Propuesta de Proyecto

Seminario de Analítica y Ciencia de Datos

# Presentación del Proyecto

# Analítica y modelo de machine learning para reportes financieros, tanto balance como P&G de una compañía de seguros, con el fin de identificar transacciones inusuales en la consolidación y presentación de los estados de resultado.

# Objetivo del Proyecto

# Desarrollar un modelo de aprendizaje automático, que permita a los directivos de una compañía reconocer transacciones e información financiera inusual, en la consolidación y presentación de los estados financieros.

# Contexto del Problema

Actualmente, la gran cantidad de información que generan y procesan las empresas, en especial en sector servicios, ha ocasionado que grandes compañías se enfrenten al desafío de asegurar la integridad, exactitud, veracidad y confiabilidad en la información financiera. Existe un historial de fraude alrededor de los estados financieros a nivel mundial. El problema ha impactado tanto la economía de los países que estados unidos creo la ley SOX en 2002, cuyo propósito es que las compañías que quieran cotizar en la bolsa de Nueva york deben garantizar una seguridad razonable sobre la confiabilidad de la información financiera brindada a los accionistas y demás grupos de interés.

Debido a lo anterior, las compañías deben buscar cada vez más herramientas para poder identificar a tiempo inconsistencias en la información financiera, y dar garantía de la veracidad de sus datos. Las cuentas contables de las empresas deben ser monitoreadas con el fin de identificar a tiempo cifras inusuales y solicitar las causas y justificaciones correspondientes.

# Propuesta de Solución y Alcance

La propuesta es identificar a través de la analítica los patrones de comportamiento de las cuentas contables del balance general y perdidas y ganancias, establecer los criterios y las reglas de negocio que determinan si la cuenta es inusual o usual y entrenar un modelo que analice las cuentas contables mes a mes y genere las alertas correspondientes.

La analítica se realizará con Python y librerías como: numpy, pandas, empiracaldist, scipy, matplotlib, seaborn y pyjanitor. Para el entrenamiento de los modelos, scikit-learn.

# Definición del Alcance

El proyecto entregara los criterios y reglas de negocio que terminaran la confiabilidad en la cifra de las cuentas contables y el algoritmo de aprendizaje para la identificación de las transacciones anuales, para los estados financieros de una compañía de seguros, a nivel de PUCU.

# Riesgos e Impacto del Negocio

La solución impacta positivamente a la compañía ya que le permite automatizar a través de aprendizaje automático el análisis de las cuentas contables, además del cumplimiento del sistema de control interno de acuerdo con COSO III y requisito de vigilancia de las cuentas contables que permite mejorar la efectividad de las auditorias que buscan encontrar fraude o errores en los estados financieros.

# Rúbricas

R01: Completitud y consistencia: el estudiante realiza una entrega del momento evaluativo con todos los elementos desarrollados, manteniendo una relación coherente entre los elementos solicitados. El estudiante usa una consistencia argumentativa que incluye el uso de referencias cuando hay mérito de autoría entre otros.

R02: Pertinencia: el proyecto del estudiante está dentro del ámbito de la analítica y la ciencia de datos y es evidente en su descripción, objetivos y contexto

R03: Solución: el estudiante es capaz de relacionar herramientas tecnologico-metodológicas con el contexto de un problema de analítica y ciencia de datos.

R04: Herramientas: Las herramientas seleccionadas son pertinentes a la solución descrita. Las herramientas descritas evidencian una relación de flujo de proceso de los datos involucrados en la solución.

R05: Delimitación: dado un contexto de solución, el estudiante define claramente los límites de la solución que propone y el impacto, consecuencias y riesgos asociados a su implementación.