



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES
CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO
EN ELECTRÓNICA Y TELECOMUNICACIONES**

TEMA:

***“Generación de datos sintéticos a partir de mediciones RSSI en una red SigFox utilizando
Data Augmentation.”***

Autores: Kevin Alexander Leiton Reina, Dennis Eduardo Tigse Pérez

Director del Proyecto: Ing. Román Alcides Lara Cueva, PhD.

Cotutor: Ing. Julio César Larco Bravo, MSc.



AGENDA

1. Introducción

2. Desarrollo

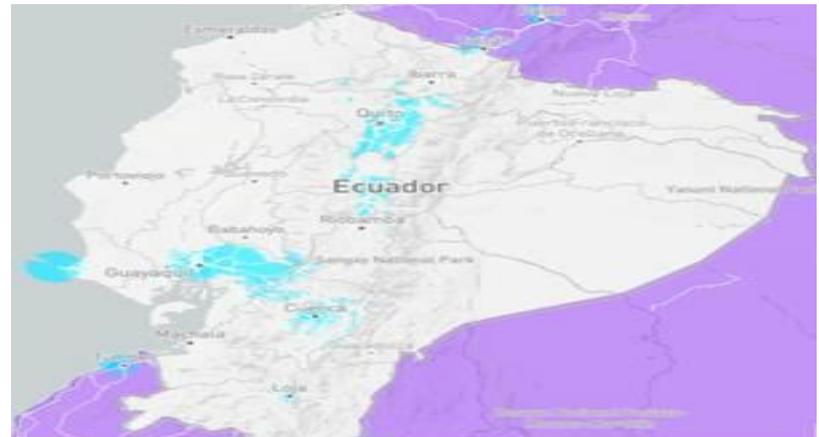
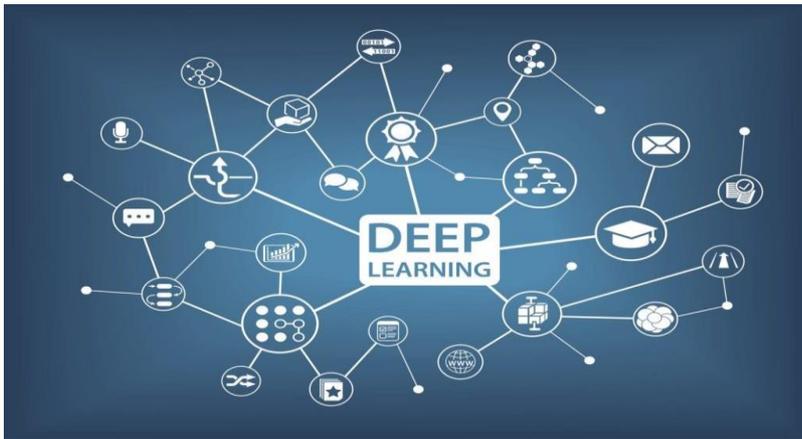
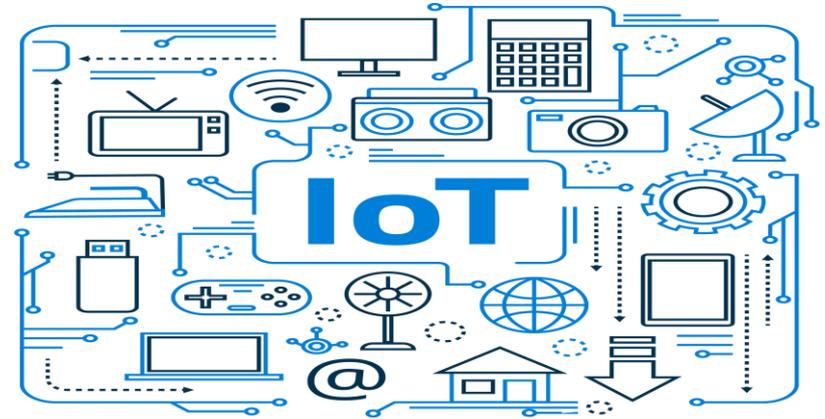
3. Pruebas y Resultados

4. Conclusiones y recomendaciones



MOTIVACIÓN

Justificación e importancia



OBJETIVOS

OBJETIVO GENERAL

- Generar de datos sintéticos a partir de mediciones RSSI en una red SigFox utilizando Data Augmentation

OBJETIVOS ESPECÍFICOS

- Analizar el estado del arte relacionado a las técnicas tradicionales del aprendizaje estadístico.
- Estimar un valor adecuado del RSSI para rellenar los datos nulos con modelos de regresión.
- Desarrollar un clasificador basado en aprendizaje supervisado del LQI.
- Aplicar técnicas asociadas a Data Augmentation para la generación de datos sintéticos empleando modelos tradicionales de Machine Learning.
- Identificar rangos del LQI para etiquetar los datos en función del RSSI.
- Evaluar el desempeño del clasificador con la base de datos aumentada.

METODOLOGÍA

MATERIALES:

Base de datos de la Red LPWAN Sigfox Cantón Quito

Software: MATLAB®, Google Earth Pro, Open Foris Collect Earth, MongoDB

Computador: Core I5, 4 GB RAM



Collect Earth - Grid
Generator



METODOLOGÍA

Diagrama de bloques de los modelos de regresión y clasificación :

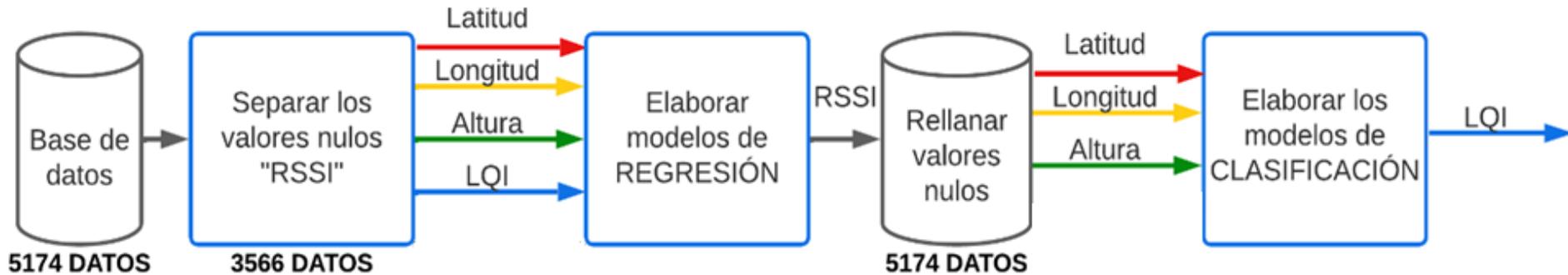
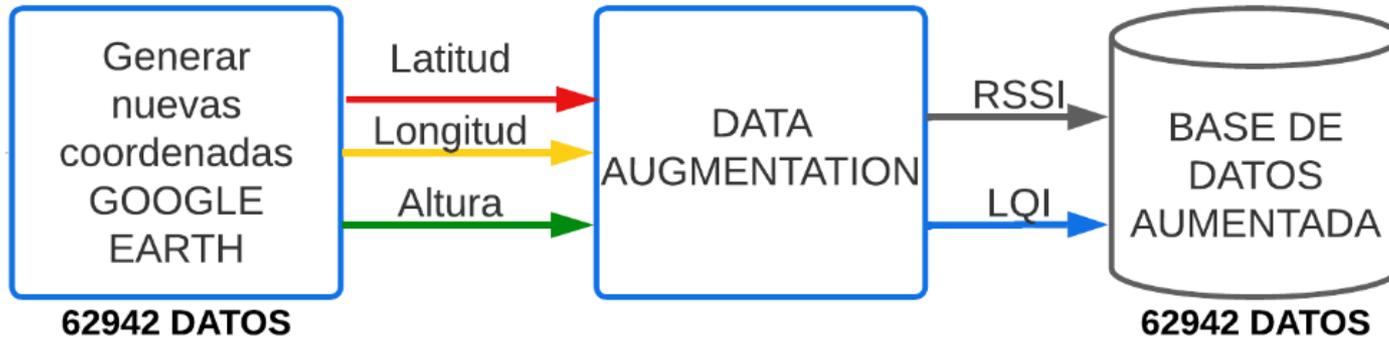


Diagrama de bloques Data Augmentation:



Base de datos original:

- La base de datos original describe las siguientes características:
- Número total de datos: 5174
- Datos con mediciones RSSI: 3566
- Datos sin mediciones RSSI: 1608
- Resolución: 200m entre parada y parada

Donde:

VARIABLES

VARIABLES	
Latitud	[grados]
Longitud	[grados]
Altura	[metros]
RSSI (Indicación de intensidad de la señal recibida)	[dBm]
LQI (Indicador de calidad del enlace)	Etiquetas

ETIQUETAS LQI

EXCELENTE
BUENO
PROMEDIO
LIMITE

METODOLOGÍA

FORMULAS PARA ENCONTRAR LAS METRICAS EN LA MATRIZ DE CONFUSION

ETAPA DE CLASIFICACIÓN

SENSIBILIDAD (%)

$$\frac{TP}{TP + FN}$$

EXACTITUD (%)

$$\frac{TP + TN}{TP + TN + FP + FN}$$

Donde:

TP o Verdadero positivo: valor real positivo y la prueba predijo también que era positivo.

TN o Verdadero negativo: valor real negativo y la prueba predijo también que el resultado era negativo.

FN o Falso negativo: valor real positivo, y la prueba predijo que el resultado es negativo.

FP o Falso positivo: valor real negativo, y la prueba predijo que el resultado es positivo.

ESPECIFICIDAD (%)

$$\frac{TN}{TN + FP}$$

PRECISIÓN (%)

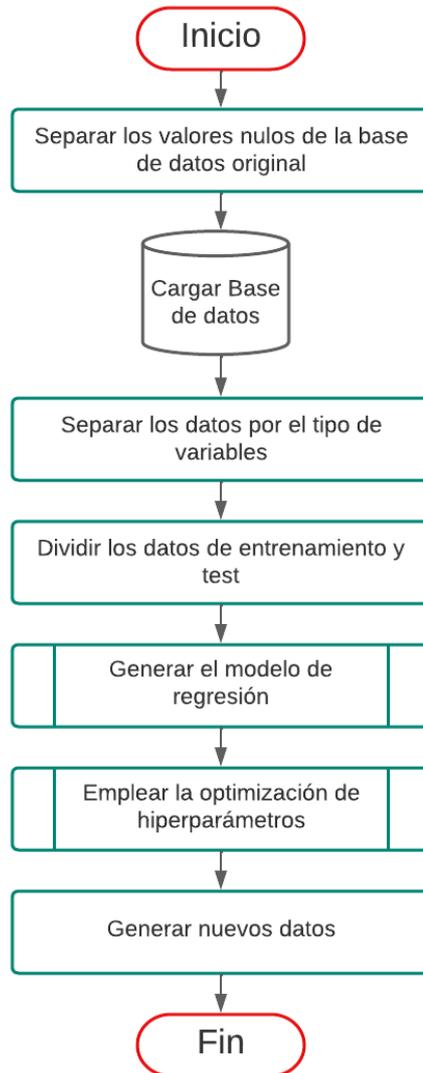
$$\frac{TP}{TP + FN}$$

BER

$$1 - \left(\frac{\text{Sensibilidad} + \text{Especificidad}}{200} \right)$$

DESARROLLO

Diagrama de flujo para los modelos de regresión:



Modelos de Regresión
Lineal
Polinomial
Support Vector Machine
Árbol de decisión

Error cuadrático medio

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Donde:

y_i : resultado real esperado

\hat{y}_i : predicción del modelo

N : número de muestras

DESARROLLO

Variables de entrada para los modelos de regresión:

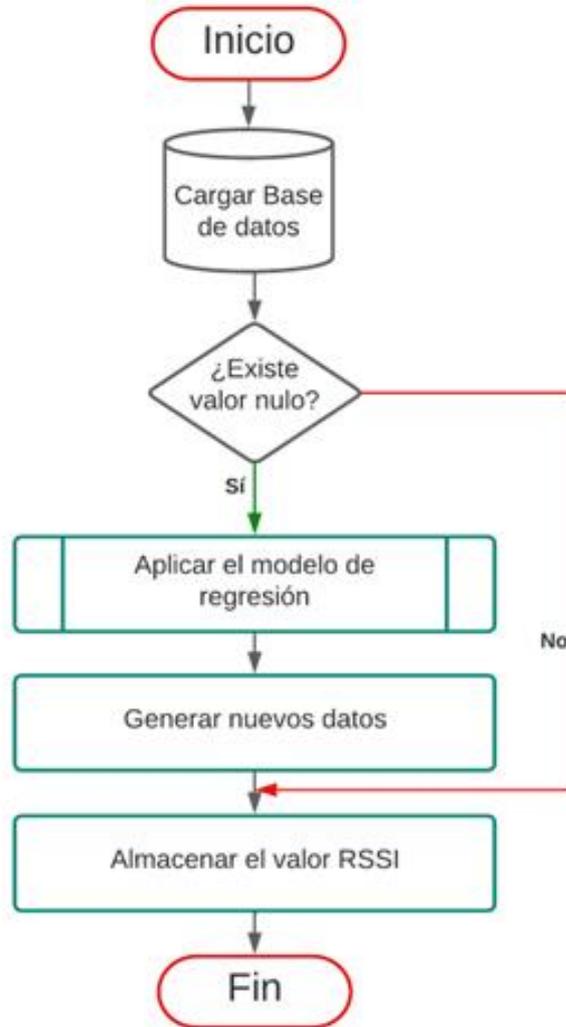
Variables de entrada	Número de variables
Latitud y longitud	2
Latitud, longitud y altura	3
Latitud, longitud, altura y LQI	4

Asignación numérica para cada etiqueta LQI:

Etiqueta LQI	Valor numérico
Excelente	4
Bueno	3
Promedio	2
Límite	1

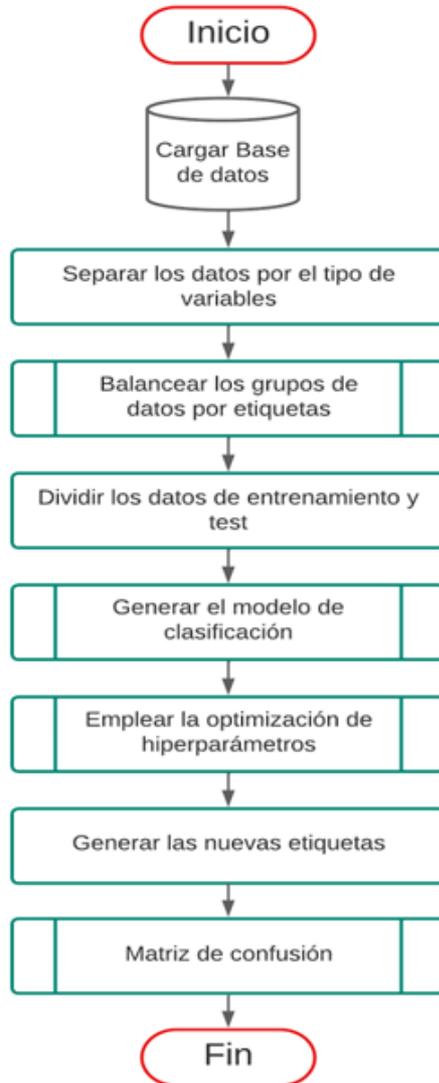
DESARROLLO

RELLENADO DE LOS VALORES NULOS



DESARROLLO

Diagrama de flujo para los modelos de clasificación:



Modelos de Clasificación
Support Vector Machine
Árbol de decisión

Variables de entrada para los modelos de clasificación:

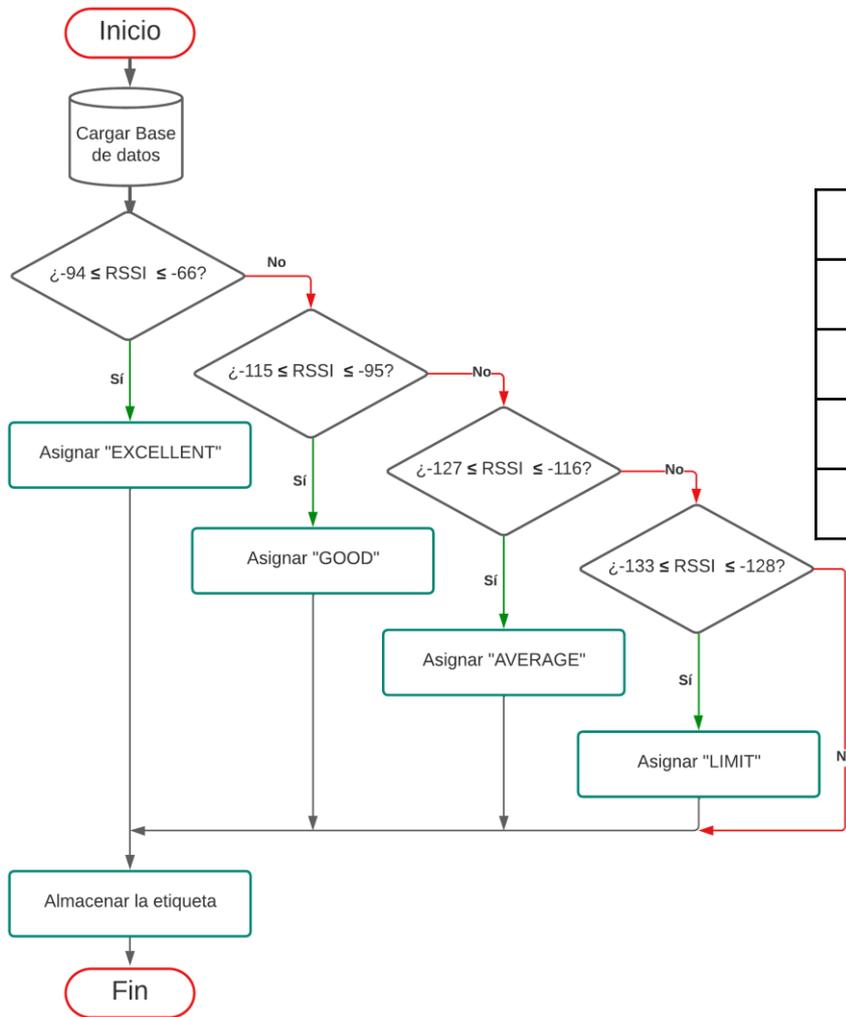
Variables de entrada	Número de variables
Latitud, longitud y altura	3

Cantidad de etiquetas:

Etiquetas	Datos originales
<i>LIMITE</i>	243
<i>EXCELENTE</i>	383
<i>PROMEDIO</i>	1218
<i>BUENO</i>	3330

DESARROLLO

Diagrama de flujo para el clasificador en función del RSSI:



Rangos LQI en función del RSSI

Etiquetas	Rangos [dBm]
<i>EXCELENTE</i>	[-94 → -66]
<i>BUENO</i>	[-115 → -95]
<i>PROMEDIO</i>	[-127 → -116]
<i>LIMITE</i>	[-133 → -128]

PRUEBAS Y RESULTADOS

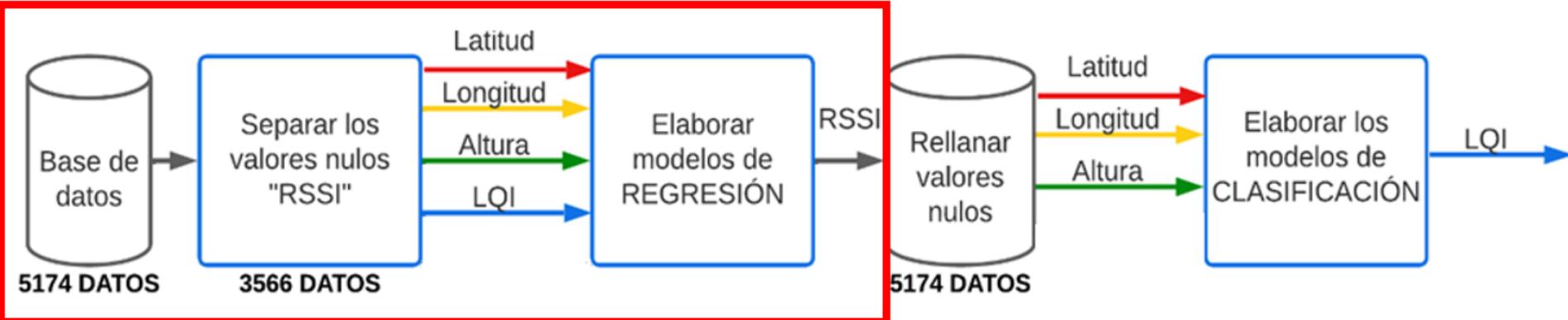
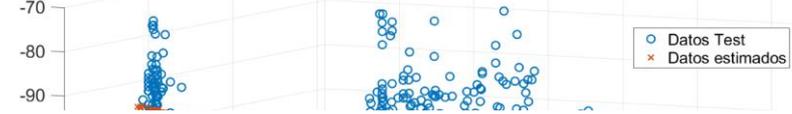
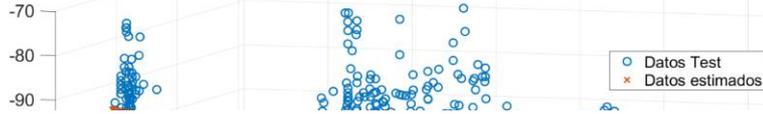
REGRESOR LINEAL

2 VARIABLES (LATITUD Y LONGITUD) 3 VARIABLES (LATITUD, LONGITUD Y ALTURA)

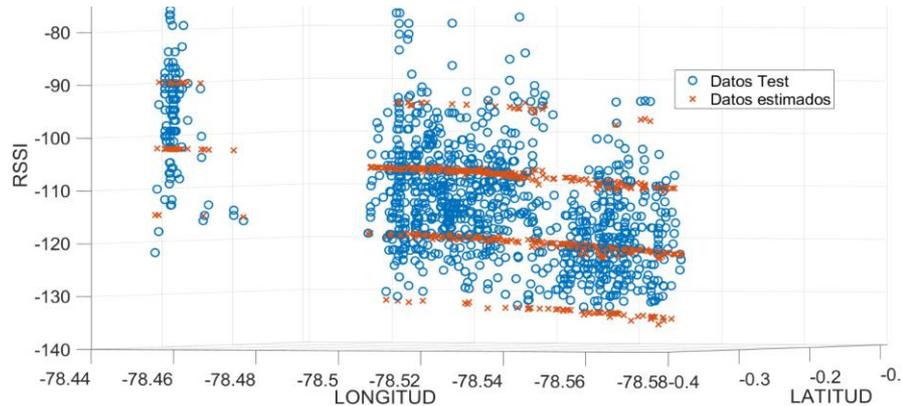
DATOS TEST VS DATOS ESTIMADOS

ETAPA DE REGRESIÓN

DATOS TEST VS DATOS ESTIMADOS



REGRESIÓN



PRUEBAS Y RESULTADOS

PARÁMETROS LIBRES

REGRESOR POLINOMIAL

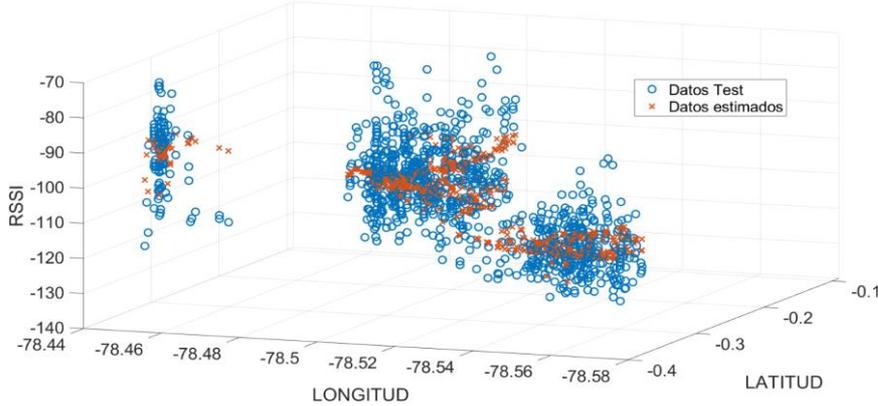
Variables de entrada	Grado del polinomio
Latitud y longitud	Latitud=7; Longitud=2
Latitud, longitud y altura	Latitud=7; Longitud=2; Altura=1
Latitud, longitud, altura y LQI	Latitud=7; Longitud=2; Altura=1; LQI=2

PRUEBAS Y RESULTADOS

REGRESOR POLINOMIAL

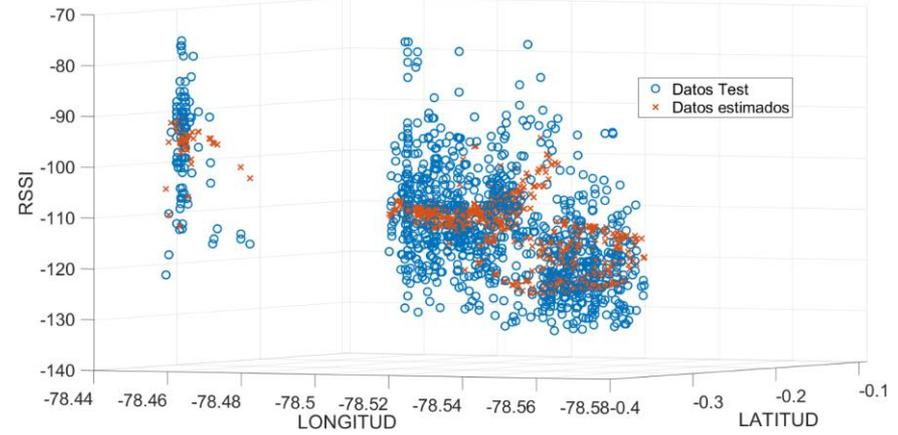
2 VARIABLES (LATITUD Y LONGITUD)

DATOS TEST VS DATOS ESTIMADOS



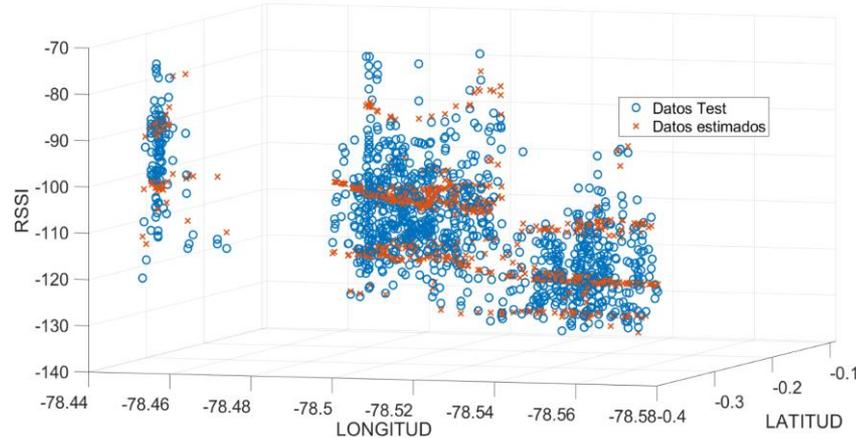
3 VARIABLES (LATITUD, LONGITUD Y ALTURA)

DATOS TEST VS DATOS ESTIMADOS



4 VARIABLES (LATITUD, LONGITUD, ALTURA Y LQI)

DATOS TEST VS DATOS ESTIMADOS



PRUEBAS Y RESULTADOS

PARÁMETROS LIBRES REGRESOR SUPPORT VECTOR MACHINE

2 VARIABLES (LATITUD Y LONGITUD)

Variable	Designación
Función Kernel	Gaussian
Escala del kernel	0.22068
Epsilon	0.012794
Restricción de caja	23.468

3 VARIABLES (LATITUD, LONGITUD Y ALTURA)

Variable	Designación
Función Kernel	Gaussian
Escala del kernel	0.21068
Epsilon	0.012794
Restricción de caja	23.468

4 VARIABLES (LATITUD, LONGITUD, ALTURA Y LQI)

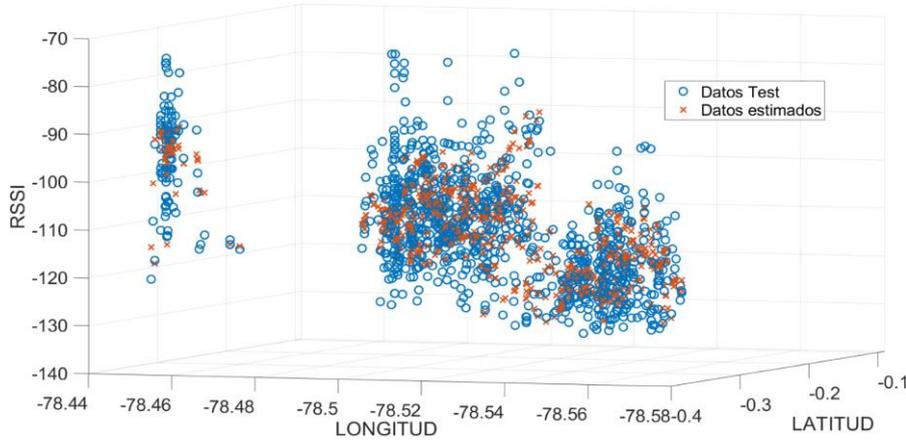
Variable	Designación
Función Kernel	Gaussian
Escala del kernel	0.21068
Epsilon	0.012794
Restricción de caja	23.468

PRUEBAS Y RESULTADOS

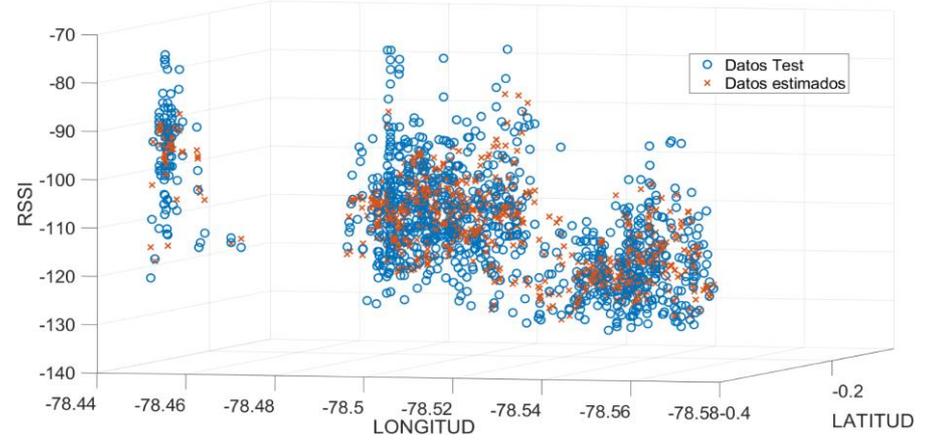
REGRESOR SUPPORT VECTOR MACHINE

2 VARIABLES (LATITUD Y LONGITUD) 3 VARIABLES (LATITUD, LONGITUD Y ALTURA)

DATOS TEST VS DATOS ESTIMADOS



DATOS TEST VS DATOS ESTIMADOS



4 VARIABLES (LATITUD, LONGITUD, ALTURA Y LQI)

DATOS TEST VS DATOS ESTIMADOS



PRUEBAS Y RESULTADOS

PARÁMETROS LIBRES REGRESOR ÁRBOL DECISIÓN

2 VARIABLES (LATITUD Y LONGITUD)

Variable	Designación
Tamaño mínimo de hoja	11

3 VARIABLES (LATITUD, LONGITUD Y ALTURA)

Variable	Designación
Tamaño mínimo de hoja	6

4 VARIABLES (LATITUD, LONGITUD, ALTURA Y LQI)

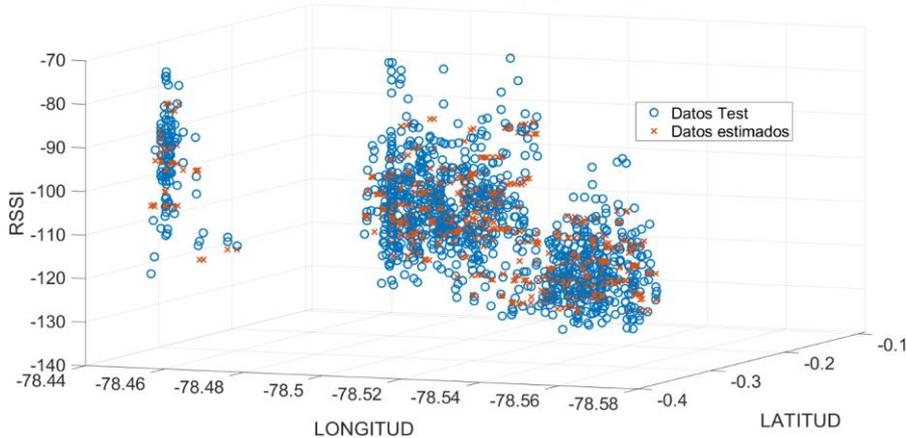
Variable	Designación
Tamaño mínimo de hoja	7

PRUEBAS Y RESULTADOS

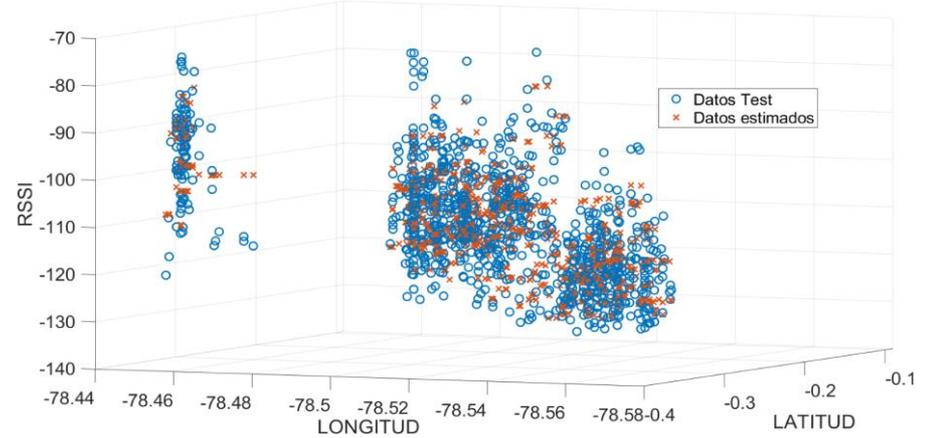
ÁRBOL DE DECISIÓN

2 VARIABLES (LATITUD Y LONGITUD) 3 VARIABLES (LATITUD, LONGITUD Y ALTURA)

DATOS TEST VS DATOS ESTIMADOS

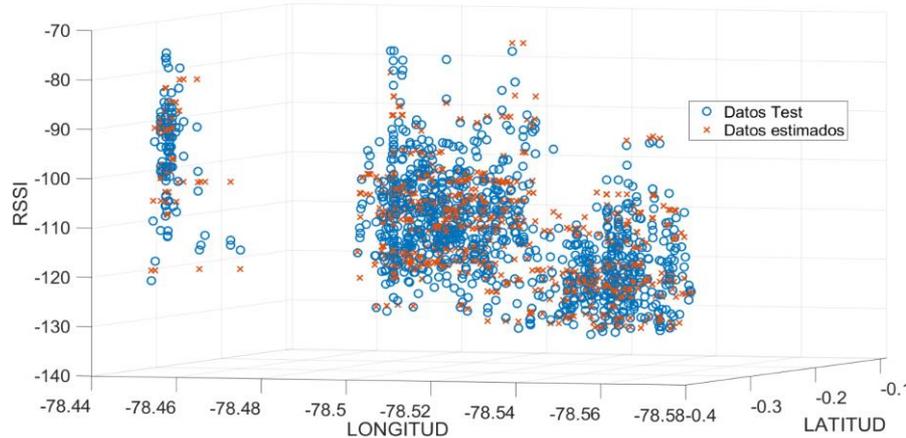


DATOS TEST VS DATOS ESTIMADOS



4 VARIABLES (LATITUD, LONGITUD, ALTURA Y LQI)

DATOS TEST VS DATOS ESTIMADOS



PRUEBAS Y RESULTADOS

MODELOS DE REGRESIÓN

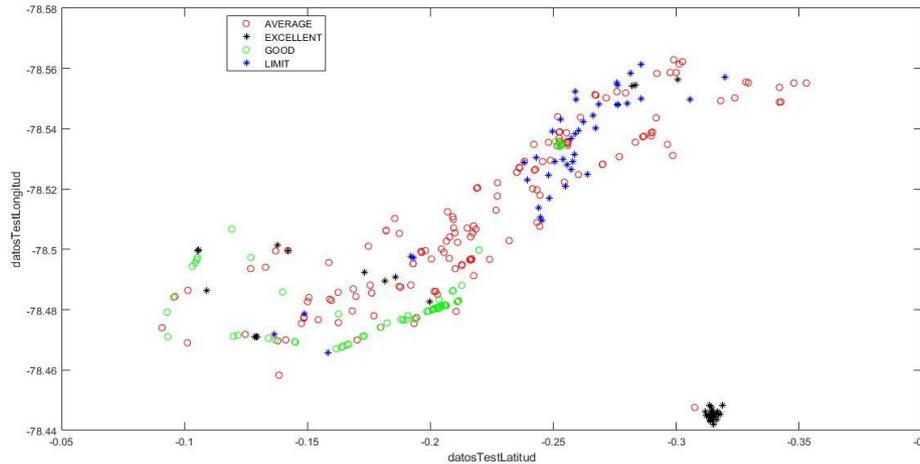
Obtención del error cuadrático medio (MSE) y la raíz del error cuadrático medio (RMSE):

	2 variables		3 variables		4 variables	
Regresor	MSE (dBm)	RMSE (dBm)	MSE (dBm)	RMSE (dBm)	MSE (dBm)	RMSE (dBm)
Lineal	85.7351	9.2593	85.6060	9.2524	24.1198	4.9112
Polinomial	69.4176	8.3317	66.2791	8.1412	20.8077	4.5615
SVM	53.8831	7.3008	53.7159	7.3291	20.2045	4.4949
DTs	57.2417	7.5658	56.7289	7.5319	17.9020	4.2311

Se emplea el modelo Support Vector Machine de 3 variables para llenar los valores nulos y en la generación de datos sintéticos.

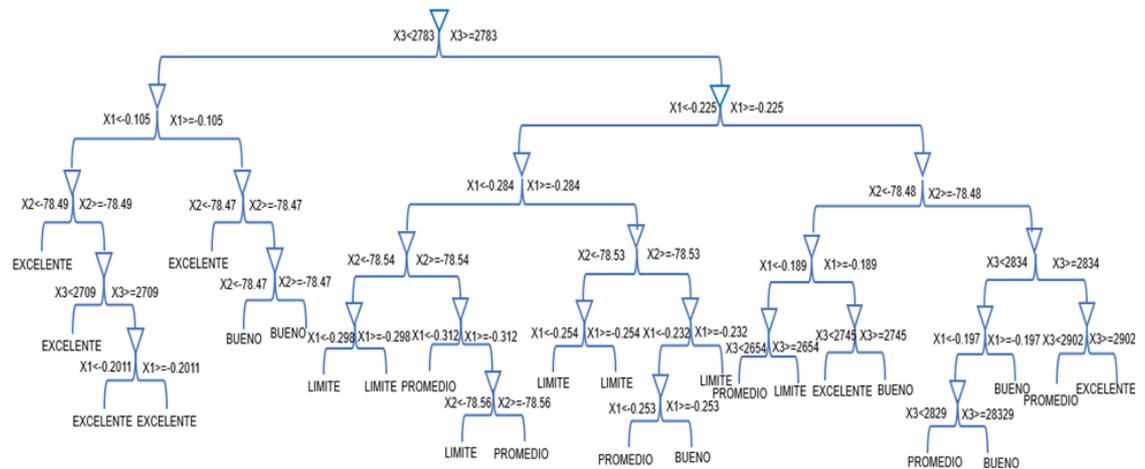
PRUEBAS Y RESULTADOS

Modelo de clasificación Support Vector Machine:



Modelo de clasificación Árbol de decisión:

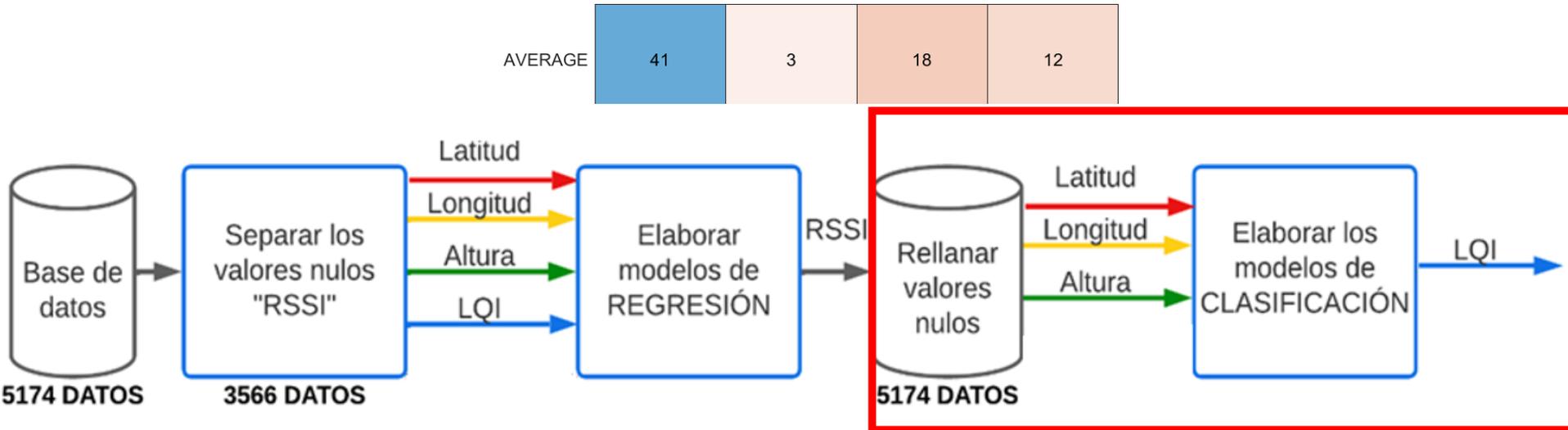
Etiquetas	Variable
$X1$	Latitud
$X2$	Longitud
$X3$	Altura



PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase **LIMITE** TAPA DE CLASIFICACIÓN



CLASIFICACIÓN

Etiquetas	Datos utilizados	Datos originales
LIMITE	243	243
EXCELENTE	243	383
PROMEDIO	243	1218
BUENO	243	3330

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase LIMITE

Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BER
PROMEDIO	55.41	90.32	81.44	66.13	0.27
EXCELENTE	84.52	93.24	90.72	83.53	0.11
BUENO	67.65	86.99	82.47	61.33	0.23
LIMITE	78.46	92.04	89	73.91	0.15
BER TOTAL	=	0.19			

PRUEBAS Y RESULTADOS

MODEL DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase EXCELENTE

True Class	AVERAGE	83	2	20	21
	EXCELLENT	4	82	5	2
	GOOD	25	10	90	6
	LIMIT	8		1	58
		AVERAGE	EXCELLENT	GOOD	LIMIT
		Predicted Class			

Etiquetas	Datos utilizados	Datos originales
LIMITE	243	243
EXCELENTE	383	383
PROMEDIO	383	1218
BUENO	383	3330

PRUEBAS Y RESULTADOS

MODEL DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase EXCELENTE

Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BER
PROMEDIO	65.87	87.29	80.82	69.17	0.23
EXCELENTE	88.17	96.3	94.48	87.23	0.08
BUENO	68.7	90.91	83.93	77.59	0.20
LIMITE	86.57	91.71	90.89	66.67	0.11
BER TOTAL	=	0.16			

PRUEBAS Y RESULTADOS

MODEL DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase PROMEDIO

True Class	AVERAGE	243	9	83	35
	EXCELLENT	4	98	17	
	GOOD	83	20	251	10
	LIMIT	34		5	26
		AVERAGE	EXCELLENT	GOOD	LIMIT
		Predicted Class			

Etiquetas	Datos utilizados	Datos originales
LIMITE	243	243
EXCELENTE	383	383
PROMEDIO	1218	1218
BUENO	1218	3330

PRUEBAS Y RESULTADOS

MODEL DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase PROMEDIO

Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BER
PROMEDIO	65.68	77.92	72.99	66.76	0.28
EXCELENTE	82.35	96.37	94.55	77.17	0.11
BUENO	69.96	81.05	76.25	70.51	0.25
LIMITE	40	94.73	90.85	36.62	0.33
BER TOTAL	=	0.24			

PRUEBAS Y RESULTADOS

MODEL DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase BUENO

True Class	AVERAGE	222	4	116	31
	EXCELLENT		83	25	
	GOOD	137	37	831	11
	LIMIT	24		5	26
		AVERAGE	EXCELLENT	GOOD	LIMIT
		Predicted Class			

Etiquetas	Datos utilizados	Datos originales
LIMITE	243	243
EXCELENTE	383	383
PROMEDIO	1218	1218
BUENO	3330	3330

PRUEBAS Y RESULTADOS

MODEL DE CLASIFICACIÓN ÁRBOL DE DECISIÓN

Datos Balanceados a clase BUENO

Mejor desempeño Datos Balanceados a clase EXCELENTE

Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BRR
PROMEDIO	69.82	88.294	80.82	56.67	0.023
EXCELENTE	88.83	97.316	95.48	66.23	0.038
BUENO	88.79	90.216	88.63	85.69	0.030
LIQUIDE	88.27	99.719	96.89	38.47	0.081
BRR TOTAL	=	00.83			

Clase	Datos utilizados
EXCELENTE	383

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase LIMITE

True Class	AVERAGE	49	11	24	20
	EXCELLENT	1	65	2	
	GOOD	11	9	48	4
	LIMIT	1		1	45
		AVERAGE	EXCELLENT	GOOD	LIMIT
		Predicted Class			

Etiquetas	Datos utilizados	Datos originales
<i>LIMITE</i>	243	243
<i>EXCELENTE</i>	243	383
<i>PROMEDIO</i>	243	1218
<i>BUENO</i>	243	3330

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase LIMITE

Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BER
PROMEDIO	47.12	93.05	76.63	79.03	0.30
EXCELENTE	95.59	91.03	92.1	76.47	0.07
BUENO	66.67	87.67	82.47	64	0.23
LIMITE	95.75	90.16	91.07	65.22	0.07
BER TOTAL	=	0.17			

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase EXCELENTE

True Class	AVERAGE	94	9	28	38
	EXCELLENT		76		
	GOOD	22	9	88	3
	LIMIT	4			46
		AVERAGE	EXCELLENT	GOOD	LIMIT
		Predicted Class			

Etiquetas	Datos utilizados	Datos originales
LIMITE	243	243
EXCELENTE	383	383
PROMEDIO	383	1218
BUENO	383	3330

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase EXCELENTE

Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BER
PROMEDIO	55.62	89.52	75.78	78.33	0.27
EXCELENTE	100	94.72	95.68	80.85	0.03
BUENO	72.13	90.51	85.13	75.86	0.19
LIMITE	92	88.83	89.21	52.87	0.1
BER TOTAL	=	0.15			

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase PROMEDIO

True Class	AVERAGE	236	5	79	26
	EXCELLENT	5	99	8	
	GOOD	94	23	261	16
	LIMIT	29		8	29
		AVERAGE	EXCELLENT	GOOD	LIMIT
		Predicted Class			

Etiquetas	Datos utilizados	Datos originales
<i>LIMITE</i>	243	243
<i>EXCELENTE</i>	383	383
<i>PROMEDIO</i>	1218	1218
<i>BUENO</i>	1218	3330

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase PROMEDIO

Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BER
PROMEDIO	68.21	77.62	74.07	64.84	0.27
EXCELENTE	88.39	96.53	95.53	77.95	0.08
BUENO	66.24	81.87	75.16	73.32	0.26
LIMITE	43.94	95.07	91.39	40.85	0.3
BER TOTAL	=	0.23			

PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase BUENO

True Class	AVERAGE	202	2	102	25
	EXCELLENT	2	83	30	
	GOOD	151	39	840	13
	LIMIT	28		5	30
		AVERAGE	EXCELLENT	GOOD	LIMIT
		Predicted Class			

Etiquetas	Datos utilizados	Datos originales
LIMITE	243	243
EXCELENTE	383	383
PROMEDIO	1218	1218
BUENO	3330	3330



PRUEBAS Y RESULTADOS

MODELO DE CLASIFICACIÓN SUPPORT VECTOR MACHINE

Datos Balanceados a clase BUENO

Mejor desempeño Datos Balanceados a clase EXCELENTE

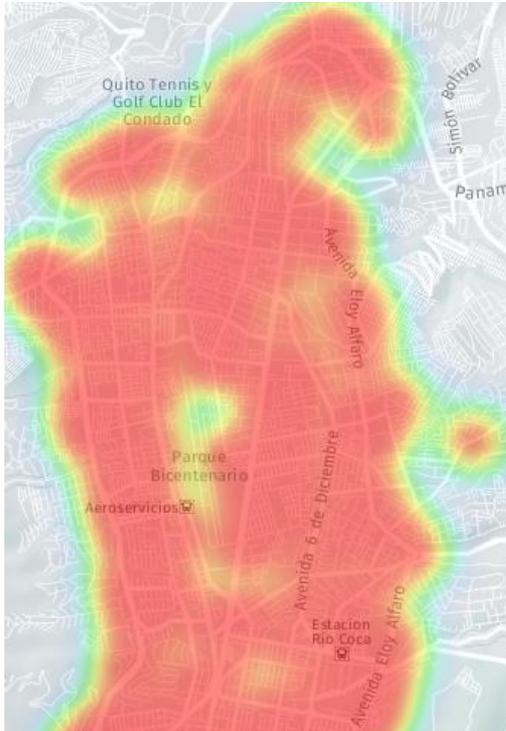
Clases	Sensibilidad (%)	Especificidad (%)	Exactitud (%)	Precisión (%)	BER
PROMEDIO	65.62	89.58	86.08	78.334	0.227
EXCELENTE	72.007	97.172	95.68	86.854	0.065
BUENO	80.53	90.58	88.09	75.88	0.128
LIMITE	49.82	88.85	89.23	52.872	0.27
BER TOTAL	=	0.15			

Clase	Datos utilizados
<i>EXCELENTE</i>	383

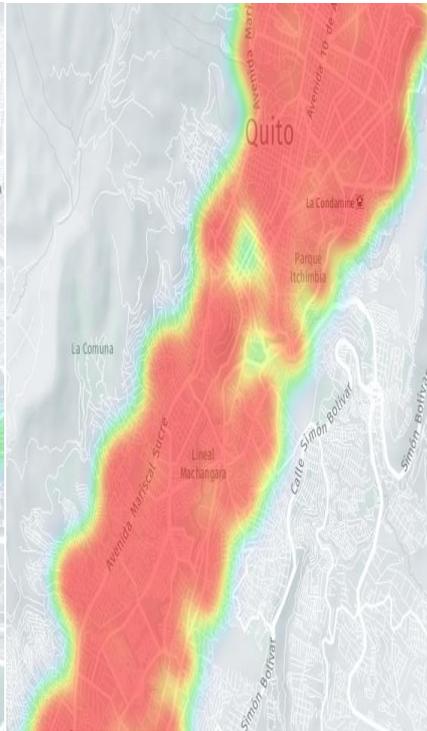


PRUEBAS Y RESULTADOS

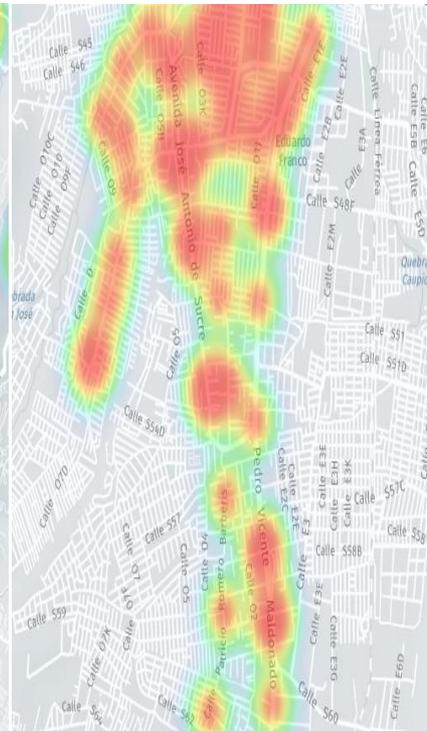
Mapas de calor con la base de datos original (5174 Datos)



Norte



Centro



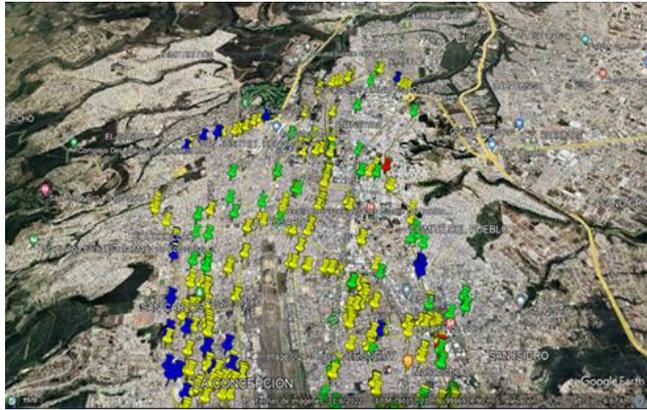
Sur



ESPE

PRUEBAS Y RESULTADOS

Categorización de forma geográfica base datos inicial (5174 Datos)



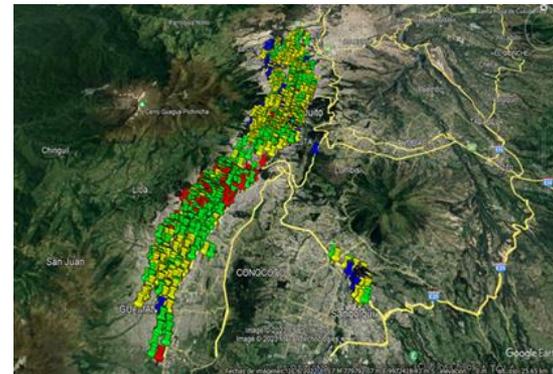
Norte



Centro



Sur



Etiqueta	Color
EXCELENTE	Blue
BUENO	Yellow
PROMEDIO	Green
LIMITE	Red

PRUEBAS Y RESULTADOS

GENERACIÓN DE NUEVAS COORDENADAS

openforis COLLECT EARTH

Buscar lugares

Earth Engine Apps

Selección de AUI: Ecuador

Selección de Provincia: Pichincha

Selección de Distrito: Quito

Use un archivo de forma (con polígonos) para generar la cuadrícula (subir a activos GEE)

ID de activo de GEE

Establezca la EPSG que mejor se adapte a tu área de estudio. Para países grandes recomendamos utilizar EPSG:4326. Busque los códigos EPSG aquí: <http://epsg.io>

EPSG:4326

Agregar de Agregar de Añadir KML Agregar de Generar gr Establezca la

Generar nuevas coordenadas GOOGLE EARTH

Latitud

Longitud

Altura

DATA AUGMENTATION

RSSI

LOI

BASE DE DATOS AUMENTADA

62942 DATOS

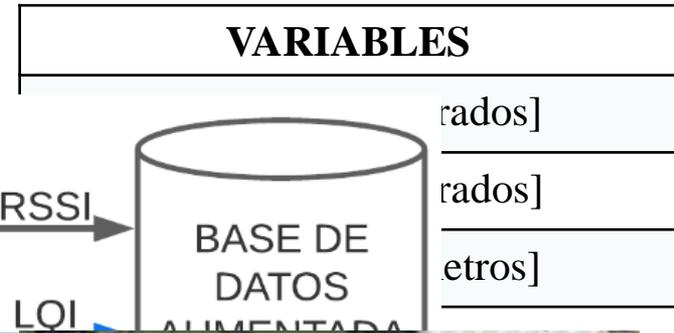
Google Earth

17 M 773975.42 m E 9982830.89 m S elevación 3966 m alt. ojo 85.35 km

Base de datos aumentada:

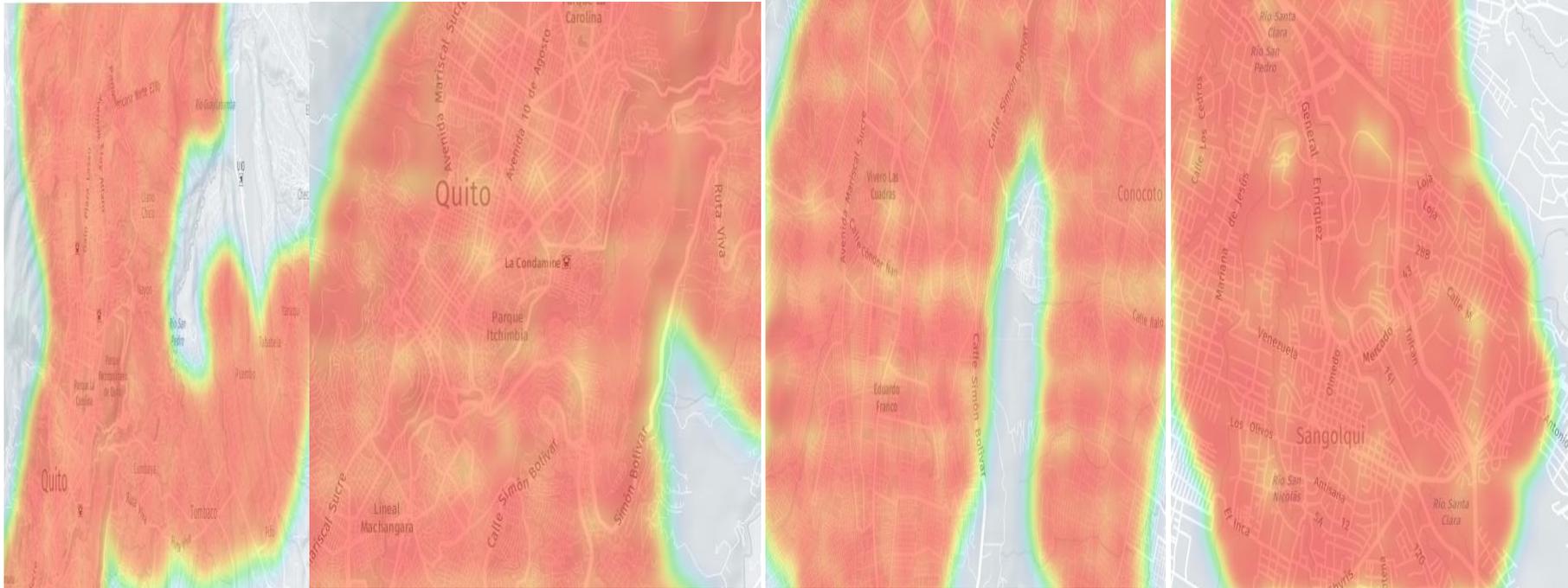
- Número total de datos: **62942**

VARIABLES



PRUEBAS Y RESULTADOS

Mapas de calor con Data Augmentation (62942 Datos)



Norte

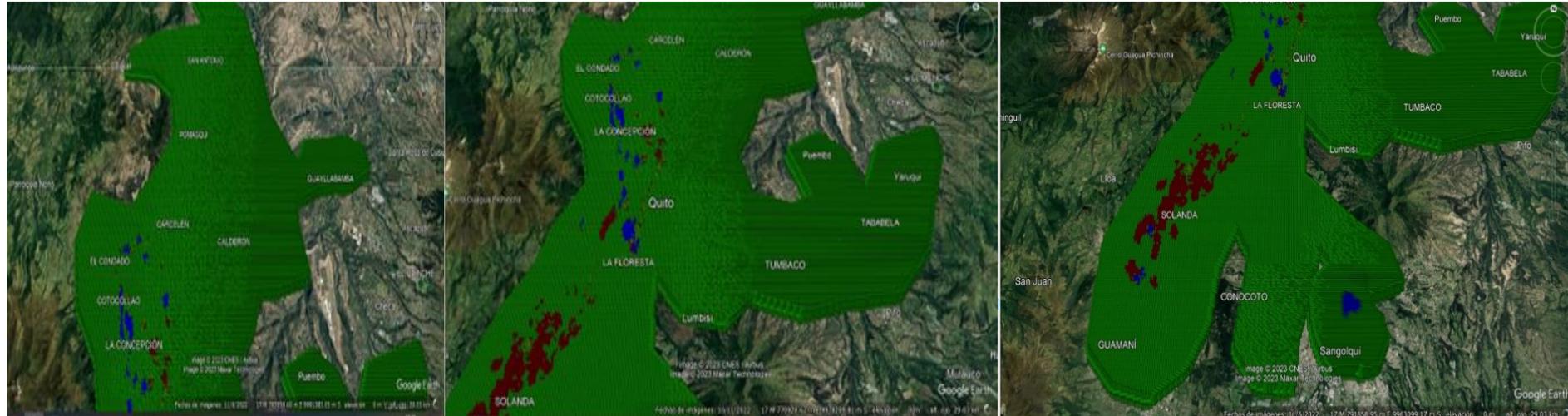
Centro

Sur

ESPE

PRUEBAS Y RESULTADOS

Categorización de forma geográfica Data Augmentation (62942 Datos)



Norte

Centro

Sur



Etiqueta	Color
EXCELENTE	Blue
BUENO	Yellow
PROMEDIO	Green
LIMITE	Red

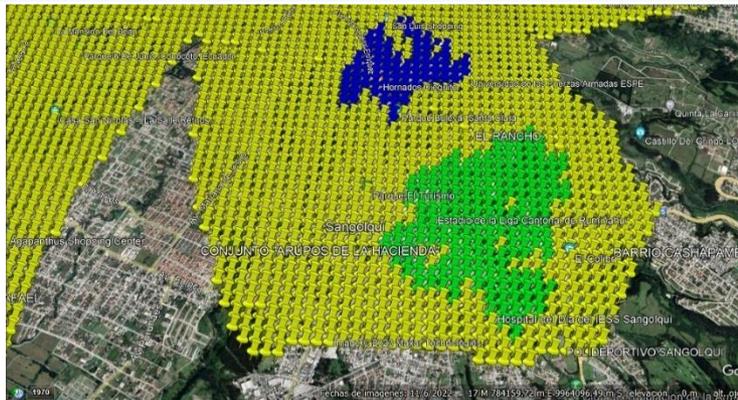
PRUEBAS Y RESULTADOS

Categorización de forma geográfica Data Augmentation en función del RSSI (62942 Datos)



TABLA DE Clasificación del LQI en función del RSSI

LQI	Rangos [dBm]
EXCELENTE	[-94 → -66]
BUENO	[-115 → -95]
PROMEDIO	[-127 → -116]
LIMITE	[-133 → -128]



ESPE

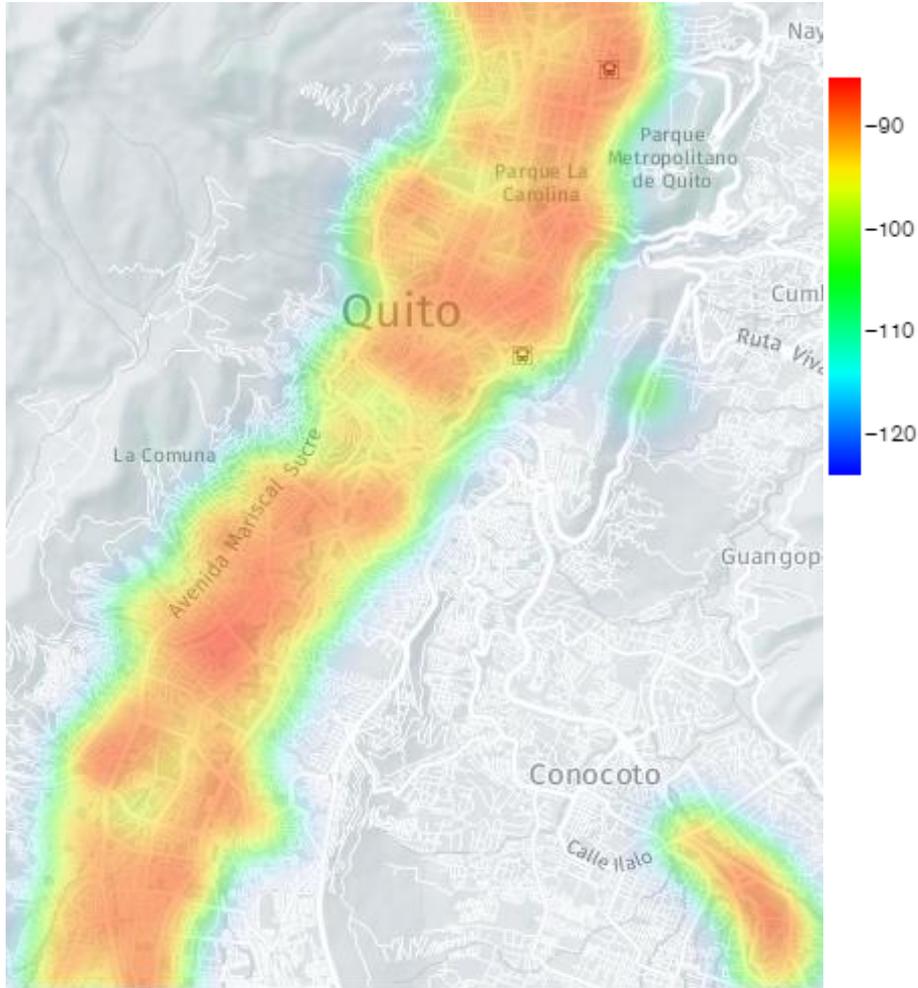


ESPE

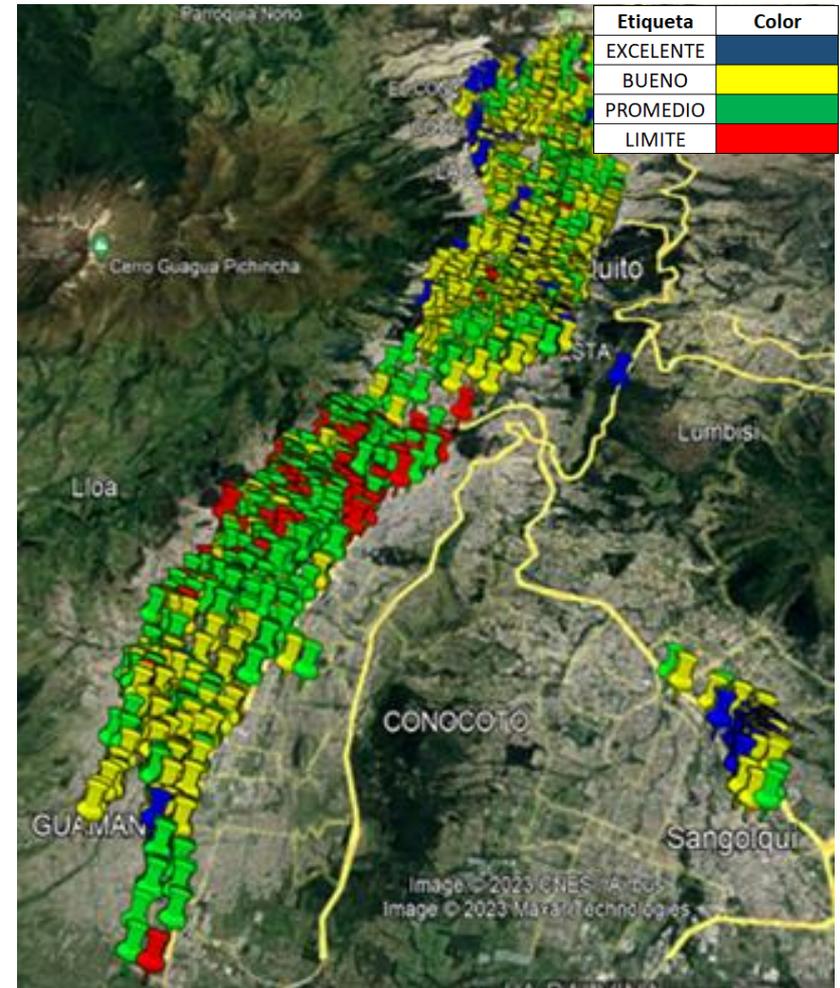
Etiqueta	Color
EXCELENTE	
BUENO	
PROMEDIO	
LIMITE	

PRUEBAS Y RESULTADOS

Mapa de calor (5174 datos) RSSI



Etiquetas LQI (5174 datos)



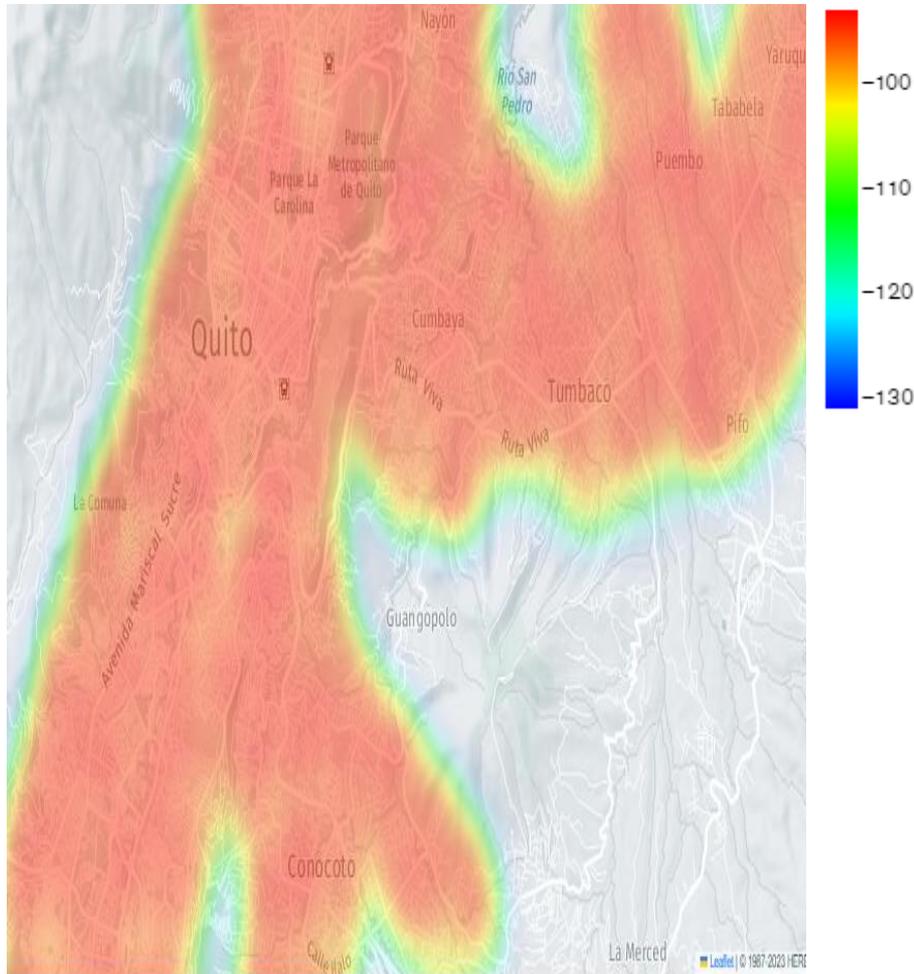
45



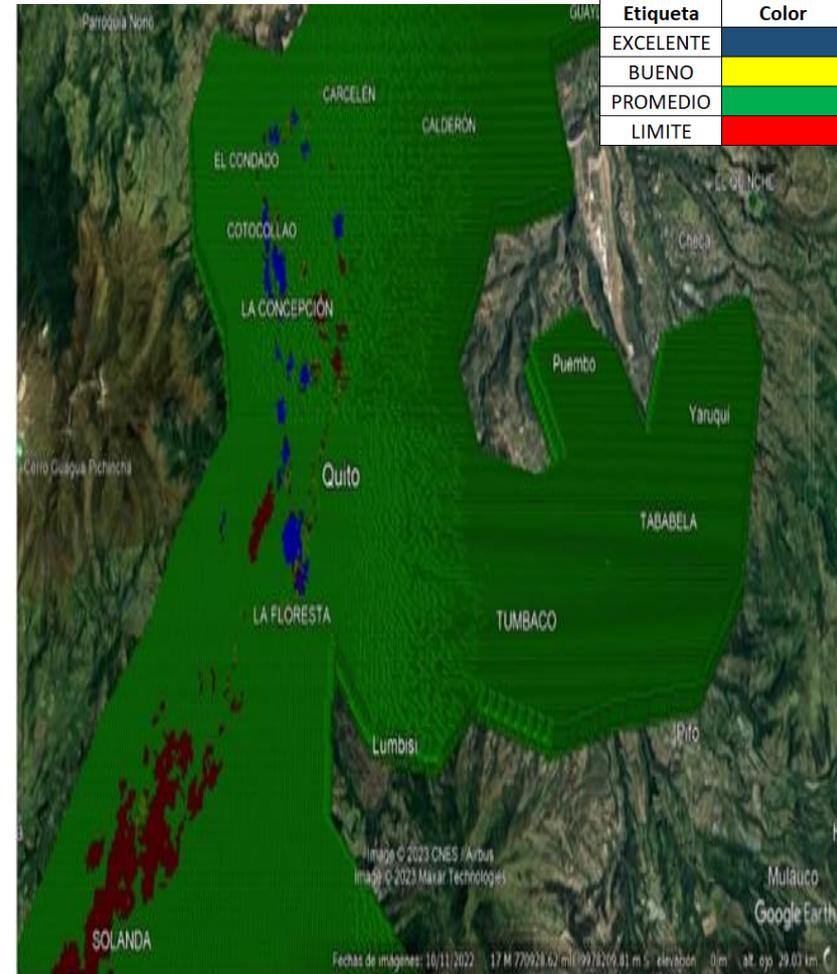
ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

PRUEBAS Y RESULTADOS

Mapa de calor (62942datos) RSSI

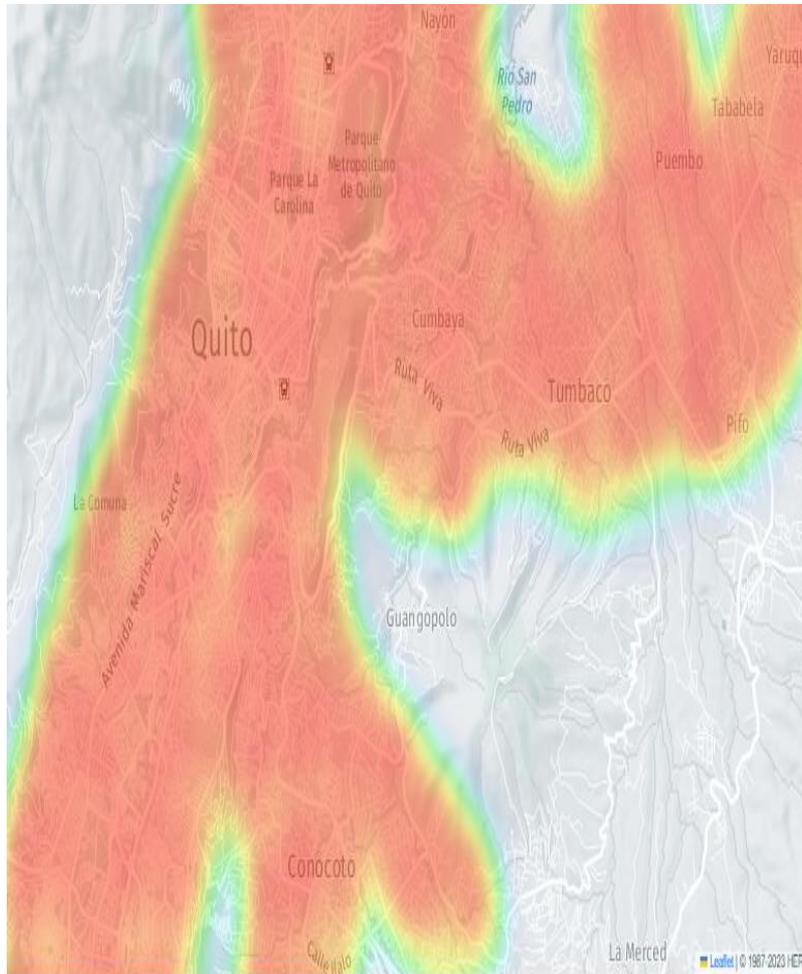


Categorización (62942datos) LQI Modelo de clasificación

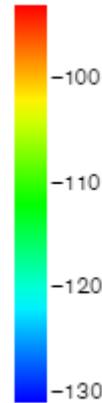


PRUEBAS Y RESULTADOS

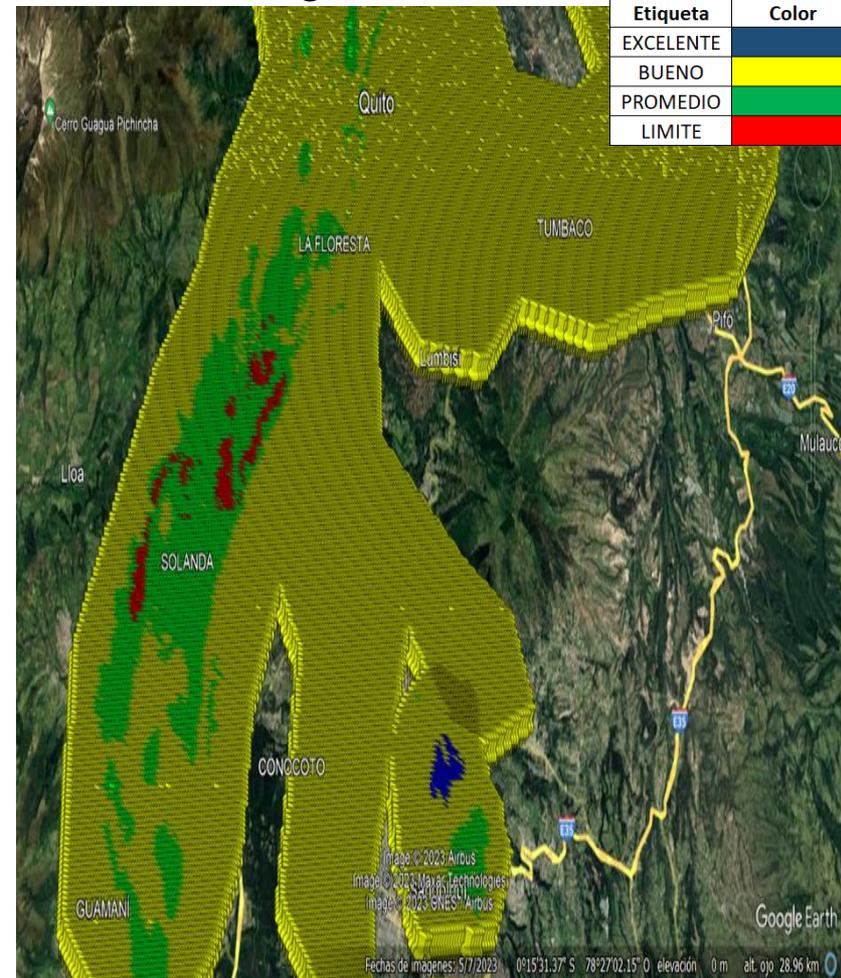
Mapa de calor (62942datos) RSSI



Categorización (62942datos) LQI LQI función del RSSI



Etiqueta	Color
EXCELENTE	Blue
BUENO	Yellow
PROMEDIO	Green
LIMITE	Red



CONCLUSIONES Y TRABAJOS FUTUROS

- Se identificó que para el regresor los mejores resultados fueron obtenidos en base a los valores de MSE y RMSE, donde el regresor de menor error corresponde al modelo de SVM con un MSE de 53.73 dBm y un RMSE de 7.33 dBm.
- Con este resultado se concluye el mejor modelo para el regresor y se completó los 2145 valores restantes o valores nulos correspondientes al RSSI, de esta manera se rellenó la base de datos inicial de 5137 datos.
- Se identificaron los mejores resultados de los clasificadores en base a Machine Learning tradicional con los valores correspondientes al 70% para entrenamiento y 30% para validación, además mediante el proceso de balanceamiento de datos se estimó que el mejor resultado alcanzado por ambos modelos de clasificación (DTs y SVM) fue el balanceamiento de la clase EXCELENTE con 383 valores de los 5137.
- De esta manera se obtuvo para el modelo de SVM balanceado a la clase EXCELENTE una exactitud del 95.68% y un BER total de 0.15, por otro lado, para el modelo de DTs balanceado a la clase EXCELENTE se obtuvo una exactitud del 94.48% y un BER total de 0.16. Por lo tanto, se concluye que el mejor modelo que se adapta mejor a la base de datos para clasificación es el SVM con un BER del 0.15.

CONCLUSIONES Y TRABAJOS FUTUROS

- Finalmente, se identificó luego de una comparación visual entre la base de datos aumentada y la base de datos original, que la zona de cobertura de la base de datos aumentada es considerablemente mayor en comparación con la base de datos original ya que ahora está abarca áreas donde antes no existía cobertura o, en su defecto era muy débil.
- Se recomienda al momento de realizar la generación de datos se debe seleccionar un software que puede generar las variables latitud, longitud y altura porque son variables necesarias en los modelos de entrenamiento, además de efectuar un trazado de área acorde al caso de estudio dentro del cantón Quito al reducir las áreas con nivel forestal alto y seleccionar las zonas urbanas.
- De acuerdo al trabajo realizado se puede realizar la publicación de un artículo científico donde se describan los modelos tradicionales de Machine Learning de regresión y clasificación desarrollados y los diferentes resultados obtenidos para la generación de datos sintéticos en una red Sigfox.
- Se pueden desarrollar otros modelos de regresión y clasificación utilizando Deep Learning para verificar el comportamiento de dichos modelos mediante el uso de imágenes en función de la base de datos aumentada, de acuerdo a esto se podrán tener modelos con diferente desempeño.



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA



**DEPARTAMENTO DE ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES
CARRERA DE INGENIERÍA EN ELECTRÓNICA Y TELECOMUNICACIONES**

**TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO
EN ELECTRÓNICA Y TELECOMUNICACIONES**

TEMA:

***“Generación de datos sintéticos a partir de mediciones RSSI en una red SigFox utilizando
Data Augmentation.”***

Autores: Kevin Alexander Leiton Reina, Dennis Eduardo Tigse Pérez

Director del Proyecto: Ing. Román Alcides Lara Cueva, PhD.

Cotutor: Ing. Julio César Larco Bravo, MSc.

