

**GENERACIÓN AUTOMÁTICA DE RESÚMENES DE RELATOS
DELICTUALES CON MODELOS TRANSFORMERS**

POR: OSCAR CABRERA VALENZUELA

Proyecto de grado presentado a la Facultad de Ingeniería de la Universidad del
Desarrollo para optar al grado académico de Magíster en *Data Science*

PROFESOR GUÍA:

Dra. MARÍA PAZ RAVEAU MORALES

DICIEMBRE DE 2025
SANTIAGO DE CHILE

Tabla de contenido

1.	RESUMEN	1
2.	INTRODUCCIÓN	2
2.1.	CONTEXTO INSTITUCIONAL	2
2.2.	PROBLEMA	3
2.3.	MOTIVACIÓN	3
2.4.	SOLUCIÓN PROPUESTA.....	4
2.5.	CONTRIBUCIONES	5
2.6.	ESTRUCTURA DEL INFORME.....	5
3.	TRABAJO RELACIONADO	6
3.1.	MODELOS DE LENGUAJE Y TAREAS DE RESUMEN AUTOMÁTICO	6
3.2.	EL MODELO T5 (<i>TEXT-TO-TEXT TRANSFER TRANSFORMER</i>)	7
3.3.	LoRA (<i>LOW-RANK ADAPTATION</i>) Y <i>FINE-TUNING</i> LIGERO	8
3.4.	EVALUACIÓN DE RESÚMENES EN PLN Y EN EL DOMINIO LEGAL	8
4.	HIPÓTESIS Y OBJETIVOS	10
4.1.	HIPÓTESIS GENERAL	10
4.2.	HIPÓTESIS ESPECÍFICAS	11
4.3.	OBJETIVO GENERAL.....	11
4.4.	OBJETIVOS ESPECÍFICOS	11
5.	DATOS Y METODOLOGÍA	13
5.1.	DATOS	13
5.1.1.	<i>Origen de los datos</i>	13
5.1.2.	<i>Estructura del dataset</i>	14
5.1.3.	<i>Proceso de limpieza</i>	14
5.1.4.	<i>División en conjuntos de entrenamiento y validación</i>	15
5.1.5.	<i>Estadísticas descriptivas</i>	15
5.1.6.	<i>Relación con el límite de tokens del modelo T5:</i>	18
5.1.7.	<i>Implicancias para el tipo de modelo</i>	18
5.1.8.	<i>Selección de sub-datasets por código de delito</i>	19
5.2.	METODOLOGÍA	22
5.2.1.	<i>Enfoque general y diseño experimental</i>	22
5.2.2.	<i>Modelo base: T5 (Text-to-Text Transfer Transformer)</i>	24
5.2.3.	<i>Tokenización y truncamiento</i>	25
5.2.4.	<i>Hiperparámetros de entrenamiento</i>	27
5.2.5.	<i>Entrenamiento de modelos LoRA</i>	29
5.2.6.	<i>Evaluación</i>	32
6.	RESULTADOS	36
6.1.	CONFIGURACIÓN DE LOS EXPERIMENTOS DE EVALUACIÓN	37
6.2.	RESULTADOS CUANTITATIVOS DEL T5 GENERAL	39
6.3.	RESULTADOS CUANTITATIVOS DE MODELOS LoRA ESPECIALIZADOS POR DELITO	40
6.4.	ANÁLISIS CUALITATIVO DE RESÚMENES	42
7.	CONCLUSIONES	46
7.1.	SÍNTESIS DE RESULTADOS.....	46
7.2.	EVALUACIÓN DE LAS HIPÓTESIS	48
7.3.	APORTE PRÁCTICO PARA LA FISCALÍA.....	50
7.4.	LIMITACIONES DEL ESTUDIO.....	51
7.5.	LÍNEAS DE TRABAJO FUTURO.....	52
8.	BIBLIOGRAFÍA	55
9.	ANEXOS	57
9.1.	ANEXO 1A - T5 GENERAL.....	57
9.2.	ANEXO 1B - LoRA 802 (ROBO CON INTIMIDACIÓN)	60
9.3.	ANEXO 1C - LoRA 803 (ROBO CON VIOLENCIA)	63
9.4.	ANEXO 1D - LoRA 804 (ROBO POR SORPRESA)	66
9.5.	ANEXO 1E - LoRA 867 (ROBO DE VEHÍCULO MOTORIZADO).....	68

1. Resumen

El presente trabajo explora la viabilidad de generar resúmenes automáticos de relatos delictuales a partir de los registros del Ministerio Público de Chile, con el fin de apoyar la revisión rápida de causas y el análisis criminal. Para ello se construyó un *dataset* a partir de RUC con relato del hecho y su correspondiente resumen operativo (“detalle”), aplicando un proceso de limpieza que elimina glosas sin contenido real, normaliza textos y exige longitudes mínimas. Sobre este corpus, del orden de decenas de miles de registros, se definió un esquema reproducible de partición en conjuntos de entrenamiento y validación y se generaron, además, *sub-datasets* para los cuatro códigos de delito más frecuentes.

A nivel de modelado, se entrenó un modelo T5-base general para transformar el relato completo en su resumen operativo, y luego se especializó este modelo mediante adaptadores LoRA por tipo de delito. La evaluación se realizó con métricas ROUGE (ROUGE-1, ROUGE-2 y ROUGE-L) sobre los conjuntos de validación, complementada con un análisis cualitativo de ejemplos de resúmenes generados.

El modelo T5 general alcanza aproximadamente $ROUGE-1 \approx 0,30$, $ROUGE-2 \approx 0,13$ y $ROUGE-L \approx 0,25$, lo que indica una capacidad razonable para recuperar el contenido esencial de los resúmenes humanos. Los modelos LoRA igualan o mejoran sistemáticamente estos resultados.

Se concluye que es técnicamente factible entrenar modelos de resumen basados en Transformers sobre los datos actualmente disponibles en la Fiscalía y que la especialización por tipo de delito mediante LoRA aporta mejoras medibles y observables en algunos casos. El trabajo deja instalado un *pipeline* de datos y modelado reproducible, identifica limitaciones y propone líneas de trabajo futuro.

2. Introducción

2.1. Contexto institucional

La labor de la persecución penal en Chile requiere que los fiscales y analistas criminales trabajen diariamente con grandes volúmenes de información proveniente de denuncias, declaraciones e informes policiales. En particular, los relatos descriptivos de hechos delictuales ingresados a través del Sistema de Apoyo a Fiscales (SAF) constituyen una fuente esencial para la caracterización inicial de los casos. Sin embargo, estos textos suelen ser extensos, heterogéneos en calidad, escritos bajo presión operativa y con un nivel significativo de ruido lingüístico (errores tipográficos, duplicaciones, mayúsculas sostenidas, variabilidad narrativa, etc.).

En la Fiscalía Regional Metropolitana Oriente, la integración entre SAF y el Sistema de Información y Registro de Análisis (SIRA) ha permitido centralizar y depurar información crítica para la toma de decisiones analíticas. Aun así, la revisión manual de los relatos sigue siendo un proceso intensivo en tiempo y altamente demandante, especialmente en unidades que deben procesar cientos de registros diarios. La posibilidad de automatizar la

generación de resúmenes precisos y consistentes representa un avance significativo tanto en eficiencia operativa como en estandarización del análisis inicial.

2.2. Problema

Actualmente, los relatos de hechos provenientes de SAF pueden superar fácilmente varios párrafos de extensión, lo que dificulta su lectura rápida y la extracción de elementos clave del incidente. Los analistas deben identificar manualmente la naturaleza del hecho, los roles involucrados, los medios comisivos, los objetos sustraídos, el contexto temporal, entre otros. Este proceso no solo consume tiempo, sino que introduce heterogeneidad entre analistas, afectando la consistencia del tratamiento de la información.

El desafío es, por tanto, desarrollar un sistema automático capaz de generar resúmenes breves, coherentes y representativos del contenido esencial del relato, preservando la información jurídica y operativa necesaria para apoyar el trabajo de fiscalía.

En la práctica, los relatos de hechos pueden ocupar varios párrafos de texto continuo, mientras que los campos de resumen asociados son comparativamente breves; esta asimetría entre entradas largas y salidas cortas sugiere la necesidad de un enfoque de resumen *abstractive*, capaz de reformular el contenido en lugar de limitarse a seleccionar oraciones aisladas.

2.3. Motivación

La motivación principal del proyecto surge de tres necesidades reales:

1. **Reducir la carga operativa** en la lectura y categorización de casos.

2. **Estandarizar la calidad** del análisis preliminar, evitando la variabilidad humana.
3. **Aprovechar los avances recientes en modelos de lenguaje basados en Transformers**, que han demostrado un desempeño sobresaliente en tareas de resumen *abstractive*.

El avance de modelos como T5 y técnicas de ajuste eficientes como LoRA permite entrenar modelos especializados utilizando infraestructura local, con costos computacionales razonables y sin comprometer la privacidad de datos sensibles.

2.4. Solución propuesta

Este proyecto propone el diseño y construcción de un ***pipeline* de procesamiento y modelado** que abarca:

1. **Extracción y limpieza** de relatos desde SAF/SIRA.
2. **Construcción de *datasets* depurados** con columnas `relato_saf` (*input*) y detalle (resumen humano).
3. **Entrenamiento de un modelo T5 ‘general’ mediante *fine-tuning*** de un modelo T5-base pre-entrenado, siguiendo un enfoque de resumen *abstractive encoder-decoder* como el descrito en la sección sobre modelos de lenguaje (sección 3.1).
4. **Entrenamiento de adaptadores LoRA especializados por tipo de delito**, reutilizando el modelo T5 general del punto anterior sobre los cuatro códigos de delito de mayor frecuencia en el *dataset*, bajo el supuesto de que las narrativas dentro de cada delito son relativamente homogéneas. Los códigos específicos seleccionados se describen en la sección 5.1.8.

5. **Evaluación automática** con métricas ROUGE.
6. **Generación de resúmenes por lote sobre conjuntos de prueba**, para comparar el desempeño del modelo T5 general y de las variantes LoRA.

Esta arquitectura apunta a desarrollar un sistema reproducible, auditable y escalable, capaz de entregar resúmenes consistentes para apoyar el análisis criminal.

2.5. Contribuciones

El proyecto aporta:

- **Un *pipeline* E2E** completo desde datos reales institucionales hasta modelos entrenados.
- **Un modelo T5 ajustado al dominio delictual**, capaz de resumir relatos policiales con vocabulario y estructuras propias del contexto.
- **Modelos LoRA por delito** que permiten especialización con bajo costo computacional.
- Una base metodológica replicable para futuras tareas de NLP dentro de la Fiscalía (clasificación, extracción de entidades, imputación de roles, etc.).

2.6. Estructura del informe

El documento se organiza de la siguiente manera:

- Capítulo 2: Introducción, donde se presenta el contexto institucional, el problema, la motivación, la solución propuesta y las contribuciones.

- Capítulo 3: Trabajo relacionado, con énfasis en modelos de lenguaje, T5, LoRA y métricas de evaluación de resúmenes.
- Capítulo 4: Hipótesis y objetivos del proyecto.
- Capítulo 5: Datos y metodología, incluyendo la construcción y limpieza del *dataset*, la definición de *splits* y el diseño del *pipeline* de modelado.
- Capítulo 6: Resultados cuantitativos y cualitativos del modelo T5 general y de las variantes LoRA por tipo de delito.
- Capítulo 7: Conclusiones, limitaciones y líneas de trabajo futuro.
- Capítulo 8: Bibliografía.
- Capítulo 9: Anexos.

3. Trabajo Relacionado

3.1. Modelos de lenguaje y tareas de resumen automático

El resumen automático de textos es una tarea clásica del procesamiento de lenguaje natural (PLN), orientada a condensar un texto preservando su significado esencial. Tradicionalmente se divide en enfoques extractivos, que seleccionan las oraciones más relevantes del texto original, y enfoques abstractivos, que generan frases nuevas para expresar la misma información (Nenkova & McKeown, 2011).

Con la expansión de los modelos de lenguaje preentrenados (PLMs), el resumen abstractivo ha alcanzado calidades cercanas a las humanas en *benchmarks* de noticias y

dominios generales, gracias a modelos basados en Transformers y entrenamiento previo masivo seguido de ajuste fino supervisado (Raffel, y otros, 2020).

En el ámbito jurídico, el problema de resumir sentencias largas y complejas ha recibido atención específica. Estudios recientes muestran que tanto métodos extractivos clásicos como modelos neuronales modernos enfrentan desafíos particulares: longitud extrema de los documentos, estructura retórica (hechos, fundamentos, decisión) y la necesidad de preservar precisión jurídica (Shukla, y otros, 2022). Estos trabajos destacan que los resúmenes legales suelen ser más extensos que en otros dominios, y que la evaluación debe considerar no sólo cobertura de contenido, sino también corrección y utilidad para el trabajo profesional.

3.2. El modelo T5 (*Text-to-Text Transfer Transformer*)

El modelo T5, propuesto por (Raffel, y otros), reformula las tareas de PLN bajo un paradigma unificado *text-to-text*, donde tanto la entrada como la salida son secuencias textuales (por ejemplo, “*summarize*: [texto fuente]” → [resumen generado]). T5 se basa en la arquitectura Transformer *encoder–decoder* (Vaswani, y otros, 2017) y se pre-entrena mediante una tarea de “*span corruption*” sobre un corpus web de gran escala (C4), lo que le permite aprender representaciones contextuales ricas.

Este diseño unificado facilita reutilizar el mismo modelo para traducción, clasificación o resumen, cambiando sólo la instrucción textual.

El *fine-tuning* consiste en ajustar los parámetros del modelo sobre un conjunto de pares (entrada → salida) adaptados al dominio. En este proyecto, se parte de un modelo T5 en

español pre-entrenado para generar resúmenes breves de relatos de hechos judiciales, entrenado con pares contruidos a partir de registros institucionales SAF/SIRA.

3.3. LoRA (*Low-Rank Adaptation*) y *fine-tuning* ligero

El ajuste fino completo de modelos grandes implica actualizar todos los pesos, lo que requiere recursos computacionales elevados y almacenamiento redundante de múltiples copias del modelo. LoRA (*Low-Rank Adaptation*) aborda este problema insertando matrices de bajo rango en las proyecciones lineales del Transformer; durante el entrenamiento sólo se actualizan esos parámetros adicionales, manteniendo fijos los pesos originales del modelo base (Hu, y otros, 2021).

Esta estrategia permite: (1) reducir el uso de memoria, (2) entrenar adaptaciones especializadas en hardware más limitado y (3) almacenar múltiples versiones del modelo (por ejemplo, una por tipo de delito) como pequeños adaptadores en lugar de duplicar todo T5. En el contexto del proyecto, se explora precisamente este esquema: un T5 “general” entrenado sobre todos los relatos y adaptaciones LoRA especializadas por `cod_delito`, reutilizando el mismo modelo base y comparando si estas especializaciones mejoran la calidad de los resúmenes.

3.4. Evaluación de resúmenes en PLN y en el dominio legal

La evaluación automática de resúmenes suele basarse en métricas de solapamiento léxico. ROUGE-1/2/L (Lin, 2004) calcula precisión, *recall* y *F-score* a partir del número de n-gramas compartidos entre el resumen generado y un resumen de referencia, y se ha

consolidado como estándar de facto en *benchmarks* como CNN/DailyMail y otros conjuntos de noticias.

Sin embargo, múltiples estudios empíricos han mostrado que las métricas de n-gramas pueden correlacionar de manera limitada con juicios humanos, especialmente cuando existen paráfrasis o reformulaciones semánticamente equivalentes ((Fabbri, y otros, 2021); (Nenkova & McKeown, 2011)). Para mitigar esta limitación, se han propuesto métricas semánticas basadas en *embeddings* contextualizados, como BERTScore, que comparan las representaciones de *tokens* generadas por modelos tipo BERT para medir similitud de significado más allá del vocabulario exacto (Zhang, Kishore, Wu, Weinberger, & Artzi, 2020).

En el dominio legal, las dificultades son aún mayores: los documentos son muy largos, las decisiones tienen estructura retórica compleja y pequeños errores pueden alterar el sentido jurídico del texto. (Shukla, y otros) muestran que, en resúmenes de sentencias, las métricas automáticas no capturan completamente criterios clave para abogados y jueces, como corrección de la *ratio decidendi* o precisión en los hechos relevantes, por lo que recomiendan combinar evaluaciones automáticas con juicios expertos segmentados por tipo de contenido (hechos, argumentos, decisión).

A la luz de esta literatura y de las restricciones de tiempo y cómputo del proyecto, se optó por:

- Utilizar ROUGE-1/2/L como métrica cuantitativa principal para comparar el modelo T5 general con las variantes LoRA, dada su amplia adopción y su bajo costo computacional.
- Complementar los resultados con una evaluación cualitativa manual de muestras de resúmenes, enfocada en claridad, fidelidad al relato original y utilidad para el análisis criminal, medida mediante una escala ordinal de tres niveles (0–2) aplicada por analistas humanos.

Métricas semánticas como BERTScore se consideran adecuadas como línea futura para profundizar el análisis, pero no se implementaron en esta versión del trabajo debido a la priorización de experimentos de modelado (T5 vs. LoRA) y al mayor costo de cómputo asociado.

4. Hipótesis y Objetivos

A continuación se presenta una hipótesis general y sus hipótesis específicas, junto con el objetivo general y los objetivos específicos del proyecto.

4.1. Hipótesis general

Se plantea que un modelo generativo de tipo *encoder–decoder*, ajustado mediante *fine-tuning* (T5) sobre relatos delictuales, es capaz de producir resúmenes breves, coherentes y fieles al contenido esencial del hecho, alcanzando una calidad suficiente para apoyar el análisis criminal cuando se evalúa mediante métricas automáticas (ROUGE) y revisión cualitativa experta.

4.2. Hipótesis específicas

- **H1.** El entrenamiento supervisado del modelo T5 sobre pares (`relato_saf` → `detalle/resumen`) extraídos desde SAF/SIRA permite obtener resúmenes de hechos delictuales nuevos que sean breves y coherentes con el relato original.
- **H2.** La especialización por tipo de delito mediante adaptaciones ligeras LoRA, construidas a partir del T5 general, mejora sistemáticamente la calidad de los resúmenes (medida con ROUGE) respecto del modelo T5 general para los códigos de delitos seleccionados.

4.3. Objetivo general

Desarrollar y evaluar un sistema automático de resumen de relatos delictuales basado en modelos de lenguaje preentrenados (T5 y adaptaciones LoRA), capaz de generar resúmenes breves y coherentes a partir de datos institucionales SAF/SIRA, y de cuantificar la mejora obtenida al especializar el modelo por tipo de delito.

4.4. Objetivos específicos

- **OE1:** Construir un *pipeline* reproducible para la extracción, limpieza y estructuración de relatos delictuales provenientes de SAF/SIRA, generando los *datasets* necesarios para el modelado (corpus general y subconjuntos por `cod_delito`).

- **OE2:** Analizar y caracterizar el *dataset* en términos de calidad textual, longitud y distribución por códigos de delito, con el fin de verificar que su tamaño, nivel de ruido y la cantidad de ejemplos por cada código de delito sean razonables para entrenar los modelos de resumen definidos en el proyecto.
- **OE3:** Entrenar un modelo T5 general mediante *fine-tuning* de un modelo T5 pre-entrenado, utilizando un corpus del orden de 41 mil pares (relato → resumen), evaluando su desempeño con métricas ROUGE y una revisión cualitativa de muestras de resúmenes.
- **OE4:** Entrenar modelos LoRA especializados para los cuatro delitos de mayor frecuencia en el corpus, reutilizando el modelo T5 general como base y respetando un flujo reproducible de entrenamiento y validación.
- **OE5:** Realizar inferencias y comparaciones cuantitativas entre el modelo T5 general y las variantes LoRA por `cod_delito` utilizando métricas ROUGE, complementadas con un análisis cualitativo de ejemplos.
- **OE6:** Documentar el pipeline extremo a extremo (E2E) y dejar una base metodológica replicable para futuras aplicaciones de PLN dentro de la Fiscalía (por ejemplo, resumen, clasificación o extracción de entidades).

5. Datos y Metodología

5.1. Datos

El *dataset* utilizado en este proyecto proviene de registros institucionales extraídos desde el Sistema de Apoyo a Fiscales (SAF) y procesados a través del Sistema de Información y Registro de Análisis (SIRA). Cada registro corresponde a un relato descriptivo de un hecho delictual, acompañado de un resumen breve elaborado manualmente por funcionarios del Ministerio Público.

5.1.1. Origen de los datos

Los datos fueron obtenidos desde SAF/SIRA mediante procedimientos internos de extracción autorizados, ejecutados a través del script `generar_dataset.py`, que genera los siguientes archivos iniciales:

- `data/raw/RUC_RESUMEN_con_RELATO_SAF.csv` (versión cruda).
- `data/processed/RUC_RESUMEN_con_RELATO_SAF_LIMPIO.csv` (versión limpia inicial).
- `data/processed/RUC_RESUMEN_con_RELATO_SAF_LIMPIO_FILTRADO.csv` (versión limpia depurada de glosas administrativas).

Este último archivo constituye la base final sobre la cual se construyeron los *splits* de entrenamiento y validación para el modelo T5 general.

5.1.2. Estructura del *dataset*

El *dataset* final contiene dos columnas centrales:

- **relato_saf**: texto extenso del relato del hecho delictual.
- **detalle**: resumen humano sobre el relato del hecho delictual que actúa como etiqueta supervisada.

Ambos campos presentan variabilidad en longitud, estilo y nivel de ruido textual, lo que hace necesaria una etapa de limpieza y normalización previa al modelado.

5.1.3. Proceso de limpieza

El proceso de limpieza implementado en el *pipeline* considera las siguientes transformaciones principales:

- Eliminación de registros con campos vacíos o claramente incompletos en *relato_saf* o *detalle* (valores nulos, solo espacios en blanco, glosas genéricas sin contenido descriptivo como “RELATO EN CARPETA DIGITAL” o “DENUNCIA DIGITALIZADA”, y casos por debajo de umbrales mínimos de longitud: < 200 caracteres en *relato_saf* o < 15 en *detalle*).
- Corrección de problemas de codificación (comillas, saltos de línea, espacios sobrantes).
- Normalización de texto mediante conversión explícita a *str* y eliminación de caracteres extraños.
- Depuración de duplicados accidentales originados desde SAF.

- Eliminación sistemática de registros cuyo `relato_saf` contiene glosas administrativas genéricas (por ejemplo, “RELATO EN CARPETA DIGITAL”, “VER RELATO EN ARCHIVO ADJUNTO” y variantes similares), que no aportan información semántica útil para el aprendizaje de resúmenes.

Tras este proceso, el *dataset* queda consolidado en `RUC_RESUMEN_con_RELATO_SAF_LIMPIO_FILTRADO.csv`, que se utiliza como fuente única para la construcción de los *splits* del modelo general y para la generación de subconjuntos por `cod_delito`.

5.1.4. División en conjuntos de entrenamiento y validación

A partir del archivo limpio y filtrado se generaron *splits* aleatorios mediante el módulo `src/splits_t5_general.py`, con una semilla fija para garantizar reproducibilidad.

El *split* principal para el modelo general se definió como:

- **80% entrenamiento** → `data/processed/DATASET_T5_GENERAL_TRAIN.csv`
- **20% validación** → `data/processed/DATASET_T5_GENERAL_VAL.csv`

Esta partición corresponde a **33.100** registros para entrenamiento y **8.276** para validación, de acuerdo con los reportes de validación generados automáticamente por el *pipeline*.

5.1.5. Estadísticas descriptivas

A partir de los reportes de validación generados por el módulo `src/validar_data.py` para los archivos `DATASET_T5_GENERAL_TRAIN.csv` y `DATASET_T5_GENERAL_VAL.csv`, se obtuvieron las siguientes estadísticas de longitud para las columnas `relato_saf` (texto de entrada) y `detalle` (resumen humano):

Longitud del relato (relato_saf) – número de caracteres por registro

- **Entrenamiento** (DATASET_T5_GENERAL_TRAIN.csv, **33.100** filas):
 - media \approx 2.906 caracteres
 - mediana \approx 2.345
 - p25 \approx 1.762
 - p75 \approx 3.318
 - mínimo \approx 202
 - máximo \approx 63.351

- **Validación** (DATASET_T5_GENERAL_VAL.csv, **8.276** filas):
 - media \approx 2.931 caracteres
 - mediana \approx 2.348
 - p25 \approx 1.760
 - p75 \approx 3.309
 - mínimo \approx 204
 - máximo \approx 54.134

En términos más intuitivos, unos 2.900 caracteres corresponden aproximadamente a del orden de 450–550 palabras en español. Es decir, un relato típico equivale aproximadamente a una página de texto continuo en formato informe, con una cola larga de casos extraordinariamente extensos (sobre 50 mil caracteres).

Longitud del resumen (detalle) – número de caracteres por registro

- **Entrenamiento:**
 - media \approx 227 caracteres

- mediana ≈ 183
- p25 ≈ 110
- p75 ≈ 265
- mínimo ≈ 15
- máximo ≈ 4.258

- **Validación:**

- media ≈ 231 caracteres
- mediana ≈ 185
- p25 ≈ 111
- p75 ≈ 272
- mínimo ≈ 15
- máximo ≈ 4.194

En promedio, por tanto, los resúmenes son alrededor de **10–13 veces más cortos** que los relatos originales, pero también presentan alta variabilidad en longitud.

En cuanto al mínimo observado (≈ 15 caracteres) corresponde a un número reducido de resúmenes extremadamente breves, pero que contienen una glosa operativa distinta de las glosas administrativas eliminadas en la etapa de limpieza, por lo que se mantuvieron en el análisis. También se considera el hecho de que al ser resúmenes que analistas profesionales asociaron al relato original, siempre deberían entregar algún valor o funcionalidad.

5.1.6. Relación con el límite de *tokens* del modelo T5:

Para el entrenamiento se utiliza el **tokenizador de subpalabras propio de T5**, que en el código se carga con `T5Tokenizer.from_pretrained("t5-base")`, y se fija un límite de **768** *tokens* para la secuencia de entrada (`max_input_length = 768`). Este valor corresponde a los *tokens* de ese tokenizador y se definió a partir de la distribución empírica de longitudes observada en el corpus, documentada en `configs/general.yaml`. En la práctica:

- un relato “típico” de ~2.900 caracteres se sitúa en el rango de 500–700 *tokens*, por lo que entra completo dentro de la ventana de 768 *tokens*;
- sólo una fracción pequeña de casos extremadamente largos (cola de la distribución) requiere truncamiento, lo que se considera aceptable dado que estos relatos son atípicos y, con frecuencia, contienen repeticiones o información administrativa.

La columna **detalle** se resume con un límite de **128** *tokens* (`max_target_length = 128`), suficiente para cubrir prácticamente todos los resúmenes observados (medias ~230 caracteres, con máximos alrededor de 4.200 caracteres; es decir, claramente por debajo de 128 *tokens* en la gran mayoría de los casos).

5.1.7. Implicancias para el tipo de modelo

Como se discutió en la sección 3.1, el problema de resumen puede abordarse con enfoques *extractive* o *abstractive*. Estas estadísticas no “obligan” por sí mismas a usar modelos *abstractive* en lugar de *extractive*, pero sí plantean desafíos relevantes que motivan la elección de un enfoque generativo *encoder–decoder* (T5):

- La gran diferencia de longitud entre relato y resumen exige condensar información dispersa a lo largo de varios párrafos; esto es más natural de abordar con modelos que puedan reformular el contenido, y no solo seleccionar oraciones.
- Los relatos presentan ruido, redundancia y glosas administrativas, por lo que un modelo generativo puede aprender a omitir segmentos poco informativos y a reescribir los hechos en un lenguaje más sintético.
- En textos tan largos, los métodos extractivos simples (por ejemplo, ranking de oraciones) tienden a seleccionar fragmentos extensos sin reescritura, lo que dificulta llegar a resúmenes tan breves como los utilizados operacionalmente en SIRA.

Por estos motivos, y en línea con la literatura reciente de resumen en dominios complejos (legal, biomédico, etc.), el proyecto se centra en modelos *abstractive* basados en T5 y sus adaptaciones LoRA, utilizando ROUGE y evaluación cualitativa experta para medir si los resúmenes generados son suficientemente breves, claros y fieles al contenido esencial del relato.

5.1.8. Selección de *sub-datasets* por código de delito

A partir del conjunto limpio general RUC_RESUMEN_con_RELATO_SAF_LIMPIO_FILTRADO.csv se construyó un resumen de frecuencias por tipo de delito, agrupando por la variable `cod_delito` y contabilizando el número de RUC asociados a cada categoría. Este resumen se materializó en el archivo RESUMEN_cod_delito_counts.csv, que contiene para cada código de delito su descripción (`delito`) y el conteo de RUC (`n_ruc`).

Con el fin de estudiar el efecto de la especialización del modelo, y en coherencia con la distribución empírica del corpus, se seleccionaron los cuatro códigos de delito con mayor número de RUC en el período analizado:

- **802:** Robo con intimidación.
- **803:** Robo con violencia.
- **804:** Robo por sorpresa.
- **867:** Robo de vehículo motorizado por sorpresa, violencia o intimidación.

La decisión de acotar la especialización LoRA a estos cuatro códigos obedece a una combinación de criterios prácticos y sustantivos: concentran una proporción significativa del volumen total de RUC del corpus general y permiten mantener el número de entrenamientos dentro de los recursos de tiempo y cómputo disponibles, dejando otros tipos de delito como línea de trabajo futuro.

Para cada uno de estos delitos se construyó un *sub-dataset* específico filtrando el conjunto general por `cod_delito` y conservando únicamente los registros que cumplieran con los criterios de calidad previamente descritos (existencia de `relato_saf` y `detalle`, limpieza de *placeholders* y longitud mínima del relato). Estos *sub-datasets* se almacenaron en la carpeta `data/processed/` bajo la forma `DATASET_cod_delito_{cod}.csv` (por ejemplo, `DATASET_cod_delito_802.csv`, `DATASET_cod_delito_803.csv`, etc.).

Cada *sub-dataset* mantiene la misma estructura de columnas que el conjunto general y es utilizado posteriormente para el entrenamiento de los modelos LoRA especializados y su evaluación comparativa frente al modelo T5 general.

Dado que estos sub-conjuntos son particiones del *dataset* limpio general, sus distribuciones de longitud de texto son análogas a las reportadas en la sección de estadísticas descriptivas globales (5.1.5); por ello, en este apartado se reportan únicamente estadísticas sintéticas (tamaños muestrales y longitudes promedio) en lugar de replicar figuras descriptivas adicionales para cada código de delito.

De igual modo y como complemento de las estadísticas globales, la Tabla 1 resume el tamaño muestral y las longitudes promedio de texto para cada *sub-dataset* por código de delito, según los reportes generados automáticamente por `validar_data.py`. En todos los casos, las longitudes se sitúan en rangos comparables al conjunto general, lo que refuerza que los modelos especializados LoRA operan sobre distribuciones de texto similares, pero restringidas a un dominio delictual específico, lo que permite concluir que los corpus por código de delito tienen un tamaño suficiente, longitudes manejables para la ventana de 768 *tokens* de T5 y un nivel de ruido controlado, lo que respalda su idoneidad para el entrenamiento de los modelos de resumen propuestos.

Tabla 1: Resumen estadístico de *sub-datasets* por `cod_delito`

cod_delito	Delito (resumen)	Registros (n)	Long. media relato_saf (carac.)	Long. media detalle (carac.)
802	Robo con intimidación	14.719	≈ 2.805	≈ 225
803	Robo con violencia	7.899	≈ 3.356	≈ 240
804	Robo por sorpresa	2.433	≈ 1.861	≈ 141
867	Robo de vehículo motorizado por sorpresa/viol./intimid.	3.015	≈ 3.481	≈ 288

5.2. Metodología

5.2.1. Enfoque general y diseño experimental

El proyecto se plantea como un problema de **resumen automático supervisado** de relatos delictuales registrados en SAF/SIRA. Para abordarlo se diseñó un *pipeline* de experimentos en dos niveles:

1. **Modelo general de propósito amplio**, basado en T5-base, entrenado sobre el conjunto completo de relatos SAF limpios.
2. **Modelos especializados por tipo de delito**, obtenidos mediante adaptación LoRA sobre el modelo general para los cuatro códigos de delito más frecuentes (802, 803, 804 y 867).

Este diseño responde directamente a la hipótesis de trabajo:

“la adaptación LoRA especializada por tipo de delito puede mejorar la calidad de los resúmenes, medida con métricas ROUGE, respecto a un modelo general entrenado sobre todos los relatos.”

En términos metodológicos, el flujo completo se organiza en etapas secuenciales y reproducibles:

(0) Construcción de *datasets*: a partir de consultas a SAF se genera un *dataset* maestro con el RUC, el relato original (*relato_saf*) y el resumen manual asociado (*detalle*). Sobre este *dataset* se obtiene, por un lado, un corpus general y, por otro, *sub-datasets* filtrados por *cod_delito* para los delitos más frecuentes.

(1) Limpieza y validación de datos: se eliminan glosas sin contenido real (por ejemplo, referencias a “carpeta digital”) y relatos demasiado breves, se normalizan campos y se validan los *splits* mediante análisis descriptivo de longitudes y revisión manual de muestras.

(2) Entrenamiento del modelo T5 general: se realiza un *split* 80/20 mediante muestreo aleatorio simple en entrenamiento y validación y se ajusta T5-base con hiperparámetros fijados en un archivo de configuración (*general.yaml*), usando como entrada el relato y como salida el resumen manual.

(3) Entrenamiento de modelos LoRA especializados: para cada código de delito seleccionado se reutiliza el T5 general como base y se entrena únicamente un adaptador LoRA sobre el *sub-dataset* correspondiente, manteniendo la misma tokenización y esquema de entrada/salida.

(4) Inferencia y generación de resúmenes: se generan predicciones del modelo T5 general y de cada modelo LoRA sobre los conjuntos de validación definidos, tanto en modo completo como en muestras reducidas para análisis cualitativo.

(5) Evaluación cuantitativa: se calcula ROUGE-1, ROUGE-2 y ROUGE-L sobre los resúmenes generados, comparando sistemáticamente el desempeño del T5 general con sus versiones LoRA especializadas por delito.

Todas estas etapas se encapsulan en un *pipeline* programático único (*pipeline.py*), que fija rutas de entrada y salida, ejecuta los distintos *scripts* (generación de *datasets*, creación de *splits*, entrenamiento, inferencia y evaluación) y permite reproducir los resultados a partir de los mismos datos crudos.

5.2.2. Modelo base: T5 (*Text-to-Text Transfer Transformer*)

Tal como se describió en la sección de trabajo relacionado, T5 es un modelo *encoder-decoder* que plantea todas las tareas de PLN bajo un esquema *text-to-text*, lo que permite reutilizar la misma arquitectura para clasificación, traducción o resumen cambiando únicamente la instrucción y los datos de entrenamiento.

En este proyecto se utiliza la variante **T5-base** (≈ 220 M parámetros), cargada desde la implementación oficial disponible en HuggingFace Transformers mediante la instrucción `T5ForConditionalGeneration.from_pretrained("t5-base")`. El modelo se emplea como base única del *pipeline*:

- En la etapa de **entrenamiento general**, T5-base se ajusta (*fine-tuning*) sobre el corpus completo de pares (relato_saf \rightarrow detalle), dando origen al modelo que en el repositorio se denomina `t5_general`.
- En las etapas de **especialización por delito**, este mismo modelo `t5_general` se reutiliza como punto de partida para inyectar los adaptadores LoRA por `cod_delito` (802, 803, 804 y 867), manteniendo congelados los pesos originales y entrenando solo los parámetros adicionales de baja dimensión.
- En la etapa de **inferencias y evaluación**, el *checkpoint* más reciente de `t5_general` se utiliza tanto como *baseline* (resúmenes generados por el modelo general) así como modelo base sobre el cual se acoplan los adaptadores LoRA para generar las predicciones especializadas.

La elección de T5-base obedece a un equilibrio entre **capacidad representacional** y **restricciones computacionales**. Modelos más pequeños (por ejemplo, *T5-small*) resultan

más rápidos pero tienden a perder capacidad para manejar relatos extensos y heterogéneos como los que se observan en SAF/SIRA, mientras que variantes mayores (*T5-large* o T5-3B) exceden los recursos disponibles en la infraestructura local. T5-base permite: (1) trabajar con una ventana de hasta 768 *tokens* de entrada y 128 *tokens* de salida (configurada en `configs/general.yaml`), suficiente para cubrir la gran mayoría de los relatos y resúmenes observados; (2) mantener tiempos de entrenamiento y de inferencia razonables en *hardware Apple Silicon*, sin sacrificar la posibilidad de capturar patrones lingüísticos complejos propios del dominio delictual.

De esta manera, T5-base actúa como núcleo común de todo el *pipeline*: primero se adapta al dominio institucional mediante *fine-tuning* general y luego sirve como “tronco” sobre el cual se montan las especializaciones LoRA, lo que simplifica la arquitectura, reduce el costo de almacenamiento de modelos y facilita la comparación directa entre el enfoque general y el especializado.

5.2.3. Tokenización y truncamiento

El *pipeline* utiliza el tokenizador de subpalabras asociado al modelo T5, cargado en el código mediante `T5Tokenizer.from_pretrained(...)`. Este tokenizador implementa un esquema tipo *SentencePiece*, en el cual las palabras se descomponen en unidades más pequeñas (*sub-words*) que permiten manejar vocabulario abierto, errores tipográficos y variación morfológica propia de los relatos delictuales. En la práctica, tanto los relatos (`relato_saf`) como los resúmenes humanos (`deta11e`) se representan como secuencias de *tokens* de este vocabulario, que luego son procesadas por el *encoder–decoder* de T5.

Los límites máximos de secuencia —768 *tokens* para la entrada y 128 *tokens* para la salida— se definen en el archivo de configuración `configs/general.yaml` y se aplican de manera consistente en todas las etapas del *pipeline*: entrenamiento del modelo general (`train_t5.py`), entrenamiento de modelos LoRA por delito (`train_lora.py`) e inferencias tanto del modelo general como de las variantes especializadas (`inference_t5_general.py` e `inference_lora.py`). En todos los casos, el tokenizador realiza truncamiento hacia la derecha cuando un texto supera el máximo permitido, de modo que sólo se conserva el prefijo inicial del relato o del resumen.

Esta decisión se apoya en el análisis previo de longitudes descrito en la sección de datos (5.1.5–5.1.6), donde se observó que un relato “típico” de aproximadamente 2.900 caracteres se sitúa en el rango de 500–700 *tokens*, por lo que entra completo dentro del límite de 768 *tokens*, y que sólo una fracción pequeña de casos de cola larga requiere truncamiento. En esos casos, el modelo ve una versión recortada del relato, pero que suele contener la parte inicial más estructurada del texto (descripción del hecho, tipo de delito, dinámica principal), lo que se considera suficiente para la tarea de resumen operativo.

En el lado del objetivo (`detaille`), el límite de 128 *tokens* garantiza que los resúmenes generados permanezcan breves y comparables a los resúmenes humanos observados en el corpus, cuya longitud media se sitúa del orden de 200–230 caracteres. Esto evita que el modelo produzca salidas excesivamente largas y refuerza la naturaleza sintética del resumen como producto final.

Finalmente, mantener los mismos parámetros de tokenización y truncamiento para el T5 general y para todos los modelos LoRA asegura que las comparaciones sean justas: tanto

el modelo *baseline* como sus versiones especializadas reciben entradas representadas con el mismo esquema de subpalabras y bajo la misma ventana de contexto. De este modo, las diferencias observadas en las métricas de evaluación (ROUGE) pueden atribuirse a la especialización del modelo, y no a cambios en la forma de codificar o recortar los textos.

5.2.4. Hiperparámetros de entrenamiento

El ajuste fino del modelo T5 general se controla a través del archivo de configuración `configs/general.yaml` y del *script* `train_t5.py`, que fijan un conjunto acotado de hiperparámetros, manteniendo el resto en sus valores por defecto de la librería Transformers. En el experimento final se utilizaron los siguientes valores:

- **Modelo base:** t5-base (≈ 220 millones de parámetros).
- **Tamaño de *batch*:** 8 ejemplos por dispositivo tanto en entrenamiento como en validación (`train_batch_size = 8`, `eval_batch_size = 8`).
- **Número de épocas:** 3 (`num_train_epochs = 3`).
- **Tasa de aprendizaje:** `learning_rate = 5e-5`.
- **Regularización:** `weight_decay = 0.01`.
- **Warmup:** sin pasos de calentamiento explícitos (`warmup_steps = 0`).
- **Semilla de reproducibilidad:** `seed = 2025`

Estos valores buscan un equilibrio entre **capacidad del modelo, tamaño del *dataset* y restricciones computacionales**. Un *batch* de tamaño 8 permite aprovechar la memoria disponible en el *hardware Apple Silicon (backend MPS de PyTorch)* sin generar errores de desbordamiento, manteniendo a la vez una estimación razonablemente estable del

gradiente en cada actualización. Tres épocas de entrenamiento resultan suficientes para que el modelo se adapte al dominio, considerando que se trabaja con un corpus relativamente homogéneo de alrededor de 41 mil pares (relato → resumen); pruebas preliminares con más épocas no mostraron mejoras significativas y aumentaban el riesgo de sobreajuste.

La tasa de aprendizaje $5e-5$ corresponde a un valor estándar recomendado para modelos T5-base en tareas supervisadas, mientras que el término de regularización `weight_decay = 0.01` ayuda a evitar que los pesos crezcan en exceso, controlando el sobreajuste sin introducir complejidad adicional en el *scheduler*. El proyecto no introduce un esquema de *warmup* sofisticado, dado que los experimentos mostraron una convergencia estable con tasa fija y sin degradaciones visibles en la pérdida de validación.

Finalmente, el entrenamiento del T5 general utiliza el optimizador Adafactor, recomendado para modelos de la familia T5 por su menor consumo de memoria y buen comportamiento en tareas de generación. En este caso se emplea con tasa de aprendizaje fija (sin *step-size* relativo) y sin *warmup* interno, lo que simplifica el control del experimento: el valor de `learning_rate` definido en `general.yaml` es el que efectivamente gobierna la magnitud de las actualizaciones. Mantener una semilla global fija (`seed = 2025`) asegura que la división `train/val`, el orden de los *batches* y los resultados reportados sean reproducibles a partir del mismo código y configuración.

En las secciones siguientes, los modelos LoRA especializados heredan estas decisiones globales (longitudes de secuencia, tamaños de *batch*, tasa de aprendizaje y semilla), modificando únicamente la arquitectura mediante la inyección de adaptadores de bajo

rango sobre el T5 general y el subconjunto de datos utilizado en cada entrenamiento por código de delito.

5.2.5. Entrenamiento de modelos LoRA

Sobre la base del modelo T5 general entrenado en la sección anterior, se entrenaron adaptadores LoRA especializados para cada uno de los cuatro códigos de delito seleccionados (802, 803, 804 y 867). El objetivo es contrastar un modelo “de propósito general”, entrenado con todos los relatos, con variantes más específicas que solo ven ejemplos de un tipo de delito, manteniendo constante la arquitectura base y el esquema de tokenización.

El procedimiento seguido es el siguiente:

- **Punto de partida común:**

Para cada código de delito se reutiliza el modelo `t5_general` como base congelada. En lugar de volver a ajustar todos sus parámetros, se insertan adaptadores LoRA de baja dimensión en las capas de atención del `Transformer`. De este modo, cada modelo especializado “aprende a corregir” el comportamiento del T5 general en su dominio (por ejemplo, robos con intimidación), sin perder el conocimiento previo adquirido en el corpus completo.

- **Construcción del *dataset* por delito:**

Para cada código se utiliza el archivo `DATASET_cod_delito_{cod}.csv` (por ejemplo, `DATASET_cod_delito_802.csv`), que contiene únicamente los registros delictuales

asociados a ese tipo de hecho. Sobre estos archivos se aplica nuevamente la función de limpieza `limpiar_relatos_sin_contenido`, que elimina:

- glosas administrativas sin contenido real (como “RELATO EN CARPETA DIGITAL” y variantes), y
- relatos demasiado breves según el umbral definido en el módulo de validación.

El resultado se guarda en archivos `DATASET_cod_delito_{cod}_FILTRADO.csv`, que actúan como fuente para el entrenamiento LoRA.

- **División interna en entrenamiento y validación:**

Cada *sub-dataset* filtrado se convierte en un DataFrame con las columnas de interés (`relato_saf` como entrada y `detalle` como salida), se limpian valores vacíos y se barajan las filas con una semilla fija. A continuación, se realiza una división interna en:

- ~80% para entrenamiento,
- ~20% para validación.

Esta partición se implementa dentro del *script* `train_lora.py`, de modo que cada modelo especializado dispone de su propio conjunto de validación, independiente del *split* del modelo general, pero construido con la misma lógica.

- **Adaptación LoRA sobre T5 general:**

El modelo T5 general se carga desde el directorio `models/t5_general` y se envuelve con una configuración LoRA que:

- agrega matrices de bajo rango en las proyecciones de atención (módulos q y v),
- fija un rango pequeño (r) y un factor de escala (α) moderado, y
- aplica un *dropout* específico para estos adaptadores.

Durante el entrenamiento, los pesos originales de T5 permanecen congelados y solo se actualizan los parámetros añadidos por LoRA. Esto reduce el costo computacional y el tamaño en disco de cada modelo especializado, permitiendo mantener una copia ligera por código de delito.

- **Esquema de entrenamiento y métricas internas:**

Los modelos LoRA heredan los hiperparámetros globales definidos en `configs/general.yaml`:

tamaño de *batch*, tasa de aprendizaje, número de épocas, longitudes máximas de entrada y salida, y semilla de reproducibilidad. El entrenamiento se orquesta con el Trainer de la librería Transformers, que:

- calcula la pérdida en el conjunto de validación al final de cada época,
- guarda periódicamente *checkpoints* de los adaptadores, y
- registra un historial de métricas (pérdida de entrenamiento y validación por época y por paso).

Al finalizar, los adaptadores LoRA y el tokenizador se almacenan en carpetas del tipo `models/lora_{cod}`, mientras que los logs de entrenamiento se guardan como archivos CSV en `results/metrics/lora_{cod}_metrics.csv`. Esto permite inspeccionar, para cada delito, cómo evoluciona la pérdida y si el modelo converge de forma estable.

En síntesis, el entrenamiento LoRA implementa un **segundo nivel de especialización**: en lugar de entrenar desde cero un T5 distinto por delito (lo que sería costoso y difícil de mantener), se parte de un modelo general ya adaptado al dominio SAF/SIRA y se añaden pequeñas capas entrenables que capturan patrones narrativos específicos de cada categoría

de robo. Este diseño permite evaluar de manera directa la hipótesis H2, comparando la calidad de los resúmenes generados por el T5 general y por sus variantes LoRA bajo el mismo esquema de datos, tokenización y métricas.

5.2.6. Evaluación

La evaluación del sistema se realizó en dos niveles complementarios: (1) métricas automáticas de resumen basadas en solapamiento léxico y (2) revisión cualitativa manual de ejemplos seleccionados. El objetivo es comparar de forma sistemática el desempeño del modelo T5 general frente a sus variantes especializadas LoRA, bajo un esquema de datos y tokenización homogéneo.

En el nivel cuantitativo se utilizó como referencia el conjunto de validación definido en la sección de datos.

Para el modelo T5 general, se generaron resúmenes sobre todo el *split* de validación (`DATASET_T5_GENERAL_VAL.csv`) mediante el *script* de inferencia (`inference_t5_general.py`, modo `full`), produciendo un archivo de predicciones con las columnas `detalle` (resumen humano) y `resumen_t5` (salida del modelo). A partir de este archivo se calcularon las métricas ROUGE-1, ROUGE-2 y ROUGE-L utilizando la implementación oficial de la librería `evaluate`, obteniendo valores promedio de F1 sobre todos los pares (`detalle`, `resumen_t5`). Estas métricas se almacenan en `results/metrics/t5_general_val_rouge.csv` y constituyen la línea base cuantitativa del experimento.

Para los modelos LoRA por delito, la evaluación se diseñó de forma análoga, pero condicionada al dominio específico de cada código de delito. En cada caso, a partir del archivo `DATASET_cod_delito_{cod}FILTRADO.csv` se reconstruyó internamente un *split* de entrenamiento y validación con la misma proporción $\sim 80/20$ y la misma semilla utilizada en el entrenamiento (`train_lora.py`). Sobre el subconjunto de validación resultante se generaron, con el *script* `inference_lora.py`, dos columnas de predicciones: `resumen_t5` (modelo general aplicado al *sub-dataset*) y `resumen_lora` (modelo LoRA especializado). El archivo comparativo `t5_vs_lora_cod{cod}val.csv` se evaluó luego con `eval_rouge_lora.py`, que calcula ROUGE-1/2/L por separado para cada columna y reporta, además, la diferencia LoRA - T5 en F1. Estas tablas se guardan en `results/metrics/lora{cod}_val_rouge.csv` y permiten verificar, para cada delito, si la especialización LoRA aporta una mejora consistente sobre el modelo general.

En todos los casos las métricas ROUGE se calcularon con la opción `use_stemmer=True` de la implementación de `evaluate`, que aplica el algoritmo de *stemming* de Porter por defecto. En este proyecto no se aplicó ningún *stemming* adicional previo: el único *stemmer* utilizado es el que incorpora internamente la implementación de ROUGE.

Dado que se trata de un *stemmer* originalmente diseñado para inglés y el corpus está en español, estos valores se interpretan como una medida aproximada y homogénea de solapamiento para comparar configuraciones entre sí, más que como estimaciones lingüísticamente óptimas del desempeño absoluto. Es decir, que estos valores se interpretan como indicadores descriptivos de desempeño promedio, no como pruebas estadísticas formales; su principal utilidad en este proyecto es comparar de manera

homogénea diferentes configuraciones de modelo (T5 general vs. LoRA 802/803/804/867) sobre conjuntos de datos equivalentes.

Como complemento a las métricas automáticas, se realizó una evaluación cualitativa manual sobre una muestra de 17 resúmenes generados (5 del modelo general y 12 de los modelos LoRA por `cod_delito`). Esta evaluación se estructuró en torno a tres dimensiones:

- **Claridad:** qué tan comprensible y bien redactado es el resumen.
- **Fidelidad al hecho:** en qué medida el resumen respeta la información relevante del relato original, sin inventar ni distorsionar elementos clave.
- **Utilidad para el análisis criminal:** cuánto aporta el resumen al trabajo cotidiano de análisis (identificación del tipo de delito, modus operandi, objeto del delito, circunstancias principales, etc.).

En cada dimensión se utilizó una escala ordinal de tres niveles (0–2):

- Claridad del texto
 - **0 – Claridad insuficiente:** El resumen es difícil de entender por contener frases truncadas o desordenadas, errores que impiden seguir la idea principal o ambigüedades fuertes o incoherencia global.
 - **1 – Claridad parcial:** Se entiende la idea general, pero hay repeticiones o frases poco naturales, errores de redacción que obligan a releer y/o se requiere cierto esfuerzo para captar el hecho principal.
 - **2 – Claridad adecuada:** El resumen es fácil de leer porque tiene frases completas y comprensibles, errores menores (mayúsculas, tildes, lemas

normalizados) que no afectan la comprensión y se entiende de inmediato qué ocurrió.

- Fidelidad al relato original

- **0 – No fiel / distorsiona el hecho:** El resumen describe algo distinto porque cambia el tipo de hecho (por ejemplo un asalto a banco vs robo a taxista), cambia el resultado central (consumado vs frustrado) y/o introduce información importante que no está en el relato o contradice lo ocurrido.
- **1 – Fidelidad parcial:** Mantiene la idea general (tipo de delito, relación básica víctima–imputado) pero omite o altera aspectos relevantes (lugar, objeto del robo, uso de arma, consumado vs frustrado). El analista podría usarlo como primera aproximación, pero tendría que revisar el relato original para detalles finos.
- **2 – Alta fidelidad** ($\approx 80\%$ o más de los elementos relevantes): Conserva correctamente el tipo de hecho, los roles principales y el resultado, las omisiones son menores (no cambian la calificación jurídica ni la dinámica central), no hay contradicciones importantes respecto del relato original.

- Utilidad para el análisis criminal

- **0 – Poco o nada útil:** No permite entender con seguridad qué pasó, o induce a una interpretación equivocada (por ejemplo, parece un asalto a banco cuando en realidad es un asalto a conductor de taxi) o no se puede usar sin volver sí o sí al relato.

- **1 – Utilidad limitada / parcial:** Sirve como orientación general (tipo de hecho, víctima, imputados) pero la falta de precisión obliga a contrastar con el relato para elementos clave (modus operandi, objeto, lugar).
- **2 – Alta utilidad:** Permite entender rápidamente el hecho en términos relevantes para análisis criminal (tipo de delito, modus operandi básico, objeto afectado, contexto mínimo como si fue en la vía pública/banco/colectivo, etc).

En los Anexos, junto a cada ejemplo se indica entre paréntesis qué se espera del resumen generado (por ejemplo, tipo de delito, acción principal, objeto del delito y circunstancias relevantes), lo que explicita el criterio utilizado para asignar los niveles de la escala. La pauta fue aplicada por un único evaluador experto con experiencia profesional en análisis criminal y en el uso de SIRA, por lo que los resultados deben interpretarse como una primera aproximación experta, complementaria a las métricas automáticas.

6. Resultados

En esta sección se presentan los resultados obtenidos con la configuración final del *pipeline* descrito en el capítulo anterior.

En primer lugar, se resumen los resultados cuantitativos del modelo T5 general entrenado sobre el conjunto completo de relatos SAF limpios. Luego, se analiza el efecto de la especialización mediante adaptadores LoRA por tipo de delito, comparando sistemáticamente su desempeño con el modelo general. Finalmente, se incluyen ejemplos

cualitativos de resúmenes generados, con el objetivo de complementar las métricas automáticas y discutir la pertinencia práctica de los resultados en el contexto de la Fiscalía.

6.1. Configuración de los experimentos de evaluación

Todos los resultados reportados en este capítulo se basan en la ejecución final y estable del *pipeline*. En esta configuración, el modelo de base corresponde a T5-base pre-entrenado, ajustado (*fine-tuning*) sobre el corpus completo de relatos SAF/SIRA, utilizando los *splits* de entrenamiento y validación generados automáticamente a partir del *dataset* limpio depurado.

El conjunto de datos se divide en aproximadamente un 80% de registros para entrenamiento y un 20% para validación, mediante un muestreo aleatorio reproducible.

En el caso del modelo T5 general:

- Sólo el 80% de entrenamiento se utiliza para actualizar los parámetros del modelo.
- El 20% restante se emplea como conjunto de validación interna durante el entrenamiento (para monitorear la pérdida y el desempeño), pero no participa en el cálculo de gradientes.

Posteriormente, y dado que en este proyecto no se definió un tercer conjunto de prueba independiente, ese mismo *split* de validación se reutiliza como *proxy* de conjunto de prueba para calcular las métricas ROUGE y reportar los resultados finales. Esto implica que el modelo ha sido evaluado varias veces sobre esos datos (aunque nunca entrenado directamente sobre ellos), por lo que las métricas deben interpretarse como una estimación razonable pero levemente optimista de su rendimiento fuera de muestra.

Un esquema análogo se utiliza para los modelos LoRA por tipo de delito. Para cada código seleccionado (802, 803, 804 y 867) se parte del *sub-dataset* correspondiente y se realiza una división interna 80/20 en entrenamiento y validación. El 20% se usa como validación durante el entrenamiento del adaptador y, posteriormente, como base para calcular ROUGE y comparar el desempeño del modelo LoRA frente al T5 general en ese mismo dominio. De nuevo, no se dispone de un conjunto de prueba adicional, por lo que las comparaciones cuantitativas se interpretan como aproximaciones consistentes dentro del marco del experimento.

La evaluación cuantitativa se realiza con métricas ROUGE-1, ROUGE-2 y ROUGE-L en su variante F1, calculadas sobre:

- El modelo T5 general aplicado al conjunto de validación completo.
- Cada modelo LoRA aplicado al conjunto de validación de su respectivo *sub-dataset*, comparando siempre contra el resumen humano (detalle) y contra el desempeño del modelo general en ese mismo conjunto.

ROUGE-1 y ROUGE-L permiten medir el solapamiento de contenido entre el resumen generado y el resumen de referencia, mientras que ROUGE-2 es más exigente y evalúa la coherencia a nivel de pares de palabras (bigramas). Estas métricas son el estándar de facto en la evaluación automática de sistemas de resumen, y han sido ampliamente utilizadas en la literatura reciente sobre modelos de tipo Transformer.

6.2. Resultados cuantitativos del T5 general

Para el modelo T5 general, entrenado sobre el 80% de los relatos limpios y evaluado en el 20% de validación descrito en la sección anterior, se calcularon las métricas ROUGE-1, ROUGE-2 y ROUGE-L en su variante F1. Los resultados agregados sobre todo el conjunto de validación se resumen en la Tabla 2.

Tabla 2: Métricas ROUGE (F1) del modelo T5 general sobre el conjunto de validación

MEDIDA	RESULTADO (F1)
ROUGE - 1 (unigramas)	F1 \approx 0,301
ROUGE - 2 (bigramas)	F1 \approx 0,133
ROUGE - L (subsecuencia común más larga)	F1 \approx 0,253

En términos cualitativos, un valor de ROUGE-1 cercano a 0,30 indica que, en promedio, el modelo recupera una fracción apreciable del vocabulario relevante de los resúmenes humanos (a nivel de palabras individuales), mientras que el ROUGE-2 en torno a 0,13 refleja que la coherencia a nivel de pares de palabras es más exigente y, como es habitual en tareas de resumen *abstractive*, presenta valores más bajos. El ROUGE-L \approx 0,25 sugiere que la estructura general de las frases generadas captura parcialmente secuencias de texto similares a las de los resúmenes de referencia, aunque sin llegar a una coincidencia completa.

Estos resultados son coherentes con lo observado en la literatura para tareas de resumen *abstractive* en dominios complejos, donde la longitud de los documentos, la variabilidad estilística y la presencia de ruido textual tienden a limitar el solapamiento léxico exacto entre los resúmenes automáticos y los humanos. En el contexto de este proyecto, las métricas obtenidas se interpretan como una primera línea base razonable: muestran que el

modelo T5 general es capaz de aproximar el contenido esencial de los resúmenes operativos generados en SIRA, pero dejan espacio para mejoras adicionales, ya sea mediante especialización por tipo de delito (modelos LoRA) o mediante evaluaciones complementarias de tipo semántico y cualitativo.

6.3. Resultados cuantitativos de modelos LoRA especializados por delito

A partir del modelo T5 general entrenado sobre todo el corpus, se ajustaron cuatro adaptadores LoRA especializados para los códigos de delito 802, 803, 804 y 867. En todos los casos, la evaluación se realizó sobre el mismo conjunto de validación utilizado para LoRA, comparando directamente el desempeño del modelo general (T5) y su versión especializada (LoRA) sobre exactamente los mismos relatos y resúmenes de referencia. De este modo, cualquier diferencia en las métricas ROUGE puede atribuirse al efecto de la especialización y no a cambios en los datos.

Tabla 3: Métricas ROUGE en validación por código de delito: comparación entre T5 general y modelos LoRA especializados

Código de delito	Métrica	T5 general (F1)	LoRA (F1)	Δ LoRA – T5
802 ROBO CON INTIMIDACIÓN	ROUGE-1	0,374	0,378	0,004
	ROUGE-2	0,194	0,196	0,002
	ROUGE-L	0,320	0,324	0,003
803 ROBO CON VIOLENCIA	ROUGE-1	0,326	0,331	0,005
	ROUGE-2	0,146	0,148	0,002
	ROUGE-L	0,272	0,276	0,004
804 ROBO POR SORPRESA	ROUGE-1	0,250	0,379	0,129
	ROUGE-2	0,072	0,190	0,118
	ROUGE-L	0,203	0,342	0,138
867 ROBO.VEHÍC. MOTORIZADO POR SORPRESA,VIOLENCIA O INTIMID.	ROUGE-1	0,376	0,381	0,005
	ROUGE-2	0,185	0,188	0,003
	ROUGE-L	0,313	0,317	0,004

La Tabla 3 resume los resultados obtenidos. Para cada código de delito se reportan los valores F1 de ROUGE-1, ROUGE-2 y ROUGE-L del modelo T5 general y del modelo LoRA correspondiente, junto con la diferencia $\Delta(\text{LoRA} - \text{T5})$. En los delitos 802, 803 y 867 se observan **mejoras pequeñas pero sistemáticamente positivas**: por ejemplo, en el código 802 ROUGE-1 pasa de aproximadamente 0,374 a 0,378 ($\Delta \approx +0,004$), ROUGE-2 de 0,194 a 0,196 ($\Delta \approx +0,002$) y ROUGE-L de 0,320 a 0,324 ($\Delta \approx +0,003$). Un patrón similar se repite en los códigos 803 y 867, donde los incrementos de LoRA sobre T5 se sitúan en el orden de milésimas, pero siempre a favor del modelo especializado.

El caso del código 804 es distinto: allí la especialización LoRA produce una **mejora sustantiva en todas las métricas**. ROUGE-1 aumenta desde aproximadamente 0,250 a 0,379 ($\Delta \approx +0,129$), ROUGE-2 desde 0,072 a 0,190 ($\Delta \approx +0,118$) y ROUGE-L desde 0,203 a 0,342 ($\Delta \approx +0,138$). Este salto sugiere que, para este tipo de delito, el patrón narrativo de los relatos es lo bastante homogéneo como para que un adaptador especializado pueda explotar regularidades adicionales que el modelo general no captura completamente.

En conjunto, los resultados cuantitativos indican que **en ningún caso la especialización LoRA degrada el desempeño del modelo general**, y que en todos los códigos de delito se observan mejoras en ROUGE-1, ROUGE-2 y ROUGE-L, aunque de magnitud variable. Desde el punto de vista de la hipótesis de trabajo, esto respalda la idea de que incorporar adaptadores LoRA por tipo de delito permite, al menos en este escenario experimental, igualar o superar sistemáticamente la calidad de los resúmenes producidos por el modelo T5 general.

6.4. Análisis cualitativo de resúmenes

Además de las métricas automáticas, se realizó una revisión cualitativa de ejemplos de resúmenes generados, tanto del modelo T5 general como de las variantes LoRA especializadas por delito. El objetivo fue observar, de manera más cercana al uso real, hasta qué punto los resúmenes producidos son claros, fieles al contenido esencial del relato y útiles para el trabajo cotidiano de análisis criminal.

La selección de casos se hizo a partir de las salidas de inferencia en modo muestra generadas por el *pipeline*. En el caso del modelo T5 general, se utilizaron registros del archivo `t5_general_muestra_val.csv`, mientras que para los modelos LoRA se trabajó con los archivos `t5_vs_lora_cod_{cod}_val_muestra.csv` para cada código de delito (802, 803, 804 y 867). En todos los casos se privilegió el análisis de ejemplos representativos de la dinámica delictual (robos con intimidación, violencia, sorpresa y robos de vehículos), evitando reproducir textualmente información sensible de los casos. A partir de esta revisión, se distinguieron tres dimensiones principales:

1. Claridad y legibilidad del texto generado.

En términos generales, tanto el T5 general como los modelos LoRA tienden a producir resúmenes gramaticalmente correctos y legibles, con frases bien formadas y un vocabulario consistente con el ámbito delictual. En algunos ejemplos, especialmente cuando el relato original es muy largo o repetitivo, el modelo general genera resúmenes algo genéricos (“la víctima fue abordada por sujetos desconocidos...”) que, si bien son comprensibles, aportan poca especificidad.

Los modelos LoRA, en particular para el código 804 (robo por sorpresa), muestran una tendencia a generar textos más directos y enfocados en la acción principal (“la víctima caminaba por la vía pública cuando un sujeto le sustrajo el teléfono celular desde el bolsillo...”), lo que mejora la rapidez de lectura y reduce rodeos narrativos.

2. Fidelidad al contenido esencial del relato original.

Siguiendo la revisión manual, en la mayoría de los ejemplos el modelo T5 general logra preservar adecuadamente elementos básicos del hecho: tipo de delito, relación víctima–imputado, medio comisivo principal y objeto sustraído. Sin embargo, se observaron varios casos en que el modelo omite detalles relevantes para el análisis (por ejemplo, el uso de armas específicas, la participación de más de un imputado o ciertas circunstancias de tiempo y lugar), especialmente cuando estos aparecen hacia el final de relatos muy extensos.

En los modelos LoRA, la fidelidad al tipo de delito tiende a ser más consistente: en los robos con intimidación (802) y con violencia (803), por ejemplo, el resumen especializado suele enfatizar de manera más explícita la presencia de amenazas o agresiones físicas, mientras que en el delito 867 (robo de vehículo motorizado) se describe con mayor precisión la sustracción de vehículos y las circunstancias asociadas (uso de intimidación, huida, lugar de estacionamiento, etc.). No obstante, persisten limitaciones propias del enfoque *abstractive*: en algunos casos el modelo infiere o reformula detalles que no aparecen literalmente en el relato, lo que refuerza la necesidad de considerar estos resúmenes como apoyo al análisis y no como reemplazo de la lectura del caso completo.

3. Utilidad operativa para el análisis criminal.

Desde una perspectiva práctica, los resúmenes generados por el T5 general ya entregan un valor agregado evidente frente al relato completo: permiten, en pocas líneas, entender “qué pasó”, con qué tipo de delito se relaciona el incidente y cuáles fueron los elementos centrales del hecho. Esto resulta especialmente útil para tareas de filtrado inicial de casos, revisión rápida de grandes volúmenes de RUC o identificación preliminar de patrones.

Aplicando la escala cualitativa descrita en la sección 5.2.6 a los ejemplos de los Anexos, se observa en primer lugar que, tanto para el modelo T5 general como para las variantes LoRA, la **claridad alcanza niveles altos en la mayoría de los casos** (valores 1 o 2). Es decir, los resúmenes generados suelen ser breves, comprensibles y con una redacción suficiente para que un analista entienda rápidamente qué ocurrió.

La variabilidad aparece sobre todo en las dimensiones de fidelidad al hecho y utilidad para el análisis criminal. En varios ejemplos el resumen automatizado omite detalles relevantes (por ejemplo, el tipo específico de bien sustraído, la forma de aproximación al afectado o una circunstancia agravante), o bien reformula el contenido en términos demasiado generales. En esos casos las puntuaciones cualitativas se sitúan en los niveles 0 o 1, reflejando que, aunque el texto sea legible, no siempre captura toda la información operativamente necesaria.

En el **código 804 (robo por sorpresa)** la especialización LoRA muestra las mejoras cualitativas más claras respecto del T5 general. En varios ejemplos el modelo especializado pasa de niveles intermedios a niveles altos (2) en fidelidad y utilidad, describiendo con mayor precisión la dinámica típica de este delito (aprovechamiento del descuido, contacto breve y sustracción de especies de fácil transporte). Esta ganancia es

coherente con el aumento sustantivo de las métricas ROUGE reportado para este código en la sección 6.3.

En los códigos **802 (robo con intimidación)**, **803 (robo con violencia)** y **867 (robo de vehículo)** el panorama es más matizado. Aunque las métricas ROUGE muestran mejoras sistemáticas de LoRA frente al modelo general, la evaluación cualitativa evidencia un comportamiento mixto: existen ejemplos en que LoRA mejora la precisión en el tipo de delito o enfatiza mejor la acción principal, pero también casos en los que el modelo especializado omite detalles que el T5 general sí recogía (por ejemplo, la forma concreta de intimidación, la participación de más de un autor o la referencia explícita al vehículo sustraído). En estos ejemplos puntuales las puntuaciones de fidelidad y/o utilidad para LoRA son iguales o incluso inferiores a las del modelo general.

En síntesis, el análisis cualitativo confirma que:

- **El modelo T5 general** es capaz de generar resúmenes breves, legibles y, en muchos casos, operativamente útiles como primera aproximación al contenido del caso.
- **Las adaptaciones LoRA** mantienen este nivel de legibilidad y aportan mejoras cualitativas claras y consistentes en el delito 804, mientras que en los códigos 802, 803 y 867 su aporte depende más del caso concreto, combinando ejemplos donde LoRA mejora la síntesis con otros donde introduce pérdidas de detalle relevantes.
- Persisten limitaciones propias de los modelos de resumen *abstractive* (omisión de matices, posible generalización excesiva o, en casos aislados, inferencias no deseadas), por lo que el uso de estos resúmenes se recomienda como herramienta

de apoyo al trabajo experto y no como sustituto de la lectura íntegra del relato cuando se deban adoptar decisiones jurídicas o estratégicas de alta relevancia.

En el siguiente capítulo se presentan las conclusiones generales del estudio, sus limitaciones y posibles líneas de trabajo futuro.

7. Conclusiones

7.1. Síntesis de resultados

El objetivo general de este trabajo fue explorar la viabilidad de generar resúmenes automáticos de relatos delictuales a partir de registros SAF/SIRA, utilizando modelos de la familia T5 y adaptaciones LoRA especializadas por tipo de delito. Para ello, se construyó un *pipeline* completo que abarca desde la extracción y depuración de datos hasta el entrenamiento, inferencia y evaluación de los modelos.

En primer lugar, se consolidó un *dataset* general de relatos y resúmenes operativos (detalle) a partir de RUC con información disponible en SAF. Sobre este conjunto se aplicaron procesos de limpieza orientados a eliminar registros sin contenido real (por ejemplo, relatos con glosas genéricas del tipo “RELATO EN CARPETA DIGITAL”) y a garantizar una longitud mínima suficiente tanto en el relato como en el resumen. A partir de este corpus limpio se generaron, de forma reproducible, los *splits* de entrenamiento y validación utilizados en el modelado.

El análisis descriptivo de longitudes y frecuencias por código de delito mostró que los relatos y resúmenes tienen tamaños compatibles con los límites de 768 y 128 *tokens* definidos para T5 y que los delitos más frecuentes concentran un volumen suficiente de

ejemplos para entrenar y validar las adaptaciones LoRA. En conjunto, estos resultados indican que el corpus disponible es razonable en términos de tamaño, nivel de ruido y distribución por códigos de delito para la tarea de resumen planteada.

Sobre el *dataset* general se entrenó un modelo T5-base pre-entrenado ajustado por *fine-tuning* a la tarea de transformar el relato completo en su correspondiente resumen operativo. Evaluado sobre el conjunto de validación, este modelo alcanzó valores de ROUGE-1 en torno a 0,30, ROUGE-2 cercanos a 0,13 y ROUGE-L alrededor de 0,25. Estas métricas indican que el modelo es capaz de recuperar una parte relevante del contenido de los resúmenes humanos, aunque con limitaciones naturales en un dominio complejo y altamente variable.

En una segunda etapa, se construyeron *sub-datasets* por código de delito para los cuatro tipos más frecuentes en el corpus (802, 803, 804 y 867), aplicando la misma lógica de limpieza y manteniendo la estructura entrada-salida (`relato_saf` \rightarrow `detalle`). A partir del modelo T5 general se entrenaron adaptadores LoRA especializados para cada uno de estos delitos, evaluando después su desempeño en los respectivos conjuntos de validación.

Los resultados cuantitativos muestran que, para los cuatro delitos analizados, los modelos LoRA **igualan o mejoran sistemáticamente** las métricas ROUGE del modelo general. En los códigos 802, 803 y 867 las mejoras son pequeñas pero consistentes (incrementos de ROUGE-1, ROUGE-2 y ROUGE-L del orden de milésimas). En el caso del código 804, en cambio, se observa una mejora sustantiva: ROUGE-L aumenta aproximadamente 0,14 puntos F1 respecto del T5 general, acompañada de incrementos similares en ROUGE-1 y ROUGE-2.

Finalmente, el análisis cualitativo de ejemplos tomados de los conjuntos de validación confirma que los resúmenes generados son, en general, **legibles, coherentes y razonablemente fieles** al contenido esencial de los relatos. El modelo T5 general ya aporta valor al condensar relatos extensos en pocas líneas, mientras que los modelos LoRA tienden a refinar esta capacidad en dominios específicos, destacando con mayor precisión el tipo de delito, el modus operandi y ciertos elementos contextuales.

7.2. Evaluación de las hipótesis

El proyecto partió de dos ideas principales (explícitas o implícitas) que pueden formularse del siguiente modo:

- **H1:** Es posible entrenar un modelo de lenguaje tipo Transformer (T5) que genere resúmenes operativos útiles a partir de relatos delictuales completos, utilizando como supervisión los resúmenes actualmente registrados en SIRA.
- **H2:** La especialización del modelo mediante adaptadores LoRA por tipo de delito permite mejorar el desempeño del modelo T5 general en la tarea de resumen, al menos en los delitos con mayor frecuencia en el corpus.

Respecto de **H1**, los resultados apoyan su validez en el contexto de este experimento. El modelo T5 general, entrenado sobre el *dataset* completo, logra métricas ROUGE compatibles con lo reportado en la literatura para tareas de resumen *abstractive* en dominios complejos, y genera resúmenes que, en la práctica, permiten entender de forma rápida el núcleo del hecho. Si bien los resúmenes no son perfectos y presentan omisiones y generalizaciones, cumplen razonablemente el rol de “primer acercamiento” al contenido de la causa.

En cuanto a **H2**, los resultados cuantitativos y cualitativos también son coherentes con la hipótesis. En ninguno de los delitos analizados la especialización LoRA empeora las métricas ROUGE respecto del modelo general; por el contrario, en todos los casos se observan mejoras, aunque de magnitud distinta. El código 804 (robo por sorpresa) constituye el ejemplo más claro, con un aumento significativo en todas las métricas, lo que sugiere que, en dominios donde la narrativa de los hechos es más homogénea, los adaptadores LoRA son especialmente eficaces. En los delitos 802, 803 y 867 las mejoras son más moderadas, pero sistemáticas, lo que refuerza la idea de que la especialización, incluso con cambios paramétricos acotados, aporta valor sobre un modelo ya entrenado de forma general.

Desde el punto de vista cualitativo, esta mejora es especialmente clara y consistente en el delito 804 (robo por sorpresa), mientras que en los códigos 802, 803 y 867 los resultados son más heterogéneos: se observan ejemplos donde LoRA aporta resúmenes más útiles que el modelo general y otros donde introduce pérdidas de detalle relevantes, lo que sugiere que la especialización por delito no garantiza una mejora uniforme en todas las dimensiones de calidad.

No obstante, es importante recordar que, por razones de tiempo y recursos, el proyecto no utilizó un conjunto de prueba completamente independiente: las métricas se calcularon sobre el mismo *split* de validación utilizado durante el entrenamiento (sin que este haya sido usado para actualizar pesos, pero sí para monitoreo). Por ello, las conclusiones deben interpretarse como evidencia fuerte dentro de este marco experimental, pero todavía abierta a confirmación adicional en evaluaciones futuras más estrictas.

7.3. Aporte práctico para la Fiscalía

Más allá de los resultados numéricos, el proyecto ofrece varios aportes concretos para la labor de la Fiscalía:

1. Demostración de factibilidad técnica.

Se demuestra que, con los datos ya disponibles en SAF/SIRA y con recursos computacionales moderados, es posible entrenar modelos de resumen basados en Transformers que produzcan salidas útiles sin necesidad de etiquetados adicionales ni cambios en los sistemas de origen.

2. Pipeline reproducible de datos y modelado.

El trabajo deja instalado un *pipeline* documentado y automatizable que:

- extrae, limpia y organiza los datos,
- genera *splits* de entrenamiento/validación reproducibles,
- entrena y evalúa modelos T5 y LoRA,
- y produce artefactos (CSV, métricas, ejemplos) que pueden ser auditados.

Esto facilita tanto la repetición de los experimentos como su extensión a nuevos periodos, regiones o tipos de delito.

3. Potencial de apoyo al análisis criminal.

Los resúmenes generados no pretenden reemplazar la lectura completa de la carpeta, pero sí pueden:

- acelerar el filtrado inicial de grandes volúmenes de casos,

- apoyar la identificación temprana de patrones (por ejemplo, series de robos con un modus operandi similar),
- y servir como insumo preliminar para otras tareas de priorización o clasificación automatizada.

En este sentido, el proyecto abre la puerta a considerar los modelos de resumen como componentes integrables en herramientas internas (por ejemplo, módulos de SIRA o paneles de análisis) que permitan a fiscales y analistas ahorrar tiempo en la revisión de textos extensos.

7.4. Limitaciones del estudio

El alcance y las conclusiones del proyecto deben leerse considerando varias limitaciones:

- **Ausencia de conjunto de prueba independiente.**

Por restricciones de tiempo y complejidad, no se generó un tercer *split* de test completamente separado. Las métricas ROUGE se calcularon sobre el mismo conjunto de validación utilizado durante el entrenamiento para monitorear el desempeño. Esto puede introducir un sesgo optimista en las estimaciones, aunque el modelo no haya sido entrenado directamente sobre esos datos.

- **Cobertura acotada de tipos de delito.**

La especialización LoRA se aplicó sólo a los cuatro códigos más frecuentes en el corpus. Si bien esto permite concentrar el análisis en delitos de alto volumen, deja fuera una parte significativa de la casuística, especialmente delitos menos frecuentes o con narrativas que pueden ser más heterogéneas o homogéneas.

- **Uso exclusivo de métricas basadas en solapamiento léxico (ROUGE).**

ROUGE es un estándar en evaluación de resumen, pero tiene limitaciones para capturar similitud semántica cuando se usan sinónimos, paráfrasis o reordenamientos de frase. No se incorporaron métricas complementarias (por ejemplo, basadas en *embeddings*) ni evaluaciones humanas sistemáticas, por lo que la calidad semántica de los resúmenes debe interpretarse también a la luz del análisis cualitativo.

- **Dependencia de reglas de limpieza y umbrales específicos.**

La construcción del *dataset* limpio implicó decisiones de diseño (umbral mínimo de longitud, patrones de “relato sin contenido”, etc.) que, si bien fueron razonadas y documentadas, podrían ajustarse en futuros trabajos. Cambios en estos parámetros podrían impactar tanto la cantidad como la calidad de los datos disponibles para entrenar.

- **Falta de integración y evaluación en entorno productivo.**

El proyecto se desarrolló en un entorno experimental aislado. No se llegó a una etapa de piloto integrado en los sistemas de la Fiscalía ni a una evaluación de uso real por parte de fiscales y analistas en su flujo de trabajo cotidiano.

7.5. Líneas de trabajo futuro

Considerando tanto los resultados como las limitaciones descritas, se identifican varias líneas de trabajo futuro:

1. Diseño de un esquema de evaluación más robusto.

- Incorporar un conjunto de prueba independiente, separado desde el inicio, para estimar con mayor rigor la capacidad de generalización de los modelos.

- Complementar ROUGE con métricas semánticas (por ejemplo, basadas en *embeddings*) y con evaluaciones humanas estructuradas, en las que fiscales y analistas valoren la utilidad de los resúmenes para tareas específicas.

2. Ampliación de la cobertura de delitos y contextos.

- Extender la generación de *sub-datasets* y adaptadores LoRA a otros tipos de delito relevantes para la Fiscalía, incluyendo aquellos menos frecuentes pero de alta complejidad.
- Explorar si la estrategia de especialización por delito se puede combinar con otros criterios, como la región geográfica, el tipo de víctima o el modus operandi.

3. Exploración de arquitecturas y configuraciones alternativas.

- Evaluar modelos más recientes o específicos para resumen (por ejemplo, variantes *encoder-decoder* de mayor tamaño o modelos instruccionales preentrenados en español), manteniendo como referencia el T5-base utilizado en este proyecto.
- Probar variantes de configuración (longitud máxima de entrada, estrategias de generación, tamaños de lote, etc.) que optimicen la relación entre calidad del resumen y costo computacional.

4. Integración gradual en herramientas internas.

- Diseñar prototipos de interfaz (por ejemplo, dentro de SIRA o de paneles analíticos) donde los resúmenes generados puedan ser consultados junto al resto de la información de la causa.
- Planificar pilotos controlados con un grupo acotado de usuarios, con seguimiento de métricas de uso y de percepción de utilidad.

5. Aspectos de gobernanza, privacidad y ética.

- Definir políticas claras sobre qué usos son apropiados para este tipo de modelos (por ejemplo, apoyo a análisis y priorización, pero no reemplazo de la lectura jurídica detallada).
- Revisar periódicamente posibles sesgos o errores sistemáticos en los resúmenes y establecer mecanismos de monitoreo y mejora continua.

En conjunto, estos desarrollos permitirían pasar desde la prueba de concepto presentada en este trabajo hacia un escenario en el que los modelos de resumen se conviertan en herramientas de apoyo confiables y auditables dentro de los procesos de análisis criminal y gestión de causas de la Fiscalía, manteniendo siempre el control experto en manos de los equipos humanos.

8. Bibliografia

- Hu, E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., . . . Chen, W. (2021). LoRA: Low-Rank Adaptation of Large Language Models.
- Lin, C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. En *Text Summarization Branches Out* (págs. 74-81). Barcelona.
- Nenkova, A., & McKeown, K. (2011). Automatic Summarization. En *Foundation and Trends in Information Retrieval* (Vol. 5, págs. 103-233).
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., . . . Liu, P. J. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. En *Journal of Machine Learning Research* 21 (págs. 1-67).
- Shukla, A., Bhattacharya, P., Poddar, S., Mukherjee, R., Ghosh, K., Goyal, P., & Ghosh, S. (2022). Legal Case Document Summarization: Extractive and Abstractive Methods and their Evaluation. En *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics* (págs. 1048-1064).
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention Is All You Need. En *31st Conference on Neural Information Processing Systems*.
- Fabrizi, A. R., Kryściński, W., McCann, B., Xiong, C., Socher, R., & Radev, D. (2021). SummEval: Re-evaluating Summarization Evaluation. En *Transactions of the Association for Computational Linguistics* (págs. 391-409).

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2020). BERTScore: Evaluating Text Generation with BERT. En *In Proceedings of the 8th International Conference on Learning Representations (ICLR)*.

9. Anexos

9.1. Anexo 1A - T5 general

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen generado T5 general	Evaluación
1	<p>DOY CUENTA A ESA FISCALIA QUE EL DIA DE HOY MIERCOLES 11 DE JUNIO DEL AO 2025 SIENDO LAS 1450 HORAS SE PRESENTO A LA CARABINERO --- DE SERVICIO GUARDIA DOTACION DE ESTA UNIDAD LA CIUDADANA AN--- 17 A OS CHILENA SOLTERA ESTUDIOS MEDIOS INCOMPLETOS ESTUDIANTE CEDULA NACIONAL DE IDENTIDAD NRO. 22.--- FECHA DE NACIMIENTO 01 -12-2007 DOMICILIADA EN EL PASAJE ESCORPION NRO. --- COMUNA DE PUENTE ALTO NUMERO DE CONTACTO 9-- QUIEN EXPUSO LO SIGUIENTEQUE EL DIA MARTES 10 DE JUNIO DEL AO 2025 SIENDO LAS 1130 HORAS EN CIRCUNSTANCIAS EN QUE LA VICTIMA SE ENCONTRABA CAMINANDO POR LAS CALLES AVENIDA VICUA MACKENNA CON INTERSECCION AVENIDA CARLOS DITTBO COMUNA DE UOA CUANDO LA AFE CTADA SIENDE DETRAS DE ELLA UNA MOTOCICLETA SE DETIENE PARA POSTERIOR A ESO ES SUJETADA POR LA NUCA CON UN INDIVIDUO MASCULINO DEL CUAL NO LOGRA RECOR DAR MAYORES ANTECEDENTES YA QUE SU ROSTRO ESTABA CUBIERTO POR UN CASCO DE SEGURIDAD DE LA MOTO CICLETA EL INDIVIDUO EMPEZO A TOCAR EL CUERPO DE LA DENU NCIANTE REGISTRANDO SU ROPA LOGRANDO SUSTRARLE SU TELEFONO CELULAR IPHONE MODELO 13 COLOR NEGRO CON CARCASA DE COLOR TRANSPARENTE COMPAIA ENTEL NR O. DE TELEFONO 9--POSTERIOR A ESO EL INDIVIDUO SE SUBE DE NUEVO A L A MOTO Y EMPIEZA A CONDUCIRLA PARA HUIR RAPIDAMENTE DEL LUGAR A RAIZ DE LO SUCEDIDO LA VICTIMA CONURRE A LA UNIDAD POLICIAL MAS CERCANA A REALIZAR LA PRESENTE DENUNCIA. HACE PRESENTE ADEMÁS QUE AL MOMENTO DE HABER A SALTADA E L SUJETO ERA REPARTIDOR DE RAPPY Y EN SUS MANOS MANTENIA UN CUCHILLO.SE HAC E PRESENTE QUE LA VICTIMA EN SU TELEFONO CELULAR MANTENIA LA CUENTA DEL BAN CO ESTADO DE SU ABUELO IDENTIFICADO COMO JUAN--- CEDULA NACIO NAL DE IDENTIDAD NRO. 7--- FECHA DE NACIMIENTO 08-10-19 57 67 AOS DESDE LA APLICACION WEB PAYU UBER EATS LE HABIAN SUSTRADO UNA SUMA DE 115.415 CIENTO QUINCE MIL CUATROCIENTOS QUINCE PESOS EN 04 GIRO S CON DIFERENTES MONTOS DOS DE 23.00 0 VEINTE Y TRES MIL PESOS UNO DE 3. 515 TRES MIL QUINIENTOS PESOS Y EL ULTIMO DE 65.900 SESENTA Y CINCO MIL NOVECIENTOS PESOS. ESPECIES SUSTRADAS01 UN TELEFONO CELULAR MARCA IPH ONE MODELO 13 COLOR NEGRO CON CARCASA DE COLOR TRANSPARENTE COMPAIA ENTE L NRO. DE TELEFONO 9--- 04 GIROS DE DINERO SUSTRADO A VALUOLA VICTIMA LO HACE EN LA SUMA DE 493.415 CUATROCIENTOS NOVENTA Y TRES MIL CUATROCI ENTOS QUINCE PESOS.SEGURONO MANTIENE.ACTUACIONES AUTONOMASART. 83 C.P.P COMO PRIMERA MEDIDA SE LE BRINDO LA ATENCION NECESARIA EN SU CALIDAD DE V ICTIMA LA CUAL SE EXPRESO LIBRE Y ESPONTANEAMENTE ACOGIENDO LA DENUNCIA RES PECTIVA HACIENDOLE ENTREGA DE LOS NUMEROS DE PLAN CUADRANTE DE SU SECTOR.DE NUNCIAS ANTERIORESNO MANTIEN E POR ESTE TIPO DE HECHO.TESTIGOSNO MANTIENE. CAMARASDESCONOCE EXISTEN EN EL LUGAR DE LOS HECHOS.CITACIONLA VICTIMA QUE DA EN ESPERA DE CITACION POR PARTE DE ESA FISCALIA.</p>	<p>SUJETO EN MOTOCICLETA ABORDA A VÍCTIMA POR LA ESPALDA, REGISTRÁNDOLA Y SUSTRAYENDO SUS PERTENENCIAS. MOMENTOS MÁS TARDE, LA VÍCTIMA SE PERCATA DE TRANSFERENCIAS FRAUDULENTAS DESDE EL DISPOSITIVO</p>	<p>VCTIMA CAMINABA POR VA PBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE SE MOVILIZABA EN MOTOCICLETA, EL QUE PORTANDO ARMA BLANCA TIPO CUCHILLO LA INTIMID Y LOGR SUSTRARLE SU TELÉFONO CELULAR PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>(Robo de celular, sujeto en motocicleta, sorpresa, luego transferencias bancarias) Claridad = 2 (texto legible y estándar). Fidelidad = 1 (describe correctamente el robo del celular por sujeto en motocicleta, omite las transferencias y añade intimidación). Utilidad = 1 (permite entender lo que pasa pero falta precisión).</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen generado T5 general	Evaluación
2	<p>RELACION DE LOS HECHOS DOY CUENTA A ESA FISCALA LOCAL DE VALPARASO, QUE EL DA DE HOY 15 DE FEBRERO 2025 A LAS 0025 HORAS, SE PRESENTA A ESTE SERVICIO DE GUARDIA, EL CIUDADANO IDENTIFICADO COMO JAVIER ---, DNI. 16---, FECHA DE NACIMIENTO 25-06-1981, 43 AOS, NACIONALIDAD VENEZOLANA, SOLTERO, ESTUDIOS MEDIOS, REPARTIDOR RAPPY, TELFONO NRO.9---, DOMICILIADO EN CALLE CANTEROS NRO. ---, DE ESTA CIUDAD VALPARASO, QUIEN EXPUSO QUE, EL DA DE HOY VIERNES 14 DE FEBRERO DEL PRESENTE AO, A LAS 2300 HORAS APROXIMADAMENTE, EN CIRCUNSTANCIAS QUE LA VCTIMA SE ENCONTRABA TRABAJANDO DE REPARTIDOR DE COMIDA RAPIDA, EN MOMENTOS QUE FUE A ENTREGAR UN PEDIDO, LA VICTIMA MANIFIESTA NO RECORDAR EL NOMBRE A QUIEN IBA LA ENTREGA. EN CIRCUNSTANCIAS QUE LA VCTIMA SE DETIENE Y ESTACIONA EN CALLE ELAS, LOGRA DIVISAR A 03 INDIVIDUOS DE SEXO MASCULINO QUE SE APROXIMAN Y A MOMENTOS QUE LE CANCELAN EL PEDIDO, LA VCTIMA SE PERCATO QUE ERA UN BILLETE DE 20.000 FALSOS, A LO QUE LE DICE QUE EL BILLETE ERA FALSO UNO DE LOS INDIVIDUOS SACA UN ARMA DE FUEGO A LO QUE LE DICE QUE ENTREGARA TODAS SUS PERTENENCIAS, QUITANDOLES LAS COSAS EMPUJANDOLO, LA VICTIMA ENTREGO LAS PERTENENCIAS QUE LOS INDIVIDUOS LE SOLICITARON Y LOS INDIVIDUOS SE FUERON DEL LUGAR EN DIRECCION DESCONOCIDA, LA VCTIMA SE ACERCA A LA UNIDAD A DEJAR LA PRESENTE DENUNCIA PARA FINES POSTERIORES. ESPECIES SUSTRADAS LA VICTIMA MANIFIESTA QUE ES UNA CELULAR MARCA REDMI 12C, COLOR NEGRO. LA DOCUMENTACION COMPLETA DE LA MOTOCICLETA MARCA LONCIN 152, AO 2023, P.P.U S---, COLOR NEGRO, SE ADJUNTA AL PRESENTE PARTE POLICIAL. LESIONES EL DENUNCIANTE MANIFIESTA NO TENER LESIONES. AVALUO 250.000 DOSCIENTOS CINCUENTA MIL PESOS TESTIGOS EL AFECTADO MANIFIESTA NO MANTENER TESTIGOS CITACION EL DENUNCIANTE QUEDO A ESPERA DE CITACION POR PARTE DE ESA FISCALA LOCAL DE VALPARASO. JULIAN --- CABO IRO. DE CARABINEROS SUBOFICIAL DE GUARDIA VO.....BO. MARIO --- TENIENTE DE CARABINEROS COMISARIO DE SERVICIO</p>	<p>JAVIER ---, VENEZOLANO RESIDENTE EN VALPARAÍSO Y TRABAJADOR COMO REPARTIDOR DE RAPPY, FUE VÍCTIMA DE UN ROBO CON VIOLENCIA. MIENTRAS ESTABA ENTREGANDO UN PEDIDO DE COMIDA RÁPIDA EN LA CALLE ELÍAS EN VALPARAÍSO, TRES INDIVIDUOS MASCULINOS SE ACERCARON A ÉL, CANCELARON EL PEDIDO CON UN BILLETE DE \$20.000 FALSOS Y, AL SER CONFRONTADOS POR BOLET, UNO DE ELLOS SACÓ UN ARMA DE FUEGO Y LE EXIGIÓ QUE ENTREGARA TODAS SUS PERTENENCIAS</p>	<p>SUJETOS DESCONOCIDOS A INGRESAN A COMIDA RPIDA Y SUSTRAN ESPECIES</p>	<p>(Víctima repartidor que fue asaltado por sujetos que hicieron pedido) Claridad = 2 (frase corta y clara). Fidelidad = 0 (cambia el contexto a "robo en local de comida rápida", lo que no coincide con el relato original). Utilidad = 0 (sirve para identificar un robo con intimidación, pero el lugar es errado, por lo que requiere revisión del relato).</p>
3	<p>DE CONFORMIDAD AL TEM CITADO EN EL RUBRO DEL OBJETIVO, EL CUAL DICE RELACION CON SITIO WEB WWW.COMISARIAVIRTUAL.CL HABILITADO PARA TODA LA CIUDADANA, QUE OTORGA LA POSIBILIDAD DE ENTREGAR INFORMACION RELACIONADA CON VENTA Y COMERCIALIZACION DE DROGAS DE FORMA ANONIMA, ES QUE ESTA UNIDAD ESPECIALIZADA TOMA CONOCIMIENTO DEL REPORTE DE INFORMACION FOLIO NRO. 593---, EN CUAL SE ENTREGAN LOS SIGUIENTES ANTECEDENTES VENTA DE DROGAS POR LA APLICACION GRINDR, EN LA CUAL OTORGA NUMERO DE WHATSAPP 9---, USA CUENTA PARA RECIBIR LOS PAGOS A NOMBRE DE CLAUDIA ---, CEDULA DE IDENTIDAD NRO. 16--- A TRAVES DE CUENTA RUT Y CUENTA MERCADO PAGO VISTA NRO. 10---, QUIEN VENDERA TODO TIPO DE DROGA. LOS HECHOS DENUNCIADOS OCURRIRAN EN LA CIUDAD ANCHUD TODOS LOS DIAS DE SEMANA EN CUALQUIER HORARIO. CONSECUENTE CON LO ANTERIOR, ESTA UNIDAD ESPECIALIZADA, SOLICITA A ESA FISCALA LOCAL, OTORGAR ORDEN DE INVESTIGAR, CON LA FINALIDAD DE CORROBORAR O DESCARTAR LA VIABILIDAD INVESTIGATIVA DE LOS HECHOS DENUNCIADOS. ES CUANTO SE INFORMA Y SOLICITA</p>	<p>DENUNCIA POR VENTA DE DROGA A TRAVES DE APLICACION GRINDR</p>	<p>VENTA DE DROGAS POR LA APLICACION GRINDR, EN LA CUAL OTORGA NUMERO DE WHATSAPP 9---, USA CUENTA PARA RECIBIR LOS PAGOS A NOMBRE DE CLAUDIA ---, CEDULA DE IDENTIDAD NRO. 16--- A TRAVES DE CUENTA RUT Y CUENTA MERCADO PAGO VISTA NRO 10---, QUIEN VENDERÁ</p>	<p>(venta de drogas por Grindr) Claridad = 2. Fidelidad = 2 (explicita la venta por Grindr y los datos de la cuenta, consistente con el detalle). Utilidad = 2.</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen generado T5 general	Evaluación
4	<p>DOY CUENTA A ESA FISCALIA QUE HOY A LAS 1600 HORAS SE PRESENTO EN LA OFICINA DE GUARDIA DE ESTA UNIDAD EL CIUDADANO LUIS --- 24 AOS CEDULA DE IDENTIDAD NRO. 19.--- FN. 12.06.1996 SOLTERO MEDIOS EMPLEADO PUBLICO CON DOMICILIO EN ESCUELA AGRICOLA NRO. --- COMUNA DE MACUL TELEFONO 2---.QUE ME DESEMPEO COMO FUNCIONARIO DE CARABINEROS Y HOY SIENDO LAS 1430 HORAS EN CIRCUNSTANCIAS QUE ME DIRIGIA A MI LUGAR DE TRABAJO UBICADO EN LA 46&DEG COMISARIA DE CARABINEROS MACUL MIENTRAS TRANSITABA POR AVDA. ESCUELA AGRICOLA CON AVDA. VICUA MACKENNA MACUL SORPRESIVAMENTE SE DETUVO UN VEHICULO COLOR BLANCO IGNORO MAS ANTECEDENTES DEL CUAL DESCENDIERON 02 INDIVIDUOS DESCONOCIDOS A ROSTRO CUBIERTO LOS CUALES ME INTIMIDARON VERBALMENTE SEALANDOME ENTREGA LAS GUEAS CULIAGO ABALANZANDOSE SOBRE MI PERSONA ANTE LO CUAL PROCEDI A DEFENDERME DE LOS INDIVIDUOS OPONIENDO TENAZ RESISTENCIA DONDE ESTOS ME AGREDIERON CON GOLPES DE PUÑO EN EL ROSTRO SIN LOGRAR SUSTRAEER NINGUNA ESPECIE PARA LUEGO ABORDAR EL MOVIL Y HUIR DEL LUGAR EN DIRECCION DESCONOCIDA.LESIONES A RAIZ DE LO ANTERIOR EL DENUNCIANTE RESULTO CON LESIONES DE CARACTER LEVES SEGUN MEDICO DE TUO DEL CESFAM SANTA JULIA CERTIFICADO QUE SE ADJUNTA.CAMARAS NO HAY EN EL LUGAR.DILIGENCIAS EN EL LUGAR EFECTUO UN AMPLIO PATRULLAJE EL CABO 2DO. DIEGO --- DE SERVICIO EN LA POBLACION Y DE ESTA DOTACION SIN RESULTADOS POSITIVOS POR E L MOMENTO.CITACION EL DENUNCIANTE QUEDO EN ESPERA DE CITACION POR PARTE DE ESA FISCALIA.</p>	<p>VÍCTIMA CAMINANDO CUANDO SUJETOS BAJAN DESDE AUTOMÓVIL BLANCO PARA INTIMIDARLO, GOLPEARLO PERO NO LOGRAN ROBAR ESPECIES.</p>	<p>VCTIMA CAMINABA POR VA PBLICA, INSTANTES EN QUE FUE ABORDADA POR 2 SUJETOS QUE DESCENDIERON DESDE VEHICULO, LOS QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMIDARON Y LOGRARON SUSTRAEERLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>(intento frustrado de robo a transeúnte por sujetos movilizados en vehículo) Claridad = 2 (igual plantilla que otros robos). Fidelidad = 1 (Se rescata la movilización de sujetos, pero añade elementos que no existen, como arma de fuego y cambia a consumado el hecho). Utilidad = 1 (sirve para ver que hay un ataque con violencia, pero induce a error sobre el resultado y dinámica; requiere leer el relato).</p>
5	<p>DOY CUENTA A ESA FISCALIA SE PRESENTA ANTE EL CABO 2&DEG IGNACIO --- LA SEORITA TANIA --- CEDULA DE IDENTIDAD NRO. 18.--- QUIEN EXPUSOQUE EL DIA DE HOY SIENDO LAS 2230 DE LA NOCHE EN CIRCUNSTANCIAS QUE TERMINABA SUS CLASES EN LA UNIVERSIDAD SAN SEBASTIAN UBICADA EN CARMEN SYLVA NRO. 2444 COMUNA DE PROVIDENCIA INSTANTES QUE ESPERABA EL UBER CON SU TELEFONO CELULAR EN LA MANO EN EL EXTERIOR DE DICHA UNIVERSIDAD MOMENTOS EN QUE UNA MOTOCICLETA DE ROJO CON BLANCO DESCONOCIENDO PLACA PATE NTE MARCA Y MODELO YA QUE LA PLACA PATENTE ESTABA TAPADA Y CON UN BOLSO DE COLOR ROJO DE LA EMPRESA DELIVERY RAPPI CON SU CONDUCTOR CON VESTIMENTA DE ROPA OSCURA ESTE SUJETO POR LA ESPALDA LA EMPUJA CON SU HOMBRO A LA VICTIMA LOGRANDO DESESTABILIZARLA Y FORCEJEANDO ENTRE AMBOS LOGRANDO EL MOTOCICLISTA QUITARLE SU TELEFONO CELULAR MARCA APPLE MODELO IPHONE 14 COLOR NEGRO DE 128 GB DE LA EMPRESA ENTEL.AVALUO LA VICTIMA LO HACE POR UNA SUMA DE 800 .000 OCHOCIENTOS MIL PESOS TESTIGOS ANTE ESA FISCALIA.CAMARAS ANTE ESA FISCALIA.CITACION LA VICTIMA QUEDA EN ESPERA DE CITACION POR PARTE DE ESA FISCALIA.</p>	<p>VÍCTIMA SE ENCONTRABA ESPERANDO UBER, CUANDO SUJETOS EN MOTO CON MOCHILA DE RAPP PASAN POR SU LADO, LA EMPUJAN, FORCEJEAN CON ELLA Y SUSTRAEEN SU CELULAR</p>	<p>VCTIMA SE ENCONTRABA EN VA PBLICA, CUANDO ES ABORDADA POR UNA MOTOCICLETA CON BLANCO Y UN BOLSO, QUIEN LA INTIMIDA Y SUSTRAE SU CELULAR</p>	<p>(robo de celular a víctima que esperaba Uber) Claridad = 2 Fidelidad = 2 (representa bien la dinámica: sujetos en moto, fuerza/intimidación, sustracción del celular). Utilidad = 2.</p>

9.2. Anexo 1B - LoRA 802 (Robo con intimidación)

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_802	Evaluación
1	<p>DOY CUENTA A ESA FISCALIA LOCAL QUE EL DIA 20 DE DICIEMBRE DEL AO 2023 A LAS 1530 HORAS SE PRESENTO ANTE ELC.P.R PABLO --- DE DOTAC ION DE ESTA UNIDAD Y DE SERVICIO DE PRIMER TUO GUARDIA LA DENUNCIANTEIVAN A BETINA ---30 AOS CEDULA DE IDENTIDAD NRO. 26.--- NACIONALIDAD VENEZOLANA SOLTERA ESTUDIOS UNIVERSITARIOS COMPLETO CONTADORA FECHA DE NACIMIENTO 17.08.1993 CON DOMICILIO CALLE DIAGONAL VICUA MACKENNA NRO. --- COMUNA DE SANTIAGO TELEFONO 9--- CORREO ELECTRONICO ---GMAIL.COM QUIEN EXPUSOQUE EL DIA MIERCOLES 20 DE DICIEMBRE DEL AO EN CURSO SIENDO LAS 0850 HORAS APROXIMADAMENTE LA DENUNCIANTE IBA ABORDA DEL BUS DE LOCOMOCION COLECTIVA DEL RECORRIDO NRO. 210 MOMENTOS EN QUE SE D ISONIA A DESCENDER POR LA PUERTA TRASERA EN LA PARADA A METROS DE AVDA. VICUA MACKENNA CON LA INTERSECCION DE AVDA. FRANCISCO BILBAO EN LA COMUNA DE PROVIDENCIA Y FUE ABORDADA POR SU ESPALDA POR 01 INDIVIDUO DESCONOCIDO D E ACENTO CHILENO P REMUNIDO CON UN ARMA DE BLANCA TIPO CUCHILLO EL CUAL LA APUNTO EN SU ESPALDA Y LE MENCIONO A VIVA VOZ INTIMIDANDOLA BAJATE CALLADI TA SI NO QUIERES QUE TE APUALE Y QUITATE EL BOLSO ANTE LO CUAL LA VICTI MA LE ENTREGO DE MANERA VOLUNTARIA SU BOL SO MARCA XCLUSIVE COLOR NEGRO QUE CONTENIA EN SU INTERIOR 01 NOTEBOOK MARCA LENOVO MODELO LO IGNORA COLOR NEG RO 01 TABLET MARCA APPLE COLOR PLATEADA 01 AUDIFONOS INALAMBRICOS MARCA APP LE COLOR BLANCO 01 LENTES DE SOL MARCA HOLBROOX COLOR AZUL 100. 000 CIEN MI L PESOS EN DINERO EN EFECTIVO MAQUILLAJE 01 TELEFONO MARCA APPLE MODELO IPH ONE 11 COLOR NEGRO ACTO SEGUIDO DICHO INDIVIDUO TRAS PERPETRAR EL ILICITO A NTES DESCRITO SE QUEDO ARRIBA DEL BUS. ESPECIES SUSTRIDAS01BOLSO MARCA XCLUSIVE C OLOR NEGRO QUE CONTENIA EN SU INTERIOR 01 NOTEBOOK MARCA LENOVO MODELO LO IGNORA COLOR NEGRO 01 TABLET MARCA APPLE COLOR PLATEADA 01 AUDIFO NOS INALAMBRICOS MARCA APPLE COLOR BLANCO 01 LENTES DE SOL MARCA HOLBROOX C OLOR AZUL 100.000 CIEN MIL PESOS EN DINERO EN EFECTIVO MAQUILLAJE 01 TELEF ONO MARCA APPLE MODELO IPHONE 11 COLOR NEGRO.AVALUO LA VICTIMA LO HARA ANT E ESA FISCALIA LOCAL. SEGUROSDESCONOCE SI MANTIENE SEGUROS PARA ESTE TIPO DE EVENTOS. PERFIL DEL DELICUENTE 01 INDIVIDUO DESCONOCI DO DE ACENTO CHILE NO NO RECUERDA MAS ANTECEDENTES CAMARAS DE VIGILANCIAS SE DESCONOCE SI EXI STEN REGISTROS DE CAMARAS DE VIGILANCIA EN EL LUGAR DE LOS HECHOS.MEDIOS DE PRUEBA ANTE ESA FISCALIA.TESTIGOS ANTE ESA FISCALIA. SOSPECHAS NO MANTI ENE EN DETERMINADAS PERSONAS.ESTADO DE LA VICTIMA EN NORMAL ESTADO DE TEMP ERANCIA.CITACION LA VICTIMA QUEDO ESPERA DE CITACION POR PARTE DE ESA FISC ALIA.</p>	<p>VÍCTIMA SE ENCONTRABA A BORDO DE RECORRIDO 210 CUANDO UN SUJETO LE PONE UN ARMA BLANCA EN LA ESPALDA Y LE DICE QUE LE ENTREGUE SU BOLSO CON ESPECIES Y BAJE CALLADA DE LA LOCOMIÓ9N, SINO, LA APUÑALA, POR LO QUE LA VÍCTIMA LE ENTREGA EL BOLSO Y BAJA DE LA MICRO Y SUJETO SE VA CON SUS PERTENENCIAS</p>	<p>VCTIMA TRANSITABA POR VA PBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE PORTANDO ARMA BLANCA TIPO CUCHILLO LA INTIMID Y LOGR SUSTRERLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VCTIMA SE ENCONTRABA EN VA PBLICA, CUANDO ES ABORDADA POR UN SUJETO QUE LA INTIMIDA CON ARMA BLANCA Y LE SUSTRAJO SUS PERTENENCIAS</p>	<p>(asalto en bus con arma blanca - recorrido 210) T5 Claridad T5 = 2. Fidelidad T5 = 1 (Detalla dinámica correcta pero erra en la ubicación) Utilidad T5 = 1 (sirve para clasificar el hecho como robo con violencia/intimidación, pero pierde matices importantes del contexto, que sea en un bus).</p> <p>LoRA-802 Claridad LoRA = 2. Fidelidad LoRA = 1 (mantiene la estructura del robo y añade el uso de arma blanca, alineado con la dinámica descrita, pero no menciona que fue en un bus). Utilidad LoRA = 1 (aprovechable para análisis al incluir la intimidación con arma blanca pero falta dinámica de robo en bus).</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_802	Evaluación
2	<p>DOY CUENTA A ESA FISCALIA QUE EL DIA DE HOY 29 DE ABRIL DEL AO 2024 A LA S 1335 HORAS SE PRESENTO ANTE LA CPR RENATE --- DE SERVICIO GUARDIA DON GIOVANNI --- 19 AOS SOLTERO MEDIOS COMPLETOS DE NACIONALIDAD CHILENA EMPLEADO CEDULA DE IDENTIDAD 21.--- FECHA DE NACIMIENTO 08.05.2004 DOMICILIADO EN AVENIDA LA ORQUESTA --- COMUNA DE LA PINTANA TELEFONO 9--- CORREO ELECTRONICO ---GMAIL.COM QUIEN EXPUSOQUE EL DIA DE HOY 29 DE ABRIL DEL AO 2024 A LAS 1320 HORAS EN CIRCUNSTANCIAS QUE SE ENCONTRABA AL EXTERIOR DEL VEHICULO PLACA PATENTE UNICA KP--- MARCA NISSAN MODELO NV350 ENTREGADO POR LA EMPRESA --- PARA REALIZAR SUS LABORES COMO PEONETA EL QUE SE ENCONTRABA ESTACIONADO EN LA INTERSECCION DE CALLE LOS CLARINES CON ANCUDE COMUNA DE LA FLORIDA CONTRA EL SENTIDO DEL TRANSITO Y AL ENCONTRARSE SUBIENDO A LA CABINA SE PERCATO QUE UN INDIVIDUO ALTO GORDO VESTIDO CON POLERON CON CAPUCHA Y PANTALON NEGRO SE ENCONTRABA SACANDO SACOS DE COMIDA DE PERRO DESDE LA CARGA DEL VEHICULO POR LO QUE SE ACERCO A EL PERCATANDOSE QUE EN EL COSTADO DEL FURGON SE ENCONTRABA UN VEHICULO PLACA PATENTE UNICA C--- MARCA CHEVROLET MODELO ASTRA COLOR NEGRO EN EL QUE SE ENCONTRABA UN INDIVIDUO COMO PILOTO DE ESTATURA BAJA CONTEXTURA DELGADA PELO CORTO NEGRO QUIEN MANTENIA UN ARMA DE FUEGO TIPO PISTOLA CON LA QUE LO APUNTO QUEDANDOSE QUIETO SIN HACER NADA MIENTRAS EL OTRO INDIVIDUO CONTINUABA SACANDO LOS SACOS. AL TERMINAR DE SUSTRAR LOS SACOS SE SUBE AL VEHICULO Y ESCAPAN EN DIRECCION A SANTA RA QUEL PERDIENDOLOS DE VISTA. PRIMERAS DILIGENCIAS. SE ACOGE DENUNCIA. SE REALIZA LA BUSQUEDA DEL VEHICULO CB---MARCA CHEVROLET MODELO ASTRA COLOR NEGRO DE PROPIEDAD DE ROBERTO -- CI 12--- DOMICILIO EN SANTIAGILLO --- DEPTO. --- COMUNA DE SANTIAGO EL QUE NO REGISTRA ENCARGO POR ROBO AL MOMENTO DE LA DENUNCIA.FISCAL CARMEN --- QUEZADA FOLIO 2024--- FISCALIA LOCAL LA FLORIDA.INSTRUCCIONES.-PRIMERAS DILIGENCIAS INVESTIGATIVAS Y TRABAJO EN EL SACA CARGO DE PERSONAL SIP DE LA UNIDAD CONCURRIR A DOMICILIO QUE REGISTRA EL VEHICULO INVOLUCRADO Y ENTREVISTARSE CON EL DUEÑO INFORMAR CUALQUIER NOVEDAD A ESTA FISCALIA INDICANDO NUMERO DE FOLIO.- EMPADRONAR TESTIGOS Y TOMARLES DECLARACION- INCAUTAR EVIDENCIA- VERIFICAR LA EXISTENCIA DE CAMARAS DE SEGURIDAD Y OBTENER SU REGISTRO- DECLARACION VICTIMA ESPECIFICANDO EN QUE CONSISTE LA INTIMIDACION CARACTERISTICAS FISICAS Y DE VESTIMENTAS DE LOS AUTORES- REGISTRAR NUMERO DE PARTE EN BITACORA WEBCARACTERISTICAS DEL INDIVIDUO. DOS INDIVIDUOS MASCULINOS EL PILOTO ESTATURA BAJA CONTEXTURA DELGADA QUIEN LO INTIMIDA INDIVIDUO ALTO CONTEXTURA GRUESA VESTIDO DE NEGRO.CARACTERISTICAS DEL LUGAR. VIA PUBLICA .MODUS OPERANDI. INDIVIDUOS DESCONOCIDOS INTERCEPTAN A VICTIMA AMENAZANDO PARA SUSTRAR ESPECIES.CANTIDAD DE VICTIMARIOS. 02 INDIVIDUOS SOSPECHAS. NO MANTIENE SOSPECHAS.ESPECIES. 06 SACOS DE COMIDA DISTINTAS MARCAS TIPOS Y TAMAOS. 03 SACOS MARCA PEDIGREE CACHORRO SENIOR Y RAZA PEQUEA 01 SACO DE CHAMPION CAT SABOR DEL MAR 01 MASTERCAT SALMON 01 MASTERDOG SABOR CAE DE TAMAOS DISTINTOS.AVALUO. 237.000ARMA O ELEMENTO UTILIZADO. UN ARMA DE FUEGO TIPO PISTOLA.LESIONES. NO MANTIENE.MOVILIZACION. INDIVIDUOS HUYEN EN VEHICULO --- MARCA CHEVROLET MODELO ASTRA COLOR NEGRO.TESTIGOS. ANTES DE ESA FISCALIA.CAMARAS. SE DESCONOCE.ELEMENTOS DE SEGURIDAD. NO HAY.SEGUROS. NO HAY.CITACION. EN ESPERA POR PARTE DE ESA FISCALIA.</p>	<p>VICTIMA SE ENCONTRABA REALIZANDO LABORES DE ENTREGA, MOMENTO EN QUE SE ENCONTRABA INGRESANDO A CABINA DE VEHICULO, SE PERCATO QUE UN SUJETO DESCONOCIDO SE ENCONTRABA SUSTRAYENDO SACOS DE COMIDA DE PERRO DESDE VEHICULO DE CARGA DE VICTIMA, PERCATANDOSE ADEMÁS QUE EN EL LUGAR SE ENCONTRABA UN VEHICULO EN EL QUE HABÍA UN SEGUNDO SUJETO QUE PORTABA ARMA DE FUEGO TIPO PISTOLA CON LA QUE INTIMIDÓ A VICTIMA, LOGRANDO SUSTRARLES LOS SACOS DE COMIDA DE PERRO Y DARSE A LA FUGA DEL LUGAR. VEHICULO VÍCTIMAS: VEHICULO PLACA PATENTE UNICA KP--- MARCA NISSAN MODELO NV350 VEHICULO IMPUTADOS: VEHICULO PLACA PATENTE UNICA --- MARCA CHEVROLET MODELO ASTRA COLOR NEGRO</p>	<p>VICTIMA SE ENCONTRABA EN SU VEHICULO EN VIA PUBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO, EL QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMID Y LOGR SUSTRARLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VICTIMA SE ENCONTRABA EN SU VEHICULO EN VIA PUBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE DESCENDI DESDE VEHICULO, EL QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMID Y LOGR SUSTRARLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>(víctima en vehículo realizando entregas, intimidado con arma de fuego. Sujetos movilizados en vehículo) T5 Claridad_T5 = 2. Fidelidad_T5 = 2 (describe bien la situación de víctima en vehículo y la sustracción con violencia/intimidación). Utilidad_T5 = 2 LoRA-802 Claridad_LoRA = 2. Fidelidad_LoRA = 2. Utilidad_LoRA = 2. (En este ejemplo ambos modelos entregan resúmenes muy similares y adecuados).</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_802	Evaluación
3	<p>DOY CUENTA A ESA FISCALIA HOY 15 DE JUNIO DE 2024 SIENDO LAS 2320 HORAS SE PRESENTA ANTE EL CARABINERO MAXIMILIANO --- ENCONTRANDOME DE 2DO PATRULLAJE GUARDIA EN LA 19&DEG COMISARIA DE PROVIDENCIA LA CIUDADANA IDENTIFICADA COMO ANIELYS -- CEDULA NRO. 27--- QUIEN EX PUSOQUE EL DIA DE HOY 15 DE JUNIO DEL PRESENTE Y SIENDO LAS 2230 HRS. EN CIRCUNSTANCIA QUE SE ENCONTRABA CAMINO A LA ESTACION DE METRO MONSEOR EY ZAGUIRRE ES CUANDO EN LA INTERSECCION DE AVENIDA SUCRE CON MANUEL MONTT FUE ABORDADA POR UN SUJETO VENEZOLANO DE APROXIMADAMENTE 25 AOS CONTEXTURA DELGADA NO RECUERDA MAYORES ANTECEDENTES EL CUAL CARGABA EN SU ESPALDA UN B OLSO DE LA EMPRESA RAPPY Y LO HACIA MONTADO EN UNA MOTOCICLETA MARCA YAMAHA MODELO XTS NEGRA. ESTE ANTISOCIAL UNA VES AL FRENTE PROCEDE A REVISAR LOS B OLSILLOS Y EXTRAER DESDE SU CHAQUETA UN TELEFONO CELULAR MARCA APPLE MODELO IPHONE 12 COLOR MORADO CARCASA TRANSPARENTE QUE EN SU INTERIOR TENIA SU CÉDULA Y LA TARJETA CTA. RUT DEL BANCO ESTADO DE NUMERO --- PERTENECIEN TE A LA COMPAIA ENTEL EL CUAL UNA VES LOGRADO SU COMETIDO HUYE HACIA IRA RRAZABAL EN DIRECCION.ES POR ESTE HECHO QUE SE DIRIGE HASTA LA 19&DEG COMI SARIA DE PROVIDENCIA PARA REALIZAR EL P ROCEDIMIENTO DE RIGOR.ESPECIES 01 T ELEFONO CELULAR MARCA APPLE MODELO IPHONE 12 COLOR MORADO CARCASA TRANSPARE NTE DE NUMERO --- PERTENECIENTE A LA COMPAIA ENTEL01 CEDULA DE IDENTIDAD01 TARJETA CTA. RUT DEL BANCO ESTADOAVALUO LAS VICTIMA LO HACE EN LA SUMA DE 600.000 SEISCIENTOS MIL PESOS SEGURO NO MANTIENE PARA ESTA CAU SAL.CAMARAS SI MANTIENE CAMARAS EN EL SECTOR. CITACION ESPERA DE CITACION POR PARTE DE ESA FISCALIA.</p>	<p>VICTIMA CAMINABA POR VÍA PÚBLICA, INSTANTES EN QUE FUE INTERCEPTADA POR UN SUJETO QUE SE MOVILIZABA EN MOTOCICLETA, EL QUE LE SUSTRAJÓ SU TELÉFONO CELULAR Y SE DIO A LA FUGA DEL LUGAR. VEHÍCULO IMPUTADO: MOTOCICLETA MARCA YAMAHA MODELO XTS NEGRA</p>	<p>VCTIMA CAMINABA POR VA PBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE SE MOVILIZABA EN MOTOCICLETA, EL QUE PORTANDO ARMA BLANCA TIPO CUCHILLO LA INTIMID Y LOGR SUSTRERLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VCTIMA CAMINABA POR VA PBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE SE MOVILIZABA EN MOTOCICLETA, EL QUE PORTANDO ARMA BLANCA TIPO CUCHILLO LA INTIMID Y LOGR SUSTRERLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>(víctima caminando, interceptada por motocicleta) T5 Claridad_T5 = 2 Fidelidad_T5 = 1 (captura el asalto en vía pública con motocicleta pero incluye arma no mencionada). Utilidad_T5 = 2 LoRA-802 Claridad_LoRA = 2. Fidelidad_LoRA = 1. (incluye arma no mencionada) Utilidad_LoRA = 2.</p>

9.3. Anexo 1C - LoRA 803 (Robo con violencia)

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_803	Evaluación
1	<p>RELACION DE LOS HECHOS SE PONE EN CONOCIMIENTO DE LA FISCALA LOCAL DE TEMUCO, QUE HOY A LAS 2230 HORAS, SE PRESENT EN ESTA UNIDAD, CARLOS --- E IDENTIFICADO EN EL RUBRO DENUNCIANTE VCTIMA, QUIEN EXPUSO LO SIGUIENTE QUE, HOY A LAS 18 30 HORAS APROXIMADAMENTE, EN CIRCUNSTANCIAS QUE INGRES AL SERVICENTRO COPEC, UBICADO EN CALLE RUDECINDO ORTEGA NRO. 1471, CON LA FINALIDAD DE SACAR DINERO EN EL CAJERO AUTOMTICO. INSTANTES QUE SE PERCAT QUE DOS INDIVIDUOS DESCONOCIDOS, UN HOMBRE Y UNA MUJER, AGREDAN A UN CLIENTE DEL SERVICENTRO Y COMO LA VCTIMA OBSERVABA LO SUCEDIDO, FUE ENCARADO POR LA MUJER QUE ACOMPAABA A INDIVIDUO, LA CUAL COMENZ A INSULTARLO CON PALABRAS GROSERAS Y DE FORMA SORPRESIVA EL HOMBRE LO AGREDI CON UN PALO EN SU CABEZA Y EN ESE INSTANTE LA VCTIMA SE LE CAY SU TELFONO, DONDE UNO DE LOS ANTISOCIALES LO TOM DEL SUELO, POR LO PARA EVITAR UN MAL MAYOR, L SE RETIR DEL LUGAR, TRASLADNDOSE HASTA EL CONSULTORIO MIRAFLORES Y POSTERIORMENTE SE TRASLAD A ESTA UNIDAD DENUNCIANDO LOS HECHOS. CARACTERISTICAS Y VESTIMENTAS LA VCTIMA SOLO MANIFEST QUE ERA UN HOMBRE Y UNA MUJER, Y QUE ESTOS GENERALMENTE SOLICITAN DINERO EN DICHO SERVICENTRO, NO RECUERDA VESTIMENTAS Y MAYORES CARACTERCTICAS. LESIONES E L MDICO DE TURNO LE DIAGNOSTIC HERIDA CUERO CABELLUDO Y CONTUSION MUSLO, DE CARCTER LEVE, FORMULARIO DE ATENCIN DE URGENCIA NRO. ---, ESPECIE SUSTRADA 01 TELFONO NRO. ---, MARCA MOTOROLA, MODELO G-20, COLOR GRIS, CON PLAN DE LA COM PAA MOVISTAR. AVALUO 150.000 CIENTO CINCUENTA MIL PESOS. SEGURO NO TIENE. DILIGENCIAS SE ENVI AL SUBOFICIAL CARLOS --- Y PERSONAL A SU CARGO, CON LA FINALIDAD DE DAR CON LA UBICACIN DE LOS ANTISOCIALES, LOS CUALES YA NO SE ENCONTR ABAN EN EL LUGAR. CAMARAS EN EL SERVICENTRO EXISTEN CMARAS DE SEGURIDAD, PERO HAY QUE SOLICITARLA MEDIANTE DOCUMENTACIN FORMAL DE LA FISCALA. TESTIGOS NO HAY, Y TAMPOCO FUERON EMPADRONADOS. CITACION EN EL EVENTO DE SER NECESARIO, LA VCTIMA SER CITADO A LA FISCALA LOCAL DE TEMUCO.--- SUBOFICIAL DE CARABINEROS SUBOFICIAL DE GUARDIA VO.....BO. FELIPE A--- TTE. CORONEL DE CARABINEROS COMISARIO</p>	<p>FUE AL SERVICENTRO COPEC A SACAR DINERO DEL CAJERO AUTOMÁTICO, Y EN EL LUGAR SE PERCATO QUE UNA PAREJA AGREDÍAN A OTRA PERSONA Y EN UN MOMENTO FUE AGREDIDO POR UNO DE ESTOS CON UN ELEMENTO CONTUNDENTE Y AL CAER SOLTÓ SU CELULAR EL CUAL FUE TOMADO POR LOS AGRESORES Y SE DIERON A LA FUGA.</p>	<p>SUJETOS INGRESAN A SERVICENTRO COPEC Y AGREDEN A CLIENTE DEL SERVICE NTRO, INTIMIDAN CON PALABRAS GROSERAS Y SUSTRAEAN TELEFONO</p>	<p>SUJETOS INGRESAN A SERVICENTRO COPEC Y AGREDEN A CLIENTE DEL SERVICE NTRO, INTIMIDAN A LA VICTIMA CON UN ARMA BLANCA PARA SUSTRAR DINERO Y ESPECIES</p>	<p>(Asalto a cliente en COPEC – golpes con palo) T5 Claridad_T5 = 2. Fidelidad_T5 = 2 (describe agresión a cliente en servicentro COPEC y sustracción de teléfono, consistente con el detalle). Utilidad_T5 = 2. LoRA-803 Claridad_LoRA = 2 Fidelidad_LoRA = 1 (similar, añadiendo arma blanca y dinero/especies). Utilidad_LoRA = 2.</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_803	Evaluación
2	<p>RELACION DE LOS HECHOS DOY CUENTA A ESA FISCALA LOCAL DE TEMUCO, QUE EL DA LUNES 02 DE SEPTIEMBRE DEL AO 2024, SIENDO ALREDEDOR DE LAS 1035 HORAS, SE PRESENTA AL CUERPO DE GUARDIA DE LA 8VA. COMISARIA TEMUCO, EL CIUDADANO FELIPE ---, CDULA DE IDENTIDAD 14---, QUIEN SE PRESENTA ANTE EL CABO 2DO. CHRISTIAN --- Y QUIEN INDICA QUE QUE, EL DA DOMINGO 01 DE SEPTIEMBRE DEL AO 2024, SIENDO ALREDEDOR DE LAS 0400 HORAS, EN MOMENTOS QUE SE ENCONTRABA TRABA JANDO FUERA DE SU RECORRIDO HABITUAL EN EL COLECTIVO DE LA LNEA 11, EL QUE ES DE PROPIEDAD DE TERESA ---, CDULA DE IDENTIDAD 5---, SE DESPLAZA HASTA CALLE CLARO SOLAR ESQUINA MANUEL BULNES, EN DONDE EN PRIMERA INSTANCIA C ONCURRE A RETIRAR DINERO EN EL CAJERO DEL BANCO ESTADO, QUE SE ENCUENTRA EN DICHA INTERSECCIN, POSTERIOR EN EL MISMO LUGAR ES CUANDO TOMA 02 PASAJEROS, UNO DE ELLOS SUBIENDO COMO COPILOTO Y EL OTRO EN LOS ASIENTOS TRASEROS, QUIENES LE SOLICITARON QU E LES REALIZARA UNA CARRERA HASTA EL SECTOR DE LAS QUILAS, ACCEDIENDO DE MANERA VOLUNTARIA A REALIZAR DICHA CARRERA. UNA VEZ AL LLEGAR HASTA EL SECTOR DE LAS QUILAS, LE INDICAN QUE LOS DEJE ENTRE CALLE LAS QUILAS ESQUINA LOS SAUCES, AL MOMENTO DE ING RESAR A CALLE LOS SAUCES, EL INDIVIDUO QUE SE ENCONTRABA EN EL ASIENTO TRASERO VOCIFERA LA PLATA, PENSANDO LA VCTIMA EN DICHO MOMENTO, QUE LE ESTABA SOLICITANDO EL DINERO DE LA CARRERA AL ACOMPAANTE QUE ABORDO EL COLECTIVO, EN DONDE SE DA VUELTA LA VICTIMA A MIRAR AL INDIVIDUO QUE SE ENCONTRABA EN LOS ASIENTOS TRASERO, QUIEN LE SEALA PASA LA PLATA, SEAL NDOLES QUE SI ERA UN BROMA, EN DONDE SIN PROVOCACIN ALGUNA, AQUEL QUE SE ENCONTRABA ATR S, LO GOLPEA EN EL HOMBRO DERECHO CON UN FIERRO QUE OCULTABA. POSTERIORMENTE A LO OCURRIDO, DESCIENDE DEL COLECTIVO FELIPE ---, CON LA FINALIDAD DE INCREPAR Y SACAR A LOS JVENES DEL INTERIOR DEL AUTOMVIL Y CUANDO VA A SACAR AL COPILOTO ESTE PREMUNIDO DE UN FIERRO, LO GOLPEA EN EL BRAZO IZQUIERDO, CONTINUAMENTE A ESTO SE BAJA EL OTRO ACOMPAANTE QUE SE ENCONTRABA EN EL VEHICULO, EL QUE APARNTEME SE ENCONTRABA PREMUNIDO DE UN ARMAMENTO DE FUEGO, TIPO ARTESANAL, QUIEN LE DISPARA, LOGRANDO LA VICTIMA EVADIR EL DISPARO, LLEGANDO ESTE DIRECTA MENTE A LA PUERTA DEL COSTADO DERECHO TRASERA, CONTINUAMENTE A LO OCURRIDO EL DENUNCIANTE SE SUBE INMEDIATAMENTE AL COLECTIVO, CON LA FINALIDAD DE RETIRARSE DEL LUGAR, EN DONDE UNO DE ELLOS CON UN FIERRO, GOLPEA EL PARABRISAS DELANTERO, CAUS NDOLE DA OS DE CONSIDERACIN. SOSPECHAS UNICAMENTE LAS MANTIENE EN LOS INDIVIDUOS QUE SE SUBIERON AL COLECTIVO, QUIENES VESTAN, EL PRIMER INDIVIDUO DE COPILOTO, MANTENDRA UN POLERON DE COLOR BURDEOS Y PANTALONES NEGROS Y EL SEGUNDO MANTENDRA COMPLETAMENT E VESTIMENTA OSCURA, UNO DE CONTEXTURA DELGADA Y EL OTRO DE CONTEXTURA M S ROBUSTA, INDICANDO QUE AMBOS NO SUPERARAN LA EDAD DE UNOS 20 AOS Y DE NACIONALIDAD CHILENA POR SU ASENT. TESTIGOS NO MANTIENE TESTIGOS DE LOS HECHOS. DAOS PERFORACIN DE PROYECTIL EN PUERTA COSTADO DERECHO TRASERA Y DAOS EN PARABRISAS DELANTERO. AVALUO DAOS LA VICTIMA LO CONSIDERA POR LA SUMA DE 300.000. TRESIENTOS MIL PESOS. ESTADO DE SALUD PENDIENTE INSTRUCCIONES FISCALIA SE PROCDE A INFORMAR LOS HECH OS OCURRIDOS AL FISCAL DE TURNO DON. MAURICIO ---, QUIEN CONFORME A FOLIO NRO. 2024---, INSTRUYE - MEDIDA DE PROTECCIN. CONTACTO TELEFNICO PREFERENTE, POR 10 DAS. - INCAUTAR EVIDENCIA SI LA HAY SIP - DECLARACIN VCTIMA . - ACTA DE PREEXISTENCIA Y AVALO. - VERIFICAR LA EXISTENCIA DE C MARAS DE SEGURIDAD Y OBTENER SU REGISTRO. SIP - FIJACIN FOTOGR FICA DE LOS DAOS DEL VEHICULO. SIP - CONCURRENCIA PERSONAL SIP - CONSTATAR LESIONES. - REGISTRO N MERO DE PARTE EN BIT CORA WEB. - REGISTRAR EL FOLIO DEL PROCEDIMIENTO EN EL RELATO DEL HECHO DEL PARTE DENUNCIA. CAMARAS EN EL LUGAR APARENTEMENTE SE ENCONTRARAN DIVERSAS C MARAS QUE PODRAN HABER REGISTRADO EL DESPLAZAMIENTO DEL COLECTIVO Y POSIBLEMENTE EL MOME NTO DE LOS HECHOS, CONFORME AL RELATO DE LA VCTIMA. REGISTRO FOTOGRAFICO SE LE CONSULTA A LA VCTIMA, SI SE LE PRESENTAN IM GENES DE POSIBLES AUTORES, SI PODRA IDENTIFICARLO, INDICANDO QUE SOLO RECONOCE LA VESTIMENTA, NO PODRA RECONOCER SUS ROSTR OS. CITACION LA VICTIMA QUEDARA EN ESPERA DE CITACIN, EN DONDE OTORGA N MERO TELEFNICO Y DOMICILIO PARTICULAR ANTE LA EVENTUALIDAD DE SER REQUERIDO POR ESA FISCALA LOCAL DE TEMUCO. SECTOR CUADRANTE 5, CORRESPONDIENTE A LA 2DA. COMISARIA TEMUCO. OFICIAL DE GUARDIA VO.BO. X X DE CARABINEROS COMISARIO</p>	<p>SUJETOS ABORDAN EL COLECTIVO Y SOLICITAN EL TRASLADO. AL LLEGAR A SU DESTINO UNO DE LOS SUJETOS GOLPEA A LA VICTIMA CON UN FIERRO Y EL OTRO LE DISPARA CON UN ARMAMENTO AL PARECER ARTESANAL</p>	<p>SUJETOS DESCONOCIDOS INGRESAN A BANCO ESTADO, INTIMIDAN A LA VICTIMA CON UN ARMA DE FUEGO, PARA LUEGO SUSTRAEER ESPECIES Y HUIR DEL LUGAR</p>	<p>SUJETOS DESCONOCIDOS INGRESAN A BANCO ESTADO, INTIMIDAN A LA VICTIMA Y SUSTRAEEN DINERO EN EFECTIVO</p>	<p>(asalto al conductor de colectivo - golpe con fierro y disparo) T5 Claridad_T5 = 2. Fidelidad_T5 = 0 (describe un asalto en BancoEstado, cuando el hecho real es sobre un conductor de colectivo, con golpes y disparo). Utilidad_T5 = 0 (lleva a interpretar erróneamente el contexto como robo a institución bancaria).</p> <p>LoRA-803 Claridad_LoRA = 2. Fidelidad_LoRA = 0 (mantiene el escenario BancoEstado, también inconsistente con el relato real). Utilidad_LoRA = 0.</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_803	Evaluación
3	<p>RELACION DE LOS HECHOS SE PONE EN CONOCIMIENTO DE LA FISCALA LOCAL DE TEMUCO, QUE HOY A LAS 0540 HORAS, SE PRESENT AL SGTO. IRO. DANIEL --- Y PERSONAL A SU CARGO, ROBERTO --- E IDENTIFICADO EN EL RUBRO DENUNCIANTE VCTIMA, QUIEN EXPUSO LO SIGUIENTE HOY, A LAS 0400 HORAS APROXIMADAMENTE, EN CIRCUNSTANCIAS QUE CAMINABA POR AVENIDA MANUEL RECARREN AL PONIENTE, AL LLEGAR A LA INTERSECCIN DE CALLE CAPRI, INGRES AL INTERIOR DE UNOS PASAJES DE LA POBLACIN EXIS TENTE EN EL LUGAR, BAJO LOS EFECTOS DEL ALCOHOL EN UN GRADO NO DETERMINADO, YA QUE MOMENTOS ANTES HABA INGERIDO BEBIDAS ALCOHOLICAS, INSTANTES QUE DE FORMA SORPRESIVA FUE INTERCEPTADO POR 3 INDIVIDUOS JVENES DESCONOCIDOS, TODOS PREMUNIDOS DE ARMAS BLANCAS CUCHILLAS, POR LO QUE DENTRO DE SU DESESPERACIN Y TEMOR POR SU VIDA, SUBI POR LA REJA DEL ANTEJARDN DE UN DOMICILIO, SUBINDOSE AL TECHO Y POR LOS RUIDOS REALIZADO A ESA HORA, SALIERON LOS VECINOS Y PROPIETARIO DEL DOMICILIO DONDE SE HABA SUBIDO, QUIENES PORTABAN PALOS Y OBJETOS CONTUNDENTES, Y AL QUERER ARRANCAR PARA NO SER AGREDIDO, PERDI EL EQUILIBRIO Y CAY AL SUELO, RESULTANDO CON LESIONES DE FACTURA EN SU TOBILLO IZQUIERDO, POR LO QUE COMENZ A DARSE A LA FUGA, SIENDO ALCANZAD O POR DOS INDIVIDUOS DESCONOCIDOS AL PARECER PROPIETARIOS DEL DOMICILIO Y VECINOS DONDE SE HABA RESGUARDADO PARA NO SER ASALTADO, QUIENES COMENZARON A AGREDIRLO CON GOLPES DE PIES, PUOS, OBJETOS CONTUNDENTES Y BASTONES RETRCTILES, EN DIVERSAS PART ES DEL CUERPO, POR LO L LE EXPLIC A ESTAS PERSONAS QUE SU INTENCIN NO ERA ENTRAR A ROBAR, SINO QUE SOLO RESGUARDARSE DEL ASALTO QUE IBA SER VCTIMA, POR LO QUE ESTOS NO LE CREYERON Y LO INGRESARON A UN VEHICULO DEL CUAL DESCONOCE PATENTE Y LO TRAS LADARON HASTA CALLE LAS PANTERAS NRO. 01247, CON LA FINALIDAD DE CORROBORAR SU IDENTIDAD, PERO EN EL CAMINO FUE AMENAZADO CON LANZARLO AL RO CAUTN Y AMARRARLO, DONDE EN EL INTERIOR DEL VEHICULO LO OBLIGARON A HACER ENTREGA DEL DINERO QUE MS ABAJO SE MENCIONAN, PARA LUEGO LOS INDIVIDUOS AL NO TENER RESPUESTAS EN SU DOMICILIO ACERCA DE SU IDENTIDAD, SE RETIRARON EN DIRECCIN DESCONOCIDA. LESIONES EN CONFORMIDAD AL ARTICULO 83 Y 130 DEL CDIGO PROCESAL PENAL, LA VCTIMA FUE TRASLADO AL CONSULTO RIO MIRAFLORES, DONDE EL MDICO DE TURNO LE DIAGNOSTIC FRACTURA DE TOBILLO IZQUIERDO, DE CARCTER GRAVE, PRODUCTO DE LA CADA DESDE LA TECHUMBRE DEL DOMICILIO Y RESPECTO A LA AGRESIN SUFRIDA, RESULT CON HERIDA SUPERFICIAL QUE IMPRESIONA POR AGAR RE EN REGION CERVICAL IZQUIERDA, MOVIL SIN DOLOR, DESTACA MULTIPLES LESIONES EN PARRILLA COSTAL IZQUIERDA, ASI COMO EQUIMOSIS EN REGION ESCAPULAR DERECHO EN SU TERCIO INFERIOR, EQUIMOSIS EN SU BRAZO IZQUIERDO, SUGERENTE CON GOLPE DE OBJETO ROMO SALID O, ASI COMO MULTIPLES EROSIONES SENSIBLE AL TACTO, EROSIONES DEN CARA DORSAL AMBAS MANOS, DE ESTAS LESIONES SON DE CARCTER LEVES, FICHA MDICA NRO. ---, SIENDO DERIVADO POSTERIORMENTE AL HOSPITAL REGIONAL DE TEMUCO, PARA EXMENES COMPLEMENTARIO S Y COLOCARLE YESO EN EL TOBILLO IZQUIERDO. ESPECIES SUSTRIDAS 200.000 DOSCIENTOS MIL PESOS, EN DINERO EFECTIVO. DILIGENCIAS SE EFECTUARON PATRULLAJES POR EL INTERIOR DE LA POBLACIN, UBICADA EN CALLE CAPRI CON MANUEL RECARREN, RECORRIENDO L OS PASAJES EN BSQUEDA DE CMARAS Y ALGUNA INFORMACIN REFERENTES A POSIBLES TESTIGOS Y VICTIMARIOS, NO SIENDO POSIBLE UBICAR EL INMUEBLE, YA QUE LA VCTIMA NO ENTREG UNA INFORMACIN MS CLARAS ACERCA DEL Nmero DE LA CASA, CALLE YO PASAJE, NI COLO R DEL INMUEBLE, ASIMISMO, NO FUE POSIBLE ENTREVISTAR ALGN TESTIGO EN LAS CERCANAS DEL PASAJE LAS PANTERAS NRO. ---, TODOS DURMIENDO Y CON LAS LUCES DE LOS DOMICILIOS APAGADAS. CAMARAS EN AMBOS SITIOS DEL SUCESO, NO HAY CMARAS. RESOLUCION DE LA FISCALIA EL SGTO. IRO. DANIEL ---, INGRES EL PROCEDIMIENTO A LA PLATAFORMA VIRTUAL DE LA FISCALA, A TRAVS DEL FOLIO NRO. ---, TOMANDO CONOCIMIENTO EL FISCAL DE TURNO MAURICIO ---, NO DANDO MAYORES INSTRUCCIONES YA QUE EL PERSONAL POLICIAL HABA REALIZADO LAS DILIGENCIA DEL ARTICULO 83 Y 130 DEL CDIGO PROCESAL PENAL, REALIZANDO TODO LO POSIBLE PARA LOGRAR OBTENER EL MXIMO DE ANTECEDENTES Y PONERLO EN CONOCIMIENTO DEL MINISTERIO PBLICO, TESTIGOS NO HAY. CITACION EN EL EVENTO DE SER NECESARIO, LA VCTIMA SER CITADO A LA FISCALA LOCAL DE TEMUCO. --- SGTO. IRO. DE CARABINEROS SUBOFICIAL DE GUARDIA VO.....BO. --- AYALA CAPITN DE CARABINEROS COMISARIO DE SERVICIO</p>	<p>SUJETOS ABORDAN A LA VICTIMA QUIEN SE TRASLADABA POR LA VIA PUBLICA, INTIMIDÁNDOLO CON ARMAS BLANCAS, LOGRANDO HUIR DEL ASALTO INGRESANDO A UN DOMICILIO, DONDE FUE SINDICADO COMO LADRÓN Y GOLPEADO.</p>	<p>3 SUJETOS INGRESAN A ANTEJARDN DE DOMICILIO, INTIMIDAN CON ARMAS BLANCAS, SUSTRAN ESPECIES Y SE DAN A LA FUGA</p>	<p>3 SUJETOS INGRESAN A LOCAL COMERCIAL, INTIMIDAN CON ARMAS BLANCAS, SUSTRAN ESPECIES Y SE DAN A LA FUGA</p>	<p>(víctima iba a ser asaltada, ingresa a un domicilio para escapar y es golpeada) T5 Claridad_T5 = 2. Fidelidad_T5 = 0 (resume como ingreso de tres sujetos a antejardín para robar especies, cambiando por completo la dinámica de confundir a la víctima con delincuente). Utilidad_T5 = 0. LoRA-803 Claridad_LoRA = 2. Fidelidad_LoRA = 0 (lo modela como ingreso a local comercial, igualmente alejado del hecho real). Utilidad_LoRA = 0.</p>

9.4. Anexo 1D - LoRA 804 (Robo por sorpresa)

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_804	Evaluación
1	<p>DOY CUENTA A ESA FISCALIA LOCAL PEALOLEN &NDASH MACUL QUE HOY SABADO 21 D E AGOSTO DEL AO 2021 A LAS 1615 HORAS SE PRESENTO EN ESTE CUERPO DE GUARDIA ANTE LA CARABINERO --- DE DOTACION DE LA 43&ORDF COMISARIA DE CARABINEROS PEALOLEN Y DE SERVICIO DE PRIMERA GUARDIA EL SR. --- LOPEZ 31 AOS CHILENO SOLTERO ESTUDIOS SUPERIORES EMPLEADO CEDULA DE IDENTIDAD NRO. 17.--- FECHA DE NACIMIENTO 20.11.19 89 DOMICILIADO EN AVDA. WALKER MARTINEZ NRO. --- COMUNA DE LA FLORIDA FONONUMERO 9.--- QUIEN EXPUSO LO SIGUIENTEQUE HOY SABADO 21 DE AGOSTO DEL AO 2021 A LAS 1435 HORAS APROXIMADAMENTE EN CIRCUNSTANCIAS QUE SE ENCONTRABA E STACIONADO EN EL PUNTO COPEC UBICADO EN LA INTERSECCION DE AVDA. ORIENTAL E SQUINA AVDA. AMERICO VESPUCCIO COMUNA DE PEALOLEN INSTANTES EN QUE COMENZO A EFECTUAR UNA LLAMADA DESDE SU CELULAR MARCA SAMSUNG MODELO S10 COLOR NEGRO CON CARCASA COLOR NEGRO ENCONTRANDOSE EN EL INTERIOR DE SU VEHICULO CON LA VENTANA DEL CONDUCTOR ABAJO MOMENTOS EN QUE UN SUJETO DEL CUAL NO MANTIENE MAYORES ANTECEDENTES LE SUSTRAJÓ DE SUS MANOS SU CELULAR PARA LUEGO DARSE A LA FUGA POR AVDA. EGAA EN DIRECCION AL PONIENTE POR LO ANTERIOR CONCURRÍO A ESTA UNIDAD A EFECTUAR LA DENUNCI A CORRESPONDIENTE.ESPECIE SUSTRAJIDA I UN CELULAR MARCA SAMSUNG MODELO S10 COLOR NEGRO CON CARCASA COLOR NEGRO .AVALUO LA VICTIMA LO HACE EN LA SUMA DE 590.000 QUINIENTOS NOVENTA MIL PESOS.CAMARAS LA VICTIMA LAS HARA PRESENTE ANTE ESA FISCALIA A.SEGUROS MANTIENE EN LA COMPAÑIA DE SEGURO ASSURANT CHILE SEGUN POLIZA NRO. ---. CITACION LA VICTIMA QUEDO EN ESPERA DE CITACION POR PARTE DE ESA FISCALIA.</p>	<p>VÍCTIMA HABLANDO POR CELULAR EN SU AUTOMÓVIL ESTACIONADO EN SERVICENTRO COPEC, CUANDO SUJETO LE SUSTRAE EL CELULAR POR LA VENTANA</p>	<p>VCTIMA CAMINABA POR VÍA PÚBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE DESCENDIÓ DESDE VEHICULO, EL QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMIDÓ Y LOGRÓ SUSTRARLE SU TELÉFONO CELULAR PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VCTIMA EN PUNTO COPEC CUANDO SUJETO LE SUSTRAE CELULAR DE LA MANO</p>	<p>(víctima en vehículo estacionado en COPEC le sustraen celular) T5 Claridad T5 = 2. Fidelidad T5 = 0 (lo presenta como víctima caminando por vía pública, perdiendo el contexto de vehículo estacionado en servicentro y añade intimidación). Utilidad T5 = 0.</p> <p>LoRA-804 Claridad LoRA = 2. Fidelidad LoRA = 2 ("víctima en punto COPEC cuando sujeto le sustrae celular de la mano"; refleja lugar y dinámica de arrebato) Utilidad LoRA = 2.</p>
2	<p>DOY CUENTA A ESA FISCALIA QUE EL DIA DE HOY 02 DE DICIEMBRE DEL AO EN CURSO SIENDO LAS 1420 HORAS SE PRESENTO ANTES ESTE CUERPO DE GUARDIA LA DENUNCIANTE VICTIMA --- CEDULA DE IDENTIDAD NRO. 9.--- QUIEN EXPUSOQUE EL DIA DE HOY 02 DE DICIEMBRE DEL AO 2020 SIENDO LAS 1315 HORAS APROXIMADAMENTE EN CIRCUNSTANCIAS QUE LA DENUNCIANTE VIAJABA EN EL BUS TRANSANTIAGO PRECISAMENTE EN LA LINEA 104 LA CUAL LO HACIA DE PIES CERCA DE LA PUERTA TRASERA DEL BUS MOMENTOS QUE AL LLEGAR A LA INTERSECCION DE AVENIDA EL SALVADOR CON AVENIDA SANTA ISABEL 01 INDIVIDUO EL CUAL VESTIA POLERA BLANCA PANTALON OSCUROS DE ESTATURA 170 APROXIMADAMENTE TEZ MORENA CONTEXTURA DELGADA AL ABRIR LA PUERTA TRASERA DEL BUS ESTE INDIVIDUO O SORPRESA ESIVAMENTE LE ARREBATA SU TELEFONO CELULAR MARCA SAMSUNG MODELO J6 COLOR NEGRO CON CARCASA CON TAPA DE COLOR BLANCO CON FLORES PERTENECIENTE A LA COMPAÑIA ENTEL CON EL NUMERO TELEFONICO --- DESDE SUS MANOS PARA POSTERIOR MENTE DESCENDIÓ DEL BUS Y HUIÓ POR AVENIDA EL SALVADOR HACIA EL NORTE PERDIÉNDOLO DE VISTA MAS ADELANTE MOTIVO POR EL CUAL CONCURRÍO A ESTA UNIDAD A REALIZAR LA DENUNCI A CORRESPONDIENTE.AVALUO LO HACE EN LA SUMA DE 80.000 OCHENTA MIL PESOS.SOSPECHAS EN EL INDIVIDUO SEALADO ANTERIORMENTE.TESTIGOS NO MANTIENE.SEGUROS NO MANTIENE SEGUROS COMPROMETIDOS.CAMARAS DESCONOCE SI EXISTEN CAMARAS EN EL LUGAR DE LOS HECHOS.CITACIONQUEDO EN ESPERA DE CITACION POR PARTE DE ESE MINISTERIO PUBLICO.</p>	<p>VÍCTIMA EN RECORRIDO 104, CUANDO SUJETO LE ARREBATA CELULAR DE LA MANO</p>	<p>VCTIMA TRANSITABA EN BUS DE TRANSANTIAGO, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE PORTANDO ARMA BLANCA TIPO CUCHILLO LA INTIMIDÓ Y LOGRÓ SUSTRARLE SU TELÉFONO CELULAR PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VCTIMA EN BUS DE TRANSANTIAGO CUANDO SUJETO LE ARREBATA CELULAR DE LA MANO</p>	<p>(víctima en bus recorrido 104) T5 Claridad T5 = 2 Fidelidad T5 = 1 (víctima en bus de Transantiago, arrebato de celular, pero incluye intimidación). Utilidad T5 = 1.</p> <p>LoRA-804 Claridad LoRA = 2. Fidelidad LoRA = 2. Utilidad LoRA = 2.</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_804	Evaluación
3	<p>DOY CUENTA A ESA FISCALIA QUE EL DIA VIEES 13DE NOVIEMBRE DEL PRESENTE A O SIENDO LAS 1730 HORAS APROXIMADAMENTE SE PRESENTO ANTE EL C.P.R. --- DE DOTACION DE ESTA UNIDAD Y DE SERVICIO SEGUNDO TUO E N LA GUARDIA --- 18 AOS CEDULA DE IDENTIDAD NRO. 9--- NACIONALIDAD CHILENA SOLTERA ESTUDIOS MEDIOS INCOMPLETO ESTUDIA NTE FECHA DE NACIMIENTO 12.07.2002 CON DOMICILIO EN AV. SALVADOR NRO. --- DEPTO. --- COMUNA DE UOA TELEFONO --- CO RREO ELECTRONICO ---GMAIL.COM QUIEN EXPUSO QUE EL DIA DE HOY SIENDO LAS 1340 HORAS APRO XIMADAMENTE LA DENUNCIANTE CAMINABA POR LA VIA PUBLICA ESPECIFICAMENTE POR AVDA. SALVADOR EN DIRECCION AL NORTE Y AL LLEGAR A LA INTERSECCION DE CA LLE LINCOYAN EN LA COMUNA DE UOA FUE INTERCEPTADA POR UN INDIVIDUO DESCONOCI DO EL CUAL SE TRASLADABA EN UN VEHICULO MARCA HYUNDAI DE COLOR GRIS NO RECU ERDA MAYORES ANTECEDENTES EL CUAL APARECIO POR SU ESPALDA SE APROXIMO Y LE ARREBATO DE MANERA S ORPRESIVA SU TELEFONO CELULAR MARCA APPLE MODELO IPHONE 7 PLUS COLOR ROSADO NUMERO --- DE LA COMPAIA MOVISTAR QUE MANIPULAB A CON SU MANO DERECHA. POSTERIORMENTE ESTE INDIVIDUO ABORDO EL AUTOMOVIL QU E SE TRASLADABA Y HUYO DEL LUGAR EN DIRECCI ON DESCONOCIDA. ESPECIES SUSTRAS IDAS 01 TELEFONO CELULAR MARCA APPLE MODELO IPHONE 7 PLUS COLOR ROSADO NUME RO --- DE LA COMPAIA MOVISTAR.AVALUO LA VICTIMA LO HACE EN LA SUMA DE 500.000 QUINIENTOS MIL PESOS SEGUROS NO MANTIENE.CAMARAS DE SEGURIDA D DESCONOCE SI EXISTEN CAMARAS DE VIGILANCIA EN EL LUGAR DE LOS HECHOS. PE RFIL DE LOS DELINCUENTES01 SUJETO 1.75 CONTEXTURA GRUESA PELO LARGO VESTIA UN POLAR COLOR CALIPSO Y UN JEAN AZUL NO RECUERDA MAYORES CARACTERISTICAS SITIO DEL SUCE SO VIA PUBLICA ABIERTO.DOCUMENTOS DE PREEXISTENCIA Y DOMINIO DE ESPECIES LA VICTIMA NO PRESENTO DOCUMENTOS QUE ACREDITEN PROPIEDAD O D OMINIO DE LAS ESPECIES SUSTRIDAS.ESTADO DE LA VICTIMA EN NORMAL ESTADO DE TEMPERANCIA.CITACION EL DENUNCIANTE QUEDO EN ESPERA DE CITACION POR PARTE DE LA FISCALIA LOCAL DE UOA</p>	<p>VÍCTIMA CAMINANDO, CUANDO SUJETO DESDE AUTO HYUNDAI GRIS BAJA PARA ROBAR CELULAR Y HUIR DEL LUGAR</p>	<p>VCTIMA CAMINABA POR VA PBLICA, INSTANTES EN QUE FUE ABORDADA POR UN SUJETO QUE DESCENDI DESDE VEHICULO, EL QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMID Y LOGR SUSTRARLE SU TELEFONO CELULAR PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VCTIMA CAMINANDO, CUANDO SUJETO LE ARREBATA CELULAR DE LA MANO.</p>	<p>(víctima caminando, sujeto baja desde auto para robar el celular)</p> <p>T5 Claridad_T5 = 2. Fidelidad_T5 = 1 (robo de celular en vía pública, por sujeto en vehículo, pero incluye intimidación). Utilidad_T5 = 1.</p> <p>LoRA-804 Claridad_LoRA = 2. Fidelidad_LoRA = 1 (“arrebata celular de la mano”; pero no mención del vehículo de los imputados). Utilidad_LoRA = 1.</p>

9.5. Anexo 1E - LoRA 867 (Robo de vehículo motorizado)

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_867	Evaluación
1	<p>DOY CUENTA A ESA FISCALIA QUE EL DIA DE HOY DOMINGO 29 DE DICIEMBRE DEL A O 2024 SIENDO LAS 2012 HORAS SE PRESENTO ANTE EL SGT. FELIPE --- Y PERSONAL A SU CARGO DESERVICIO DE 2DO TUO EN LA POBLACION MOVILIZADOS EN EL VEHICULO POLICIAL Z--- EL CIUDADANO IDENTIFICADO COMO SEB ASTIAN --- 37 AOS CHILENO SOLTERO ESTUDIOS UNIVERSITARIOS COMPLETO MEDICO VETERINARIO FECHA DE NACIMIENTO 30.01.1987 CEDULA DE IDENTIDAD NRO. 16. --- DOMICILIADO EN MANQUELPI NRO. --- COMUNA DE LA FLORIDA TELEFONO 9--- Y EXPUSO QUE EL DIA DE HOY 29.12.2024 SIENDO LAS 1940 HORAS APROXIMADAMENTE EN CIRCUNSTANCIAS QUE ENCONTRABA EN SU DOMICILIO UBICADO EN MANQUELPI NRO. --- COMUNA DE LA FLORIDA ESPECIFICAMENTE EN EL PATIO DELANTERO DE SU CASA MOMENTO EN EL CUAL SE DISPONE A SACAR EL VEHICULO DE PROPIEDAD DE SU PADRE AL EXTERIOR DE LA PROPIEDAD PARA PODER MOVER SU AUTO POR LO QUE ABRE EL PORTON METALICO SACANDO EL AUTOMOVIL A LA VIA PUBLICA MARCA STATION WAGON RENAULT CAPTUR HB EXPRESSION COLOR CREMA CON TECHO NEGRO AO 2016 PLACA PATENTE UNICA ---55 MOTOR NRO. --- CHASIS NRO. --- COMO CARACTERISTICA EN EL TECHO MANTENIA UNA MANCHA NOTORIA JUSTO EN MEDIO EN LA PARTE DE ARRIBA DEL PARABRISAS MIENTRAS EL ACOMODABA EL VEHICULO SE ACERCARON ALREDEDOR DE TRES A CUATRO SUJETOS TODOS ELLOS ENCAPUCHADOS A ROSTRO CUBIERTO UNO DE ELLOS SUBIO POR EL COSTADO DEL COPILOTO EL CUAL VESTIA UNA POLERA DEPORTIVA NO RECUERDA MAYORES CARACTERISTICAS RECONOCIENDO EL EQUIPO DE FUTBOL DE BOCA JUNIOR O DE EVERTON DE COLOR AZUL CON FRANJAS AMARILLAS QUIEN LO HACE PORTANDO UN ARMA DE FUEGO TIPO PISTOLA A ROSTRO TAPADO INTIMIDANDOLO CON EL ARMA SOBRE SU CABEZA AMENAZANDOLO DICIE ENDOLE BAJATE CONCHETUMADRE ENTREGA LAS LLAVES Y AL MISMO TIEMPO OTRO SUJETO SUBIO POR EL COSTADO DEL CONDUCTOR EL SUJETO USABA UN POLERON NEGRO TAPADO O COMPLETO CON VESTIMENTAS OSCURAS EL CUAL SE ENCONTRABA CON UN ARMA ELECTRICA TIPO ELECTROCHOQUE QUIEN LO INTIMIDA DICIENDOLE BAJATE CONCHETUMADRE MI ENTRAS ACTIVABA EL ELECTROSHOCK Y SE LO ACERCABA AL CUELLO SIN CAUSAR NINGUNA LESION POR LO CUAL PARA EVITAR UN MAL MAYOR LA VICTIMA DESCENDE DEL AUTOMOVIL DICIENDOLE AL SUJETO ENCAPUCHADO CALMATE CALMATE QUE ESTA MI HIJA EN EL PATIO HACIENDO PRESENTE QUE POR LA FORMA DE HABLAR DE LOS SUJETOS ERAN DE NACIONALIDAD CHILENA. POSTERIORMENTE ESTOS SUJETOS SE SUBEN AL VEHICULO Y HUYEN DEL LUGAR POR CALLE VOZ DE ARAUCO SIENDO SEGUIDOS POR EL CIUDADANO IDENTIFICADO COMO JUAN PABLO VECINO DE LA VICTIMA DESCONOCIENDO MAYORES ANTECEDENTES QUIEN LOS PIERDE DE VISTA A LA ALTURA DE WALKER MARTINEZ CON AVENIDA LA FLORIDA EN DONDE ESTOS SUJETOS PASAN CON LUZ ROJA EN DIRECCION AL PONTONTE. CAMARAS SI MANTIENE. TESTIGOS JUAN PABLO --- CEDULA DE IDENTIDAD NRO. 20. ---. AVALUO LO HACE EN LA SUMA DE 11.000.000 O CINCO MILLONES DE PESOS. ENCARGO SE REALIZO ENCARGO DE VEHICULO EL CUAL ARROJO EL CODIGO EUN ---. SEGURO NO MANTIENE. GPS NO MANTIENE. TAG SI MANTIENE. FISCAL DE TURO DE LOS HECHOS TOMO CONOCIMIENTO MEDIANTE BITACORA WEB EL FISCAL DE TURO --- QUIEN MEDIANTE EL FOLIO 2024 --- QUIEN INSTRUYO LO SIGUIENTE. 1. DERIVAR PROCEDIMIENTO AL SEBV PARA DILIGENCIAS INVESTIGATIVAS CAMARAS TESTIGOS PASADAS POR PORTICOS ETC. 2. RASTREAR ESPERANCIAS SUSTRIDAS POR GPS. 3. ENCARGO VEHICULO. 4. ACTA DE PREEXISTENCIA Y AVALUO. 5. VERIFICAR LA EXISTENCIA DE CAMARAS DE SEGURIDAD Y OBTENER SU REGISTRO. 6. INCAUTAR EVIDENCIA. 7. EMPADRONAR TESTIGOS Y TOMARLES DECLARACION. 8. DE CLARACION DE LA VICTIMA. 9. REGISTRAR NUMERO DE PARTE EN BITACORA WEB. 10. REGISTRAR EL FOLIO DEL PROCEDIMIENTO EN EL RELATO DEL HECHO DEL PARTE DENUNCIANTE. SEBV SE TOMO CONTACTO POR VIA TELEFONICA CON LA OFICINA DE GUARDIA DEL DEPARTAMENTO SEBV CABO RIVERO CAMILA --- QUIEN ENTREGA PROCEDIMIENTO A PATRULLA DE SERVICIO NOCTUO. CITACION EN ESPERA DE CITACION POR PARTE DE ESA FISCALIA.</p>	<p>VICTIMA SE ENCONTRABA EN SU VEHICULO, INSTANTES EN QUE FUE ABORDADA POR 4 SUJETOS, LOS QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA Y ELECTROSHOCK INTIMIDARON A VICTIMA Y LOGRARON SUSTRARLE SU VEHICULO PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VICTIMA SE ENCONTRABA EN SU VEHICULO EN VIA PUBLICA, INSTANTES EN QUE FUE ABORDADA POR 3 SUJETOS, LOS QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMIDARON Y LOGRARON SUSTRARLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VICTIMA SE ENCONTRABA EN SU VEHICULO EN VIA PUBLICA, INSTANTES EN QUE FUE INTERCEPTADA POR OTRO SUJETO, EL QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMID Y LOGRARON SUSTRARLE SUS PERTENENCIAS, PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>(robo de vehículo, tipo portonazo, uso de arma de fuego y electroshock) T5 Claridad_T5 = 2. Fidelidad_T5 = 1 (refiere a "sustraerle sus pertenencias" sin enfatizar el vehículo como objeto principal). Utilidad_T5 = 1. LoRA-867 Claridad_LoRA = 2. Fidelidad_LoRA = 0 (similar, sin destacar claramente el robo de vehículo pero erra también en cantidad de sujetos). Utilidad_LoRA = 1.</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_867	Evaluación
2	<p>RELACION DE LOS HECHOS DOY CUENTA A ESA FISCALA LOCAL DE VICTORIA QUE HOY A LAS 1210 HORAS, SE PRESENTO ANTE EL SERVICIO DE GUARDIA LA CIUDADANA MARIA ---, INDIVIDUALIZADA EN EL RUBRO DE DENUNCIANTEVCTIMA, QUIEN EXPUSO QUE HOY A LAS 1145 HORAS, MIENTRAS QUE ESTACIONO SU JEEP, MARCA TOYOTA, MODELO LAND CRUISER PRADO, AO 2011, DE COLOR GRIS, PLACA PATENTE NICA ---50, EN EL ESTACIONAMIENTO DEL SUPERMERCADO SANTA ISABEL, UBICADO EN AV. PRAT NRO. 1472, DE ESTA COMUNA. INST ANTES EN QUE AL DESCENDER DEL VEHUCULO, EN COMPAA DE SU MADRE EMA ---, CDULA DE IDENTIDAD NRO. 6--- Y SU HIJO PABLO ---, DE 8 AOS, MOMENTOS EN QUE SE DIRIGI HACIA EL MALERO CON LA FINALIDAD DE SACAR UNA BOLSA PARA REALIZAR COMPRAS EN ESTE SUPERMERCADO, SE LE ACERCO UNA PERSONA DE SEXO MASCULINO, QUIEN VESTA CON UNA CHAQUETA DE COLOR NEGRO, JEANS DE COLOR NEGRO, JOCKEY DE COLOR NEGRO, TS DE COLOR MORENA, ROBUSTO, QUIEN LE SEALO LA AYUDO, NE GNDOSE A ESTA PETICIN, SEALNDOME ENTREGA EL AUTO, INTIMIDNDOLA CON UN CUCHILLO, RAZN POR LA CUAL LE ENTREGO LAS LLAVES DEL MVIL, DNDOSE A LA FUGA POR CALLE MUOZ VARGAS EN DIRECCIN AL CENTRO. HACE PRESENTE QUE EN EL MALETERO DEL VEHUCULO S E ENCONTRABA UNA BICICLETA, MARCA FOCUS, MODELO O1E, DE COLOR NEGRA, ARO 29, DE FIBRA CARBN, CON UNA HORQUILLA CON EL NRO. 34. TESTIGOS SE LE TOMO DECLARACIN A AMANDA ---, INDIVIDUALIZADA EN EL RUBRO QUE SE MENCIONA. CAMARAS EN EL LUGAR DE LOS HECHOS SI EXISTEN CMARAS DE SEGURIDAD EN DONDE EL PERSONAL POLICIAL A CARGO DEL SGTO. 2DO. PATRICIO ACEVEDO ALVARES Y PERSONAL A SU CARGO, QUIENES CONCURRIERON HASTA EL SUPERMERCADO EN DONDE SE ENTREVISTAN CON EVELYN ---, CDULA DE IDENTIDAD NRO. 15---, FONO CONTACTO ---, QUIEN SEALO QUE LAS VIDEO GRABACIONES SOLO LAS MANIPULA LA CENTRAL DE SANTIAGO Y PARA OBTENER SUS REGISTROS DEBE SER POR UN MANDATO JUDICIAL. AVALUO LA VCTIMA LO REALIZA EN LA SUMA DE 21.000.000 VEINTE UN MILLN DE PESOS. SEGUROS LA VCTIMA NO MANTIENE SEGUROS COMPROMETIDOS. ENCARGO SE EFECTU EL ENCARGO EN ENCARGO UNICO NACIONAL, ARROJANDO EN LA PLATAFORMA ENCARGO NRO. ---. GPS EL AUTOMVIL NO MANTIENE GPS. DILIGENCIAS PERSONAL DE SERVICIO EN LA POBLACIN A CARGO DEL SGTO. 2DO. PATRICIO ---, ACOMPAADO DEL CABO IRO. JOS ---, MOVILIZADOS EN EL Z---, REALIZARON PATRULLAJES PREVENTIVOS POR EL LUGAR DE LOS HECHOS Y CALLES ALEDA AS CON LA FINALIDAD DE UBICAR AL AUTOR DEL ILCITO Y EL MVIL, REALIZANDO A SU VEZ UNA DIFUSIN RADIAL ENCARGANDO EL AUTOMVIL, NO OBTENIENDO RESULTADOS POSITIVOS, ACTO SEGUIDO SE ENTREVISTO CON EL ADMINISTRADOR DEL SUPERMERCADO DE NOMBRE MIGUEL ---, CDULA DE IDENTIDAD NRO. 13---, QUIEN MANIFIESTA QUE MANTIENEN CMARAS DE SEGURIDAD LAS QUE SOLO ENFOCAN EL INGRESO AL SUPERMERCADO. PERSONAL SIP SE LE INFORMO AL CABO IRO. CARLOS ---, QUIEN REALIZA DILIGENCIAS CONFORME A SU ESPECIALIDAD, REMITIENDO POSTERIORMENTE INFORME DE CONCURRENCIA DIRECTAMENTE A ESA FISCALA LOCAL DE VICTORIA. FISCAL DE TURNO DE ESTE HECHO SE LE DIO CUENTA AL FISCAL DE TURNO ---, MEDIANTE LA PLATAFORMA DE BITACORA WEB, EN DON DE CONFORME AL FOLIO NRO. ---, QUIEN INSTRUYO CONCURRENCIA DE PERSONAL SIP. VERIFICACIN DE CAMARAS DE SEGURIDAD. ENCARGO DEL VEHUCULO. CITACION LA VCTIMA QUEDO EN ESPERA DE CITACIN POR PARTE DE ESA FISCALA LOCAL DE VICTORIA GERSON A. MIRANDA MAYORGA CABO 2DO. DE CARABINEROS SUBOFICIAL DE GUARDIA VTO.....BNO. CRISTOPHER --- DE CARABINEROS COMISARIO DE SERVICIO</p>	<p>VICTIMA SE ENCONTRABA ABRIENDO SU VEHÍCULO, INSTANTES EN QUE UN SUJETO LA INTIMIDA CON UN ARMA BLANCA PARA SUSTRARLE EL VEHÍCULO</p>	<p>SUJETO ABORDA A VICTIMA EN SUPERMERCADO SANTA ISABEL, INTIMIDANDOLA CON UN ARMA BLANCA PARA SUSTRAR SU VEHICULO</p>	<p>VICTIMA ESTACIONA SU VEHICULO EN SUPERMERCADO SANTA ISABEL, CUANDO ES ABORDADO POR UN SUJETO QUE LO INTIMIDA CON UN ARMA BLANCA Y SUSTRAE SU BICICLETA</p>	<p>(víctima abriendo su vehículo en supermercado, amenaza con arma blanca, un sujeto, robo de vehículo) T5 Claridad_T5 = 2. Fidelidad_T5 = 2. (describe todo). Utilidad_T5 = 2. LoRA-867 Claridad_LoRA = 2 Fidelidad_LoRA = 0 (cambia el objeto robado a "bicicleta", lo que contradice el relato de robo de vehículo que es el núcleo del código 867). Utilidad_LoRA = 0.</p>

ID	Relato SAF (entrada)	Resumen original (detalle)	Resumen T5 general	Resumen LoRA_867	Evaluación
3	<p>DOY CUENTA A ESA FISCALIA LOCAL QUE HOY LUNES 17 DE MARZO DEL AO 2025 SI ENDO LAS 2145 HORAS SE PRESENTA ANTE EL SERVICIO DE SEGUNDA GUARDIA LA DEN UNCIANTEVICTIMA IDENTIFICADA COMO BARBARA --- CEDULA DE IDENTIDAD NRO. 15. --- QUIEN EXPUSO QUE HOY LUNES 17 DE MARZO DEL AO 2025 A LAS 2145 HORAS APROXIMADAMENTE EN CIRCUNSTANCIAS QUE SE ENCON TRABA EN SU DOMICILIO PARTICULAR UBICADO EN CALLE JUAN MOYA MORALES NRO. --- COMUNA DE UO4 MOMENTOS QUE SE BAJA DE SU VEHICULO PLACA PATENTE UN ICA ---26 MARCA CHERY MODELO TIGGO 7PRO GLX 1.5 AUT COLOR ROJO METALIZAD O AO 2023 TIPO STATION WAGON. INSTANTES EN QUE AL QUERER ABRIR EL PORTON SE ACERCAN 04 SUJETOS DE SHORT CON POLERON DONDE FORCEJEAN TIRONEA NDOLA A LA VICTIMA EL CUAL LA TOMAN DEL BRAZO POR LO QUE RECUERDA LA VESTIMENTA DE SOLO 01 INDIVIDUO EL CUAL VESTIA COMPLETAMENTE DE NEGRO CABELLO COLORIN LO CUAL LOS SUJETOS LOGRAN EXTRAER EL VEHICULO HUYENDO POR CALLE JUAN MOYA MOR ALES EN DIRECCIO N AL SUR PERDIENDOLOS DE VISTA DESCONOCIENDO MAS DETALLES. DEBIDO A LO OCURRIDO LA VICTIMA CONCURRE A LA UNIDAD POLICIAL A REALIZAR LA DENUNCIA CORRESPONDIENTE. AVALUO LA VICTIMA LO HACE POR LA SUMA TOTAL DE 16.000.000 DIECISEIS MILLONES DE PESOS . SEGUROS NO MANTIENE SEGUROS AS OCIADOS AL HECHO. TESTIGOSNO MANTIENE. CAMARAS SI MANTIENE DOMICILIO.EN CARGO ---.LESIONES NO MANTIENE.CITACION LA VICTIMA QUEDA EN ESPERA DE CITACION POR PARTE DE ESA FISCALIA LOCAL DE UOA.FISCAL --- FOLIO 2025----5INSTRUCCIONES 1. DILIGENCIAS INVESTIGATIVAS SE RAN DERIVADAS A BIRO.2. REVISION DE PASO DEL VEHICULO POR PORTICOS MUNICIPA LES O DE AUTOPISTAS. VERIFICAR SI EL VEHICULO PARTICIPA EN OTROS ILLICITOS D URANTE LA NOCHE.3. VERIFICAR LA EXISTENCIA DE C&65533MARAS DE SEGURIDAD Y OBTENER SU REGISTRO4. INCAUTAR EVIDENCIAS. ENCARGO VEH&65533&65533CULO 6. EMPADRONAR TESTIGOS Y TOMARLES DECLARACI&65533N7. ACTA DE PREEXISTECIA Y AVAL&65533O8. DECLARACI&6553 3N V&65533CTIMA ESPECIFICANDO EN QU&6 5533 CONSISTI&65533 LA VIOLENCIA CARACTERISTICAS F&65533SICAS Y DE VES TIMENTAS DE LOS AUTORES9. REGISTRAR EL FOLIO DEL PROCEDIMIENTO EN EL RELATO DEL HECHO DEL PARTE DENUNCIA.10. REGISTRAR N&65533MERO DE PARTE EN BIT& 65533CORA WEB</p>	<p>VÍCTIMA DESCIENDE DE SU VEHÍCULO PARA ABRIR EL PORTÓN DE SU DOMICILIO, MOMENTOS EN LOS CUALES APARECEN 4 SUJETOS QUE PROCEDEN A INTIMIDARLO PARA LUEGO SUSTRAER SU VEHÍCULO Y HURTI DEL LUGAR</p>	<p>VCTIMA SE ENCONTRABA EN SU VEHICULO EN VA PBLICA, INSTANTES EN QUE FUE ABORDADA POR 4 SUJETOS, LOS QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMIDARON Y LOGRARON SUSTRAERLE SU TELÉFONO PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>VCTIMA SE ENCONTRABA EN SU VEHICULO EN VA PBLICA, INSTANTES EN QUE FUE INTERCEPTADA POR 4 SUJETOS, LOS QUE PORTANDO ARMA DE FUEGO TIPO PISTOLA LA INTIMIDARON Y LOGRARON SUSTRAERLE SU TELÉFONO PARA LUEGO DARSE A LA FUGA DEL LUGAR</p>	<p>(víctima abriendo portón, 4 sujetos, intimidación y forcejeo y le roban el vehículo) T5 Claridad_T5 = 2. Fidelidad_T5 = 0 (lo resume como robo de teléfono celular, cambiando completamente el objeto del delito). Utilidad_T5 = 0. LoRA-867 Claridad_LoRA = 2. Fidelidad_LoRA = 0 (misma distorsión hacia robo de celular). Utilidad_LoRA = 0.</p>