Manual Didáctico de la asignatura: "Introducción al Aprendizaje Automático para Ciberseguridad"

Máster Universitario en Ciberseguridad

Código de asignatura: 31109097
Curso: 2019-20

Orientaciones para el estudio,
erratas, aclaraciones sobre contenido del libro base,
y material complementario.

José Ramón Álvarez Sánchez - Enrique J. Carmona Suárez
Dpto. de Inteligencia Artificial - E.T.S.I. Informática - UNED

Versión del 2020-02-10
(la versión actualizada estará siempre en el curso virtual de la asignatura)

UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA

Índice

Lista de cambios y correcciones en este manual

Se listan a continuación los cambios realizados en cada versión respecto a la anterior (orden de más reciente a más antigua)[1] de este manual. Solamente se incluyen aquí los cambios o modificaciones significativos, relevantes respecto al contenido. No se incluyen en esta lista las erratas tipográficas, ortográficas o de puntuación no-significativas ya corregidas que no induzcan a error ni afecten a la comprensión.

2020-02-10 Versión inicial para el comienzo del curso

- Se incluye la información sobre las orientaciones para el estudio específicas de cada tema, erratas, aclaraciones y comentarios sobre el libro base, y también material complementario.

—

La composición tipográfica de este texto se ha realizado por sus autores, utilizando LyX y LaTeX.

Nota: *En este documento, todas las palabras de género masculino que se apliquen a personas, se deben entender referidas en genérico a cualquiera de los sexos indistintamente.*

[1] En el curso virtual de la asignatura siempre estará accesible la última versión más moderna.

Información general

Asignatura: Introducción al Aprendizaje Automático para Ciberseguridad (Código: 31109097)
(de la titulación: "Máster Universitario en Ciberseguridad")

Equipo docente: José Ramón Álvarez Sánchez <jras@dia.uned.es> (coordinador)
Enrique J. Carmona Suárez <ecarmona@dia.uned.es>

Toda la información básica sobre la contextualización, recomendaciones y requisitos previos, competencias, resultados, metodología, plan de trabajo y **evaluación** de la asignatura está especificada en la **Guía de Estudio** que está accesible en su versión completa desde un enlace en el curso virtual de la asignatura:

https://2020.cursosvirtuales.uned.es/dotlrn/posgrados/asignaturas/31109097-20/

La versión pública de la Guía de la Asignatura también está accesible antes del inicio de curso desde el listado de asignaturas de la titulación en el portal de la UNED. Es importante que se consulte dicha guía, en especial la información sobre la forma de **evaluación** de la asignatura.

Este manual didáctico contiene indicaciones específicas para el estudio y preparación de la asignatura. En este documento, se proporciona una descripción detallada del temario y de los correspondientes capítulos del libro base, así como erratas, comentarios y aclaraciones sobre cada tema, y también posibles materiales y recursos complementarios adicionales o para ampliación de conocimientos.

Esquema de contenidos

- Tema 1: La Inteligencia Artificial en ciberseguridad:
  - Aprendizaje Automático en el dominio de la seguridad.
- Tema 2: Algoritmos de agrupamiento:
  - Algoritmo K-Means.
  - Algoritmo DBSCAN.
  - Aplicación a análisis de seguridad.
- Tema 3: Clasificación:
  - Regresión Logística.
  - Algoritmos de Árboles de Decisión.
  - Aplicación a clasificación de ataques.
- Tema 4: Modelos probabilísticos:
  - Clasificador Bayesiano ingenuo (naïve Bayes).
  - Modelo de mezclas Gaussianas.
  - Aplicación en clasificación y agrupamiento.
- Tema 5: Arquitecturas de aprendizaje profundo (Deep Learning):
  - Redes Neuronales Recurrentes (Long Short-Term Memory).
  - Redes Neuronales Convolucionales (CNN).

Material de estudio

El libro base utilizado para los contenidos de esta asignatura se citará en este Manual Didáctico de la siguiente forma: The Cylance Data Science Team [2017] y sus datos son los siguientes:

> "Introduction to Artificial Intelligence for Security Professionals"
> *The Cylance Data Science Team* (Cylance Press, 2017)
> (PDF gratuito descargable dentro del repositorio https://www.cylance.com/intro-to-ai
> o solicitándolo en: https://pages.cylance.com/en-us-introduction-to-ai-book.html
> y en formato electrónico para Kindle de Amazon o bien para Nook de Barnes & Noble).

La mayor parte de los contenidos de los temas de la asignatura se corresponden con los capítulos de este libro base. En cada capítulo, después del planteamiento teórico, se pueden encontrar ejemplos de aplicaciones prácticas a casos reales, que ayudarán a entender los conceptos explicados en los contenidos. En algunos capítulos los ejercicios y ejemplos resueltos contienen información más detallada que amplía los contenidos principales. En el libro base también se proporcionan ejemplos con datos y código que están disponibles en un repositorio de GitHub accesible desde la dirección: `https://www.cylance.com/intro-to-ai`. Al final de cada capítulo del libro base, también se incluye un breve resumen de los contenidos expuestos en el mismo.

Adicionalmente a estas orientaciones que ya van incorporadas en el propio libro base, se especifican en este Manual Didáctico (dentro de las orientaciones de cada tema) los siguientes apartados:

- Introducción, objetivos y contexto del tema.
- Contenidos: especificación del capítulo del libro base y otro posible material adicional necesario.
- Recomendaciones de estudio.

Además de los ejemplos y ejercicios presentados en cada capítulo del libro base, en este Manual Didáctico pueden añadirse sugerencias para algunas otras actividades voluntarias no evaluables, tanto de tipo teórico como práctico, así como también otras fuentes para extender conocimientos sobre cada tema. Las referencias que se citan en este Manual Didáctico están recopiladas al final en la sección de Referencias de la pág. 7.

Las actividades evaluables, correspondientes a las pruebas de evaluación continua y el trabajo final, se especificarán y detallarán en los enunciados correspondientes que estarán disponibles en el curso virtual de la asignatura, que además será el único medio para su entrega.

Aparte del libro base y de las lecturas recomendadas específicas de cada capítulo, el estudiante puede encontrar otros ejercicios y explicaciones alternativas o más extensas sobre algunos de los temas en los libros: [Freeman and Chio, 2018, Deisenroth et al., 2020, Dua and Du, 2011, Mena, 2011, Russell and Norvig, 2004, Nilsson, 2001].

A continuación, se incluyen las orientaciones de estudio, junto con erratas, avisos y comentarios, específicos en apartados numerados correspondientes a cada tema en las páginas de este Manual Didáctico: Tema 1 en la página siguiente, Tema 2 en la página siguiente, Tema 3 en la página 6, Tema 4 en la página 6 y Tema 5 en la página 7.

1. La Inteligencia Artificial en ciberseguridad.

1.1. Introducción, objetivos y contexto del tema

Se trata de dar una visión general de lo que habitualmente se conoce como Inteligencia Artificial, para centrar la atención sobre la aplicación de algunos métodos más concretos de la misma en los problemas de ciberseguridad. En este tema también se hace una presentación del resto de temas.

1.2. Contenidos

Los contenidos teóricos de este tema se corresponden con la introducción del libro base [The Cylance Data Science Team, 2017]. En este tema se presentan los diferentes percepciones habituales sobre la Inteligencia Artificial en general, y también el planteamiento más específico de las técnicas que se encuadran en el Aprendizaje Automático, que en inglés se denomina Machine Learning, aplicadas en el dominio de la ciberseguridad.

Adicionalmente a los contenidos en la introducción del libro base, también es importante recordar que los métodos de aprendizaje automático se basan en datos. Los datos disponibles deberían permitir resolver el problema de forma significativa. Dados unos datos, se pueden aplicar diferentes tipos de aprendizaje automático. La agrupación (clustering) es útil para encontrar similitudes en el conjunto de datos, para descubrir tendencias u obtener otros conocimientos. Usando datos etiquetados se puede utilizar un método de clasificación para construir modelos de predicción. Habitualmente es necesario transformar de alguna manera los datos en bruto para puedan ser utilizados por los algoritmos de aprendizaje automático. El proceso típico es extraer características de los datos, y convertir esas características en vectores. Los vectores se introducen en los algoritmos de aprendizaje automático para el entrenamiento o ajuste de parámetros.

1.3. Recomendaciones de estudio

Después de la lectura y comprensión de los contenidos teóricos en el capítulo correspondiente, es muy interesante preparar el entorno y herramientas para poder probar los ejemplos prácticos del resto de capítulos. Por ello se recomienda preparar un entorno de ejecución del lenguaje Python y que será el utilizado en los ejemplos. En este sentido se deben centrar los esfuerzos también en probar el entorno en Python de aprendizaje automático: Scikit-Learn `https://scikit-learn.org`, que está usualmente disponible en casi todos los sistemas como un paquete o software adicional.

Después es muy recomendable acceder al repositorio de código y ejemplos en GitHub: `https://www.cylance.com/intro-to-ai`, para familiarizarse con su estructura y contenidos que se usarán en el resto de capítulos.

2. Algoritmos de agrupamiento

2.1. Introducción, objetivos y contexto del tema

Los algoritmos de agrupamiento utilizan funciones de similaridad para agrupar trozos de información. Estos algoritmos son de gran utilidad para identificar tipos de comportamientos y detectar anomalías en ciberseguridad. En el tema también se estudia la alta dimensionalidad y las características, que son esenciales para la comprensión de los temas posteriores.

2.2. Contenidos

Este tema se corresponde completamente con el capítulo 1 con el título "Clustering: Using the K-Means and DBSCAN Algorithms" del libro base [The Cylance Data Science Team, 2017]. En el capítulo se cubren con cierta profundidad los algoritmos de agrupación k-medias (k-means), del tipo basado en centroides, y DBSCAN (Density-based spatial clustering of applications with noise), del tipo basado

en densidad. Ambos son de los tipos básicos de algoritmos de agrupamiento más conocidos y usados. En el capítulo del libro base se utiliza un ejemplo práctico de la agrupación de registros HTTP para identificar grupos de comportamientos así como comportamientos anómalos.

2.3. Recomendaciones de estudio

Una vez estudiado el contenido teórico, es importante que el estudiante intente seguir y comprender bien el ejemplo práctico desarrollado al final del capítulo.

3. Clasificación

3.1. Introducción, objetivos y contexto del tema

Los algoritmos de clasificación son de los más importantes en aprendizaje automático y muy útiles para ayudar a decidir si un conjunto de datos indica un posible ataque, o determinar si un archivo es malicioso o benigno, etc. En el tema también se estudia la diferencia entre aprendizaje supervisado y no-supervisado, así como entre clasificadores lineales y no-lineales. Además se incluye el uso de las matrices de confusión, de la precisión, y de la sensibilidad (exhaustividad o recall) para evaluar y validar los modelos producidos por los algoritmos de clasificación.

3.2. Contenidos

Los contenidos de este tema se corresponden por entero con el capítulo 2 del libro base [The Cylance Data Science Team, 2017], que se titula: "Classification: Using the Logistic Regression and Decision Tree Algorithms". En este capítulo, aparte del análisis paso a paso de los algoritmos de regresión logística y de árboles de decisión para clasificación, también se incluye un ejemplo que se centra en la identificación de los paneles de comando y control (C&C) de botnets, utilizando esos algoritmos para poder hacer esa clasificación en muy pocas solicitudes, minimizando así el ruido que podría notar un operador del comando y control de la botnet.

3.3. Recomendaciones de estudio

Después del estudio de lo contenidos teóricos es muy importante el seguimiento detallado del ejemplo planteado en el capítulo.

4. Modelos probabilísticos

4.1. Introducción, objetivos y contexto del tema

Se pueden utilizar modelos probabilísticos básicos también para agrupamiento y para clasificación en el aprendizaje automático. Estas son herramientas básicas disponibles en Aprendizaje Automático que, por tanto, también serán muy útiles en aplicaciones a ciberseguridad.

4.2. Contenidos

El estudio de este tema se corresponde íntegramente con el capítulo 3, con el título "Probability", del libro base [The Cylance Data Science Team, 2017]. En este capítulo, después de la explicación paso a paso de los algoritmos probabilísticos básicos (Bayesiano ingenuo o naïve Bayes, y modelo de mezclas Gaussianas), se aborda el clásico problema del aprendizaje automático en seguridad: la identificación de los mensajes de spam, aunque aplicando un pequeño giro al problema, para identificar los mensajes SMS de spam.

4.3. Recomendaciones de estudio

Aparte de los aspectos teóricos presentados en el capítulo 3 en el libro base, es importante que cada estudiante repase y comprenda el ejemplo que se plantea al final del mismo.

4.4. Ampliación de conocimientos

Un estudio más matemático de la estimación de densidad con modelos de mezclas Gaussianas se puede encontrar en el capítulo 11 del libro [Deisenroth et al., 2020].

5. Arquitecturas de aprendizaje profundo (Deep Learning)

5.1. Introducción, objetivos y contexto del tema

Las herramientas de redes neuronales (convolucionales y long short-term memory) pueden utilizan capacidades de computación muy elevadas para obtener buenos resultados en ciberseguridad. Es una herramienta poderosa que conviene entender y usar, pero puede ser bastante compleja.

5.2. Contenidos

Se estudia este tema en el capítulo 4, "Deep Learning", completo del libro base [The Cylance Data Science Team, 2017]. Se explican los fundamentos de los métodos de aprendizaje automático mediante redes neuronales y más concretamente el desarrollo paso a paso de dos de los tipos más utilizados recientemente, LSTM (Long Short-Term Memory, o redes de gran memoria de corto plazo) y CNN (Convolutional Neural Networks o redes neuronales convolucionales). En el ejemplo planteado en este capítulo, se intentan aplicar ambos métodos para predecir la longitud de una clave de cifrado XOR sobre el texto cifrado, dejando espacio para que el lector cree su propio modelo predictivo para predecir la clave de cifrado.

5.3. Recomendaciones de estudio

Aparte del contenido teórico, en el capítulo correspondiente a este tema se incluye un ejemplo de aplicación desarrollado que es necesario para afianzar los conocimientos.

Referencias

Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. *Mathematics for Machine Learning*. Cambridge University Press, 2020. ISBN 9781108455145. PDF available in `https://mml-book.com`.

S. Dua and X. Du. *Data mining and machine learning in cybersecurity*. CRC press, 2011. ISBN 9781439839423.

David Freeman and Clarence Chio. *Machine Learning and Security: Protecting Systems with Data and Algorithms*. O'Reilly Media, Inc., 2018. ISBN 9781491979891. Available online in Safari-Books.

Jesus Mena. *Machine learning forensics for law enforcement, security, and intelligence*. Auerbach Publications, 2011. ISBN 9781439860694. Available online in Safari-Books.

Nils J. Nilsson. *Inteligencia Artificial: Una nueva síntesis*. McGraw-Hill, 2001. ISBN 978-84-481-2824-1.

Stuart Russell and Peter Norvig. *Inteligencia Artificial: Un Enfoque Moderno*. Pearson, 2 edition, 2004. ISBN 9788420540030.

The Cylance Data Science Team. *Introduction to Artificial Intelligence for Security Professionals*. Cylance Press, 2017. Request free copy in `https://pages.cylance.com/en-us-introduction-to-ai-book.html`.