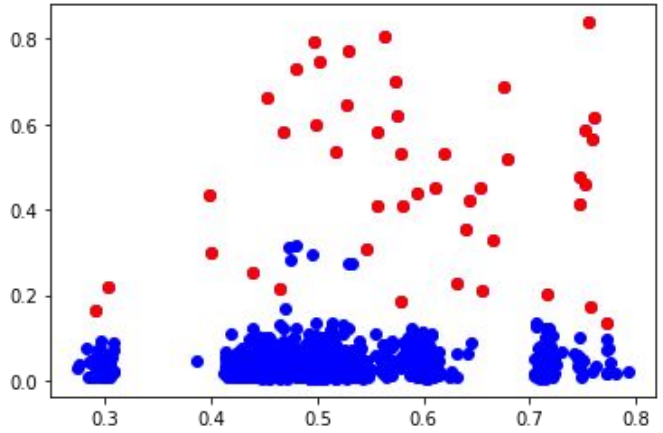
*A futuro

K (7) Nearest Neighbours



Cpu

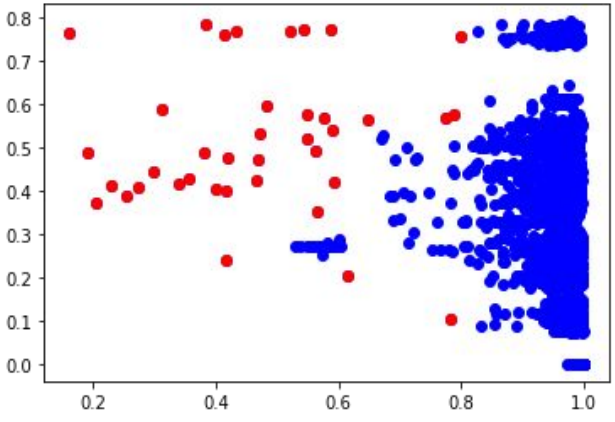
Mem libre



Mem uso

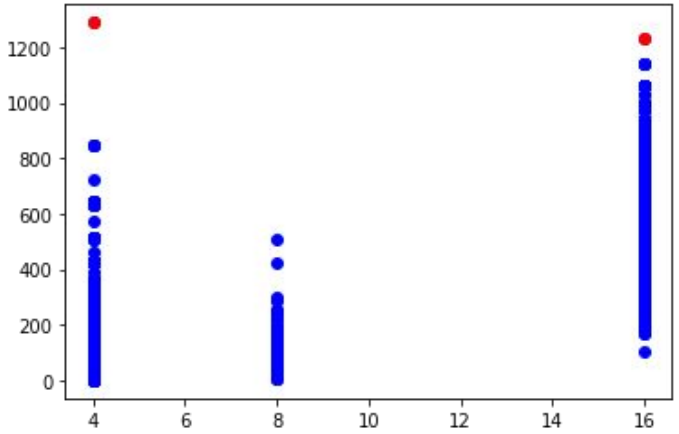
Cpu

K (7) Nearest Neighbours



Swap

Idle

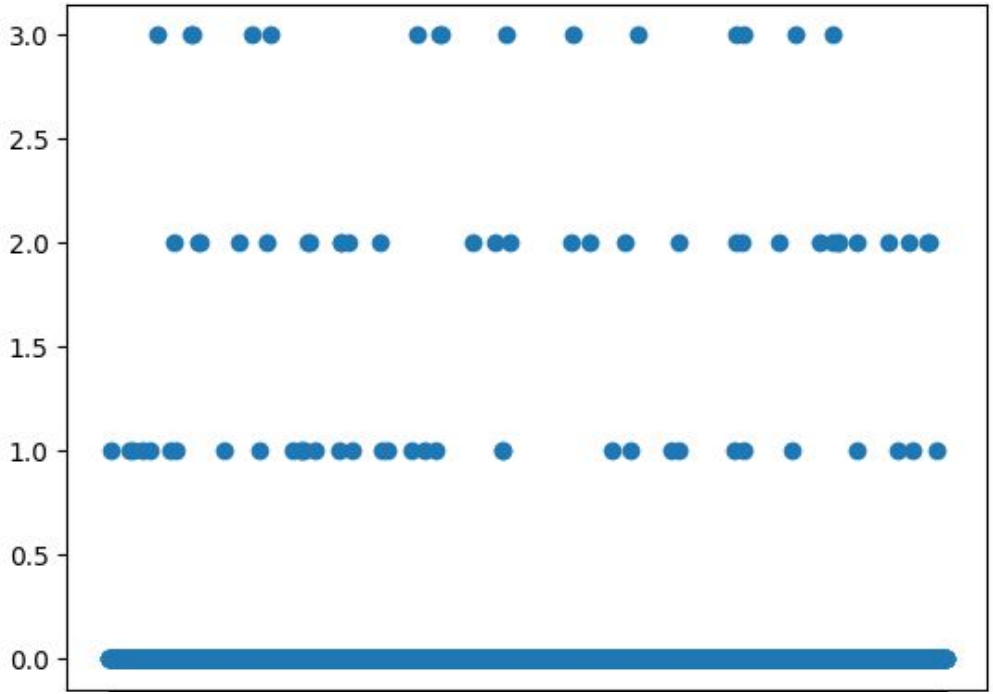


Cpu usage

Cores

Anomalías en el tiempo

Anomalías



Tiempo

Resultados

- Método KNN: más simple y adecuado a los indicadores
- Las anomalías tienen comportamiento aleatorio
- El indicador más útil fue el de memoria/cpu

Impacto

- Encontramos conexiones
- No podemos reportar qué tan bien se comporta con otros errores, como los no reportados

Bibliografía

- Lindh, F. (2019). Machine learning to detect anomalies in datacenter.
- Mehrotra, K. G., Mohan, C. K., & Huang, H. (2017). Anomaly detection principles and algorithms (Vol. 1). New York, NY, USA:: Springer International Publishing.