**Resumen:**

Nuestro innovador modelo Contractor se basa en la sinergia de dos potentes modelos de Inteligencia Artificial: uno especializado en Procesamiento del Lenguaje Natural (NLP) y otro basado en técnicas de Ensamble. Estos modelos trabajan conjuntamente para analizar minuciosamente las descripciones y características de los contratos, permitiéndonos determinar el precio de referencia de cada contrato.

Las ventajas de nuestro modelo son múltiples y significativas: Nos permite identificar contratos con sobrecostos y sobrecostos, proporcionando una herramienta valiosa para la gestión eficiente de los recursos. Además, nuestro modelo puede sugerir precios de referencia a los nuevos contratistas, facilitando su integración en el proceso de licitación.

Pero quizás lo más importante es que nuestro modelo puede detectar señales de corrupción mediante la identificación de factores anómalos. Esto nos permite mantener la integridad y transparencia de todo el proceso de contratación, reafirmando nuestro compromiso con la ética empresarial y la responsabilidad social.

**Abstract:**

Our innovative Contracting model leverages the synergy of two powerful Artificial Intelligence models: one specialized in Natural Language Processing (NLP) and another based on Ensemble techniques. These models work together to meticulously analyze the descriptions and characteristics of contracts, allowing us to determine the reference price of each contract.

The benefits of our model are multiple and significant. It enables us to identify contracts with over-costs and under-costs, providing a valuable tool for efficient resource management. In addition, our model can suggest reference prices to new contractors, facilitating their integration into the bidding process.

Perhaps most importantly, our model can detect signals of corruption by identifying anomalous factors. This allows us to maintain the integrity and transparency of the entire contracting process, reaffirming our commitment to business ethics and social responsibility.

**Productos del modelo:**

El modelo provee información sobre el contrato, En primer lugar reporta un valor de referencia del precio, además da dos valores entre los que es mas probable que se encuentren este tipo de contratos. Por último si se le da un valor sugerido (o se prueba un contrato real) identifica que tanta verosimilitud tiene frente a las referencia del modelo.

**Utilidades de competencia:**

El modelo permite a nuevos contratistas tener información de precios asociados

**Utilidades para corrupción:**

Este modelo permite identificar sobrecostos y subcostos en algunos contratos públicos. La intuición que tenemos es que muchas de las estratégias que se usan para llegar a corrupción (contratos por debajo de puerta, licitaciones faltas, contratación directa) dan la libertad para que los contratos sean en valores más convenientes. De tal manera que queden recursos para repartir entre el licitador y el contratante.

Por otro lado también se puede identificar contratos a más bajo precio de lo normal. en estos casos de corrupción, aunque distintos. los contratos más bajos se pueden lograr para sacar a la competencia, recibir los primeros pagos e incumplir el contrato.

**Limitations:**

Hay una cantidad importante de limitaciones en este proceso. En primer lugar la información sobre la que se entrenó el modelo contiene casos de sobrecostos y subcostos. Aún así el modelo es capaz de reconocer estos casos (utilizando estratégias de regularización y con un error de entrenamiento específico, ver documentación).

Además la información en la que se constituyo es pasada, el mercado puede tener cambios importantes en los precios por información, tecnología, entre otros. Por lo tanto puede que el modelo de referencias de precios un poco atrasadas. Esperamos reducir este problema con reentrenamientos periódicos sobre nuevos contratos.

Tal vez lo más importante es que el modelo no es perfecto, van a haber casos en los que el modelo levanta una alerta, o sugiere un precio y es errado. Estas herramientas nunca tienen una capacidad absoluta ni perfecta de predicción y por lo tanto deben verse como herramientas para priorizar casos de corrupción y acercarse a la realidad. Sobre el porcentaje de error que tienen realizaremos experimentos con expertos para determinarlo.

**Características técnicas:**

modelo base: bert-base-multilingual-cased en descripción del contrato

Capa de finetuning: Entrada(768) Relu(15000) Sigmoid(15000) Sigmoid(15000) Linear (1500) dropout(p=0.3 en entrenamiento)

capa xgboost:entrada(1500 pesos de la capa de finetuning,"compiledRelease/tender/awardCriteria","compiledRelease/tender/status","compiledRelease/tender/mainProcurementCategory","compiledRelease/tender/numberOfTenderers","compiledRelease/tender/procurementIntention/category", "compiledRelease/tender/submissionMethod","compiledRelease/tender/awardCriteriaDetails",'compiledRelease/tender/mainProcurementCategoryDetails','compiledRelease/tender/hasEnquiries','compiledRelease/planning/budget/amount/currency'), error (pseudo huber error) reg\_alpha:50, reg\_lambda:50, "rate\_drop":0.1)

R2 en datos menores a 500 millones de dolares en validación: 0.547

R2 en validación: 0.054219614264203075