ITUPublications
Recomendaciones

Unión Internacional de Telecomunicaciones
Sector de la Normalización

Recomendación

UIT-T M.3387 (03/2024)

SERIE M: Gestión de las telecomunicaciones, incluida la RGT y el mantenimiento de redes

Red de gestión de las telecomunicaciones

Requisitos de gestión de sistemas de aprendizaje automático federado

RECOMENDACIONES UIT-T DE LA SERIE M

Gestión de las telecomunicaciones, incluida la RGT y el mantenimiento de redes

Introducción y principios generales de mantenimiento y organización del mantenimiento	M.10-M.299
Sistemas internacionales de transmisión	M.300-M.559
Circuitos telefónicos internacionales	M.560-M.759
Sistemas de señalización por canal común	M.760-M.799
Circuitos internacionales utilizados para transmisiones de telegrafía y de telefotografía	M.800-M.899
Enlaces internacionales arrendados en grupo primario y secundario	M.900-M.999
Circuitos internacionales arrendados	M.1000-M.1099
Sistemas y servicios de telecomunicaciones móviles	M.1100-M.1199
Red telefónica pública internacional	M.1200-M.1299
Sistemas internacionales de transmisión de datos	M.1300-M.1399
Designaciones e intercambio de información	M.1400-M.1999
Red de transporte internacional	M.2000-M.2999
Red de gestión de las telecomunicaciones	**M.3000-M.3599**
Redes digitales de servicios integrados	M.3600-M.3999
Sistemas de señalización por canal común	M.4000-M.4999

Para más información, véase la Lista de Recomendaciones del UIT-T.

Recomendación UIT-T M.3387

Requisitos de gestión de sistemas de aprendizaje automático federado

Resumen

La Recomendación UIT-T M.3387 es aplicable al diseño arquitectónico, la investigación y el desarrollo de modelos de aprendizaje automático federado (FMLM). La privacidad de los datos y la seguridad de la información plantean importantes retos a las comunidades de los macrodatos y la inteligencia artificial (IA), ya que estas se ven cada vez más presionadas para cumplir requisitos reglamentarios. Muchas operaciones rutinarias en sistemas y aplicaciones de macrodatos, como la fusión de datos de usuarios procedentes de diversas fuentes para construir un modelo de aprendizaje automático, se consideran ilegales en marcos reglamentarios actuales.

El objetivo del aprendizaje automático federado (FML) es ofrecer una solución viable que permita a las aplicaciones de aprendizaje automático utilizar los datos de forma distribuida. En un marco del FML, los propietarios de los datos no intercambian datos brutos directamente y no permiten que ninguna parte infiera la información privada de otras partes. Para facilitar la construcción y el uso de FMLM y mejorar la calidad del servicio FML, la Recomendación UIT-T M.3387 especifica los requisitos de gestión de los sistemas de aprendizaje automático federados (FMLS), incluida la arquitectura funcional de los FMLS, así como los requisitos del dominio de gestión básica, el dominio de gestión de modelos y el dominio de gestión de datos.

Historia*

Edición	Recomendación	Aprobación	Comisión de Estudio	ID único
1.0	UIT-T M.3387	11-03-2024	2	11.1002/1000/15786

Palabras clave

Servicio de aprendizaje automático federado, sistema de aprendizaje automático federado, requisitos de gestión.

* Para acceder a la Recomendación, introduzca el URL https://handle.itu.int/ en el campo dirección del navegador, seguido por el identificador único de la Recomendación.

PREFACIO

La Unión Internacional de Telecomunicaciones (UIT) es el organismo especializado de las Naciones Unidas en el campo de las telecomunicaciones y de las tecnologías de la información y la comunicación. El Sector de Normalización de las Telecomunicaciones de la UIT (UIT-T) es un órgano permanente de la UIT. Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Asamblea Mundial de Normalización de las Telecomunicaciones (AMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución 1 de la AMNT.

En ciertos sectores de la tecnología de la información que corresponden a la esfera de competencia del UIT-T, se preparan las normas necesarias en colaboración con la ISO y la CEI.

NOTA

En esta Recomendación, la expresión "Administración" se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

La observancia de esta Recomendación es voluntaria. Ahora bien, la Recomendación puede contener ciertas disposiciones obligatorias (para asegurar, por ejemplo, la aplicabilidad o la interoperabilidad), por lo que la observancia se consigue con el cumplimiento exacto y puntual de todas las disposiciones obligatorias. La obligatoriedad de un elemento preceptivo o requisito se expresa mediante las frases "tener que, haber de, hay que + infinitivo" o el verbo principal en tiempo futuro simple de mandato, en modo afirmativo o negativo. El hecho de que se utilice esta formulación no entraña que la observancia se imponga a ninguna de las partes.

PROPIEDAD INTELECTUAL

La UIT señala a la atención la posibilidad de que la utilización o aplicación de la presente Recomendación suponga el empleo de un derecho de propiedad intelectual reivindicado. La UIT no se pronuncia en lo que respecta a la existencia, validez o aplicabilidad de los derechos de propiedad intelectual reclamados, ya sea por los miembros de la UIT o por terceros ajenos al proceso de elaboración de Recomendaciones.

En la fecha de aprobación de la presente Recomendación, la UIT no ha recibido notificación de propiedad intelectual, protegida por patentes/derechos de autor de *software*, que puede ser necesaria para aplicar esta Recomendación. Sin embargo, debe señalarse a los usuarios que puede que esta información no se encuentre totalmente actualizada al respecto, por lo que se les insta encarecidamente a consultar las correspondientes bases de datos del UIT-T disponibles en el sitio web del UIT-T en http://www.itu.int/ITU-T/ipr/.

ÍNDICE

Recomendación UIT-T M.3387

Requisitos de gestión de sistemas de aprendizaje automático federado

1 Alcance

La presente Recomendación especifica los requisitos de gestión de los sistemas de aprendizaje automático federados (FMLS). El ámbito de aplicación de la presente Recomendación incluye los aspectos siguientes:

– Arquitectura funcional general de los FMLS.

– Requisitos del dominio de gestión básica, que supervisa la configuración de las propiedades del sistema, la configuración de permisos de los nodos, la gestión de solicitudes de servicio, etc.

– Requisitos del dominio de gestión del modelo, que supervisa el despliegue del modelo inicial, la gestión de algoritmo de aprendizaje, la gestión de mecanismos de agregación, etc.

– Requisitos del dominio de gestión de datos, que supervisa el almacenamiento seguro, la recuperación y la transmisión de recursos de datos, etc.

– Caso de uso.

2 Referencias

Las siguientes Recomendaciones del UIT-T y otras referencias contienen disposiciones que, mediante su referencia en el presente texto, constituyen disposiciones de la presente Recomendación. A la fecha de esta publicación, las ediciones citadas están en vigor. Todas las Recomendaciones y demás referencias son objeto de revisión, por lo que se alienta a los usuarios de la presente Recomendación a utilizar la edición más reciente de las Recomendaciones y demás referencias que se indican a continuación. Se publica periódicamente una lista de las Recomendaciones UIT-T actualmente vigentes. En esta Recomendación, la referencia a un documento, en tanto que autónomo, no le otorga el rango de una Recomendación.

[IEEE 3652.1] IEEE standard 3652.1-2020, *IEEE Guide for Architectural Framework and Application of Federated Machine Learning*

3 Definiciones

3.1 Términos definidos en otros documentos

En la presente Recomendación se utilizan los siguientes términos definidos en otros documentos:

3.1.1 encriptación [b-UIT-T X.1367]: transformación criptográfica de datos para producir textos cifrados.

3.1.2 mecanismo de incentivos intrínsecos [b-UIT-T Y.4205]: mecanismo que ofrece una recompensa originada internamente como resultado de la contribución a una actividad o la participación en la misma; por ejemplo, experimentar autorrealización, alegría o contribuir a una causa mayor.

3.2 Términos definidos en la presente Recomendación

En esta Recomendación se definen los términos siguientes:

3.2.1 auditor: nodo responsable de supervisar el funcionamiento de un proceso de aprendizaje automático federado para verificar que el proceso cumple los requisitos reglamentarios.

3.2.2 coordinador: nodo responsable de construir modelos de aprendizaje automático federado a partir de diferentes propietarios de datos y proporciona modelos a los usuarios del aprendizaje automático federado.

3.2.3 propietario de los datos: nodo que posee el conjunto de datos utilizado en el aprendizaje automático federado y realiza el entrenamiento del modelo local, garantizando al mismo tiempo la privacidad de los datos.

3.2.4 calidad de los datos: métrica utilizada para evaluar la validez y utilidad de los conjuntos de datos.

3.2.5 conjunto de datos: recopilación de puntos o elementos de datos se utilizan con fines de entrenamiento, pruebas o evaluación. Cada punto de datos de un conjunto representa una muestra, incluidos el identificador de datos, las características de los datos (nombres y valores) o la etiqueta de clase (en el caso del aprendizaje supervisado).

3.2.6 aprendizaje automático federado (FML): marco de aprendizaje automático que facilita la construcción colaborativa de modelos de aprendizaje automático entre múltiples nodos de entrenamiento distribuidos sin revelar los datos privados de los propietarios de los datos.

3.2.7 modelo de aprendizaje automático federado (FMLM): resultado del proceso de entrenamiento de un sistema de aprendizaje automático federado. El modelo entrenado se utiliza para realizar tareas de inferencia sobre nuevos datos.

3.2.8 sistema de gestión del aprendizaje automático federado (FMLMS): sistema de gestión capaz de administrar los recursos de los nodos y los servicios para el entrenamiento de sistemas de aprendizaje automático federados.

3.2.9 servicio de aprendizaje automático federado: servicio de entrenamiento de modelos de inteligencia artificial que utiliza un método de aprendizaje automático federado y cuyo resultado es un modelo entrenado a escala global.

3.2.10 cliente del servicio de aprendizaje automático federado (FMLSC): entidad de aplicación que inicia la solicitud de servicio de aprendizaje automático federado y recibe el modelo de aprendizaje automático federado entrenado.

3.2.11 sistema de aprendizaje automático federado (FMLS): sistema con múltiples nodos de entrenamiento que construyen y utilizan de forma colaborativa modelos de aprendizaje automático sin revelar los datos brutos y privados propiedad de los participantes.

3.2.12 datos brutos: recopilación de conjuntos de datos obtenida, almacenada y mantenida por los propietarios de los datos. Los datos brutos contienen información privada de los usuarios y los propietarios de los datos.

3.2.13 entrenamiento: proceso de aprendizaje automático federado que incluye el entrenamiento local de los datos brutos y la agregación de parámetros actualizados intermedios para optimizar la eficiencia de los modelos de aprendizaje automático federado.

4 Abreviaturas y acrónimos

En la presente Recomendación se utilizan las siguientes abreviaturas y acrónimos:

AUC Superficie bajo la curva (*area under the curve*)

CAP Punto de acceso a la capacidad de computación (*computing access point*)

CNN Red neural convolucional (*convolutional neural network*)

CPU Unidad central de procesamiento (*central processing unit*)

FML Aprendizaje automático federado (*federated machine learning*)

FMLM	Modelo de aprendizaje automático federado (*federated machine learning model*)
FMLMS	Sistema de gestión del aprendizaje automático federado (*federated machine learning management system*)
FMLS	Sistema de aprendizaje automático federado (*federated machine learning system*)
FMLSC	Cliente del servicio de aprendizaje automático federado (*federated machine learning service client*)
GPU	Unidad de procesamiento gráfico (*graphics processing unit*)
IA	Inteligencia artificial
ID	Identificación
IoV	Internet de los vehículos (*Internet of vehicles*)
MEC	Computación periférica móvil (*mobile edge computing*)
MSE	Error cuadrático medio (*mean squared error*)

5 Convenios

En esta Recomendación:

– La expresión "**se requiere**" indica que el requisito es absolutamente obligatorio y debe aplicarse sin excepción si se pretende declarar la conformidad con este documento.

– La expresión "**se recomienda**" indica que se trata de un requisito recomendado pero que no es absolutamente obligatorio. Su cumplimiento no es, por tanto, indispensable para poder declarar la conformidad.

– La expresión "**puede opcionalmente**" indica que un requisito opcional que se permite, sin que ello signifique que se recomienda. No implica que el fabricante deba ofrecer esta opción y que el operador de red/proveedor de servicio tenga la posibilidad de activarla. Significa, más bien, que el fabricante tiene la opción de proporcionar esta función sin que ello afecte a la conformidad con la presente especificación.

6 Visión general

Según se señala en [IEEE 3652.1] "IEEE Guide for Architectural Framework and Application of Federated Machine Learning", la privacidad de los datos y la seguridad de la información plantean importantes retos a las comunidades de los macrodatos y la inteligencia artificial (IA), ya que estas se ven cada vez más presionadas para cumplir los requisitos reglamentarios. Muchas operaciones rutinarias en sistemas y aplicaciones de macrodatos, como la fusión de datos de usuarios procedentes de diversas fuentes para construir un modelo de aprendizaje automático, se consideran ilegales en los marcos reglamentarios actuales. El objetivo del aprendizaje automático federado (FML) es ofrecer una solución viable que permita a las aplicaciones de aprendizaje automático utilizar los datos de forma distribuida. En un marco FML, los propietarios de los datos no intercambian datos brutos directamente y no permiten que ninguna parte infiera la información privada de otras partes.

La norma internacional [IEEE 3652.1] define el marco arquitectónico del FML para promover y facilitar la colaboración entre múltiples partes. Sin embargo, los distintos escenarios empresariales tienen diferentes requisitos para los servicios de entrenamiento de los modelos. Por lo tanto, el sistema federado de aprendizaje automático (FMLS) necesita coordinar diferentes nodos de entrenamiento FML para proporcionar un servicio FML seguro y estable mediante la gestión de las propiedades del sistema, las propiedades del modelo y las propiedades de los datos.

El FMLS es un sistema que proporciona el servicio FML a clientes del servicio de aprendizaje automático federado (FMLSC). Para proporcionar un servicio FML seguro y eficiente, deben definirse la configuración de las propiedades del sistema, la configuración de las propiedades de los nodos de entrenamiento, la gestión de las solicitudes de servicio FML y otras funciones de gestión. Estas funciones de gestión se alojan en el sistema de gestión del aprendizaje automático federado (FMLMS).

Basándose en el principio de funcionamiento de los FMLS descrito en [IEEE 3652.1], esta Recomendación especifica los requisitos de gestión de los FMLS, incluidos el dominio de gestión básica, el dominio de gestión de modelos y el dominio de gestión de datos.

7 Escenario de gestión del sistema de aprendizaje automático federado

En la Figura 1 se muestra el escenario de la gestión del FMLS. En un FMLS, los nodos de entrenamiento del FML representan distintos papeles con arreglo a su función en la tarea en curso del FML, incluidos el de coordinador, auditor y propietario de los datos.

El coordinador se encarga de coordinar la tarea FML dentro del FMLS y aplicar el modelo de aprendizaje automático federado (FMLM) aprendido. El auditor es responsable de supervisar todo el proceso FML para garantizar que los datos son fiables y seguros. El propietario de los datos es responsable de entrenar y actualizar los modelos localmente. En [IEEE 3652.1] figuran funciones más específicas de estos papeles.

Las interfaces relacionadas con la gestión del FMLS se muestran en la Figura 1. Se trata principalmente de dos interfaces:

– La **interfaz I1** es la interfaz situada entre el FMLMS y el FMLSC, utilizado para remitir los requisitos del servicio de aprendizaje automático federado desde el FMLSC y devolver el FMLM entrenado al FMLSC.

– La **interfaz I2** es la interfaz situada entre el FMLMS y el FMLS, utilizado para gestionar los recursos de los nodos y las tareas de entrenamiento.

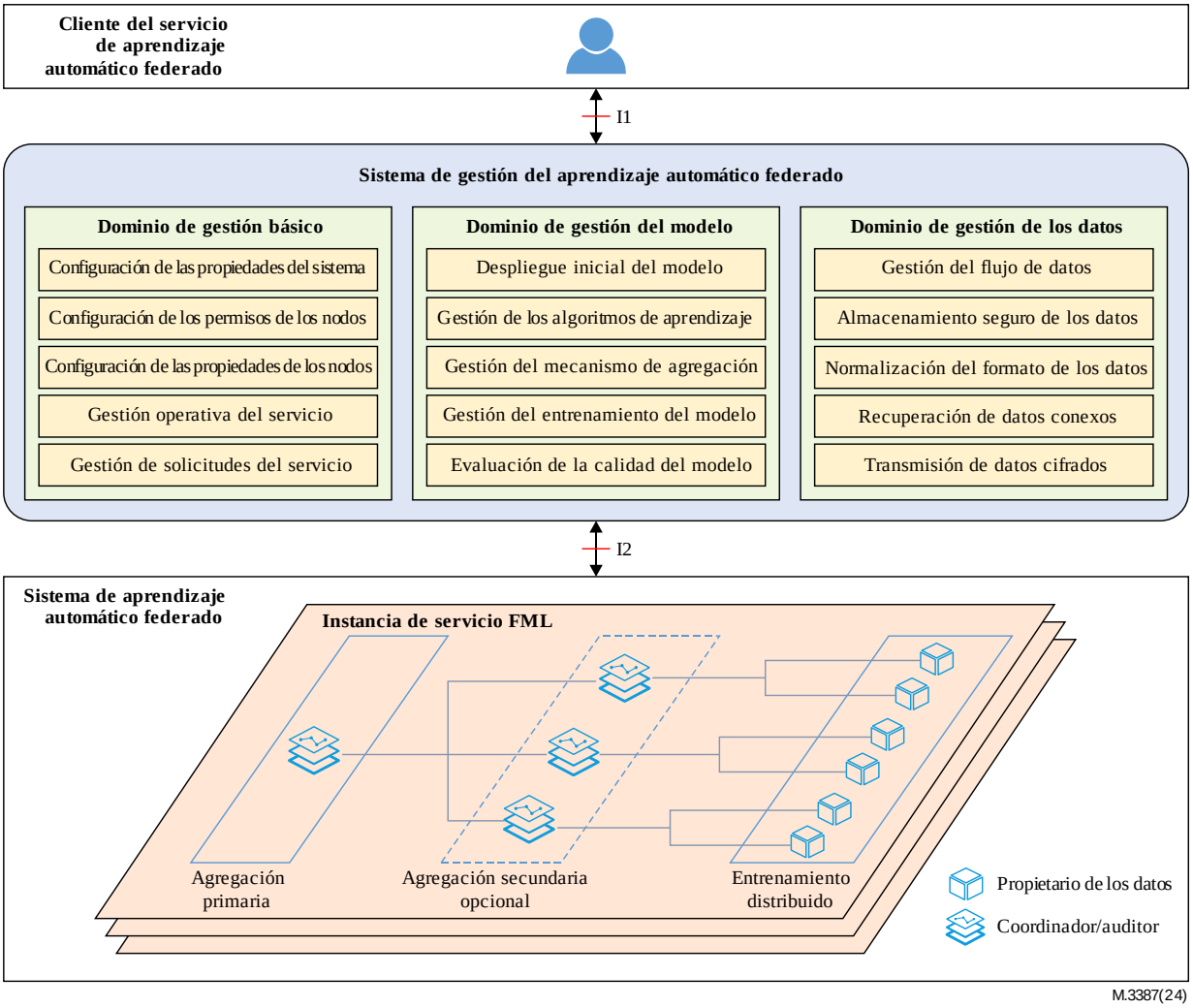

Figura 1 – Escenario de gestión de un sistema federado de aprendizaje automático

El dominio de gestión básico incluye la configuración de las propiedades del sistema, la configuración de los permisos de los nodos, la configuración de las propiedades de los nodos, la gestión del funcionamiento del servicio y la gestión de las solicitudes del servicio.

– **Configuración de las propiedades del sistema**: inicialización y modificación de las propiedades de los FMLS.

– **Configuración de los permisos de los nodos**: gestión los permisos de los nodos de entrenamiento FML según las reglas de seguridad establecidas por el FMLMS.

– **Configuración de las propiedades de los nodos**: inicialización y modificación de las propiedades de los nodos de entrenamiento del FML.

– **Gestión del funcionamiento del servicio**: gestión de la topología del servicio FML y evaluar la calidad del servicio FM.

– **Gestión de las solicitudes de servicio**: clasificación, procesamiento y respuesta a las solicitudes de servicio FML.

El dominio de gestión del modelo, que incluye el despliegue inicial del modelo, la gestión del algoritmo de aprendizaje, la gestión de mecanismos de agregación, la gestión del entrenamiento de modelos y la evaluación de la calidad del modelo.

– **Despliegue inicial del modelo**: despliegue del FMLM inicial en el coordinador conforme a los requisitos del servicio.

– **Gestión de los algoritmos de aprendizaje**: selección los algoritmos de aprendizaje automático adecuados en función de los requisitos del servicio FM.
– **Gestión de los mecanismos de agregación**: selección o diseño de estrategias de agregación adecuadas en función de los requisitos del servicio FML, la capacidad de los recursos y las características de los datos.
– **Gestión del entrenamiento del modelo**: control y supervisión del proceso de entrenamiento del modelo, incluyendo la transmisión y actualización del FMLM.
– **Evaluación de la calidad del modelo**: evaluación de la calidad del FMLM basándose en las métricas de evaluación.

El dominio de gestión de los datos, incluida la gestión del flujo de datos, el almacenamiento seguro de datos, la normalización del formato de los datos, la recuperación de datos conexos y la transmisión cifrada de datos.

– **Gestión del flujo de datos**: control del flujo de metadatos de los datos brutos y del flujo de datos de FMLM.
– **Almacenamiento seguro de datos**: almacenamiento de los metadatos de los datos brutos utilizando diversos métodos de cifrado.
– **Normalización del formato de los datos**: estandarización del formato de los metadatos de los datos brutos, por ejemplo, la estructura de la tabla.
– **Recuperación de datos relacionados**: recuperación de los datos relacionados con la tarea FML como los conjuntos de datos para llevar a cabo el entrenamiento del FMLM.
– **Transmisión cifrada de datos**: seleccionar los algoritmos de cifrado para cifrar los datos transmitidos y los canales de comunicación.

8 Requisitos del dominio de gestión básico

8.1 Requisitos de la configuración de las propiedades del sistema

Se requiere que el sistema de gestión de aprendizaje automático federado (FMLMS) configure las propiedades del FMLS para soportar la implementación funcional de los servicios FML, incluidas las propiedades de las tareas y las propiedades de los recursos.

8.1.1 Propiedades de las tareas

Se requiere que el FMLMS configure las propiedades de las tareas en función de las solicitudes de servicio FML de los FMLSC.

– Tipo de tarea: la categoría de la tarea FML, por ejemplo, tarea de clasificación de imágenes y tarea de generación de texto.
– Prioridad de la tarea: la importancia de la tarea FML, por ejemplo, prioridad alta, prioridad media y prioridad baja.

8.1.2 Propiedades de los recursos

Se requiere que el FMLMS configure las capacidades de recursos de una tarea FML, incluyendo la capacidad de cálculo, la capacidad de comunicación y la capacidad de almacenamiento.

– Capacidad de cálculo: recursos totales de cálculo necesarios para las tareas FML, por ejemplo, el número de unidades centrales de procesamiento (CPU) y unidades de procesamiento gráfico (GPU).
– Capacidad de comunicación: recursos de comunicación totales necesarios para las tareas del FML, por ejemplo, potencia de transmisión y ancho de banda.

– Capacidad de almacenamiento: recursos totales de almacenamiento necesarios para las tareas de FML, por ejemplo, capacidad de disco libre y espacio de almacenamiento preasignado.

8.2 Requisitos de la configuración de los permisos de los nodos

Se requiere que el FMLMS seleccione nodos de entrenamiento FML adecuados con recursos suficientes de acuerdo con la actual tarea de FML, y posteriormente autorice y controle el permiso de acceso basado en la seguridad y fiabilidad de los nodos.

8.3 Requisitos de la configuración de las propiedades de los nodos

Se requiere que el FMLMS soporte la configuración de las propiedades de los nodos de entrenamiento del FML, incluyendo las propiedades de las funciones, propiedades de computación, propiedades de comunicación y propiedades de almacenamiento.

8.3.1 Propiedades de los papeles desempeñados

Se requiere que el FMLMS configure el papel que desempeñan los nodos de entrenamiento del FML Los papeles de los nodos de entrenamiento del FML incluyen el de coordinador, auditor y propietario de los datos. Las funciones específicas de estos tres papeles están referidas en [IEEE 3652.1].

8.3.2 Propiedades de computación

Se requiere que el FMLMS soporte los nodos de entrenamiento del FML en la configuración de sus atributos computacionales, por ejemplo, el número de CPU y GPU.

8.3.3 Propiedades de comunicación

Se requiere que el FMLMS soporte nodos de entrenamiento de la FML en la configuración de sus atributos de comunicación, por ejemplo, potencia de transmisión y ancho de banda.

8.3.4 Propiedades de almacenamiento

Se requiere que el FMLMS soporte los nodos de entrenamiento FML en la configuración de sus atributos de almacenamiento, por ejemplo, capacidad de disco libre y espacio de almacenamiento preasignado.

8.4 Requisitos de la gestión operativa del servicio

Se requiere que el FMLMS gestione la calidad de funcionamiento del servicio FML, incluida la gestión de la topología del servicio y la evaluación de la calidad del servicio.

8.4.1 Gestión de la topología del servicio

Se requiere que el FMLMS gestione la topología del servicio FML, incluyendo principalmente la generación y la reconstrucción de la topología de servicio.

– Generación de la topología del servicio: generación de la topología del servicio FML para la actual tarea de entrenamiento FML, incluidas las funciones y las relaciones de conexión entre los nodos de entrenamiento FML, y envía la topología a todos los nodos de entrenamiento del FML.
– Reconstrucción de la topología del servicio: reconstruye la topología del servicio FML en caso de fallo de un nodo de servicio, agotamiento de recursos, etc.

8.4.2 Evaluación de la calidad del servicio

Se recomienda que el FMLMS evalúe la calidad del servicio del FML, incluida la evaluación de la calidad de funcionamiento de la red y la evaluación de los incentivos del servicio.

– Evaluación de la calidad de funcionamiento de la red: evaluación de la eficiencia de los sistemas de aprendizaje automático federado (FMLS), como el consumo de recursos y el retardo.
– Evaluación de los incentivos del servicio: evaluación de la contribución global de los nodos de entrenamiento del FML al servicio FML, y establecer un mecanismo de incentivos intrínsecos en función de la contribución a fin de motivar la participación de cada nodo de entrenamiento FML en el servicio FML. La contribución global incluye el consumo de recursos, la contribución a la mejora de la calidad del modelo, etc.

8.5 Requisitos de la gestión de solicitudes de servicios

Se requiere que el FMLMS clasifique y asigne cada solicitud de servicio cuando coexistan varias solicitudes de servicio.

– Clasificación del servicio: clasificación de las solicitudes de servicio en función de su importancia, prioridad, plazos, etc.
– Asignación del servicio: respuesta a las solicitudes de servicios con métodos de programación, como la programación por prioridades.

9 Requisitos del dominio de gestión del modelo

9.1 Requisitos del despliegue inicial del modelo

Se requiere que el FMLMS soporte la distribución del FMLM y el inicio del entrenamiento de acuerdo con las solicitudes de servicio del FML de los FMLSC.

9.2 Requisitos de la gestión del algoritmo de aprendizaje

Se requiere que el FMLMS seleccione algoritmos de aprendizaje automáticos y configure los parámetros pertinentes en función de los requisitos del servicio del FML.

– Selección de algoritmos: selección de algoritmos de aprendizaje automático adecuados, como redes neuronales y árboles de decisión.
– Configuración de parámetros: configuración de los parámetros e hiperparámetros del algoritmo de aprendizaje automático, como la tasa de aprendizaje, el tamaño del lote y el coeficiente de regularización.

9.3 Requisitos de la gestión del mecanismo de agregación

Se requiere que el FMLMS seleccione un mecanismo de agregación y configure los parámetros pertinentes de conformidad con los requisitos del servicio del FML, las capacidades de los recursos y las características de los datos.

– Selección del mecanismo: selección o diseño de un mecanismo de agregación adecuado (por ejemplo, síncrono, asíncrono, semisíncrono).
– Configuración de parámetros: configuración de los parámetros del mecanismo de agregación (por ejemplo, número de rondas de agregación, número de agrupaciones, pesos aplicados en la agregación de modelos).

9.4 Requisitos de la gestión del entrenamiento del modelo

Se requiere que el FMLS preste su apoyo a los propietarios de los datos en la realización del entrenamiento local del FMLM y en la carga del modelo aprendido al coordinador.

Se requiere que el FMLS preste su apoyo al coordinador en la gestión del proceso de entrenamiento del FMLM, incluyendo la difusión, agregación y actualización.

Se requiere que el FMLS preste su apoyo al auditor en la supervisión del proceso de entrenamiento del FMLM sobre la base de las disposiciones reglamentarias, por ejemplo, detectando si el nodo de entrenamiento del FML es fiable, y midiendo la contribución de los nodos de entrenamiento del FML.

9.5 Requisitos de la evaluación de la calidad del modelo

Se requiere que el FMLMS evalúe la calidad del FMLM según las métricas de calidad de funcionamiento del modelo.

NOTA – Las métricas de calidad de funcionamiento del modelo incluyen precisión, recuperación, superficie bajo la curva (AUC) para modelos de clasificación, y error cuadrático medio (MSE) para modelos de regresión [b-UIT-T Y.3179].

Se recomienda que el FMLMS ajuste el proceso de entrenamiento mediante la evaluación del modelo para mejorar su eficiencia.

10 Requisitos del dominio de gestión de datos

10.1 Requisitos de la gestión del flujo de datos

Se requiere que el FMLS soporte la generación, almacenamiento, transmisión y actualización de los metadatos de datos brutos.

NOTA – Los metadatos de los datos brutos (por ejemplo, identificación (ID) de datos, característica de los datos) pueden intercambiarse opcionalmente entre los propietarios de los datos durante el proceso de FML.

Se requiere que el FMLS soporte la transmisión, agregación y actualización de datos de FMLM.

Se requiere que el FMLS soporte la gestión de la seguridad de los datos del modelo para evitar que la privacidad personal quede expuesta a nodos externos y maliciosos durante el proceso FML.

10.2 Requisitos del almacenamiento seguro de datos

Se requiere que el FMLMS soporte el almacenamiento seguro de los metadatos de los datos brutos y los datos del modelo.

10.3 Requisitos de la normalización del formato de los datos

Se requiere que el FMLMS recopile las características de los metadatos de los datos brutos y proporcione un formato estándar y unificado de base de datos.

NOTA – Según la norma [IEEE 3562.1], los datos brutos para FML se almacenan normalmente en un formato de base de datos estándar, en el que cada fila representa una muestra de datos y cada columna representa una característica o etiqueta de esa muestra. Habitualmente, un conjunto de atributos de datos se representa como vectores propios (X_1, X_2, …, X_n). En el aprendizaje supervisado, el conjunto completo de datos de entrenamiento está formado por características representadas por X y etiquetas representadas por Y.

10.4 Requisitos de la recuperación de datos conexos

Se requiere que el FMLS preste apoyo a los propietarios de los datos en la recuperación de los datos brutos relacionados con el entrenamiento del modelo como conjuntos de datos. En un FMLS, pueden solaparse los ID de las muestras y las características de los atributos. Se distinguen tres casos en función del grado de solapamiento de los ID de muestras o de las características:

– FML horizontal: construcción de un modelo en el que los conjuntos de datos presentan solapamientos significativos en el espacio de características pero no en el espacio de los ID. El coordinador es responsable de alinear las características de los distintos propietarios de datos.

– FML vertical: construcción de un modelo en el que los conjuntos de datos presentan solapamientos significativos en el espacio muestral pero no en el espacio de características. El coordinador se encarga de alinear las muestras entre los propietarios de los datos.

– Aprendizaje por transferencia federada: construcción de un modelo en el que los conjuntos de datos no presentan solapamientos significativos ni en el espacio muestral ni en el espacio de características. El coordinador se encarga de explotar los conocimientos reutilizables en los distintos dominios de características.

10.5 Requisitos de la transmisión cifrada de datos

Se requiere que el FMLS soporte tecnologías de protección de la privacidad de los datos, como la computación multipartita segura, el cifrado homomórfico y la privacidad diferencial, garantizando que otros nodos de entrenamiento del FML no puedan inferir la información de los datos en bruto a partir de los datos del modelo.

Se recomienda que el FMLS soporte tecnologías de cifrado de canales para mantener un entorno seguro durante los procesos de transmisión de datos.

Apéndice I

Ejemplo de caso de uso del FMLMS para gestionar el entrenamiento del FMLM para servicios de detección de anomalías viales en la Internet de los vehículos

(Este apéndice no forma parte integrante de la presente Recomendación.)

Este apéndice presenta un ejemplo típico de aplicación y servicio para el uso del sistema de gestión de aprendizaje automático federado (FMLMS) a fin de gestionar el sistema de aprendizaje automático federado (FMLS) en la Internet de los vehículos (IoV). También se describen las funciones del FMLMS para el servicio FML de la presente Recomendación.

I.1 Introducción

Las aplicaciones emergentes basadas en dispositivos inteligentes, como los vehículos inteligentes, tienen requisitos estrictos de latencia y privacidad. Esto hace que la computación en la nube no sea adecuada para estos escenarios por lo que se plantea el uso del FML basado en la computación periférica móvil (MEC). El FML basado en MEC entrena los modelos de aprendizaje automático de forma distribuida en dispositivos móviles con recursos limitados de computación, almacenamiento, energía y ancho de banda, conservando los datos brutos a nivel local. Los parámetros del modelo se agregan en los puntos de acceso de computación (CAP) más cercanos. En la Internet de los vehículos (IoV), la compartición de datos entre vehículos para un análisis colaborativo mejora la experiencia de conducción y la calidad del servicio. Por lo tanto, el diseño de una arquitectura de computación colaborativa de FML asistida por MEC facilita la detección de anomalías viales al tiempo que se salvaguarda la privacidad de los datos.

I.2 Arquitectura FML colaborativa nube-periferia-terminal

En la IoV, un FMLS utiliza FML para aprender un FMLM global que se aplica a servicios de detección de anomalías viales. Los componentes del FMLS incluyen a los propietarios de los datos desplegados en vehículos inteligentes, los coordinadores desplegados en puntos de acceso de computación, y un servidor de agentes FML (representado como AS en la Figura I-1) desplegado en el servidor en la nube (representado como CS en la Figura I.1). El FMLMS está desplegado en la nube y gestiona el proceso de entrenamiento del FML. Las tecnologías MEC se aplican para garantizar la calidad del servicio FML mediante la programación de las tareas y la agregación regional del modelo. En la Figura I.1 se representa el escenario de entrenamiento del FML en una arquitectura nube-periferia-terminal.

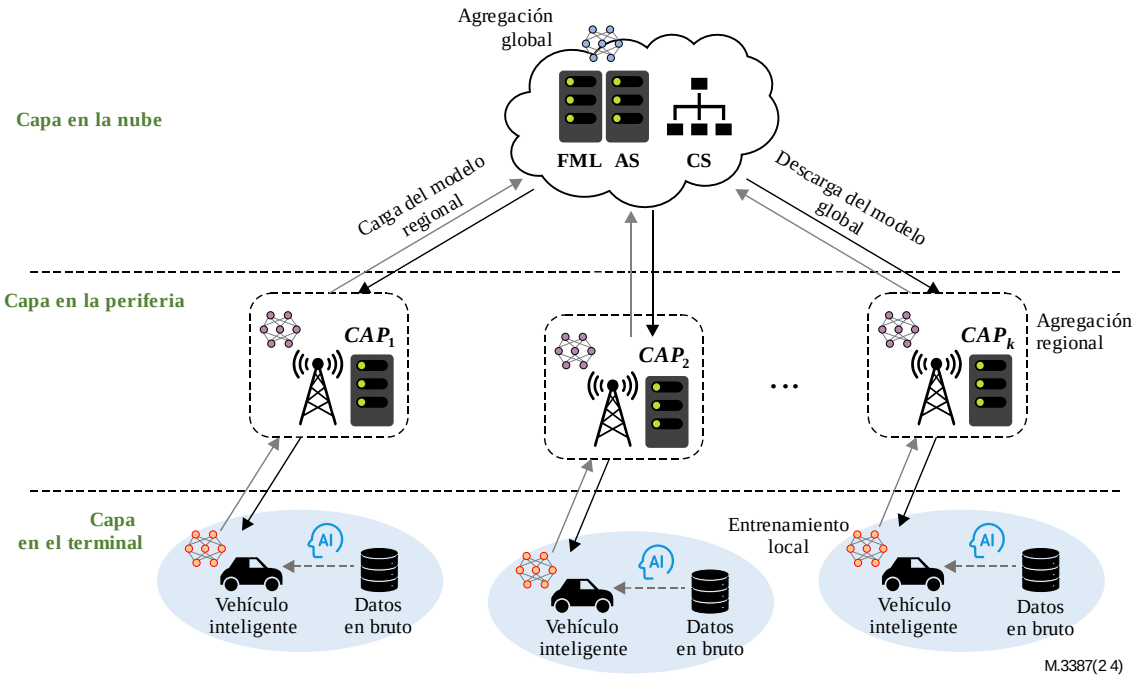

Figura I.1 – Escenario de FML colaborativo nube-periferia-terminal en la IoV

Como se ilustra en la Figura I.1, las responsabilidades de cada una de las capas de los nodos de entrenamiento del FML son las siguientes:

– Capa terminal: los propietarios de los datos, que son vehículos inteligentes, utilizan los datos generados localmente para entrenar un FMLM local y transmiten los parámetros del modelo a los CAP para su agregación regional (agregación secundaria).

– Capa periférica: los CAP, normalmente unidades ubicadas junto a los viales en escenarios vehiculares, son responsables individualmente de la recopilación de parámetros del modelo local de los propietarios de datos en una región específica. A continuación, actualizan el FMLM regional mediante agregación regional (agregación secundaria) y envían los parámetros del modelo actualizados al servidor de agentes. Además, los CAP recopilan metadatos de datos brutos e información sobre el estado de los dispositivos, que luego envían al sistema de gestión de aprendizaje automático federado (FMLMS).

– Capa en la nube: en la nube se despliega un FMLMS para gestionar el funcionamiento del FML. Simultáneamente, el servidor de agentes agrega todos los parámetros del modelo regional para aprender el FMLM global a través de la agregación global (agregación primaria). Además, el FMLMS evalúa la calidad de funcionamiento del FMLS.

I.3 Proceso de gestión del FMLS en el entrenamiento de modelos de detección de anomalías viales

En los escenarios de FML colaborativos nube-periferia-terminal, se despliega un FMLMS en la nube para gestionar los recursos de red y la calidad del servicio del FML, garantizando la eficiencia, sostenibilidad y seguridad del servicio de FML en la IoV. El ejemplo siguiente ilustra el entrenamiento de un modelo de detección de anomalías viales para vehículos inteligentes.

Paso 1: todos los propietarios potenciales de datos (es decir, los vehículos inteligentes) acceden a la red de FML. El cliente del servicio de aprendizaje automático federado (FMLSC) solicita al FMLMS, a través de la interfaz 1, un servicio de entrenamiento del modelo de detección de anomalías viales, es decir, el entrenamiento de un modelo de detección de anomalías viales.

Paso 2: el FMLMS configura los atributos de la tarea de acuerdo con la solicitud de servicio de FML, identificándola como una tarea de reconocimiento de imágenes y asignándole prioridad. Sobre la base de la situación de los recursos de la red, el FMLMS selecciona una estrategia de incentivos para alentar que propietarios de datos adicionales se unan a la tarea de entrenamiento del modelo de detección de anomalías viales.

Utilizando la estrategia de incentivos generada, el FMLMS determina los nodos de entrenamiento del FML para la actual tarea de FML. A continuación, el FMLMS asigna los cometidos de todos los nodos de entrenamiento del FML. Después, el FMLMS determina las correspondientes propiedades de los recursos para todos los nodos de entrenamiento del FML, incluyendo el papel y las propiedades de computación, comunicación y almacenamiento.

Sobre la base del papel de los nodos de entrenamiento del FML, el FMLMS genera una topología de servicio del FML, incluidas las relaciones de vinculación entre los nodos de entrenamiento y las configuraciones de atributos pertinentes.

A continuación, el FMLMS determina el algoritmo de aprendizaje y el mecanismo de agregación para la tarea de detección de anomalías viales, por ejemplo, una red neuronal convolucional (CNN) y un algoritmo de agregación asíncrono.

Paso 3: el FMLMS envía al FMLS, a través de la interfaz 2, el modelo CNN inicial, las propiedades del sistema, los permisos del nodo, las propiedades del nodo, la topología del servicio, el algoritmo de aprendizaje y el mecanismo de agregación. Además, el FMLMS envía al FMLS a través de la interfaz 2, parámetros relacionados con los algoritmos de privacidad y los algoritmos de cifrado de canales, logrando el cifrado de los datos del modelo y el cifrado del canal para garantizar la privacidad y la seguridad del FMLM.

Paso 4: el FMLS despliega el modelo CNN inicial en todos los nodos de entrenamiento del FML y exige a los propietarios de los datos que preprocesen los datos brutos de su propiedad en un formato estandarizado. A continuación, todos los propietarios de datos recuperan los datos de imágenes pertinentes de acuerdo con la solicitud de servicio para entrenar el FMLM local.

Paso 5: los propietarios de los datos cargan el modelo local entrenado en un CAP cercano. A continuación, los CAP generan modelos regionales basados en los datos de los modelos locales recopilados y envían los modelos regionales al servidor de agentes para su agregación global.

Paso 6: el FMLMS supervisa la eficiencia operativa del FMLS y la calidad de los FMLM. El FMLMS puede ajustar la topología del servicio en función de dicha eficiencia operativa y determinar si debe finalizar la tarea de entrenamiento basándose en la calidad del FMLM global. Cuando la precisión del modelo cumple los requisitos del servicio, el FMLMS envía el FMLM entrenado globalmente al FMLSC a través de la interfaz 1. Si la precisión del modelo no cumple los requisitos del servicio, la tarea de entrenamiento continúa.

Bibliografía

[b-UIT-T X.1367] Recomendación UIT-T X.1367 (2020), *Formato normalizado para los registros de errores de Internet de las cosas (IoT) utilizado en las operaciones de incidentes de seguridad*.

[b-UIT-T Y.3179] Recomendación UIT-T Y.3179 (2021), *Marco arquitectónico para el aprendizaje automático en redes futuras, incluidas las IMT-2020*.

[b-UIT-T Y.4205] Recomendación UIT-T Y.4205 (2019), *Requisitos y modelo de referencia de los sistemas de externalización masiva relativos a la IoT*.

SERIES DE RECOMENDACIONES DEL UIT-T

Serie A	Organización del trabajo del UIT-T
Serie D	Principios de tarificación y contabilidad y cuestiones económicas y políticas de las telecomunicaciones/TIC internacionales
Serie E	Explotación general de la red, servicio telefónico, explotación del servicio y factores humanos
Serie F	Servicios de telecomunicación no telefónicos
Serie G	Sistemas y medios de transmisión, sistemas y redes digitales
Serie H	Sistemas audiovisuales y multimedia
Serie I	Red digital de servicios integrados
Serie J	Redes de cable y transmisión de programas radiofónicos y televisivos, y de otras señales multimedia
Serie K	Protección contra las interferencias
Serie L	Medio ambiente y TIC, cambio climático, ciberdesechos, eficiencia energética, construcción, instalación y protección de los cables y demás elementos de planta exterior
Serie M	**Gestión de las telecomunicaciones, incluida la RGT y el mantenimiento de redes**
Serie N	Mantenimiento: circuitos internacionales para transmisiones radiofónicas y de televisión
Serie O	Especificaciones de los aparatos de medida
Serie P	Calidad de la transmisión telefónica, instalaciones telefónicas y redes de líneas locales
Serie Q	Conmutación y señalización, y mediciones y pruebas asociadas
Serie R	Transmisión telegráfica
Serie S	Equipos terminales para servicios de telegrafía
Serie T	Terminales para servicios de telemática
Serie U	Conmutación telegráfica
Serie V	Comunicación de datos por la red telefónica
Serie X	Redes de datos, comunicaciones de sistemas abiertos y seguridad
Serie Y	Infraestructura mundial de la información, aspectos del protocolo Internet, redes de próxima generación, Internet de las cosas y ciudades inteligentes
Serie Z	Lenguajes y aspectos generales de soporte lógico para sistemas de telecomunicación

Publicado en Suiza
Ginebra, 2024