

Malware de detección de Android

El problema predictivo que se busca resolver es la detección de malware en dispositivos Android a través de técnicas de aprendizaje automático.

enlace

El conjunto de datos utilizado sería el proporcionado en el enlace de Kaggle:

<https://www.kaggle.com/datasets/subhajournal/android-malware-detection>

Este dataset contiene características extraídas de 5560 aplicaciones de Android, donde se han etiquetado como malware o no malware.

Métricas de desempeño

- Métricas de machine learning: precisión (accuracy), sensibilidad (recall), especificidad (specificity) y F1-Score.
- Métricas de negocio: tiempo de respuesta, tasa de falsos positivos, tasa de falsos negativos y costo asociado a la implementación de la solución.

Criterio de desempeño

El desempeño deseable en producción sería lograr una alta precisión (accuracy) en la clasificación de aplicaciones de Android como malware o no malware, al mismo tiempo que se logra un bajo costo asociado a la implementación de la solución y una tasa de falsos positivos y falsos negativos aceptable para el negocio. En este sentido, un valor de precisión por encima del 95% y una tasa de falsos positivos y falsos negativos inferior al 5% podrían ser indicadores de un buen desempeño en producción.