



TRABAJO DE GRADO
Opción Seminario-Diplomado.

**ALGORITMO COMPUTACIONAL PARA EL ANÁLISIS Y TOMA DE DECISIONES
DE PUNTOS CALIENTES EN EL AÑO 2021 Y 2022 UTILIZANDO ESTRATEGIAS DE
MACHINE LEARNING**

Corporación Universitaria Remington.
Facultad de Ingeniería
Ingeniería de Sistemas

Estudiante:
Edwin Ferney Cruz Santos.
Tutor: Juan Carlos Briñez de León
Opción de Trabajo de grado Seminario-Diplomado.
2024.

Dedicatoria

A mi esposa, padre y madre, quienes me han apoyado en toda mi formación académica, Respaldandome en momento de dificultad, y por mi hijo quien es el motor de luchar a diario en mi vida personal y profesional para seguir adelante

Agradecimientos

Como principio universal un agradecimiento a Dios, mis compañeros, el docente de seminario por su paciencia y buenas prácticas de enseñanzas, a mis demás docentes que enseñan a diario con esmero y dedicación y a la universidad la cual al tener su facilidad de trabajar y estudiar lo cual me permiten superar mis capacidades y ser un estudiante dispuesto a recibir la mejor información de cada uno de mis mentores que me ayudaron durante mi proceso de formación universitaria.

Tabla de contenido

Tablas de Ilustraciones.....	5
Tablas de Tablas	6
Resumen.....	7
Palabras clave.....	7
Marco conceptual y contextual	8
Pregunta problema:	9
Acercamiento a los datos:	9
Descripción de variables.	9
Aproximaciones con gráficos.	12
Objetivos	14
Objetivo general.....	14
Objetivos específicos.	14
Desarrollo e implementación del aprendizaje.....	15
Procesamiento de los datos	15
Modelo de toma de decisiones	17
Resultados adicionales	21
Conclusiones	22
Referencias.....	23

Tablas de Ilustraciones

Ilustración 1. Distribución de toma de lectura	12
Ilustración 2. Gráfico de dispersión	14
Ilustración 5. Cargue de datos del Incendios Forestales	15
Ilustración 6. Eliminar datos innecesarios.	16
Ilustración 8. Mostrando datos a utilizar.....	16
Ilustración 11. Dividir datos de entrada y de salida.....	18
Ilustración 12. Entrenando y validando datos.....	19
Ilustración 13. Normalización de los datos.....	19
Ilustración 14. Evaluando los clasificadores.....	20
Ilustración 15. Práctica del modelo.....	21

Tablas de Tablas

Tabla 2. Acercamiento de los datos	9
Tabla 3.Descripción de variables	10
Tabla 4. Clave y descripción de los números del datasets	10

Resumen

Durante la ejecución del proyecto actual de Machine learning en tiempo de datos, se inicia con la identificación de variables para su análisis, recurriendo a los historiales de empresas u organizaciones dedicadas a recopilar la información necesaria en formato de conjuntos de datos. Estos conjuntos de datos son esenciales para facilitar el aprendizaje del modelo y la toma de decisiones adecuadas, así como para fomentar un aprendizaje continuo a partir de la información que se recopila diariamente. Para llevar a cabo este proceso, se han seleccionado las condiciones climáticas que podrían haber desencadenado posibles incendios forestales en Colombia durante el año 2022.

A partir de estos mismos datos, es posible identificar las zonas más afectadas durante el período mencionado. Esto, a su vez, permite desarrollar estrategias para abordar la situación en dichas zonas, teniendo en cuenta la magnitud del impacto. La extensión geográfica de los incendios y su intensidad son variables clave que se tienen en cuenta al diseñar acciones preventivas y de respuesta.

Es importante destacar que el análisis de datos no se limita únicamente a la identificación de áreas afectadas, sino que también implica la evaluación de factores subyacentes que contribuyen a la propagación y la severidad de los incendios forestales. Esto puede incluir factores como la densidad forestal, la presencia de vías de acceso, la disponibilidad de recursos para combatir incendios y las condiciones climáticas locales.

En resumen, el proyecto actual de Machine learning en tiempo de datos se centra en la identificación y análisis de variables clave relacionadas con los incendios forestales en Colombia en el año 2022, con el objetivo de desarrollar estrategias efectivas para la prevención y respuesta ante futuros eventos similares. Este enfoque integrado, que combina la recopilación y análisis de datos con técnicas avanzadas de modelado predictivo, es fundamental para mejorar la capacidad de anticipación y gestión de riesgos en el contexto de la gestión forestal y la protección del medio ambiente.

Palabras clave

Puntos calientes, atención de emergencias, Análisis de datos.

Marco conceptual y contextual

Durante los últimos años los incendios forestales han generado un aumento en Colombia generado desastres a causa de la interacción equivocada del ser humano y la naturaleza y la ve origina una fuerte preocupación teniendo en cuentas las afectaciones de fauna y flora que se presenta debido a este tipo de emergencias por ello es de gran inquietud observar cómo año a año se afecta el suelo colombiano y el entorno que lo rodea como se puede ver reflejado en la paradoja del fuego (Referencia 1).

Para los cuerpos de bomberos que son las entidades a nivel nacional en territorio colombiano responsables de atender este tipo de emergencias es de vital importancia tener esa estadística o aquellos indicadores que señalen cuales son los puntos más calientes o propensos a sufrir este tipo de emergencias como se encuentra reflejado en (Referencia 2).

Teniendo en cuenta las circunstancias anteriormente mencionadas se ha tratado de generar ponencias de investigación que logre crear un sistema de predicción de ocurrencias de incendios a nivel mundial tomando información de registros de atención de emergencias de cuerpos de bomberos voluntarios, servicios meteorológicos para Colombia el (IDEAM) o incluso imágenes satelitales suministradas por la Nasa a través de la plataforma firms moad como es el caso de (Referencia 3)

Pregunta problema:

¿Cómo desarrollar una estrategia computacional para ayudar a la toma de decisiones de campaña de sensibilización en los puntos donde existen mayor ocurrencia o probabilidad de generarse incendios forestales?

Acercamiento a los datos:

Descripción	Creadores	Enlace
Este conjunto de datos contiene los resultados de la toma de puntos calientes tomados por el satélite de la Nasa Modis	FIRMS Fire Information For Resource Management System	https://firms.modaps.eosdis.nasa.gov/map/#d:24hrs;@0.0,0.0,3.0z

Tabla 1. Acercamiento de los datos

Descripción de variables.

Total, de datos analizados por Collaboratory en Python: 30.246

#	Column	Non-Null Count	Dtype
0	latitude	30246 non-null	float64
1	longitude	30246 non-null	float64
2	brightness	30246 non-null	float64
3	scan	30246 non-null	float64
4	track	30246 non-null	float64
5	acq_date	30246 non-null	object
6	acq_time	30246 non-null	int64
7	satellite	30246 non-null	object
8	instrument	30246 non-null	object
9	confidence	30246 non-null	int64
10	version	30246 non-null	float64
11	bright_t31	30246 non-null	float64
12	frp	30246 non-null	float64
13	daynight	30246 non-null	object
14	type	30246 non-null	int64

Tabla 2. Descripción de variables

CLAVE	DESCRIPCIÓN
1	D
2	N

Tabla 3. Clave y descripción de los números del datasets

De acuerdo con la tabla de clave y descripción de las variables se hace necesario generar una breve explicación de algunas variables que serán tomadas para el desarrollo del proceso:

1. **brightness:** hace referencia al brillo y en donde esté a su vez nos puede dar la probabilidad de una incidencia de incendio forestal.
2. **Scan:** nos determina el tipo de escaneo que se realizo
3. **daynight:** Muestra si la lectura fue tomada durante Horas del Dia o Noche.

La recopilación de datos a través de estas plataformas nos ofrece una valiosa oportunidad para realizar análisis exhaustivos y precisos. Este proceso nos permite profundizar en la información disponible y obtener un conocimiento más completo de los datos en cuestión. Al aprovechar esta información detallada, podemos enriquecer nuestra comprensión y obtener perspectivas más claras sobre diversos aspectos.

Además, al analizar estos datos de manera más detallada y precisa, podemos identificar patrones, tendencias y correlaciones que de otra manera podrían pasar desapercibidos. Esto nos brinda la capacidad de anticipar y prever eventos futuros con

mayor precisión, lo que puede ser invaluable en una variedad de contextos, desde la planificación estratégica hasta la gestión de riesgos.

La riqueza de datos históricos almacenados en estas plataformas también juega un papel crucial en este proceso. Al aprovechar el vasto conjunto de datos históricos disponibles, podemos construir modelos predictivos más robustos y precisos. Estos modelos nos permiten generar escenarios hipotéticos y evaluar posibles resultados con mayor confianza, lo que a su vez nos ayuda a tomar decisiones informadas y estratégicas.

En resumen, la toma de información de datos a través de estas plataformas nos brinda una ventaja significativa al permitirnos realizar análisis más detallados y precisos, prever eventos futuros con mayor precisión y tomar decisiones más informadas y estratégicas. Esto no solo aumenta nuestra eficacia y eficiencia en una variedad de campos, sino que también nos ayuda a estar mejor preparados para enfrentar los desafíos y aprovechar las oportunidades que se presentan.

Posibles aplicaciones.

Dentro de las múltiples funcionalidades que nos ofrece el machine learning al generar un análisis calificado de la información, permitiendo así una toma de decisión en factores decisivos como lo puede la prevención en aquellos lugares donde la ocurrencia de un incendio forestales es bastante significativa.

O por otra parte la creación de sistemas de predicción que ayuden a identificar de manera oportuna aquellos donde por su material combustible, topografía o condiciones atmosféricas sean propensas para afectaciones de grandes extensiones de tierra afectando de esta manera la flora, fauna y múltiples ecosistemas que allí se desarrollen como podemos observar en la siguiente referencia (Referencia 4)

Aproximaciones con gráficos.

```
#Cargando librerías
import seaborn as sns
import matplotlib.pyplot as plt

# Crear un gráfico de barras con Seaborn
sns.countplot(data=Conjunto_Datos, x='daynight')
plt.title('Distribución Toma de Lectura')
plt.xlabel('Dia o Noche')
plt.ylabel('Frecuencia')
plt.show()
```

Diagrama de Torta de Distribución Toma de informacion

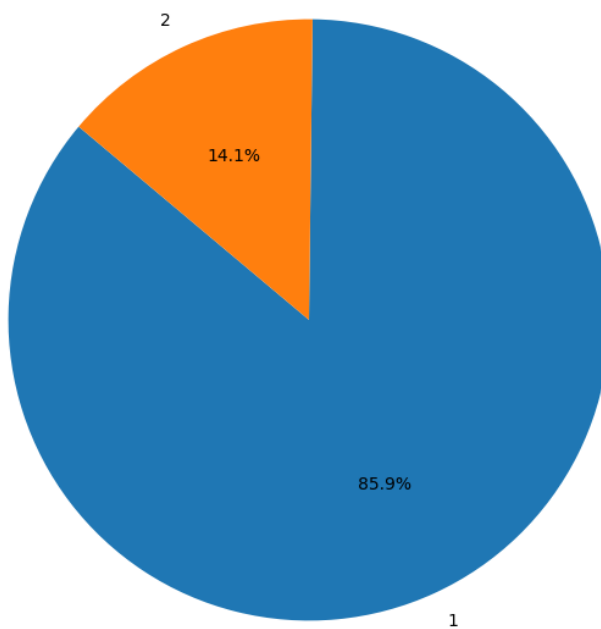


Ilustración 1. Distribución de toma de lectura

De acuerdo al diagrama de torta de distribución de lectura de Información requiere demostrar que la mayor cantidad de toma de información se realiza durante jornadas Diurnas con un 85.9 % teniendo en cuenta que durante esta jornada por las condiciones climatológicas existe mayores probabilidades que se desencadenen emergencias de tipo incendios forestales frente al 14.1% de las tomas realizadas durante la jornada nocturna el cual puede indicar brillos de emergencias forestales que se extendieron de la jornada Diurna hacia la jornada nocturna

```
#Cargando librerías
import seaborn as sns
import matplotlib.pyplot as plt

# Crear un gráfico de dispersión con Seaborn
sns.scatterplot(data = Conjunto_Datos, x='acq_date', y='brightness')
plt.title('Gráfico de Dispersión')
plt.xlabel('fecha')
plt.ylabel('Brillo')
plt.show()
```

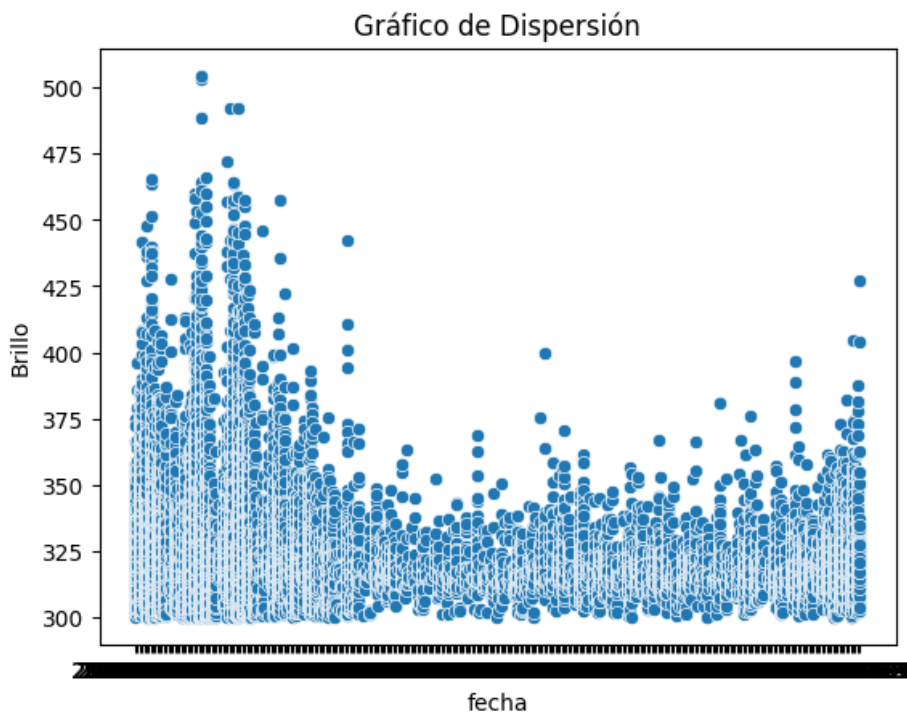


Ilustración 2. Gráfico de dispersión

En la anterior tabla se logra ilustrar como los mayores brillos que pueden desencadenar emergencias forestales se producen sobre los primeros meses del año donde el nivel de sequía teniendo en cuenta que la temporada de verano tiene mayor fuerza y hacia los meses de mediados baja su brillo generando menos incidencias de emergencias forestales en el territorio colombiano.

Objetivos:**Objetivo general.**

Implementar un algoritmo computacional para el análisis y toma de decisiones a partir de datos de posibles incendios forestales con el fin de generar predicciones, utilizando estrategias de machine learning.

Objetivos específicos.

- Desarrollar un algoritmo de detección temprana de posibles incendios forestales mediante el análisis de datos históricos y en tiempo real, utilizando técnicas de machine learning.
- Implementar un algoritmo de Machine learning que permita una correcta toma de decisión en cuanto a incendios forestales.
- Validar la practicidad de los algoritmos implementados para la toma de decisión.

Desarrollo e implementación del aprendizaje

Procesamiento de los datos

1.1. Carga de datos de incendios Forestales.

```
#Para cargar los datos
import pandas as pd
from google.colab import files
uploaded = files.upload()
for filename in uploaded.keys():
    Conjunto_Datos = pd.read_csv(filename, sep=',')
Conjunto_Datos.head()
```

Examinar... modis_2022_Colombia.csv
 modis_2022_Colombia.csv(application/vnd.ms-excel) - 2367432 bytes, last modified: n/a - 100% done
 Saving modis_2022_Colombia.csv to modis_2022_Colombia (1).csv

	latitude	longitude	brightness	scan	track	acq_date	acq_time	satellite	instrument	confidence	version	bright_t31	frp	daynight	type
0	4.5698	-69.2416	303.0	1.1	1.0	2022-01-01	252	Terra	MODIS	51	6.03	292.6	3.7	N	0
1	5.0267	-70.1586	307.8	1.2	1.1	2022-01-01	252	Terra	MODIS	72	6.03	292.9	7.7	N	0
2	4.7468	-69.6832	306.3	1.1	1.0	2022-01-01	252	Terra	MODIS	63	6.03	293.7	5.2	N	0
3	4.7374	-69.6819	306.0	1.1	1.0	2022-01-01	252	Terra	MODIS	53	6.03	293.7	4.9	N	0
4	4.5431	-70.6714	305.4	1.2	1.1	2022-01-01	252	Terra	MODIS	61	6.03	293.2	5.5	N	0

Ilustración 3. Carga de datos del Incendios Forestales

Como se observa en el Datasets (“modis_2022_Colombia.csv”) son significativos los datos que se toman para la identificación de los puntos calientes y que desencadenan incendios forestales por ello se resume y en la muestra y de igual manera se hace necesario eliminar algunos datos que no generan un aporte circunstancial para el desarrollo del presente proyecto:

1.2. Limpieza de datos.

```
#Quitando columnas que no aportan al proceso
Dataset = Dataset.drop(['acq_time', 'satellite', 'instrument', 'version', 'type'], axis=1)
Dataset.head(10)
```

	latitude	longitude	brightness	scan	track	acq_date	confidence	bright_t31	frp	daynight
0	4.5698	-69.2416	303.0	1.1	1.0	2022-01-01	51	292.6	3.7	N
1	5.0267	-70.1586	307.8	1.2	1.1	2022-01-01	72	292.9	7.7	N
2	4.7468	-69.6832	306.3	1.1	1.0	2022-01-01	63	293.7	5.2	N
3	4.7374	-69.6819	306.0	1.1	1.0	2022-01-01	53	293.7	4.9	N
4	4.5431	-70.6714	305.4	1.2	1.1	2022-01-01	61	293.2	5.5	N
5	4.7453	-69.6931	304.7	1.1	1.0	2022-01-01	53	293.6	4.1	N
6	4.7555	-70.0588	311.6	1.1	1.1	2022-01-01	82	293.6	10.3	N
7	4.5430	-70.6656	305.4	1.2	1.1	2022-01-01	61	293.2	5.5	N
8	4.9970	-70.2279	342.8	1.2	1.1	2022-01-01	100	293.9	58.2	N
9	4.9955	-70.2382	310.9	1.2	1.1	2022-01-01	81	292.6	10.5	N

Ilustración 4. Eliminar datos innecesarios.

No se muestran valores nulos en la base de datos, ni se demuestra falencias de datos.

```
# Corregir errores de formato (por ejemplo, convertir a mayúsculas)
Conjunto_Datos['daynight'] = Conjunto_Datos['daynight'].str.upper()
Conjunto_Datos.head(20)
```

	latitude	longitude	brightness	scan	track	acq_date	confidence	bright_t31	frp	daynight
0	4.5698	-69.2416	303.0	1.1	1.0	2022-01-01	51	292.6	3.7	N
1	5.0267	-70.1586	307.8	1.2	1.1	2022-01-01	72	292.9	7.7	N
2	4.7468	-69.6832	306.3	1.1	1.0	2022-01-01	63	293.7	5.2	N
3	4.7374	-69.6819	306.0	1.1	1.0	2022-01-01	53	293.7	4.9	N
4	4.5431	-70.6714	305.4	1.2	1.1	2022-01-01	61	293.2	5.5	N
5	4.7453	-69.6931	304.7	1.1	1.0	2022-01-01	53	293.6	4.1	N
6	4.7555	-70.0588	311.6	1.1	1.1	2022-01-01	82	293.6	10.3	N
7	4.5430	-70.6656	305.4	1.2	1.1	2022-01-01	61	293.2	5.5	N
8	4.9970	-70.2279	342.8	1.2	1.1	2022-01-01	100	293.9	58.2	N
9	4.9955	-70.2382	310.9	1.2	1.1	2022-01-01	81	292.6	10.5	N
10	5.0081	-70.2189	312.9	1.2	1.1	2022-01-01	85	293.1	11.9	N
11	5.0162	-70.2307	302.7	1.2	1.1	2022-01-01	23	292.3	3.8	N
12	5.0066	-70.2293	334.1	1.2	1.1	2022-01-01	100	293.1	40.5	N
13	5.0282	-70.1483	310.5	1.1	1.1	2022-01-01	80	293.0	9.8	N
14	5.0177	-70.2204	316.2	1.2	1.1	2022-01-01	93	292.8	15.8	N
15	5.6559	-68.5403	304.9	1.0	1.0	2022-01-01	55	294.8	3.7	N
16	5.0053	-70.2302	302.6	2.3	1.5	2022-01-01	50	289.9	15.3	N
17	5.2208	-72.6162	300.6	3.7	1.8	2022-01-01	28	286.4	32.2	N
18	5.2144	-72.6131	300.0	3.7	1.8	2022-01-01	2	286.2	29.4	N
19	5.0076	-70.2161	302.0	2.3	1.5	2022-01-01	22	290.1	13.0	N

Ilustración 5. Mostrando datos a utilizar.

Aquí podemos observar los datos con los cuales se procede a trabajar en el desarrollo del presente proyecto faltando convertir la casilla de Día o noche a un valor mas factible para la toma de decisión.

```
# Cambiando rango de valores del dataframe
Estrategia = {'D':1,'N':2}
Conjunto_Datos['daynight']=Conjunto_Datos['daynight'].map(Estrategia)
Conjunto_Datos.head(20)
```

	latitude	longitude	brightness	scan	track	acq_date	confidence	bright_t31	frp	daynight
0	4.5698	-69.2416	303.0	1.1	1.0	2022-01-01	51	292.6	3.7	2
1	5.0267	-70.1586	307.8	1.2	1.1	2022-01-01	72	292.9	7.7	2
2	4.7468	-69.6832	306.3	1.1	1.0	2022-01-01	63	293.7	5.2	2
3	4.7374	-69.6819	306.0	1.1	1.0	2022-01-01	53	293.7	4.9	2
4	4.5431	-70.6714	305.4	1.2	1.1	2022-01-01	61	293.2	5.5	2
5	4.7453	-69.6931	304.7	1.1	1.0	2022-01-01	53	293.6	4.1	2
6	4.7555	-70.0588	311.6	1.1	1.1	2022-01-01	82	293.6	10.3	2
7	4.5430	-70.6656	305.4	1.2	1.1	2022-01-01	61	293.2	5.5	2
8	4.9970	-70.2279	342.8	1.2	1.1	2022-01-01	100	293.9	58.2	2
9	4.9955	-70.2382	310.9	1.2	1.1	2022-01-01	81	292.6	10.5	2
10	5.0081	-70.2189	312.9	1.2	1.1	2022-01-01	85	293.1	11.9	2
11	5.0162	-70.2307	302.7	1.2	1.1	2022-01-01	23	292.3	3.8	2
12	5.0066	-70.2293	334.1	1.2	1.1	2022-01-01	100	293.1	40.5	2
13	5.0282	-70.1483	310.5	1.1	1.1	2022-01-01	80	293.0	9.8	2
14	5.0177	-70.2204	316.2	1.2	1.1	2022-01-01	93	292.8	15.8	2
15	5.6559	-68.5403	304.9	1.0	1.0	2022-01-01	55	294.8	3.7	2
16	5.0053	-70.2302	302.6	2.3	1.5	2022-01-01	50	289.9	15.3	2
17	5.2208	-72.6162	300.6	3.7	1.8	2022-01-01	28	286.4	32.2	2
18	5.2144	-72.6131	300.0	3.7	1.8	2022-01-01	2	286.2	29.4	2
19	5.0076	-70.2161	302.0	2.3	1.5	2022-01-01	22	290.1	13.0	2

Ilustración 9. Mostrando datos limpios.

Aquí se observa los datos limpios que permite un análisis de gráficos anteriormente mencionados, son columnas manejables porque son datos enteros y la interpretación es más rápida y sencilla.

Modelo de toma de decisiones

1.1. Mapeando los datos de entrada y de salida.

```
import pandas as pd
from google.colab import files
uploaded = files.upload()
for filename in uploaded.keys():
    Dataset = pd.read_csv(filename)
```

Dataset.head(7)

Examinar... modis_2022_Colombia.csv

modis_2022_Colombia.csv(application/vnd.ms-excel) - 2367432 bytes, last modified: n/a - 100% done
Saving modis_2022_Colombia.csv to modis_2022_Colombia.csv

	latitude	longitude	brightness	scan	track	acq_date	acq_time	satellite	instrument	confidence	version	bright_t31	frp	daynight	type
0	4.5698	-69.2416	303.0	1.1	1.0	2022-01-01	252	Terra	MODIS	51	6.03	292.6	3.7	N	0
1	5.0267	-70.1586	307.8	1.2	1.1	2022-01-01	252	Terra	MODIS	72	6.03	292.9	7.7	N	0
2	4.7468	-69.6832	306.3	1.1	1.0	2022-01-01	252	Terra	MODIS	63	6.03	293.7	5.2	N	0
3	4.7374	-69.6819	306.0	1.1	1.0	2022-01-01	252	Terra	MODIS	53	6.03	293.7	4.9	N	0
4	4.5431	-70.6714	305.4	1.2	1.1	2022-01-01	252	Terra	MODIS	61	6.03	293.2	5.5	N	0
5	4.7453	-69.6931	304.7	1.1	1.0	2022-01-01	252	Terra	MODIS	53	6.03	293.6	4.1	N	0
6	4.7555	-70.0588	311.6	1.1	1.1	2022-01-01	252	Terra	MODIS	82	6.03	293.6	10.3	N	0

```
#Ajustando los datos a una matriz numérica
import numpy as np
Dataset_Matriz = np.array(Dataset)
#Dividiendo los datos en entrada y salida
Entradas = Dataset_Matriz[:,0:3]
Salida = Dataset_Matriz[:, -1]
```

enando los modelos (En este caso entrenamos dos modelos)

```
#Modelo para predicciones de series de tiempo (KNN: Buen o si sabemos que los datos no se salen del rango)
from sklearn.neighbors import KNeighborsRegressor as KNNR
Modelo_1=KNNR()
Modelo_1.fit(Entradas,Salida)
```

```
+ KNeighborsRegressor
KNeighborsRegressor()
```

```
#Modelo para series de tiempo complicadas (Sólo se conoce la dinámica)
from sklearn.neural_network import MLPRegressor
Modelo_2 = MLPRegressor()
Modelo_2.fit(Entradas, Salida)
```

```
+ MLPRegressor
MLPRegressor()
```

Ilustración 6. Dividir datos de entrada y de salida.

- Se determina una matriz para los datos de entrada.

1.2. Dividir los datos de entrenamiento.

```
#Se ingresan los datos de entrada para generar una predicción
Nueva_entrada = np.zeros((1,3))
Nueva_entrada[0,0]=float(input('Ingrese el valor de la brightness actual: '))
Nueva_entrada[0,1]=float(input('Ingrese el valor de la brightness pasada: '))
Nueva_entrada[0,2]=float(input('Ingrese el valor de la brightness hace dos meses: '))

Proyeccion_1 = Modelo_1.predict(Nueva_entrada)
Proyeccion_2 = Modelo_2.predict(Nueva_entrada)

print('')
print('')
print('Según los datos ingresados, la proyección del brightness para el siguiente mes usando KNN será: ',Proyeccion_1[0])
print('')
print('Según los datos ingresados, la proyección del brightness el siguiente mes usando ANN será: ',Proyeccion_2[0])

Ingrese el valor de la brightness actual: 334.4
Ingrese el valor de la brightness pasada: 323.2
Ingrese el valor de la brightness hace dos meses: 312.2

Según los datos ingresados, la proyección del brightness para el siguiente mes usando KNN será: 0.0
Según los datos ingresados, la proyección del brightness el siguiente mes usando ANN será: 75.17868999655013
```

Ilustración 7. Entrenando y validando datos.

Se selecciona los datos del Brightness (brillo) con el fin de generar un modelo de predicción que me permita establecer posibles puntos calientes o de demasiado brillo y que estos a u vez desencadenen un incendio forestal.

Normalizar datos.

```
import pandas as pd
from google.colab import files
uploaded = files.upload()
for filename in uploaded.keys():
    Dataset = pd.read_excel(filename)

Dataset.head(7)
```

Elegir archivos brightness.xlsx

- **brightness.xlsx**(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 62597 bytes, last modified: 10/4/2024 - 100% done
Saving brightness.xlsx to brightness (2).xlsx

	brightness actual	brightness mes pasado	brightness hace dos meses	brightness mes siguiente
0	321.2	311.8	304.6	305.6
1	335.1	317.0	320.2	307.0
2	354.1	304.1	314.0	302.8
3	316.8	309.1	309.3	316.5
4	314.7	305.5	313.9	317.9
5	310.0	319.4	306.5	317.7
6	366.6	320.4	310.2	315.4

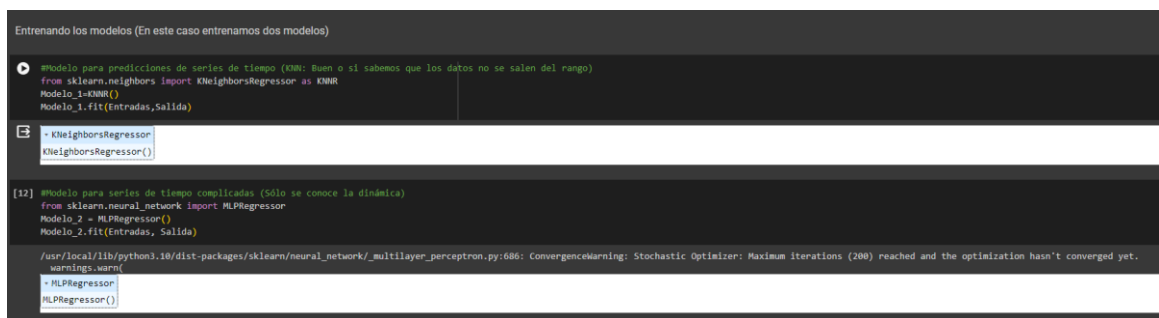
Ilustración 8. Normalización de los datos

Transformar los datos para las pruebas necesarias, donde se pueda interpretar los resultados adyacentes al dataset.

```
[10] #Ajustando los datos a una matriz numérica
import numpy as np
Dataset_Matriz = np.array(Dataset)
#Dividiendo los datos en entrada y salida
Entradas = Dataset_Matriz[:,0:3]
Salida = Dataset_Matriz[:, -1]
```

Ilustración 14. Ajuste de los datos Matriz Numérica

1.3. Entrenando los modelos de predicción.



Entrenando los modelos (En este caso entrenamos dos modelos)

```
#Modelo para predicciones de series de tiempo (KNN: Buen o si sabemos que los datos no se salen del rango)
from sklearn.neighbors import KNeighborsRegressor as KNNR
Modelo_1 = KNNR()
Modelo_1.fit(Entradas, Salida)
```

• KNeighborsRegressor
KNeighborsRegressor()

```
[12] #Modelo para series de tiempo complicadas (Sólo se conoce la dinámica)
from sklearn.neural_network import MLPRegressor
Modelo_2 = MLPRegressor()
Modelo_2.fit(Entradas, Salida)
```

/usr/local/lib/python3.10/dist-packages/sklearn/neural_network/multilayer_perceptron.py:686: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (200) reached and the optimization hasn't converged yet.
warnings.warn(
• MLPRegressor
MLPRegressor())

Ilustración 9. Evaluando los clasificadores.

En este espacio se procede a entrenar los dos modelos de predicción propuestos en clases con el fin de generar valores para el siguiente mes

1.4. Resultados con los mismos datos.

```

#Se ingresan los datos de entrada para generar una predicción
Nueva_entrada = np.zeros((1,3))
Nueva_entrada[0,0]=float(input('Ingrese el valor de la brightness actual: '))
Nueva_entrada[0,1]=float(input('Ingrese el valor de la brightness pasada: '))
Nueva_entrada[0,2]=float(input('Ingrese el valor de la brightness hace dos meses: '))

Proyeccion_1 = Modelo_1.predict(Nueva_entrada)
Proyeccion_2 = Modelo_2.predict(Nueva_entrada)

print('')
print('')
print('Según los datos ingresados, la proyección de la brightness para el siguiente mes usando KNN será: ',Proyeccion_1[0])
print('')
print('Según los datos ingresados, la proyección de la brightness para el siguiente mes usando ANN será: ',Proyeccion_2[0])

Ingrese el valor de la brightness actual: 334.4
Ingrese el valor de la brightness pasada: 316.6
Ingrese el valor de la brightness hace dos meses: 317.6

Según los datos ingresados, la proyección de la brightness para el siguiente mes usando KNN será: 314.48
Según los datos ingresados, la proyección de la brightness para el siguiente mes usando ANN será: 316.28603717669733

```

Ilustración 10. Práctica del modelo.

Teniendo en cuenta las pruebas realizadas podemos establecer un proceso de predicción que nos ayudara a determinar cuales posibles puntos puedan llar a causar un incendio forestal.

Resultados adicionales

Durante el desarrollo del presente proyecto se hace necesario una correcta normalización de la información con el fin de generar procesos de aprendizaje más óptimos, a su vez cuando se necesita la toma de decisión asertiva y ello facilita la compresión de los datos llegando optimizar de forma eficaz y

Conclusiones

- Gracias al seminario de machine learning en tiempo de datos, se observa un beneficio para el estudiantado que estuvo como asistente a estos temas, permitiendo la comprensión de los diferentes datos de muchos sitios Web, que dan a conocer la información, permitiendo que estos a su vez se puedan manipular y generar optimizaciones en procesos o en múltiples tareas que presentan cada una de las organizaciones.
- Los datos Suministrados para incendio forestales permiten la identificación de incendios forestales a través del brillo que se pueda ver reflejado a través de los satélites que generan las tomas de medición.
- El dataset que seleccione permite una correcta adecuación de los datos, aunque se debe anotar que para el ultimo proceso visto en el seminario no se permitió debido a mas datos que permitan mejorarse aún más.
- Mediante las herramientas de machine learning en tiempos de datos se logra Establecer resultados aplicables, los cuales permiten indicar, si puede desarrollar un incendio forestal o por el contrario no existe la posibilidad, como a su vez generar labores de prevención adecuadas para atender cada año estas emergencias.
- Se logra observar con el mismo sitio web de donde se obtuvo la información que tienen un proceso de formación que permite tomar esta información y optimizarlas mediante modelos que ellos han determinado o permite generar nuestros procesos de optimizacion

Referencias

la, p. contexto general de los incendios forestales en Colombia y perspectivas de los cuerpos de bomberos para la reducción de incendios forestales. *del fuego*, 73.

Colombia, C. D. (2012). Ley 1575 de 2012. Obtenido de viewsource: https://www.secretariasenado.gov.co/senado/basedoc/ley_1575_2012.html.

Inchausti, P. E., Martínez Saucedo, A., Amet, L. J., Blanco, P. M., Nieves, G. M., & Giusto, L. (2022). Aplicaciones de machine learning para el uso sustentable de recursos naturales. In XXIV Workshop de Investigadores en Ciencias de la Computación (WICC 2022, Mendoza).

Curbelo, A., Alric, J. C., & Inchausti, P. E. (2023). StopFire: Alertas de incendios forestales en Argentina usando IoT y machine learning. In XXVIII Congreso Argentino de Ciencias de la Computación (CACIC)(La Rioja, 3 al 6 de octubre de 2022).