

Manual de programador para el prototipo de algoritmo para la predicción de precipitaciones a partir de datos estadísticos recolectados de la estación meteorológica de la Universidad De Cundinamarca Seccional Ubaté (MCPP)



# Tabla de contenido

Introducción
Acerca de este Manual
Objetivos del Software
Especificaciones del software
Desarrollo y Estructura del Software
Arquitectura del Software
Componentes del sistema
Casos de uso del funcionamiento del sistema
Tecnologías utilizadas
Configuración del entorno de desarrollo
Requisitos del entorno de desarrollo
Versiones de dependencias utilizadas
Instalación de herramientas y dependencias
Configuración de la base de datos
Estructura del código fuente
Organización de carpetas
Comentarios y documentación dentro del código
Codificación de los algoritmos
Descripción de los algoritmos utilizados
Desarrollo de la interfaz de usuario
Creación de rutas para los módulos
Diseño y estructura de la interfaz



Preguntas Frecuentes (FAQ)	20
Glosario	21
Giosario	21
Referencias	22



#### Introducción

#### Acerca de este Manual

El presente manual contiene los aspectos técnicos del aplicativo MCPP y aspectos a cumplir necesarios para su uso, además de información para los usuarios interesados en conocer la predicción de precipitaciones mediante un modelo desarrollado para proporcionar información relevante sobre las posibles futuras lluvias, basándose en la toma de datos de la estación meteorológica de la Universidad de Cundinamarca Seccional Ubaté, dicho modelo esta implementado en un aplicativo diseñado para ser intuitivo, agradable y útil para los usuarios que lo utilicen.

Este manual proporciona una guía importante para los usuarios que deseen conocer el funcionamiento del aplicativo MCPP (Modelo Comparativo Predictivo de Precipitaciones).

#### **Objetivos del Software**

El objetivo principal del presente software es proporcionar información a los usuarios acerca de los porcentajes de efectividad y predicciones de los modelos Random Forest, Naive Bayes, Support Vector Machine y MCPP, los cuales manejan técnicas de predicción diferentes, por lo cual permiten a los usuarios tomar decisiones basados en sus necesidades. Dicho software también tiene como objetivo proporcionar una herramienta para probar los modelos con datos que el usuario posea, para así darle a conocer porcentajes de efectividad y predicciones futuras con base en los datos proporcionados.



### Especificaciones del software

Es necesario cumplir con los siguientes requisitos previos a la utilización de la página para poder hacer uso óptimo de todas sus funciones:

- Contar con el aplicativo instalado: Es necesario contar con el aplicativo instalado para el funcionamiento del mismo.
- Navegadores web: El aplicativo puede ser visitado desde cualquier navegador moderno, entre ellos Opera, Microsoft Edge, Brave, Google Chrome, etc.
- Sistema operativo: Para poder acceder al aplicativo solo es necesario contar con un navegador, por lo cual no se requiere un sistema operativo en especial.
- Cuenta de usuario: Para acceder a las funcionalidades principales del aplicativo no es
  necesario contar con una cuenta de usuario, a excepción de contar con privilegios de
  administrador para acceder a la base de datos del modelo, en dicho caso se asignará
  una cuenta preestablecida a quienes tengan el derecho de acceso.
- Tamaño de pantalla: El sistema cuenta con un diseño adaptable para cualquier tamaño de pantalla, por lo cual puede ser utilizado tanto en dispositivos móviles como computadores.



#### Desarrollo y Estructura del Software

### **Arquitectura del Software**

La arquitectura del aplicativo de predicción de precipitaciones está diseñada con el fin de ofrecer una experiencia agradable y precisa para los usuarios, esta se basa en tres módulos principales: el modulo principal, el módulo de probar modelo y el módulo de administrador. La información detallada sobre el diseño de estos módulos puede ser encontrado en el Manual de Usuario.

#### Componentes del sistema

Módulo Principal. Este módulo sirve como punto de entrada para los usuarios al aplicativo. Dentro del cual se pueden visualizar las predicciones dadas por el modelo MCPP, Random Forest, Naive Bayes y Support Vector Machines. Además, contiene información sobre los datos alojados en la base de datos, con los cuales es posible realizar las predicciones, widgets de información general sobre el clima y un apartado inferior de información de interés acerca del proyecto y el aplicativo.

Módulo de probar modelo. Este módulo proporciona una herramienta sencilla e intuitiva para el usuario tenga la posibilidad de probar la efectividad del modelo MCPP y también de los modelos Random Forest, Naive Bayes y SVM, una forma útil de comprobar el funcionamiento del aplicativo principal, brindando la posibilidad de obtener resultados con datos que el usuario tenga y así mismo asignar un porcentaje de prueba para el debido entrenamiento de los datos ingresados.

**Módulo de administrador.** Este módulo está enfocado exclusivamente a los usuarios con el rol especial de administrador, a los cuales se les dará una cuenta única previamente para



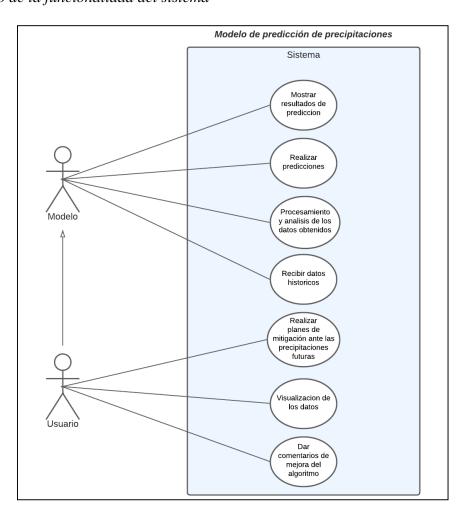
que tengan la posibilidad de visualizar, actualizar, ingresar o eliminar datos dentro de la base de datos del aplicativo (MongoDB).

### Casos de uso del funcionamiento del sistema

Para mayor entendimiento y claridad del funcionamiento del sistema se generó un diagrama de casos de uso, con el fin de identificar las funciones tanto de los usuarios como del modelo MCPP.

Figura 1

Casos de uso de la funcionalidad del sistema



*Nota*. Diagrama de casos de uso del sistema, mostrando los roles tanto del modelo como del usuario final.



### Tecnologías utilizadas

El desarrollo de los modelos y del aplicativo estuvo representado por la utilización de diversas tecnologías que garantizan un rendimiento eficiente y una experiencia de usuario idónea. Entre dichas tecnologías se encuentran:

**Lenguajes de programación.** Python para la funcionalidad de los modelos y el backend general del aplicativo.

**Frameworks.** Flask para el backend y conexión con el diseño del aplicativo, Bootstrap para el diseño de la interfaz de usuario.

Bases de Datos. Mongo para el almacenamiento de los datos meteorológicos del entrenamiento de los modelos y de la estación meteorológica de la Universidad de Cundinamarca Seccional Ubaté.

**Bibliotecas y librerías.** Scikit-learn para el modelado y entrenamiento de los datos, Pandas y Numpy para la manipulación y análisis de los datos recolectados, matplotlib para graficas e interpretación de los datos.

**Lenguajes de diseño.** HTML para la estructuración del aplicativo MCPP y CSS para el diseño y estilo del aplicativo.

**Entorno de desarrollo.** Jupyter Notebook para el análisis y desarrollo de los algoritmos y Visual Studio Code para la codificación de los modelos finales.



### Configuración del entorno de desarrollo

#### Requisitos del entorno de desarrollo

Se recomienda el uso del entorno de desarrollo Visual Studio Code, pero cualquier IDE compatible con el lenguaje Python y permita la instalación de las librerías mencionadas es aceptable.

### Versiones de dependencias utilizadas

A continuación, se detallas las versiones de las dependencias utilizadas en el desarrollo del aplicativo. Bibliotecas de Python:

- certifi 2023.7.22
- Flask 2.3.3
- Jinja2 3.1.2
- matplotlib 3.8.0
- numpy 1.26.0
- pandas 2.1.1
- pip 23.3.1
- pymongo 4.5.0
- scikit-learn 1.3.0
- scipy 1.11.2
- seaborn 0.12.2

### Instalación de herramientas y dependencias

Para la instalación de las librerías necesaria fue necesaria la creación de la carpeta del proyecto y abrirla dentro del IDE VSCode, de esta manera se abrió la terminal dentro del proyecto y se instalaron las librerías, empezando por pip, luego mediante el comando "pip



install" seguido del nombre de la librería requerida fue posible instalar las herramientas necesarias para el desarrollo, para esto fue necesaria la creación de un entorno virtual; el cual se creó con el comando "py -3 -m venv env".

Figura 2

Creación del entorno virtual en VSCode

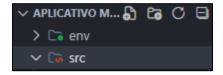
```
\U-2023-2\Tesis de Grado\Aplicativo MCPP> py -3 -m venv env
```

Nota. La imagen muestra el comando para la creación del entorno virtual en el proyecto.

De esta manera se creó una carpeta denominada env, luego se creó una carpeta llamada src, la cual contiene todo el desarrollo del aplicativo y los modelos predictivos.

Figura 3

Creación de la carpeta src en VSCode



Nota. La imagen muestra la creación de la carpeta src junto al env en el proyecto.

Para el inicio del desarrollo dentro del entorno virtual se ingresó al mismo mediante la terminal con los comandos "Set-ExecutionPolicy Unrestricted -Scope Process" y "env\Scripts\activate", de esta manera se instalaron las librerías necesarioas, en este caso fueron las encontradas en el apartado de Versiones de dependencias utilizadas, esto mediante el comando "pip install + dependencia"



Figura 4

Instalación de la librería de Flask

\U-2023-2\Tesis de Grado\Aplicativo MCPP> pip install Flask

*Nota*. La imagen muestra la instalación de la librería de Flask como ejemplo de cómo se desarrolló la instalación de todas las dependencias.

### Configuración de la base de datos

Para la base de datos se utilizó el motor de Mongo Atlas, el cual permite manejar gran cantidad de datos de forma flexible y sencilla. Se creó una colección denominada "Datos" que contiene los datos utilizados para las predicciones y otra colección llamada "Usuarios" que contiene la cuenta de usuario para el modulo administrador.

Figura 5

Base de datos en Mongo Atlas



*Nota*. La imagen muestra las colecciones Datos y Usuarios dentro de la base de datos del aplicativo.

Para la conexión dentro del aplicativo se utilizó la librería pymongo y certifi para permitir la conexión del código fuente con la base de datos.



Figura 6

Conexión del aplicativo con Mongo

```
from pymongo import MongoClient
import certifi
from pymongo.collection import ReturnDocument
from pymongo import MongoClient

# Conexión con MongoDB

MONGO =

# Utilización del certificado
certificado = certifi.where()

# Función para la conexión con la DB
def Conexion():
    try:
        client = MongoClient(MONGO, tlsCAFile = certificado)
        bd = client["bd_Precipitaciones"]
    except ConnectionError:
        print("Error de Conexión")
    return bd
```

Nota. La imagen muestra la conexión del aplicativo con Mongo.

### Estructura del código fuente

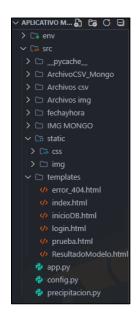
### Organización de carpetas

Para la estructuración del código fuente se organizaron los archivos tipo Python dentro de la carpeta src, los archivos de tipo HTML dentro de una carpeta llamada "templates", además se creó una carpeta llamada "static" que contiene los archivos tipo CSS dentro de una carpeta "css" y las imágenes del aplicativo dentro de "img", sumado a esto se crearon carpetas específicas para alojar los gráficos, archivos css e imágenes generales del aplicativo.



Figura 7

Organización de carpetas para el aplicativo MCPP



Nota. La imagen muestra la organización de las carpetas para el aplicativo MCPP

### Comentarios y documentación dentro del código

Para un correcto entendimiento del código se utilizaron comentarios dentro del mismo para entender todos los aspectos posibles en el desarrollo del aplicativo.

### Figura 8

Ejemplo de comentarios dentro del código

```
# Se carga el conjunto de datos spam.data

## sep== separacion por comas

## Header== no tiene

df=pd.read_csv("./src/ArchivoCSV_Mongo/filename.csv", header=None, skiprows=1)

# separa los datos en "KRF" y en "YRF"

# "XRF" representa todas las columnas menos la ultima la cual es la de respuesta

# "YRF" representa la columna de respuesta

XRF = df.iloc[:,-1]

YRF = df.iloc[:,-1]

# Normalizamos los datos con logaritmos neperianos

XRF_normalizada=np.log1p(df.iloc[:,:-1])

# configuracion dl modelo con datos normalizados mediante a logaritmos neperianos

XRF_train, XRF_test, YRF_train, YRF_test = train_test_split(XRF_normalizada,YRF, test_size=0.2, random_state=1)
```

*Nota*. La imagen muestra un ejemplo de los comentarios dejados a lo largo del código fuente para mayor entendimiento del lector.



De igual manera el proyecto está debidamente documentado para comprender razones, objetivos, importancia, desarrollo y conclusiones del presente desarrollo.

### Codificación de los algoritmos

### Descripción de los algoritmos utilizados

Se utilizaron 3 algoritmos principales, los cuales son Random Forest, Naive Bayes y Support Vector Machines, lo cuales fueron debidamente entrenados y adaptados al código fuente del aplicativo.

**Figura 9**Algoritmo Random Forest implementado en el aplicativo

Nota. La imagen muestra la codificación del algoritmo de Random Forest



### Figura 10

Algoritmo Support Vector Machines implementado en el aplicativo

Nota. La imagen muestra la codificación del algoritmo de Support Vector Machines

### Figura 11

Algoritmo Naive Bayes implementado en el aplicativo

```
# Modelo Naive Bayes

# Separa los datos en "NN" y en "Y"

# "200" representa todos las columnas menos la ultima la cual es la de respuesta

# "Y" representa la columna de respuesta

df_data = pd.read_csv('./src/ArchivoCSV_Mongo/filename.csv', header=None, skiprows=1)

# Haremos Feature Selection para mejorar los resultados del algoritmo. Utilizando la clase SelectKBest para sels

XNW = df_data.iloc[:,:-1]

y = df_data.iloc[:,:-1]

best = SelectKBest(k = 4)

XNV_new = best.fit_transform(XNV,y)

XNV_new.shape

selected = best.get_support(indices = True)

used_features = XNV.columns[selected]

# División de los datos en conjuntos de prueba y entrenamiento, 80% para entrenamiento y 20% para prueba

# X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.20, random_state = 6)

XNV_train, XNV_test, y_train, y_test = train_test_split(XNV,y, test_size = 0.20, random_state = 6)

# Normalizacion de la data

bayes_naive = GaussianNB()

bayes_clasif = bayes_naive.fit(XNV_train[used_features].values, y_train)

y_pred = bayes_naive.predict(X)V_test[used_features])

# Junice de ocurrencia

ocurrenciaNV-accuracy_score(y_test, y_pred)

fl_N8-fl_score(y_test, y_pred)

print(fl_N8)

print(fl_N8)

print(fl_N8)

print(fl_N8)

RTANV = pd.DataFrame({'Real': y_test, 'Predicho': y_pred})
```

Nota. La imagen muestra la codificación del algoritmo de Naive Bayes.



Seguido de esto de implemento la creación del algoritmo basado en la comparación de los tres anteriores, el cual más adelante fue denominado modelo MCPP.

Figura 12
Fragmento de código de la creación del modelo MCPP

Nota. La imagen muestra un fragmento del código utilizado para la creación del modelo MCPP

De esta manera fue posible obtener los resultados de predicciones lanzados por los tres modelos desarrollados en un proceso de algoritmia.

### Desarrollo de la interfaz de usuario

Para el desarrollo de la interfaz del aplicativo se hizo uso de herramientas como Flask, Bootstrap, HTML, CSS y el mismo Python para la creación de cada ruta que contiene los tres módulos mencionados anteriormente.



### Creación de rutas para los módulos

La creación de dichas rutas se realiza mediante el código "@app.route('/')", como se puede visualizar en la figura 13.

Figura 13

Creación de la ruta prueba para el módulo de probar modelo

```
# Ruta para la pantalla de probar modelo
@app.route('/prueba')
def prueba():
    return render_template('prueba.html')
# Funcion para recibir el archivo y enviarlo a una carpeta llamada
# Archivos csv creada en la carpeta src
# Aqui se hara todo el proceso para leer el archivo seleccionado y aplicar los modelos
@app.route('/RTA', methods=['POST'])
def procesoAlgoritmo():
    if request.method == 'POST':
    # Obtenemos el archivo del input "archivo"
    f = request.files['archivo']
    filename = secure_filename(f.filename)
# Guardamos el archivo en el directorio "Archivos PDF"
    f.save(os.path.join(app.config['UPLOAD_FOLDER'], filename))
# test_size- valor de entrenamiento
    traine=int(request.form['test_size'])
    trainee=traine*0.01
```

*Nota*. La imagen muestra la creación de la ruta para probar modelo.

De esta manera fue posible crear las rutas necesarias y realizar la configuración que cada una de ellas requería para su correcto funcionamiento.

#### Diseño y estructura de la interfaz

Para el diseño de la interfaz se utilizó esencialmente HTML y CSS, pero también fueron implementadas librerías de Bootstrap y tablas para mejorar el aspecto del aplicativo.



### Figura 14

Fragmento de código en HTML del módulo principal

```
<
```

Nota. La imagen muestra un fragmento de código del módulo principal del aplicativo MCPP

Para la mejora de los estilos se utilizó CSS.

Figura 15 Fragmento de código en CSS de los estilos del módulo principal

*Nota*. La imagen muestra un fragmento de código de los estilos del módulo principal del aplicativo MCPP.



Así, fue posible el desarrollo del aplicativo, combinando la unión del Backend que contiene el funcionamiento de los algoritmos para el resultado en forma de modelos con el Frontend, el cual está basado en brindar una experiencia de usuario idónea a la hora de consumir los servicios que ofrece el aplicativo.



### **Preguntas Frecuentes (FAQ)**

### 1) La aplicación no se inicia correctamente ¿Cuáles podrían ser las causas?

### Respuesta:

- Verifica que todas las dependencias estén instaladas.
- Asegúrate de que la configuración de la base de datos sea la correcta.
- Visualiza los mensajes de error lanzados por la terminal o el aplicativo para obtener soluciones adicionales.

### 2. El entorno virtual del aplicativo me da error ¿Cómo puedo solucionarlo?

Respuesta: Asegúrate de instalar el entorno virtual del proyecto en tu máquina, la carpeta de src puede copiarla y pegarla dentro de tu proyecto, de esta manera el aplicativo funcionara correctamente.

## 3. ¿Cómo puedo contribuir al desarrollo del proyecto?

Respuesta: Para contribuir con el proyecto puedes comunicarte con los desarrolladores Luis Rincón o Jeison Polanco, de esta manera podrás crear tu rama en el repositorio de GitHub y realizar contribuciones propias.

.



#### Glosario

MCPP: Modelo Comparativo Predictivo de Precipitaciones.

**Porcentaje de prueba:** Cantidad de datos que se desea dar a los modelos para el entrenamiento de los datos.

Cadena de datos: Lista de elementos dados a el aplicativo para la constante actualización de las predicciones.

RF: Random Forest.

**SVM:** Support Vector Machine.

NB: Naive Bayes.

Po: Presión Atmosférica.

**T:** Temperatura.

**U:** Humedad Relativa.

**Ff:** Velocidad del viento.

**RRR:** Precipitación.

VSCode: Visual Studio Code



### Referencias

Estación meteorológica de la Universidad de Cundinamarca Seccional Ubaté. (s. f.). [Conjunto de datos].

rp5.ru. (2023, 15 julio). rp5.ru. https://rp5.ru/

Python. (2023). Python.org. https://www.python.org/

Mark Otto Jacob Thornton, And Bootstrap. (s. f.). Bootstrap. getbootstrap.com.

https://getbootstrap.com/

Flask Documentation. (s. f.). *flask.palletsprojects.com*. https://flask.palletsprojects.com/en/3.0.x/

**VERSION** 1.0 | Fecha: 05/11/2023