

SISTEMA DE PREDICCIÓN METEOROLÓGICA

PROYECTO 3





ÍNDICE

1. NEGOCIO
2. ENTENDER DATOS
3. PREPARAR DATOS
4. MODELACIÓN
5. EVALUACIÓN
6. DEMO

The Team



Aakriti Guerrero Bertolín



Oriol Fernández Font



Joan Pau Grau Homs



NEGOCIO



¿Quiénes somos?

En PREMATEO S.L. estamos comprometidos con el almacenamiento y análisis de millones de datos meteorológicos. Estos datos se utilizan exclusivamente para desarrollar sistemas avanzados de predicción meteorológica, el producto principal de nuestra empresa, diseñado para satisfacer la creciente demanda de diversas industrias.



Nuestros objetivos y criterios

Objetivos de Negocio	KPI (Éxito empresarial)
Precisión del sistema	Aumentar la precisión de los datos y predicciones del sistema en comparación a los sistemas actuales.
Velocidad de procesamiento	Disminuir el tiempo de espera a la hora de obtener un resultado y el tiempo de ejecución de los modelos entrenados.
Gestión de datos	Debido a la gran cantidad de datos que se reciben, es importante saber tratar los datos y obtener variables significativas.
Visualización de datos	Crear una interfaz user friendly a la vez que los datos se muestren de forma detallada.



NEGOCIO



Situación de partida

Riesgos	Contingencias	Supuestos	Restricciones
Datos históricos faltantes o poco fiables	Controles rigurosos de calidad y procesos de limpieza de datos	Datos históricos de calidad	Plazo ajustado para el desarrollo e implementación
Complejidad técnica	Desarrollo del modelo por fases con validaciones intermedias	Equipo con habilidades técnicas requeridas	Presupuesto limitado para infraestructura y personal
Restricciones de recursos computacionales	Provisión de infraestructura en la nube escalable	Demanda de mercado para predicciones meteorológicas	
Competencia	Monitoreo continuo del mercado y actualización ágil de la hoja de ruta del producto		



Objetivos y criterios en la minería de datos

ENTENDER DATOS

● Descripción de los datos

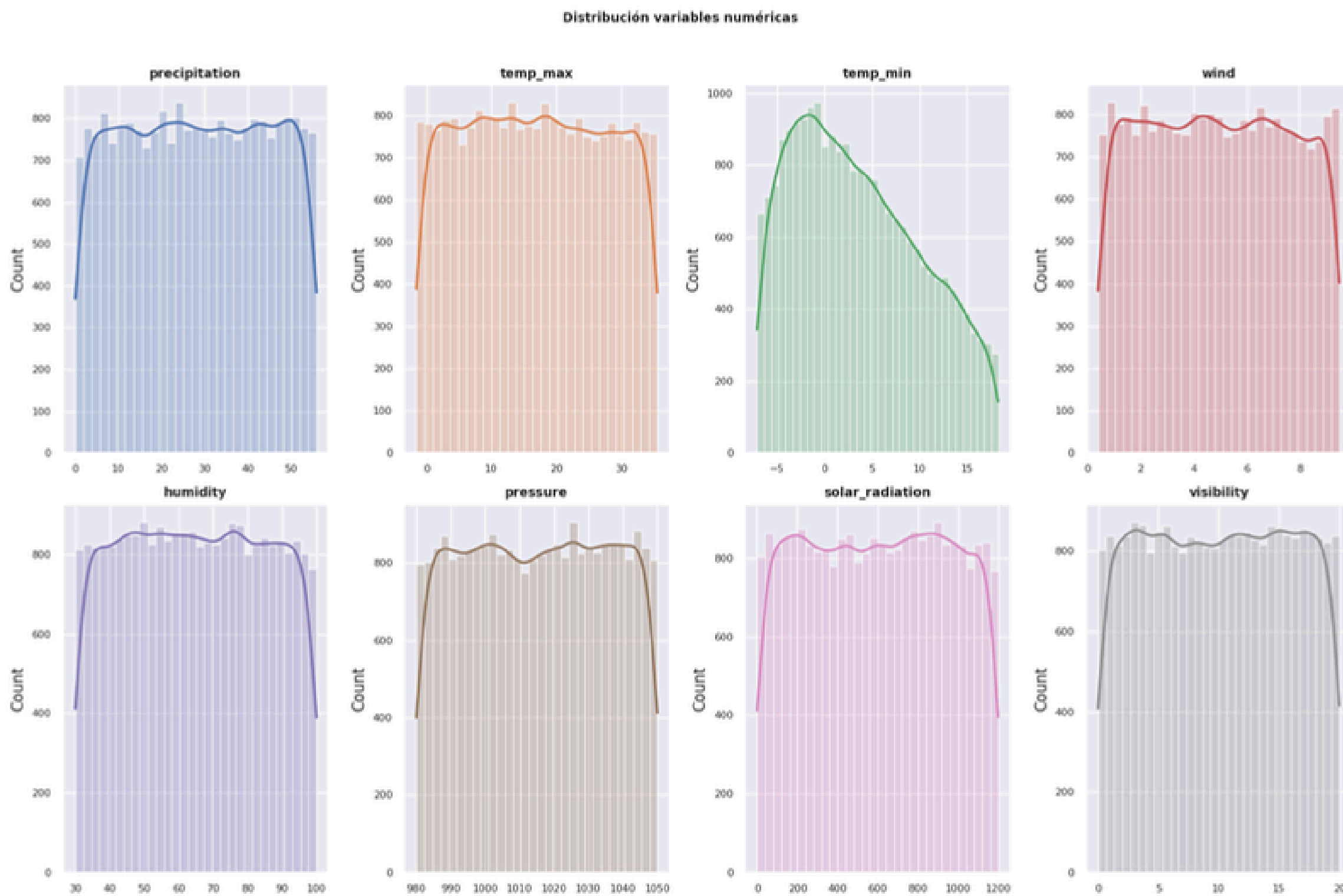
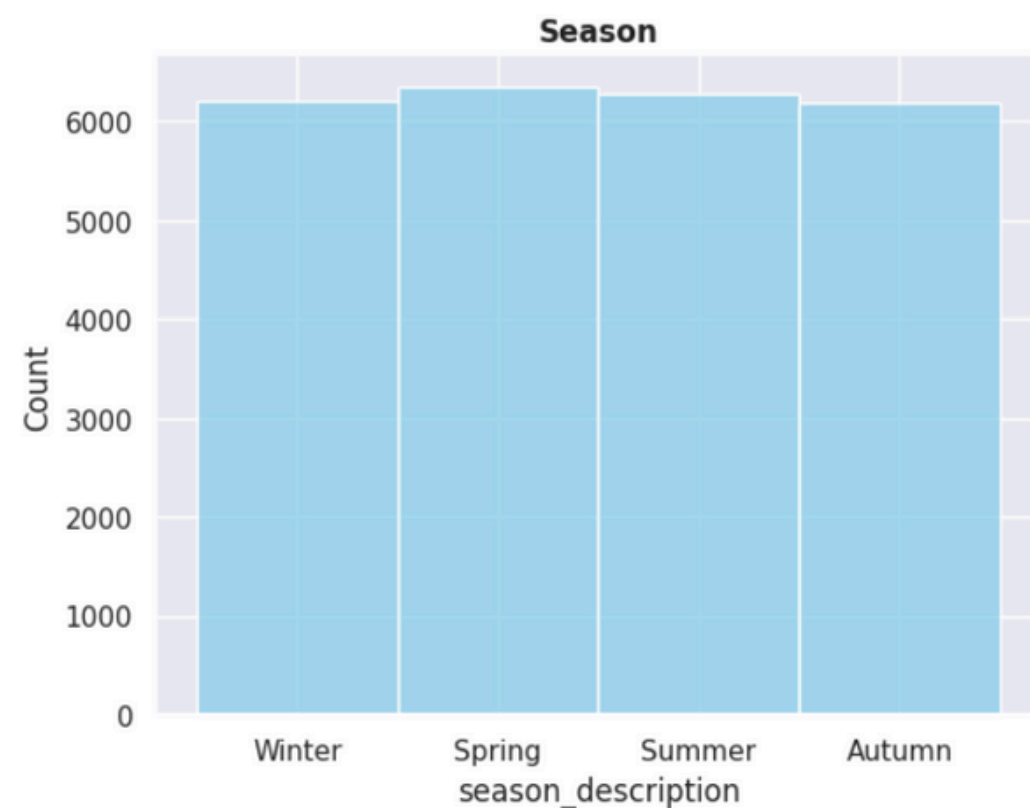
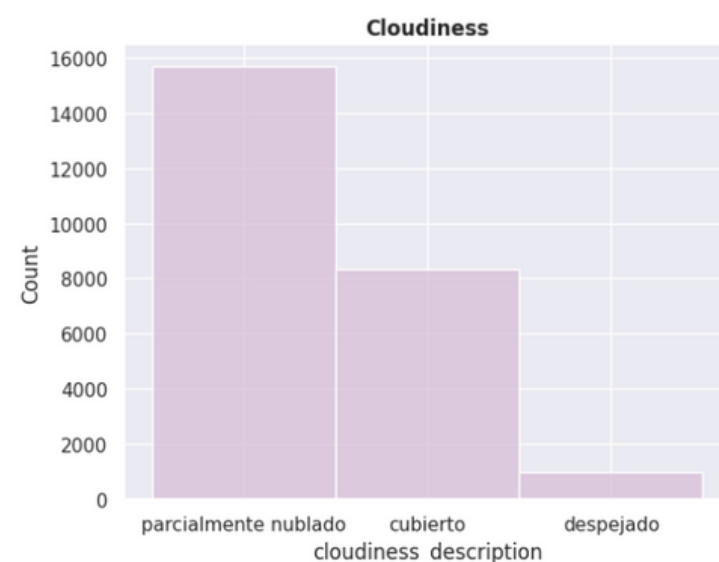
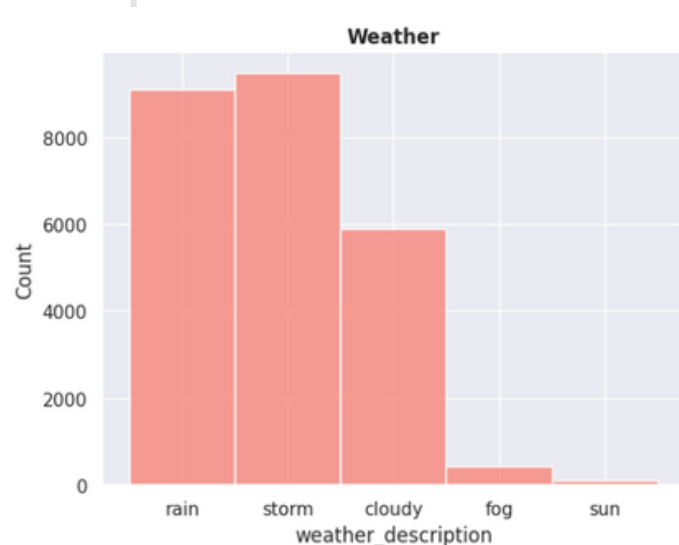
● Verificación de la calidad los datos

Variables	Description	Data Type
<u>cloudiness_id</u>	Identificador único que representa un estado específico de nubosidad	<u>int64</u>
<u>cloudiness</u>	Descripción del nivel de nubosidad	<u>object</u>
<u>date_id</u>	Identificador único para una fecha específica	<u>int64</u>
<u>date</u>	Fecha para identificar cuándo se realizó la observación	<u>object</u>
<u>weather_id</u>	Identificador único asociado a un tipo específico de condición meteorológica	<u>int64</u>
<u>weather</u>	Descripción de las condiciones meteorológicas observadas	<u>object</u>
<u>season_id</u>	Identificador único para la estación del año	<u>int64</u>
<u>season</u>	Estación del año	<u>object</u>
<u>observation_id</u>	Identificador único para cada registro individual de observación	<u>int64</u>
<u>precipitation</u>	Cantidad de precipitación registrada durante el periodo (mm)	<u>float64</u>
<u>temp_max</u>	Temperatura máxima registrada durante el periodo, medida en grados Celsius (°C)	<u>float64</u>
<u>temp_min</u>	Temperatura mínima registrada durante el periodo, medida en grados Celsius (°C)	<u>float64</u>
<u>wind</u>	Velocidad del viento registrada (km/h)	<u>float64</u>
<u>humidity</u>	Nivel de humedad relativa en el aire	<u>float64</u>
<u>pressure</u>	Presión atmosférica registrada (<u>hPa</u>)	<u>float64</u>
<u>solar_radiation</u>	Cantidad de radiación solar recibida en la superficie (W/m²)	<u>float64</u>
<u>visibility</u>	Distancia máxima a la que se puede ver un objeto claramente (km)	<u>float64</u>

ENTENDER DATOS

● Exploración de los datos

POWERBI

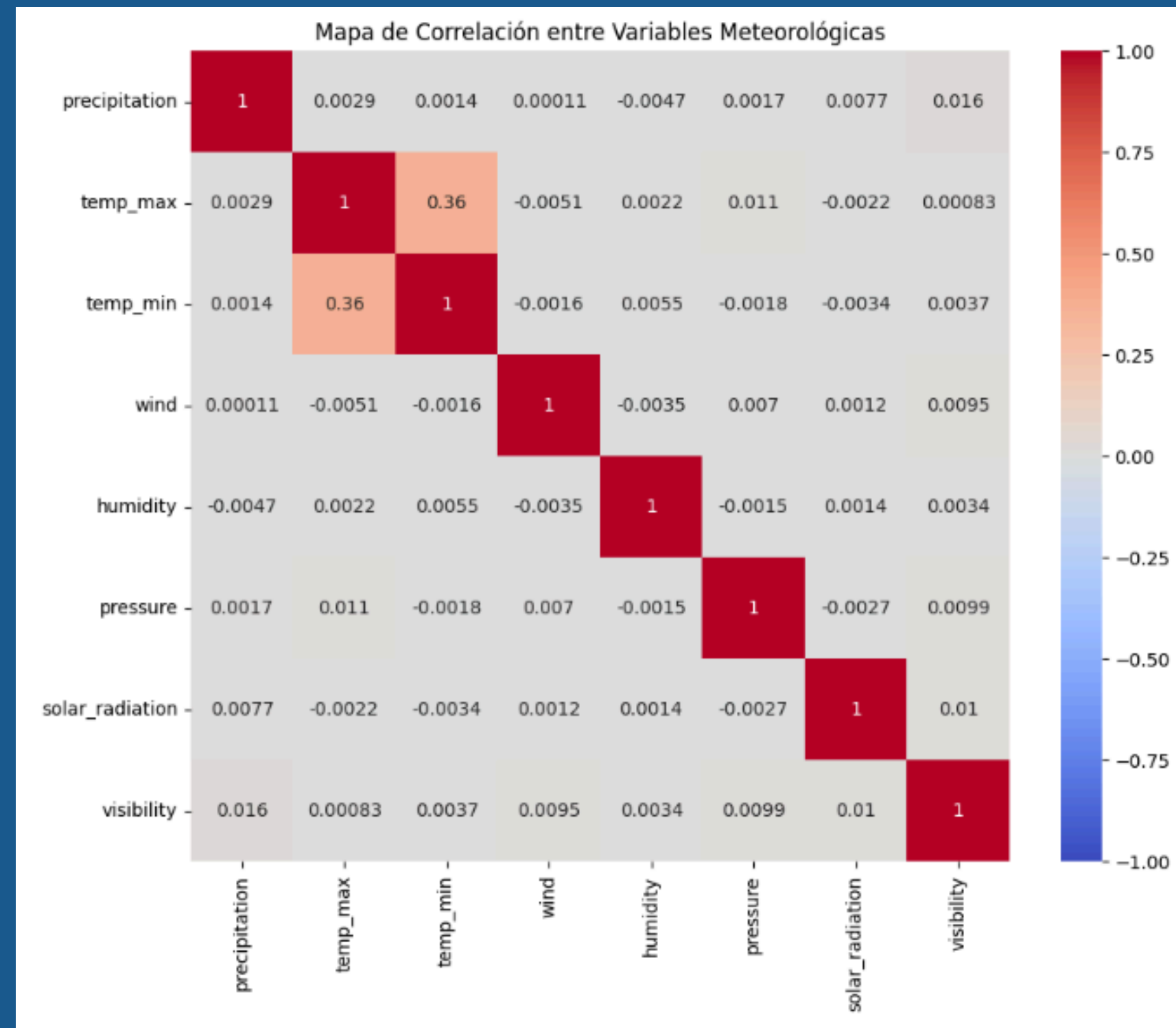


PREPARAR DATOS

SELECCIÓN

Hemos seleccionado los siguientes datos:

- Viento
- Humedad
- Temp. Min.
- Temp. Max.



PREPARAR DATOS

LIMPIEZA

Antes de crear el modelo es necesario limpiar los datos para garantizar la calidad de los mismos.

Para la limpieza:

- **Eliminar valores innecesarios y duplicados**
- **Convertir tipos de datos**
- **Imputar datos faltantes**
- **Balancear la data**

PREPARAR DATOS

CONSTRUCCIÓN

Calculamos la media e imputamos los datos substituyendo los nulos:

```
observation_date    0
precipitation       0
temp_max            0
temp_min            0
wind                0
humidity            0
pressure            0
solar_radiation     0
visibility           0
weather_description 0
cloudiness_description 0
season_description   0
```

PREPARAR DATOS

Normalizamos y transformamos los datos para evitar que estén desbalanceados.

Aplicamos en:

- **Precipitation: Transformación logarítmica**
- **Wind: Transformación de raíz cuadrada**
- **Temp_max, Temp_min, Precipitation y Wind: Escalar min-max**

MODELACIÓN

Modelo utilizado

Modelo SVM (Support Vector Machines)

¿Porqué?

- Manejo de datos no lineales
- Robustez
- Manejo de datos desbalanceados
- Flexibilidad en las características del núcleo

Características del modelo

- Class_weight='balanced'
- C = 10
- Gamma = scale
- Kernel = poly

Reporte de Clasificación:

	precision	recall	f1-score	support
1	0.97	0.89	0.93	2704
2	0.97	0.85	0.91	2852
3	0.74	0.88	0.80	1773
4	0.38	0.93	0.54	142
5	0.39	0.90	0.55	29
accuracy			0.87	7500
macro avg	0.69	0.89	0.75	7500
weighted avg	0.90	0.87	0.88	7500

Accuracy: 0.8741333333333333

```
# Crear el modelo SVM 01 MAXMINSCALER 3
svm_model1M3 = SVC(class_weight='balanced', kernel='poly', C=10, gamma='scale', random_state=42)

# Entrenar el modelo
svm_model1M3.fit(X_train_std, y_train)
```



EVALUACIÓN

Para la evaluación se ha analizado el Reporte de Clasificación y la Accuracy.

El modelo ha mostrado un rendimiento sólido y con una exactitud global del 87.41%.

También muestra ciertos puntos dónde el modelo falla, como las clases 4 y 5, que refleja la necesidad de un mejor manejo del desequilibrio de clases.

PERMETEO SL demo



Contact us to get more info



PermeteoSL@gmail.com



www.PermeteoSL.com