Proyecto 1

Inteligencia de negocios

Gregorio Ospina 201631760, andres Varon, Juan sebastian vaca

2019-2

.

|  |  |  |
| --- | --- | --- |
| **Oportunidad/problema Negocio** | Seguridad de los Alpes, según nos cuenta el enunciado, es una compañía de seguridad informática que se dedica a desarrollar sistemas para la prevención de ataques y amenazas a una gran variedad de ataques cibernéticos. Con base en la tarea que nos han puesto los directivos de esta empresa, podemos afirmar que el objetivo de esta es ser capaces de desarrollar sistemas que, basado en la minería de datos pueda detectar ataques de manera temprana basado en patrones en estos ataques. Por ende, la medida de éxito de nuestra tarea, desde un punto de vista de negocio es si pudimos o no encontrar una correlación entre algún comportamiento de navegación y un ataque en específico. | |
|
|
|
|
|
|
|
|
| **Descripción del requerimiento desde el punto de vista de minería de datos** | Los recursos que nos compartió Seguridad de los Alpes son unos logs de comportamiento en la red, con una cantidad de columnas y descriptores amplía que tiene marcado cuales de estas actividades fueron ataques y de cual tipo. Basado en los instructivos que nos dieron se puede derivar que las metas y los objetivos para esta tarea de minería seria utilizar la minería de datos para encontrar patrones en estos ataques, para esto lo primero que se debería hacer es intentar entender el tipo de ataques, entender cual es su relación con los datos para ver si somos capaces de armar categorías, esto se debería hacer con una tarea de clustering utilizar de manera correcta tanto un árbol de decisión para ver cuales son las variables y los comportamientos que son indicativos de un ataque llegando a un nivel de aciertos mayor a 75% con bajos niveles de falsos-positivos y positivos-falsos, ademas de eso seria pertinente hacer un análisis por afinidad utilizando reglas de asociación consiguiendo métricas de confidence mayores a 0.5. | |
|
|
|
|
| **Detalle sobre la actividad de minería de datos** | | |
| **Tarea** | **Tecnica** | **Algoritmo y parámetros** |
| Predictivo | Arbol de decision | C4.5 |
| Predictivo | Reglas de asociacion | Apriori |
| Clasificación | Clustering | K-means |