

Portafolio de Análisis

Jessica Nicole Copado Leal A01637876

December 6, 2022

1 SMA0104C Análisis de información

1.1 Trabaja con grandes volúmenes de datos

- Justificación: Se demuestra esta subcompetencia al trabajar con más de 40,000 imágenes que fueron utilizadas para la creación de modelos y 42,059 datos con 59 variables las cuales fueron analizadas para usar las de mayor correlación con las variables dependientes.
- Evidencia: [Análisis de los datos](#)

1.2 Procesa grandes volúmenes de datos de manera eficiente

- Justificación: Se demuestra esta subcompetencia al haber transformado las variables de manera continua para crear diferentes modelos de predicción y clasificación. De igual forma, se editaron las imágenes al recortar ciertas partes como el borde o enfocarse en la espuma del río.
- Evidencia: [Recorte de imagen](#)

1.3 Genera tableros útiles y correctos que apoyen a la toma de decisiones

- Justificación: Se demuestra esta subcompetencia al trabajar con tablas que contenían descripciones de ciertos datos para la toma de decisiones respecto a que variables deberíamos escoger dependiendo de su correlación con las demás variables.
- Evidencia: [Análisis de los datos](#)

```
Variables to remove [[0.002509042942147199, 'fNumber'], [0.00287906006790003, 'isoSpeed'],  
[0.0028828648469837043, 'height'], [0.0036953419003759973, 'vSigma'], [0.014486264183031072, '  
WwRawLineMin'], [0.027829165782506576, 'hSigma'], [0.029456574645584575, 'vSigma 0'], [0.032135583859325745,  
'graySigma'], [0.044859798318310856, 'entropyMean 0'], [0.04837940193890762, 'entropySigma 0'],  
[0.06740790739535044, 'shutterSpeed'], [0.0674254108852751, 'exposure'], [0.08530155850829349, 'graySigma  
0'], [0.10729929716586555, 'WwRawLineMean'], [0.10798271025459476, 'WwCurveLineMean']]
```

Figure 1: Correlación entre variables.

1.4 Mide correctamente el desempeño del modelo y sus métricas son correctas

- Justificación: Se demuestra esta subcompetencia al analizar los datos y en la creación de los modelos se calculan métricas para ver su precisión y error (MSE, RMSE, MAE)
- Evidencia: [Obtención de métricas](#)

MSE: 0.007437074183154802
RSMSE: 0.08623847275523149
MAE: 0.0671800927704174
Error estandar: 0.086245088327837

Figure 2: Errores obtenidos en el modelo.

1.5 Interpreta los resultados de las predicciones de los modelos y los interpreta en el contexto del problema de manera correcta

- Justificación: Se demuestra esta subcompetencia ya que al crear modelos y modificar los datos, se obtenían diferentes resultados, algunos mejores que otros. Las predicciones no fueron tan buenas pero usando diferentes métodos de trabajar con las imágenes se mejoraron los modelos.
- Evidencia: [Resultados obtenidos en cada modelo](#)

2 SEG0403B - Resuelve situaciones mediante el cumplimiento de leyes, normas y principios éticos.

2.1 Verifica que los datos estén anonimizados para no violar normas o leyes de privacidad de la información

No fue necesario anonimizar los datos ya que la organización socioformadora nos dio los datos de su propio estudio.

2.2 Especifica el proceso o estándar a seguir para validar el manejo de los datos y garantizar que solo el equipo tenga acceso a ellos en cumplimiento con las normativas.

No se necesitó llevar a cabo un estándar para validar el manejo de datos.

2.3 Presenta registros claros sobre el seguimiento del proceso y las autorizaciones para tener acceso a los datos

- Justificación: Los datos que se nos dieron están disponibles en línea para todo el público para poder hacer uso de ellos sin restricciones.
- Evidencia: [River Water Segmentation Segmentation Code](#)

3 Correcciones

A lo largo de los bloques, trabajamos con los profesores para poder entender el set de datos con el que trabajamos. También nos hicieron sugerencias para empezar a trabajar con estos mismo y cambiar ciertas cosas de nuestros modelos para optimizarlos (ARIMA, Librería Prophet).

En nuestro escrito también recibimos retroalimentación para ver si el orden era correcto y corregir la redacción.