

UNIVERSIDAD CATÓLICA DE LA SANTÍSIMA CONCEPCIÓN

FACULTAD DE INGENIERÍA
DEPARTAMENTO DE MEDIO AMBIENTE Y ENERGÍA



UCSC

Aplicación de ciencia de datos para identificación de suciedad y sombreado en MicroRed UCSC.

Boris Josue Ortiz Alarcón

Informe de Habilitación Profesional para optar al título de:
Ingeniero Civil Eléctrico

Profesor Patrocinante:
Dr. Samuel Vergara.
Mg. Roberto Aedo

Profesores Guía:
Dr. Guillermo Ramírez .

Concepción, Agosto 2023

UNIVERSIDAD CATÓLICA DE LA SANTISIMA CONCEPCION
Facultad de Ingeniería
Departamento de Medio Ambiente y Energía

Profesor Patrocinante:
Dr. Guillermo Ramírez .

Aplicación de ciencia de datos para identificación de suciedad y sombreado en MicroRed UCSC.

Boris Josue Ortiz Alarcón

Informe de Habilitación Profesional
para optar al Título de

Ingeniero Civil Eléctrico

Agosto 2023

Resumen

En este estudio se desarrollaron varios algoritmos de aprendizaje automático para la detección de sombreado y suciedad en paneles solares en la Microred UCSC. El objetivo principal fue identificar posibles fallas en el sistema fotovoltaico y comprender los efectos del sombreado y la suciedad en dichos sistemas.

Se analizaron y compararon diferentes algoritmos de clasificación, como regresión logística, vecinos cercanos y redes neuronales, en términos de sus métricas de rendimiento. Los resultados mostraron que el modelo de vecinos cercanos obtuvo un rendimiento superior al 90% tanto en la detección de sombreado como en la detección de suciedad, en comparación con el modelo de regresión logística.

Además, se entrenaron modelos de salida multiclase para detectar ambas condiciones simultáneamente. Tanto el modelo de red neuronal como el de vecinos cercanos mostraron un rendimiento similar, con la métrica de exactitud de 96.0 % y 95,5% respectivamente. Estos demuestran la eficacia de los algoritmos de aprendizaje automático en la detección de sombreado y suciedad en sistemas fotovoltaicos.

“Si quieres encontrar los secretos del universo, piensa
en terminos de energía, frecuencia y vibración”
- **Nikola Tesla**

A los alumnos del pasado, presente y futuro de la carrera de Ingeniería Civil Eléctrica de la U.C.S.C.

Agradecimientos

Primero agradecer a mi familia por apoyarme en todo mi proceso universitario siempre con buenos deseos, así también a los diferentes profesores que me acompañaron en este proceso orientándome y apoyándome. Especialmente el Dr Guillermo Ramirez.

Es imposible no recordar todos los buenos momentos que pase durante mi proceso de estudiante de la UCSC, por eso quiero agradecer a todos los profesores los cuales me hicieron clases y con los que comparti en esta etapa.

A mis amigos que conoci y estuvieron conmigo, como Braulio Silva quien me ayudo durante la realizacion de varias pruebas, Esteban Concha, Cinthia Altamirano, Eduardo Carrasco, Sebastian Freude, Diego Marin, Maverick Muñoz, Evelyn Pedreros, Benjamin Fernandez y Crhistoper Carrasco los Unstoppables Machine.

Asi tambien agradecer a las personas con las que forme equipo para liderar el centro de estudiantes, Juan Acevedo, Cristian Monsalve, Tabata Ponce, Felipe Barrera, Jorge Nuñez e Ignacio Quiroz, con quienes ademas de ser un equipo de trabajo entablamos amistad.

Asi tambien a todos los estudiantes de Ingenieria Civil Electrica y de la UCSC con los que comparti haciendose parte de mi dia a dia.

Tabla de Contenidos

LISTA DE TABLAS	IX
LISTA DE FIGURAS	X
ABREVIACIONES	XIII
CAPÍTULO 1. INTRODUCCIÓN	1
1.1. INTRODUCCIÓN GENERAL	1
1.2. TRABAJOS PREVIOS	2
<i>Funcionamiento de sistemas fotovoltaicos ante condiciones de suciedad y sombreado y su detección con variables eléctricas.....</i>	<i>2</i>
<i>Aplicación de la Ciencia de Datos en la energía Solar FV</i>	<i>3</i>
<i>Investigaciones previas realizadas en la MicroRed UCSC</i>	<i>7</i>
<i>Discusión</i>	<i>8</i>
1.3. HIPÓTESIS DE TRABAJO	9
1.4. OBJETIVOS	9
<i>Objetivo General</i>	<i>9</i>
<i>Objetivos Específicos</i>	<i>9</i>
1.5. ALCANCES Y LIMITACIONES	9
1.6. TEMARIO Y METODOLOGÍA	10
CAPÍTULO 2. IDENTIFICACIÓN DE FALLAS EN SISTEMAS FOTOVOLTAICOS	11
2.1. INTRODUCCIÓN	11
2.2. FALLAS EN LOS SISTEMAS FOTOVOLTAICOS	11
2.2.1 <i>Fallas internas del Sistema.....</i>	<i>12</i>
2.2.2 <i>Fallas en los módulos Fotovoltaicos</i>	<i>13</i>
2.2.3 <i>Fallas Eléctricas</i>	<i>16</i>
2.2.4 <i>Fallas externas al sistema.....</i>	<i>18</i>
CAPÍTULO 3. PRUEBAS EXPERIMENTALES DE SOMBREADO Y SUCIEDAD EN MICRORED UCSC..	19
3.1. INTRODUCCIÓN	19
3.2. EQUIPOS ELÉCTRICOS UTILIZADOS	19
3.3. PRUEBAS DE SOMBREADO	24
3.3.1 <i>Metodología</i>	<i>24</i>
3.3.2 <i>Sombreado de un solo panel Fotovoltaico.....</i>	<i>27</i>
3.3.3 <i>Sombreado de dos o más paneles Fotovoltaicos.....</i>	<i>34</i>
3.4. PRUEBAS DE SUCIEDAD	39
3.4.1 <i>Metodología</i>	<i>39</i>
3.4.2 <i>Pruebas de suciedad en un panel Fotovoltaico</i>	<i>43</i>
3.4.3 <i>Pruebas de suciedad en dos o más paneles Fotovoltaicos.....</i>	<i>49</i>
CAPÍTULO 4. ALGORITMOS DE CLASIFICACION	55
4.1. INTRODUCCIÓN	55
4.2. MACHINE LEARNING	55
4.2.1 <i>Algoritmo de regresión logística.....</i>	<i>56</i>
4.2.2 <i>Algoritmo de k vecinos más cercanos</i>	<i>59</i>
4.2.3 <i>Redes Neuronales</i>	<i>61</i>
4.3. FUNCIONES DE ACTIVACIÓN	64
4.3.1 <i>Función de activación ReLU</i>	<i>64</i>
4.3.2 <i>Función de activación Leaky ReLU</i>	<i>65</i>
4.3.3 <i>Función de activación Lineal.....</i>	<i>66</i>
4.3.4 <i>Función de activación Softmax</i>	<i>67</i>
4.4. MÉTRICAS DE RENDIMIENTO DE ALGORITMOS DE CLASIFICACIÓN	69

CAPÍTULO 5. ALGORITMOS DE CLASIFICACIÓN PARA DETECCIÓN DE SOMBREADO Y SUCIEDAD	72
5.1. INTRODUCCION	72
5.2. OBTENCION DE DATOS	72
5.3. DATOS UTILIZADOS	73
5.3.1 Base de datos Sombreado	74
5.3.2 Base de datos Suciedad.....	80
5.3.3 Base de Datos Sombreado y Suciedad.....	82
5.4. MODELOS UTILIZADOS	83
5.4.1 Modelos de Clasificación Sombreado.....	84
5.4.2 Modelos de Clasificación Suciedad	87
5.4.3 Modelos de Clasificación de Sombreado y Suciedad.....	92
5.5 VALIDACIÓN DE RESULTADOS CON DATOS DE OTOÑO	98
5.5.1 Algoritmos de clasificación de sombreado	99
5.5.2 Algoritmos de clasificación de suciedad.....	101
5.5.3 Algoritmos de clasificación de sombreado y suciedad	102
5.6 VALIDACIÓN DE RESULTADOS CON DATOS PRIMAVERA Y/O VERANO.	106
5.6.1 Algoritmos de clasificación de sombreado	106
5.6.2 Algoritmos de clasificación de suciedad.....	107
5.6.3 Algoritmos de clasificación de sombreado y suciedad	107
CAPÍTULO 6. CONCLUSIONES	109
6.1. SUMARIO	109
6.2. CONCLUSIONES.....	109
6.3. TRABAJO FUTURO.....	110
BIBLIOGRAFÍA.....	111
ANEXOS 113	
ANEXO 1: CÓDIGOS EJEMPLO DE LOS DISTINTOS MODELOS UTILIZADOS	113
ANEXO 2: GRÁFICAS LIMPIEZA BASE DE DATOS SUCIEDAD.	117
ANEXO 3: CORRELACIÓN DE VARIABLES EN LAS DIFERENTES BASES DE DATOS	121

Lista de Tablas

Tabla 2.1: Resumen de fallas en paneles FV y sus consecuencias	17
Tabla 3.1: Resumen de las secciones utilizadas	20
Tabla 3.2: Especificaciones técnicas Panel Monocristalino.....	22
Tabla 3.3: Especificaciones técnicas Panel Policristalino.....	22
Tabla 3.4: Especificaciones técnicas inversor Fronius Symo 10 kW	23
Tabla 3.5 Pruebas de Sombreado Realizadas.....	26
Tabla 3.6 Comparación casos de sombreado de un panel FV.....	33
Tabla 3.7 Comportamiento de las variables electricas en condiciones de sombreado de dos o mas modulos FV	38
Tabla 3.8 Pruebas de Suciedad Realizadas	41
Tabla 3.9 Resultados de Pruebas de suciedad en un panel.....	48
Tabla 3.10 Resultados de los casos de suciedad (Caso 5 – Caso 8).....	53
Tabla 4.1: Distribución Matriz de confusión	70
Tabla 5.1: Tabla Resumen de las Variables a utilizar en los algoritmos de clasificación	74
Tabla 5.2 Clasificación de base de datos de Sombreado.....	76
Tabla 5.3: Clasificación de base de datos de Suciedad	82
Tabla 5.4: Clases de la base de datos de sombreado y suciedad	82
Tabla 5.5:Distribución Matriz de confusión Regresión logística sombreado	85
Tabla 5.6: Métricas de rendimiento Algoritmo Regresión Logística (Sombreado).....	85
Tabla 5.7: Distribución Matriz de confusión KNN para sombreado	87
Tabla 5.8: Métricas de rendimiento Algoritmo Vecinos cercanos (Sombreado).....	87
Tabla 5.9: Distribución Matriz de confusión RL para suciedad	89
Tabla 5.10: Métricas de rendimiento Algoritmo Regresión Logística (Suciedad)	89
Tabla 5.11: Distribución Matriz de confusión KNN para suciedad.....	91
Tabla 5.12: Métricas de rendimiento Algoritmo Vecinos cercanos (Suciedad)	91
Tabla 5.13: Distribución Matriz de confusión Vecinos Cercanos para suciedad y sombreado	93
Tabla 5.14: Métricas de rendimiento Algoritmo Vecinos cercanos (Sombreado y Suciedad)	93
Tabla 5.15: Comparación resultado de RN para diferentes estructuras	95
Tabla 5.16 Rendimiento de los algoritmos con 3000 iteraciones	96
Tabla 5.17: Distribución Matriz de confusión Red Neuronal para suciedad y sombreado.....	97
Tabla 5.18: Métricas de rendimiento Red Neuronal Estructura 2 (Sombreado y Suciedad)	97
Tabla 5.19: Detalle de las pruebas realizadas en construcción de base de datos para validación.....	98
Tabla 5.20: Descripción de las bases de datos para validación.....	99
Tabla 5.21: Comparación de rendimientos de algoritmos usados en detección de sombreado en otoño.....	99
Tabla 5.22:Distribución Matriz de confusión KNN otoño validación de resultados sombreado	100
Tabla 5.23: Comparación de rendimientos de algoritmos usados en detección de suciedad en otoño	101
Tabla 5.24: Distribución Matriz de confusión KNN otoño validación de resultados suciedad.....	101
Tabla 5.25: Métricas de rendimiento KNN Otoño Validación de Sombreado y Suciedad.....	103
Tabla 5.26: Distribución Matriz de confusión KNN Otoño validación sombreado y suciedad.....	103
Tabla 5.27: Métricas de rendimiento Red Neuronal Otoño validación de sombreado y suciedad ..	104
Tabla 5.28: Distribución Matriz de confusión RN Otoño validación de resultados sombreado y suciedad.....	104

Tabla 5.29: Comparación de rendimientos de Validación algoritmos para detección de sombreado primavera.....	106
Tabla 5.30: Comparación de rendimientos de Validación algoritmos para detección de suciedad verano	107
Tabla 5.31: Métricas de rendimiento de Validación algoritmo KNN Sombreado y Suciedad (primavera - verano).....	107
Tabla 5.32: Métricas de rendimiento de Validación Red Neuronal Sombreado y Suciedad (primavera - verano)	107

Lista de Figuras

Figura 2.1 Estructura interna de un panel fotovoltaico	13
Figura 2.2 Interior de caja de conexiones panel fotovoltaico	14
Figura 2.3: Comportamiento de los Diodos By-Pass	14
Figura 2.4: Configuración de Diodos By-Pass y Diodo de Bloqueo	15
Figura 3.1: Distribución Física de Paneles FV por Strings	19
Figura. 3.2: Módulos Fotovoltaicos utilizados en la investigación,	21
Figura. 3.3: Inversor Fronius Symo 10 kW	23
Figura 3.4 Elementos Utilizados para emular sombreado,	24
Figura 3.5: Sombreado de manera experimental de paneles fotovoltaicos	25
Figura 3.6 Representación de prueba de sombreado total de un panel FV	27
Figura 3.7 Muestra el comportamiento del voltaje y la corriente sin sombreado	27
Figura 3.8 Gráficas de corriente y voltaje con un panel FV completamente sombreado 06 de Octubre,	28
Figura 3.9 Representación de pruebas de sombreado parcial dos paneles FV del mismo string.....	29
Figura 3.10 Gráficas de corriente y voltaje con dos paneles FV sombreados a la mitad 13 de Octubre,	29
Figura 3.11: Representación de pruebas de sombreado parcial dos paneles FV de diferentes string	30
Figura 3.12: Representación de pruebas de sombreado parcial dos paneles FV de diferentes strings	31
Figura 3.13 Gráficas de corriente y voltaje con dos paneles FV sombreados la mitad inferior 16 de Octubre,	31
Figura 3.14 Gráficas de corriente y voltaje con dos paneles FV sombreados la mitad Superior 17 de Octubre	32
Figura 3.15 Representación de prueba de sombreado total dos paneles FV de diferentes strings.....	34
Figura 3.16 Variables eléctricas en condiciones de dos paneles FV sombreados distintas strings,...	34
Figura 3.17 Representación de pruebas de sombreado total dos paneles FV del mismo string.....	35
Figura 3.18 Variables eléctricas en condiciones de dos paneles FV sombreados del mismo string..	35
Figura 3.19 Representación de pruebas de sombreado parcial cuatro paneles FV de distintos strings	36
Figura 3.20 Comportamiento variables eléctricas en condiciones de sombreado,	36
Figura 3.21 Representación de pruebas de sombreado total de cuatro paneles FV del distintos strings	37
Figura 3.22 Variables eléctricas en condiciones de sombreado de 4 paneles FV de dos distintos strings,	37
Figura 3.23: Elementos utilizados para emular la suciedad en los Paneles,	40
Figura 3.24 Muestra la comparación entre un panel limpio y otro ensuciado intencionalmente.....	42
Figura 3.25 Comparación entre un panel limpio y uno sucio por condiciones del entorno.....	42
Figura 3.26 Comportamiento de las variables eléctricas cuando los paneles no presentan suciedad, 43	43
Figura 3.27 Comportamiento de las variables eléctricas Caso 1 Suciedad	44
Figura 3.28 Comportamiento Variables eléctricas Caso 2 Suciedad	45
Figura 3.29 Comportamiento Variables eléctricas Caso 3	46
Figura 3.30 Comportamiento Variables eléctricas Caso 4	47
Figura 3.31 Comportamiento Variables eléctricas Caso 5,	49
Figura 3.32 Comportamiento Variables eléctricas Caso 6,	50

Figura 3.33 Comportamiento Variables eléctricas Caso 7,	51
Figura 3.34 Comportamiento Variables eléctricas Caso 8,	52
Figura. 4.1: Gráfica de la función sigmoid.....	56
Figura 4.2: Estructura de redes neuronales	61
Figura. 4.3: Comportamiento interno de red Neuronal	62
Figura 4.4: Gráfica de Función de activación ReLU	64
Figura 4.5: Gráfica de Función de activación leaky ReLU	65
Figura 4.6 Gráfica de Función de activación Lineal	66
Figura 5.1 Proceso de limpieza de datos base de datos de sombreado,	75
Figura 5.2 Base de Datos de sombreado posterior a la limpieza de los datos.....	75
Figura 5.3 Comparación Corriente Continua base de datos de sombreado,	76
Figura 5.4 Comparación Voltaje en CC base de datos de sombreado,	77
Figura 5.5 Comparación Potencia en CC base de datos de sombreado,	77
Figura 5.6 Comparación Temperatura del panel FV base de datos de sombreado,	78
Figura 5.7 Comparación Temperatura Ambiente base de datos de sombreado,	78
Figura 5.8 Comparación Velocidad del Viento base de datos de sombreado,	79
Figura 5.9 Comparación Radiación base de datos de sombreado,	79
Figura 5.10 Proceso de limpieza de datos base de datos de suciedad	81
Figura 5.11: Base de datos de suciedad posterior a la depuración.....	81
Figura 5.12: Base de datos Mixta de suciedad y Sombreado.....	82
Figura 5.13: Matriz de confusión Algoritmo de Regresión Logística en detección de sombreado. ..	84
Figura 5.14: Selección de valor K para el algoritmo de clasificación sombreado	86
Figura 5.15: Matriz de confusión Algoritmo de Vecinos Cercanos en detección de sombreado.	86
Figura 5.16: Matriz de confusión normalizada Algoritmo de Regresión Logística en detección de suciedad.	88
Figura 5.17: Selección de valor K para el algoritmo de clasificación suciedad.....	90
Figura 5.18: Matriz de confusión Algoritmo de Vecinos Cercanos en detección de suciedad.	90
Figura 5.19: Selección de valor K para el algoritmo de clasificación suciedad y suciedad.....	92
Figura 5.20: Matriz de confusión Algoritmo de Vecinos Cercanos en detección de suciedad y sombreado.	93
Figura 5.21: Matriz de confusión normalizada Red Neuronal detección de Sombreado y Suciedad	97
Figura 5.22: Matriz de confusión normalizada Algoritmo KNN otoño en validación resultados de sombreado	100
Figura 5.23: Matriz de confusión normalizada Algoritmo KNN Otoño validación de resultados suciedad	102
Figura 5.24: Matriz de confusión normalizada KNN Otoño validación de resultados sombreado y suciedad	103
Figura 5.25: Matriz de confusión normalizada RN otoño validación resultados sombreado y suciedad Fuente: Elaboración Propia	105
Figura 0.1 Comparación Corriente Continua base de datos de suciedad,	117
Figura 0.2 Comparación Voltaje en CC base de datos de suciedad,	117
Figura 0.3 Comparación Potencia en CC base de datos de suciedad,	118
Figura 0.4 Comparación Temperatura del panel FV base de datos de suciedad,.....	118
Figura 0.5 Comparación Temperatura Ambiente base de datos de suciedad,.....	119
Figura 0.6 Comparación Velocidad del Viento base de datos de suciedad,.....	119
Figura 0.7 Comparación Radiación base de datos de suciedad,.....	120
Figura 0.8: Grafica de correlación de variables datos de sombreado (Primavera).....	121
Figura 0.9: Grafica de correlación de variables datos de suciedad (Verano).....	121

Figura 0.10: Grafica de correlación de variables datos de sombreado y suciedad (Primavera -
Verano)..... 122

Abreviaciones

Mayúsculas

AC	: Siglas en ingles de Corriente Alterna (Alternating Current)
CEN	: Cordinador Electrico Nacional
CNE	: Comision Nacional de Energia
CC	: Corriente continua
DC	: Siglas en ingles de Corriente Continua (Continuous Current)
ERNC	: Energias Renovables No Convencionales
FV	: Fotovoltaico
FN	: Falso Negativo
FP	: Falso Positivo
FDA	: Funcion de Activacion
KNN	: Siglas en ingles del algoritmo k vecinos cercanos (K- nearest neighbors)
RN	: Redes Neuronales
T	: Temperatura
VP	: Verdadero Positivo
VN	: Verdadero Negativo

Capítulo 1. Introducción

1.1. Introducción General

En los últimos años, se ha observado un crecimiento significativo, tanto a nivel mundial como nacional, en el uso de energías renovables no convencionales, tales como la hidroeléctrica, eólica, geotérmica y solar fotovoltaica. Este aumento se debe principalmente a la preocupación por el cambio climático y la necesidad de reducir el impacto ambiental asociado al consumo de energía, tanto en el ámbito doméstico como en el industrial. Entre las energías renovables, la solar fotovoltaica ha adquirido gran relevancia debido a su capacidad para aprovechar un recurso inagotable: la radiación solar.

En la actualidad, en el sistema eléctrico nacional chileno, la energía solar fotovoltaica representa aproximadamente el 23.6% de la potencia total instalada, con 7.714,77 [MW], según lo informado en el reporte mensual del sector energético de julio de 2023 por el CEN.

El Machine Learning ha adquirido gran relevancia en diversos sectores industriales y, en particular, en el sector energético. Se ha documentado cómo su implementación puede mejorar la eficiencia y la producción en diferentes áreas. En el caso de la tecnología solar fotovoltaica, su aplicación ha ido en aumento, y el objetivo principal es maximizar la producción de energía.

En trabajos anteriores, se demostró la posibilidad de crear un modelo predictivo de energía AC y DC para un conjunto de paneles fotovoltaicos en la MicroRed UCSC utilizando herramientas basadas en Machine Learning. En esta investigación, se busca ampliar ese trabajo previo y detectar posibles fallas en el sistema de la MicroRed. Las fallas más comunes que afectan a los paneles FV son el sombreado y la suciedad. Estas condiciones en los paneles FV impiden que la radiación solar llegue de manera óptima al módulo, lo que conlleva a una disminución en la producción de energía. Además, si estas condiciones perduran, pueden llegar a producir daños permanentes en los paneles, afectando la vida útil del proyecto. Esto hace necesario realizar un mantenimiento constante e idealmente monitorizar el comportamiento del sistema solar FV.

El enfoque de este estudio se centra en utilizar algoritmos de Machine Learning para identificar el sombreado y la suciedad en la MicroRed UCSC, lo que permitirá tomar medidas correctivas de manera oportuna, evitando daños mayores en el sistema disminuyendo las pérdidas y optimizando la producción de energía.

1.2. Trabajos Previos

Funcionamiento de sistemas fotovoltaicos ante condiciones de suciedad y sombreado y su detección con variables eléctricas

- ♣ Cordero, R. R., Damiani, A., Laroze, D., MacDonell, S., Jorquera, J., Sepúlveda, E., Feron, S., Llanillo, P., Labbe, F., Carrasco, J., Ferrer, J., & Torres, G. (2018). “Effects of soiling on photovoltaic (PV) modules in the Atacama Desert”. *Scientific Reports*, vol. 8, no 13943.

En este trabajo evalúan como afecta la suciedad al rendimiento de los paneles fotovoltaicos en seis ciudades diferentes de Chile (Arica, Iquique, Calama, Copiapó, La Serena y Santiago) evaluando como afecta su ubicación geográfica y analizando sus pérdidas tanto en potencia y el valor económico, además de realizar una discusión de cómo afectan diferentes factores tales como: la velocidad del viento, la humedad de la zona, aerosoles naturales, el ángulo de inclinación de los paneles, los efectos de las precipitaciones y la frecuencia con la que se limpian los paneles FV.

- ♣ Kim, J., Rabelo, M., Padi, S., Yousuf, H., Cho, E.-C., Yi, J. (2021) “A Review of the Degradation of Photovoltaic Modules for Life Expectancy” *Energies 2021*, vol. 14, no 4278.

En esta investigación establecen que para saber la fiabilidad de un módulo solar fotovoltaico se debe esperar 25 años lo cual es indeseable, por ende, realizan pruebas de estrés acelerado en laboratorio para determinar la fiabilidad de estos. Explicando las diferentes fallas que por lo general presentan los paneles, si no se les realiza una mantención u control constante. Además, identificando como las diversas condiciones ambientales afectan a los módulos.

- ♣ Al Abri, W., Al Abri, R., Yousef, H., & Al-Hinai, A. (2021). “A Simple Method for Detecting Partial Shading in PV Systems”. *Energies 2021*, vol. 14, no 16.

En este trabajo crean un sistema simple para detectar condiciones de sombreado parcial (PSC) ya que estas pueden reducir la energía total generada haciendo que el sistema no funcione a su máximo rendimiento, para crear este sistema de detección de sombreado parcial mediante la detección del cambio de signo de la derivada del voltaje. El método predice un sombreado parcial si el signo del

voltaje es el mismo durante un cierto número de cambios sucesivos debido al sombreado. Probando la efectividad de este método en diferentes configuraciones de paneles serie y serie – paralelo.

- ♣ Pei, T., & Hao, X. (2019). “A Fault Detection Method for Photovoltaic Systems Based on Voltage and Current Observation and Evaluation”. *Energies* 2019, vol. 12, no. 9.

En este trabajo establecen un método de detección de fallas para los sistemas FV basado en la observación y evaluación tanto del voltaje como la corriente, para detectar los fallos más comunes de los sistemas fotovoltaicos (falla de cortocircuito, circuito abierto, degradación y sombreado). Para ello utilizan valores de voltaje y corriente característicos de las fallas como valores umbrales y al compararlos con los valores de operación se pueden identificar las posibles fallas en el sistema.

- ♣ Khonkar, H., Alyahya, A., Aljuwaied, M., Halawani, M., Al Saferan, A., Al-khaldi, F., Alhadlaq, F., & Wacaser, B. A. (2014). "Importance of cleaning concentrated photovoltaic arrays in a desert environment." *Solar Energy*, vol. 110, pp. 268-275.

En esta investigación demuestran algunos efectos que produce la suciedad en los paneles fotovoltaicos y en los ultra high concentration PV, haciendo la comparación cuando estos están limpios y cuando presentan suciedad, para ver el contraste entre ambos casos y demostrar la importancia de realizar un mantenimiento periódico de estos. Esto lo realizan mediante la comparación de las curvas I-V, que precisamente son en las cuales uno ve el desgaste de los módulos fotovoltaicos.

Aplicación de la Ciencia de Datos en la energía Solar FV

- ♣ Mekki, H., Mellit, A., & Salhi, H. (2016). "Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules." *Simulation Modelling Practice and Theory*, vol. 67, pp. 1-13.

Esta investigación presenta un método de detección de fallos para módulos FV en condiciones de sombreado parcial utilizando una Red Neuronal Artificial (RNA) para la estimación de las variables eléctrica de corriente y voltaje de salida. Al comparar los valores simulados con los datos reales es capaz de proporcionar información sobre panel FV presenta fallas de sombreado parcial.

- ♣ Theocharides, S., Makrides, G., Kyprianou, A., & Georghiou, G. E. (2018). "Machine Learning Algorithms for Photovoltaic System Power Output Prediction." *IEEE International Energy Conference (ENERGYCON)*.

Este trabajo realiza una comparativa de diferentes algoritmos de machine learning que predicen la energía producida de un sistema FV compuesto por cinco paneles policristalinos de una potencia instalada de 1175 Wp. Los algoritmos utilizados son una Red Neuronal Artificial (RNA), Super Vector Machine (SVMs) Árbol de Regresión (Regression Trees) y los compararon con un modelo de persistencia (PM). Para realizar la comparativa utilizaron diferentes métricas, tales como Error absoluto medio (MAE), Error absoluto medio Porcentual (MAPE), Error cuadrático medio (RMSE), Error cuadrático medio normalizado (Nrmse) y Skill Score (SS). Siendo el algoritmo con mejores rendimientos la Red Neuronal.

- ♣ Laarabi, B., May Tzuc, O., Dahlioui, D., Bassam, A., & Barhdadi, A. (2019). "Artificial neural network modeling and sensitivity analysis for soiling effects on photovoltaic panels in Morocco." *Superlattices and Microstructures*, vol. 127, pp. 139-150.

En este trabajo se presenta una metodología de estudio de la modelización del efecto de la suciedad en los módulos de paneles FV utilizando redes neuronales artificiales (RNA). Para llevar a cabo esta investigación se dejó en el exterior un vidrio solar FV al cual se le realizaron mediciones periódicas para controlar la tasa de ensuciamiento en el tiempo. Donde a través de un estudio de sensibilidad se determinaron los efectos de los parámetros de entrada en la salida del modelo, las variables de entrada más influyentes en la tasa de ensuciamiento del módulo fueron la humedad relativa y la dirección del viento.

- ♣ Fuchang, Zhicong, Long, Jinling, Peijie, & Shuying Cheng. (2019). "An intelligent fault diagnosis method for PV arrays based on an improved rotation forest algorithm." *Energy Procedia*, vol. 158, pp. 613-618.

En este trabajo de investigación trabajan con algoritmos de Rotation Forest (RoF) y Extreme learning Machine (ELM). Fusionando ambos algoritmos crean (RoF-ELM) y utilizan estos para clasificar fallas en los sistemas FV. La parte experimental la llevan a cabo en un laboratorio a pequeña

escala un conjunto pequeño de paneles fotovoltaicos de 2kW en el cual generan fallas de cortocircuito, circuito abierto, sombreado y deterioro de estos. Mediante la utilización de los algoritmos mencionados anteriormente y son capaces de clasificar las distintas fallas, donde comparan el rendimiento entre los algoritmos siendo el con mejor rendimiento el algoritmo (RoF-ELM) llegando a tener un 98.63 % de precisión.

- ♣ Gosumbonggot, J., & Fujita, G. (2019). "Partial Shading Detection and Global Maximum Power Point Tracking Algorithm for Photovoltaic with the Variation of Irradiation and Temperature." *Energies*, vol. 12, no. 2.

En esta investigación formulan un nuevo sistema de seguimiento Máximo Power Point Tracking (MPPT) puesto que los métodos convencionales no distinguen la potencia máxima global de la curva P-V en condiciones de sombreado. Por ende, proponen un método de detección de sombreado con las curvas P-V usando diferentes configuraciones en la conexión de los paneles. Con este nuevo sistema MPPT son capaces de maximizar la energía obtenida por los paneles fotovoltaicos aumentando la potencia generada en un 8,55 %.

- ♣ Heinrich, M., Meunier, S., Samé, A., Quéval, L., Darga, A., Oukhellou, L., & Multon, B. (2020). "Detection of cleaning interventions on photovoltaic modules with machine Learning." *Applied Energy*, vol. 263, no. 114642.

En este trabajo plantean la opción de tener un sistema de monitoreo de la limpieza de los paneles fotovoltaicos de bajo costo para sistemas fotovoltaicos autónomos en localidades rurales. Para ello, utilizan tres módulos fotovoltaicos con una potencia total instalada de 620 Wp. La finalidad de esta investigación es identificar si es necesario realizar trabajos de limpieza adicionales esto se plantea el problema como uno de clasificación para saber si la suciedad se elimina con regularidad. Para ello utilizan herramientas de Machine Learning basadas en Regresión Logística (RL), Red Neuronal Artificial (RNA), Super Vector Machine (SVMs) y Ramdon Forest (RF) analizando variables tales como: temperatura ambiente, temperatura de las celdas FV, voltaje y corriente del sistema FV para la identificar cuando es necesario realizar limpieza en los paneles.

- ♣ Zitouni, H., Azouzoute, A., Hajjaj, C., El Ydrissi, M., Regragui, M., Polo, J., Oufadel, A., Bouaichi, A., & Ghennioui, A. (2021). "Experimental investigation and modeling of photovoltaic soiling loss as a function of environmental variables: case study of semi-arid climate." *Solar Energy Materials and Solar Cells*, vol. 221, no.110874

Este trabajo de investigación se genera un modelo para determinar la tasa de suciedad en los paneles FV en función de variables ambientales en un clima que cuenta con un periodo seco y otro lluvioso usando modelos de regresión lineal múltiple (RLM), regresión lineal múltiple con modelo de interacción (RLMWI) y redes neuronales artificiales (RNA) realizando mediciones durante un año en una planta solar FV de 2,16 kW de tecnología Amorfa o Capa fina. Con estos modelos son capaces de determinar cuanta energía no se produce por causa de la suciedad presente en los módulos FV.

- ♣ Hanafy, W. (2019). "Machine Learning Approach for Photovoltaic Panels Cleanliness Detection." 2019 15th International Computer Engineering Conference (ICENCO).

Este trabajo presenta un método para determinar el porcentaje de suciedad que tiene un módulo fotovoltaico mediante la utilización de machine learning y el procesamiento de imágenes. Esto lo hacen para identificar cuanto porcentaje del panel fotovoltaico presentan suciedad. identificando el color de los pixeles de las fotografías usando distintas imágenes de paneles FV con porcentajes de suciedad distintos, utilizando diferentes inclinaciones y posicionamiento de las cámaras para ver las diferencias entre estos casos.

- ♣ Starzyński, J., Zawadzki, P., & Harańczyk, D. (2022). "Machine Learning in Solar Plants Inspection Automation." *Energies*, vol. 15, no. 16, art. 59661.

Este trabajo se plantea que debido al crecimiento que tendrá la energía producida por plantas solares fotovoltaicas es necesario poder tener un sistema de vigilancia de bajo costo, para asegurar un buen funcionamiento de la planta de generación. Realizando un análisis cuantitativo del efecto que tienen tener fallas en los paneles, falta de mantenimiento en la producción total de energía. Proponiendo un método de vigilancia que combina un vehículo aéreo no tripulado (UAV), imágenes térmicas y aprendizaje automático. Usando redes neuronales profundas para el análisis de las

fotografías térmicas usando la red YOLO (You Only Look Once) que es una red Neuronal la cual se utiliza para la detección y localización de defectos en los módulos.

Investigaciones previas realizadas en la MicroRed UCSC

- ♣ Alderete, L. (2019). “Diseño e implementación de experiencias en redes On – Grid en laboratorio Microred UCSC [Título profesional Universidad Católica de la Santísima Concepción].”

En este trabajo muestra la construcción y puesta en marcha de la MicroRed UCSC. Realizando una recopilación de instalaciones FV de otros recintos educativos y luego especificando los distintos equipos y plataformas que componen la MicroRed UCSC, para finalmente proponer distintos experimentos en la MicroRed tanto On-grid como Off-grid.

- ♣ Rebolledo, F. (2021). “Análisis de desempeño en Generadores FV Microred UCSC [Título profesional Universidad Católica de la Santísima Concepción].”

En este trabajo se evalúan distintas instalaciones fotovoltaicas tanto instalaciones fijas como con seguimiento solar. Para luego analizar estos mismos tipos de instalaciones, pero presentes en la MicroRed UCSC utilizando métricas de rendimiento estandarizadas para sistemas fotovoltaicos.

- ♣ Rebolledo, F., Mendoza-Araya, P., Carvajal, G., & Ramírez, G. (2022). "Performance evaluation of different solar modules and mounting structures on an on-grid photovoltaic system in south-central Chile." *Energy for Sustainable Development*, vol. 68, pp. 65-75.

En esta investigación tiene un enfoque en demostrar cómo funcionan los distintos tipos de tecnologías de energía solar fotovoltaica. Dando a conocer qué tipo de sistema de paneles funciona mejor ya sea un sistema fijo o de seguimiento teniendo en cuentas distintos tipos tecnología de paneles fotovoltaico (monocristalinos, policristalinos o capa fina) dependiendo determinados climas o condiciones ambientales. Para esto utilizan métricas de rendimientos estandarizadas para sistemas fotovoltaicos.

- ♣ Ortiz, B. (2022) “Modelo Predictivo de Potencia para Inversor On-grid - MicroRed UCSC [Proyecto de iniciación de la ingeniería Eléctrica]”

En este trabajo se dan a conocer los conceptos básicos del Machine Learning, para luego utilizar modelos de regresión lineal para predecir cuanta potencia en corriente continua y corriente alterna debería generar un determinado Inversor On-grid la MicroRed UCSC. Además de aplicar métricas de rendimiento a los resultados simulados para comparar entre la potencia generada simulada y el comportamiento real.

Discusión

En los últimos años, la ciencia de datos ha ganado popularidad como una herramienta de investigación utilizada por muchas personas y profesionales. En el campo de las plantas de generación de paneles fotovoltaicos, esta tendencia se ha mantenido, ya que se reconoce la importancia de analizar y detectar problemas en estas instalaciones. Entre los problemas más comunes se encuentran el sombreado parcial y la suciedad [1].

Los documentos mencionados previamente hacen hincapié en la necesidad de desarrollar sistemas capaces de detectar estos problemas. Para ello, se han propuesto diversos métodos de detección que emplean diferentes algoritmos, como Redes Neuronales Artificiales (RNA), Máquinas de Vectores de Soporte (SVM), Árboles de Regresión, Rotation Forest (RoF), Extreme Learning Machine (ELM), Regresión Logística (RL) y Random Forest (RF), así como la regresión lineal múltiple (RLM). [2], [3], [4], [5],[6].

En general, los estudios mencionados utilizan laboratorios compuestos por una cantidad reducida de paneles fotovoltaicos. Estos laboratorios incorporan distintas variables de entrada en sus sistemas, como imágenes, fotografías térmicas, variables eléctricas y variables ambientales. Estas variables permiten obtener información detallada para el análisis y la detección de problemas en las plantas de generación fotovoltaicas.

El sombreado parcial es uno de los problemas que afecta significativamente la producción de energía en estas instalaciones. Por otro lado, es fundamental comprender el impacto de la suciedad en los paneles, ya que esto permitiría establecer un cronograma de mantenimiento adecuado para optimizar los recursos utilizados en el cuidado de la MicroRed.

Con la incorporación de una estación meteorológica en la MicroRed UCSC, se pueden obtener datos precisos sobre la radiación solar, velocidad del viento, humedad, temperatura ambiente y otras variables ambientales relevantes. Estos datos brindan la oportunidad de llevar a cabo un estudio preciso sobre cómo el sombreado parcial y la suciedad afectan a la MicroRed. Para ello, se pueden emplear diferentes algoritmos de Machine Learning, los cuales permiten analizar y extraer patrones

de los datos recopilados para identificar y predecir el impacto de estos problemas en la eficiencia y el rendimiento de la MicroRed.

La utilización de algoritmos de Machine Learning en este contexto proporciona una herramienta eficaz para el análisis de datos y la detección de anomalías en los paneles solares, lo que a su vez facilita la toma de decisiones informadas en cuanto al mantenimiento y optimización de la MicroRed UCSC.

1.3. Hipótesis de Trabajo

El trabajo consiste en utilizar algoritmos de Machine Learning para identificar casos anómalos de sombreado y suciedad, por ende:

- Es posible identificar las diferencias entre suciedad y sombreado.
- Los algoritmos son capaces detectar y diferenciar entre ambos casos.

1.4. Objetivos

Objetivo General

Aplicar técnicas de machine learning para la detección de suciedad y/o sombreado en sistemas FV de MicroRed UCSC, a partir de las variables disponibles en el sistema de monitoreo y registro existente.

Objetivos Específicos

- Estudiar componentes y funcionamiento de los sistemas FV de la MicroRed UCSC.
- Estudiar e identificar las principales fallas que ocurren en los paneles fotovoltaicos.
- Estudio y aplicación de técnicas de ciencia de datos para la identificación de condiciones de suciedad y sombreado.
- Simular condiciones de sombreado y suciedad en sistemas FV de MicroRed UCSC, para la evaluación de las técnicas de identificación de ciencia de datos.
- Evaluar la efectividad de los modelos creados para detección de fallas, considerando las distintas condiciones experimentales realizadas.

1.5. Alcances y Limitaciones

Para este proyecto, se emplearán tecnologías de paneles fotovoltaicos monocristalinos y policristalinos con una inclinación fija, los cuales se encuentran disponibles en la MicroRed UCSC.

Las técnicas que se esperan utilizar para la identificación de estas fallas son modelos de machine learning de aprendizaje supervisado específicamente de clasificación tales como; regresión logística, vecinos cercanos y redes neuronales.

1.6. Temario y Metodología

El Capítulo 2 de la tesis describe la composición de los paneles fotovoltaicos y enumera las fallas más comunes a las que están expuestos. Se proporciona información detallada sobre la estructura y los componentes de los paneles, así como las posibles fuentes de fallos, como el desgaste, la corrosión, los problemas de cableado y los defectos en las células solares.

En el Capítulo 3 se presentan las pruebas experimentales realizadas en la MicroRed UCSC, abarcando tanto pruebas de suciedad como de limpieza. Se describen los procedimientos llevados a cabo para simular condiciones de suciedad en los paneles y se detallan los resultados obtenidos.

Los Capítulos 4 y 5 se centran en la explicación de los diferentes algoritmos que se emplearán en el estudio. Se comienza con modelos básicos, como la regresión lineal, y se avanza hacia métodos más avanzados, como las redes neuronales. Se proporciona información sobre la teoría y el funcionamiento de cada algoritmo, y se detalla su implementación en el contexto específico de este trabajo.

Finalmente, el Capítulo 6 corresponde a las conclusiones del estudio realizado. Se presentan los hallazgos más relevantes, se discuten los resultados obtenidos y se formulan recomendaciones basadas en los resultados de la investigación. Además, se pueden incluir reflexiones sobre posibles áreas de mejora y futuras investigaciones relacionadas con el tema.

Por último, los Anexos contienen información adicional que complementa y respalda los contenidos de la tesis. El Anexo 1 muestra la comparativa de las gráficas de la base de datos de suciedad antes y post procesado de datos. El Anexo 2 códigos de programación básicos de los algoritmos utilizados para tener un punto de referencia de la programación utilizada en esta investigación.

Capítulo 2. Identificación de Fallas en Sistemas Fotovoltaicos

2.1. Introducción

En esta parte del trabajo se explicaran diferentes fallas que pueden ocurrir en los sistemas fotovoltaicos y tambien los que se producen en los paneles fotovoltaicos, haciendo énfasis en los daños que estos pueden generar y que medidas se pueden tomar para evitar estas fallas.

2.2. Fallas en los Sistemas Fotovoltaicos

Las fallas en los sistemas fotovoltaicos, se pueden producir debido a un mal funcionamiento o daño de un componente interno del sistema, como por alguna condición externa al sistema en sí, que afecte el desempeño del sistema y/o produzca el deterioro de este. Los sistemas FV están formados por distintos componentes tales como:

- Paneles solares: Son el componente esencial de un sistema FV estos son los que, mediante el efecto fotoeléctrico, producen energía eléctrica en DC a partir de la irradiancia solar. En la actualidad las principales tecnologías de estos son: monocristalinos, policristalinos y capa fina o amorfos.
- Inversor: Este equipo eléctrico transforma la energía DC generada por los paneles FV en corriente alterna (AC), siendo fundamentales para poder distribuir la energía del sistema. Estos equipos pueden ser monofásicos o trifásicos dependiendo de la necesidad del sistema.
- Cableado: Estos son utilizados para el transporte de la energía generada de los paneles FV hasta el inversor y luego desde este hacia la red eléctrica. Cabe resaltar que en cada uno de estos tramos el tipo de cable y la sección de este dependerá de la corriente que transporte y el lugar donde se utilice.
- Sistema de protecciones: Estos componentes son los que mantienen la seguridad tanto de la instalación solar, como de las personas que transitan y/o trabajan en las instalaciones.
- Baterías o sistema de almacenamiento: Estos componentes son opcionales dependiendo que tipo de sistema solar fotovoltaico se tenga, están presentes en sistemas Off-grid y en los sistemas híbrido. Son utilizadas para almacenar la energía producida, para utilizarla cuando se requiera.
- Estructura de montaje: Son utilizados para instalar los paneles solares en ellos. Estos pueden ser de seguimiento fijos, los de seguimiento pueden ser de 1 o 2 ejes y de manera automática

se mueven buscando la mayor producción de energía. Mientras que los fijos, se puede ajustar su ángulo de inclinación, el cual se elige dependiendo el lugar del planeta donde se instalen los módulos FV.

2.2.1 Fallas internas del Sistema

El sistema fotovoltaico este compuesto por distintos equipos y materiales eléctricos, específicamente la Microred UCSC tiene: Paneles Fotovoltaicos, Inversores, sistema de protecciones eléctricas, cables y baterías. Por ende, alguna falla en alguno de estos equipos afectaría directamente el funcionamiento y la producción de energía de la MicroRed.

- Paneles fotovoltaicos: Las fallas de los paneles fotovoltaicos pueden ser físicas, medioambientales (dependiendo del entorno) y eléctricas. Estas se detallarán más adelante.
- Inversor: La falla de un inversor afectaría directamente la producción de energía del sistema ya que la energía generada por los paneles fotovoltaicos (DC) no se podría inyectar a la red, debido a que no se realizaría la conversión a (AC), además puede producir aumentos de temperatura en el sistema, afectando a los paneles, cables y baterías conectadas al inversor.
- Cables: La falla de uno o más cables afecta directamente la producción de energía debido a que estos son los que transportan la energía producida por los paneles fotovoltaicos hacia el inversor y desde el inversor a la barra donde se inyecta la energía a la red, además es una falla de las más peligrosas ya que puede producirse un cortocircuito, una descarga eléctrica e incluso un incendio.
- Sistema de Protecciones: una falla en el sistema de protecciones, puede producir la no desconexión del sistema frente a una falla de sobre tensión o de sobre corriente el cual puede afectar a los módulos fotovoltaicos con un aumento de temperatura pudiéndose producir incendios y poniendo en riesgo la vida de los operadores del sistema.
- Baterías: presentes en sistema Off-grid o Híbrido la falla de una batería puede afectar el rendimiento del sistema fotovoltaico, ya que afectaría directamente la energía que pueden entregar cuando los paneles fotovoltaicos salen de operación, además puede afectar a otras baterías que estén conectadas entre sí, afectando su capacidad de almacenamiento como a su tiempo de carga y descarga. Provocando fallas de inestabilidad en el sistema.

Para evitar este tipo de fallas, en primer lugar, se debe realizar un buen dimensionamiento de los diferentes elementos del sistema y se debe realizar un mantenimiento regular, revisar si se tiene un correcto funcionamiento de los diferentes elementos del sistema, corroborar que las conexiones estén

en buen estado, además reemplazar o reparar lo antes posible un elemento que este en mal estado, para mitigar las pérdidas del sistema fotovoltaico.

2.2.2 Fallas en los módulos Fotovoltaicos

Como se indicó anteriormente los paneles fotovoltaicos pueden fallar por diferentes motivos, tales como; daños físicos, fallas por causas del ambiente como es la suciedad y el sombreado, además de fallas eléctricas que se pueden producir en estos.

A. *Fallas Físicas*

Para entender cómo se producen las fallas físicas hay que saber cómo está construido un panel solar fotovoltaico.

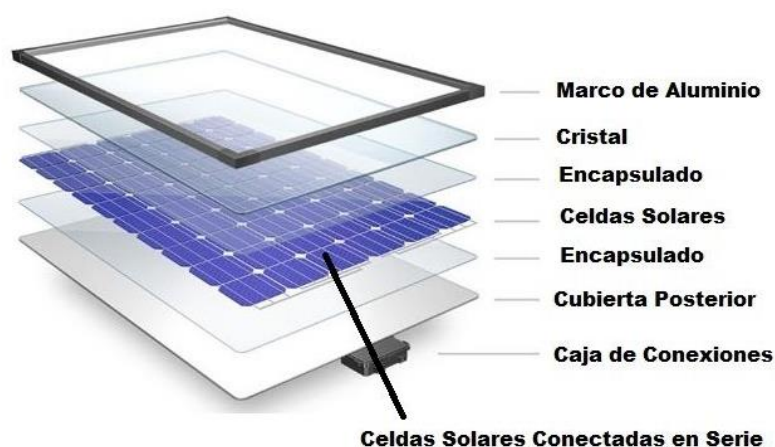


Figura 2.1 Estructura interna de un panel fotovoltaico

Fuente: {1} Web (Bibliografía)

En la figura 2.1 se muestra la estructura interna de un panel fotovoltaico el cual consta de varias capas. El marco de aluminio, la cubierta posterior y el cristal tienen como finalidad proteger a las celdas solares, en conjunto con las capas de encapsulado. Estas reducen o mitigan los daños producidos por el entorno u ambientales a las celdas. Por otro lado, la caja de conexiones contiene diodos de derivación también conocidos como diodos By-pass. Estos se pueden observar en la figura 2.2 donde también se aprecian diodos y las salidas positiva y negativa del panel.

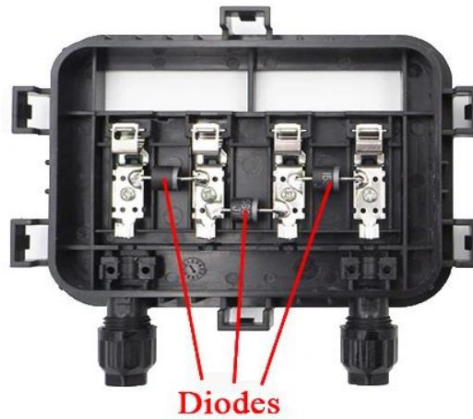


Figura 2.2 Interior de caja de conexiones panel fotovoltaico

Fuente: {2} Web (Bibliografía)

Los diodos By-pass tienen la función de aislar a las celdas que no estén produciendo energía lo que se puede producir con condiciones de sombreado parcial o suciedad, aislando de esta manera a la celda. Lo explicado anteriormente sucedería si cada celda en particular tuviera un diodo By-pass conectado a ella, pero eso aumenta el costo de producción de los paneles. La solución a esto es conectar varios grupos de celdas a un diodo By-pass. Como indica la figura 2.3. de esta manera se disminuye el costo del módulo, pero también frente a una eventual condición de las anteriormente mencionadas la pérdida de generación de energía será mayor.

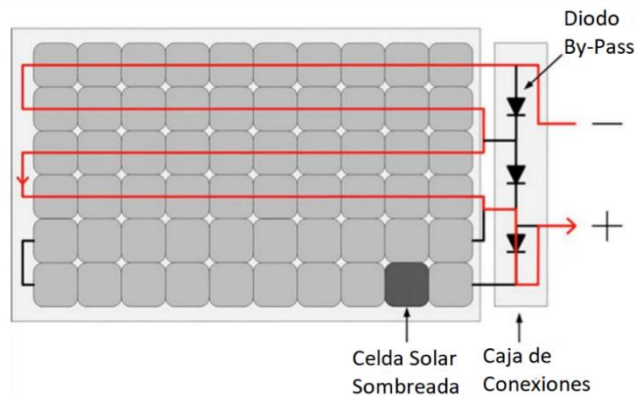


Figura 2.3: Comportamiento de los Diodos By-Pass

Fuente: {3} Web (Bibliografía)

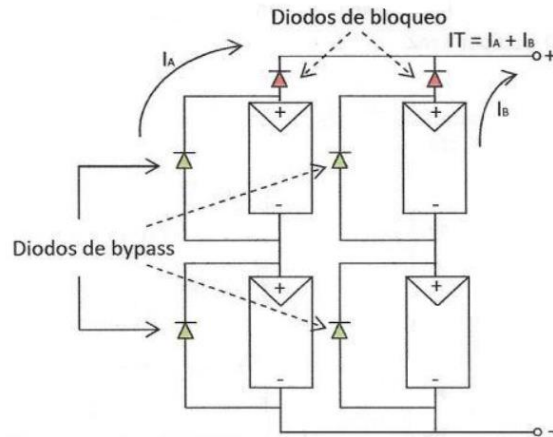


Figura 2.4: Configuración de Diodos By-Pass y Diodo de Bloqueo

Fuente: {4} Web (Bibliografía)

En la figura 2.4 se muestra como sería la conexión ideal de los diodos en los módulos fotovoltaicos teniendo un diodo By-Pass conectado en paralelo a cada celda fotovoltaica y un diodo de bloqueo conectado en serie a un conjunto de celdas fotovoltaicas.

Los diodos de bloqueo cumplen la función de proteger a los paneles de corrientes inversas, que circulan en el sistema cuando este no percibe radiación o cuando la luz es baja evitando de esta manera que aumente la temperatura en los módulos fotovoltaicos lo que puede producir daños permanentes en el sistema.

- **Puntos Calientes (Hot Spot):** esta falla se produce cuando se produce un aumento de temperatura en las celdas fotovoltaicas. Esto puede llevar a fallos permanentes en las celdas disminuyendo su eficiencia y vida útil. [7]
- **Falla en los diodos By-pass y diodo de Bloqueo:** estas fallas se pueden producir por un aumento de temperatura, condiciones ambientales ya que estos se pueden sulfurar si hay mucha humedad e incluso romperse. Esto puede hacer que un panel disminuya su producción de energía a cero y dependiendo como esté conectado este con otros módulos hacer que un string no genere energía eléctrica. Por otra parte, si falla el diodo de bloqueo, además de lo anteriormente mencionado disminuirá la protección a corrientes que lleguen al panel aumentando los riesgos de un cortocircuito o de una falla de circuito abierto. [8]
- **Daños en el Panel Fotovoltaico:** estas fallas pueden producirse por alguna condición ambiental, como por una mala instalación y/o un posible error de fabricación. Estamos

hablando de una ruptura del cristal, delaminación, burbujas de agua debido a un deterioro del encapsulado de las celdas FV.

B. Fallas por condiciones ambientales

Este tipo de fallos que se pueden producir dependen mucho del entorno en el cual se instale ya sea una planta solar fotovoltaica o sistemas mas pequeños de esta tecnología de producción de energía. Ya que dependiendo de esto los paneles FV pueden estar expuesto a distintos tipos de suciedad o de sombreado. [9]

- Suciedad: esta puede ser polvo, nieve y suciedad propiamente tal como desechos de animales, plástico, etc. Estas condiciones reducen la superficie donde incide la radiación disminuyendo la energía producida y si esta condición persiste disminuyendo la vida útil y eficiencia del panel fotovoltaico.
- Sombreado: esta condición puede ser parcial o permanente, que pueden llegar a producir los mismos fallos en los módulos fotovoltaicos, pero en menor o mayor superficie. Además, disminuye de manera significativa la cantidad de energía producida ya que los diodos By-pass asilan las celdas que están sombreadas.

2.2.3 Fallas Eléctricas

Este tipo de fallas principalmente ocurren debido a algún tipo de imperfección de los materiales eléctricos o debido al desgaste de los equipos, así también producto de la falta de mantención de los equipos.

- Circuito Abierto: esta falla se produce cuando una parte del circuito de esta se desconecta de manera abrupta dejando de funcionar. Disminuyendo la potencia generada ya que se reduce de manera significativa la corriente.
- Corto Circuito: esta se produce cuando entran en contacto conexiones de polaridades diferentes que están separadas, generando un aumento en la corriente y temperatura lo que puede ocasionar daños graves en el sistema.
- Línea a Línea: esta falla ocurre cuando dos conductores con diferente potencial entran en contacto generando un cortocircuito accidental esto puede ser permanente o momentáneo.

A continuación, se muestra una tabla de resumen de las diferentes fallas, que parte del sistema y/o panel fotovoltaico afecta.

Tabla 2.1: Resumen de fallas en paneles FV y sus consecuencias

Tipos de falla	Parte afectada	Causas	Consecuencias
Puntos Calientes	- Celdas Fotovoltaicas	- Suciedad.	- Daño a las celdas fotovoltaicas
	- Panel Fotovoltaico	- Sombreado. - Conexión entre módulos de diferentes parámetros.	- Reducción de la eficiencia del módulo.
Fallas en los Diodos	- Diodos By-Pass	- Sombreado	- Daño en los diodos
	- Diodo de Bloqueo	- Aumento de temperatura - Corrosión	- Corto circuito y circuito Abierto - Daños en los módulos y riesgo de incendios - Disminución de rendimiento y eficacia.
Falla en las celdas y Paneles Fotovoltaicos	Paneles Fotovoltaicos y sus celdas	- Ruptura del Cristal	
		- Errores de fabricación	
		- Daños en la instalación	- Disminución de eficiencia y rendimiento de los módulos
		- Daños en el encapsulado de las celdas	- Reducción de la potencia generada
		- Condiciones Ambientales	
Falla Línea a Línea	Conjunto de Paneles	- Contactos accidentales de Cables	- Daño en los conductores
		- Deterioro del cableado	- Daño al panel Fotovoltaico - Riesgo de Incendio

Fuente: Elaboración Propia

Respecto a lo expuesto en la tabla 2.1, la mayoría de las fallas se producen debido al sombreado y a la suciedad. Por lo que resulta fundamental contar con un sistema de monitoreo de la placa fotovoltaica capaz de identificar si una de estas condiciones está presente en el funcionamiento diario del sistema, para evitar daños en los módulos y obtener una mayor producción de energía y por ende una mayor rentabilidad del proyecto.

2.2.4 Fallas externas al sistema

Las fallas externas al sistema fotovoltaico son fallas debido a factores que pueden afectar el funcionamiento de los equipos eléctricos, es decir son condiciones que se dan por el entorno en el que está el sistema fotovoltaico que afectan el rendimiento del sistema.

- Orientación o inclinación no optimas del sistema fotovoltaico: Elegir mal la orientación de los paneles fotovoltaicos y el ángulo de inclinación afecta directamente la producción de energía del sistema, ya que el tiempo que la irradiancia incidirá en la superficie de los paneles fotovoltaicos disminuirá de manera considerable.
- Sombreado: el sombreado puede ser parcial o permanente afectando directamente la corriente de salida del panel y del arreglo fotovoltaico en general dependiendo si la conexión de los paneles es en serie o paralelo, ya que al estar en serie la corriente máxima del conjunto fotovoltaico será la mínima del conjunto (Panel Sombreado).
- Acumulación de Suciedad: si la superficie de un panel fotovoltaico tiene suciedad esta disminuye el área donde incide la irradiación disminuyendo la potencia generada por el sistema solar, pero la reducción de la energía generada es mucho menor en comparación al sombreado, pero puede llevar a daños permanentes en el módulo fotovoltaico.
- Altas temperaturas: los paneles fotovoltaicos están diseñados para operar en determinados rangos de temperatura, el incremento de esta tiene directa relación con el aumento de la corriente lo que conlleva a tener mayor cantidad de pérdidas en el sistema, puede afectar también al cableado del sistema aumentando la resistencia eléctrica de estos reduciendo la energía que son capaces de transmitir.
- Alta humedad: si el sistema solar está en una zona de alta humedad puede afectar directamente a los módulos fotovoltaicos reduciendo la eficiencia de estos, debido a la acumulación de agua en el sistema bloqueando la irradiancia que llega al arreglo fotovoltaico.
- Baja Irradiancia: si se tiene una irradiancia baja esto afecta directamente la producción de energía, pero también puede afectar la eficiencia de los módulos, dependiendo las condiciones para las que fue diseñado el panel fotovoltaico.

Estas fallas pueden evitarse eligiendo un lugar optimo que no tenga edificios o arboles cerca, con una temperatura ambiente dentro del rango de operación de los módulos, realizándole limpiezas a los paneles y eligiendo bien tanto el tipo de paneles a utilizar como su orientación e inclinación de estos.

Capítulo 3. Pruebas Experimentales de Sombreado y Suciedad en MicroRed UCSC

3.1. Introducción

En este capítulo se mostrarán las pruebas experimentales realizadas en la MicroRed UCSC realizadas en distintas estaciones del año. Siendo las pruebas de sombreados realizadas en primavera y las pruebas de suciedad durante el verano. El objetivo de esto es comprobar de manera experimental el comportamiento de las variables eléctricas (Voltaje, Corriente) frente a condiciones anómalas. Para monitorear las el comportamiento de las variables eléctricas se utilizó la plataforma “meteo-control” la cual muestra el comportamiento de estas con una tasa de muestreo de cinco minutos.

3.2. Equipos Eléctricos Utilizados

Para realizar las pruebas experimentales se utilizaron las dependencias de la Microred UCSC a continuación se describirán los equipos utilizados.

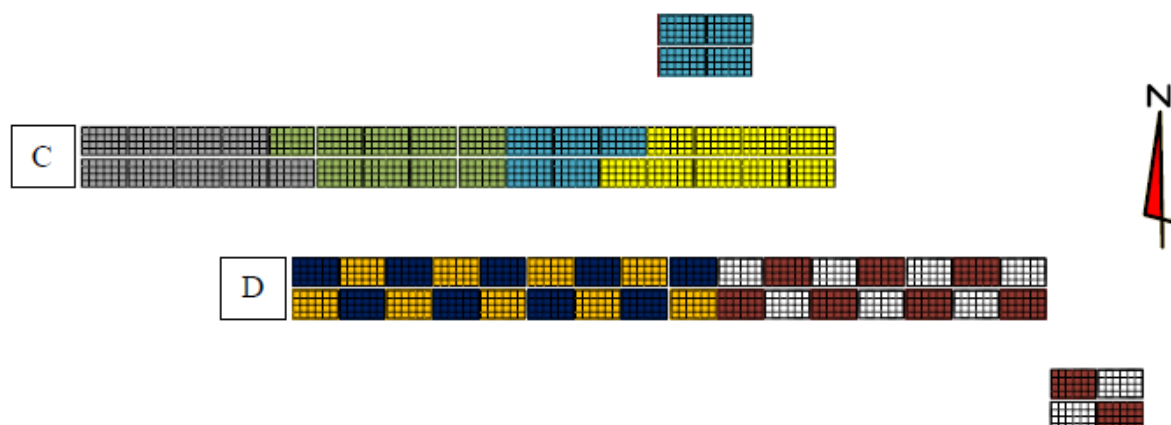


Figura 3.1: Distribución Física de Paneles FV por Strings

La sección C está compuesta por paneles policristalinos. La sección D es una combinación de paneles Monocristalinos y policristalinos

Fuente: Elaboración Propia

La figura 3.1 nos muestra la distribución de los paneles de la MicroRed UCSC que se utilizaron en las pruebas. Los Strings que se utilizarán para la construcción de la base de datos son la sección C y D a continuación se presenta una descripción de estos. [10]

Sección C

En esta sección se presentan 4 strings señalados por los colores gris, verde, celeste y amarillo. Estos 4 Strings se encontrarán conectados al inversor Fronius SYMO n°2 de 10 [kW]. Cada string está compuesto por 9 módulos fotovoltaicos policristalino conectados en serie; los Strings de color celeste y verde se conectan al MPPT1 de dicho inversor, y los strings amarillo y gris se conectan a la entrada MPPT2. Para esa configuración de conexión de los paneles el inversor tendrá condiciones similares en cada entrada MPPT, se tendrán los mismos parámetros, $V_{mp} = 285,3$ [V], $I_{mp} = 8,52$ [A] en cada string, la potencia en cada entrada MPPT será de 4,86 [kWp]. La potencia total instalada en este inversor será 9,72 [kWp].

Sección D

Esta sección tiene otros 4 strings señalados con los colores naranja, azul, blanco y color café. Estos strings están compuestos por series de 9 módulos fotovoltaicos en serie y se conectarán a inversor Fronius SYMO n°1 de 10 [kW], de manera que el MPPT1 recibirá la potencia generada de 2 strings policristalinos (String color café y azul), con un voltaje V_{mp} de 285.3 [V], un corriente I_{mp} de 8.52 [A] de igual manera que los strings de la sección C y la potencia en la entrada MPPT 1 es de 4.86 [kWp], mientras el MPPT2 recibirá la potencia de 2 series de strings monocristalino (naranja y blanco), un voltaje V_{mp} de 282.6 [V], una corriente I_{mp} de 8.91[A] y una potencia en la entrada MPPT 2 de 5.04 [kWp]. En total, el inversor tiene la una potencia instalada de 9,9 [kWp].

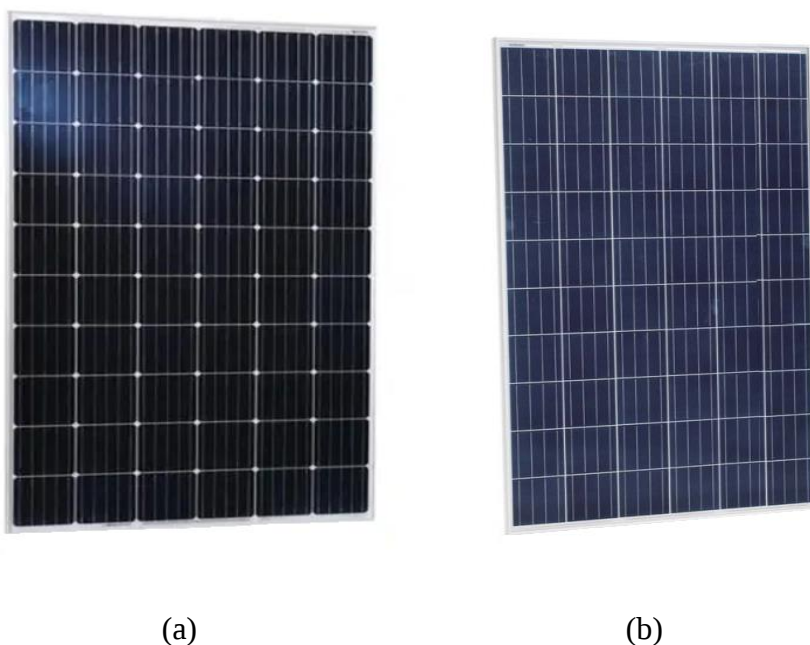
Tabla 3.1: Resumen de las secciones utilizadas

Seccion	Inversor	N° Fases	Tipo de Módulos	N° Módulos	Tipo de Estructura	N° String	N° Entradas MPPT	Potencia MPPT [kWp]	Potencia Instalada [kWp]
C	Fronius	3	Policristalino	36	Fija	4	2	4,86	9,72
	Symo 10[kW]							4,86	
D	Fronius	3	Policristalino	18	Fija	2	2	4,86	9,9
	Symo 10[kW]		Monocristalino	18		2		5,04	

Fuente: Elaboración Propia

La tabla 3.1 detalla los elementos que se utilizaron para realizar esta investigación, tales como: inversores, tipo de tecnología de los paneles FV, cantidad de Strings, potencia por entrada MPPT potencia total instalada.

Los paneles fotovoltaicos presentes en la MicroRed UCSC son de diferentes tecnologías, monocristalinas, policristalina y capa fina o amorfo. De las cuales utilizaremos solo 2, monocristalino y policristalinos presentes en la Sección C y D



**Figura. 3.2: Módulos Fotovoltaicos utilizados en la investigación,
(a) Panel Monocristalino, (b) Panel Policristalino**

Fuente: Elaboración Propia

La figura 3.2 muestra los dos tipos de tecnologías de paneles fotovoltaicos a la izquierda está el panel Monocristalino y a la derecha el Policristalino los cuales se diferencian es su aspecto físico principalmente en su color y en la forma de la celda ya que el monocristalino tiene los bordes de la celda redondeados, mientras que las celdas del policristalino son más cuadradas los detalles de estos dos modelos se indican en las tablas 3.2 y 3.3.

Tabla 3.2: Especificaciones técnicas Panel Monocristalino

Modulo Monocristalino YL 280D	
Potencia nominal (P _{máx})	280 [Wp]
Voltaje VMPP	31,4 [V]
Corriente IMPP	8,91 [A]
Voltaje Circuito Abierto VOC	39,3 [V]
Corriente Circuito Abierto ISC	9,38 [A]
Eficiencia del Modulo	17,25 %
Coeficiente de Temperatura de VOC	-0,32 %/°C
Coeficiente de Temperatura de ISC	0,05 %/°C

Fuente: Elaboración Propia

La tabla 3.2 describe las características que posee el panel solar fotovoltaico YL 280D, este modelo de panel está presente la sección D de la MicroRed UCSC.

Tabla 3.3: Especificaciones técnicas Panel Policristalino

Modulo Policristalino JKM 270 PP	
Potencia nominal (P _{máx})	270 [Wp]
Voltaje VMPP	31,7 [V]
Corriente IMPP	8,52 [A]
Voltaje Circuito Abierto VOC	38,8 [V]
Corriente Circuito Abierto ISC	9,09 [A]
Eficiencia del Modulo	16,50 %
Coeficiente de Temperatura de VOC	-0,30 %/°C
Coeficiente de Temperatura de ISC	0,06 %/°C

Fuente: Elaboración Propia

La tabla 3.3 describe las características que posee el panel solar fotovoltaico JMK 270 PP, este panel está presente en la sección C y D.

Tanto la sección de paneles fotovoltaicos C y D se conectan al mismo modelo de inversor Fronius Symo 10.0-3-M de 10 kW el cual se puede observar en la figura 3.3 y sus especificaciones se muestran en la tabla 3.4.



Figura. 3.3: Inversor Fronius Symo 10 kW

Fuente: Web Fronius

Tabla 3.4: Especificaciones técnicas inversor Fronius Symo 10 kW

Inversor Fornius Symo 10 kW	
Potencia nominal AC	10.000 [W]
Maxima Potencia de Salida	10.000 [VA]
Cantidad de Entradas MPPT	2
Corriente Maxima de entradas MPPT	27,0 / 16,5 [A]
Voltaje minimo de entrada	200 [V]
Voltaje Maximo de entrada	900 [V]
Eficiencia	98,0 %
Frecuencia	50 [Hz]
Factor de Potencia ($\cos \varphi$)	0 - 1 [Ind / Cap]

Fuente: Elaboración Propia

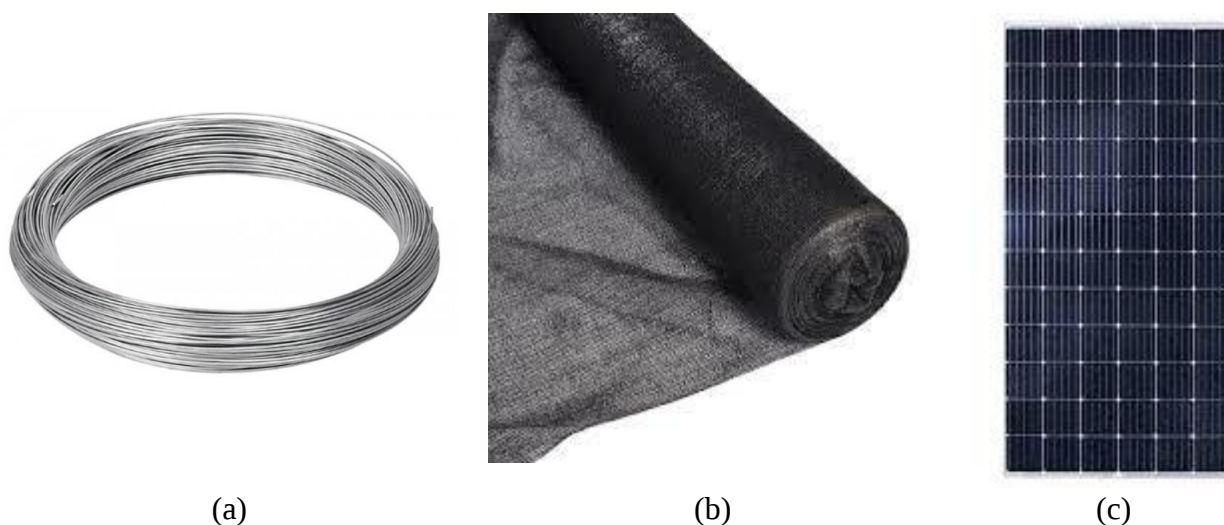
Para la realización de las pruebas se utilizó la sección C, específicamente los strings Celeste y Verde conectados a la entrada MPPT 1 del inversor Fronius Symo 2, mientras que la sección D se usaran los strings Café y Azul que están conectados a la entrada MPPT 1 del inversor Fronius Symo 1. Estos strings se utilizarán como punto de comparación para las diferentes pruebas de sombreado y suciedad.

3.3. Pruebas de sombreado

Como se indicó anteriormente estas pruebas fueron realizadas durante la estación de primavera específicamente durante el mes de Octubre. Para realizar las pruebas se cubrió a los paneles fotovoltaicos con una malla negra que simulaba el sombreado, esta malla se utilizó para simular sombreado de distintas formas para apreciar el comportamiento de las variables eléctricas ante esta condición.

Elementos utilizados

- Malla Negra
- Alambre
- Paneles Fotovoltaicos



**Figura 3.4 Elementos Utilizados para emular sombreado,
(a) Alambre, (b) Malla Rachel negra, (c) Paneles FV**

Fuente: Elaboración Propia

3.3.1 Metodología

La metodología utilizada consistió en utilizar una malla con dimensiones similares a las de un panel fotovoltaico Policristalino JKM 270 PP, es decir, 1,65 metros de largo y 0,992 metros de ancho. Se otorgó un margen adicional a la malla para asegurar la cobertura completa de la superficie del panel. Con esta configuración, se llevaron a cabo diversas pruebas variando la cantidad de paneles

sombreados y el porcentaje de superficie cubierta. El objetivo era simular diferentes escenarios de sombreado que pueden ocurrir en condiciones reales de funcionamiento.

Para realizar estas pruebas, se esperó a que los paneles dejaran de recibir radiación solar y su producción de energía alcanzara cero. En ese momento, se cubrieron la cantidad de paneles del string 1 correspondiente al estudio. Estas pruebas se llevaron a cabo en las tardes y noches para evitar interferencias en las mediciones y garantizar condiciones estables.

Esta metodología permitió evaluar el impacto del sombreado en la producción de energía de los paneles fotovoltaicos en situaciones reales. Al variar la cantidad de paneles sombreados y el porcentaje de superficie cubierta, se obtuvieron datos relevantes que posteriormente se utilizaron en el análisis y entrenamiento de los algoritmos de Machine Learning.

El resumen de todas las pruebas realizadas se muestra en la tabla 3.5 donde se indica la fecha en las cuales se realizaron las pruebas, la cantidad de superficie cubierta y como identificar que paneles se cubrieron según el código de colores mostrado en la figura 3.1.



Figura 3.5: Sombreado de manera experimental de paneles fotovoltaicos

Fuente: Elaboración Propia

La figura 3.5 muestra una de las pruebas de sombreado realizadas, donde se cubrió totalmente un panel fotovoltaico en la Microred UCSC.

Tabla 3.5 Pruebas de Sombreado Realizadas

Fecha	Cantidad de Paneles con Sombreado	Superficie del Panel Sombreado	Color del String Sombreado (Figura 3.1)
06 – Octubre – 2022	1	Totalmente Sombreado	1 – Celeste
07 – Octubre – 2022	2	Totalmente Sombreado	1 - Celeste 1 – Verde
10 – Octubre – 2022	2	Totalmente Sombreado	2 – Celeste
11– Octubre – 2022	4	Totalmente Sombreado	2 - Celeste 2 – Verde
13 – Octubre – 2022	2	1 - Mitad Superior 1 - Mitad Inferior	2 – Celeste
14 – Octubre – 2022	4	2 - Mitad Superior 2 - Mitad Inferior	2 - Celeste 2 – Verde
15 – Octubre – 2022	4	2 - Mitad Superior 2 - Mitad Inferior	2 - Celeste 2 – Verde
16 – Octubre – 2022	2	2 - Mitad Inferior	1 - Celeste 1 – Verde
17 – Octubre – 2022	2	2 - Mitad Superior	1 - Celeste 1 – Verde

Fuente: Elaboración Propia

De la tabla 3.5 es interesante analizar el comportamiento de las variables eléctricas cuando se producen escenarios similares por ejemplo el comparar el comportamiento de tapar dos paneles fotovoltaicos siendo el caso del día 07 de octubre que se cubre un panel del string celeste y otro del string verde y compararlo con el día 10 de octubre donde se tapan dos paneles, pero del mismo conjunto de paneles (String Celeste).

3.3.2 Sombreado de un solo panel Fotovoltaico

En este punto se comparan casos de sombreado realizados un día en que se realizó la prueba de sombreado total de un panel FV en comparación con un día sin sombreado. El día antes mencionado es el 06 de octubre. El esquema del caso se muestra en la figura 3.6.

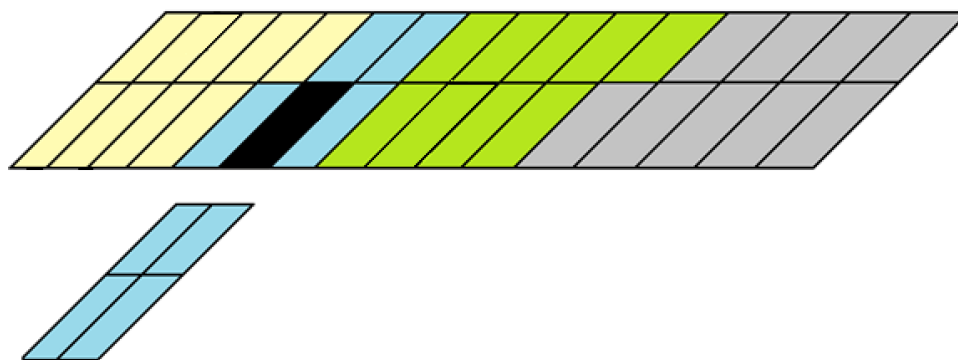


Figura 3.6 Representación de prueba de sombreado total de un panel FV

Fuente: Elaboración Propia

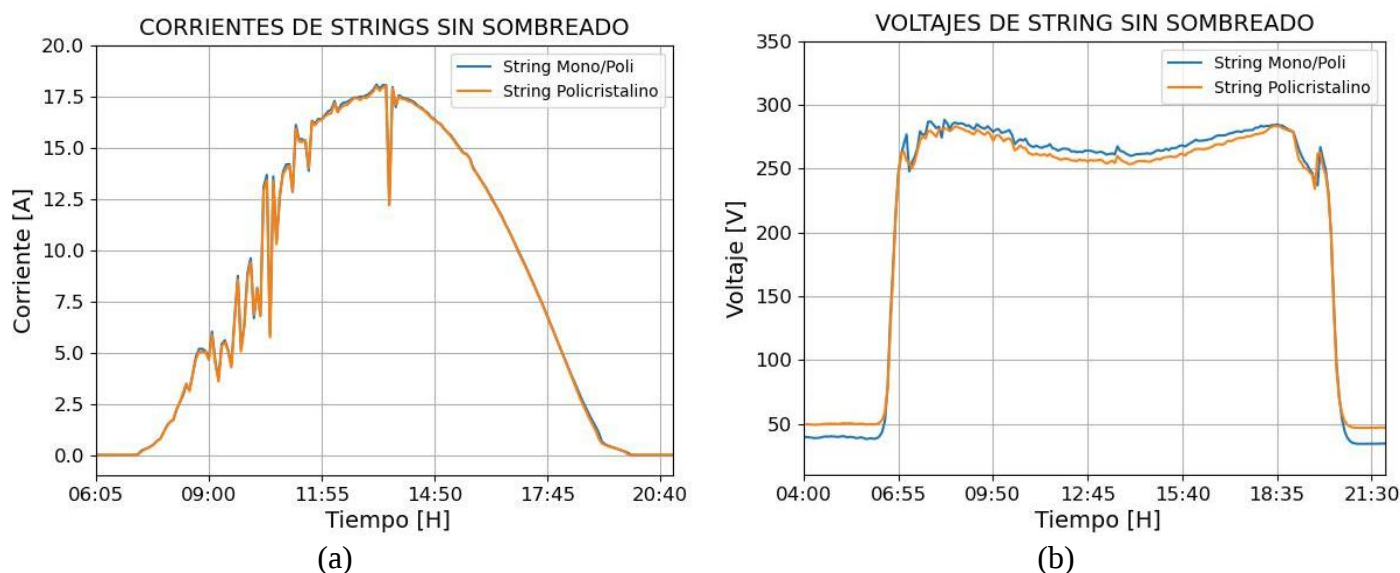


Figura 3.7 Muestra el comportamiento del voltaje y la corriente sin sombreado

(a) Corriente en [A] de los Strings sin sombreado, (b) Voltaje en [V] de los Strings sin sombreado

Fuente: Elaboración Propia

En la figura 3.7 podemos ver cómo se comportan las variables eléctricas en un día normal sin condiciones de sombreado específicamente el día 12 de Octubre, en esta se puede apreciar que la corriente del string Policristalino prácticamente es igual a la corriente del string Mono/ Poli que es el que se está utilizando como referencia. Por otro lado, se aprecia que el voltaje del string Policristalino

es un poco menor al voltaje de referencia aproximadamente 6 [V] menor en los datos obtenidos es decir es aproximadamente un 2% menor que el voltaje de referencia.

A. Sombreado Caso 1

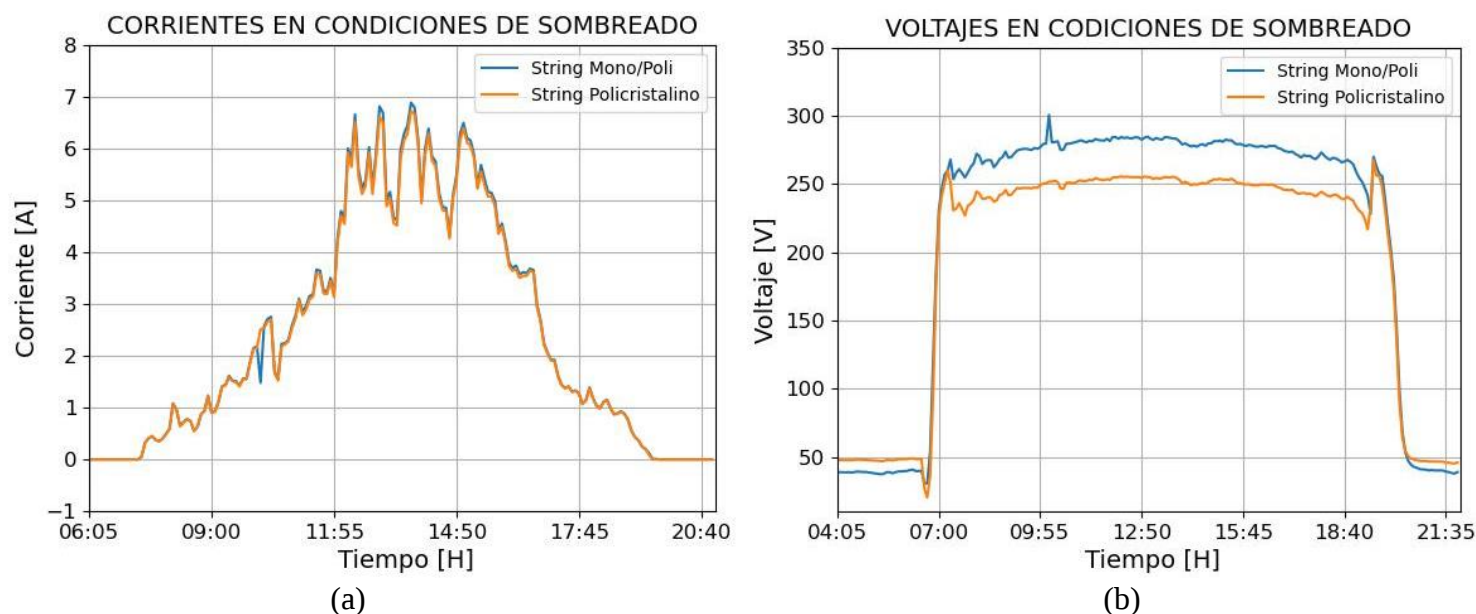


Figura 3.8 Gráficas de corriente y voltaje con un panel FV completamente sombreado 06 de Octubre,

(a) Corriente en [A] de los Strings en el caso 1 (Somb), (b) Voltaje en [V] de los Strings en el caso 1 (Somb).

Fuente: Elaboración Propia

En la figura 3.8 se presenta el comportamiento de la corriente y el voltaje en relación a la presencia de sombreado. Se observa que la corriente se mantiene constante y es prácticamente idéntica a la corriente del string utilizado como referencia, lo cual también se evidencia en la figura 3.7 donde la corriente del string evaluado es similar a la del string Mono/Poli.

Por otro lado, se aprecia una disminución en el valor del voltaje cuando hay sombreado, aunque esta disminución es menor en comparación con el voltaje de referencia. En la figura 3.8, el voltaje se mantiene cerca de los 255 [V] a lo largo del día, mientras que, en ausencia de sombreado, el voltaje alcanza valores superiores, llegando hasta los 282 [V]. Esto implica una diferencia de aproximadamente 27 [V] en cada punto de datos del día, lo que representa una disminución de alrededor del 10% en comparación con el voltaje de referencia.

B. Sombreado Caso 2

Otro caso que se puede comparar es cuando se cubren dos paneles fotovoltaicos, pero de manera parcial cubriendo en uno la mitad superior y en otro la mitad inferior todo esto en el mismo string. Este experimento se puede apreciar en la figura 3.9 donde la superficie total cubierta en ambos paneles es la de un panel sombreado de manera completa.

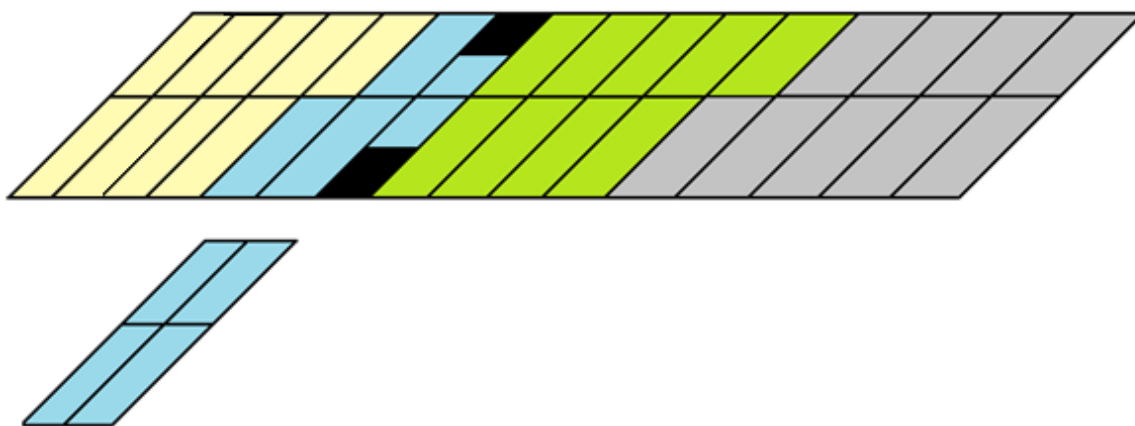


Figura 3.9 Representación de pruebas de sombreado parcial dos paneles FV del mismo string

Fuente: Elaboración Propia

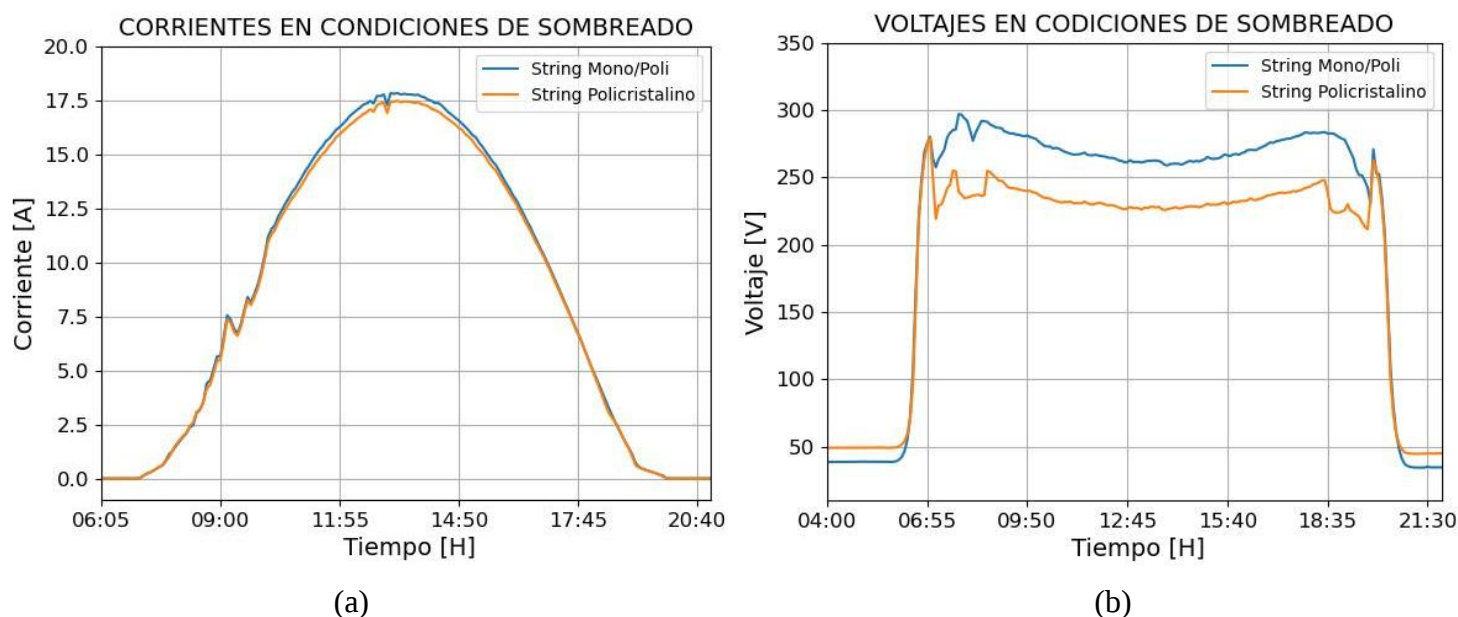


Figura 3.10 Gráficas de corriente y voltaje con dos paneles FV sombreados a la mitad el 13 de Octubre,

(a) Corriente en [A] de los Strings en el caso 2 (Somb), (b) Voltaje en [V] de los Strings en el caso 2 (Somb).

Fuente: Elaboración Propia

En la figura 3.10 se representa el comportamiento del voltaje y la corriente en respuesta al sombreado, como se muestra en la figura 3.9. Se observa que la corriente sigue un patrón similar al de los casos anteriores, siendo casi idéntica a la corriente de referencia utilizada. Sin embargo, se aprecia una mayor diferencia en los picos de corriente entre las 11:00 a.m. y las 3:00 p.m.

En cuanto al voltaje, al sombrear dos mitades de paneles solares, se produce una disminución aún mayor en comparación con el voltaje de referencia. En el día 13 de octubre, se observa una diferencia de 36 [V] en todos los datos de voltaje obtenidos, lo que representa una disminución del 13% en relación al voltaje de referencia utilizado.

Estos resultados indican que, al sombrear una mayor cantidad de paneles, el voltaje generado por el sistema solar experimenta una disminución más pronunciada en comparación con el sombreado de un solo panel. Por otro lado, la corriente se mantiene en niveles cercanos a la corriente de referencia en la mayoría de los casos.

C. *Sombreado Caso 3 y Caso 4*

Otros dos casos que son similares al del día 13 de octubre son los de las figuras 3.11 y 3.12 en los cuales igual se sombrean los paneles fotovoltaicos, pero en primer caso se hace en la parte inferior de estos y en la segunda imagen la parte superior, en ambos casos los paneles sombreados son de diferentes string siendo la superficie total sombreada equivalente a un módulo fotovoltaico. Estas pruebas se realizaron los días 16 de octubre y 17 de octubre respectivamente.

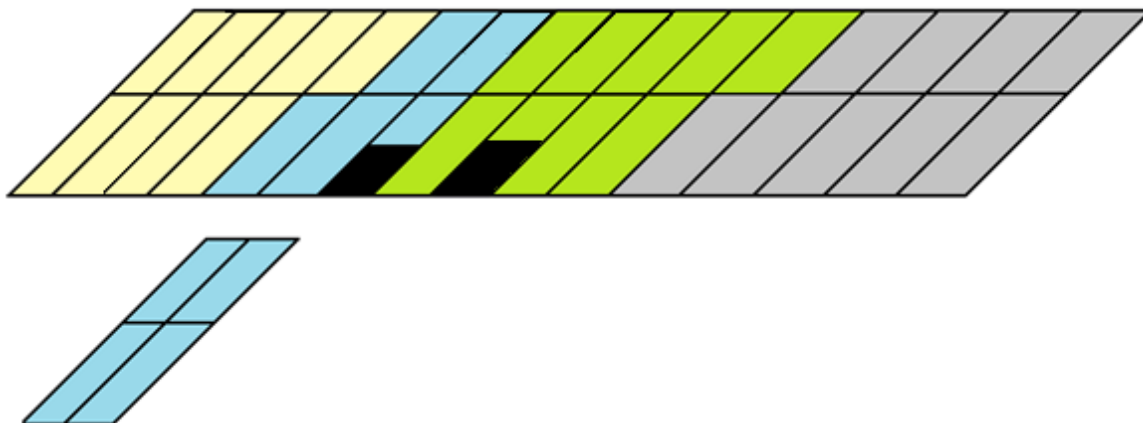


Figura 3.11: Representación de pruebas de sombreado parcial dos paneles FV de diferentes string

Fuente: Elaboración Propia

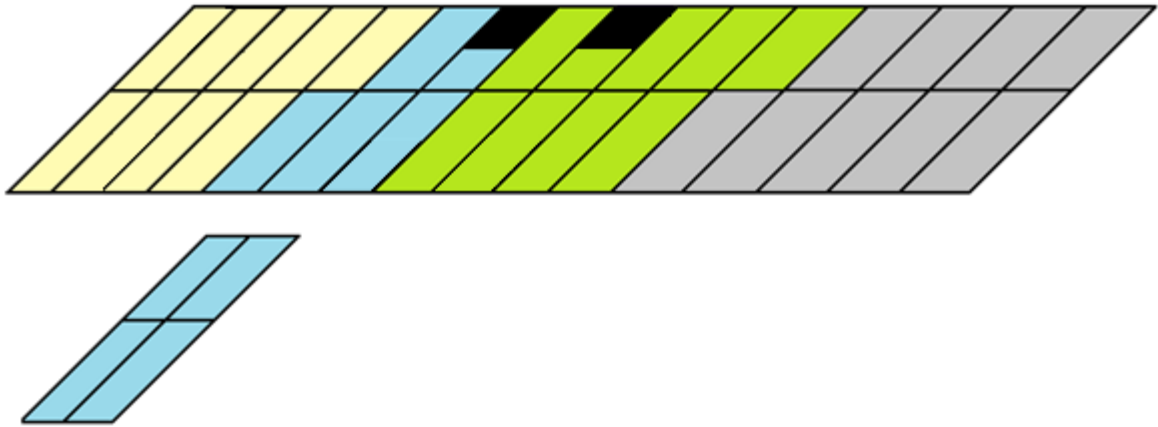


Figura 3.12: Representación de pruebas de sombreado parcial dos paneles FV de diferentes strings

Fuente: Elaboración Propia

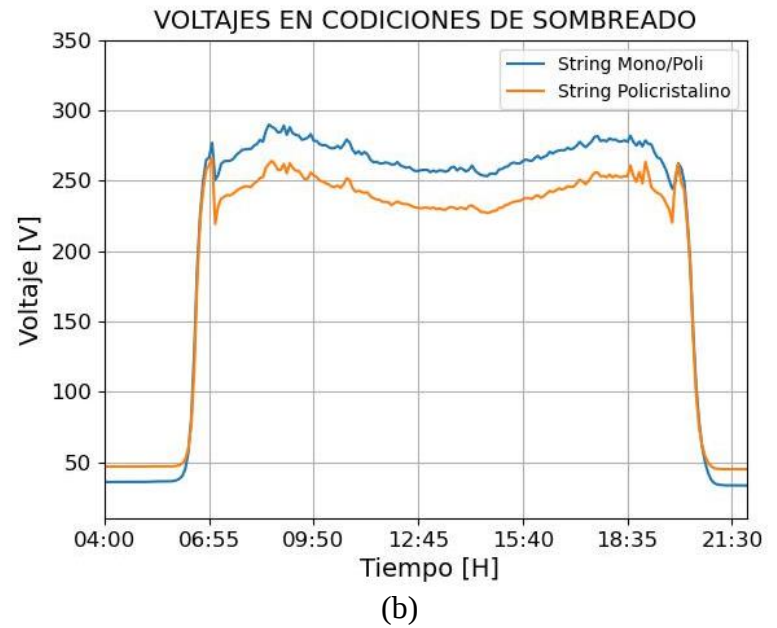
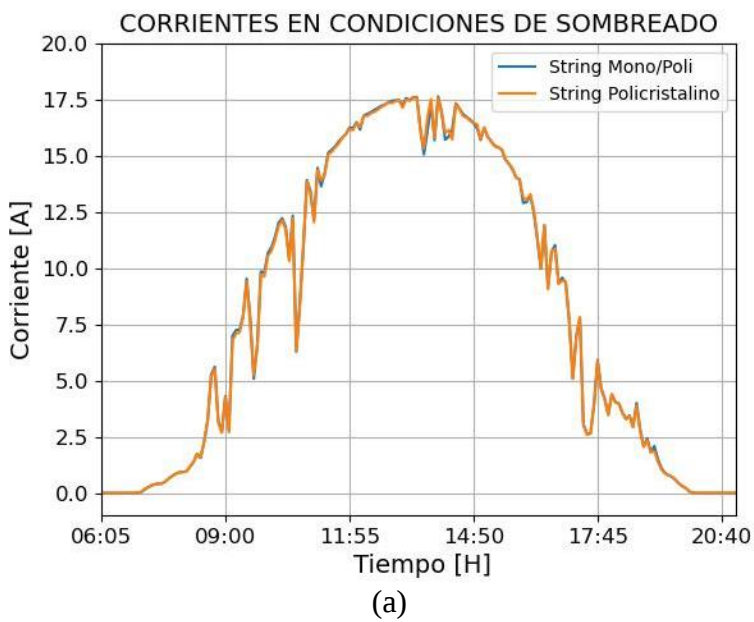


Figura 3.13 Gráficas de corriente y voltaje con dos paneles FV sombreados la mitad inferior 16 de Octubre,

(a) Corriente en [A] de los Strings en el caso 3 (Somb), (b) Voltaje en [V] de los Strings en el caso 3 (Somb).

Fuente: Elaboración Propia

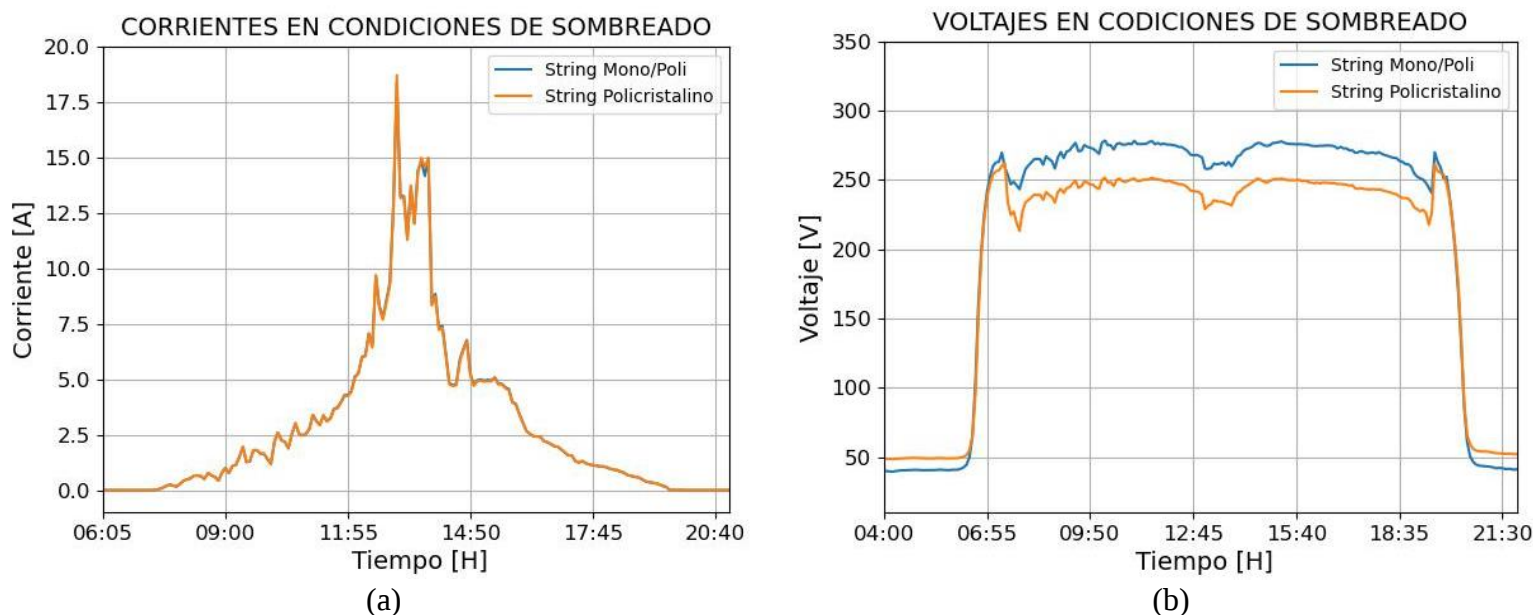


Figura 3.14 Gráficas de corriente y voltaje con dos paneles FV sombreados la mitad Superior 17 de Octubre
(a) Corriente en [A] de los Strings en el caso 4 (Somb), (b) Voltaje en [V] de los Strings en el caso 4 (Somb).

Fuente: Elaboración Propia

En las figuras 3.13 y 3.14 se representa el comportamiento de la corriente y el voltaje bajo diferentes condiciones de sombreado, que resultan ser bastante similares entre sí en términos del comportamiento de las variables eléctricas. Se observa que las corrientes se mantienen muy cercanas a la corriente de referencia, al igual que en los casos anteriores analizados.

Respecto al voltaje, se observa que las curvas de ambos casos son muy similares entre sí. Si bien disminuyen en comparación con el voltaje de referencia, esta disminución es menor en comparación con los casos presentados en las figuras 3.6 y 3.9. En ambos casos, el voltaje experimenta una reducción de aproximadamente 26 [V], lo que representa una disminución del 9,7% en relación a la curva de referencia.

Estos resultados indican que, bajo las condiciones de sombreado analizadas en las figuras 3.13 y 3.14, principalmente el voltaje se ve afectado, aunque en menor medida en comparación con los casos previamente estudiados. Esto sugiere que las condiciones específicas de sombreado presentadas en estas figuras pueden tener un impacto menos significativo en el rendimiento de los paneles solares en comparación con otros escenarios de sombreado.

Tabla 3.6 Comparación casos de sombreado de un panel FV

Caso Experimental	Cantidad total de superficie sombreada del panel	Variación del voltaje respecto a la referencia	Porcentaje de Variación de voltaje	Variación de corriente respecto a la referencia	Porcentaje de Variación de Corriente
Sin Sombreado	0	-6 [Volts]	2 %	-0,05 [A]	0,6 %
1 panel sombreado	1 panel Entero	-27 [Volts]	10%	-0,035 [A]	1,44%
1 mitad superior					
1 mitad inferior del mismo String	1 panel Entero	-36 [Volts]	13,5%	-0,16 [A]	1,93%
2 mitades inferiores de distintos String	1 panel Entero	-26,24 [Volts]	9,76 %	-0,018 [A]	0,22%
2 mitades superiores de distintos String	1 panel Entero	-26,11 [Volts]	9.6%	-0,008 [A]	0,27%

Fuente: Elaboración Propia

En la tabla 3.6 se presentan los resultados de sombrear un panel completo de diferentes maneras, y se puede concluir que el sombreado tiene un impacto mayor en el voltaje en comparación con otras variables como la corriente. Es interesante observar que cuando se sombrea la misma superficie en un mismo string, la disminución del voltaje es más pronunciada que cuando la sombra se distribuye en diferentes strings. Este fenómeno se explica por el hecho de que, al estar los strings conectados en paralelo, el voltaje se reduce en ambos strings cuando una parte de la superficie se encuentra sombreada. Esto significa que el impacto en el voltaje no se limita solo al string afectado por la sombra, sino que también se propaga al otro string conectado en paralelo.

Estos resultados sugieren que la ubicación específica del sombreado puede tener un impacto significativo en el voltaje, lo que debe ser tenido en cuenta al diseñar y optimizar los sistemas fotovoltaicos. En conclusión, los resultados de la tabla 3.6 evidencian que el sombreado afecta de manera más notable al voltaje, especialmente cuando la sombra cubre la misma superficie distintos string, debido a la configuración en paralelo de estos.

3.3.3 Sombreado de dos o más paneles Fotovoltaicos

D. Sombreado Caso 5

En este caso se somborean dos paneles completos del string celeste y verde uno de cada strings, el diagrama de esta prueba se aprecia en la figura 3.15

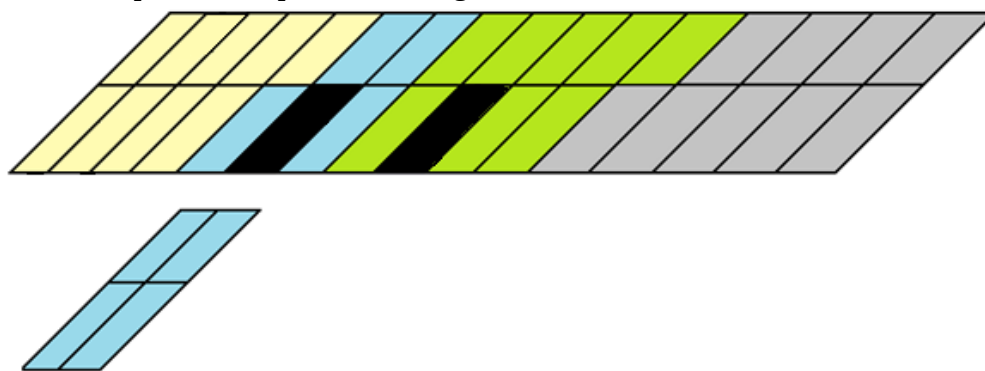


Figura 3.15 Representación de prueba de sombreado total dos paneles FV de diferentes strings

Fuente: Elaboración Propia

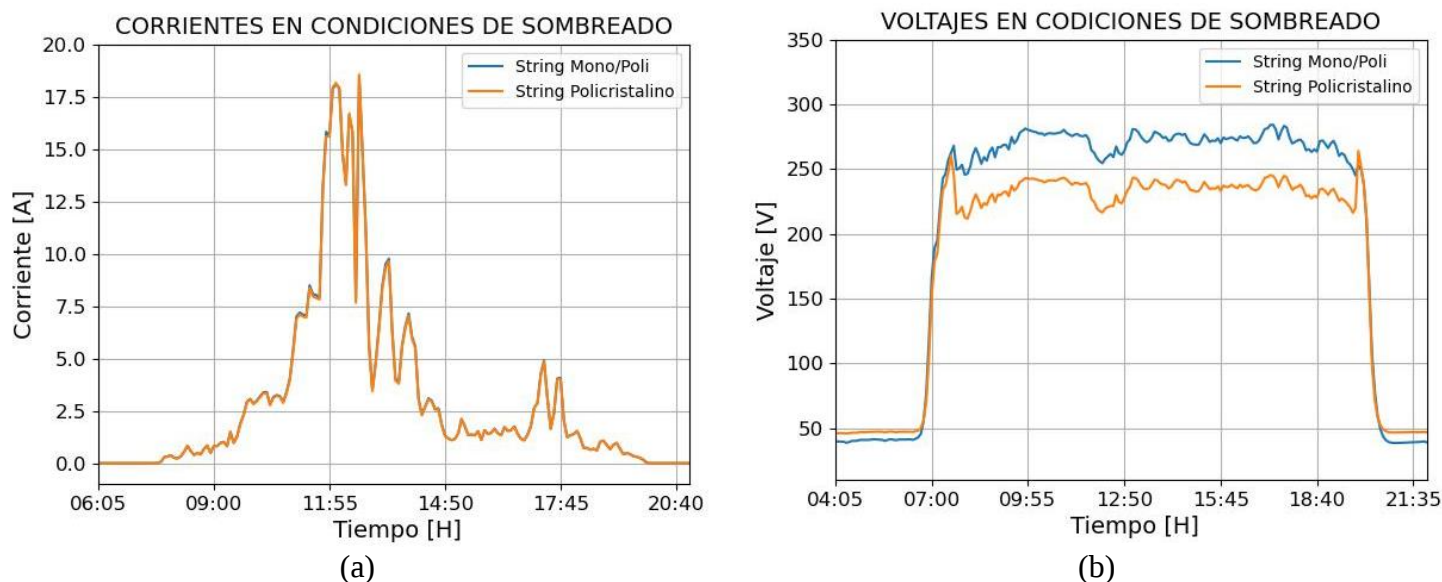


Figura 3.16 Variables eléctricas en condiciones de dos paneles FV sombreados distintas strings,

(a) Corriente en [A] de los Strings en el caso 5 (Somb), (b) Voltaje en [V] de los Strings en el caso 5 (Somb).

Fuente: Elaboración Propia

En la figura 3.16 se observa el comportamiento de las variables eléctricas. En cuanto a la corriente, se puede apreciar que se mantiene prácticamente constante y no presenta variaciones significativas en comparación con la corriente utilizada como referencia. Esto indica que el sombreado no tiene un impacto considerable en la corriente generada por el panel. Por otro lado, en el caso del voltaje, se observa una disminución promedio de 35 [V] en comparación con el voltaje de referencia. Esta reducción representa aproximadamente un 13.1% de disminución en el voltaje. Este

comportamiento es similar al observado cuando se sombrean dos paneles, uno cubriendo la mitad superior y el otro la parte inferior.

E. Sombreado Caso 6

En este caso se sombrearon de manera artificial dos paneles del string Celeste como se ilustra en la figura 3.17.

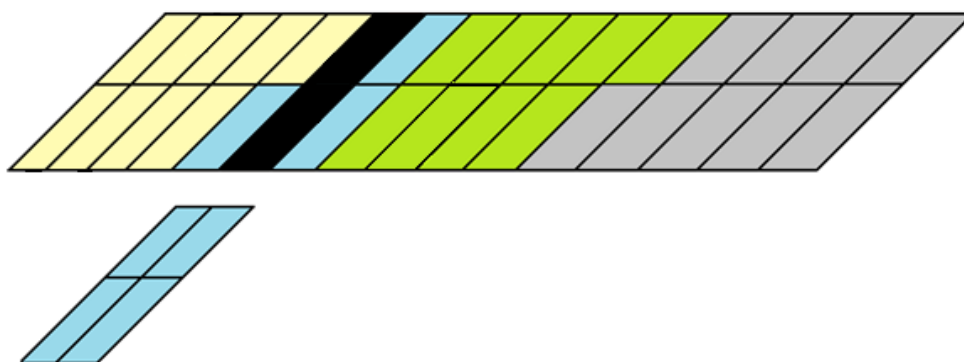


Figura 3.17 Representación de pruebas de sombreado total dos paneles FV del mismo string

Fuente: Elaboración Propia

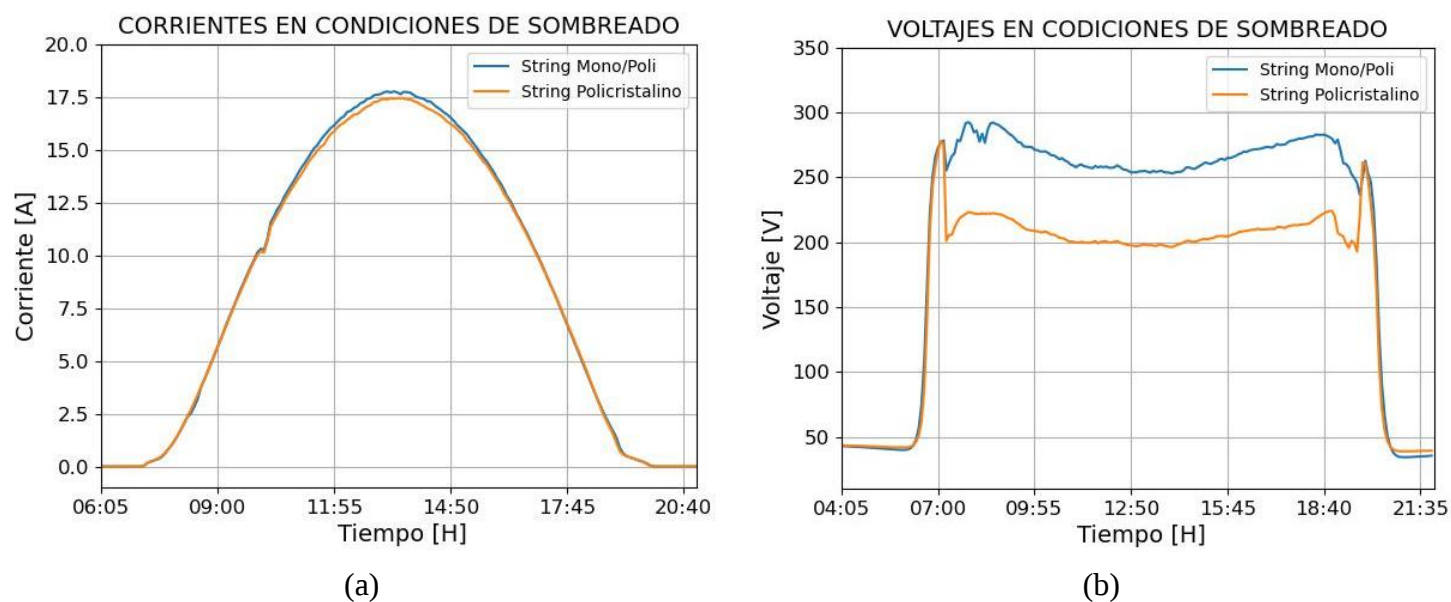


Figura 3.18 Variables eléctricas en condiciones de dos paneles FV sombreados del mismo string

(a) Corriente en [A] de los Strings en el caso 6 (Somb), (b) Voltaje en [V] de los Strings en el caso 6 (Somb).

Fuente: Elaboración Propia

En comparación con el caso anterior y la configuración sin sombreado, se observan cambios significativos en las variables eléctricas. La corriente presenta una disminución leve de 0.11 [A] en comparación con la corriente utilizada como referencia. Por otro lado, el voltaje experimenta una reducción más significativa, disminuyendo en 57.1 [V], lo que representa aproximadamente un 21.7%

de reducción. Esta disminución en el voltaje es casi el doble de la observada en el caso anterior. Además, en este caso se presenta la mayor disminución de corriente por lo que se puede intuir que a medida que mayor superficie este sombreada más disminuirá la corriente.

F. Sombreado Caso 7

En este caso se sombrearon cuatro mitades de distintos paneles FV, tapando una parte superior y la parte inferior de dos paneles diferentes del string celeste, repitiendo esto mismo en el string verde. Esto se puede apreciar en la figura 3.19

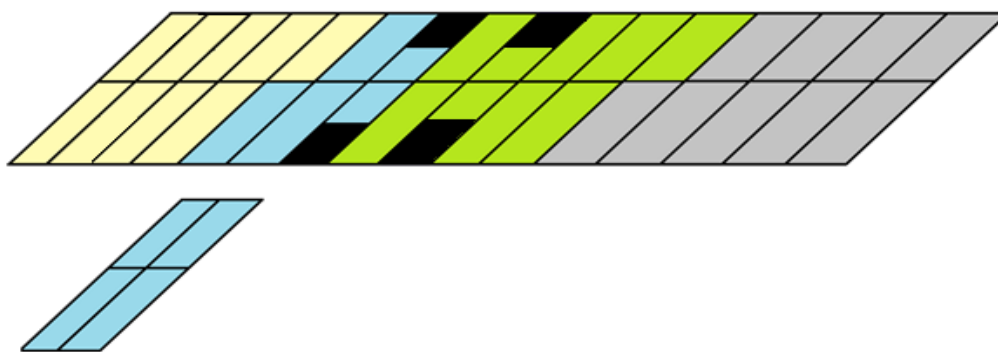


Figura 3.19 Representación de pruebas de sombreado parcial cuatro paneles FV de distintos strings

Fuente: Elaboración Propia

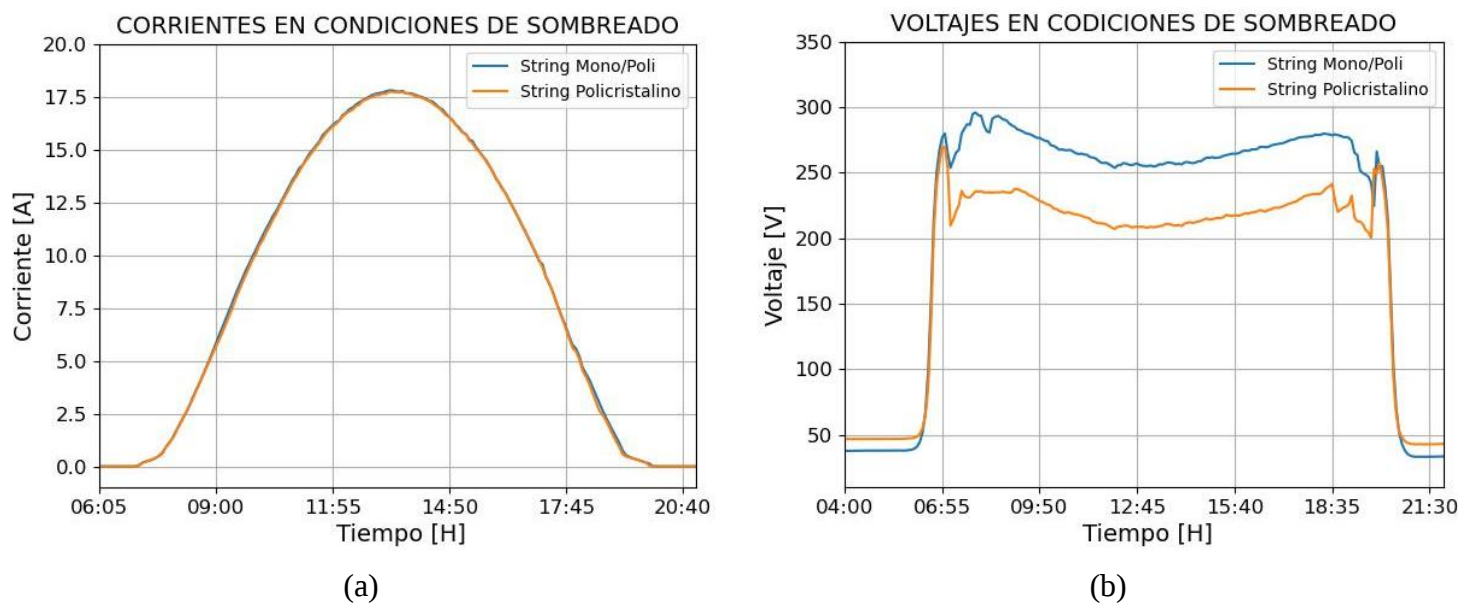


Figura 3.20 Comportamiento variables eléctricas en condiciones de sombreado,

(a) Corriente en [A] de los Strings en el caso 7 (Somb), (b) Voltaje en [V] de los Strings en el caso 7 (Somb).

Fuente: Elaboración Propia

En relación al caso anterior, se observa un comportamiento similar en la curva de voltaje, donde se presenta una disminución de 47.7 [V], lo cual representa un cambio del 17.7% en comparación con el voltaje de referencia. Por otro lado, la corriente se mantiene muy similar a la que se obtuvo cuando no hay presencia de sombreado, presentando una variación de tan solo -0.06 [A] en comparación con el caso mencionado anteriormente.

G. Sombreado Caso 8

En este caso se somborean cuatro paneles en su totalidad, dos del string celeste y dos del string verde como se ilustra en la figura 3.21.

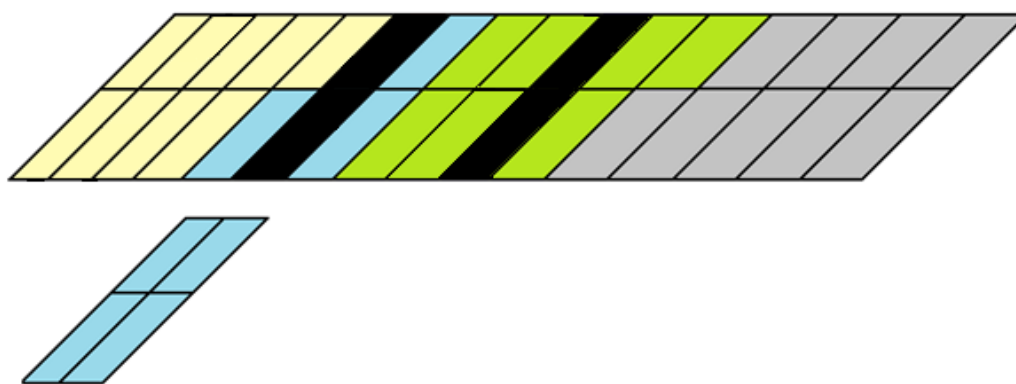
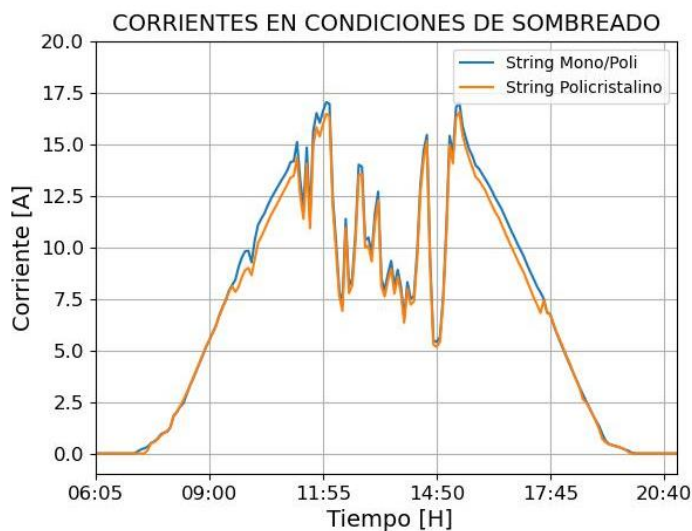
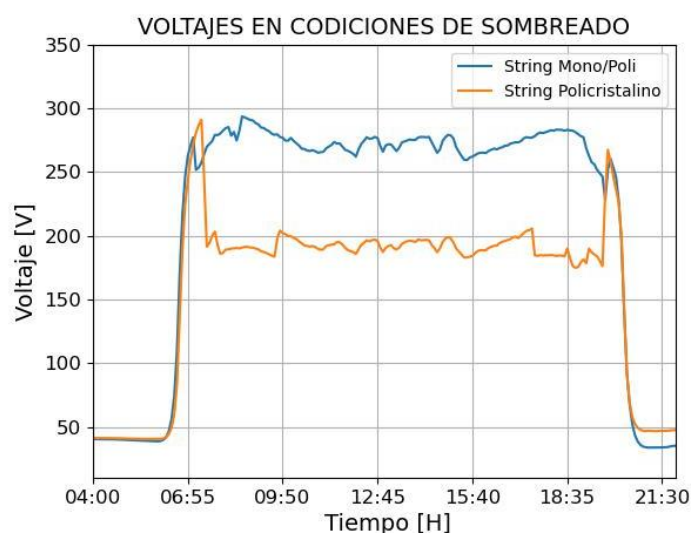


Figura 3.21 Representación de pruebas de sombreado total de cuatro paneles FV del distintos strings

Fuente: Elaboración Propia



(a)



(b)

Figura 3.22 Variables eléctricas en condiciones de sombreado de 4 paneles FV de dos distintos strings,

(a) Corriente en [A] de los Strings en el caso 8 (Somb), (b) Voltaje en [V] de los Strings en el caso 8 (Somb).

Fuente: Elaboración Propia

El comportamiento de las variables eléctricas se muestra en la figura 3.22, donde se observa una disminución significativa tanto en la corriente como en el voltaje en comparación con los casos anteriores. Este caso particular muestra la mayor reducción en la curva de voltaje y también en la corriente. En cuanto al voltaje, se aprecia una disminución de 82.12 [V], lo cual representa una reducción aproximada del 30% en comparación con el voltaje de referencia. Esta disminución es considerable y puede tener un impacto significativo en el rendimiento y la eficiencia del sistema fotovoltaico. Por otro lado, la corriente también presenta una disminución de 0.31 [A], lo que equivale a una variación aproximada del 4.3% en comparación con la corriente de referencia. Aunque la disminución en la corriente es menor que la del voltaje, sigue siendo significativa y debe tenerse en cuenta al evaluar el rendimiento del sistema.

Tabla 3.7 Comportamiento de las variables electricas en condiciones de sombreado de dos o mas modulos FV

Caso Experimental	Cantidad total de superficie sombreada del panel	Variación del voltaje respecto a la referencia	Porcentaje de Variación de voltaje	Variación de corriente respecto a la referencia	Porcentaje de Variación de Corriente
Sin Sombreado	0	-6 [Volts]	2 %	- 0,05 [A]	0,6 %
2 paneles sombreados diferentes Strings	2 paneles enteros	-35 [Volts]	13,01%	-0,025 [A]	0,82 %
2 paneles sombreados del mismo String	2 paneles enteros	-57,1 [Volts]	21,7 %	-0,11 [A]	1,36 %
2 mitades superiores de diferentes String	2 paneles enteros	-47,7 [Volts]	17,7 %	-0,06 [A]	0,76 %
2 paneles sombreado 2 por Strings	4 paneles enteros	-82,12 [Volts]	30 %	-0,31 [A]	4,36 %

Fuente: Elaboración Propia

En la tabla 3.7 se muestra el comportamiento de las variables eléctricas del sistema en relación con la cantidad de paneles fotovoltaicos que se encuentran en condición de sombreado. Podemos

observar que a medida que aumenta la cantidad de paneles sombreados, el voltaje experimenta una disminución significativa. En cuanto al voltaje, se puede apreciar que a medida que más paneles están sombreados, hay una disminución progresiva en el valor de voltaje. Esto indica que el sombreado tiene un impacto directo en la producción de voltaje del sistema fotovoltaico. La disminución del voltaje es más pronunciada en los casos donde se somborean una mayor cantidad de paneles, lo que implica una mayor superficie sombreada. Por otro lado, la corriente muestra un comportamiento más estable en comparación con el voltaje. En la mayoría de los casos, la corriente se mantiene constante, con una disminución menor al 1% en la mayoría de los casos. Sin embargo, en los casos donde se somborean dos y cuatro paneles completamente, se observa una disminución en la corriente del 1.36% y 4.36% respectivamente. Estos casos representan situaciones donde una mayor superficie de paneles está sombreada, lo que resulta en una reducción más significativa en la corriente generada.

3.4. Pruebas de Suciedad

En estas pruebas se replicó lo realizado en el sombreado, es decir se ensuciaron los paneles FV de la misma forma que con el sombreado, pero con una mezcla entre tierra del mismo sector de la MicroRed y agua, agua en determinadas cantidades para que sea una prueba fácil de replicar.

3.4.1 Metodología

Para simular el efecto de suciedad en los módulos fotovoltaicos, se llevaron a cabo pruebas utilizando diferentes mezclas de agua y tierra en distintas cantidades para encontrar la mezcla que mayor similitud tuviera con la suciedad real que se acumula en los módulos FV de la MicroRed UCSC. Después de realizar varias pruebas, se decidió utilizar una cantidad específica de 45 gramos de tierra molida mezclada con 500 ml de agua, la cual crea una capa de suciedad similar al soiling que se acumula naturalmente, pero con una densidad un poco mayor ya que tenía que generar una variación en las variables eléctricas. Como esta mezcla se tuvo que realizar varias veces para llevar a cabo todas las pruebas, se empleó un vaso milimétrico y una balanza para asegurar la precisión en la cantidad de cada elemento utilizado. Esto en mayor detalle se puede ver en la figura 3.23 donde se muestran las cantidades utilizadas. Si bien en la figura a se muestra una balanza marcando 68 gramos esto se debe a que en frasco el cual se utilizaba para poder poner la tierra sobre la balanza tiene un peso de 23 gramos.

Para aplicar la suciedad a los módulos fotovoltaicos, se utilizó un paño que se impregnaba con la mezcla preparada. Este paño se aplicó en las áreas específicas donde se deseaba simular la presencia

de suciedad y se pasaba las veces que fuera necesaria hasta crear la capa de suciedad que se deseaba analizar, replicando así las condiciones de sombreado previamente establecidas. Es importante destacar que estas pruebas de suciedad se llevaron a cabo en el mes de enero. Al igual que en las pruebas de sombreado las pruebas se dejaban preparadas de un día para otro por lo que se tenía que esperar que los paneles FV dejaran de producir energía para impregnarlos con la mezcla antes descrita. Posterior a esto al día siguiente se revisaba que la capa de suciedad estuviera aun intacta, de esta manera se cercioraba que la prueba que se llevaba a cabo fuera correcta.



(a)



(b)

**Figura 3.23: Elementos utilizados para emular la suciedad en los Paneles,
(a) cantidad de gramos de tierra 45g, (b) Cantidad de agua utilizada 500 ml**

Fuente: Elaboración Propia

La tabla 3.8 muestra los detalles de todas las pruebas realizadas, las cuales son las mismas que se presentan en la tabla 3.5, pero en un orden diferente ya que muchas pruebas tuvieron que ser repetida ya que el rocío y la humedad de algunos días hacía que la suciedad artificial se saliera de los paneles fotovoltaicos teniendo que esperar hasta el otro día para volver a ensuciar los paneles FV.

Estas pruebas permitieron evaluar el impacto de la suciedad en las variables eléctricas de los módulos fotovoltaicos durante el mes de enero.

Tabla 3.8 Pruebas de Suciedad Realizadas

Fecha	Cantidad de Paneles con Suciedad	Superficie del Panel Ensuciado	Color del String con Suciedad (Figura 3.1)
04 – Enero – 2023	1	Totalmente Sucio	1 – Celeste
05 – Enero – 2023	2	Totalmente Sucio	1 – Celeste 1 – Verde
08 – Enero – 2023	4	2 - Mitades Superiores sucias 2 - Mitades Inferiores sucias	2 – Celeste 2 – Verde
09 – Enero – 2023	4	2 - Mitades Superiores sucias 2 - Mitades Inferiores sucias	2 – Celeste 2 – Verde
10 – Enero – 2023	2	1 - Mitad Superior sucia 1 - Mitad Inferior sucia	2 – Celeste
15 – Enero – 2023	2	2 - Mitades Inferiores sucias	1 – Celeste 1 – Verde
16 – Enero – 2023	2	2 - Mitades Superiores sucias	1 – Celeste 1 – Verde
17 – Enero – 2023	4	Totalmente Sucios	2 – Celeste 2 – Verde
18 – Enero – 2023	2	Totalmente Sucias	2 – Celeste

Fuente: Elaboración Propia



Figura 3.24 Muestra la comparación entre un panel limpio y otro ensuciado intencionalmente

Fuente: Elaboración Propia



Figura 3.25 Comparación entre un panel limpio y uno sucio por condiciones del entorno

Fuente: Elaboración Propia

Si comparamos las figuras 3.24 y 3.25 vemos que la suciedad natural es mucho menor que la suciedad artificial, pero en ambos casos la distribución es la misma ya que cubre toda superficie del panel FV.

3.4.2 Pruebas de suciedad en un panel Fotovoltaico

Como se indicó anteriormente las pruebas realizadas son las mismas que para el sombreado, pero en este caso con suciedad en esta parte del capítulo se analizarán las pruebas mostradas en las figuras 3.6, 3.9, 3.11 y 3.12.

En primera instancia se verá cómo se comportan las variables de corriente y voltaje en un día normal del mes de enero

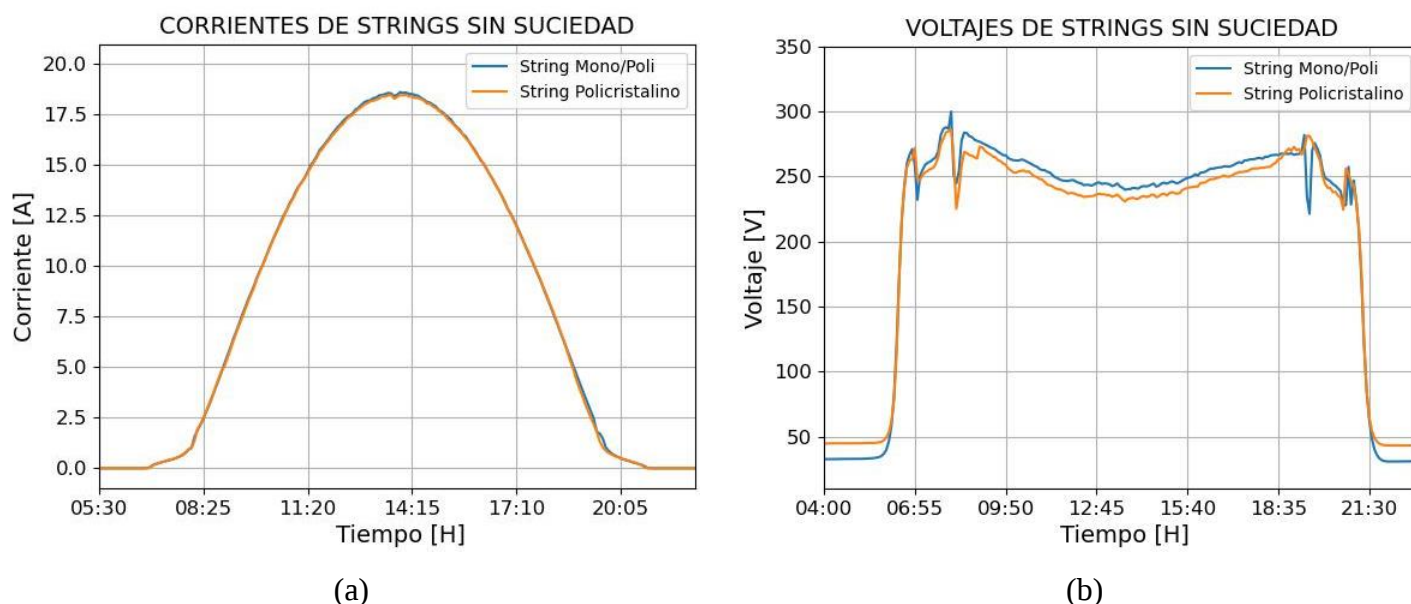


Figura 3.26 Comportamiento de las variables eléctricas cuando los paneles no presentan suciedad,

(a) Comportamiento de la corriente, (b) Comportamiento del Voltaje

Fuente: Elaboración Propia

De la figura 3.26, podemos apreciar que la corriente del string que se está evaluando (String Policristalino), es igual con respecto al string de referencia (string Mono/Poli), donde prácticamente no se aprecian diferencias entre ellas, lo que es bueno ya que se podrán apreciar de mejor manera las posibles variaciones de corriente cuando se analicen las pruebas de suciedad. Por otro lado, el voltaje del conjunto de paneles en el cual se realizarán las pruebas de suciedad es un poco menor respecto a la referencia durante casi todo el día es aproximadamente 6 volts menor lo que es aproximadamente un 2,5 % de diferencia respecto al voltaje de referencia. Respecto de esa referencia veremos si el voltaje aumenta o disminuye en las pruebas de suciedad.

A. Suciedad Caso 1

En esta prueba utilizaremos como referencia la figura 3.6 indicada en este mismo capítulo donde se ensucia un panel del string Celeste. Esta prueba se realizó el día 04 de enero del 2023 y en la figura 3.27 se puede apreciar el comportamiento de las variables eléctricas de corriente y voltaje.

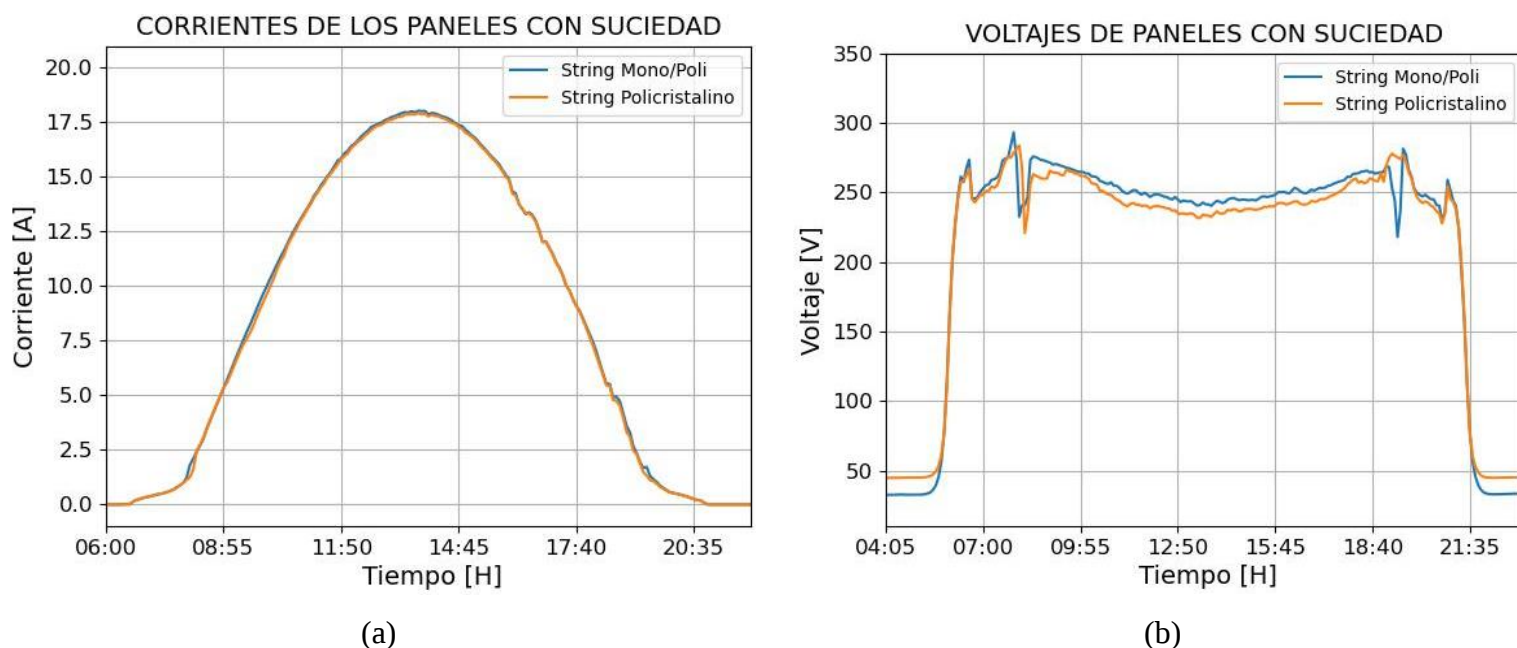


Figura 3.27 Comportamiento de las variables eléctricas Caso 1 Suciedad

Corriente en [A] de los Strings en el Caso 1 (Suci), (b) Voltaje en [V] de los Strings en el Caso 1 (Suci).

Fuente: Elaboración Propia

Al comparar las figuras 3.26 y 3.27, se observa que presentan una similitud notable en su comportamiento. La corriente generada por el string evaluado se mantiene prácticamente idéntica a la corriente de referencia, con una pequeña variación promedio de tan solo 0,01 [A]. Esto indica que la presencia de suciedad en los paneles fotovoltaicos no tiene un impacto significativo en la corriente producida por el sistema. En cuanto al voltaje, se aprecia un comportamiento similar al caso en el que el panel se encuentra limpio. En promedio, el voltaje aumenta en 0,5 [V], presentando una diferencia promedio de 5,8 [V] en comparación con el voltaje de referencia. Esta variación representa aproximadamente un 2,3% de diferencia respecto al voltaje de referencia.

Estos resultados sugieren que, a pesar de la presencia de suciedad en los paneles fotovoltaicos, la corriente generada se mantiene estable y la variación en el voltaje es mínima. Esto indica que la

suciedad no tiene un impacto sustancial en las características eléctricas del sistema, al menos en las condiciones y configuraciones evaluadas.

B. Suciedad Caso 2

Para esta prueba se toma como referencia la figura 3.9 la cual muestra un caso donde se encuentra sucia una mitad superior y la mitad inferior de dos paneles distintos ambos del mismo string.

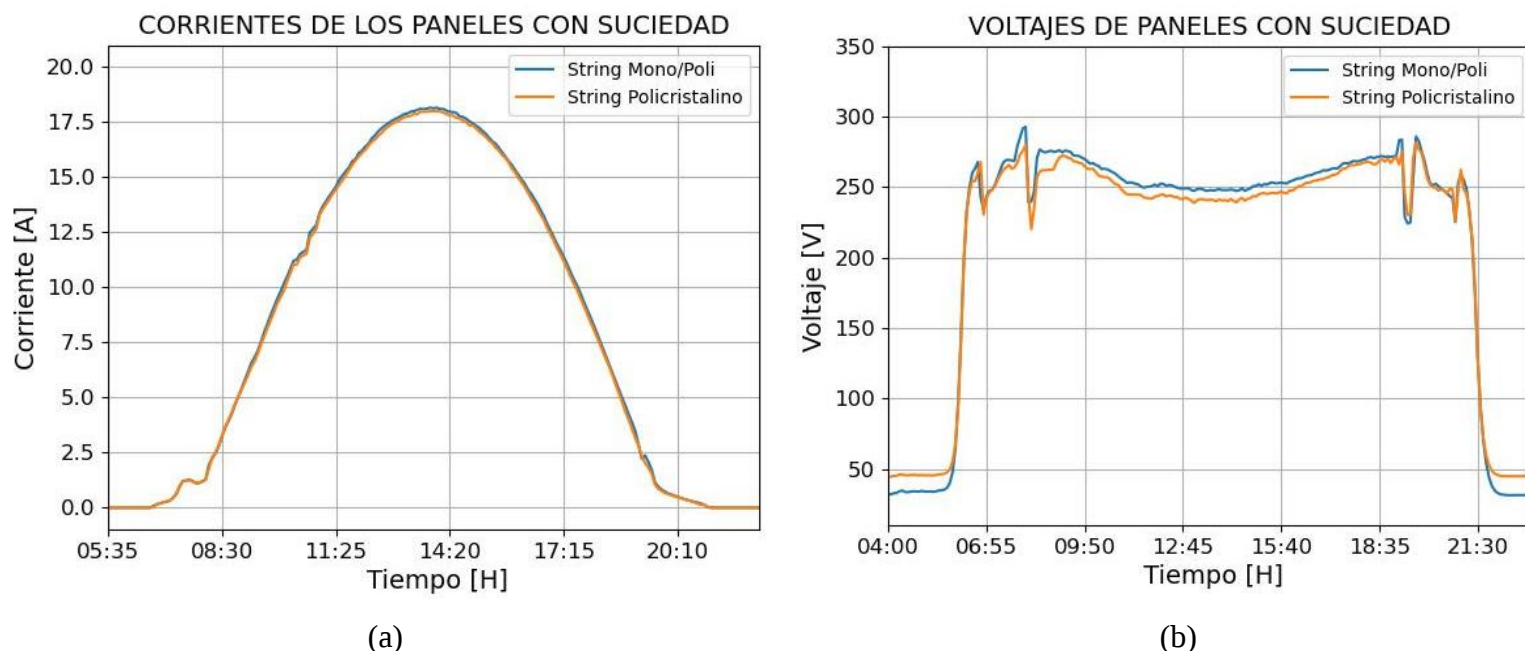


Figura 3.28 Comportamiento Variables eléctricas Caso 2 Suciedad

Corriente en [A] de los Strings en el Caso 2 (Suci), (b) Voltaje en [V] de los Strings en el Caso 2 (Suci).

Fuente: Elaboración Propia

En la figura 3.28 se muestra el comportamiento de las variables eléctricas para el caso dos, donde se evalúa la presencia de suciedad en los paneles solares. En este caso, se observa que la corriente generada por el string evaluado es en promedio 0.1 [A] menor a la corriente de referencia, presentando una variación mínima en esta variable. Esto indica que la presencia de suciedad en los paneles tiene un impacto mínimo en la corriente producida por el sistema.

En cuanto al voltaje, al compararlo con el caso sin suciedad, se aprecia una diferencia de 6,5 [V] en relación al valor de referencia. Esta diferencia representa una desviación del 2,5%, similar al caso sin suciedad donde la diferencia aproximada era de 2,43%. Estos resultados indican que la

suciedad en los paneles solares tiene un impacto leve en el voltaje generado por el sistema, pero no afecta significativamente su funcionamiento.

C. Suciedad Caso 3

Para esta prueba se toma como referencia la figura 3.11, en la que muestra el caso donde se encuentran cubiertos por suciedad la parte inferior de dos paneles FV ambos son de distintos strings (celeste y verde) esta prueba se realizó el día 15 de Enero.

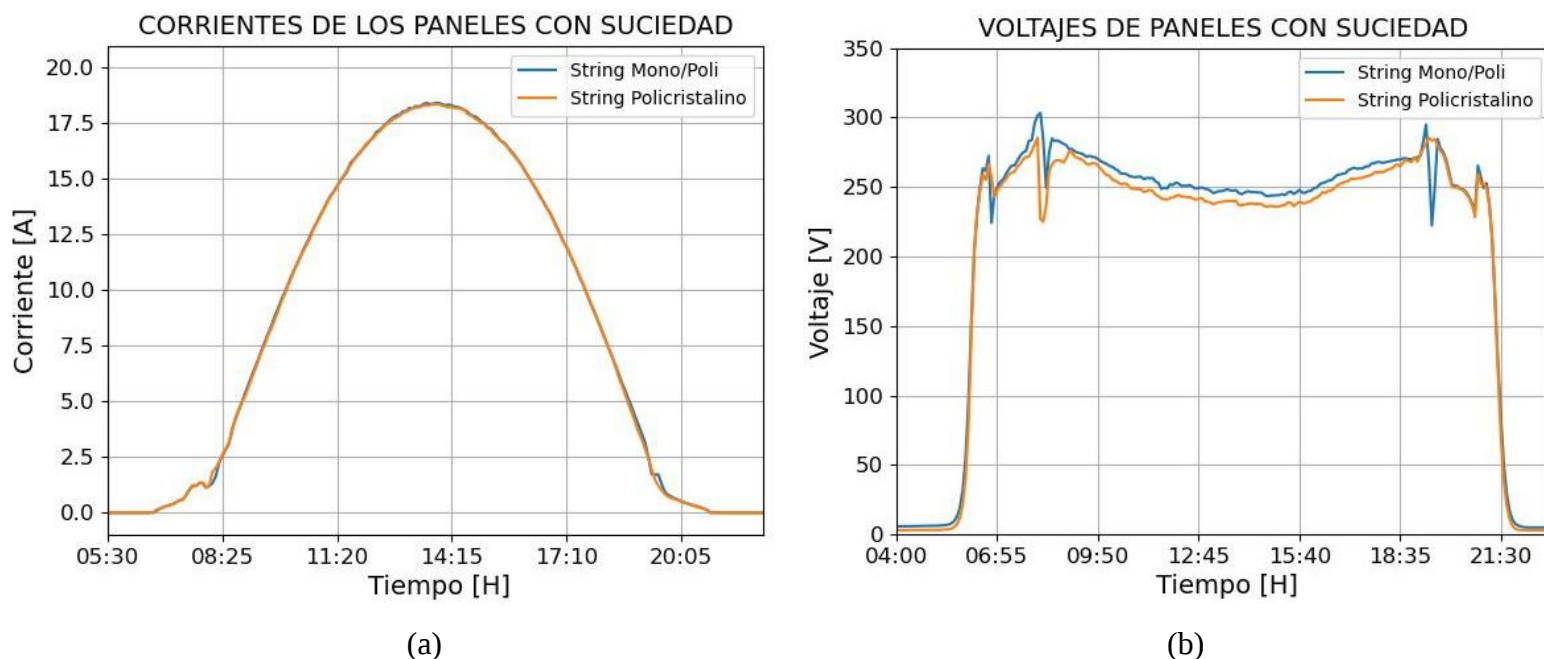


Figura 3.29 Comportamiento Variables eléctricas Caso 3

Corriente en [A] de los Strings en el Caso 3 (Suci), (b) Voltaje en [V] de los Strings en el Caso 3 (Suci).

Fuente: Elaboración Propia

En la figura 3.29 se muestra el comportamiento de las variables eléctricas en el caso analizado. La corriente se mantiene constante y prácticamente igual a la curva de referencia, sin presentar variaciones significativas solo disminuyendo en 0.03 [A] respecto a la referencia. Por otro lado, el voltaje exhibe una curva muy similar a la de referencia, pero con pequeñas variaciones. En este caso, se observa una disminución promedio de 8,2 Volts, lo que representa una reducción del 3,19 % con respecto al voltaje de referencia. Esta disminución del voltaje tiene similitud con el caso de sombreado donde a medida que mayor cantidad de paneles eran sombreados aumentaba la variación del voltaje del string. Es importante destacar que este caso presenta la mayor variación en los parámetros analizados entre los casos de suciedad estudiados.

D. Suciedad Caso 4

Para esta prueba se toma como referencia la figura 3.12 la cual muestra un caso donde se encuentran sucias las mitades superiores de dos paneles de distintos string. Esta prueba se realizó el día 16 de Enero el comportamiento de las variables eléctricas se puede ver en la figura 3.30

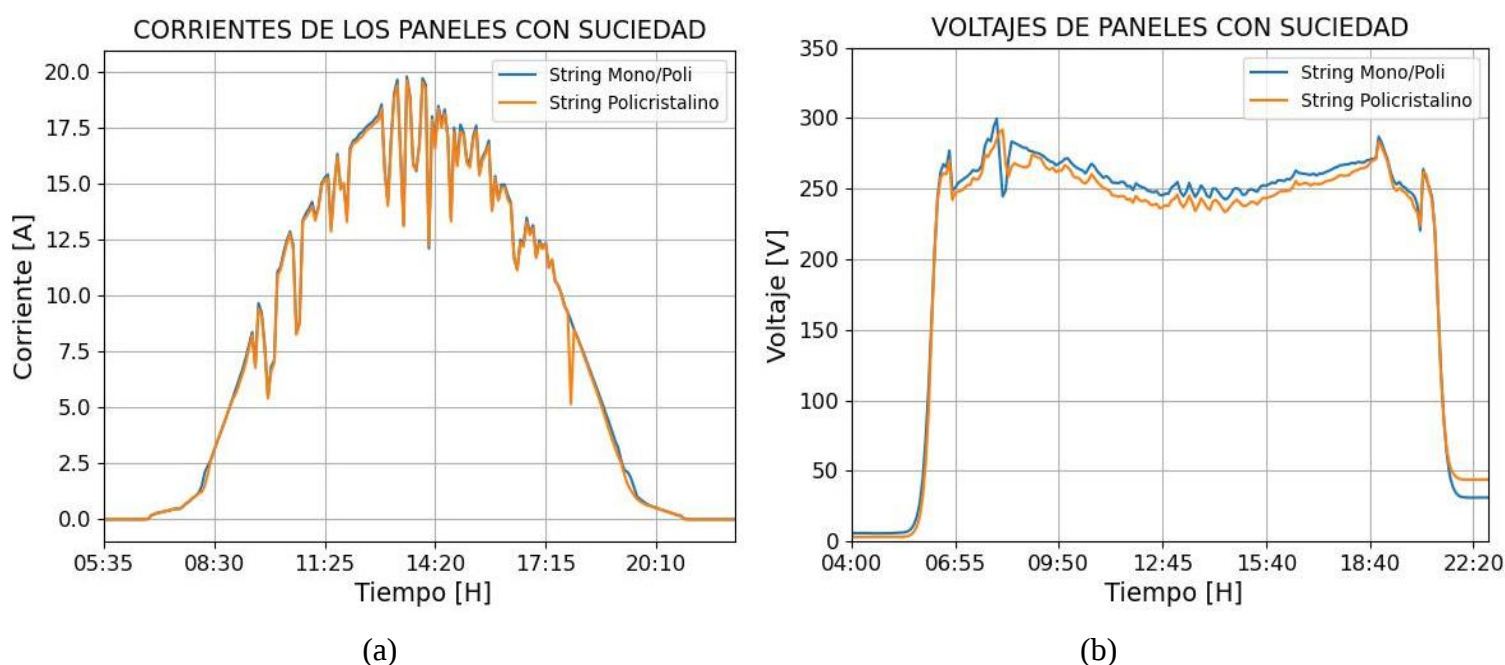


Figura 3.30 Comportamiento Variables eléctricas Caso 4

Corriente en [A] de los Strings en el Caso 4 (Suci), (b) Voltaje en [V] de los Strings en el Caso 4 (Suci).

Fuente: Elaboración Propia

En la figura 3.30 se puede observar el comportamiento de las variables eléctricas en el caso analizado. La corriente del string Policristalino varía respecto a valores de referencia, como se puede ver gráficamente en (a), donde hay una disminución de esta en casi todos los datos de corriente teniendo en promedio una diferencia de 0.11 [A] siendo la mayor variación registrada en estos casos de estudio. Por otro lado, el voltaje presenta un comportamiento similar al de referencia, especialmente si se compara con la figura 3.27. Sin embargo, al analizar los datos, se pueden apreciar pequeñas diferencias. La diferencia promedio del voltaje en relación al valor de referencia es de 7,5 Volts, lo que representa un 2,9% de variación por lo que sigue la tendencia de mientras más paneles FV presenten suciedad, mayor es la variación de voltaje.

Tabla 3.9 Resultados de Pruebas de suciedad en un panel

Caso Experimental	Cantidad total de superficie sucia del panel	Variación del voltaje respecto a la referencia	Porcentaje de Variación de voltaje	Variación de corriente respecto a la referencia	Porcentaje de Variación de Corriente
Sin Suciedad	0	6,3 [Volts]	2,43 %	0,04 [A]	0,57 %
1 panel sucio	1 panel Entero	5,8 [Volts]	2,3 %	0,085[A]	1 %
1 mitad superior sucia					
1 mitad inferior sucia del mismo String	1 panel Entero	6,5 [Volts]	2,5%	0,1 [A]	1,25 %
2 mitades inferiores sucias de distintos String	1 panel Entero	8,3 [Volts]	3,19 %	0,03 [A]	0,37 %
2 mitades superiores sucias de distintos String	1 panel Entero	7,5 [Volts]	2,92%	0,11 [A]	1,45 %

Fuente: Elaboración Propia

En la tabla 3.9 se muestra un resumen de los datos obtenidos en los cuatro casos analizados de suciedad en un panel. Se destaca que cuando se ensucia un solo panel por completo, se observa un ligero aumento en el voltaje en comparación con el día en que el panel estaba limpio, con un aumento de solo 0,5 [V] respecto al día sin suciedad en los paneles.

Es importante mencionar que en ninguno de los casos se observó un aumento en la corriente siendo los casos 2 y 4 cuando mayor disminución de corriente se pudo apreciar. En todos los casos en los que se cubrió más de un panel, equivalente a ensuciar un módulo FV completo, se observó una disminución en el voltaje en comparación con el valor de referencia. Los casos 3 y 4 presentaron las mayores diferencias, con una reducción de voltaje de 8,3 [V] y 7,5 [V], respectivamente.

3.4.3 Pruebas de suciedad en dos o más paneles Fotovoltaicos

En esta sección se analizarán las pruebas realizadas en los cuales se ensuciaban 2 o más paneles FV replicando las pruebas realizadas de sombreado por lo que los casos mostrados serán los de las figuras 3.15, 3.17, 3.19 y 3.21.

E. Suciedad Caso 5

Este caso analiza lo mostrado en la figura 3.15 donde se ensucia intencionalmente dos paneles FV de dos strings diferentes cubriendo totalmente su superficie con la combinación de tierra y agua mencionada al comienzo de este capítulo. Esta prueba fue realizada el día 05 de enero.

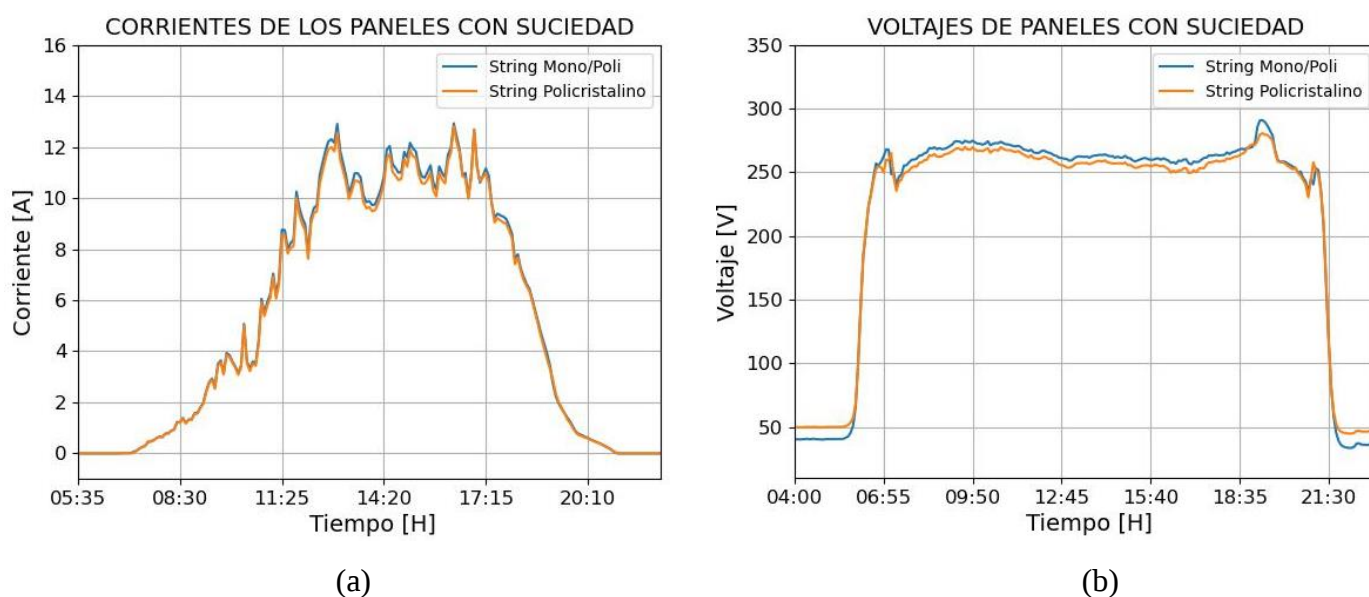


Figura 3.31 Comportamiento Variables eléctricas Caso 5,
(a) Corriente en [A] de los Strings en el Caso 5 (Suci), (b) Voltaje en [V] de los Strings en el Caso 5 (Suci).

Fuente: Elaboración Propia

En la figura 3.31 se muestra el comportamiento de las variables eléctricas del caso 5, donde se observa una ligera disminución en la corriente del string Policristalino en comparación con su referencia. En promedio, la corriente disminuyó en 0,15 [A], lo que representa una reducción promedio del 2%. Siendo la mayor disminución de corriente en comparación a todos los casos anteriores en los que se cubría la superficie de un panel FV.

Por otro lado, en cuanto al voltaje, se observa un aumento en comparación con el caso sin suciedad. La diferencia promedio de voltaje entre el caso 5 y el voltaje de referencia del string Mono/Poli es de 4,93 [V]. Esto representa una diferencia promedio del 1,86% en relación al voltaje de referencia.

F. Suciedad Caso 6

Este caso analiza lo mostrado en la figura 3.17 donde se ensucia intencionalmente dos paneles FV del mismo String en este caso el de color celeste cubriendo totalmente su superficie.

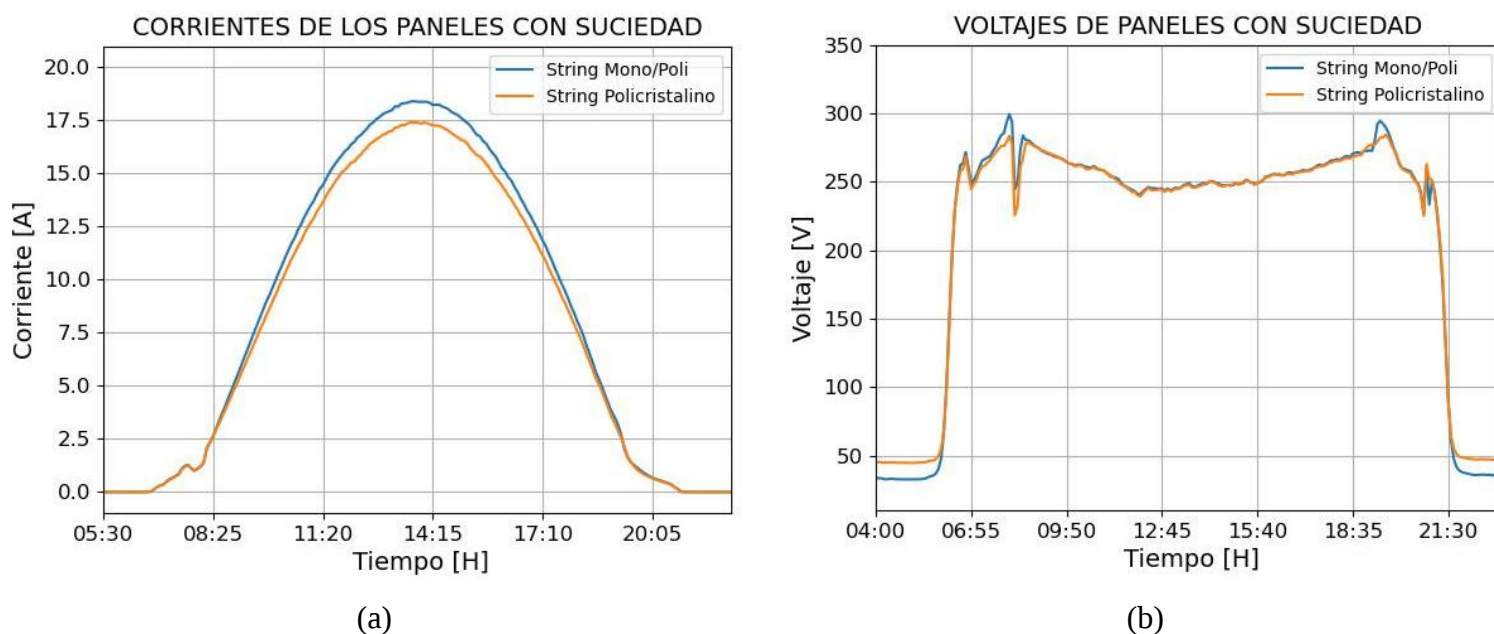


Figura 3.32 Comportamiento Variables eléctricas Caso 6,

(a) Corriente en [A] de los Strings en el Caso 6 (Suci), (b) Voltaje en [V] de los Strings en el Caso 6 (Suci).

Fuente: Elaboración Propia

En la figura 3.32 se muestra el comportamiento de las variables eléctricas del caso 6, correspondiente a las pruebas realizadas el día 18 de enero. En esta figura se aprecia un cambio notable en la reducción de la corriente y el aumento del voltaje en comparación con los valores de referencia.

En cuanto a la corriente, se observa una disminución promedio de 0,7 [A], lo que representa una diferencia del 5,73% en relación a la curva de corriente utilizada como referencia. Esta reducción en la corriente se debe a la presencia de suciedad en dos paneles del mismo string, lo que afecta la generación de corriente.

Por otro lado, el voltaje muestra un comportamiento muy similar al voltaje de referencia, pero con valores ligeramente superiores a la curva del caso sin suciedad analizado en la figura 3.26. La variación promedio del voltaje en el caso 6 es de 1,78 [V] en comparación con la señal de referencia, lo que equivale a una variación del 0,68%. Este aumento en el voltaje puede estar relacionado con la menor generación de corriente debido a la presencia de suciedad.

G. Suciedad Caso 7

Este caso es el mostrado en la figura 3.19 donde se les aplica suciedad artificial a 4 paneles 2 del string celeste y dos del string verde donde a un panel del string se le ensucia la parte superior y al otro modulo la parte inferior, esta prueba fue realizada el día 09 de enero

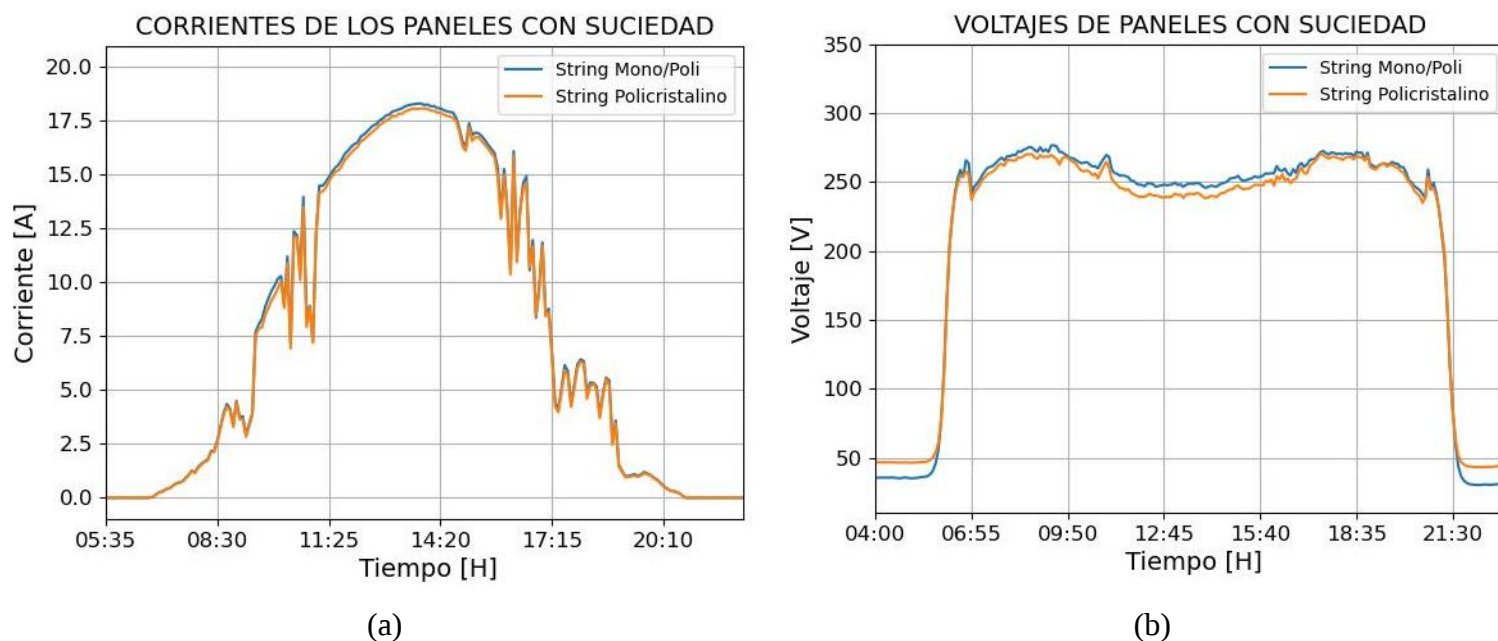


Figura 3.33 Comportamiento Variables eléctricas Caso 7,

(a) Corriente en [A] de los Strings en el Caso 7 (Suci), (b) Voltaje en [V] de los Strings en el Caso 7 (Suci).

Fuente: Elaboración Propia

En la figura 3.33 se muestra el comportamiento de las variables eléctricas del caso 7, donde se puede apreciar similitudes con la figura 3.31. En este caso, también se observa una disminución de la corriente en comparación con la curva de referencia, y el voltaje se encuentra cercano a la curva de referencia, mostrando un aumento respecto al caso sin suciedad. En el caso 7, al igual que en el caso 5, la corriente disminuye en promedio 0,15 [A], lo que representa una diferencia similar en ambos casos. Por otro lado, el voltaje presenta una diferencia de 5,17 [V] respecto al voltaje de referencia, lo que equivale a una diferencia promedio del 2%.

Estas similitudes entre los casos 5 y 7 indican que la presencia de suciedad en el panel fotovoltaico tiene un impacto similar en la generación de corriente y en el comportamiento del voltaje en ambos casos. La disminución en la corriente puede atribuirse a la obstrucción parcial de la superficie del panel debido a la suciedad, y el aumento del voltaje puede ser consecuencia de una menor carga y menor generación de corriente.

H. Suciedad Caso 8

Este caso es el mostrado en la figura 3.21 donde se les aplica suciedad artificial a 4 paneles dos del string celeste y dos del string verde cubriendo completamente la superficie de estos con la suciedad artificial. Esta prueba fue realizada el día 17 de enero

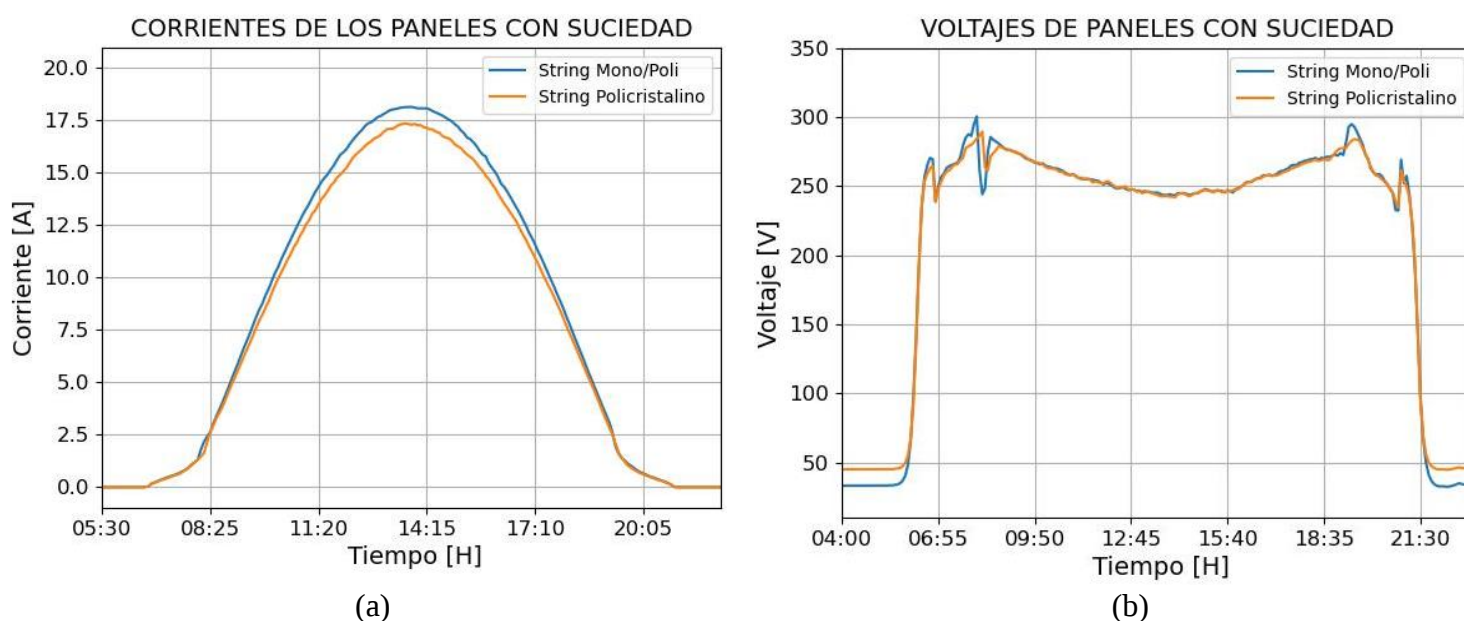


Figura 3.34 Comportamiento Variables eléctricas Caso 8,

(a) Corriente en [A] de los Strings en el Caso 8 (Suci), (b) Voltaje en [V] de los Strings en el Caso 8 (Suci).

Fuente: Elaboración Propia

En la figura 3.34, se observa el comportamiento de la corriente y el voltaje en el caso de sombreado analizado. En este caso, la corriente muestra un claro descenso en comparación con la curva de referencia, con una diferencia promedio de 0,66 [A]. Esta disminución de corriente representa aproximadamente un 5,3 % de reducción en comparación con la situación sin suciedad presente en el sistema.

Por otro lado, el voltaje en este caso presenta un aumento en comparación con la situación sin suciedad. La diferencia promedio de voltaje respecto al voltaje de referencia es de 0,89 [V], lo que indica que cuando hay suciedad en varios paneles a la vez el voltaje aumenta ya que este mismo comportamiento se pudo apreciar en los casos 5, 6, 7 y 8.

Estos resultados indican que la suciedad tiene un impacto significativo en la generación de corriente, con una disminución considerable, mientras que el voltaje experimenta un ligero aumento

en comparación al caso estudiado en que los paneles no presentaban suciedad. Este comportamiento puede deberse a la obstrucción parcial de la superficie del panel, no permitiendo que la irradiancia llegue de manera directa al sistema FV, lo que limita la generación de corriente, pero no afecta de manera significativa el voltaje.

De la tabla 3.10 podemos comentar que cuando se aplica suciedad artificial en la superficie de los paneles FV el voltaje aumenta en todos los casos independiente cuantos de estos presenten suciedad en la superficie, por otro lado la corriente presenta una disminución respecto a la curva de esta variable que es usada de referencia, siendo los casos donde se somborean más de dos paneles donde mayor disminución de esta se aprecia, en los casos 6 y 8. Disminuyendo en 0,7 [A] y 0,66 [A] respectivamente, lo que equivale a una disminución del 5,7 % y 5,3 %. Precisamente en esos casos antes mencionados es donde más aumento el voltaje, en el caso 6 tiene un aumento de 4.22 [V] respecto al caso donde no se presenta suciedad. Mientras que en el caso 8 hay un aumento de 5,11 [V] respecto a la situación sin suciedad en los paneles FV.

Tabla 3.10 Resultados de los casos de suciedad (Caso 5 – Caso 8)

Caso Experimental	Cantidad total de superficie sombreada del panel	Variación del voltaje respecto a la referencia	Porcentaje de Variación de voltaje	Variación de corriente respecto a la referencia	Porcentaje de Variación de Corriente
Sin Suciedad	0	-6 [Volts]	2.2 %	-0,04 [A]	0,57 %
2 paneles Sucios diferentes Strings	2 paneles enteros	-4,93 [Volts]	1,86%	-0,15 [A]	2%
2 paneles Sucios del mismo String	2 paneles enteros	-1,78 [Volts]	0,68 %	-0,7 [A]	5,73%
2 mitades superiores y 2 mitades inferiores sucias de diferentes String	2 paneles enteros	-5,17 [Volts]	2%	-0,15 [A]	2%
2 paneles sombreado 2 por Strings	4 paneles enteros	-0,89 [Volts]	0,34%	-0,66 [A]	5,3%

Fuente: Elaboración Propia

La tabla 3.10 proporciona un resumen de los resultados obtenidos en los diferentes casos experimentales de suciedad en los paneles fotovoltaicos. Se observa que cuando se aplica suciedad artificial en la superficie de los paneles, el voltaje tiende a aumentar en todos los casos, independientemente de la cantidad de paneles sucios. Por otro lado, la corriente presenta una disminución en comparación con la curva de referencia utilizada.

En los casos donde se ensucian dos paneles completos en diferentes strings (caso 2) o en el mismo string (caso 4), se aprecia una disminución de la corriente de 0,15 [A], lo que equivale a una reducción del 2% en promedio. En el caso 6, donde se somborean dos mitades superiores e inferiores de diferentes strings, la disminución de la corriente es de 0,7 [A], representando un - 5,73% de variación. En el caso 8, donde se ensucian dos paneles en cada string, la disminución de la corriente es de 0,66 [A], lo que corresponde a una disminución del 5,3% siendo el caso donde se registra la mayor variación de corriente.

En cuanto al voltaje, se observa un aumento en todos los casos. En el caso 2, el aumento promedio es de 4,93 [V], lo que representa un incremento del 1,86%. En el caso 4, el aumento promedio es de 1,78 [V], equivalente a un 0,68% de variación. En el caso 6, el aumento promedio es de 5,17 [V], lo que implica un incremento del 2%. En el caso 8, el aumento promedio es de 0,89 [V], correspondiendo a un 0,34% de variación respecto a la curva de referencia.

En conclusión, la suciedad en los paneles fotovoltaicos provoca una disminución en la corriente y un aumento en el voltaje en comparación con la curva de referencia. La magnitud de estas variaciones depende de la cantidad de paneles sucios y de su distribución en los distintos strings del sistema Fotovoltaico.

Lo contrario ocurre con el sombreado ya que este afecta directamente al voltaje, el cual mientras mayor superficie del panel este sombreada menor será el valor de este siendo fácil de identificar a simple vista, por lo que se espera que los algoritmos que se entrenen les sea más fácil clasificar correctamente los casos de sombreado que los de suciedad, ya que en esto últimos en varias pruebas realizadas se tuvo que analizar dato por dato, para encontrar variaciones.

Capítulo 4. Algoritmos de Clasificación

4.1. Introducción

En este capítulo se presentará un resumen sobre qué es el aprendizaje automático (machine learning) y algunos de los algoritmos de aprendizaje supervisado, en particular, los algoritmos de clasificación. Estos últimos serán utilizados en el desarrollo de esta investigación, y se explicará su funcionamiento interno, así como las diversas métricas que se emplean para evaluar su rendimiento.

4.2. Machine Learning

El estudio del Machine Learning ha ido en aumento estos últimos años, teniendo cabida en diferentes industrias y en parte de nuestra vida, y la industria eléctrica no es la excepción. Existen distintas definiciones respecto al machine learning entre las más destacadas encontramos las siguientes

Tom Michell: "un conjunto de métodos que pueden detectar automáticamente patrones en los datos y utilizarlos para hacer predicciones o tomar decisiones con mínima intervención humana".

Arthur Samuel: "el campo de estudio que le da a las computadoras la habilidad de aprender sin ser programadas explícitamente".

Andrew Ng: "la ciencia de hacer que las computadoras aprendan a partir de los datos, al igual que los humanos lo hacen, y luego utilizar esa experiencia para hacer predicciones o tomar decisiones".

El Machine learning se divide en dos grandes grupos el aprendizaje supervisado y no supervisado. Estos se diferencian en el tipo de información se le entrega al modelo, ya que el aprendizaje supervisado al algoritmo se le proporcionan las entradas y se indica la salida deseada la cual puede ser una clase o una salida deseada. De esta forma existe una subdivisión dentro de este tipo de aprendizaje supervisado, algoritmos de regresión al cual se le indica la salida deseada y existen los algoritmos de clasificación a los que a cada dato se le asigna una clase. Por otro lado, el aprendizaje no supervisado no se le indica nada y es el algoritmo es que con los datos entregados busca e identifica patrones y estructuras en los datos.

Esta investigación se centrará en el aprendizaje supervisado específicamente en los algoritmos de clasificación. Específicamente se utilizarán 4 diferentes algoritmos.

4.2.1 Algoritmo de regresión logística

La regresión logística es un algoritmo de clasificación el cual es capaz de predecir la probabilidad de una variable dependiente categórica es decir puede tomar un valor entre determinados rangos de números, en el caso específico de la regresión logística la salida “y” son datos binarios los cuales pueden ser 0 - 1. Por ende, la regresión logística entrega la probabilidad de que los valores obtenidos sean 1 o 0.

Para entregar la respuesta binaria se basa en una o más variables independientes “x”

El algoritmo de regresión logística utiliza una función llamada “sigmoid” o “función logística” la cual viene dada por la siguiente ecuación

$$g(z) = \frac{1}{1 + e^{-z}}$$

Ec. 4.1

$$0 < g(z) < 1$$

Donde $g(z)$ tendrá valores entre 0 y 1 si “z” toma un valor positivo muy grande $g(z)$ será un valor cercano a 1, por otro lado, si “z” es un valor negativo muy grande $g(z)$ tendera a 0. Esto se puede apreciar en la figura 4.1 la cual muestra el comportamiento de la función sigmoid.

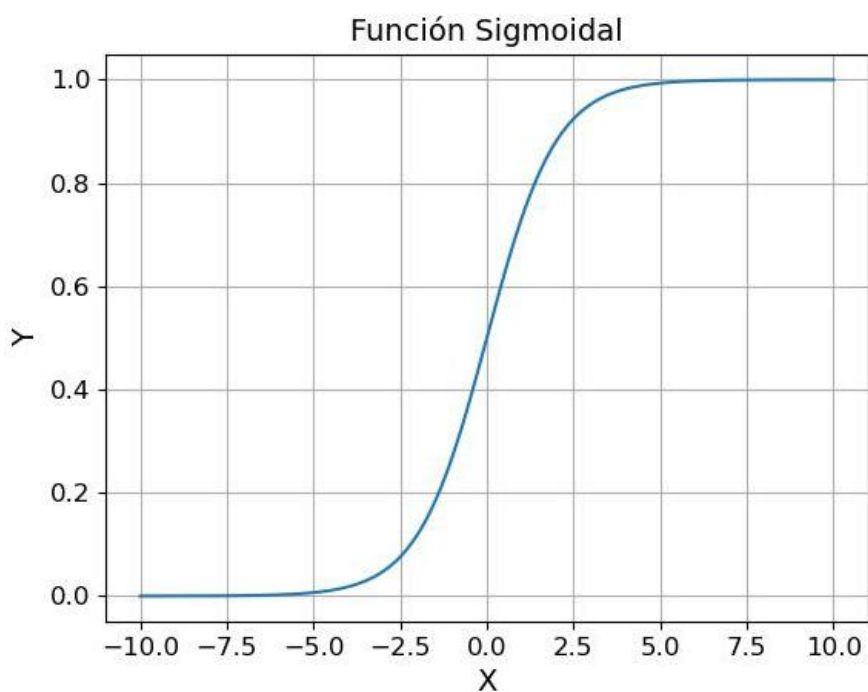


Figura. 4.1: Gráfica de la función sigmoid

Fuente: Elaboración Propia

Para poder construir el modelo se tiene que

$$z = \vec{w} \cdot \vec{x} + b \quad \text{Ec. 4.2}$$

Donde $\vec{w}$ y $\vec{x}$ son vectores y la variable b es un numero

$\vec{x}$: Es el vector que contiene las diferentes variables independientes de entrada del modelo

$\vec{w}$: son los pesos que se le asignan a cada variable de entrada en la construcción del modelo

Tanto $\vec{w}$, como b son los valores que se pueden variar para optimizar el modelo.

Con esta información podemos volver a escribir la ecuación de la función sigmoid

$$f_{\vec{w},b}(\vec{x}) = g(\vec{w} \cdot \vec{x} + b) \quad \text{Ec. 4.3}$$

Remplazando la Ec. 4.1 en Ec. 4.3

$$f_{\vec{w},b}(\vec{x}) = \frac{1}{1 + e^{-(\vec{w} \cdot \vec{x} + b)}} \quad \text{Ec. 4.4}$$

La probabilidad de que “y” sea cero o uno debe venir dada por la siguiente expresión

$$P(y = 0) + P(y = 1) = 1 \quad \text{Ec. 4.5}$$

La probabilidad que “y” sea cero o uno viene dada por las siguientes ecuaciones respectivamente

$$f_{\vec{w},b}(\vec{x}) = P(y = 0 | \vec{x}; \vec{w}, b) \quad \text{Ec. 4.6}$$

$$f_{\vec{w},b}(\vec{x}) = P(y = 1 | \vec{x}; \vec{w}, b) \quad \text{Ec. 4.7}$$

Como la función sigmoid entrega una predicción en un valor entre 0 y 1 se debe establecer un límite de decisión, para definir cuál será el resultado final. Por lo general se establece un umbral por encima del cual predices y lo usual es usar un límite en 0.5.

Si se cumple esa condición “y” será 1. En caso contrario “y” será 0

$$f_{\vec{w},b}(\vec{x}) \geq 0.5$$

$$g(z) \geq 0.5$$

$$z \geq 0.5$$

$$\vec{w} \cdot \vec{x} + b \geq 0.5$$

$$y = 1$$

Por otro lado

$$\begin{aligned}
 f_{\vec{w},b}(\vec{x}) &< 0.5 \\
 g(z) &< 0.5 \\
 z &< 0.5 \\
 \vec{w} \cdot \vec{x} + b &< 0.5 \\
 y &= 0
 \end{aligned}$$

Es importante señalar que dependiendo el conjunto de datos “x” del problema el umbral de decisión puede ser lineal y también no lineal. Ya que los valores del vector “x” pueden tener un grado de polinomio mayor a 1.

La función de costo es una función matemática que permite estimar que tan bien se ajustan el conjunto de parámetros de entrada a los datos de entrenamiento lo que ayuda a elegir los mejores parámetros de w y b para un mejor funcionamiento del algoritmo de clasificación logística.

Función de costo del algoritmo de regresión logística viene dada por la siguiente ecuación

$$J(w, b) = -\frac{1}{m} \sum_{i=1}^m \left[y^{(i)} \log \left(f_{\vec{w},b}(\vec{x}^{(i)}) \right) + (1 - y^{(i)}) \log \left(f_{\vec{w},b}(\vec{x}^{(i)}) \right) \right] \quad \text{Ec. 4.8}$$

Donde

$J(w, b)$: es la Función de Costo.

w : son los parámetros con lo que esta aprendiendo el modelo.

b : es una constante con la cual esta aprendiendo el modelo.

m : es el número de datos de entrenamiento del modelo.

$y^{(i)}$: son valores reales de la salida.

$f_{\vec{w},b}(\vec{x}^{(i)})$: es la función de la regresión logística

El objetivo con esta función es minimizar su valor durante el entrenamiento esto se hace encontrando los valores óptimos de los parámetros. Esto se realiza con el algoritmo del descenso gradiente.

El algoritmo del descenso gradiente se utiliza para minimizar la función de costo del modelo de regresión logística esto se realiza ajustando iterativamente los parámetros w y b .

$$w_j = w_j - \alpha \frac{\partial}{\partial w_j} J(w, b) \quad \text{Ec. 4.9}$$

$$b = b - \alpha \frac{\partial}{\partial b} J(w, b) \quad \text{Ec. 4.10}$$

Donde

α : es la tasa de aprendizaje que es un parámetro que se puede elegir para que el aprendizaje sea más rápido o más lento.

$\frac{\partial}{\partial w_j} J(w, b)$: es la derivada parcial de la función de coste respecto a w .

$\frac{\partial}{\partial b} J(w, b)$: es la derivada parcial de la función de coste respecto a b .

Para la correcta ejecución del algoritmo del descenso gradiente se deben actualizar ambas variables (w , b) simultáneamente.

4.2.2 Algoritmo de k vecinos más cercanos

Algoritmo de vecinos cercanos KNN o k-NN es un algoritmo clasificador que utiliza la proximidad para realizar las clasificaciones. Para ello es necesario etiquetar los datos, ya que el algoritmo predice con las etiquetas de los datos que están más cercanas al punto al cual se está evaluando utilizando un sistema de votos por mayoría. Es decir, elige dependiendo la frecuencia de las etiquetas alrededor del punto evaluado, siendo en el caso de una etiquetación binaria necesario una frecuencia superior al 50 % para definir que etiqueta clasificara (0 ó 1).

Se puede decir que el algoritmo de vecinos más cercanos simplemente almacena los ejemplos de entrenamiento etiquetados ya que este algoritmo no pasa por un proceso de entrenamiento. Este algoritmo forma parte los modelos denominados de aprendizaje perezosos. Por lo que la predicción dependerá en parte de la memoria.

La siguiente ecuación se utiliza para describir los datos de ejemplos del algoritmo

$$(x^{[i]}, y^{[i]}) \in D \quad (|D| = n) \quad \text{Ec. 4.11}$$

Donde

$x^{[i]}$: son los datos de entrenamientos las variables atributos.

$y^{[i]}$: son las etiquetas de los datos de entrenamiento.

D : es el conjunto de datos de entrenamiento.

n : es la numero de ejemplos de entrenamiento.

Los datos de entrenamiento que se está evaluando se nombra como “ i ” el cual va desde $i = 1, 2, \dots, n$.

Algo a considerar es que se debe determinar la cantidad de vecinos que se quiere considerar esto se realiza seleccionando un valor de “ k ” donde k será la cantidad de ejemplos cercanos que se consideraran para predecir que etiqueta tendrá el punto evaluado.

Una vez seleccionado la cantidad de vecinos que se considerara se debe calcular la distancia entre el punto que se está evaluando y demás puntos de entrenamiento para saber cuáles son esos vecinos y dependiendo de estos será la etiqueta que recibirá el punto evaluado.

Si bien hay varias formas de calcular la distancia las más ocupadas son:

Distancia euclidiana

Mide una línea recta entre el punto de consulta y el otro punto que se mide

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad \text{Ec. 4.12}$$

Distancia Manhattan

Mide el valor absoluto entre dos puntos

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i| \right) \quad \text{Ec. 4.13}$$

Distancia Minkowsky

Esta es la forma generalizada de las métricas de distancia euclidiana y la distancia de Manhattan

$$d(x, y) = \left(\sum_{i=1}^n |x_i - y_i| \right)^{\frac{1}{p}} \quad \text{Ec. 4.14}$$

Donde p es un parámetro del grado de la distancia

Si p = 1 se tiene la distancia euclidiana

Si p = 2 se tiene la distancia euclidiana

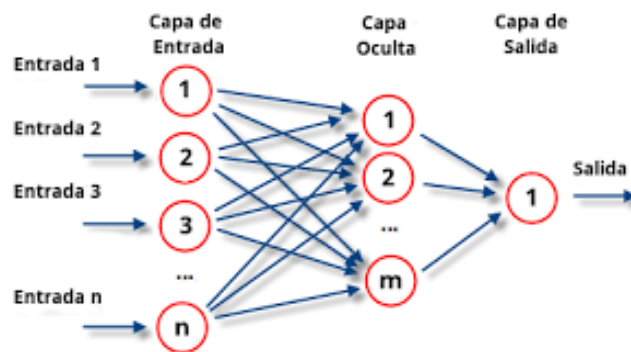
Por lo general se utiliza la Distancia Minkowsky ya que se puede adaptar para obtener los mejores resultados dependiendo el problema que se trabaje.

4.2.3 Redes Neuronales

Las redes neuronales artificiales (RNA) Son algoritmos los cuales en sus inicios que trataban de emular el comportamiento de las neuronas del cerebro humano. Este tipo de algoritmo funciona bastante bien con datos tanto estructurados y no estructurados es ampliamente utilizado para trabajos con imágenes, datos de voz y datos numéricos.

Una red neuronal está compuesta por varias neuronas que a su vez forman capas (Layers), el conjunto de Layers forman las redes neuronales

Figura 4.2: Estructura de redes neuronales



Fuente: {5} Web (Bibliografía)

Como se muestra en la figura 4.2 las redes neuronales tienen una capa de entrada, una capa de salida y una capa oculta. Las capas ocultas son las que están entre la capa de entrada y la capa de salida, una red neuronal debe tener a lo menos una capa oculta y no tiene un limite en la cantidad de estas, por eso la imagen muestra que las capas ocultas van desde 1 a m.

Los valores de entrada vienen dados por un vector $\vec{x}$ (valores de la capa de entrada) la capa oculta luego de recibir los datos de la capa de entrada, realiza cálculos y genera un vector de activación $\vec{a}$, el cual llega a la capa de salida que finalmente genera un vector $\vec{a}$ de activación final o predicción final.

Tanto las capas ocultas, como la capa de salida pueden estar compuesta por una gran cantidad de neuronas, esto dependerá de que tipo de problema se esté abordando. Cada una de estas capas debe tener una función de activación.

La función de activación es una función matemática que se le aplica a la salida de cada capa de la red neuronal. Esta función determina que información se entrega a la siguiente capa o la salida del algoritmo. En el caso de tener un problema de clasificación binario, se recomienda utilizar en la capa de salida la función de activación sigmoid, esta fue explicada anteriormente en el algoritmo de regresión logística.

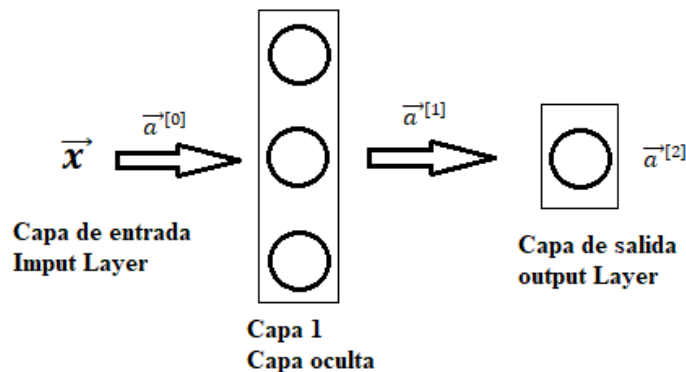


Figura. 4.3: Comportamiento interno de red Neuronal

Fuente: Elaboración Propia

En la figura 4.3 se aprecia una red neuronal de 2 capas, 1 capa oculta y la capa de salida las cuales tienen 3 y 1 neurona respectivamente.

Considerando que tienen una función de activación sigmoid, cada una de las neuronas presentes en la figura será una pequeña función logística por lo que tendrán valores de w y b la representación será la siguiente:

Recordando las ecuaciones de la función sigmoid de la ecuación 4.1

$$g(z) = \frac{1}{1 + e^{-z}}$$

Donde z es

$$z = \vec{w} \cdot \vec{x} + b$$

La ecuación general para representar los valores del vector $\vec{a}$ será

$$a_j^{[l]} = g(\vec{w}_j^{[l]} \cdot \vec{a}^{[l-1]} + b_j^{[l]}) \quad \text{Ec. 4.15}$$

Donde:

l : es el numero de la capa la cual se está describiendo.

j : es el numero de la neurona de la capa que se está describiendo

$\vec{a}^{[l-1]}$: es el vector de datos que se recibe de la salida de la capa anterior por eso es $l - 1$.

La primera neurona de la capa 1

$$a_1^{[1]} = g(w_1^{[1]} \cdot \vec{a}^{[0]} + b_1^{[1]})$$

La segunda neurona de la capa 1

$$a_2^{[1]} = g(w_2^{[1]} \cdot \vec{a}^{[0]} + b_2^{[1]})$$

La tercera neurona de la capa 1

$$a_3^{[1]} = g(w_3^{[1]} \cdot \vec{a}^{[0]} + b_3^{[1]})$$

Ahora el vector de activación $\vec{a}^{[1]}$ que se ve en la figura contendrá los valores de $a_1^{[1]}$, $a_2^{[1]}$ y $a_3^{[1]}$.

$$\vec{a}^{[1]} = \begin{bmatrix} a_1^{[1]} \\ a_2^{[1]} \\ a_3^{[1]} \end{bmatrix} \quad \text{Ec. 4.16}$$

4.3. Funciones de Activación

4.3.1 Función de activación ReLU

Si para un determinado problema la salida de este debe ser un valor positiva, la función de activación que se debe utilizar es unidad lineal rectificada (ReLU) esta es una función no lineal que se representa con la siguiente ecuación:

$$ReLU(z) = g(z) \quad \text{Ec. 4.17}$$

$$g(z) = \max(0, z)$$

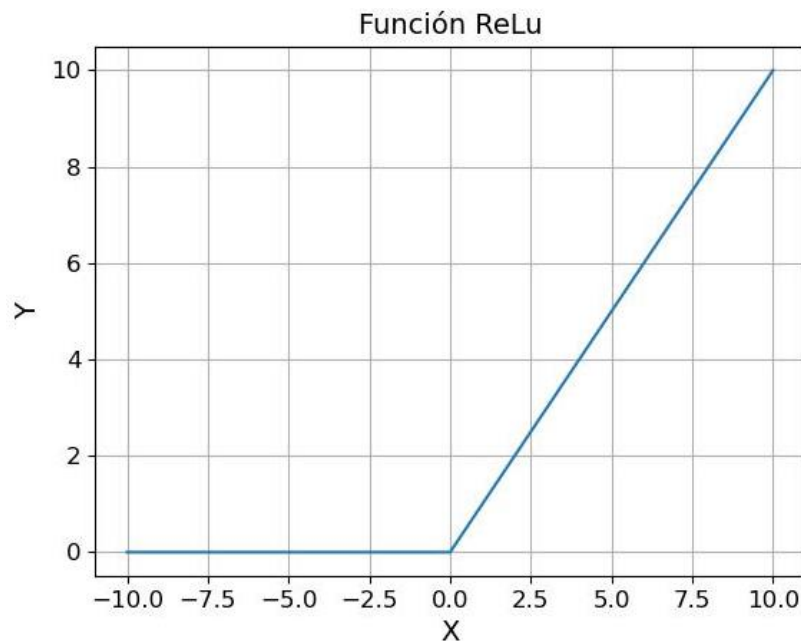


Figura 4.4: Gráfica de Función de activación ReLU

Fuente: Elaboración Propia

Si z es cualquier valor menor a 0

$$z < 0$$

$$g(z) = 0$$

Si z es cualquier valor positivo mayor o igual a cero

$$z \geq 0$$

$$g(z) = z$$

En la figura 4.4 se muestra cómo se comporta la función de activación ReLU la que se puede interpretar de mejor manera con las ecuaciones expuestas.

4.3.2 Función de activación Leaky ReLU

Una variante de la función ReLU es leaky ReLU unidad lineal rectificada con escape, la diferencia entre ambas es que leaky ReLU tiene un valor de pendiente que hace que los datos menores a 0 sean distinto a cero.

Esta se representa con la siguiente ecuación

$$\begin{aligned} \text{Leaky ReLU}(z) &= g(z) \\ g(z) &= \max(az, z) \end{aligned} \quad \text{Ec. 4.18}$$

Donde es una a constante que puede tomar un valor entre 0.01 y 0.3

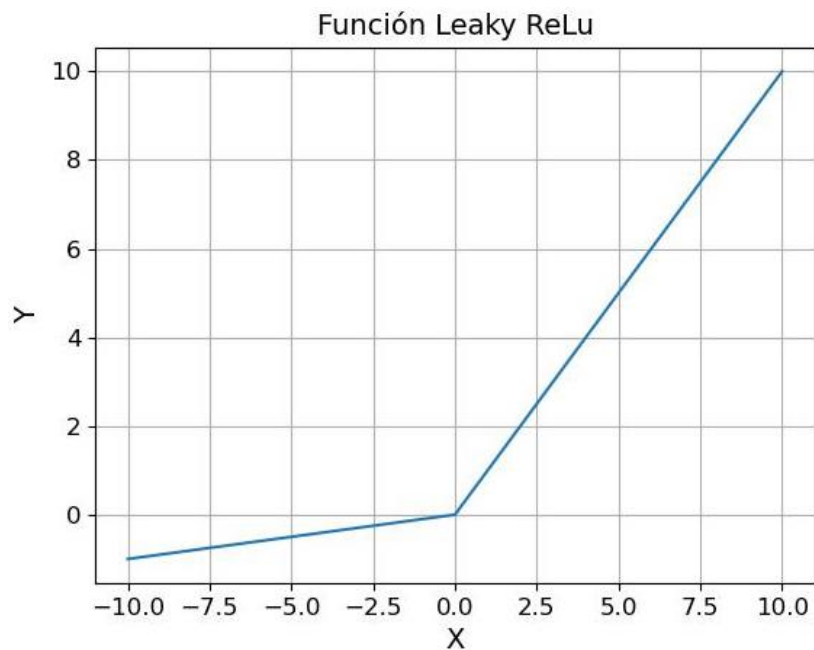


Figura 4.5: Gráfica de Función de activación leaky ReLU

Fuente: Elaboración Propia

En la Figura 4.5 se muestra cómo funciona la función leaky ReLU teniendo $a = 0.1$
Si z es cualquier valor menor a 0

$$z < 0$$

$$g(z) = az$$

Si z es cualquier valor positivo mayor o igual a cero

$$z \geq 0$$

$$g(z) = z$$

4.3.3 Función de activación Lineal

Función de activación lineal también llamada “no función de activación” esta es una función matemática que no realiza ninguna operación, no tiene un efecto no lineal a la salida de la red neuronal. Puede producir cualquier valor numérico continuo positivo y negativo. Su definición matemática es

$$g(z) = z$$

Ec. 4.19

$$z = \vec{w} \cdot \vec{x} + b$$

$$g(z) = \vec{w} \cdot \vec{x} + b$$

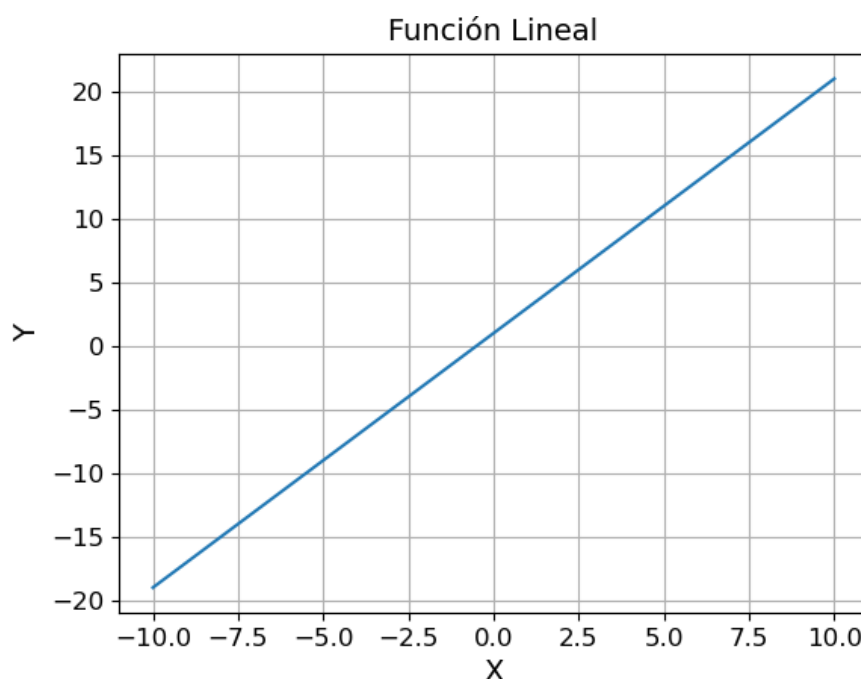


Figura 4.6 Gráfica de Función de activación Lineal

Fuente: Elaboración Propia

La Figura 4.6 muestra como es el comportamiento de la función lineal.

Ya revisadas diferentes funciones de activación para las redes neuronales la correcta implementación de estas resulta fundamental para un correcto funcionamiento del algoritmo.

Algoritmo de clasificación: si se trabaja con una salida binaria lo recomendable es utilizar la función de activación ReLU en las capas ocultas o Leaky ReLU si se conoce que las variables de entrada tienen valores negativos, finalmente en la capa de salida lo ideal es utilizar la función de activación sigmoid.

Algoritmo de regresión: lo recomendable es utilizar la función de activación ReLU en las capas ocultas o Leaky ReLU si se conoce que las variables de entrada tienen valores negativos finalmente en la capa de salida dependerá si la salida deseada tiene valores positivos y negativos lo recomendable sería utilizar la función de activación lineal, si la salida son valores netamente positivos la función de activación a utilizar será ReLU.

4.3.4 Función de activación Softmax

Esta función de activación es una generalización de la regresión logística pero adaptada a multiclases es decir más de dos posibles clasificaciones.

Recordando la salida de la regresión logística tenía dos posibilidades (1 ó 0)

$$g(z)$$

$$z = \vec{w} \cdot \vec{x} + b$$

La probabilidad de que la salida sea 1

$$a_1 = g(z) = \frac{1}{1 + e^{-z}} = P(y = 1|\vec{x}) \quad \text{Ec. 4.20}$$

La probabilidad de que la salida sea 0

$$a_2 = 1 - a_1 = P(y = 0|\vec{x}) \quad \text{Ec. 4.21}$$

Si generalizamos esto a una salida para 3 posibles valores de salida tendremos ($y = 0, 1, 2$)

$$z_1 = \vec{w}_1 \cdot \vec{x} + b_1$$

$$z_2 = \vec{w}_2 \cdot \vec{x} + b_2$$

$$z_3 = \vec{w}_3 \cdot \vec{x} + b_3$$

La probabilidad de que la salida $y = 0$ viene dada por la ecuación 4.22

$$a_1 = \frac{e^{z_1}}{e^{z_1} + e^{z_2} + e^{z_3}} = P(y = 0|\vec{x}) \quad \text{Ec. 4.22}$$

La probabilidad de que la salida $y = 1$ viene dada por la ecuación 4.23

$$a_2 = \frac{e^{z_2}}{e^{z_1} + e^{z_2} + e^{z_3}} = P(y = 1|\vec{x}) \quad \text{Ec. 4.23}$$

La probabilidad de que la salida $y = 2$ viene dada por la ecuación 4.23

$$a_3 = \frac{e^{z_3}}{e^{z_1} + e^{z_2} + e^{z_3}} = P(y = 2|\vec{x}) \quad \text{Ec. 4.24}$$

La fórmula general de la función de activación Softmax para N posibles salidas será

$$a_j = \frac{e^{z_j}}{\sum_{k=1}^N e^{z_k}} \quad \text{Ec. 4.25}$$

Donde z_j es una adaptación de la ecuación 4.2

$$z_j = \vec{w}_j \cdot \vec{x} + b_j \quad \text{Ec. 4.26}$$

Por otro lado “j” es el numero de la clase que se está evaluando, y k es el número de todas las posibles clases

Ecuación general para N posibles clasificaciones

$$a_n = \frac{e^{z_n}}{e^{z_1} + \dots + e^{z_n}} = P(y = N|\vec{x}) \quad \text{Ec. 4.27}$$

Las perdidas serán

$$Loss(\vec{a}, y) = \begin{cases} -\log a_1 & si \ y = 1 \\ -\log a_2 & si \ y = 2 \\ \vdots & \vdots \\ -\log a_n & si \ y = N \end{cases} = \begin{cases} -\log \frac{e^{z_1}}{e^{z_1} + e^{z_2} + e^{z_3}} & si \ y = 1 \\ -\log \frac{e^{z_2}}{e^{z_1} + e^{z_2} + e^{z_3}} & si \ y = 2 \\ \vdots & \vdots \\ -\log \frac{e^{z_n}}{e^{z_1} + \dots + e^{z_n}} & si \ y = N \end{cases} \quad \text{Ec. 4.28}$$

La Función de costo es la suma de todas las perdidas, por lo que se debe considerar lo siguiente

$$1\{y == j\} == \begin{cases} 1, & si \ y == j \\ 0, & en \ otro \ caso \end{cases} \quad \text{Ec. 4.29}$$

Si la clasificación de la salida es igual a la que clasificación que se está evaluando, se multiplica por 1, en caso de que la clasificación no sea la misma que se evalúa se multiplica por 0.

Esto implica que la función de coste solo aumentara si la salida real concuerda con la clase que se está evaluando.

$$J(w, b) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^N 1\{y^{(i)} == j\} \log \frac{e^{z_j^{(i)}}}{\sum_{k=1}^N e^{z_k^{(i)}}} \right] \quad \text{Ec. 4.30}$$

Donde

i : es el dato que se esta evaluando de todos los datos de entrada

j : es el numero de la clase que se está evaluando

k : son todas las posibles clases

Es necesario evaluar el rendimiento de los algoritmos que hemos mencionado, como la regresión logística y las redes neuronales. Esta evaluación es fundamental para determinar si su desempeño es adecuado, y para ello se emplean métricas de rendimiento. Cabe destacar que las métricas utilizadas varían según el tipo de algoritmo que se esté evaluando.

4.4. Métricas de rendimiento de algoritmos de clasificación

Matriz de confusión: es una de las métricas más intuitivas que se utiliza para encontrar la precisión y exactitud del modelo se puede utilizar para casos de estudio donde la salida puede ser dos o más clases. La matriz de confusión es una tabla en la cual se puede ver el número de veces que el modelo entrenado clasifico correctamente e incorrectamente cada clase.

Verdaderos Positivos (VP): son los casos donde los datos de entrenamiento son verdaderos y la predicción del modelo también es verdadera (1 - 1).

Verdaderos Negativos (VN): son los casos donde los datos de entrenamiento son falsos y la predicción del modelo también indica falso (0 - 0).

Falsos Positivos (FP): son los casos donde los datos de entrenamiento son falsos y la predicción del modelo es verdadera (0 - 1).

Falsos Negativos (FN): son los casos donde los datos de entrenamiento son verdaderos y la predicción del modelo indica que es falso (1 - 0).

Tabla 4.1: Distribución Matriz de confusión

	Predicción		
Datos de Entrenamiento		Positivo	Negativo
	Positivo	Verdaderos Positivos	Falsos Negativos
	Negativo	Falsos Positivos	Verdaderos Negativos

Fuente: Elaboración Propia

La diagonal de la matriz de confusión indicara los casos donde el modelo predijo bien mientras que los demás valores de la matriz serán la cantidad casos donde el modelo no predijo correctamente. Cabe mencionar que la tabla trazada es de un problema de clasificación con solo dos posibles clases de salida.

Exactitud (Accuracy): es número de predicciones realizadas correctamente por la modelo dividida en el número total de predicciones realizadas.

$$Exactitud = \frac{VP + VN}{VP + FP + FN + VN} \quad \text{Ec. 4.31}$$

Precisión (Precision): Con esta métrica se evalúa el desempeño de las predicciones positivas la que nos indicara el porcentaje de predicciones verdaderas que son correctas.

$$Precisión = \frac{VP}{VP + FP} \quad \text{Ec. 4.32}$$

Sensibilidad (Recall): Es el número de predicciones positivas correctas respecto del número real de salidas verdaderas del conjunto de entrenamiento.

$$Sensibilidad = \frac{VP}{VP + FN} \quad \text{Ec. 4.33}$$

Especificidad (Specificity): Con esta metrica se evalúa el desempeño de las predicciones negativas correctamente realizada respecto al total de casos donde la salida es negativa.

$$Especificidad = \frac{VN}{VN + FP} \quad \text{Ec. 4.34}$$

Valor F1 (f-Score): Esta metrica combina la precisión y la sensibilidad haciendo un balance entre ellas teniendo una puntuación que representa a ambas métricas.

$$\text{Valor F1} = \frac{2 * \text{Precisión} * \text{Sencibilidad}}{\text{Precisión} + \text{Sencibilidad}} \quad \text{Ec. 4.35}$$

Para los casos de clasificación multiclase las métricas de rendimiento son distintas, la única metrica de las antes vistas que se mantiene es **Precision (Accuracy)** y la **Matriz de Confusión** las demás métricas se presentan a continuación

Precision por clase: Es la precisión del modelo para del modelo por cada clase. Se calcula dividiendo el número de verdaderos positivos (VP) de la clase i entre la suma de TP y falsos positivos (FP) de la clase i.

$$\text{Precision clase}_i = \frac{\text{VP}_i}{\text{VP}_i + \text{FP}_i} \quad \text{Ec. 4.36}$$

Recall por clase: Es la proporción de casos positivos que fueron identificados correctamente por el modelo. Se obtiene dividiendo el número de verdaderos positivos de la clase “i” TP_i entre la suma de los verdaderos positivos de la clase y los falsos negativos de la misma clase “i” NF_i

$$\text{Recall clase}_i = \frac{\text{TP}_i}{\text{TP}_i + \text{NF}_i} \quad \text{Ec. 4.37}$$

F1-Score por clase: Es la media armónica de la precisión y el Recall por cada clase su ecuación.

$$\text{F1 Score clase i} = \frac{2 * \text{Precision clase}_i * \text{Recall clase}_i}{\text{Precision clase}_i + \text{Recall clase}_i} \quad \text{Ec. 4.38}$$

Con todas estas métricas de rendimiento explicadas se puede proceder a evaluar los distintos tipos de algoritmos que se utilizaran para clasificar los casos de suciedad y sombreado que se presentaran en el siguiente capítulo, pasando desde clasificadores binarios a clasificadores multiclase. Es importante mencionar que depende mucho la distribución de las clases en los datos con los que se entrenaran los modelos, para elegir cual es la mejor metrica para evaluarlos para ellos se utilizaran todas estas métricas para comprobar de forma experimental cual es mejor en determinados casos.

Capítulo 5. Algoritmos de Clasificación para detección de sombreado y suciedad

5.1. Introduccion

Un proyecto de Machine Learning cuenta con diferentes etapas, se comienza con la recolección de datos, para construcción de la base de datos los cuales posteriormente se preparan eliminando los datos que no aportan información, luego se selecciona el o los algoritmos que se van a utilizar, se entrena el modelo, acto seguido se evalúa con métricas que son distintos dependiendo que tipo de modelo se esté utilizando, se configuran los parámetros para un mejor rendimiento y finalmente se utiliza para realizar predicciones, seguido de un aprendizaje continuo para el modelo elegido.

Para este proyecto se tiene como objetivo entrenar modelos que sean capaces de identificar diferentes condiciones en los paneles fotovoltaicos como se indicó anteriormente es fundamental poder identificar si un sistema fotovoltaico presenta sombreado ya sea parcial o permanente y de igual manera detectar la suciedad en los módulos fotovoltaicos, para esto fue necesario crear diferentes bases de datos.

5.2. Obtencion de datos

La obtención de datos para este estudio se llevó a cabo utilizando las dependencias de la Microred UCSC. Una Microred se define como un sistema eléctrico a pequeña escala que puede abastecer la demanda de cargas, operar de manera coordinada con la red eléctrica y tener la capacidad de almacenar energía y operar de forma aislada de la red.

En el caso específico de la Microred UCSC, esta cuenta con una potencia instalada de 43,3 [kWp] y opera como una MicroRed On-grid, lo que significa que trabaja coordinadamente con la red eléctrica. Sin embargo, una desventaja de este tipo de operación es que si se produce un corte en el suministro eléctrico general, la Microred se desconectará también.

Las pruebas realizadas en el capítulo 3 tienen un doble propósito: por un lado, observar el comportamiento de las variables eléctricas en situaciones de sombreado y suciedad; y por otro, construir bases de datos que serán utilizadas para entrenar los algoritmos de clasificación.

5.3. Datos Utilizados

En el proceso de entrenamiento del modelo, se seleccionaron cuidadosamente las variables ambientales y eléctricas que se considera que tienen un impacto significativo en el funcionamiento de los sistemas eléctricos, específicamente en relación al sombreado y la suciedad. Las variables eléctricas elegidas incluyen la corriente continua, el voltaje y la potencia generada en el string 1 de la configuración de paneles solares.

Las variables seleccionadas para el modelo son las siguientes:

- Corriente CC String 1 (Fronius Symo (Poli)): Esta variable representa la corriente continua generada en el string 1 de la configuración de paneles solares. Se mide en Amperios (A).
- Potencia CC String 1 (Fronius Symo (Poli)): Esta variable indica la potencia en corriente continua generada por el string 1. Su unidad de medida es en Vatios (W).
- Tensión CC String 1 (Fronius Symo (Poli)): Esta variable representa el voltaje en corriente continua del string 1 y se mide en Voltios (V).

Además de las variables eléctricas, se consideraron las siguientes variables ambientales:

- Temperatura ambiente: Se obtiene a través de un termostato instalado en la Microred y se registra en grados Celsius (°C).
- Lluvia: Esta variable indica la cantidad de agua precipitada en forma de lluvia en un intervalo de tiempo determinado. Se registra cada cinco minutos y su unidad de medida es en milímetros (mm).
- Dirección del Viento: Representa la dirección del viento y se expresa en grados, utilizando como origen el norte. Los datos correspondientes a esta variable fueron obtenidos de la estación meteorológica de la UCSC.
- Velocidad del Viento: Se refiere a la velocidad del viento y se mide en metros por segundo (m/s). Los datos correspondientes a esta variable también fueron obtenidos de la estación meteorológica de la UCSC.
- Humedad Relativa: Es la medida de la cantidad de vapor de agua presente en el aire, expresada en porcentaje (%HR). Los datos correspondientes a esta variable fueron obtenidos de la estación meteorológica de la UCSC.

- Temperatura Celda (Policristalina): Representa la temperatura de las celdas del conjunto de paneles solares policristalinos. Se obtiene mediante un termómetro adhesivo y se registra en grados Celsius ($^{\circ}\text{C}$).
- Radiación (Sensor de Irradiación): Esta variable indica la radiación solar que reciben los paneles solares y se registra en Vatios por metro cuadrado (W/m^2). La medición se realiza mediante un sensor de radiación instalado en la Microred.

Estas variables fueron seleccionadas debido a su influencia en el comportamiento de los paneles solares y se utilizaron en el entrenamiento del modelo para lograr clasificar la condición de los paneles en relación al sombreado y la suciedad.

La correlación de las variables de las diferentes bases de datos se puede encontrar en el anexo 3.

Tabla 5.1: Tabla Resumen de las Variables a utilizar en los algoritmos de clasificación

Variables		
X_1	=	Corriente CC
X_2	=	Potencia CC
X_3	=	Tension CC
X_4	=	Temperatura ambiente
X_5	=	Lluvia
X_6	=	Dirección del Viento
X_7	=	Velocidad del Viento
X_8	=	Humedad Relativa
X_9	=	Temperatura Celda
X_{10}	=	Radiación

Fuente: Elaboración Propia

5.3.1 Base de datos Sombreado

La base de datos de los paneles fotovoltaicos bajo condiciones de sombreado se realizó desde el día 4 al 17 de octubre del 2022, la cual tiene una cantidad de 4032 filas y 12 Columnas. Para poder utilizar estos datos para entrenar los modelos es necesario realizar una limpieza de la base de datos eliminando las filas y/o columnas que no entreguen información. En este caso al evaluar un sistema fotovoltaico se da la condición que la información que se obtiene cuando no hay radiación es irrelevante por lo que es información que se debe eliminar. Lo anteriormente mencionado se puede observar en la figura 5.1 donde se pueden apreciar que hay varias filas que no aportan información siendo estas las de las horas de la noche donde no hay producción de energía.

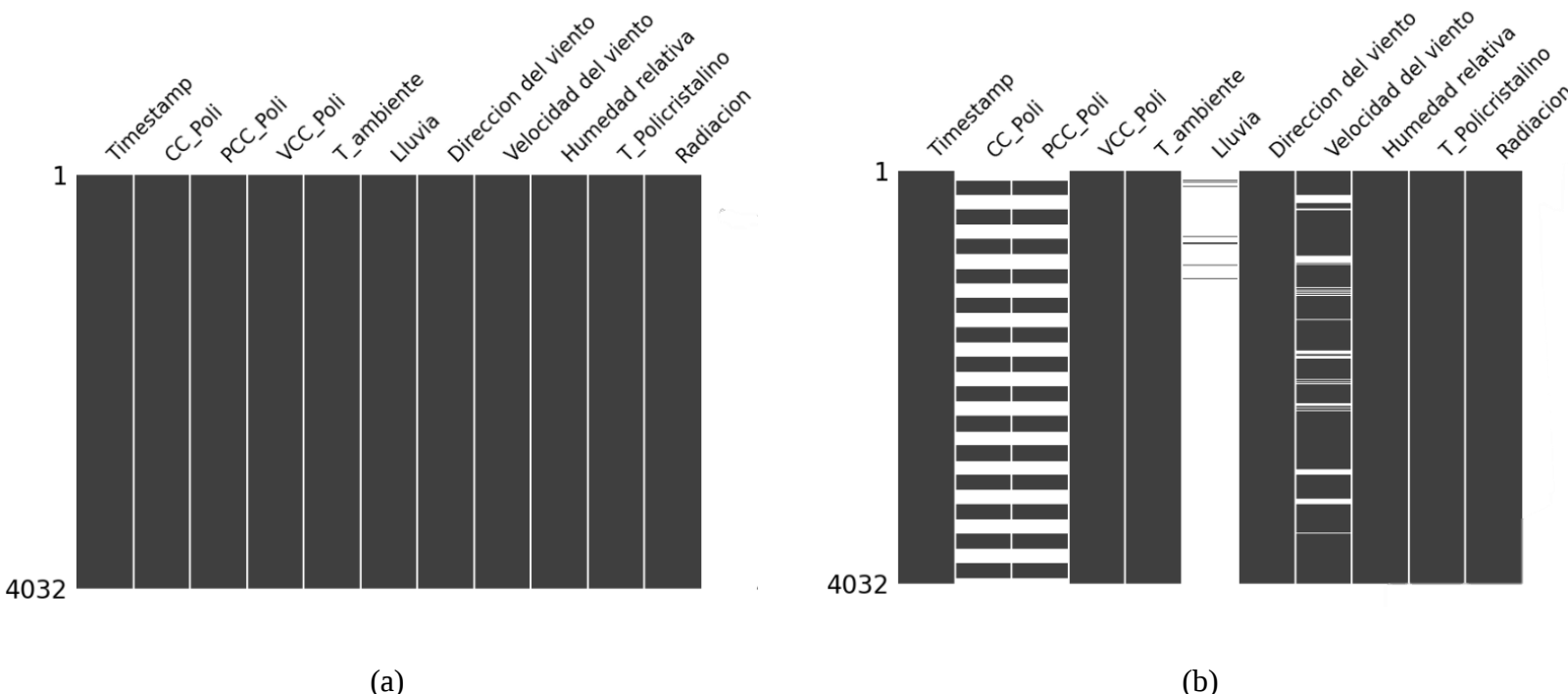


Figura 5.1 Proceso de limpieza de datos base de datos de sombreado,
(a) Base de Datos antes de la limpieza (b) Base de datos Durante la limpieza

Fuente: Elaboración Propia

De la imagen 5.1 (b) se muestran espacios en blanco lo que representa datos que son nulos (NaN). En esta limpieza lo que se hace es eliminar las filas con datos que son cero en las columnas corriente y potencia. Ya que son datos que no aportan información del comportamiento de los paneles FV. Posterior a la limpieza de los datos, la cantidad de datos de con los que se entrenara el modelo son 2095 filas y 11 columnas esto se puede apreciar en la figura 5.2.

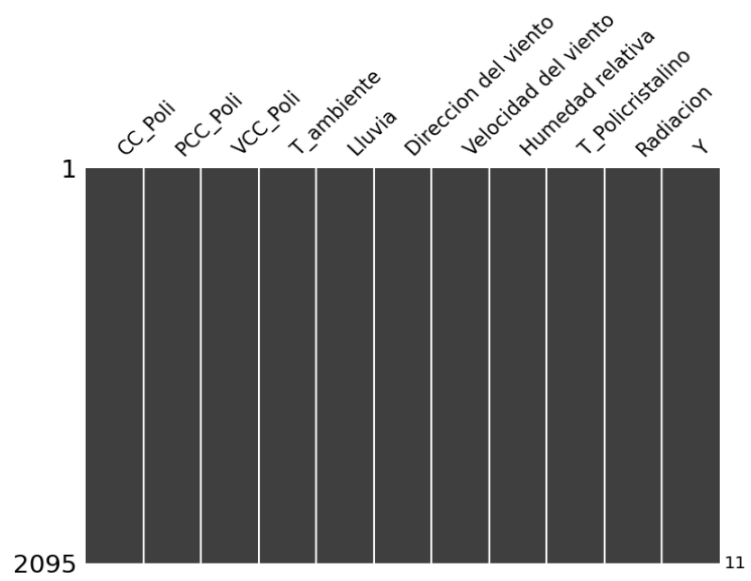


Figura 5.2 Base de Datos de sombreado posterior a la limpieza de los datos

Fuente: Elaboración Propia

En la figura 5.2 se puede apreciar que la cantidad de datos que si aportan información al futuro modelo era un poco más de la mitad del total de datos obtenidos, lo cual muestra la importancia de realizar una buena limpieza a las bases de datos.

Tabla 5.2 Clasificación de base de datos de Sombreado

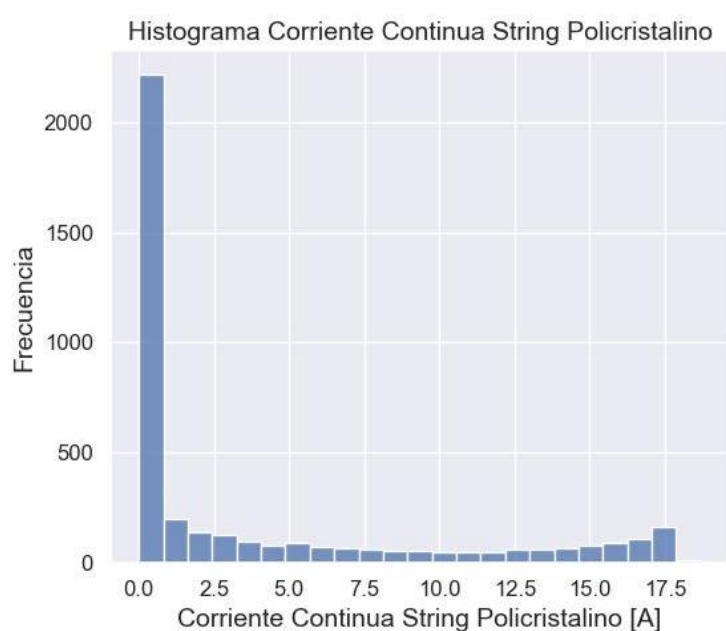
Tipo de prueba	Clases “Y”
Panel/es Limpio sin suciedad	0
Panel/es Sombreado	1

Fuente: Elaboración Propia

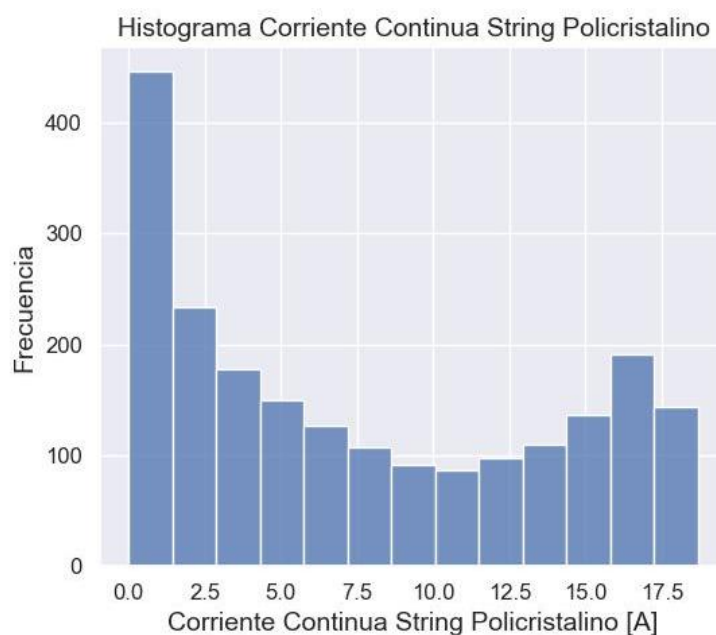
La tabla 5.2 muestra como es clasificación de la salida del algoritmo es decir la variable “Y”.

Mediante esta limpieza de datos las distintas variables tanto eléctricas como ambientales vieron modificadas su distribución de información. A continuación, se mostrará cual es el efecto de realizar la limpieza de la base de datos en las distintas variables.

Corriente Continua



(a)

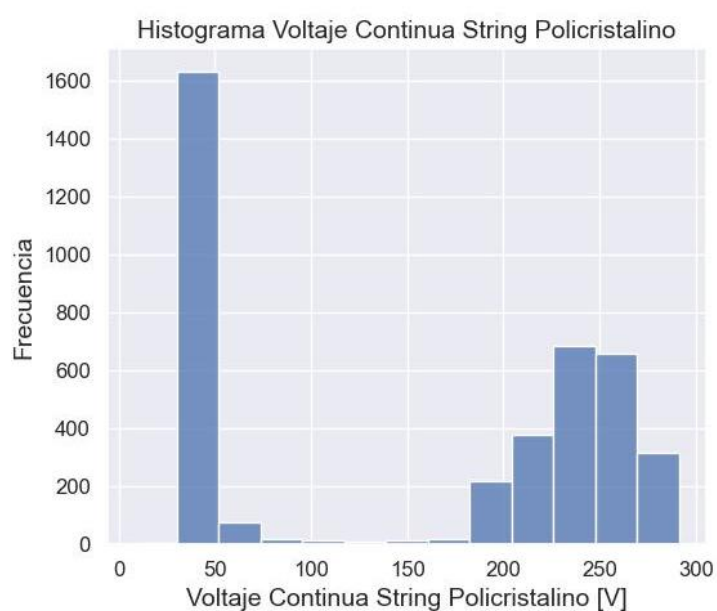


(b)

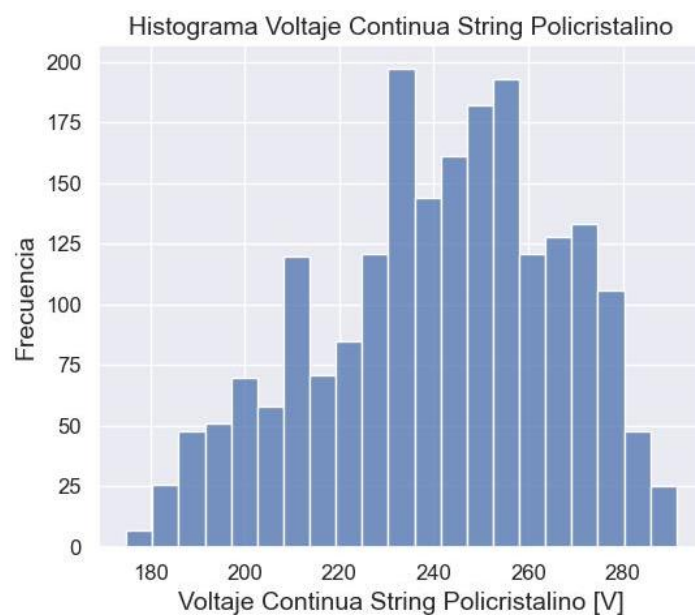
Figura 5.3 Comparación Corriente Continua base de datos de sombreado,
(a) Datos de Corriente sin Depuración, (b) Datos de Corriente post Depuración de Datos

Fuente: Elaboración Propia

Voltaje en corriente continua



(a)

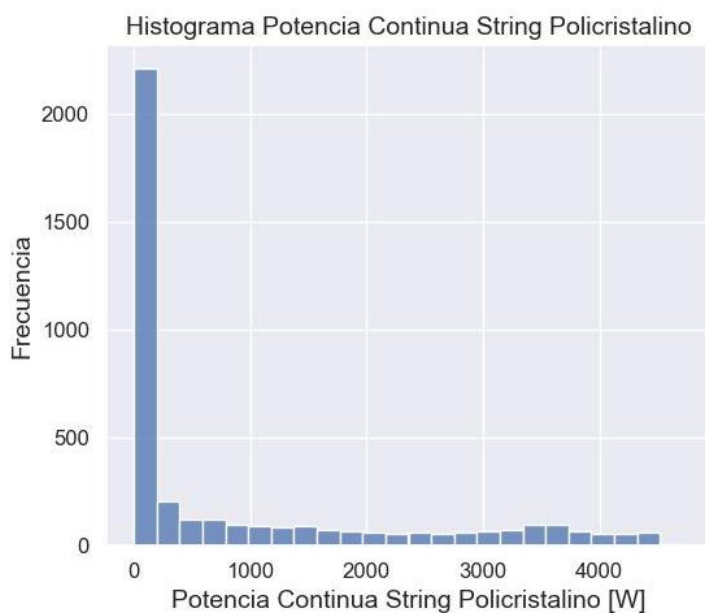


(b)

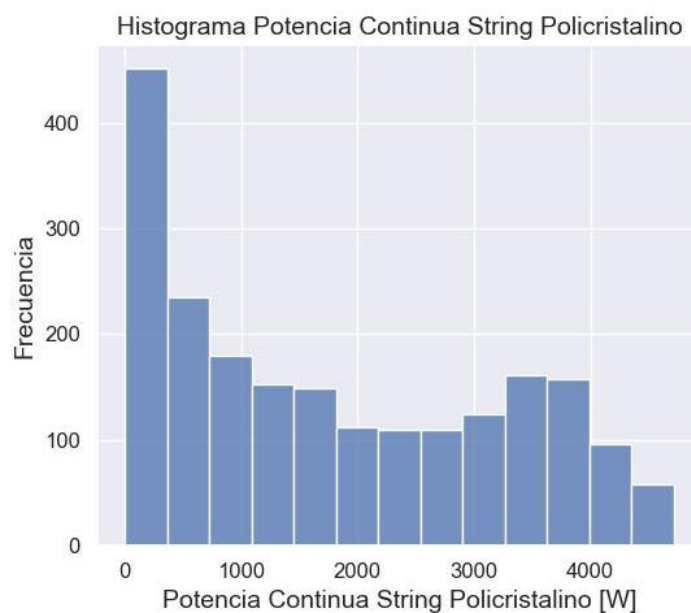
Figura 5.4 Comparación Voltaje en CC base de datos de sombreado,
(a) Datos de Voltaje sin Depuración, (b) Datos de Voltaje post Depuración de Datos.

Fuente: Elaboración Propia

Potencia en corriente continua



(a)

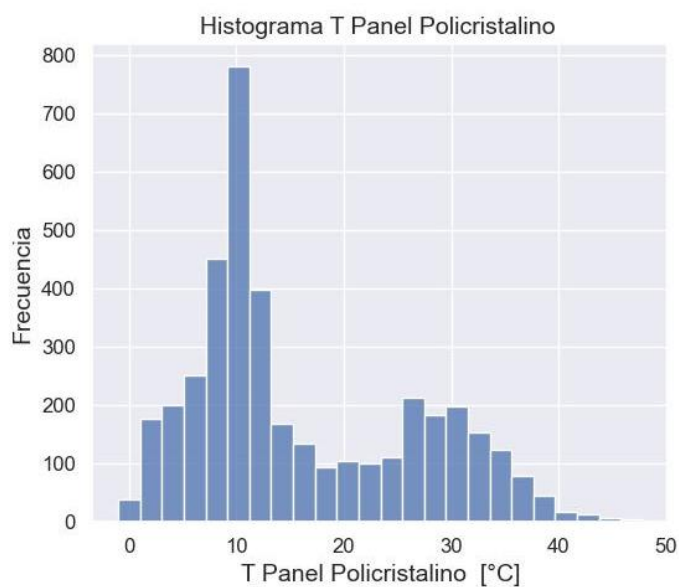


(b)

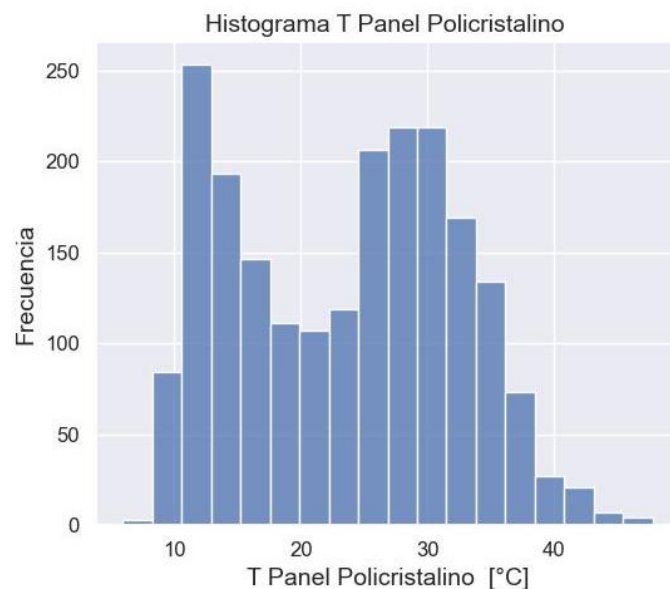
Figura 5.5 Comparación Potencia en CC base de datos de sombreado,
(a) Datos de Potencia sin Depuración, (b) Datos de Potencia post Depuración de Datos.

Fuente: Elaboración Propia

Temperatura del Panel Fotovoltaico



(a)

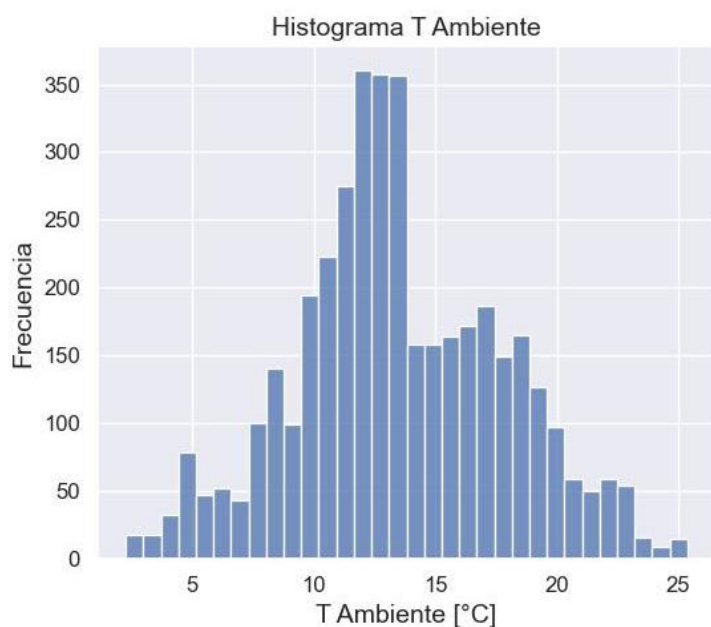


(b)

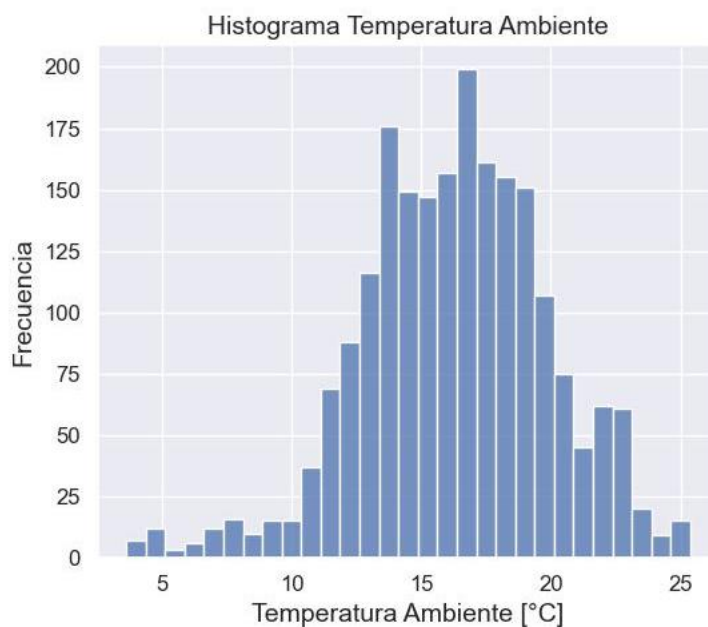
Figura 5.6 Comparación Temperatura del panel FV base de datos de sombreado,
(a) Datos de T. del panel FV sin Depuración, (b) Datos de T. del panel FV post Depuración de Datos

Fuente: Elaboración Propia

Temperatura ambiente



(a)

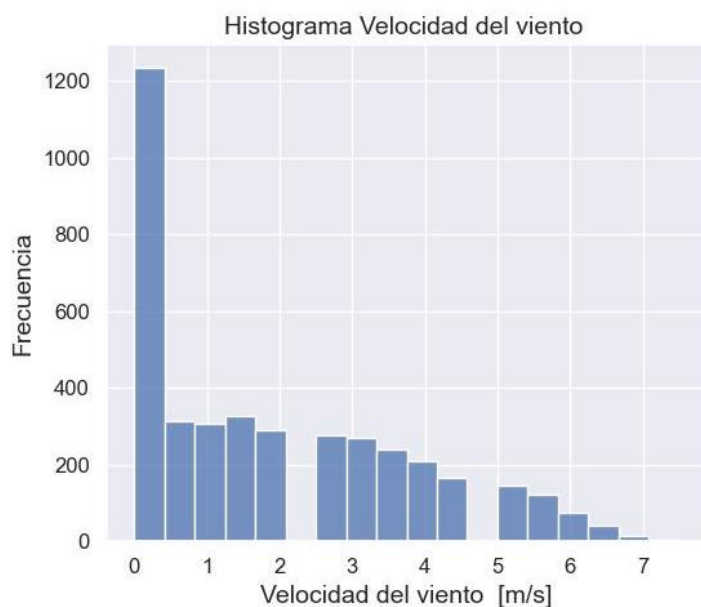


(b)

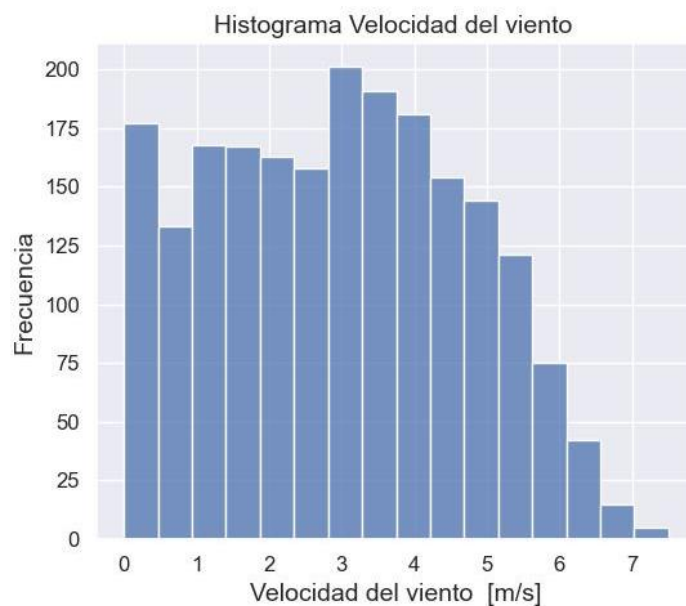
Figura 5.7 Comparación Temperatura Ambiente base de datos de sombreado,
(a) Datos de T. Ambiente sin Depuración, (b) Datos de T. Ambiente post Depuración de Datos

Fuente: Elaboración Propia

Velocidad del Viento



(a)



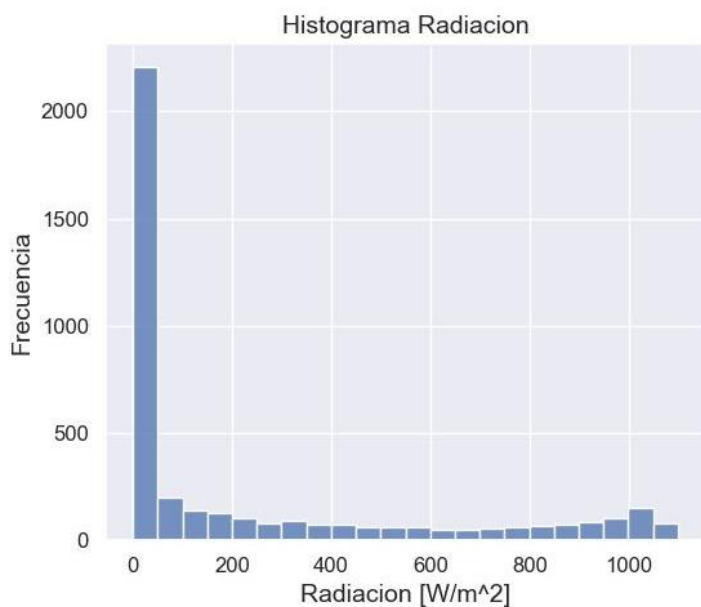
(b)

Figura 5.8 Comparación Velocidad del Viento base de datos de sombreado,

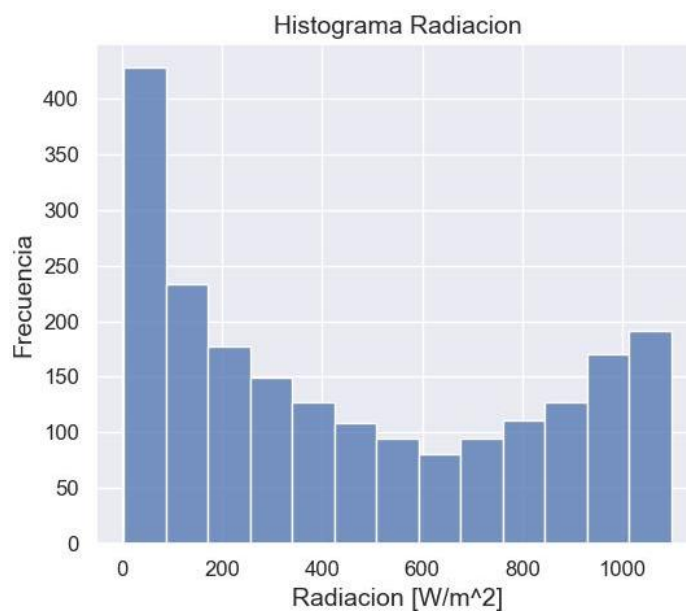
(a) Datos de Velocidad del Viento sin Depuración, (b) Datos de Velocidad del Viento post Depuración de Datos

Fuente: Elaboración Propia

Radiación



(a)



(b)

Figura 5.9 Comparación Radiación base de datos de sombreado,

(a) Datos de Radiación sin Depuración, (b) Datos Radiación post Depuración de Datos

Fuente: Elaboración Propia

A través de la representación gráfica de los histogramas de las distintas variables utilizadas para este problema plasmadas desde la figura 5.3 a la figura 5.9, se puede notar la mejoría de los datos con los que entrenará al modelo. En las figuras 5.3, 5.5 y 5.9 que corresponden a la Corriente Continua, Potencia CC y la Radiación respectivamente se ve una disminución de los valores “Cero” presentes en la base de datos, esto se debe que los datos que corresponden a la noche no entregan ninguna información valiosa, por lo que posterior a la depuración de los datos su cantidad se ve reducida considerablemente. En esas figuras se puede apreciar que el histograma tiene una concentración de datos en valores positivos cercanos a 0, que corresponden a cuando amanece y al atardecer. La distribución de estos gráficos es prácticamente simétrica pero como se explicó anteriormente con una tendencia a los valores cercanos a cero.

De la figura 5.4 correspondiente al Voltaje CC, se puede apreciar como al igual que en las figuras anteriores al eliminar los valores igual a cero, se obtiene una distribución asimétrica izquierda.

Respecto a las figuras 5.6 y 5.7 correspondiente a la temperatura de los paneles y la temperatura ambiente respectivamente, se puede apreciar que tienen una forma similar solo que el panel fotovoltaico presenta mayores temperaturas alcanzando valores que rondan los 40° C esto tiene sentido ya que los paneles tienen a calentarse debido a que siempre están funcionando incluso en la noche presentan voltaje, por otro lado al medio día debido al aumento de la temperatura ambiente y también al aumento de la corriente alcanza mayores temperaturas.

La figura 5.8 correspondiente a la velocidad del viento muestra como con la limpieza de datos quedo una distribución asimétrica izquierda y sus valores van desde 0 m/s a una máxima de 8 m/s.

5.3.2 Base de datos Suciedad

Los datos para esta base de datos fueron los obtenidos de las pruebas de suciedad las cuales fueron mostradas en el capítulo 3. Estas pruebas fueron realizadas entre el día 4 y 18 de enero del 2023. En este caso la base de datos está compuesta por 4320 filas y 12 columnas y al igual que en la base de datos de sombreados se debe realizar la limpieza de los datos en la figura 5.10 se puede apreciar la comparación de la base de datos cuando se realiza un procesamiento de los datos para apreciar la cantidad de datos los cuales no aportan información, que como se explicó anteriormente son cuando no hay irradiancia es decir en la tarde noche donde no hay luz solar.

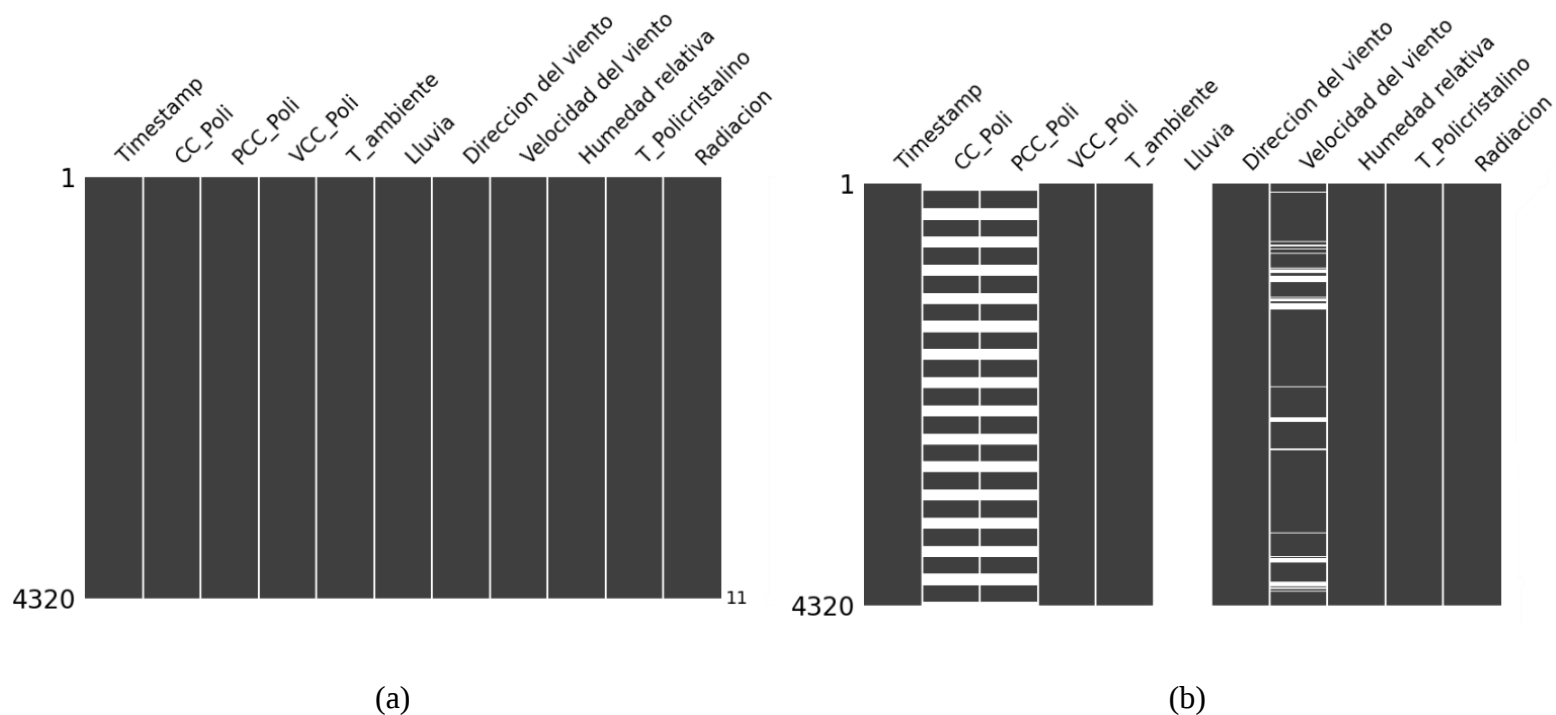


Figura 5.10 Proceso de limpieza de datos base de datos de suciedad

(a) Base de Datos Suciedad antes de la limpieza (b) Base de Datos Suciedad Durante la limpieza

Fuente: Elaboración Propia

Luego de realizar la limpieza de los datos mostrados en la figura 5.10 la base de datos para entrenar a los algoritmos para identificar la suciedad es de 2608 filas y 11 columnas esta se puede ver en la figura 5.11.

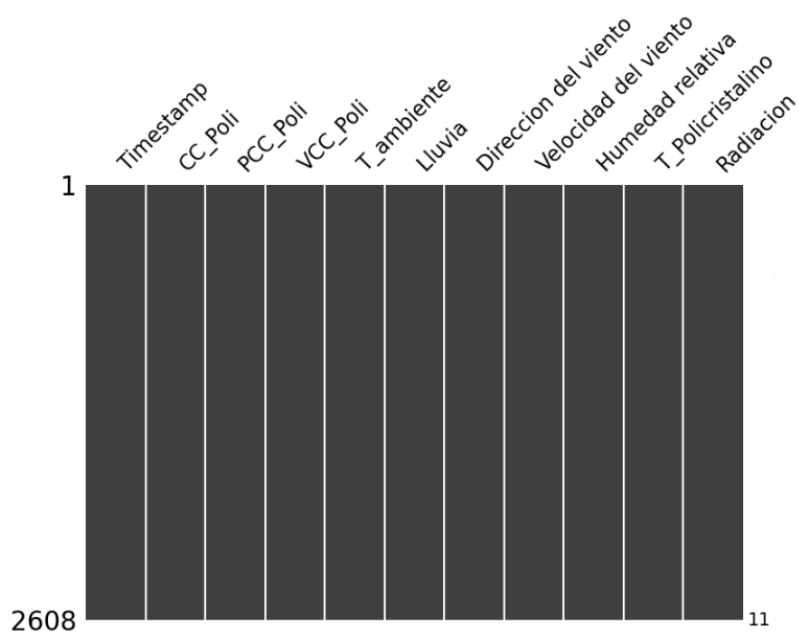


Figura 5.11: Base de datos de suciedad posterior a la depuración

Fuente: Elaboración Propia

Tabla 5.3: Clasificación de base de datos de Suciedad

Tipo de prueba	Clases “Y”
Panel/es Limpio sin suciedad	0
Panel/es Sombreado	1

Fuente: Elaboración Propia

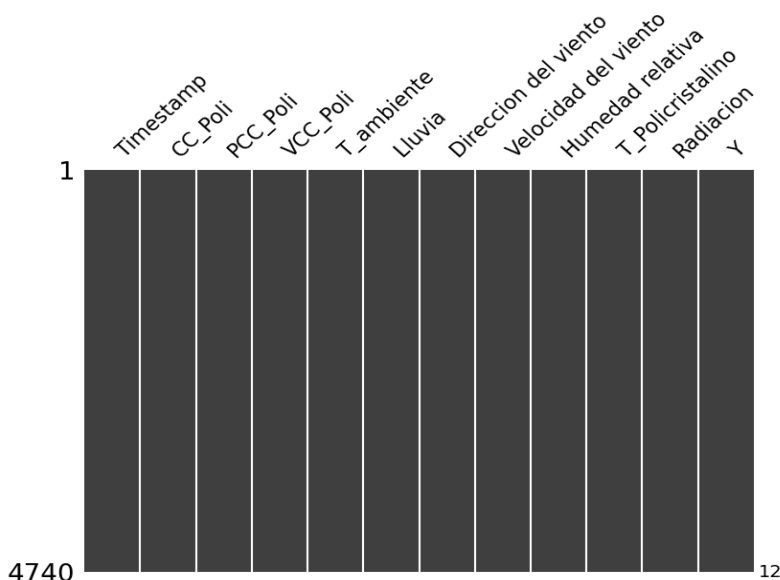
La tabla 5.3 muestra como es clasificación de la salida del algoritmo es decir la variable “Y”. Cabe destacar que la descripción de las variables utilizadas posterior a la limpieza de datos, en el entrenamiento de los algoritmos para detectar suciedad, fue bastante similar a las expuestas ya en las figuras 5.3 a 5.9 por lo que si desea visualizar el resultado de la limpieza de datos se encuentran en el Anexo 2.

5.3.3 Base de Datos Sombreado y Suciedad

La construcción de esta base de datos se realiza juntando ambas bases de datos anteriores y modificando su salida, ya que se tendrán 3 clases las cuales se muestran a continuación en la tabla 5.4.

Tabla 5.4: Clases de la base de datos de sombreado y suciedad

Tipo de prueba	Clases “Y”
Panel/es Limpio sin suciedad ni sombreado	0
Panel/es Sombreado	1
Panel/es Sucios	2

Fuente: Elaboración Propia**Figura 5.12: Base de datos Mixta de suciedad y Sombreado****Fuente: Elaboración Propia**

De la figura 5.12 podemos apreciar que la base de datos está compuesta por 12 columnas y 4740 filas, sin ningún dato nulo ya que es la combinación de las dos bases de datos anteriores.

5.4. Modelos Utilizados

En esta sección se presentarán los diferentes algoritmos utilizados y los resultados obtenidos al tratar de clasificar los casos de suciedad y/o sombreado, según corresponda. Se aplicarán los algoritmos mencionados en el Capítulo 4, tanto para clasificación binaria como para clasificación multiclase. Se utilizarán dos algoritmos diferentes en cada caso y se comparará su rendimiento utilizando las métricas explicadas en el Capítulo 4.

Es importante destacar que, para el entrenamiento de los modelos, se realizaron diversas modificaciones en los datos para poder ingresarlos a los algoritmos. A continuación, se detallan los pasos realizados:

- Escalado o normalización de las variables: se lleva a cabo para evitar tener diferentes escalas entre los datos. Esto puede ocurrir al combinar variables con diferentes unidades de medida. El objetivo es facilitar la comparación y combinación de estas variables. Para realizar esto, se utilizó el comando "StandardScaler" de la biblioteca "scikit-learn" que también es conocida como "sklearn" todos los datos quedaron en un rango de $[-2 \text{ a } 2]$. Posterior a esto, se aplica el método "scaler.fit ()" para ajustar todas las variables a esta escala.
- División de la base de datos: antes de ingresar los datos a los modelos, es importante separar la base de datos en conjuntos de entrenamiento y prueba. Se divide la base de datos en dos grupos, siendo el primero (generalmente el más grande) utilizado para entrenar el modelo. El segundo grupo, denominado conjunto de prueba, se utiliza para evaluar el rendimiento del modelo. Esto se logra comparando las salidas esperadas con las predicciones realizadas por el modelo, utilizando los datos de "x_test".

En resumen, esta sección se enfoca en la presentación de los algoritmos utilizados y los resultados obtenidos en la clasificación de los casos de suciedad y sombreado. Se destacan los pasos previos necesarios para preparar los datos, como el escalado de variables y la división de la base de datos, para todos los modelos que se entrenen la base de datos se dividirá en 60% de los datos para entrenamiento, 30 % de los datos para test y un 10 % para validación de resultados. Estos pasos son fundamentales para garantizar un análisis adecuado y una evaluación correcta del rendimiento de los algoritmos.

5.4.1 Modelos de Clasificación Sombreado

En esta sección se mostrarán los algoritmos utilizados para realizar la clasificación binaria por lo que habrá una división entre los algoritmos entrenados para identificar la suciedad y otros para el sombreado.

A. Regresión Logística (Sombreado)

Para poder utilizar este algoritmo es necesario conocer su ecuación general la cual se utilizará para definir la estructura del algoritmo la cual se obtiene luego de haber ajustado los hiperparámetros del modelo con los datos utilizados en el entrenamiento

$$y = w_1x_1 + w_2x_2 + w_3x_3 + w_4x_4 + \dots + w_nx_n + \beta \quad \text{Ec. 5.1}$$

En este caso los hiperparámetros obtenidos fueron los siguientes

$$y = 0.3376x_1 - 1.9856x_2 - 5.7437x_3 + 0.0726x_4 - 0.2544x_5 - 0.2489x_6 + 0.9710x_7 + 1.7514x_8 + 1.9344x_9 + 0.7384x_{10} + 2.8596 \quad \text{Ec. 5.2}$$

De la ecuación 5.2 y teniendo en cuenta la información de la tabla 5.1 podemos decir que la variable que más peso tiene en este modelo es el voltaje del string en conjunto con la temperatura de la celda, lo que hace sentido respecto a los resultados mostrados en el capítulo 3.

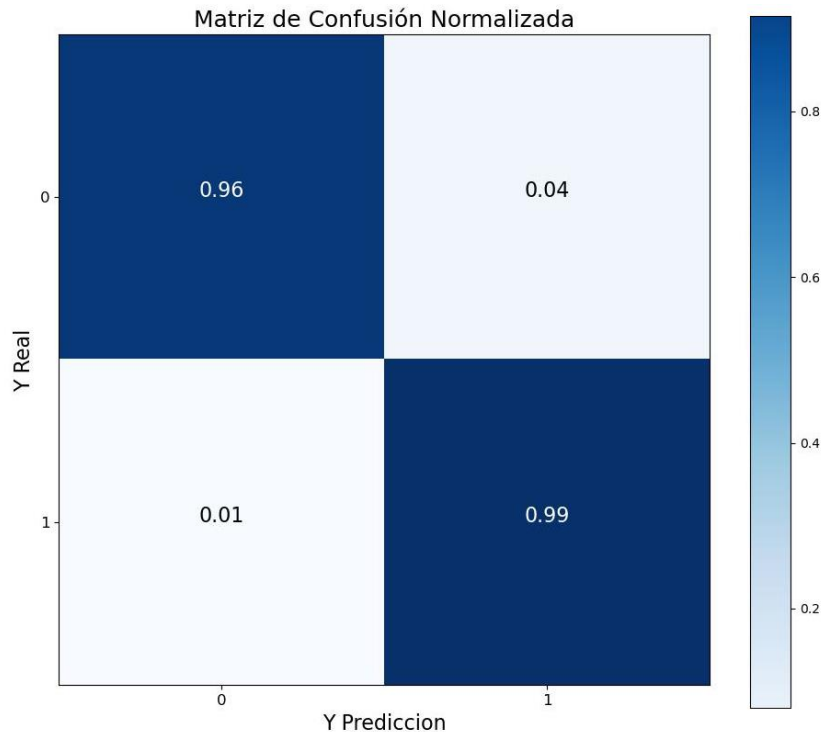


Figura 5.13: Matriz de confusión Algoritmo de Regresión Logística en detección de sombreado.

Fuente: Elaboración Propia

Tabla 5.5: Distribución Matriz de confusión Regresión logística sombreado

Predicción		
	Negativo	Positivo
Negativo	VN = 200	FP = 9
Positivo	FN = 4	VP = 427

Fuente: Elaboración Propia

Tabla 5.6: Métricas de rendimiento Algoritmo Regresión Logística (Sombreado)

	Exactitud	Precision	Sensibilidad	F1 Score
Regresión Logística	97.96%	97.93%	99.01%	98.50%

Fuente: Elaboración Propia

Los resultados obtenidos por el algoritmo presentados en las tablas 5.4 y 5.5. Este tiene una exactitud del 97.96%, una precisión del 97.93%, una sensibilidad del 99.01% y un puntaje F1 del 98.5%, se evidencia un rendimiento del algoritmo. Estas métricas demuestran que el algoritmo es capaz de clasificar correctamente la mayoría de las instancias de sombreado, logrando identificar tanto los verdaderos positivos como los verdaderos negativos.

La precisión indica que el algoritmo tiene una baja tasa de falsos positivos, es decir, minimiza los errores al clasificar una instancia como positiva cuando en realidad es negativa. Por otro lado, la alta sensibilidad muestra que el algoritmo tiene una baja tasa de falsos negativos, es decir, es capaz de detectar correctamente la mayoría de las instancias positivas esto se ve reflejado en la figura 5.13 donde se puede apreciar de manera Gráfica que el algoritmo clasifica de manera correcta el 96% de los casos donde no hay sombreado en el sistema y con un 99% de precisión los casos en que si hay sombreado.

El puntaje F1, que combina la precisión y la sensibilidad en un solo valor, refleja el equilibrio entre ambas métricas. Un puntaje F1 del 98.5% indica que el algoritmo tiene un rendimiento balanceado, logrando una alta precisión y sensibilidad al mismo tiempo.

B. Modelo de Vecinos cercanos (Sombreado)

Una de las cosas más importantes a definir para usar este algoritmo es definir el valor de “K” que especifica los datos cercanos que evaluará para al momento de predecir una solución.

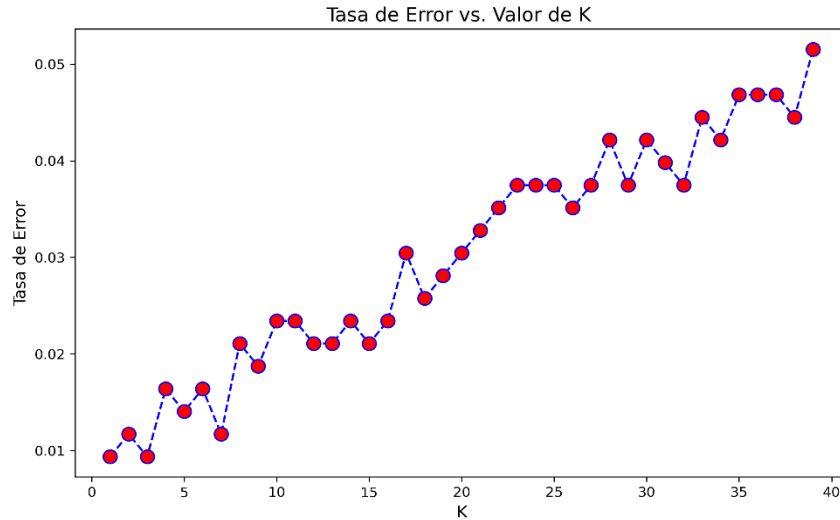


Figura 5.14: Selección de valor K para el algoritmo de clasificación sombreado

Fuente: Elaboración Propia

De la figura 5.14 podemos concluir que los mejores valores de K para minimizar el error son $k = 1$ y $k = 3$. El valor de k seleccionado en este caso fue de $K = 3$.

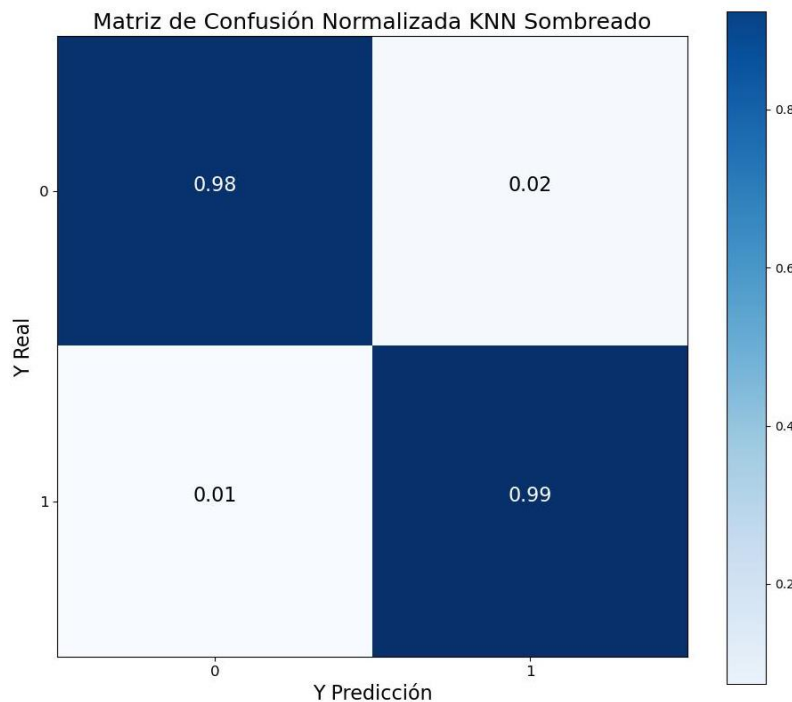


Figura 5.15: Matriz de confusión Algoritmo de Vecinos Cercanos en detección de sombreado.

Fuente: Elaboración Propia

Tabla 5.7: Distribución Matriz de confusión KNN para sombreado

Predicción		
	Negativo	Positivo
Negativo	VN = 205	FP = 4
Positivo	FN = 4	VP = 427

Fuente: Elaboración Propia

Tabla 5.8: Métricas de rendimiento Algoritmo Vecinos cercanos (Sombreado)

	Exactitud	Precision	Sensibilidad	F1 Score
Vecinos cercanos	98.75%	99.07%	99.07%	99.07%

Fuente: Elaboración Propia

Los resultados obtenidos por el algoritmo de vecinos cercanos utilizado para detectar sombreado se aprecian en la tabla 5.8. este modelo tiene una exactitud del 98.75%, una precisión del 99.07%, una sensibilidad del 99.07% y un puntaje F1 del 99.07%, se evidencia un excelente rendimiento del algoritmo.

La exactitud indica que el algoritmo es capaz de clasificar correctamente la gran mayoría de las instancias, logrando un alto nivel de acierto en la detección de sombreado. Además, la alta precisión y sensibilidad, ambas con un valor del 99.07%, indican que el algoritmo es altamente efectivo en la identificación de los verdaderos positivos y verdaderos negativos. Lo que se puede apreciar en la tabla 5.7 donde clasifica de manera incorrecta solo 8 datos de un total de 640 datos utilizados para realizar el test de este algoritmo.

El puntaje F1 del 99.07% refleja el equilibrio entre la precisión y la sensibilidad del algoritmo. Este valor tan alto indica que el algoritmo tiene un rendimiento sobresaliente en la detección de sombreado. Esto se puede complementar con lo que muestra la figura 5.15 donde se puede ver que el algoritmo identifica de manera correcta que el sistema no presenta sombreado en un 98 % de las veces, por otro lado, cuando hay sombreado en los paneles FV lo detecta en un 99 % de las ocasiones.

5.4.2 Modelos de Clasificación Suciedad

De igual manera que en los modelos para detectar el sombreado los modelos que se utilizaran para la detección de sombreado son el algoritmo de regresión logística y el de vecinos cercanos.

C. Regresión Logística (Suciedad)

Su ecuación general con los hiperparámetros ajustados para la detección de suciedad se muestra a continuación. Remplazando los valores de los hiperparámetros obtenidos la ecuación 5.1

$$y = -0.7334x_1 - 2.9877x_2 - 0.1709x_3 + 0.5455x_4 - 0x_5 - 0.2183x_6 + 0.3923x_7 + 1.0683x_8 + 0.5394x_9 + 3.1348x_{10} + 0.4743 \quad \text{Ec. 5.3}$$

De la ecuación 5.3 se puede decir que para el modelo de clasificación de suciedad las variables que más importancia tienen son la potencia en CC y la radiación seguidos en menor medida por la humedad relativa y la corriente es importante destacar que la variable x_5 que corresponde a la lluvia no aporta al modelo, esto tiene lógica ya que en la base de datos no había ningún registro de lluvia dado a que estas pruebas se realizaron en el mes de enero (Verano). Pese a esto se mantuvo esta variable en la base de datos, para poder tener la misma cantidad de variables que la data de sombreado.

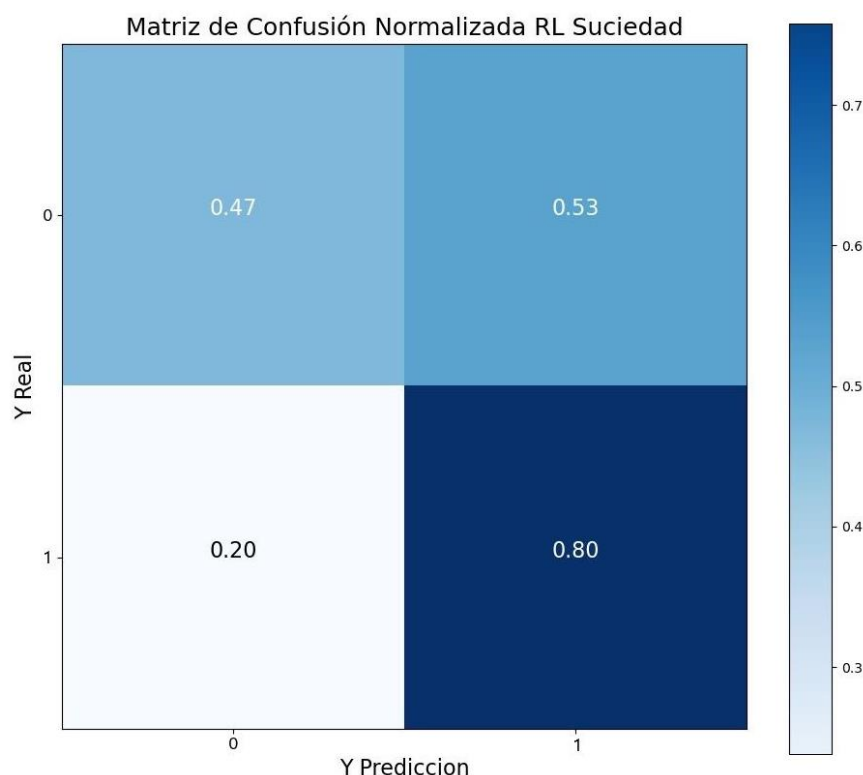


Figura 5.16: Matriz de confusión normalizada Algoritmo de Regresión Logística en detección de suciedad.

Fuente: Elaboración Propia

Tabla 5.9: Distribución Matriz de confusión RL para suciedad

Predicción		
	Negativo	Positivo
Negativo	VN = 154	FP = 174
Positivo	FN = 89	VP = 366

Fuente: Elaboración Propia

Tabla 5.10: Métricas de rendimiento Algoritmo Regresión Logística (Suciedad)

	Exactitud	Precision	Sensibilidad	F1 Score
Regresión Logística	66.41%	67.77%	80.43%	73.56%

Fuente: Elaboración Propia

Los resultados obtenidos por el algoritmo de regresión logística utilizado para detectar suciedad son moderados. Con una exactitud del 66.41%, una precisión del 67.77%, una sensibilidad del 80.43% y un puntaje F1 del 73.56%, se evidencia un rendimiento promedio del algoritmo.

La exactitud del 66.41% indica que el algoritmo clasifica correctamente una mayoría de las instancias, aunque aún queda margen para mejorar. La precisión del 67.77% y la sensibilidad del 80.43% sugieren que el algoritmo es más efectivo en la identificación de verdaderos positivos que de verdaderos negativos. Esto se ve reflejado en la tabla 5.9 donde se aprecian la cantidad de datos clasificados correctamente en cada categoría.

Además, en la figura 5.16 se puede ver como el algoritmo clasifica correctamente el 80% de los casos en los que hay suciedad presente en el sistema y solo clasifica correctamente el 47% de los casos donde no hay suciedad en el sistema.

El puntaje F1 del 73.56% muestra un equilibrio razonable entre la precisión y la sensibilidad, aunque aún hay margen de mejora. En general, estas métricas indican que el algoritmo de regresión logística utilizado para detectar suciedad tiene un rendimiento moderado y podría beneficiarse de mejoras adicionales para aumentar su efectividad y confiabilidad. Una de las cosas a mejorar es entregarle mayor cantidad de datos de la variable que le cuesta identificar que en este caso es cuando los paneles no presentan suciedad.

D. *Modelo de Vecinos cercanos (Suciedad)*

Como se explicó anteriormente lo primero que se debe hacer con este algoritmo de clasificación es determinar el valor de K.

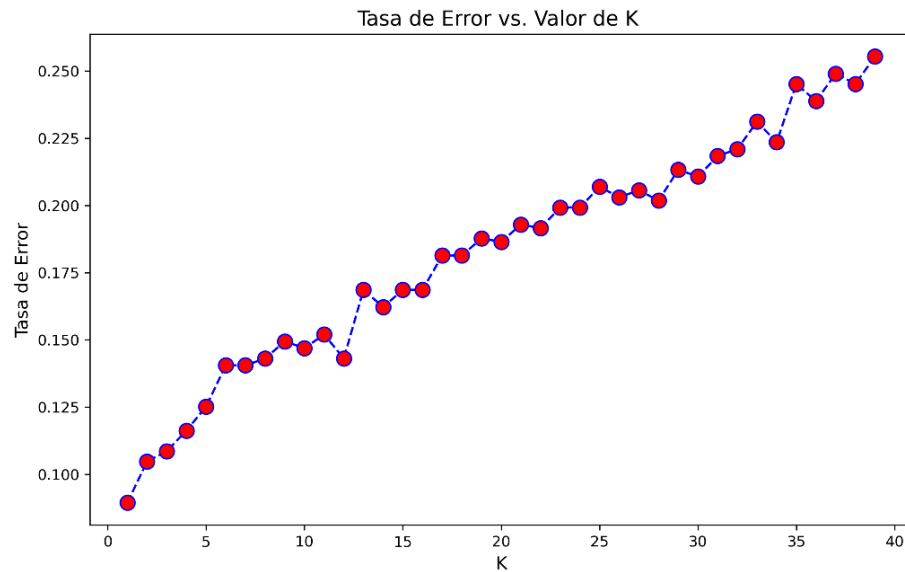


Figura 5.17: Selección de valor K para el algoritmo de clasificación suciedad

Fuente: Elaboración Propia

De la figura 5.17 seleccionamos el valor $k = 1$ ya que es el que menor error presenta a comparación de los otros valores de k .

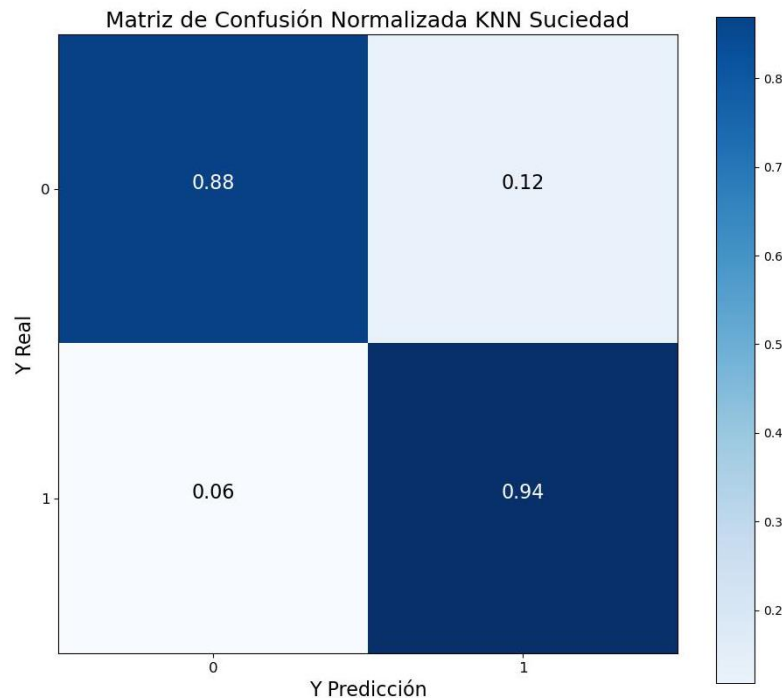


Figura 5.18: Matriz de confusión Algoritmo de Vecinos Cercanos en detección de suciedad.

Fuente: Elaboración Propia

Tabla 5.11: Distribución Matriz de confusión KNN para suciedad

Predicción		
	Negativo	Positivo
Negativo	VN = 287	FP = 41
Positivo	FN = 29	VP = 426

Fuente: Elaboración Propia

Tabla 5.12: Métricas de rendimiento Algoritmo Vecinos cercanos (Suciedad)

	Exactitud	Precision	Sensibilidad	F1 Score
Vecinos cercanos	87.16%	88.29%	89.19%	88.71%

Fuente: Elaboración Propia

Los resultados obtenidos mediante la aplicación del algoritmo de vecinos cercanos para detectar suciedad se pueden apreciar en las tablas 5.11 y 5.12 respecto a esta última podemos concluir lo siguiente. El algoritmo muestra una exactitud del 87.16%, una precisión del 88.29%, una sensibilidad del 89.19% y un puntaje F1 del 88.71%. Estas métricas indican que el algoritmo es efectivo en la identificación precisa tanto de verdaderos positivos como de verdaderos negativos.

Lo anterior se puede apreciar de mejor manera en la tabla 5.11 donde se aprecia la cantidad de datos bien etiquetados, siendo mayor los verdaderos positivos lo que concuerda con un mejor métrica de sensibilidad. En general, los resultados obtenidos respaldan la efectividad y confiabilidad del algoritmo de vecinos cercanos utilizado para detectar suciedad. Esto se respalda con la figura 5.18 donde podemos apreciar que el algoritmo es capaz de detectar correctamente la suciedad en un 94% de los casos. Además identifica correctamente cuando el sistema FV está limpio en un 88% de las ocasiones.

En conclusión, podemos decir que el algoritmo es capaz de identificar de manera correcta la mayoría de los casos. aun que al igual que el modelo de regresión logística utilizado para detectar la suciedad tienden a predecir falsos positivos es decir confunden casos en que el panel está limpio clasificándolo como sucio.

5.4.3 Modelos de Clasificación de Sombreado y Suciedad

Para poder realizar la clasificación tanto de sombreado y suciedad hay que utilizar algoritmos de clasificación multiclase, para eso se utilizarán los algoritmos de vecinos cercanos y la de red neuronal.

E. Vecinos cercanos KNN (Suciedad y Sombreado)

En este caso al igual que en los modelos de clasificación de salida binaria es necesario definir cuál es el valor de k que menor error produce a la hora de predecir la salida.

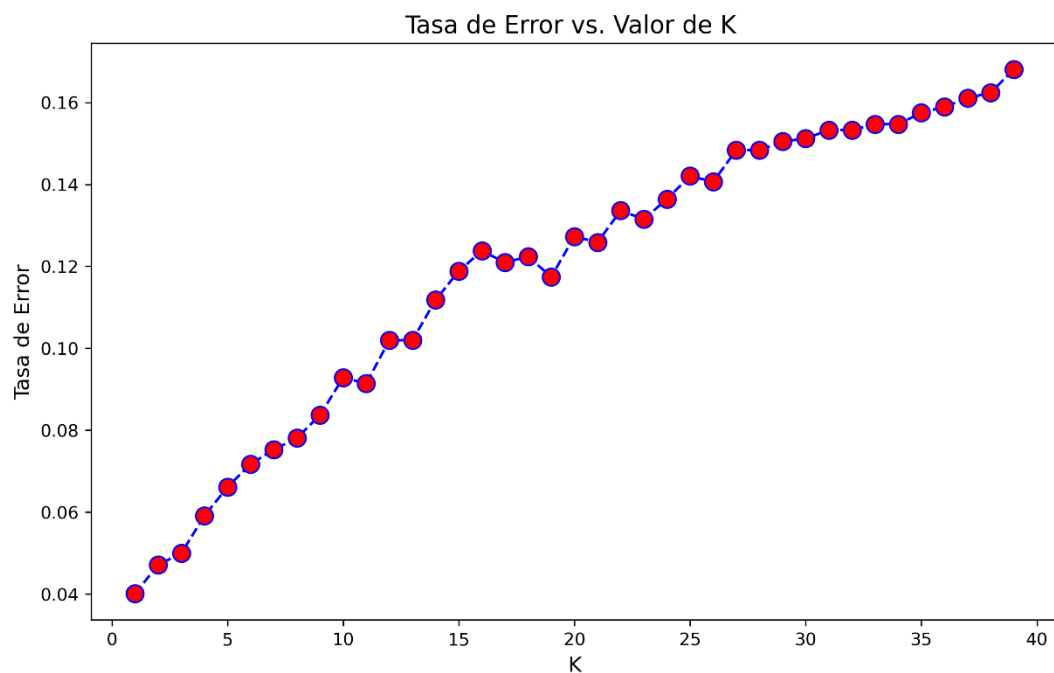


Figura 5.19: Selección de valor K para el algoritmo de clasificación suciedad y suciedad

Fuente: Elaboración Propia

En la figura 5.19 muestra que el valor de K que menor error produce es considerando el parámetro $K = 1$. Ya seleccionado el valor de k podemos evaluar el modelo con las métricas de rendimiento. A diferencia de los modelos entrenados anteriormente las métricas de rendimiento te evalúan la capacidad del modelo de predecir cada una de las clases mostradas en la tabla 5.4. solo la metrica de exactitud puede evaluar el algoritmo en general, los demás indicadores de rendimiento como la precisión, sensibilidad y F1 Score se evalúa por clase. Los resultados de este algoritmo se pueden apreciar en la tabla 5.14.

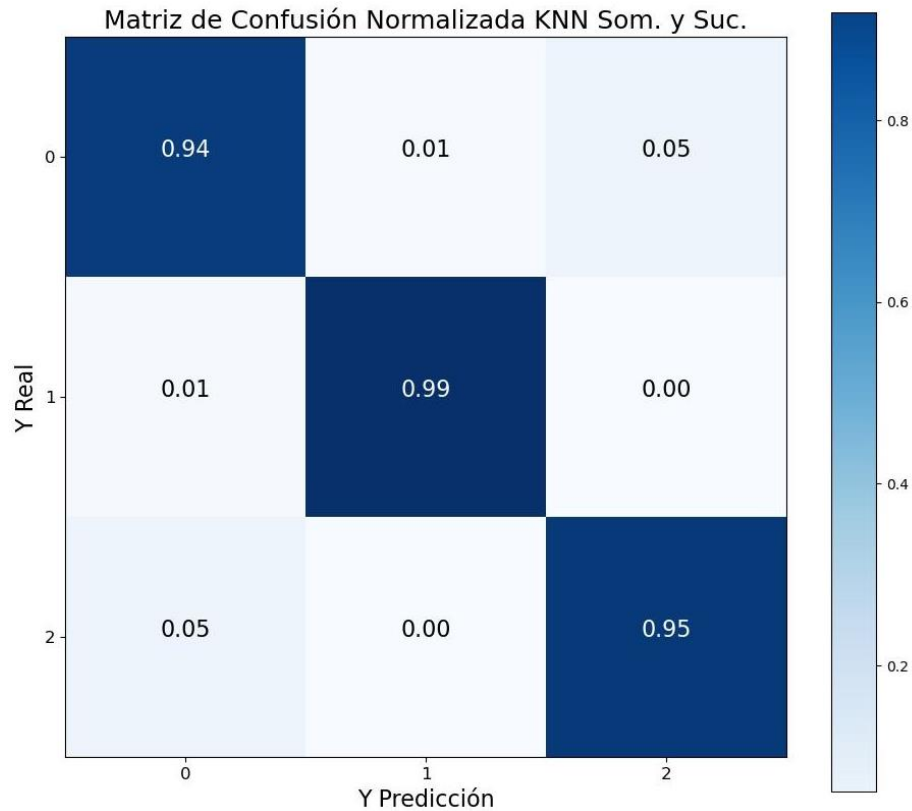


Figura 5.20: Matriz de confusión Algoritmo de Vecinos Cercanos en detección de suciedad y sombreado.

Fuente: Elaboración Propia

Tabla 5.13: Distribución Matriz de confusión Vecinos Cercanos para suciedad y sombreado

Predicción			
	Clase 0 (Limpio)	Clase 1 (Sombreado)	Clase 2 (Sucio)
Clase 0 (Limpio)	VP Clase 0 = 511	FP Clase 0 = 4	FP Clase 0 = 27
Clase 1 (Sombreado)	FP Clase 1 = 4	VP Clase 1 = 410	FP Clase 1 = 0
Clase 2 (Sucio)	FP Clase 2 = 22	FP Clase 2 = 0	VP Clase 2 = 444

Fuente: Elaboración Propia

Tabla 5.14: Métricas de rendimiento Algoritmo Vecinos cercanos (Sombreado y Suciedad)

Métrica	Clase Y = 0	Clase Y = 1	Clase Y = 2
Exactitud		95.99 %	
Precisión	95.1 %	99.03 %	94.26 %
Sensibilidad	94.2 %	99.03 %	95.27 %
F1 Score	94.7 %	99.03 %	94.7 %

Fuente: Elaboración Propia

El algoritmo de vecinos cercanos de clasificación utilizado para identificar si los paneles están limpios (clase 0), sombreados (clase 1) o sucios (Clase 2) muestra un rendimiento excelente en la detección de condiciones específicas de los paneles solares.

Los resultados mostrados en la tabla 5.14 obtenidos revelan un alto desempeño del algoritmo en términos de las métricas de rendimiento. Se logró una exactitud del 95.99%, lo que indica que el algoritmo clasifica correctamente la mayoría de las instancias en las tres clases. Lo cual se puede apreciar en la figura 5.20 donde se aprecia que el algoritmo es capaz de clasificar correctamente la clase 0 en un 94% de las ocasiones, la clase 1 en un 99% de las veces y la clase 3 en un 95% de las veces. Además, se obtuvo una precisión del 95.1% para la clase 0 (paneles limpios), 99.03% para la clase 1 (sombreado) y 94.26% para la clase 2 (suciedad). Esto muestra la capacidad del algoritmo para identificar correctamente los paneles en diferentes condiciones.

En cuanto a la sensibilidad, se obtuvieron valores del 94.2% para la clase 0, 99.03% para la clase 1 y 95.27% para la clase 2. Estos resultados indican la capacidad del algoritmo para detectar correctamente los verdaderos positivos en cada clase, los cuales se pueden ver a detalle en la tabla 5.13. El puntaje F1 también fue alto, con valores del 94.7% para la clase 0, 99.03% para la clase 1 y 94.7% para la clase 2. Esto refleja un buen equilibrio entre la precisión y la sensibilidad en la clasificación de las diferentes condiciones de los paneles solares.

F. Red Neuronal (Suciedad y Sombreado)

Para el entrenamiento de un modelo de red neuronal, es necesario seleccionar una estructura adecuada que determine el número de capas y neuronas en cada capa. En el caso de una red neuronal de salida multiclase, la función de activación utilizada en la capa de salida es Softmax. En este estudio, se variarán la cantidad de capas, la cantidad de neuronas y la función de activación (FDA) en busca del mejor modelo. Además, se considerarán diferentes cantidades de iteraciones del modelo, específicamente 100, 500 y 1000 iteraciones.

Para seleccionar el mejor modelo, se utilizará la métrica de exactitud, que indica la proporción de etiquetas correctas en relación al total de etiquetas. Con el objetivo de combinar la cantidad de iteraciones con el rendimiento del modelo, se establecerán parámetros de exactitud (Exactitud 1, Exactitud 2 y Exactitud 3) en una tabla.

Parámetros de exactitud:

- Exactitud 1 (Metrica exactitud después de 100 iteraciones).
- Exactitud 2 (Metrica exactitud después de 500 iteraciones).
- Exactitud 3 (Metrica exactitud después de 1000 iteraciones).

Estos parámetros permitirán evaluar y comparar el desempeño de los modelos de red neuronal entrenados, con el objetivo de seleccionar el mejor modelo en términos de exactitud en la clasificación multiclase. Otra cosa a considerar es que todas las estructuras que se analicen tendrán 10 entradas que son las variables con las que se entrena el modelo, estas entradas no se considera como una capa es decir no tiene función de activación y su última capa tendrá 3 neuronas de salidas con la función de activación Softmax. Por con siguiente todas las estructuras que se analicen comenzaran con el valor 10 y terminaran con el número 3.

Tabla 5.15: Comparación resultado de RN para diferentes estructuras

Estructura	FDA Capa 1	FDA Capa 2	FDA Capa 3	FDA Capa 4	Exactitud 1	Exactitud 2	Exactitud 3
10-10-32-3	Relu	Relu	Softmax	--	90.36 %	90.92 %	91.15%
10-20-32-3	Relu	Relu	Softmax	--	92.75%	94.79 %	94.65%
	Relu	Sigmoid	Softmax	--	93.10 %	93.61 %	94.11%
	Sigmoid	Relu	Softmax	--	89.94 %	92.19 %	93.03 %
10-32-20-3	Relu	Relu	Softmax	--	93.53 %	95.14%	95.35%
	Relu	Sigmoid	Softmax	--	38.11 %	38.11 %	38.11 %
10-64-32-3	Relu	Relu	Softmax	--	93.35 %	95.42%	95.21%
	Tanh	Relu	Softmax	--	38.11 %	38.11 %	38.11 %
10-64-128-3	Relu	Relu	Softmax	---	94.51 %	95.49 %	96.00 %
	Sigmoid	Relu	Softmax	--	91.27 %	95.63%	95.63 %
	Relu	Sigmoid	Softmax	--	95.00 %	95.21 %	94.58 %
10-64-128-32-3	Relu	Relu	Relu	Softmax	94.44 %	94.93 %	95.63 %
	Relu	Sigmoid	Sigmoid	Softmax	94.44 %	94.09 %	94.93%
	Relu	Sigmoid	Relu	Softmax	94.23 %	95.21 %	94.30 %
10-128-64-32-3	Relu	Relu	Relu	Softmax	94.65 %	95.42 %	95.85 %
	Relu	Sigmoid	Relu	Softmax	94.16 %	93.45 %	94.86 %
	Relu	Relu	Sigmoid	Softmax	91.91%	92.98 %	90.64 %

Fuente: Elaboración Propia

En la tabla 5.15 podemos ver diferentes estructuras de RN y con distintas funciones de activación, donde se aprecia que cuando se ocupaban las funciones de activación “Relu” y “Softmax,

es cuando mejores valores de exactitud se obtuvieron en todas las estructuras probadas. Para esta evaluación de estructuras no se utilizó la función “Leaky ReLU” ya que ninguna de las variables de entrada utilizadas para entrenar al modelo tenía valores negativos por lo que si se utilizara se obtendría el mismo resultado que con la función “Relu” solo que el gasto de recursos computacionales sería mayor, tampoco se utilizó la función lineal ya que es como no utilizar ninguna función de activación.

Las estructuras que mejor rendimiento tuvieron fueron las siguientes

- Estructura 1: 10 – 64 – 32 – 3. Con las funciones de activación Relu, Relu, Softmax.
- Estructura 2: 10 – 64 – 128 – 3. Con las funciones de activación Relu, Relu, Softmax
- Estructura 3: 10 – 128 – 64 – 32 – 3. Con las funciones de activación Relu, Relu, Relu, Softmax

De estas tres estructuras, se seleccionará el modelo de red neuronal que se utilizará para la clasificación de sombreado y suciedad. Se aumentará la cantidad de iteraciones para evaluar cada modelo y determinar cuál de ellos proporciona los mejores resultados. En esta ocasión, se utilizarán todas las métricas de rendimiento disponibles para elegir la opción más adecuada. Es importante destacar que no siempre aumentar la cantidad de iteraciones garantiza un mejor rendimiento para la red neuronal, ya que otros factores también pueden influir en su desempeño.

Tabla 5.16

Modelo	Exactitud	Precisión			Sensibilidad			F1 Score		
		C - 0	C - 1	C - 2	C - 0	C - 1	C - 2	C - 0	C - 1	C - 2
Estructura 1	94.93 %	94.5%	99.0%	91.9%	92.6%	97.5%	95.2%	93.5%	98.2%	93.5%
Estructura 2	95.49 %	94.4%	98.5%	94.0%	94.0%	98.7%	94.2%	94.2%	98.6%	94.1%
Estructura 3	94.93 %	95.3%	99.2%	90.8%	91.3%	99.0%	95.4%	93.3%	99.1%	93.0%

Fuente: Elaboración Propia

Respecto a la tabla 5.16 considerando las distintas estructuras de RN analizadas la que mejor rendimiento presenta es la estructura 2 la cual es una red neuronal compuesta la cual tiene la siguiente estructura: 10 – 64 – 128 – 3. Es importante destacar que esta estructura es la que mejor porcentaje de precisión muestra en la detección de la clase 1 y clase 2 que corresponde a suciedad y sombreado respectivamente.

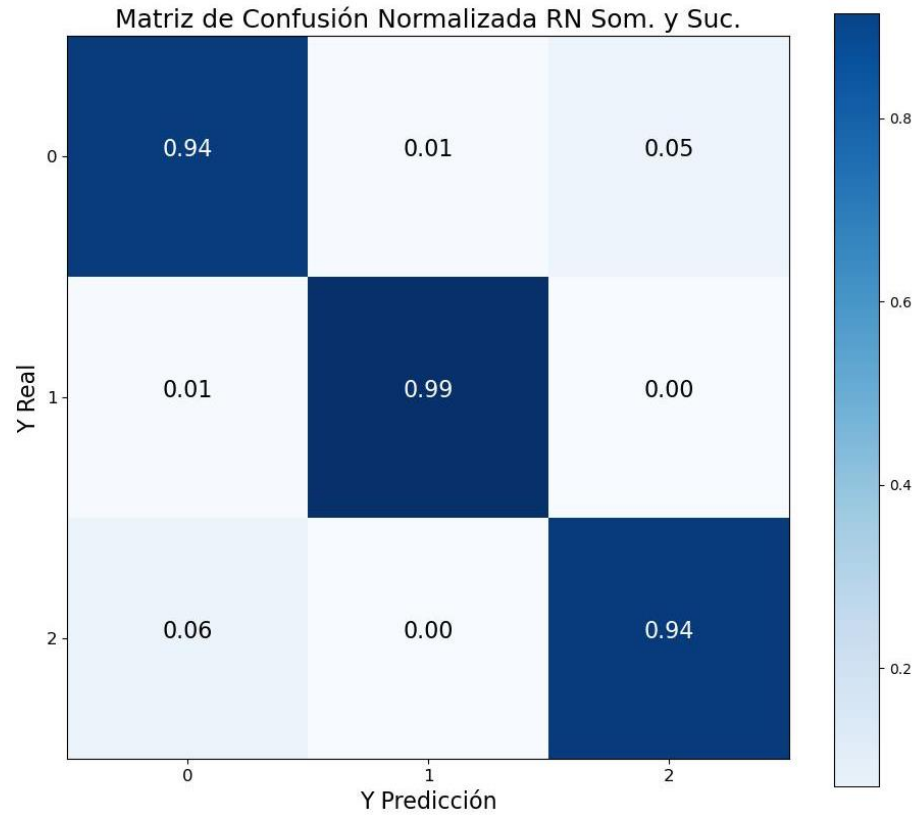


Figura 5.21: Matriz de confusión normalizada Red Neuronal detección de Sombreado y Suciedad
Fuente: Elaboración Propia

Tabla 5.17: Distribución Matriz de confusión Red Neuronal para suciedad y sombreado

Predicción			
	Clase 0 (Limpio)	Clase 1 (Sombreado)	Clase 2 (Sucio)
Clase 0 (Limpio)	VP Clase 0 = 510	FP Clase 0 = 5	FP Clase 0 = 27
Clase 1 (Sombreado)	FP Clase 1 = 4	VP Clase 1 = 409	FP Clase 1 = 1
Clase 2 (Sucio)	FP Clase 2 = 26	FP Clase 2 = 1	VP Clase 2 = 445

Fuente: Elaboración Propia

Tabla 5.18: Métricas de rendimiento Red Neuronal Estructura 2 (Sombreado y Suciedad)

Metrica	Clase Y = 0	Clase Y = 1	Clase Y = 2
Exactitud	95.49 %		
Precision	94.4%	98.5%	94.0%
Sensibilidad	94.0%	98.7%	94.2%
F1 Score	94.2%	98.6%	94.1%

Fuente: Elaboración Propia

La estructura de la red neuronal seleccionada muestra un excelente rendimiento con una precisión del 95.49%. Esta red es capaz de identificar de manera efectiva tanto el sombreado como la suciedad en los paneles, logrando una precisión mayor al 94% al detectar cuando los paneles están limpios. Es notable que este algoritmo muestra un desempeño destacado en la clasificación de casos con sombreado, manteniendo un equilibrio al identificar los verdaderos positivos, lo que se refleja en una métrica F1 Score del 98.5%. Esto se puede apreciar en la figura 5.21, donde se muestra en porcentaje cómo el algoritmo etiqueta correctamente cada clase. Las clases 0 y 1 son etiquetadas correctamente el 84% de las ocasiones. En la tabla 5.17, se puede ver en detalle que la red neuronal clasifica correctamente 510 de 542 casos de la clase 0, mientras que de la clase 2 etiqueta correctamente 445 de 472 casos totales. Es importante destacar que la mayoría de los errores de clasificación ocurren porque la red neuronal no es capaz de identificar correctamente los casos donde el sistema fotovoltaico no presenta suciedad ni sombreado (clase 0) de los casos donde los paneles FV presentan suciedad (clase 2). Esto se evidencia de mejor manera en la tabla 5.17, donde se muestra que hay 27 ocasiones en las que la etiqueta correcta era 2, pero la red neuronal la clasificó como 0. Además, hay 26 casos en los que la etiqueta correcta es 0, pero se clasificaron como clase 2.

5.5 Validación de resultados con datos de otoño

Para realizar una evaluación final de los modelos entrenados y los cuales en su mayoría tienen métricas de rendimiento bastante buenas, se les entregaran datos nuevos, que no estuvieron presente en su etapa de entrenamiento. Para ello se creó una nueva base de datos. Los valores para la construcción de esta fueron obtenidos realizando una mezcla de pruebas iguales a las explicadas en el capítulo 3 repitiendo 3 pruebas de los 8 antes vistas, esto se hizo tanto para el sombreado como para la suciedad. Posterior a comparar las métricas de los algoritmos se seleccionará el mejor de los algoritmos para cada caso analizando este en mayor detalle.

Tabla 5.19: Detalle de las pruebas realizadas en construcción de base de datos para validación

Día Sin pruebas Clase 0	Pruebas Realizadas Clase 1	Fecha	Pruebas Realizadas Clase 2	Fecha
07/05/2023	Sombrear 1 panel FV	05/05/2023	Ensuciar 1 panel FV	09/05/2023
	Sombrear 2 Paneles		Ensuciar 2 Paneles FV	
	FV del mismo String		del mismo String	
	Sombrear 2 Paneles		Ensuciar 2 Paneles FV	
	FV en cada String		en cada String	

Fuente: Elaboración Propia

Para realizar la validación de los diferentes algoritmos se construyeron 3 bases de datos una para los algoritmos que detectan sombreado, otra para detección de suciedad y la última para detección de ambas condiciones. Los detalles de estas bases de datos se muestran en la tabla 5.2.

Tabla 5.20: Descripción de las bases de datos para validación

Base de datos	Cantidad de Filas	Cantidad de Columnas	Cantidad de valores por Clase		
Validación Sombreado	252	11	“0” = 125	“1” = 127	
Validación Suciedad	250	11	“0” = 125	“1” = 125	
Validación Sombreado y Suciedad	377	11	“0” = 125	“1” = 127	“2” = 125

Fuente: Elaboración Propia

De la tabla 5.20 se puede decir que las bases de datos utilizadas para la validación de resultado están bastante equilibradas ya que cada una de las clases tiene aproximadamente 125 datos, siendo el sombreado el con mayor cantidad con 127 casos de estudio.

5.5.1 Algoritmos de clasificación de sombreado

Los resultados que muestra la tabla 5.22 muestran las métricas de rendimiento de los algoritmos de clasificación de sombreado siguen, podemos decir que estos continúan siendo efectivos.

Tabla 5.21: Comparación de rendimientos de algoritmos usados en detección de sombreado en otoño

Algoritmo	Exactitud	Precision	Sensibilidad	F1 Score	Tiempo de entrenamiento	Tiempo de Validación
Regresión Logística	78.17 %	70.69%	96.85%	81.72 %	5.81 Segundos	2.68 Segundos
Vecinos cercanos	90.07 %	83.33%	98.42%	90.9%	6.98 Segundos	3.41 Segundos

Fuente: Elaboración Propia

Entre los modelos de clasificación analizados, el modelo de vecinos cercanos se destaca por su rendimiento en la detección de sombreado y suciedad en sistemas fotovoltaicos. Aunque sus métricas de rendimiento disminuyeron ligeramente en comparación con los valores mostrados en la tabla 5.8, la exactitud, sensibilidad y puntuación F1 aún se mantienen por encima del 90%. Si se tuviera que elegir un modelo para implementar, el modelo de vecinos cercanos sería el más adecuado.

Es importante destacar que el modelo de regresión logística es más rápido de aplicar, ya que su tiempo de ejecución es de 6.47 segundos, lo cual es 2.51 segundos más rápido que el algoritmo de vecinos cercanos. Sin embargo, en términos de rendimiento y precisión en la detección de sombreado y suciedad, el modelo de vecinos cercanos sigue siendo la mejor opción.

A continuación, se analizará en mayor detalle el modelo seleccionado para la detección de sombreado.

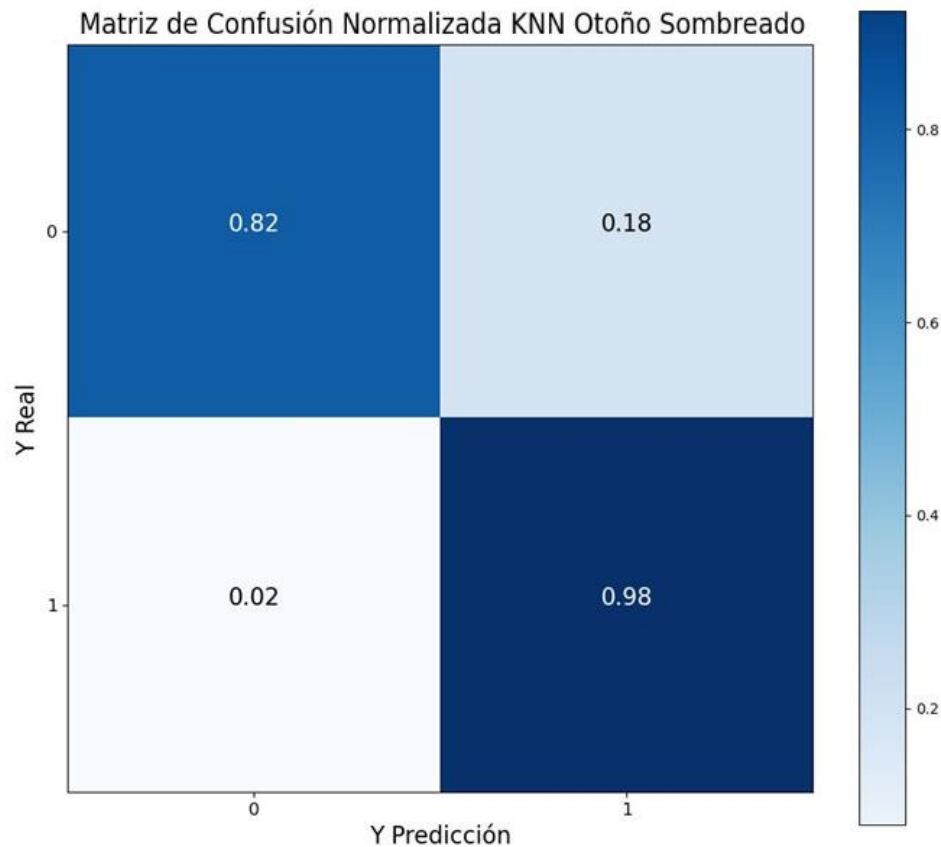


Figura 5.22: Matriz de confusión normalizada Algoritmo KNN otoño en validación resultados de sombreado

Fuente: Elaboración Propia

Tabla 5.22: Distribución Matriz de confusión KNN otoño validación de resultados sombreado

Predicción		
	Negativo	Positivo
Negativo	VN = 102	FP = 23
Positivo	FN = 2	VP = 125

Fuente: Elaboración Propia

Si analizamos más los resultados del algoritmo de vecinos cercanos, podemos apreciar en la matriz de confusión de la figura 5.22 el algoritmo es capaz de identificar el sombreado correctamente el 98 % de las ocasiones si apreciamos la tabla 5.23 podemos darnos cuenta que etiqueta correctamente 125 de 127 datos totales de sombreado. Por otro identifica que no hay presencia de sombreado correctamente un 82 % de las veces es decir 102 veces de 125 ocasiones, lo que también es un rendimiento bastante bueno. Si consideramos que el modelo fue entrenado con datos de los meses de octubre y la base de validación fue creada en mayo donde estamos en climas distintos, podemos concluir que el algoritmo de vecinos cercanos realiza una muy buena clasificación.

5.5.2 Algoritmos de clasificación de suciedad

Los resultados de los algoritmos de clasificación de suciedad con los datos de validaciones, que se detallan en la tabla 5.4. Son bastante regulares tendiendo a malos, en especial el algoritmo de regresión logística ya que si bien es un modelo básico su exactitud es menor al 10 %. Por otro lado, el algoritmo de vecinos cercanos tiene una exactitud superior al 50% en sus resultados, siendo el mejor algoritmo entre las dos opciones que se evalúan.

Tabla 5.23: Comparación de rendimientos de algoritmos usados en detección de suciedad en otoño

Algoritmo	Exactitud	Precision	Sensibilidad	F1 Score	Tiempo de entrenamiento	Tiempo de Validación
Regresión Logística	5.6 %	10 %	11.2 %	10.6 %	6.32 Segundos	3.58 Segundos
Vecinos cercanos	53.2 %	52.24 %	71.2 %	60.33 %	8.69 Segundos	4.01 Segundos

Fuente: Elaboración Propia

Analizando en mayor profundidad el modelo de vecinos cercanos lo primero a remarcar es que los resultados que entrega con la base de datos de validación son muy distintos a los valores posterior a su entrenamiento, que se muestran en la tabla 5.12. ya que en esta las métricas de algoritmo superaban el 85%.

Tabla 5.24: Distribución Matriz de confusión KNN otoño validación de resultados suciedad

	Predicción	
	Negativo	Positivo
Negativo	VN = 44	FP = 81
Positivo	FN = 36	VP = 89

Fuente: Elaboración Propia

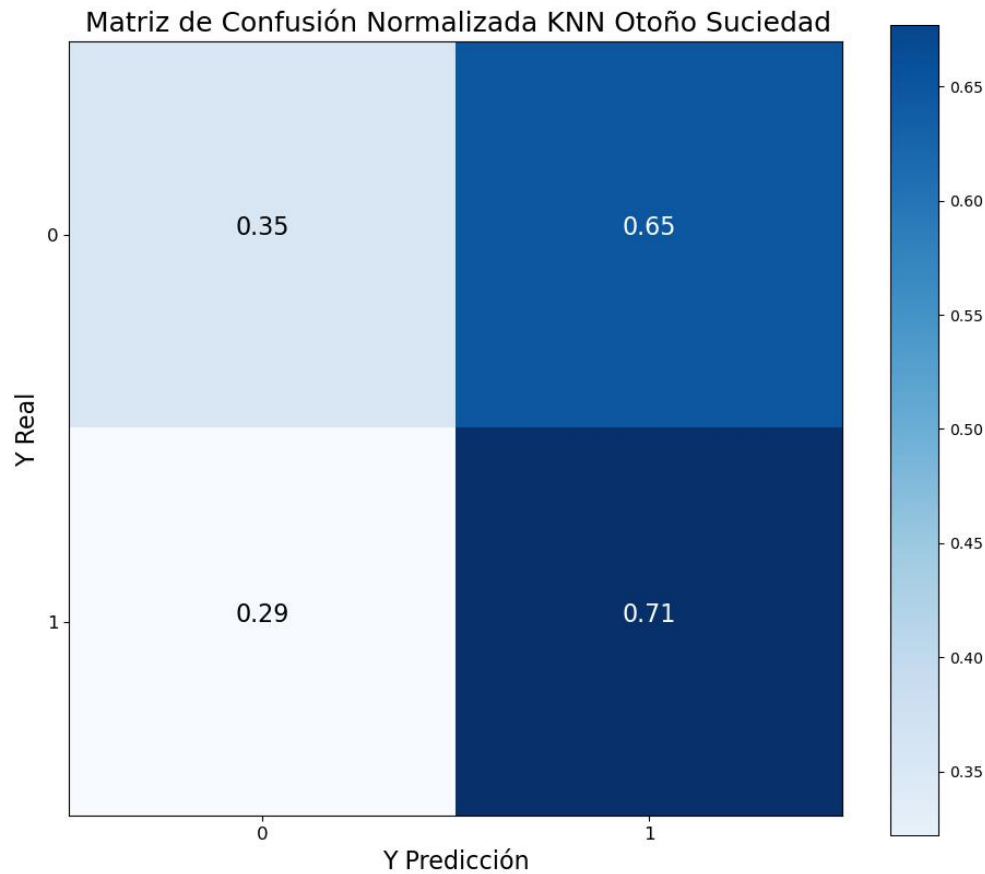


Figura 5.23: Matriz de confusión normalizada Algoritmo KNN Otoño validación de resultados suciedad

Fuente: Elaboración Propia

La figura 5.23 muestra en porcentajes cómo clasifica el algoritmo de vecinos cercanos. Podemos observar que su mayor dificultad radica en identificar cuando el sistema fotovoltaico no presenta suciedad, logrando acertar solo en el 35% de los casos, lo que resulta en 81 casos de falsos positivos, como indica la tabla 5.24.

Por otro lado, el algoritmo detecta correctamente la suciedad en los paneles fotovoltaicos en el 71% de los casos, es decir, 89 de 125 casos. Esto es un resultado favorable. No obstante, debido a la gran cantidad de falsos positivos, el algoritmo indica que la MicroRed UCSC tiene los paneles con suciedad más veces de lo que realmente sucede.

5.5.3 Algoritmos de clasificación de sombreado y suciedad

Los resultados de los algoritmos de clasificación múltiple de sombreado y suciedad se presentan en las tablas 5.25 – 5.28. comenzado por el modelo de vecinos cercanos

Tabla 5.25: Métricas de rendimiento KNN Otoño Validación de Sombreado y Suciedad

Metrica	Clase Y = 0	Clase Y = 1	Clase Y = 2	Tiempo de entrenamiento	Tiempo de validación
Exactitud		49.07 %			
Precision	40.7%	68.2%	31.9%	9.12 Segundos	3.76 Segundos
Sensibilidad	44.0%	77.9%	24.8%		
F1 Score	42.3%	72.7%	27.9%		

Fuente: Elaboración Propia

Tabla 5.26: Distribución Matriz de confusión KNN Otoño validación sombreado y suciedad

Predicción			
	Clase 0 (Limpio)	Clase 1 (Sombreado)	Clase 2 (Sucio)
Clase 0 (Limpio)	VP Clase 0 = 55	FP Clase 0 = 15	FP Clase 0 = 55
Clase 1 (Sombreado)	FP Clase 1 = 17	VP Clase 1 = 99	FP Clase 1 = 11
Clase 2 (Sucio)	FP Clase 2 = 63	FP Clase 2 = 31	VP Clase 2 = 31

Fuente: Elaboración Propia

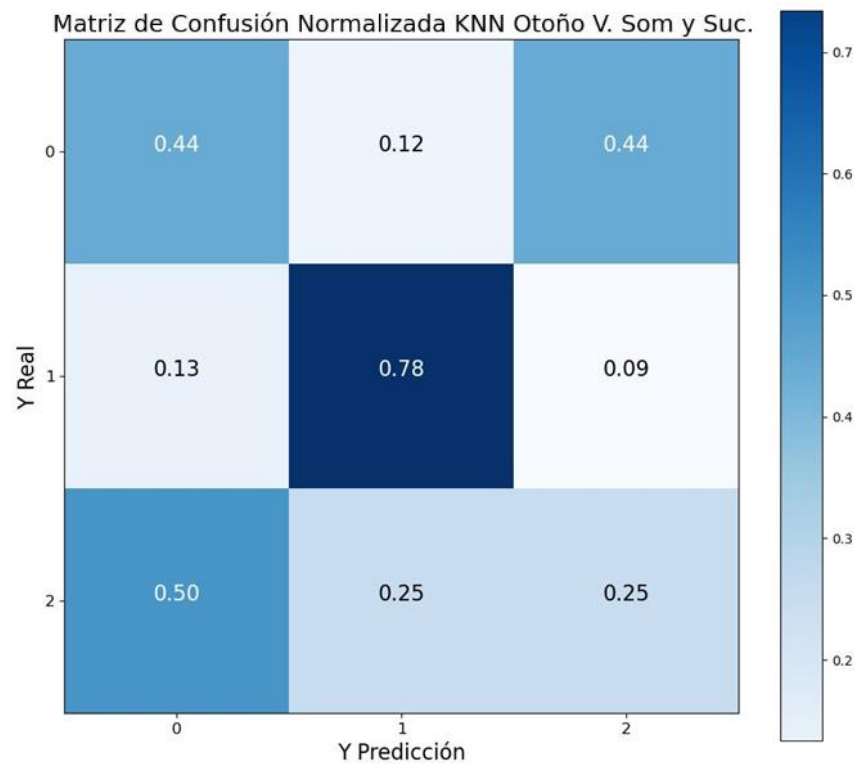


Figura 5.24: Matriz de confusión normalizada KNN Otoño validación de resultados sombreado y suciedad

Fuente: Elaboración Propia

La figura 5.24 muestra en porcentaje cómo clasifica el algoritmo de vecinos cercanos los datos de prueba de sombreado y suciedad del mes de mayo, siendo el caso de sombreado de los paneles FV lo que mejor clasifica.

Los resultados que muestran las tablas 5.25 y 5.26 muestran métricas bastante promedio del modelo de vecinos cercanos teniendo una precisión del 49 % lo que indica que clasifica de manera correcta la mitad de los datos de validación, siendo más efectivo en identificar los casos de sombreado ya que su F1 Score indica un 72.7 %. Cuando se trata de identificar la suciedad en los paneles este algoritmo tiene lo identifica correctamente en un 31.4 % de las veces, lo que es un mal resultado para el modelo. Como se puede ver en la tabla 5.26 la mayoría de los datos de suciedad el algoritmo los identifica como si los paneles estuvieran limpios en 63 de 125 de los casos. En los casos donde los paneles están completamente limpios el sistema clasifica de manera correcta el 40.7 % de las veces.

A continuación, se darán a conocer los resultados obtenidos del algoritmo red neuronal para clasificación de sombreado y suciedad.

Tabla 5.27: Métricas de rendimiento Red Neuronal Otoño validación de sombreado y suciedad

Metrica	Clase Y = 0	Clase Y = 1	Clase Y = 2	Tiempo de entrenamiento	Tiempo de Validación
Exactitud		45.88 %		10:52 Minutos	4.02 Segundos
Precision	32.8 %	82.67 %	21.6 %		
Sensibilidad	27.33 %	82.03 %	27.27 %		
F1 Score	29.81 %	82.35 %	24.10 %		

Fuente: Elaboración Propia

Tabla 5.28: Distribución Matriz de confusión RN Otoño validación de resultados sombreado y suciedad

Predicción			
	Clase 0 (Limpio)	Clase 1 (Sombreado)	Clase 2 (Sucio)
Clase 0 (Limpio)	VP Clase 0 = 41	FP Clase 0 = 16	FP Clase 0 = 93
Clase 1 (Sombreado)	FP Clase 1 = 18	VP Clase 1 = 105	FP Clase 1 = 5
Clase 2 (Sucio)	FP Clase 2 = 66	FP Clase 2 = 6	VP Clase 2 = 27

Fuente: Elaboración Propia

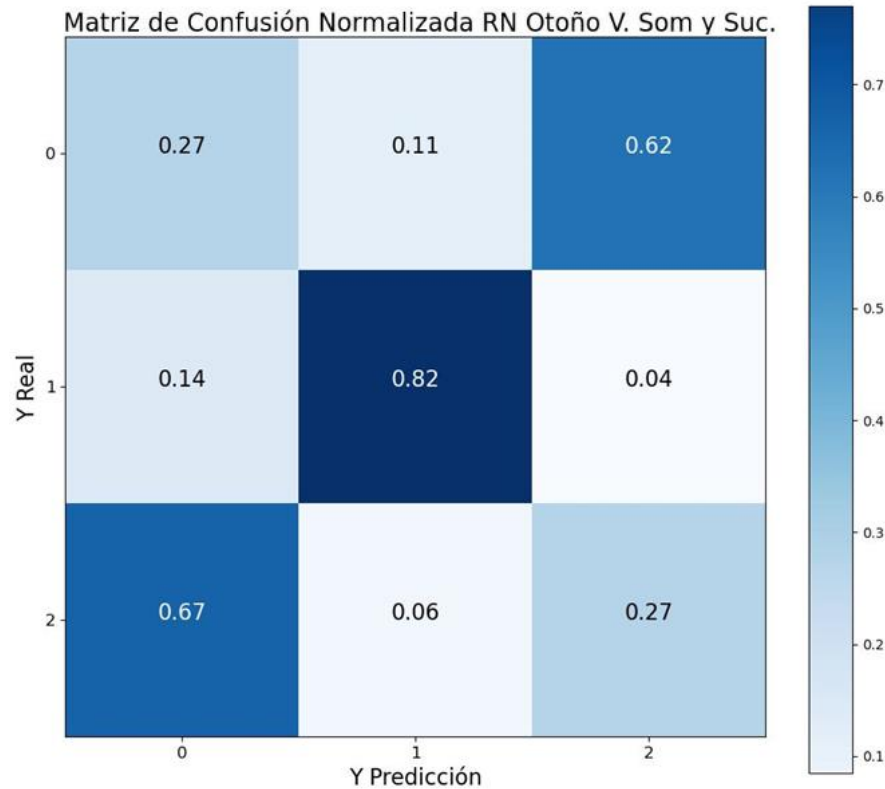


Figura 5.25: Matriz de confusión normalizada RN otoño validación resultados sombreado y suciedad

Fuente: Elaboración Propia

Los resultados de la red neuronal con la base de datos de validación, se muestran en las tablas 5.27 y 5.28, muestran un mal rendimiento, teniendo una precisión de un 45.88 %. Al igual que con el modelo de vecinos cercanos, lo que mejor identifica el modelo son los casos de sombreado teniendo una precisión de 82.67%, identificando correctamente 105 de 127 casos. Por otro lado en la tabla 5.27 se puede apreciar que en las clases 0 y 2 tienen métricas bastante más bajas llegando teniendo un F1 Score del 29.81% y 24.10% respectivamente, si vemos en detalle la tabla 5.28 podemos fijarnos que el algoritmo tiende a etiquetar mal los casos donde los paneles están limpios, clasificándolos como sucios en 66 oportunidades y algo similar sucede en el caso contrario ya que identifica situaciones en las que el sistema presenta suciedad como si este estuviera limpio en 93 casos, lo que si bien no es un buen resultado para el algoritmo, indica que este identifica de mejor manera las diferentes situaciones.

Si recordamos lo visto en el capítulo 3, cuando se sombreada el sistema FV la anomalía era evidente en los gráficos de voltaje ya que este disminuía en gran medida en comparación del voltaje de referencia. Por otro lado, cuando se analizaban los casos de suciedad en los gráficos en muchos de

esto no se podía apreciar diferencia, por lo que se tuvo que evaluar punto a punto las curvas de voltaje y corriente para notar las variaciones que se presentaban. Por lo que, si bien algunas de sus métricas de rendimiento son menores, logra comprender de mejor manera los datos que se le entregaron a comparación del modelo de vecinos cercanos que clasificó datos de las diferentes clases incorrectamente generando falsos positivos en las distintas clases.

Si bien los resultados con esta validación de resultados no son muy favorables hay varias consideraciones que mencionar

- Los modelos fueron entrenados con datos de octubre y enero meses donde las condiciones ambientales son muy diferentes, en especial la radiación, temperatura ambiente y la humedad relativa.
- La cantidad de datos de la base de datos son pocos, además en la construcción de esta se realizaron 3 pruebas diferentes en un solo día, por lo que está presente el error humano, al interferir cada cierto tiempo en la irradiancia que percibían los paneles y posiblemente generando sombreado en algunos momentos.

5.6 Validación de resultados con datos primavera y/o verano.

Para llevar a cabo esta validación de resultados se separaron los datos recolectados en las pruebas realizadas en octubre y enero, en tres partes. Para entrenar los modelos se utilizó un 60 % de los datos para el test se utilizaron el 25 % de estos y finalmente para la validación se utilizaron el 15% restante.

5.6.1 Algoritmos de clasificación de sombreado

Los resultados de los algoritmos que identifican el sombreado en los paneles FV se muestran en la siguiente tabla.

Tabla 5.29: Comparación de rendimientos de Validación algoritmos para detección de sombreado primavera

Algoritmo	Exactitud	Precision	Sensibilidad	F1 Score
Regresión Logística	97.66 %	98.65%	98.00%	98.32 %
Vecinos cercanos	99.06 %	100 %	98.66%	99.32%

Fuente: Elaboración Propia

Si comparamos la tabla 5.29 con la tabla 5.21 los resultados varían bastante el algoritmo de regresión logística, pasa de etiquetar el 71% de los casos bien a un 97.66%, mientras que el modelo de vecinos cercanos aumenta de un 91% a un 99%.

5.6.2 Algoritmos de clasificación de suciedad

Los resultados de los algoritmos que identifican la suciedad en los módulos fotovoltaicos se muestran en la siguiente tabla.

Tabla 5.30: Comparación de rendimientos de Validación algoritmos para detección de suciedad verano

Algoritmo	Exactitud	Precision	Sensibilidad	F1 Score
Regresión Logística	71.26 %	69.39 %	86.98 %	77.20 %
Vecinos cercanos	91.57 %	90.25 %	95.20 %	92.66 %

Fuente: Elaboración Propia

En los algoritmos de identificación de suciedad es donde mayor diferencia hay respecto al caso de validación anterior mostrado en la tabla 5.23 donde el modelo de regresión logística tiene una exactitud del 5.6%, para la validación, con los nuevos datos llega al 71.26% de exactitud y un F1 Score del 77.20 %. Por otro lado, el modelo de vecinos identifico correctamente un 91.57 % de los casos de suciedad mientras que en el otro escenario solo lo hizo el 51% de las ocasiones.

5.6.3 Algoritmos de clasificación de sombreado y suciedad

Los resultados de los algoritmos que identifican el sombreado y la suciedad en los módulos fotovoltaicos se muestran en las siguientes tablas

Tabla 5.31: Métricas de rendimiento de Validación algoritmo KNN Sombreado y Suciedad (primavera - verano)

Metrica	Clase Y = 0	Clase Y = 1	Clase Y = 2
Exactitud		95.7 %	
Precision	95.6%	100 %	92.4%
Sensibilidad	93.54%	100 %	94.83%
F1 Score	94.45%	100 %	93.63%

Fuente: Elaboración Propia

Tabla 5.32: Métricas de rendimiento de Validación Red Neuronal Sombreado y Suciedad (primavera - verano)

Metrica	Clase Y = 0	Clase Y = 1	Clase Y = 2
Exactitud		95.14 %	
Precision	93.01%	97.74%	95.48%
Sensibilidad	94.53%	100 %	91.92%
F1 Score	93.76%	98.85%	93.67%

Fuente: Elaboración Propia

Si comparamos los nuevos resultados obtenidos con los datos de validación de la misma fecha que los datos con los que se entrenaron los modelos, podemos ver que existe una diferencia muy grande y una mejoría notable en las métricas de rendimiento. Algoritmos que antes no superaban el 50 % de exactitud, tienen una precisión de superior al 90 %. Por lo que se concluye que a los diferentes algoritmos les falta ser entrenados con datos de las diferentes épocas del año.

Si bien la suciedad es un problema que se presenta generalmente en la época de verano en la región del Bio-Bio, debido a las pocas lluvias que hay durante esos meses. No es el caso del sombreado, Además existe la posibilidad de tener pocas precipitaciones en el invierno lo que afectaría en la limpieza natural de estos módulos. la detección del sombreado es llevaba a cabo con métricas de rendimiento bastante buenas, en la mayoría de los casos igual queda un margen para la mejora.

El mayor desafío está en la identificación de la presencia de suciedad en el sistema a través del algoritmo, ya que no es una condición fácil de detectar, a menos que se analizan los datos en detalle.

La incorporación del machine learning en los sistemas fotovoltaicos a demostrado ser bastante útil, ya que facilita la identificación de estas condiciones de falla en los paneles FV de la Microred UCSC.

Capítulo 6. Conclusiones

6.1. Sumario

En este trabajo, se construyeron diversos algoritmos para la detección de sombreado y/o suciedad, utilizando modelos de regresión logística (RL), vecinos cercanos (KNN) y redes neuronales (RN). Además, se abordan las distintas fallas que pueden afectar a un sistema fotovoltaico, tanto en sus componentes como en los paneles solares. Se pone un enfoque especial en las fallas de sombreado y suciedad, proporcionando una descripción de las principales causas de estas fallas y las posibles consecuencias que pueden generar.

Para identificar los efectos reales del sombreado y la suciedad, se llevaron a cabo pruebas en la Microred UCSC, evaluando el impacto concreto de estas condiciones en el sistema, en particular en las variables eléctricas de corriente y voltaje en el lado de corriente continua. Los datos recolectados durante estas pruebas se emplearon para construir una base de datos.

Se obtuvieron cuatro conjuntos de datos: sombreado obtenido durante el mes de octubre (primavera), suciedad obtenida en enero (verano), suciedad y sombreado obtenidos de la combinación de los datos de primavera y verano. El último conjunto de datos corresponde a sombreado y suciedad obtenidos en el mes de mayo (otoño). Estos conjuntos de datos se utilizaron en el entrenamiento y la validación de los modelos de clasificación desarrollados.

Se presenta una explicación teórica de los distintos algoritmos utilizados, con el objetivo de comprender su funcionamiento y los resultados que ofrecen. Se detallan las estructuras y los hiperparámetros de los diferentes algoritmos implementados. Finalmente, se realiza un análisis exhaustivo de los resultados obtenidos por los modelos entrenados, tanto en la etapa de entrenamiento como en la validación de los resultados.

6.2. Conclusiones

A través de esta investigación se pudo ver el efecto real que genera el sombreado y la suciedad en la Microred UCSC. El sombreado es el que mayor impacto produce en la generación de energía eléctrica, disminuyendo el voltaje en cerca de 10 voltios cuando se sombrea un panel y hasta en 82 voltios cuando hay sombreado en 4 paneles FV de un total de 18 que se conectan a la misma entrada del inversor, lo que representa una reducción del 30% de su valor de referencia. Además, genera una disminución del 4% en la corriente del string 1.

La suciedad es un caso distinto, ya que para poder identificar variaciones en la corriente o el voltaje, es necesario evaluar punto a punto respecto a la señal de corriente y voltaje usadas como referencia (string 1 del inversor Fronius SYMO n°1). En los casos donde tiene mayor impacto, es cuando se ensucian 2 o más módulos fotovoltaicos, generando un pequeño aumento en el voltaje de alrededor de 5 [V] respecto a los valores de voltaje registrados cuando el string no presentaba suciedad. También se produce una disminución en la corriente, la cual es más perceptible mientras mayor cantidad de paneles presenten suciedad, llegando a disminuir en 0.66 [A], que es un 5% de la corriente de referencia. Si bien los efectos de la suciedad no afectan mucho la producción de energía, el no realizar un correcto mantenimiento puede producir fallas internas en los módulos fotovoltaicos, como puntos calientes y fallas en los módulos.

Los algoritmos entrenados fueron capaces de clasificar de manera correcta las situaciones de sombreado y/o suciedad, sobre todo en el tiempo de primavera y verano, llegando a tener una precisión del 95% con exactitud, y alrededor de un 50% de exactitud con datos del mes de mayo, es decir, en temporada de otoño - invierno.

Si bien las redes neuronales utilizadas para identificar la suciedad y el sombreado entregan un excelente rendimiento, el algoritmo de vecinos cercanos tiene métricas de rendimiento iguales o mejores en algunos casos. Dado el tiempo que se demora en entrenar el modelo y el costo computacional que tiene, es preferible utilizar el algoritmo KNN en este caso.

6.3. Trabajo Futuro

Es necesario entrenar los modelos construidos con nuevos datos, especialmente de las épocas de otoño e invierno, para crear un algoritmo robusto capaz de identificar los casos de sombreado y suciedad en cualquier época del año. De esta manera, obtendremos un algoritmo listo para su implementación.

Mediante un método de procesamiento de imágenes, podremos identificar la presencia de suciedad en los paneles fotovoltaicos y cuantificar el porcentaje de superficie que cubre, así como el efecto que provoca.

El objetivo es implementar un algoritmo capaz de identificar tanto el sombreado como la suciedad en toda la Microred UCSC durante todo el año.

Bibliografía

- [1] H. Khonkar *et al.*, «Importance of cleaning concentrated photovoltaic arrays in a desert environment», *Sol. Energy*, vol. 110, pp. 268-275, dic. 2014, doi: 10.1016/j.solener.2014.08.001.
- [2] H. Mekki, A. Mellit, y H. Salhi, «Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules», *Simul. Model. Pract. Theory*, vol. 67, pp. 1-13, sep. 2016, doi: 10.1016/j.simpat.2016.05.005.
- [3] S. Theocharides, G. Makrides, G. E. Georghiou, y A. Kyprianou, «Machine learning algorithms for photovoltaic system power output prediction», en *2018 IEEE International Energy Conference (ENERGYCON)*, Limassol: IEEE, jun. 2018, pp. 1-6. doi: 10.1109/ENERGYCON.2018.8398737.
- [4] M. Heinrich *et al.*, «Detection of cleaning interventions on photovoltaic modules with machine learning», *Appl. Energy*, vol. 263, p. 114642, abr. 2020, doi: 10.1016/j.apenergy.2020.114642.
- [5] F. Han *et al.*, «An intelligent fault diagnosis method for PV arrays based on an improved rotation forest algorithm», *Energy Procedia*, vol. 158, pp. 6132-6138, feb. 2019, doi: 10.1016/j.egypro.2019.01.498.
- [6] H. Zitouni *et al.*, «Experimental investigation and modeling of photovoltaic soiling loss as a function of environmental variables: A case study of semi-arid climate», *Sol. Energy Mater. Sol. Cells*, vol. 221, p. 110874, mar. 2021, doi: 10.1016/j.solmat.2020.110874.
- [7] J. Kim, M. Rabelo, S. P. Padi, H. Yousuf, E.-C. Cho, y J. Yi, «A Review of the Degradation of Photovoltaic Modules for Life Expectancy», *Energies*, vol. 14, n.º 14, p. 4278, jul. 2021, doi: 10.3390/en14144278.

- [8] A. Mellit, G. M. Tina, y S. A. Kalogirou, «Fault detection and diagnosis methods for photovoltaic systems: A review», *Renew. Sustain. Energy Rev.*, vol. 91, pp. 1-17, ago. 2018, doi: 10.1016/j.rser.2018.03.062.
- [9] R. R. Cordero *et al.*, «Effects of soiling on photovoltaic (PV) modules in the Atacama Desert», *Sci. Rep.*, vol. 8, n.º 1, p. 13943, sep. 2018, doi: 10.1038/s41598-018-32291-8.
- [10] Alderete, L. (2019) «Diseño e implementación de experiencias en redes On-grid en laboratorio Microred UCSC» (Tesis de grado Ingeniero Civil Eléctrico) Universidad Católica de la santísima Concepción.

Imágenes utilizadas

- {1} <https://www.areatecnologia.com/electricidad/paneles-solares.html>
- {2} <https://www.dsisolar.com/info/pv-junction-box-s-bypass-diode-for-solar-panel-54221810.html>
- {3} <https://diysolarforum.com/threads/discussion-pv-series-vs-parallel-is-parallel-really-better-for-partial-shading-situations.59840/>
- {4} <https://www.aulafacil.com/cursos/medio-ambiente/energia-solar-fotovoltaica-1/sombras-diodos-y-bypass-l37871>
- {5} https://rstudio-pubs-static.s3.amazonaws.com/402754_6cbdea25a79d43f4895cfd7df0a8bd07.html

Anexos

Anexo 1: Códigos ejemplo de los distintos modelos utilizados

Regresión Logística

```
# Primero se importan los modulos con los que se trabajaran
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
# Segundo paso importar la base de datos con la que se trabajara
X = Variables independientes
y = Variable dependiente
# Tercero dividir los datos en conjunto de entrenamiento y conjunto de prueba,
también se puede dividir en datos de entrenamiento, validación cruzada y datos de
prueba como se estime conveniente
X_train, X_test, y_train, y_test = train_test_split(X, y,)
# Cuarto Crear el modelo en este caso de regresión logística y ajustar el modelo a los
datos de entrenamiento
Modelo = LogisticRegression()
Modelo.fit(X_train, y_train)

# Quinto predecir la variable dependiente con el conjunto de datos de prueba
y_pred = Modelo.predict(X_test)

# Sexto calcular la precisión del modelo utilizando los datos de prueba
# Matriz de confusión
from sklearn.metrics import confusion_matrix
matriz = confusion_matrix(Y_test, Y_pred)
# Exactitud
from sklearn.metrics import accuracy_score
exactitud = accuracy_score(Y_test, Y_pred)
```

Vecinos Cercanos

```
# Primero se importan los modulos con los que se trabajaran
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split

# Segundo paso importar la base de datos con la que se trabajara
X = Variables independientes
y = Variable dependiente

# Tercero dividir los datos en conjunto de entrenamiento y conjunto de prueba
X_train, X_test, y_train, y_test = train_test_split(X, y,)

# Cuarto Crear el modelo en este caso de vecinos cercanos y ajustar el modelo a los
datos de entrenamiento donde se elige cuantos vecinos se quiere considerar para
realizar la clasificacion en este código se especificara considerar 3
Modeloknn = KNeighborsClassifier(n_neighbors=3)
Modeloknn.fit(X_train, y_train)

# Quinto predecir la variable dependiente con el conjunto de datos de prueba
y_pred = Modeloknn.predict(X_test)

# Sexto calcular la precisión del modelo utilizando los datos de prueba utilizaremos
parámetros diferentes a los mostrados en el código de regresion logística

# Precision
from sklearn.metrics import precision_score
precision= precision_score(Y_test, Y_pred)

# Sencibilidad
from sklearn.metrics import recall_score
sencibilidad = recall_score(Y_test, Y_pred)

# Valor F1 (
from sklearn.metrics import f1_score
Valor_F1 = f1_score(Y_test, Y_pred)
```

Redes Neuronales (Salida Binaria)

```
# Primero se importan los modulos con los que se trabajaran
import tensorflow as tf
from tensorflow.keras.layers import Dense, Input
from tensorflow.keras import Sequential
from tensorflow.keras.losses import BinaryCrossentropy
from tensorflow.keras.activations import sigmoid, relu
from sklearn.model_selection import train_test_split

# Segundo paso importar la base de datos con la que se trabajara una diferencia de
trabajar con redes neuronales a diferencia de los otros modelos es que ambas variables
deben se bidimensionales
X = Variables independientes
y = Variable dependiente

# Tercero dividir los datos en conjunto de entrenamiento y conjunto de prueba
X_train, X_test, y_train, y_test = train_test_split(X, y,)

# Cuarto crear el modelo de red neuronal con la forma secuencial especificando la
cantidad de neuronas por capa, la funcion de activación y el nombre de las capas.
model = Sequential([
    Dense(3, activation = 'relu', name = "Capa1"),
    Dense(1, activation = 'sigmoid', name = "Capa2"),])

# Quinto ajustar el modelo a los datos de entrenamiento definiendo la función de perdida
(BinaryCrossentropy), optimizador (Adam) en el cual se define la tasa de aprendizaje
0.01.
model.compile(
    loss = tf.keras.losses.BinaryCrossentropy(from_logits=False), # funcion de perdida
    optimizer = tf.keras.optimizers.Adam(learning_rate=0.01),) # optimizador
model.fit(X_train_scaled, Y_train, epochs = 200) # Ajuste del
modelos

# Sexto predecir la variable dependiente con el conjunto de datos de prueba
y_pred = model.predict(X_test)
```

Red Neuronal (Salida Multiclase)

```
# Primero se importan los modulos con los que se trabajaran
import tensorflow as tf
from tensorflow.keras.layers import Dense, Input
from tensorflow.keras import Sequential
from tensorflow.keras.losses import SparseCategoricalCrossEntropy
from tensorflow.keras.activations import relu, softmax
from sklearn.model_selection import train_test_split

# Segundo paso importar la base de datos con la que se trabajara
X = Variables independientes
y = Variable dependiente

# Tercero dividir los datos en conjunto de entrenamiento y conjunto de prueba
X_train, X_test, y_train, y_test = train_test_split(X, y,)

# Cuarto crear el modelo de red neuronal con la forma secuencial especificando la
cantidad de neuronas por capa, la funcion de activación y el nombre de las capas. Al ser
una RN con salida multivariable la capa de salida debe tener la cantidad de neuronas igual
a las posibles salidas.
model = Sequential([
    Dense(16, activation = 'relu', name = 'Capa1'),
    Dense(32, activation = 'relu', name = 'Capa2'),
    Dense(3, activation = 'softmax', name = 'Capa3')])

# Quinto ajustar el modelo a los datos de entrenamiento definiendo la función de perdida
(SparseCategoricalCrossEntropy), optimizador (Adam) en el cual se define la tasa de
aprendizaje.
model.compile(
    loss = tf.keras.losses.SparseCategoricalCrossEntropy (from_logits = true), # funcion
de perdida
    optimizer = tf.keras.optimizers.Adam(learning_rate=0.01),) # optimizador
model.fit(X_train_scaled, Y_train, epochs = 200) # Ajuste del modelos

# Quinto predecir la variable dependiente con el conjunto de datos de prueba
y_pred = model.predict(X_test)
```

Anexo 2: Gráficas limpieza base de datos suciedad.

Corriente Continua

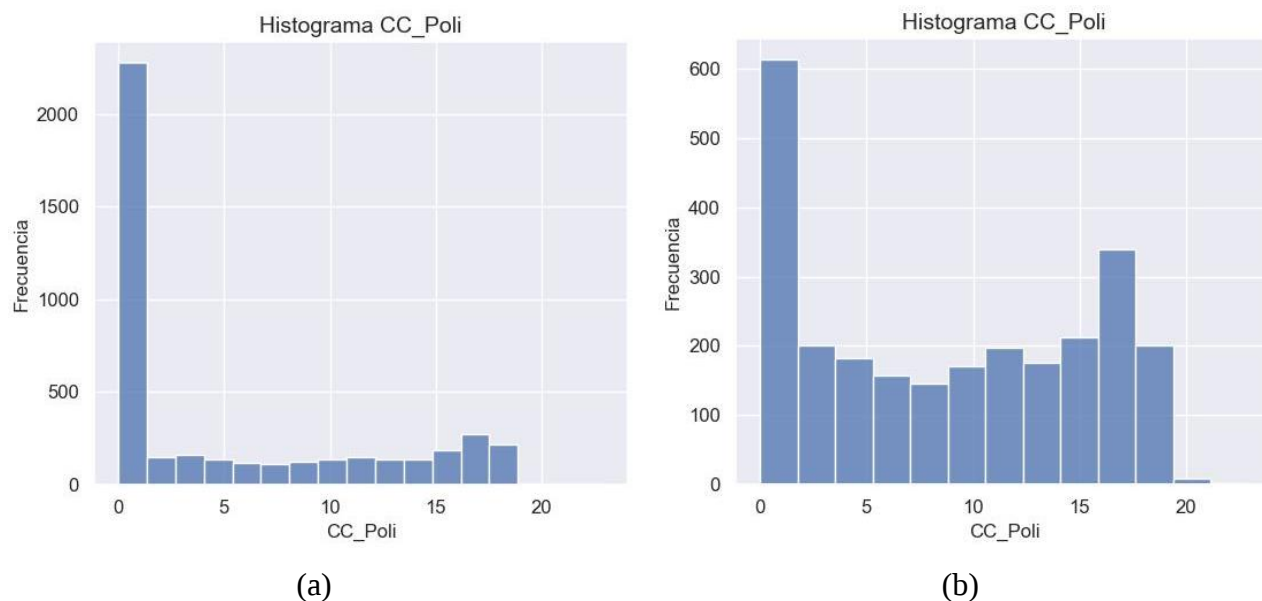


Figura 0.1 Comparación Corriente Continua base de datos de suciedad,
(a) Datos de Corriente sin Depuración, (b) Datos de Corriente post Depuración de Datos

Fuente: Elaboración Propia

Voltaje CC

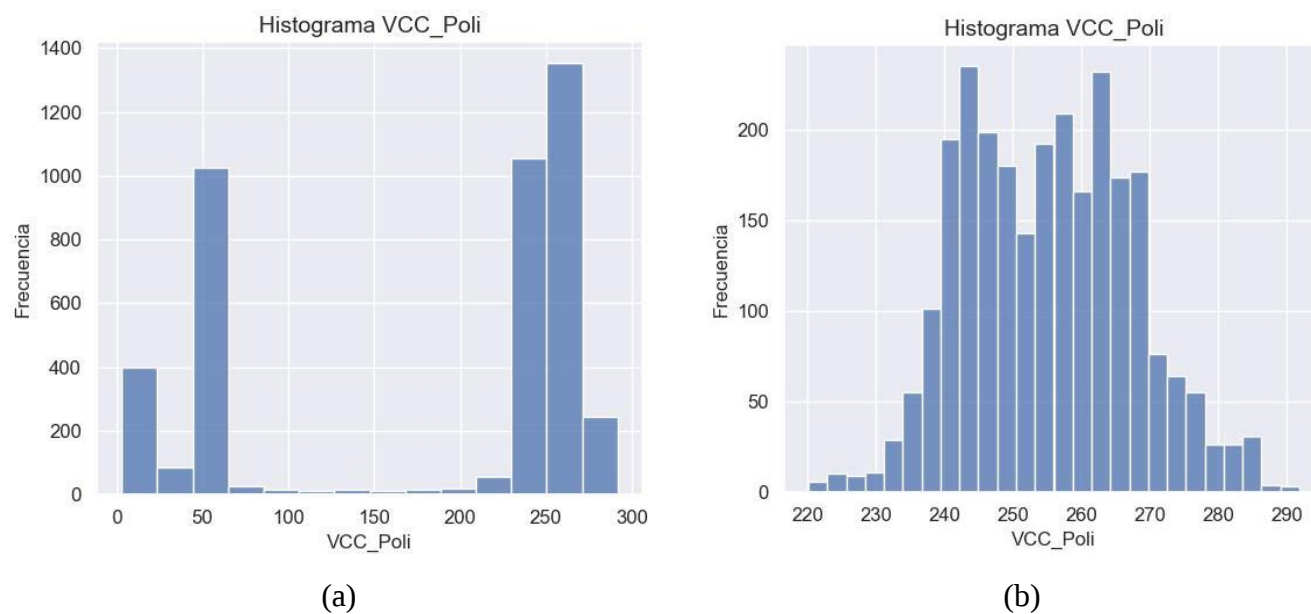


Figura 0.2 Comparación Voltaje en CC base de datos de suciedad,
(a) Datos de Voltaje sin Depuración, (b) Datos de Voltaje post Depuración de Datos.

Fuente: Elaboración Propia

Potencia CC

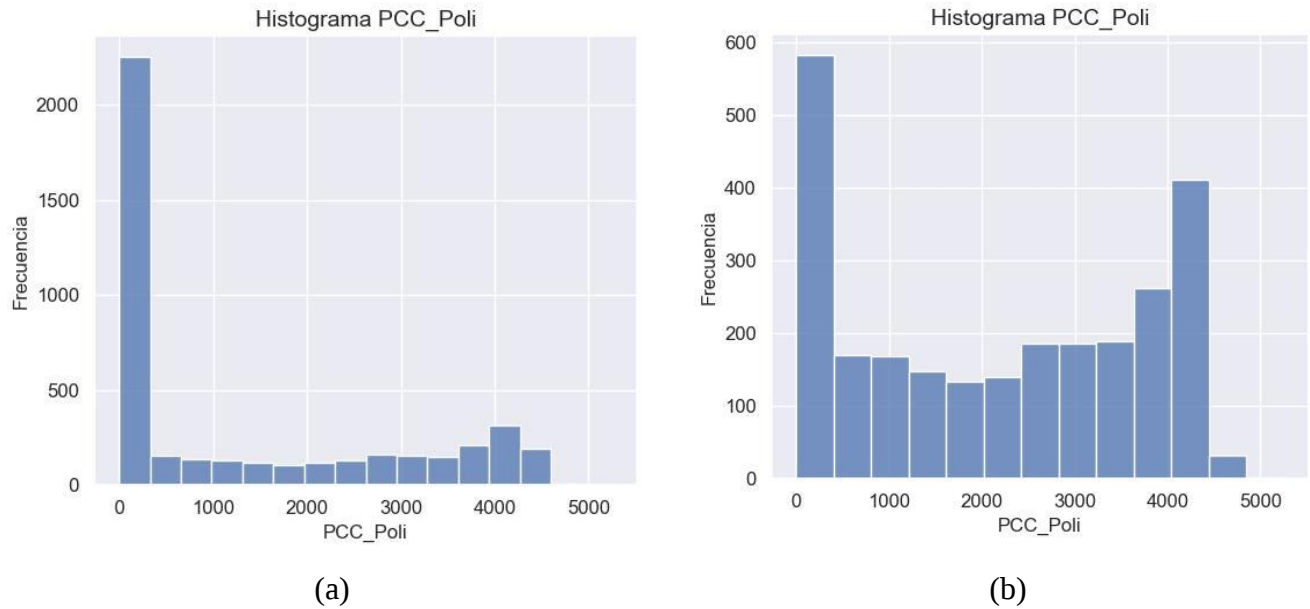


Figura 0.3 Comparación Potencia en CC base de datos de suciedad,
(a) Datos de Potencia sin Depuración, (b) Datos de Potencia post Depuración de Datos.

Fuente: Elaboración Propia

Temperatura del Panel Fotovoltaico

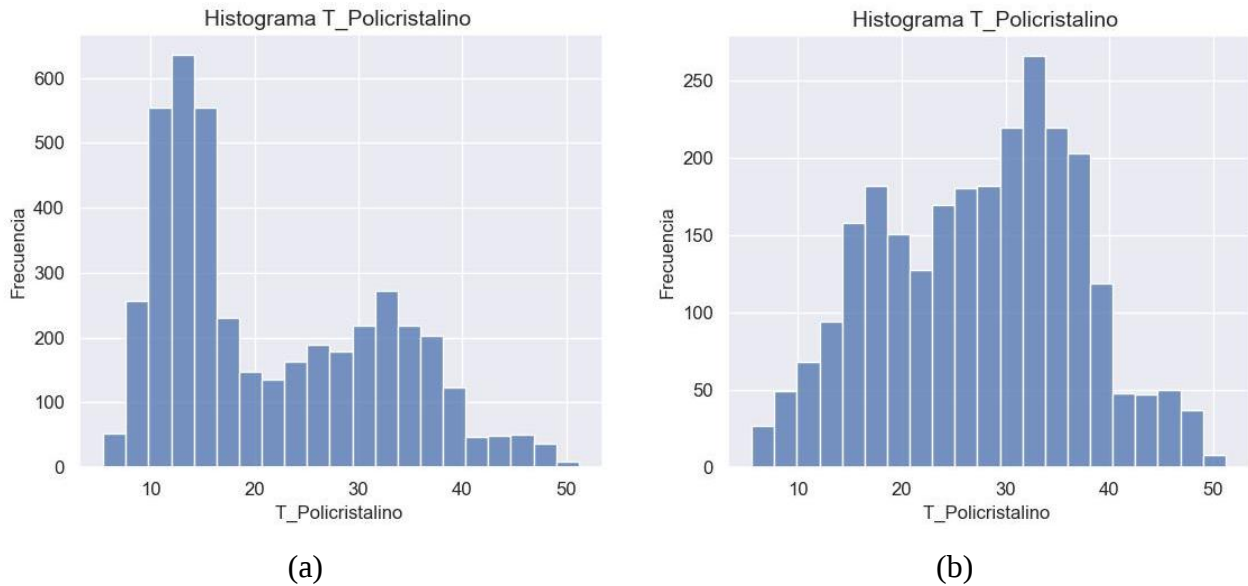
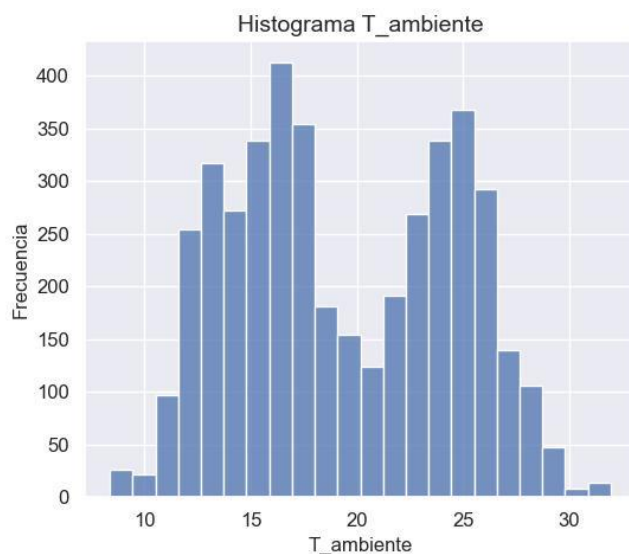


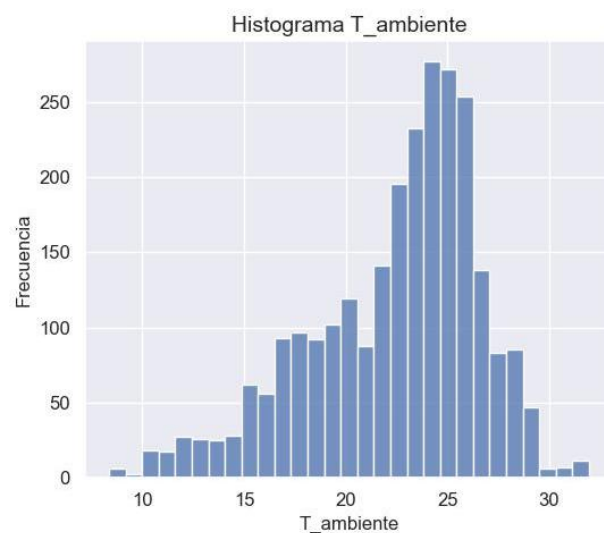
Figura 0.4 Comparación Temperatura del panel FV base de datos de suciedad,
(a) Datos de T. del panel FV sin Depuración, (b) Datos de T. del panel FV post Depuración de Datos

Fuente: Elaboración Propia

Temperatura ambiente



(a)

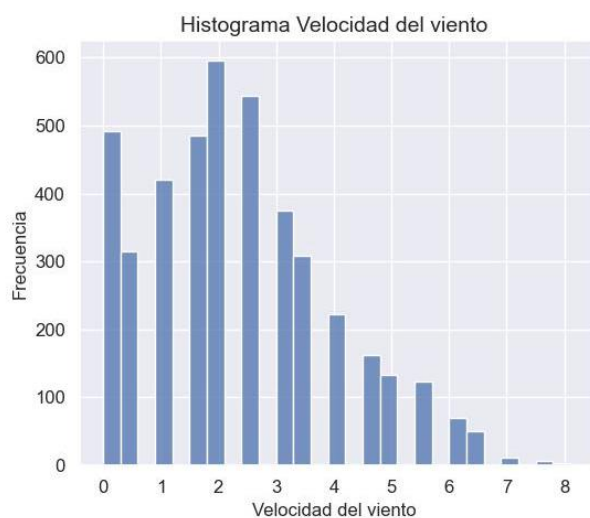


(b)

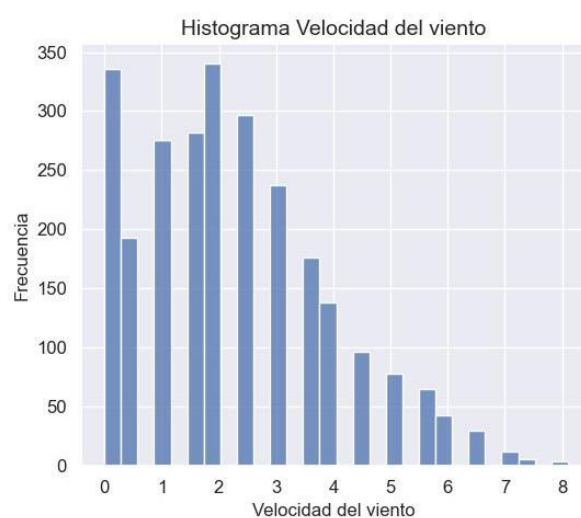
Figura 0.5 Comparación Temperatura Ambiente base de datos de suciedad,
(a) Datos de T. Ambiente sin Depuración, (b) Datos de T. Ambiente post Depuración de Datos

Fuente: Elaboración Propia

Velocidad del Viento



(a)



(b)

Figura 0.6 Comparación Velocidad del Viento base de datos de suciedad,
(a) Datos de Velocidad del Viento sin Depuración, (b) Datos de Velocidad del Viento post Depuración

Fuente: Elaboración Propia

Radiación

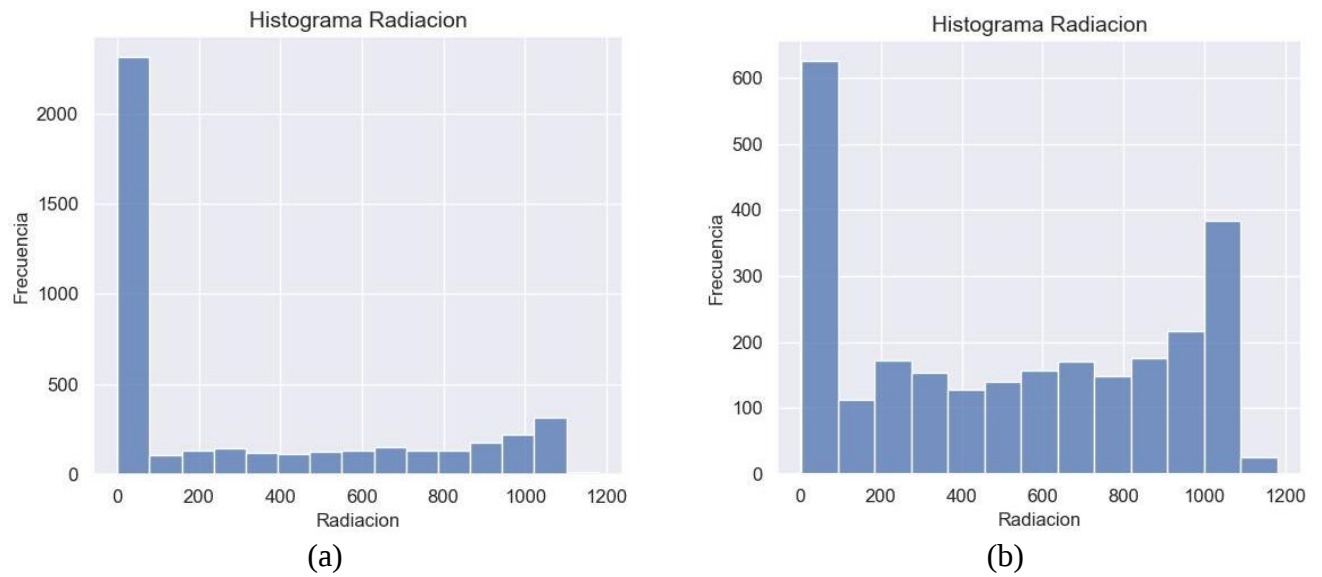


Figura 0.7 Comparación Radiación base de datos de suciedad,
(b) Datos de Radiación sin Depuración, (b) Datos Radiación post Depuración de Datos
Fuente: Elaboración Propia

Anexo 3: Correlación de variables en las diferentes bases de datos

Base de datos sombreado (Primavera)

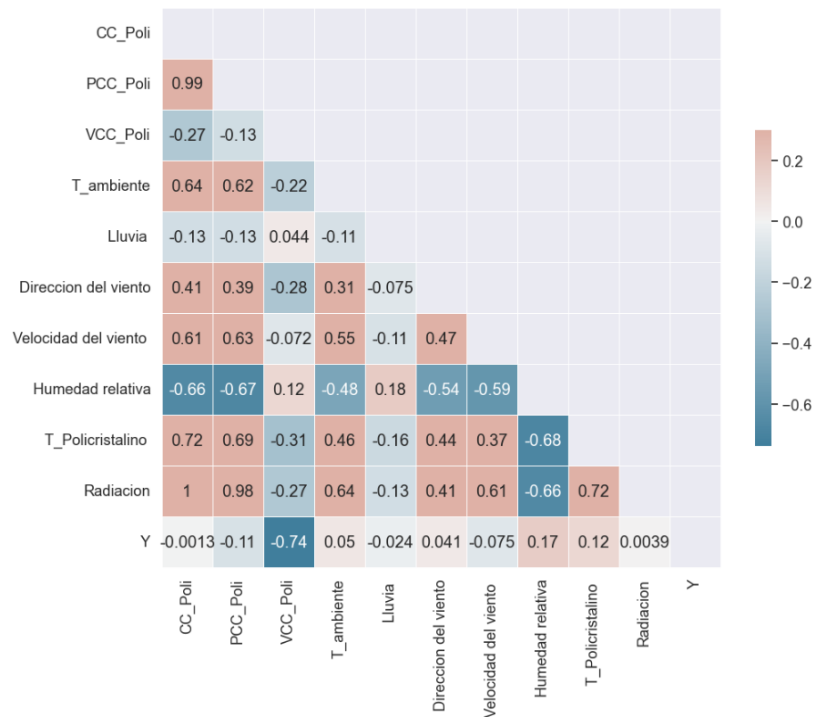


Figura 0.8: Grafica de correlación de variables datos de sombreado (Primavera)

Base de datos suciedad (Verano)

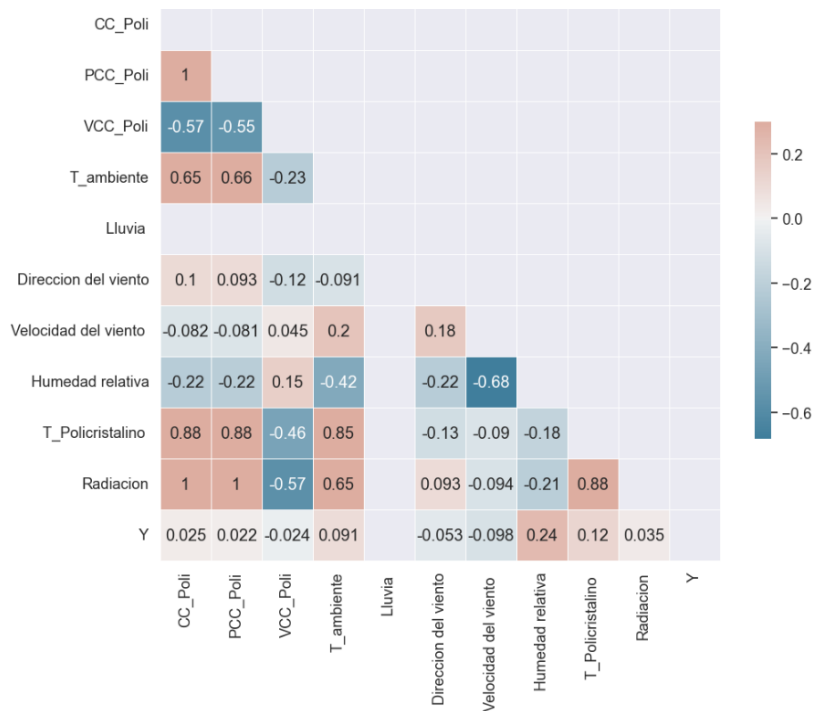


Figura 0.9: Grafica de correlación de variables datos de suciedad (Verano)

Base de datos sombreado y suciedad (Primavera - Verano)

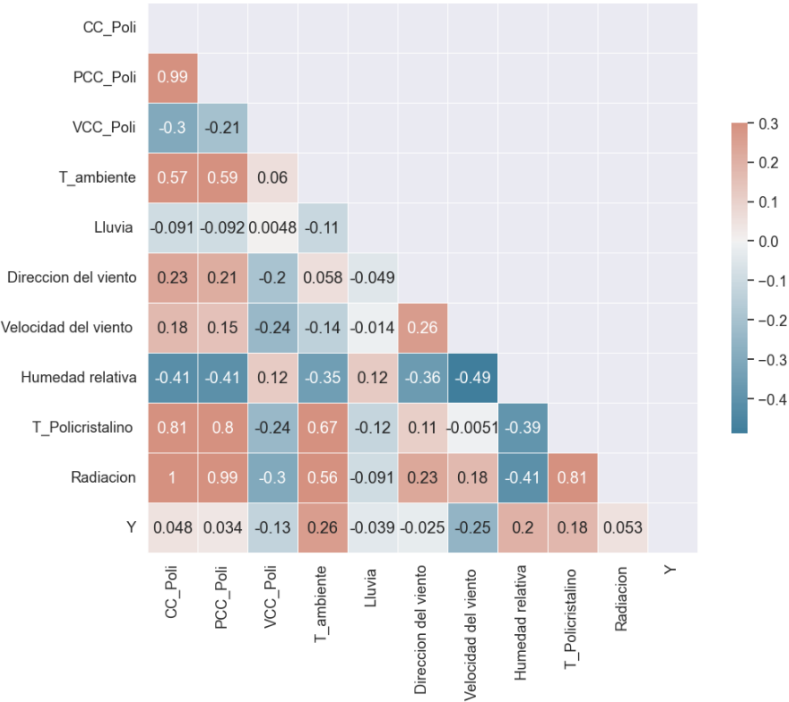


Figura 0.10: Grafica de correlación de variables datos de sombreado y suciedad (Primavera - Verano)