
 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Semestre: 2019-20</p>	
--	---	---

## ENUNCIADO PRIMER PROYECTO

### **Objetivo.**

- Aplicar la metodología de minería de datos para la construcción de soluciones alineadas con los objetivos del negocio en un contexto de aplicación.



### **Descripción.**

El Internet de las Cosas (IoT, por sus siglas en inglés) está desempeñando un papel cada vez más importante en nuestras actividades cotidianas, al habilitar la conexión de los objetos que nos rodean con servicios digitales. Esto está conduciendo a cambios significativos en los ámbitos laboral, productivo y personal. Ya son muchas las aplicaciones IoT, entre ellas: la automatización del hogar (*Smart Home*), que convierte una casa en “inteligente” al facilitar el control de los dispositivos domésticos mediante la voz de los usuarios; las ciudades inteligentes (*Smart Cities*) que, a través del uso de sensores adecuados, apoya la resolución de problemas como la delincuencia y el tráfico automotor; la agricultura de precisión, que utiliza sensores para la monitorización de cultivos, lo cual permite identificar patrones inusuales y condiciones del suelo que al ser analizados junto con información meteorológica, mejora la producción agrícola; Retail, sector económico que ya está usando sensores con tecnología *bluetooth* que se comunican con dispositivos inteligentes, para suministrar tanto información de localización en el interior de las tiendas como ofertas a los usuarios.

Últimamente, IoT se ha visto afectado por una variedad de *botnets* (grupos de computadoras infectadas y controladas por un atacante de forma remota), los cuales intentan explotar vulnerabilidades en los protocolos de aplicación (como DNS y HTTP) que interactúan directamente con los sistemas, provocando fugas de datos y violaciones de seguridad. Una solución a este tipo de ataques la representan los sistemas de detección de intrusos. Estos pueden identificar, de manera automática, eventos maliciosos contra protocolos de comunicación utilizados en redes IoT y rastrear actividades sospechosas.

Seguridad Alpes, una empresa que desarrolla sistemas para la identificación temprana de estas amenazas, desea aplicar la minería de datos en la construcción de una solución para la detección de intrusos, en particular los ataques de *botnets* y sus huellas. Para ello cuenta con un conjunto de datos<sup>1</sup> conformado por paquetes de red, que describen actividades normales y comportamientos de nueve tipos de ataques, a saber: *Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode* y *Worms*. Los diferentes registros han sido caracterizados por 38

<sup>1</sup> Datos UNSW-NB 15, creados por la herramienta IXIA PerfectStorm en el Cyber Range Lab del Centro Australiano de Seguridad Cibernética (ACCS).

 <b>Universidad de los Andes</b> Colombia	<b>Ingeniería de Sistemas y Computación</b> <b>Pregrado</b> <b>ISIS-3301 – Inteligencia de Negocios</b> <b>Semestre: 2019-20</b>	
---	---	---

atributos y una etiqueta de clase (0: normal, 1: ataque). Además, Seguridad Alpes desea identificar patrones recurrentes que puedan caracterizar los grupos de actividades maliciosas.

### **Entregas.**

a) Documento que describa cómo fue aplicada la metodología de minería de datos y los resultados obtenidos en cada una de las fases:



1. **(15%) Comprensión del negocio y enfoque analítico.** Definición de los objetivos y criterios de éxito desde el punto de vista del negocio. Determinación de las tareas de minería de datos que se considera son las adecuadas para alcanzar los objetivos del negocio. Descripción de cómo el requerimiento de negocio es resuelto con el o los requerimientos de minería de datos propuestos, para lo cual debe diligenciar la tabla que se presenta a continuación:

Oportunidad/problema Negocio		
Descripción del requerimiento desde el punto de vista de minería de datos		
Detalles de la actividad de minería de datos		
Tarea	Técnica	Algoritmo y parámetros utilizados

2. **(30%) Comprensión de los datos y preparación de los datos.** Perfilamiento de los datos. Análisis de la calidad de los datos. Tratamiento de los datos sobre la base del conocimiento del dominio y de los algoritmos seleccionados para resolver las tareas.
3. **(35%) Modelado y evaluación.** Aplicación de las técnicas de minería de datos para la construcción de los modelos y presentación de los resultados de la evaluación.
4. **(20%) Análisis de resultados.** Descripción de los resultados obtenidos que permita a la organización comprenderlos, haciendo énfasis en el análisis de las medidas arrojadas por los modelos utilizados y cómo aportan en la consecución de los objetivos del negocio. Incluir posibles estrategias que la organización debe plantear relacionadas con los resultados obtenidos en los modelos y una justificación de por qué esa información es útil para ellos.

**Importante: justifique cada decisión tomada en cada una de las fases.**

b) Conjunto de datos resultado de la fase de comprensión y preparación de los datos.

 <p>Universidad de los Andes Colombia</p> <p>Acreditación institucional de alta calidad <b>10 años</b> MinEducación Resolución 385 del 9 de enero de 2015</p>	<p>Ingeniería de Sistemas y Computación</p> <p><b>Pregrado</b></p> <p>ISIS-3301 – Inteligencia de Negocios</p> <p>Semestre: 2019-20</p>	
--	---	---

c) Deben entregar los modelos analíticos desarrollados (Flujo de trabajo construido en knime o su equivalente si fue utilizado otro software) y todo lo relacionado con el proyecto para poder ejecutarlos en la sustentación.

## EVALUACIÓN

- El proyecto se realiza en grupos de 3 estudiantes.
- El documento a entregar tiene máximo 8 páginas (sin incluir portada, tabla de contenido, ni referencias), a una columna y con letra arial, tamaño 12.

La fecha máxima de entrega es el **viernes 20 de septiembre a las 10:00 p.m.**