



FINANCIAMIENTO DE BECAS DE INVESTIGACIÓN (SEGUNDA CONVOCATORIA)

Análisis de requerimientos para el desarrollo de un modelo de predicción de sentencias en la Sala Penal de la Corte Suprema de Justicia de Paraguay basado en aprendizaje supervisado
Universidad Nacional Autónoma de México
Helena Montserrat Gómez Adorno – helena.gomez@iimas.unam.mx

RESUMEN

Este proyecto de investigación se enfoca en el análisis de requerimientos para desarrollar un modelo de predicción de sentencias, de casos de la Sala Penal de la Corte Suprema de Justicia de Paraguay, utilizando aprendizaje supervisado. La Sala Penal de la Corte Suprema de Justicia de Paraguay (CSJ) maneja una gran cantidad de casos penales cada año, cada caso es único en términos de circunstancias, hechos y pruebas presentadas. El análisis de los requerimientos para el uso de aprendizaje supervisado, permite la identificación de patrones y tendencias en la jurisprudencia de la Sala Penal, lo que a su vez, proporcionaría una mejor comprensión de las decisiones judiciales y la posibilidad de entrenar un clasificador para predecir el resultado de un caso determinado. Primero recopilamos y preprocesamos 5000 casos de la jurisprudencia de la CSJ desde 2011 hasta 2023. A cada expediente se le realizó un proceso de limpieza y normalización de datos. De esta manera pudimos identificar y codificar las variables relevantes para realizar el análisis descriptivo. Las variables analizadas fueron al preopinante, el tribunal o juzgado de origen, la acción resuelta, el resultado de la acción, y la fecha de la resolución judicial agrupada a diferentes granularidades. Normalizamos y estandarizamos las variables a analizar para que sean comparables entre sí y puedan ser utilizadas en el análisis descriptivo.

OBJETIVOS

- Recopilar datos históricos de los últimos cinco mil expedientes de acuerdos y sentencias de la Sala Penal de la Corte Suprema de Justicia de Paraguay, además de obtener la metadata de dichos expedientes.
- Preparar y limpiar estos datos para entrenar un modelo de predicción y crear un corpus.
- Identificación de características representativas de la información textual de los casos penales consultados.
- Proporcionar recomendaciones para la implementación del modelo de predicción en la práctica judicial..

APORTES DE LA INVESTIGACIÓN

Creamos una base de datos de 5000 casos de la jurisprudencia de la Sala Penal de la CSJ. Cada caso cuenta con datos normalizados de el preopinante, el tribunal o juzgado de origen, la acción resuelta, el resultado de la acción, y la fecha de la resolución judicial agrupada a diferentes granularidades. Este recurso se puede utilizar para implementar un modelo de predicción de sentencias que ayudará a identificar características relevantes, una mayor consistencia y previsibilidad en las decisiones judiciales.

Identificamos posibles desafíos y limitaciones en la implementación del modelo de predicción en la práctica judicial para el cual tuvimos en cuenta 3 problemas principales que encontramos en la jurisprudencia: tendencia de aumento de casos en el segundo semestre, desbalance en las acciones resueltas y sesgo en el resultado de la acción. Finalmente, proponemos recomendaciones para la implementación del modelo de predicción: realizar un balance de recursos en el segundo semestre del, diversificación de tipos de acciones resueltas, revisión de procedimientos para declarar inadmisibilidad y validación de modelos.

ACTIVIDADES REALIZADAS

- 1- Se desarrolló un programa en Python que realiza lo siguiente:
 - Descarga de archivos de casos de la sala penal de la CSJ.
 - Extracción de textos de los archivos (pdf, doc, docx, html y txt) y procesamiento OCR de imágenes.
 - Corrección y eliminación de caracteres raros, así como eliminación de saltos de línea.
 - Normalización de variables para el análisis descriptivo de los casos.
- 2- Análisis descriptivo de la base de datos recopilada.

RESULTADOS OBTENIDOS

La Figura 1 muestra el número de casos (expedientes) agrupados por semestre del año. Podemos observar que hemos descargado casos desde el 2011 hasta diciembre de 2023.

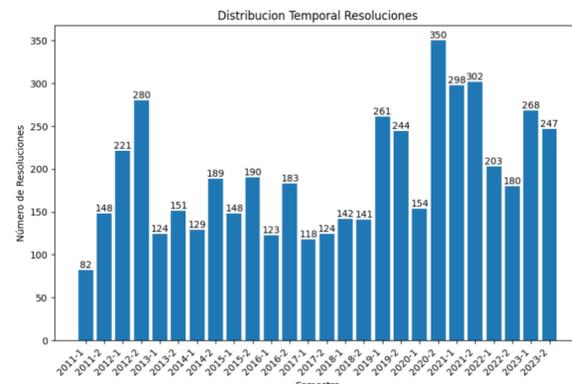


Figura 1. Histograma de casos por año.

La Figura 2 muestra el número de casos agrupados por "Acción resuelta". Vemos que el 77% de los casos corresponden a "Recurso de Casación", lo cual muestra un desbalance en los tipos de acciones resueltas en la sala penal.

La Figura 3 muestra el número de casos de acuerdo al resultado de la acción binaria que fue normalizado de acuerdo a la tabla presentada previamente. Podemos observar que el 83% de los casos recibieron una sentencia negativa a la solicitud de la demanda. Mientras que el el restante 17% de los casos obtuvo una sentencia positiva.

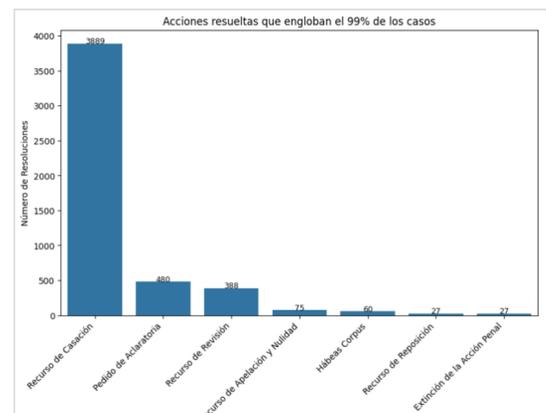


Figura 2. Histograma de casos por acción resuelta

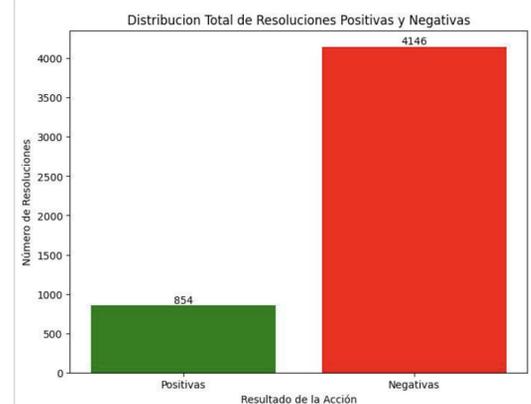


Figura 3. Distribución de casos con sentencia positiva vs. negativa

CONCLUSIÓN

La creación de esta base de datos así como su análisis, brinda un amplio panorama para aplicaciones futuras, así como para entender desde una visión general, el posicionamiento de los últimos años de la jurisprudencia en Paraguay.

VISIÓN Y PLANES FUTUROS

Extraer variables embebidas en el texto del caso penal y construir modelos de aprendizaje supervisado para predicción del resultado de la acción.

"Esta estancia de investigación fue cofinanciada por el Consejo Nacional de Ciencia y Tecnología (CONACYT) con recursos del FEEI"