



Lección 4.

Aplicaciones de detección de anomalías



La detección de anomalía

Aplicaciones

UNIVERSIDAD DE CÓRDOBA

La detección de anomalías se ha aplicado a muchos dominios. Entre algunos de estos campos destacan la detección de fraude, la detección de intrusiones, aplicaciones médicas, análisis en redes, mantenimiento de maquinaria industrial y procesamiento de imágenes. Lo que comparten los diferentes dominios es que normalmente la anomalía en estos dominios suele suponer una situación crítica que debe ser detectada y solucionada lo antes posible.

Aunque los métodos de detección de anomalías se definen para dominios específicos. A veces una metodología de detección de anomalías exitosa en un dominio también puede tener éxito en un área de estudio completamente nueva, y el conocimiento del primero puede permitir avances más rápidos en el segundo. Es por tanto importante tener una comprensión firme de los principios y algoritmos de detección de anomalías y comprender el alcance de su aplicabilidad.

1. Aplicaciones

Aunque se encuentra casi inadvertida, la detección de anomalías tiene una gran cantidad de áreas de aplicación. Prácticamente cualquier tarea que implique una gran recopilación de datos podría aplicar la detección de anomalías. En esta sección discutimos varias aplicaciones reales [1,2]. De cada aplicación se dará, de acuerdo con la información que ya hemos visto, la noción de anomalía, la naturaleza de los datos y los retos asociados con el problema de la detección de anomalías.

1.1 Medicina y salud

En general, en el ámbito de la salud, la detección de anomalías desempeña un papel muy importante. Concretamente, con la detección de anomalías se pueden analizar muchos procesos. Por ejemplo, los modelos pueden detectar sutiles irregularidades en los latidos del corazón de un paciente para clasificar enfermedades, o pueden medir la actividad de las ondas cerebrales para ayudar a los médicos a diagnosticar ciertas enfermedades. Más allá de eso, pueden ayudar a analizar datos de diagnóstico y procesarlos rápidamente para diagnosticar cualquier posible enfermedad. También en estudios epidemiológicos para determinar medicamentos que anteriormente habían tenido éxito, pero que dejan de ser útiles para los pacientes, lo que significa la aparición de una mutación resistente a los medicamentos del patógeno responsable. También pueden obtenerse datos de pacientes individuales, cuya respuesta a un medicamento puede seguir un camino inusual, por ejemplo, inicialmente mejora y luego empeora rápidamente.

Finalmente, puede usarse en imágenes médicas para determinar si la imagen contiene objetos anómalos o no. Por ejemplo, si un modelo solo estuvo expuesto a imágenes de resonancia magnética de huesos normales y se mostró una imagen de un hueso roto, marcaría la nueva imagen como una anomalía. Del mismo modo, la detección de anomalías puede incluso extenderse a detección de tumores, lo que permite que el modelo analice cada imagen en una resonancia magnética de cuerpo completo y buscar la presencia de crecimiento o patrones anormales.

Los datos en este ámbito suelen consistir en registros que pueden tener varios tipos diferentes de características como la edad del paciente, el grupo sanguíneo, el peso. Los datos también pueden tener aspectos temporales y espaciales. Muchas de las aplicaciones están orientadas a detectar registros anómalos (anomalías puntuales). Por lo general, los datos etiquetados pertenecen a pacientes sanos, por lo que la mayoría de las técnicas adoptan un enfoque semisupervisado. También se trabaja con datos de series temporales, como electrocardiogramas y en este caso, se busca la detección colectiva de anomalías.

El aspecto más desafiante del problema de detección de anomalías en este dominio es que las consecuencias de clasificar una anomalía como normal, cuando realmente no lo es, pueden ser catastróficas.

1.2 Industria

La llegada de la Industria 4.0 y el proceso de digitalización está creando un clima muy propicio para este tipo de técnicas, ya que gracias a los sensores instalados en las maquinarias se obtiene una gran cantidad de información en tiempo real que puede ser procesada por técnicas de detección de anomalías.

Una de las áreas más relevantes es la relativo al mantenimiento predictivo. Las maquinarias industriales sufren daños debido al uso continuo y al desgaste normal. Dichos daños deben detectarse a tiempo para evitar que la máquina se pare durante un tiempo debido a un fallo, que supondría grandes pérdidas para la empresa. Los datos en este dominio generalmente se obtienen de sensores. Las técnicas de detección de anomalías se han aplicado ampliamente en este dominio para detectar tales daños y llevar a cabo un mantenimiento predictivo, se predice cuándo va a ocurrir un fallo, para que la máquina se repare justo cuando sea necesario. La detección de daños industriales se puede clasificar en dos dominios, uno que se ocupa de los defectos en componentes mecánicos tales como motores y el otro que se ocupa de los defectos en las estructuras físicas [1].

Algunas de las principales aplicaciones que se pueden encontrar en este entorno son:

- **Detección de fallos en los dispositivos.** Las técnicas de detección de anomalías en este dominio monitorean el desempeño de componentes industriales tales como motores, turbinas, flujo de aceite en tuberías u otros componentes mecánicos y detectan posibles averías que pueden ocurrir debido al uso y desgaste o por otras circunstancias imprevistas. Los datos en este dominio tienen típicamente un aspecto temporal y un análisis de series de tiempo. Las anomalías ocurren principalmente debido a una observación en un contexto específico (anomalías contextuales) o como una secuencia anómala de observaciones (anomalías colectivas). Típicamente, los datos normales (pertenecientes a componentes sin defectos) suelen estar disponibles y, por lo tanto, se trabaja con técnicas semi-supervisadas. Se requiere que las anomalías se detecten en línea ya que se requieren medidas preventivas que eviten paradas en el sistema.
- **Detección de defectos estructurales.** Las técnicas de detección de daños y defectos estructurales detectan anomalías estructurales en estructuras, por ejemplo, grietas en vigas o tensiones en fuselajes. Los datos recopilados en este dominio tienen un aspecto temporal. Las técnicas de detección de anomalías son similares a las técnicas de detección de novedades o detección de puntos de cambio ya que intentan detectar cambios en los datos recopilados de una estructura. Los datos normales, no suelen presentar cambios a lo largo del tiempo y, por lo tanto, los modelos aprendidos suelen ser estáticos. Los datos podrían tener correlaciones espaciales.

1.3 Finanzas

El principal ámbito en esta área está relacionado con la detección de fraude y se refiere a la detección de actividades delictivas que ocurren en organizaciones como bancos, agencias de seguros, empresas telefónicas o mercado de valores. Los usuarios maliciosos pueden ser los clientes reales de la organización o podría estar haciéndose pasar por un cliente (también conocido como robo de identidad).

El fraude se produce cuando estos usuarios consumen los recursos proporcionados por la organización de forma no autorizada. Las organizaciones están interesadas en la inmediata detección de tales fraudes para prevenir pérdidas económicas. El enfoque típico de las técnicas de detección de anomalías en este ámbito es mantener un perfil de usuario para cada cliente y monitorear los perfiles para detectar cualquier desviación de su perfil [2].

Algunas de las principales aplicaciones que se pueden encontrar en este entorno son:

- **Detección de fraude en tarjetas de crédito.** En este dominio, la tecnología de detección de anomalías se aplican técnicas para detectar solicitudes de tarjetas de crédito fraudulentas o el uso de tarjetas de crédito fraudulentas (asociado con robos de tarjetas de crédito). Los datos con los que se trabaja generalmente se componen de registros definidos en varias dimensiones, como la identificación del usuario, la cantidad gastada, el tiempo entre el uso consecutivo de la tarjeta, entre otros. Los fraudes normalmente se reflejan en registros transaccionales (anomalías puntuales) y corresponden a pagos altos, compra de artículos no comprados anteriormente por el usuario o alta tasa de compra. Las empresas de crédito disponen de datos completos y también registros etiquetados. Además, los datos se dividen en distintos perfiles según el usuario. Por lo tanto, las técnicas basadas en el perfilado y la agrupación se utilizan típicamente en este dominio.

El desafío asociado con la detección del uso no autorizado de tarjetas de crédito es que requiere la detección en línea de fraude tan pronto como se realice la transacción fraudulenta. Las técnicas de detección de anomalías se han aplicado de dos maneras diferentes para abordar este problema. El primero se conoce como de propietario en el que cada usuario de tarjeta de crédito se perfila en función del historial de uso de su tarjeta de crédito. Cualquier transacción nueva es comparada con el perfil del usuario y etiquetada como una anomalía si no coincide con el perfil. Este enfoque suele ser costoso, ya que requiere consultar un repositorio de datos cada vez que un usuario realiza una transacción. Otros enfoques, detectan anomalías entre las transacciones que tienen lugar en una ubicación geográfica específica, se centran más en la operación en sí. Tanto las técnicas por usuario como por operación detectan el contexto de las anomalías. En el primer caso el contexto es un usuario, mientras que en el segundo caso el contexto es la ubicación geográfica.

- **Detección de fraude en telefonía.** La detección de fraude en móviles es un problema típico de monitoreo de actividad. La tarea es escanear un gran conjunto de cuentas, examinar el comportamiento de llamada de cada uno y emitir una alarma cuando una cuenta parece haber sido mal utilizado.

La actividad de llamadas se puede representar de varias maneras, pero generalmente se describe con registros de llamadas donde se almacena información del número, de la duración y de la ciudad. Además, las llamadas se pueden agregar por tiempo, por ejemplo, en horas de llamada o días de llamada o usuario o área dependiendo de la granularidad deseado. Las anomalías corresponden a alto volumen de llamadas o llamadas realizadas a improbables destinos.

- **Detección de fraude de seguros.** Las empresas de seguros de accidentes reciben muchos fraudes en los partes que le suponen grandes pérdidas financieras. Los datos disponibles en este dominio son los documentos presentados por los reclamantes donde se utilizan diferentes características tanto categóricas como continuas de estos documentos. Por lo general, los evaluadores de los partes y los investigadores evalúan estas reclamaciones por fraude. Estos casos investigados manualmente se utilizan como instancias etiquetadas mediante técnicas supervisadas y semisupervisadas. La detección de fraudes en las reclamaciones de seguros se suele tratar como un problema genérico de seguimiento de actividades.

1.4 Ciberseguridad

A diario, y cada día más, utilizamos las tecnologías de la información para comunicarnos. Esto nos lleva a poner cada vez más el foco de atención en la seguridad informática, ya que, junto con la cantidad de usuarios, crece también la cantidad de ciberdelincuentes que pretenden obtener beneficio a costa de dichos usuarios. Por ello, es primordial priorizar en la detección ciberataques, para así poder defendernos de estos ciberdelincuentes. Esta detección la podemos enfocar y enmarcar dentro del campo de la detección de anomalías en los sistemas en los que hacemos uso de las tecnologías de la información, que en algunos casos están relacionados con los fraudes que se producen también en las finanzas [3].

En este dominio la detección de intrusos es uno de los retos más tratados en la seguridad de las redes en sistemas informáticos y su objetivo es la identificación de actividad inusual o ataques a la seguridad de redes internas. Con este fin se han desarrollado sistemas de detección de intrusiones que proveen tempranas alertas ante intrusiones que permiten prevenir o minimizar el daño.

Algunas de las principales aplicaciones que se pueden encontrar en este entorno son:

- **La detección de malware.** Un malware es un software malicioso que causa daños a través de virus informáticos, gusanos y troyanos. Si no se cuenta con seguridad en Internet, tus ordenadores, redes, dispositivos móviles y datos pueden verse afectados. En la última década, la detección de malware en redes se ha convertido en una tarea crucial, con el uso incipiente de las redes de computadores los nuevos tipos de malware están a la orden del día. Mediante técnicas de anomalías se realiza un análisis estadístico de la carga útil del tráfico que se ha convertido en una medida esencial para la identificación de nuevos tipos de malware.

El funcionamiento de estos sistemas se basa en la creación de patrones que representen el modo de uso habitual y legítimo de la red y se crea un modelo estadístico del tráfico legítimo. Comparando dicho modelo con el de varios ataques conocidos es capaz de establecerse un conjunto de reglas que permitan detectar las anomalías presentes en el tráfico de la red a proteger. La fase de detección compara el tráfico a analizar con el modelo de tráfico legítimo generado en el entrenamiento, y aplica las reglas generadas a partir de muestras de ataques para determinar si contiene algún tipo de amenaza para el sistema protegido.

- **La detección de ataques de denegación de servicio.** Estos ataques envían varias solicitudes al recurso atacado, con la intención de desbordar la capacidad del sitio de administrar varias solicitudes y haciendo que este deje de funcionar correctamente. Los ataques de denegación de servicio plantean una amenaza en constante crecimiento. Esto es debido principalmente a la tendencia al incremento en su sofisticación, facilidad de implementación, mejora de su capacidad de ofuscación y la existencia de métodos cada vez más eficaces para ocultar la identidad del atacante. En este problema se suele trabajar con los datos de sensores distribuidos con capacidad de adaptación a cambios producidos en el escenario a monitorizar. El tráfico monitorizado es analizado por medio del estudio de la entropía de su distribución. Con este fin se buscan variaciones inesperadas en el volumen de datos transmitidos que descubran comportamientos discordantes. El uso de esta técnica permite detectar cualquier movimiento, incluidos aquellos que no han sido reconocidos con anterioridad.

En este dominio existen muchas otras amenazas donde la detección de anomalías resulta una técnica muy efectiva, ya que todos ellos se basan en encontrar un uso no frecuente en la red, que permita adelantarse a los ataques.

Finalizamos esta sección indicando que se tratan de algunos dominios de aplicación, para tomar un primer contacto de la utilidad de esta técnica en diferentes ámbitos. No obstante, debido a su funcionamiento, les hace poder estar presente en una gran cantidad de aplicaciones y dominios donde resulta crucial detectar cualquier cambio relevante que se produzca con respecto a una situación considerada normal.

Referencias

- [1] Alla, S., & Adari, S. K. (2019). *Beginning anomaly detection using python-based deep learning*. New Jersey: Apress.
- [2] Chandola, V., Banerjee, A., & Kumar, V. (2009). *Anomaly detection: A survey*. ACM computing surveys (CSUR), 41(3), 1-58.
- [3] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman. (2019). *Survey of intrusion detection systems: techniques, datasets and challenges*. Cybersecurity, 2(1), 1-22.