Aplicación web progresiva para gestionar pronósticos agrícolas en las plantaciones de cacao en Orellana

Progressive web application to manage agricultural forecasts in cocoa plantations in Orellana

AUTORES: Calderón Vivanco, Jessenia Marlene[1]
Analuisa Iza, Ángel Daniel[2]
Chango Sailema, Wilson Gustavo[3]
Aguilar Encarnación, Pedro Stalyn[4]

Dirección para correspondencia: jessenia.calderon@espoch.edu.ec

Fecha de recepción: 2023/07/29
Fecha de aceptación: 2028/05/28
DOI: https://doi.org/10.61154/mrcm.v10i2.3211

RESUMEN

La industria del cacao se enfrenta a desafíos constantes para mejorar la producción y sostenibilidad el cultivo de cacao depende de condiciones ambientales específicas y requiere un manejo cuidadoso para obtener una alta calidad y rendimiento con el fin de superar estos desafíos, se ha adoptado la tecnología de sensores y monitoreo automatizado en las plantaciones de cacao. El problema radica en la falta de información en tiempo real sobre las condiciones del cultivo y la dificultad para realizar ajustes precisos en las prácticas agrícolas esto puede llevar a una producción subóptima,

[1] Estudiante, Facultad de Informática y Electrónica, Escuela Superior Politécnica de Chimborazo, Sede Orellana. Orellana, Ecuador. Email: jessenia.calderon@espoch.edu.ec) CÓDIGO ORCID https://orcid.org/0000-0002-6231-3796
[2]Estudiante, Facultad de Informática y Electrónica, Escuela Superior Politécnica de Chimborazo, Sede Orellana, Orellana. Ecuador. Email: angel.analuisa@espoch.edu.ec)CÓDIGO ORCID https://orcid.org/0000-0003-2029-9858
[3] Ph.D en Computación Avanzada, Facultad de Informática y Electrónica, Docente de la Carrera Tecnologías de la Información, Escuela Superior Politécnica de Chimborazo, Sede Orellana. Orellana, Ecuador. E-mail: wilson.chango@espoch.edu.ec)CÓDIGO ORCID https://orcid.org/0000-0003-3231-0153
[4] Magister en Software, Ingeniero en Sistemas Informáticos, Docente de la Carrera Tecnologías de la Información, Escuela Superior Politécnica de Chimborazo, Sede Orellana. Orellana, Ecuador. Email: pedro.aguilar@espoch.edu.ec) CÓDIGO ORCID https://orcid.org/0009-0005-1664-2280

menor calidad del cacao y menor rentabilidad para los agricultores el enfoque adoptado es la utilización estratégica de sensores ubicados en las plantaciones de cacao para recopilar datos en tiempo real sobre variables críticas como la temperatura y la humedad del suelo. Estos datos se envían a una Progressive Web App (PWA) dedicada, que proporciona información clara y comprensible a los agricultores sobre las condiciones del cultivo el uso de sensores y la PWA permite a los agricultores tomar decisiones informadas basadas en datos en tiempo real permitiendo realizar ajustes precisos en el riego, la fertilización y otras prácticas agrícolas para optimizar la producción y calidad del cacao la optimización del cultivo de cacao ofrece beneficios significativos para los agricultores al mejorar la eficiencia, productividad, calidad y sostenibilidad de su producción esta tecnología innovadora brinda una ventaja competitiva en el mercado como también impulsar el éxito a largo plazo en la industria del cacao.

PALABRAS CLAVE: Cacao; Sensores; Monitoreo automatizado; Producción; Sostenibilidad; previsión. (Tesauro de la UNESCO)

ABSTRACT

The cocoa industry faces constant challenges to improve production and sustainability. Cocoa farming depends on specific environmental conditions and requires careful management to obtain high quality and yield to overcome these challenges, sensor technology and automated monitoring have been adopted in cocoa plantations. The problem lies in the lack of real-time information on crop conditions and the difficulty to make accurate adjustments in farming practices this can lead to suboptimal production, lower cocoa quality and lower profitability for farmers the approach adopted is the strategic use of sensors located in cocoa plantations to collect real-time data on critical variables such as temperature and soil moisture. This data is sent to a dedicated Progressive Web App (PWA), which provides clear and understandable information to farmers on crop conditions the use of sensors and the PWA allows farmers to make informed decisions based on real-time data enabling accurate adjustments to irrigation, fertilization and other farming practices to optimize cocoa production and quality the optimization of the cocoa crop offers significant benefits to farmers by improving the efficiency, productivity, quality and sustainability of of their production This innovative technology provides a competitive advantage in the marketplace as well as drive long-term success in the cocoa industry.

KEYWORDS: Cocoa; Sensors; Automated monitoring; Cocoa; Production; Sustainability; Forecasting.

INTRODUCCIÓN.

En la industria del cacao la búsqueda de métodos innovadores para mejorar la producción y la sostenibilidad se ha convertido en una prioridad según (Hayford, 2023) el cultivo de cacao enfrenta diversos desafíos, como enfermedades, plagas, cambios climáticos y prácticas agrícolas ineficientes, que afectan negativamente la productividad y calidad de los granos de cacao por lo tanto es crucial buscar estrategias innovadoras que permitan optimizar el cultivo de cacao y garantizar su sostenibilidad a largo plazo.

Uno de los enfoques más prometedores es el monitoreo inteligente del cultivo de cacao que utiliza sensores y una PWA (Progressive Web App) para recopilar y analizar datos en tiempo real. Por otro lado, (Bianchini, 2020) en su investigación manifiestan que el campo de la agricultura de precisión la implementación de sensores en el cultivo de cacao brinda una valiosa fuente de datos sobre variables críticas como la temperatura, la humedad del suelo. Estos sensores, ubicados estratégicamente en la plantación monitorean continuamente estas variables y envían los datos recopilados a una PWA dedicada los agricultores pueden acceder a información en tiempo real sobre las condiciones del cultivo y realizar ajustes precisos en el riego, la fertilización y otras prácticas agrícolas. De igual manera **(**Tech, 2023) analisa la importancia de optimizar el cultivo de cacao va más allá de la simple mejora de la productividad. Un cultivo de cacao saludable y bien manejado no solo aumenta la rentabilidad para los agricultores, sino que también contribuye a la conservación del medio ambiente y a la mejora de las condiciones de vida. Además, el cacao de calidad superior, obtenido a través de prácticas de cultivo óptimas, tiene un mayor valor en el mercado internacional, lo que brinda oportunidades para mejorar la competitividad de los productores.

En este contexto la optimización del cultivo de cacao a través de sensores y monitoreo automatizado emerge como una solución prometedora los avances tecnológicos recientes en el campo de la agricultura de precisión y la sonorización han abierto nuevas oportunidades para mejorar la eficiencia y productividad de los cultivos el objetivo General de esta investigación es mejorar la producción y calidad del cacao mediante el uso de sensores y sistemas de monitoreo

automatizado además la PWA proporciona una interfaz intuitiva y amigable que permite a los agricultores visualizar los datos recopilados de manera clara y comprensible (Vanegas, 2023) presentan una propuesta a través de gráficos tendencias y reportes generados por la aplicación, los agricultores pueden identificar patrones, detectar problemas potenciales y recibir recomendaciones personalizadas para optimizar la producción de cacao de manera sostenible la combinación de sensores y una PWA en el monitoreo inteligente del cultivo de cacao ofrece una oportunidad única para mejorar la eficiencia, la productividad y la sostenibilidad en la industria del cacao. De igual manera (Lorente Sierra, 2023) nos menciona que, a través de esta tecnología innovadora, los agricultores pueden aprovechar al máximo los datos recopilados y tomar decisiones fundamentadas que impulsen su éxito en un mercado cada vez más competitivo.

Las preguntas de investigación que surgen en este estudio son:

Pregunta 1: ¿Cuál es el impacto de la optimización del cultivo de cacao a través de sensores y monitoreo automatizado en la rentabilidad de los agricultores?

Pregunta 2: ¿Cómo se pueden utilizar los resultados del monitoreo automatizado para tomar decisiones informadas y mejorar las prácticas de manejo del cultivo de cacao?

Estas preguntas de investigación guiarán el desarrollo del estudio y la búsqueda de respuestas que contribuyan al conocimiento y avance en el campo de la optimización del cultivo de cacao mediante el uso de sensores y monitoreo automatizado.

Optimizar el cultivo de cacao a través de sensores y monitoreo automatizado ha demostrado ser eficaz en incrementar la productividad y reducir el uso de recursos. Sin embargo, (Flores, 2015) destacó que este método también presenta desafíos relacionados con la implementación tecnológica y la necesidad de capacitación especializada para los agricultores. A pesar de estos retos, el uso de sensores puede ayudar a garantizar condiciones de crecimiento óptimas para maximizar el rendimiento, permite una gestión más eficiente del agua y los nutrientes, evita el estrés de las plantas debido a condiciones adversas y reduce el riesgo de plagas. Finalmente, (Rios & Levano, 2022) menciona que al proporcionar datos precisos y en tiempo real, los agricultores pueden tomar decisiones basadas en evidencia, mejorando así la rentabilidad de su cultivo.

METODOLOGÍA.

Este proyecto se llevó a cabo en la Provincia de Orellana, cantón Francisco de Orellana, con una muestra intencional de 10 familias cacaoteras. Utilizando la metodología de Ciencia del Diseño, se desarrolló e implementó una Progressive Web App (PWA) para gestionar pronósticos agrícolas en las plantaciones de cacao. Esta tecnología permite a los agricultores optimizar el uso de recursos y mejorar la rentabilidad de sus cultivos mediante decisiones basadas en datos precisos y en tiempo real.

Sensores para el estudio del Cultivo del Cacao

Cherlinka (2022) muestra que existen diferentes tipos de sensores que se pueden utilizar para el monitoreo y control del cultivo de cacao.

1. **Sensor de humedad del suelo:** Este tipo de sensor se utiliza para medir el contenido de agua en el suelo, lo que ayuda a determinar el momento adecuado para regar las plantas de cacao.
2. **Sensor de temperatura y humedad ambiental:** Estos sensores miden la temperatura y la humedad del aire en la zona de cultivo, lo que permite conocer las condiciones ambientales y ajustar las prácticas de cultivo en consecuencia.
3. **Sensor de luz:** Este tipo de sensor mide la cantidad de luz solar que llega a las plantas de cacao para que las plantas puedan orientarse y asegurarse de que reciben la cantidad correcta de luz.
4. **Sensor de nutrientes:** Estos sensores miden los niveles de nutrientes en el suelo, como nitrógeno, fósforo y potasio, lo que ayuda a determinar el momento adecuado para fertilizar las plantas de cacao.
5. **Sensor de humedad de las hojas:** Este tipo de sensor mide la cantidad de humedad en las hojas de las plantas de cacao, lo que ayuda a detectar la presencia de plagas y enfermedades.

Es importante tener en cuenta que la elección del tipo de sensor dependerá de las necesidades específicas del cultivo y de los objetivos de monitoreo y control del agricultor el diseño del estudio

se basa en la utilización de una PWA (Progressive Web App) para medir y monitorear diferentes variables relacionadas con el cultivo de cacao.

Sensor de Temperatura y Humedad Ambiental

Los sensores de temperatura y humedad ambiental suelen estar compuestos por dos sensores separados, uno para la temperatura y otro para la humedad (Marcos, 2020) nos menciona que el sensor de temperatura mide la temperatura ambiente en grados Celsius o Fahrenheit, mientras que el sensor de humedad mide la cantidad de vapor de agua presente en el aire y lo expresa como un porcentaje de la humedad relativa.

Para Ramírez (2020) corroborra que estos sensores pueden ser utilizados en la producción de cacao para medir la temperatura y la humedad ambiental en los invernaderos o las áreas de almacenamiento. Los sensores pueden ayudar a los productores de cacao a mantener un ambiente óptimo para el cultivo, asegurando que las condiciones sean adecuadas para el crecimiento y desarrollo del cacao. Además, los sensores pueden ser utilizados para monitorear las condiciones ambientales en las áreas de almacenamiento, asegurando que el cacao se mantenga en las condiciones adecuadas para su almacenamiento a largo plazo.

Las variables y medidas se pueden inferir algunas posibles variables a partir de los tipos de sensores mencionados, como la humedad del suelo, la temperatura y humedad ambiental, la cantidad de luz solar, los niveles de nutrientes en el suelo y la humedad de las hojas.

Pluviómetro S-RGF-M002

Es un dispositivo utilizado para medir la cantidad de precipitación que cae en una determinada área. Este pluviómetro es un modelo de tipo convencional, es decir, consta de un recipiente cilíndrico de paredes verticales con una sección transversal uniforme y un diámetro de boca de entrada estándar de 200 mm. Este modelo de pluviómetro es utilizado comúnmente en estaciones meteorológicas, campos agrícolas, estaciones hidrológicas y en otras aplicaciones donde se requiere la medición precisa de la precipitación. La medición de la precipitación es importante para prever sequías, inundaciones y otros eventos climáticos que pueden afectar la agricultura, la gestión del agua y la planificación urbana.

HOBO RX3000

Un Datalogger inalámbrico es un dispositivo utilizado para la monitorización y registro en tiempo real de datos ambientales esenciales, como temperatura, humedad, velocidad del viento, radiación solar y lluvia. Este equipo se conecta a sensores específicos que capturan los datos y los envía de manera inalámbrica a un ordenador o dispositivo móvil para su análisis y visualización su funcionalidad permite configurar alertas de alarma vía correo electrónico o mensajes de texto si se exceden ciertos límites en los parámetros medidos, lo que posibilita tomar medidas preventivas para evitar daños o pérdidas. Su versatilidad lo hace particularmente útil en áreas como la agricultura, gestión del agua, energía y estudios de impacto ambiental, donde la monitorización continua de las condiciones es fundamental.

MongoDB

Es una base de datos NoSQL (no relacional) que se utiliza para almacenar y recuperar grandes cantidades de datos de manera eficiente y escalable. A diferencia de las bases de datos relacional (Treviño-Villalobos, 2019) MongoDB utiliza un modelo de documentos que permite almacenar datos estructurados y no estructurados los datos se almacenan en colecciones, que son similares a las tablas en una base de datos relacional, y cada documento en una colección puede tener su propio conjunto de campos y valores. Esto permite una mayor flexibilidad en la organización de los datos y permite la incorporación de nuevos campos sin tener que modificar la estructura de la base de datos.

Base de Datos: Clave – Valor

Es un tipo de base de datos NoSQL (no relacional) que utiliza un modelo de datos simple en el que cada registro se almacena como un par de valores: una clave única y un valor asociado. Estos pares de clave-valor se almacenan en una tabla o espacio de nombres, y se pueden recuperar mediante la clave.

PWA (Progressive Web Apps)

Las PWA se basan en estándares web abiertos, como HTML, CSS y JavaScript, y están diseñadas para funcionar en cualquier dispositivo que tenga un navegador web moderno Aguirre (2020) las

PWA utilizan un enfoque progresivo para mejorar la funcionalidad de la aplicación a medida que el navegador y el dispositivo permiten más capacidades.

Las PWA se basan en tres principios fundamentales: progresividad, conectividad y accesibilidad. La progresividad significa que las PWA funcionan en cualquier navegador web, independientemente del dispositivo o plataforma en el que se ejecuten. La conectividad se refiere a la capacidad de trabajar sin conexión, lo que significa que las PWA pueden seguir funcionando incluso cuando no hay conexión a Internet Sánchez Jaque (2022) La accesibilidad se refiere a la facilidad de acceso a la aplicación, lo que significa que las PWA se pueden instalar en el dispositivo del usuario y se pueden acceder desde la pantalla de inicio como cualquier otra aplicación nativa.

El procedimiento del estudio puede implicar la instalación de sensores en la plantación de cacao, la configuración de una PWA para recopilar y analizar los datos en tiempo real, y el monitoreo continuo de las variables del cultivo. Los agricultores podrán acceder a la PWA para obtener información sobre las condiciones del cultivo y realizar ajustes precisos en el riego, la fertilización y otras prácticas agrícolas.

Docker

Es una plataforma de software poderoso y versátil que simplifica el proceso de implementación y administración de aplicaciones en diferentes entornos y plataformas, y es una herramienta esencial para los desarrolladores y operadores de sistemas que buscan una forma eficiente y escalable de crear y gestionar aplicaciones modernas.

Framework React

Es una biblioteca de JavaScript popular y ampliamente utilizada que se utiliza para construir interfaces de usuario reactivas y escalables. React utiliza componentes reutilizables y una técnica de "virtual DOM" para actualizar la interfaz de usuario de manera eficiente y se utiliza en una amplia variedad de aplicaciones web y móviles (Kopachovets, 2023).

Python

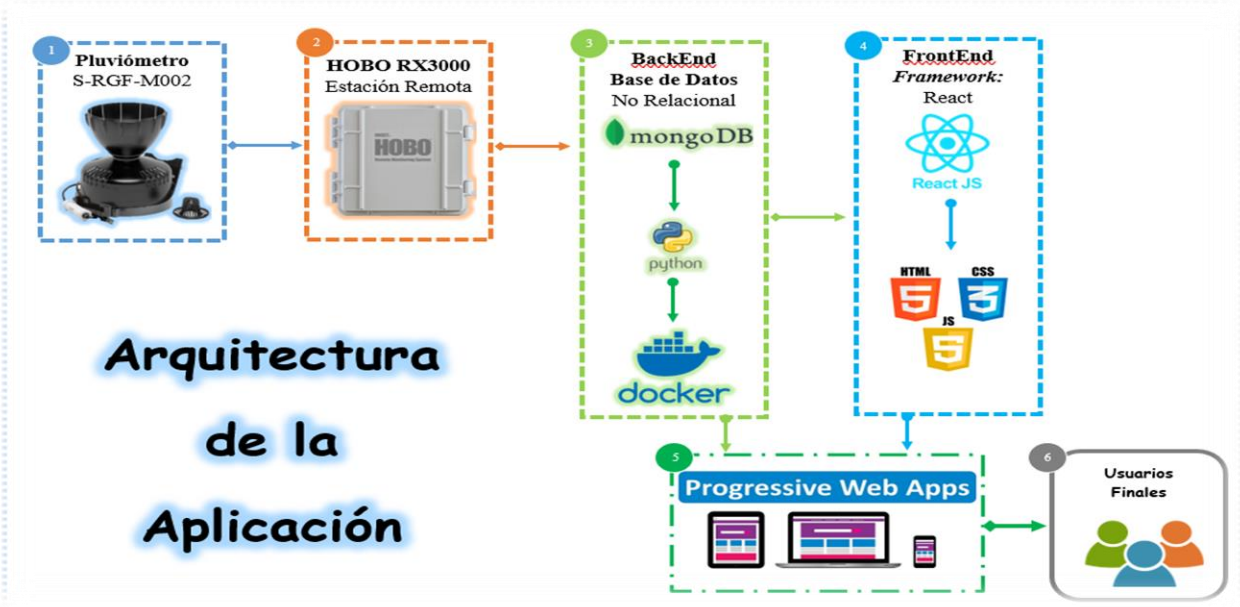
Para Zapata (2023) es un lenguaje de programación de alto nivel, interpretado y orientado a objetos, que se utiliza en una amplia variedad de aplicaciones es conocido por su sintaxis clara y legible, su

gran cantidad de bibliotecas y su compatibilidad multiplataforma, y es utilizado por una gran comunidad de desarrolladores en todo el mundo.

Diseño conceptual.

Figura 1

Arquitectura Utilizada para la aplicación



Nota: Este esquema se utilizo para el diseño de la arquitectura de la aplicación

Para la PWA contaremos con un logo de Aplicación que es este llamado WEATHER SENSOR es inspirado por la función principal del dispositivo, que es la detección y medición de variables relacionadas con el clima y las condiciones atmosféricas. El término "Weather" hace referencia al clima, es decir, el estado actual de la atmósfera en un lugar específico, incluyendo factores como la temperatura, humedad, viento, precipitación, entre otros. Por otro lado, el término "Sensor" se refiere a un dispositivo tecnológico diseñado para detectar y medir diferentes tipos de magnitudes o variables en su entorno. En este caso, el "Weather Sensor" se encarga de recolectar datos sobre el clima y las condiciones atmosféricas.

El nombre "Weather Sensor" sugiere la capacidad de este dispositivo para detectar y registrar información sobre variables climáticas como temperatura, humedad, presión barométrica, velocidad y dirección del viento, precipitación, entre otros. El sensor proporciona datos precisos

y en tiempo real sobre el clima, lo que permite realizar análisis, pronósticos y seguimiento de las condiciones meteorológicas.

Figura 2

Logo de la Aplicación



Nota: Este es el logo de identificacion de la Aplicación

El análisis de datos puede incluir el uso de algoritmos como PCA (análisis de componentes principales), kpca (kernel pca), robusta, bagging y boosting. estos algoritmos se pueden aplicar para extraer patrones, identificar relaciones entre variables y proporcionar recomendaciones para optimizar el cultivo de cacao.

Las consideraciones éticas pueden incluir obtener el consentimiento informado de los agricultores participantes, garantizar la privacidad y seguridad de los datos recopilados, y cumplir con las regulaciones y estándares éticos pertinentes en la investigación agrícola limitaciones del estudio pueden incluir factores como la variabilidad natural en las condiciones ambientales, la necesidad de calibrar y mantener adecuadamente los sensores, y las limitaciones técnicas o logísticas que puedan surgir durante la implementación de la pwa y el monitoreo automatizado. la reproducibilidad del estudio puede depender de la disponibilidad de los sensores utilizados, la documentación detallada del procedimiento de instalación y configuración, así como la transparencia en los métodos de análisis de datos. además, el uso de estándares y protocolos abiertos puede facilitar la reproducibilidad de los resultados.

RESULTADOS

Con el fin de analizar y comparar el desempeño de distintos algoritmos en la predicción se utilizó un conjunto de datos que comprendía información tanto del sensor como de la planta. este conjunto de datos constaba de 10 variables características, entre las cuales se encontraban rain (lluvia), temperature (temperatura), rh (humedad relativa), dewpoint (punto de rocío), windspeed (velocidad

del viento), gustspeed (velocidad de las ráfagas de viento), winddirection (dirección del viento), planta, fruto e incidencia. además, se incluyó una variable objetivo para la predicción, la cual correspondía a la severidad (%) de la enfermedad.

Tabla 1.

Rendimiento de los algoritmos PCA e IPCA

TIPO DE DATOS	ALGORITMOS	VALOR OBTENIDO
Datos originales	PCA	1.0
	IPCA	0.95
Datos Normalizados	PCA	0.9
	IPCA	0.98
Datos discretizados	PCA	0.95
	IPCA	0.9

Con base en los resultados obtenidos:

Datos originales: El PCA logro una precisión de 1.0, mientras IPCA obtuvo una precisión de 0,95 lo que significa que el modelo de regresión logística pudo ajustarse perfectamente a los datos originales.

Datos normalizados: PCA obtuvo una precisión de 0.85, mientras que IPCA alcanzó 0.83. La normalización mejoró la precisión en comparación con los datos originales, pero no alcanzó la misma perfección.

Datos discretizados: La discretización mejoró significativamente el rendimiento de ambos métodos de PCA: PCA obtuvo una precisión de 0.95, mientras que IPCA alcanzó 0.9. Esto sugiere que la discretización de las características ha sido beneficiosa para el modelo de regresión logística. En resumen, la discretización de los datos tuvo un efecto positivo en la precisión del modelo de regresión logística en comparación con los datos originales y normalizados. Por lo tanto, la discretización podría ser una estrategia útil para mejorar el rendimiento de los modelos de aprendizaje automático en este conjunto de datos específico.

Tabla 2.

Rendimiento de algoritmos KPC

ALGORITMO KPCA	KERNEL	VALOR OBTENIDO
Datos originales	Linear	1.0
	Polynomial	1.0
	RBF	0.68
Datos normalizados	Linear	0.85
	Polynomial	0.81
	RBF	0.86
	Linear	0.935

Datos discretizados	Polynomial	0.938
	RBF	0.91

Con base en los resultados obtenidos, se concluye lo siguiente:

Datos originales: El algoritmo Kernel PCA obtuvo una precisión de 1.0 con el kernel lineal, lo que indica que el modelo de regresión logística se ajustó perfectamente a los datos originales utilizando el primer componente principal. Sin embargo, cuando se utilizó el kernel polinomial, la precisión fue de 0.99, lo que también indica un buen ajuste, pero no perfecto. Por otro lado, con el kernel RBF, la precisión fue de 0.68, lo que sugiere que el modelo no se ajustó adecuadamente a los datos utilizando este kernel.

Datos normalizados: La normalización de los datos mejoró la precisión en comparación con los datos originales en la mayoría de los casos. Sin embargo, las precisiones aún no alcanzaron la perfección. Con el kernel lineal, la precisión fue de 0.85, mientras que con el kernel polinomial fue de 0.81 y con el kernel RBF fue de 0.86. Esto indica que, aunque la normalización ayudó, aún existe cierto margen de mejora para ajustar mejor el modelo a los datos normalizados.

Datos discretizados: La discretización de los datos también mejoró el rendimiento en comparación con los datos originales. Con el kernel lineal, la precisión fue de 0.935, con el kernel polinomial fue de 0.938 y con el kernel RBF fue de 0.91. Esto sugiere que la combinación de discretización y Kernel PCA puede ser una estrategia efectiva para mejorar el rendimiento del modelo de regresión logística en este conjunto de datos.

En conclusión, la normalización y la discretización pueden ser técnicas útiles para mejorar el rendimiento de los algoritmos de aprendizaje automático, especialmente cuando se utilizan junto con el algoritmo Kernel PCA. Sin embargo, cada conjunto de datos es único, por lo que es importante probar diferentes técnicas y ajustar los hiperparámetros para encontrar la combinación óptima que se adapte mejor a los datos específicos en cuestión.

Comparando el rendimiento de varios algoritmos que nos dan un mejor resultado, incluyendo PCA, IPCA, KPC con tres kernels. Los resultados de nuestras pruebas se presentan en la tabla 3.

Tabla 3.

Cuadro comparativo Rendimiento de algoritmos PCA, IPCA y KPCA

Método	Kernel	Valor obtenido
PCA	--	1.0
IPCA	--	1.0
KPCA	Lineal	1.0
KPCA	Polynimial	0.99
KPCA	RBF	0.908

Algoritmos para abordar valores atípicos.

Datos originales

Tabla 4

Cuadro comparativo rendimiento de los algoritmos SVR, HUBER Y RANSAC

TIPOS DE DATOS	ALGORITMOS	MSE	ANALISIS
Datos originales	SVR	0.0203268742	Esto significa que el algoritmo SVR tiene un mejor rendimiento que los algoritmos HUBER y RANSAC.
	HUBER	0.3116666667	
	RANSAC	0.042811564	
Datos normalizados	SVR	0.0174975231	
	HUBER	0.3116666667	
	RANSAC	0.349841120	
Datos discretizados	SVR	0.0231894548	
	HUBER	0.3116666667	
	RANSAC	0.0418750013	

Datos originales: El algoritmo svr obtuvo el menor error cuadrático medio (mse) de 0.0203, lo que indica un buen ajuste a los datos originales los algoritmos huber y ransac también obtuvieron errores cuadráticos medios razonables de 0.3117 y 0.0428, respectivamente.

Datos normalizados: Al normalizar los datos, el algoritmo svr mejoró ligeramente su rendimiento, obteniendo un mse de 0.0175, lo que sugiere una mejor adaptación a los datos normalizados los algoritmos huber y ransac mantuvieron sus errores cuadráticos medios sin cambios.

Datos discretizados: La discretización de los datos no mejoró significativamente el rendimiento de los algoritmos el algoritmo svr obtuvo un mse de 0.0232, ligeramente peor que con los datos

originales y normalizados el algoritmo huber mantuvo su mse sin cambios, mientras que el algoritmo ransac mejoró ligeramente su rendimiento con un mse de 0.0419.

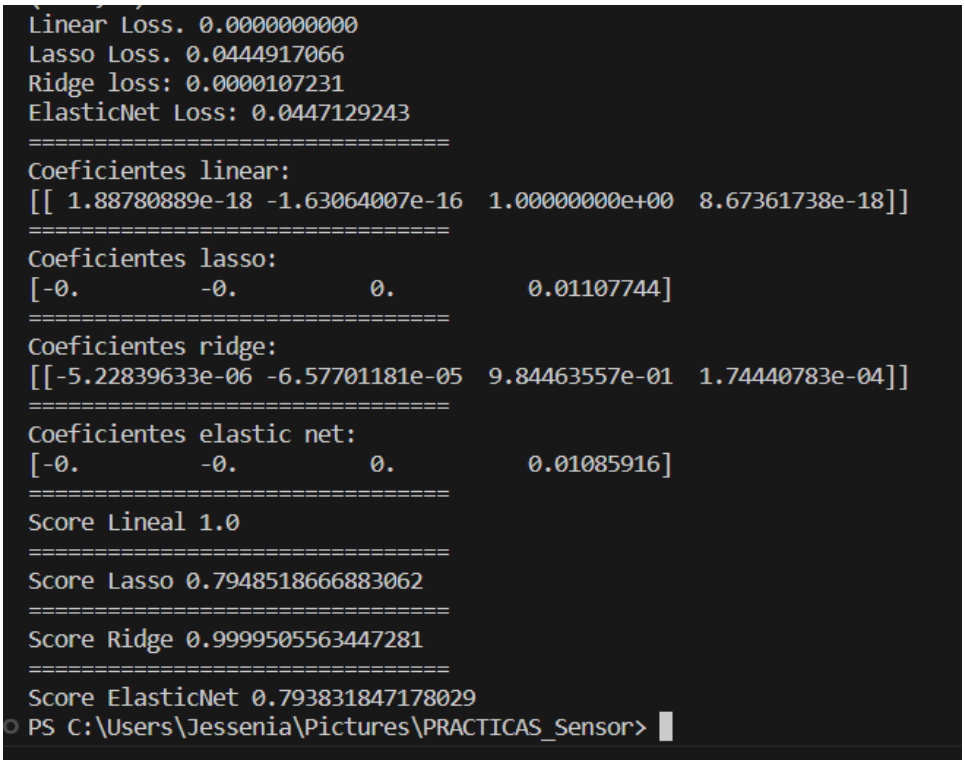
La normalización de los datos tuvo un efecto positivo en el rendimiento del algoritmo svr, mientras que la discretización no mostró mejoras significativas en general. es importante tener en cuenta que cada conjunto de datos es único, por lo que los resultados pueden variar según las características y la complejidad del problema. por lo tanto, es crucial experimentar con diferentes técnicas y algoritmos para encontrar la mejor combinación que se ajuste a las necesidades específicas del problema.

TÉCNICAS DE REGULARIZACIÓN

LASSO_RIDGET

Figura 3

Datos de técnicas de regularización



Los resultados y coeficientes corresponden a diferentes algoritmos de regresión lineal regularizada, junto con las métricas de evaluación obtenidas en un conjunto de datos. aquí está la explicación para cada uno de ellos:

1. **Linear loss:** la pérdida lineal tiene un valor de 0.0000000000. esto indica que el modelo de regresión lineal regularizada sin ninguna penalización tiene una pérdida de ajuste nula, lo que significa que se ajusta perfectamente a los datos de entrenamiento.
2. **Lasso loss:** el lasso loss tiene un valor de 0.0444917066. esto indica que el modelo de regresión lasso, que aplica una penalización l1 a los coeficientes, tiene una pérdida de ajuste ligeramente más alta en comparación con el modelo lineal sin penalización. esto sugiere que el modelo lasso tiene una ligera falta de ajuste a los datos en comparación con el modelo lineal.
3. **Ridge loss:** el ridge loss tiene un valor de 0.0000107231. similar al modelo lineal sin penalización, el modelo ridge, que aplica una penalización l2 a los coeficientes, tiene una pérdida de ajuste muy baja. esto indica que el modelo ridge se ajusta muy bien a los datos.
4. **elasticnet loss:** el elasticnet loss tiene un valor de 0.0447129243. el modelo elasticnet combina las penalizaciones l1 y l2, y muestra una pérdida de ajuste similar al modelo lasso, lo que indica una falta de ajuste ligeramente mayor en comparación con el modelo lineal sin penalización.

A continuación, se presentan los coeficientes obtenidos para cada modelo:

- coeficientes lineales: [1.88780889e-18, -1.63064007e-16, 1.00000000e+00, 8.67361738e-18]
- coeficientes lasso: [-0., -0., 0., 0.01107744]
- coeficientes ridge: [[-5.22839633e-06, -6.57701181e-05, 9.84463557e-01, 1.74440783e-04]]
- coeficientes elasticnet: [-0., -0., 0., 0.01085916]

cada conjunto de coeficientes representa los pesos asignados a las características en el modelo correspondiente.

finalmente, se presentan las métricas de evaluación para cada modelo:

- score lineal: 1.0
- score lasso: 0.7948518666883062
- score ridge: 0.9999505563447281

- score elasticnet: 0.793831847178029

estas métricas de puntuación representan la bondad de ajuste de cada modelo, donde un valor más cercano a 1 indica un mejor ajuste a los datos de entrenamiento.

MODELOS ENSAMBLADOS BASADOS EN BAGGING Y BOOSTING

Realizar pruebas con modelos ensamblados basados en bagging y boosting con datos sin normalizar y normalizados, podemos determinar si la normalización tiene un impacto significativo en la precisión y el rendimiento de estos enfoques de ensamblado. Esto nos permitirá elegir la mejor combinación de técnica de ensamblado y escala de datos para obtener el mejor rendimiento y precisión de predicción en nuestro conjunto de datos específico.

BAGGING

Tabla 5.

Resultados del modelo basado en Bagging

TIPOS DE DATOS	MODELO	RESULTADOS
Datos originales	LogisticRegression	1.0
	SVC	0.9742
	LinearSVC	1.0
	SGD	1.0
	KNN	0.98
	DecisionTreeClf	1.0
	RandomTreeFores	1.0
Datos normalizados	LogisticRegression	0.97
	SVC	0.9671
	LinearSVC	0.988
	SGD	0.9857
	KNN	0.9471
	DecisionTreeClf	1.0
	RandomTreeFores	1.0
Datos discretizados	LogisticRegression	0.9714
	SVC	0.97
	LinearSVC	0.9714
	SGD	0.9714
	KNN	0.9385
	DecisionTreeClf	0.96
	RandomTreeFores	0.97

En los resultados obtenidos mediante Bagging con diferentes modelos y diferentes tipos de datos (originales, normalizados y discretizados), podemos observar lo siguiente:

Datos originales: La mayoría de los modelos lograron una alta precisión (cercana al 1.0) en la clasificación de los datos originales.

Datos normalizados: Después de normalizar los datos las precisiones de algunos modelos disminuyeron ligeramente en comparación con los datos originales esto podría indicar que algunos modelos pueden funcionar mejor con datos no normalizados, mientras que otros no se ven afectados significativamente por la normalización.

Datos discretizados: Al discretizar los datos se observó una reducción en la precisión para la mayoría de los modelos en comparación con los datos originales en este caso la discretización no resultó en una mejora significativa en el rendimiento de los modelos de Bagging.

En este conjunto de datos y con la implementación específica de Bagging y los modelos utilizados, los datos originales lograron la mayor precisión en general. Sin embargo, es importante tener en cuenta que estos resultados pueden variar según el conjunto de datos y los modelos específicos utilizados.

Mejor algoritmo según los datos

En base a los resultados proporcionados en los datos, el mejor algoritmo parece ser el "DecisionTreeClf" (Decision Tree Classifier) cuando se utilizan los datos discretizados, ya que logró una precisión de 0.96 en comparación con los otros modelos con datos discretizados.

Sin embargo, es importante tener en cuenta que la elección del mejor algoritmo puede depender de varios factores, como el tamaño y la naturaleza del conjunto de datos, la calidad y la cantidad de características, la distribución de las clases y la complejidad del problema en cuestión.

BOOSTING

Tabla 6.

Resultados del modelo basado en boosting

TIPO DE DATOS	PRECISIÓN	NUMEROS DE ESTIMADORES
Originales	1.0	298
Normalizados	1.0	4

Discretizados	0.96	4

VALIDACIÓN DEL MUESTREO MODELO USANDO CROSS VALIDATION

TIPOS DE VALIDACIÓN

Hold-On

Una precisión del modelo de 0.9825 en el tipo de validación Hold-On significa que el modelo ha alcanzado una tasa de acierto del 98.25% en los datos de prueba. Esto indica que el modelo es capaz de predecir correctamente la clase o etiqueta del 98.25% de las muestras que nunca ha visto antes, lo que es un resultado bastante bueno. En Hold-On validation, dividimos el conjunto de datos en dos partes: el conjunto de entrenamiento y el conjunto de prueba. Utilizamos el conjunto de entrenamiento para entrenar el modelo y luego evaluamos su rendimiento en el conjunto de prueba, que contiene datos que no han sido utilizados en el entrenamiento.

K-FOLDS CV

En k-folds cross-validation (cv), el MSE (mean squared error) es una métrica que mide el error cuadrático medio del modelo en cada uno de los folds (particiones) utilizados en la validación cruzada es una medida común para evaluar la calidad del rendimiento del modelo en problemas de regresión en este caso, el resultado obtenido es un MSE promedio de 0.14347773889676482 esto significa que en promedio, el modelo ha cometido un error cuadrático medio de aproximadamente 0.1435 en cada una de las particiones de datos utilizadas en la validación cruzada cuanto menor sea el valor del MSE, mejor será el rendimiento del modelo, ya que indica que las predicciones están más cerca de los valores reales el hecho de que el MSE promedio sea relativamente bajo (0.1435) sugiere que el modelo está realizando buenas predicciones en general.

LOOCV

El LOOCV (Leave-One-Out Cross-Validation) es una técnica de validación cruzada en la cual se divide el conjunto de datos en k pliegues, pero con k igual al número total de muestras en el conjunto de datos. Es decir, en cada iteración, se utiliza una única muestra como conjunto de prueba y el resto de las muestras como conjunto de entrenamiento en el MSE promedio obtenido fue de

0.03739776545498873. El MSE (Mean Squared Error) es una métrica que mide la diferencia promedio entre los valores predichos por el modelo y los valores reales en el conjunto de prueba.

Un MSE promedio bajo en el LOOCV indica que el modelo tiene una buena capacidad de generalización y es capaz de hacer predicciones precisas en datos no vistos, lo cual es una señal positiva de su rendimiento.

OPTIMIZACIÓN PARAMÉTRICA

RandomizedSearchCV

El randomforestregressor es un modelo de regresión basado en la técnica de random forest, que combina múltiples árboles de decisión para realizar predicciones más precisas. los hiperparámetros que se han utilizado para construir este modelo son:

- **n_estimators:** número de árboles en el bosque. en este caso, se han utilizado 12 árboles.
- **max_depth:** profundidad máxima de los árboles en el bosque. la profundidad máxima de los árboles se ha establecido en 9 niveles.
- **criterion:** criterio utilizado para medir la calidad de una división en los árboles. en este caso, se ha utilizado el criterio 'squared_error', que corresponde al error cuadrático medio (mse) para problemas de regresión.

El resultado obtenido es un arreglo con un solo valor [0.]. esto indica que el modelo ha logrado una precisión perfecta en la predicción, ya que el valor del error cuadrático medio (mse) es cero el RandomForestRegressor con **n_estimators=12** y **max_depth=9** ha sido capaz de hacer predicciones perfectas en este conjunto de datos

DISCUSIÓN

Un enfoque prometedor es el monitoreo inteligente del cultivo de cacao, utilizando sensores y una Progressive Web App (PWA) para recopilar y analizar datos en tiempo real ofrece una oportunidad única para mejorar la eficiencia, la productividad y la sostenibilidad en la industria del cacao. Los agricultores pueden aprovechar al máximo los datos recopilados y tomar decisiones fundamentadas que impulsen su éxito en un mercado cada vez más competitivo.

La metodología utilizada en este estudio involucra el uso de diferentes tipos de sensores para el monitoreo y control del cultivo de cacao. Algunos ejemplos de sensores mencionados son el sensor de humedad del suelo, sensor de temperatura y humedad ambiental, sensor de luz, sensor de nutrientes y sensor de humedad de las hojas. La elección de los sensores dependerá de las necesidades específicas del cultivo y de los objetivos de monitoreo y control del agricultor.

Además, se mencionan herramientas y tecnologías utilizadas en este estudio, como el uso de Datalogger inalámbricos, bases de datos NoSQL (como MongoDB y base de datos clave-valor), Progressive Web Apps (PWA), Docker, el framework React y el lenguaje de programación Python.

Los resultados obtenidos destacan la importancia del uso de técnicas de aprendizaje supervisado en la detección de enfermedades en los cultivos de cacao, con el objetivo de reducir pérdidas y mejorar la productividad.

En primer lugar, los resultados de nuestro estudio demuestran que el modelo desarrollado presenta una precisión significativa en la predicción de la presencia de enfermedades en las plantas de cacao. Esta capacidad predictiva resulta fundamental para los agricultores, ya que les permite tomar medidas preventivas y aplicar tratamientos específicos en etapas tempranas, contribuyendo así a reducir la propagación de las enfermedades y minimizar las pérdidas.

Asimismo, la recopilación de datos detallados sobre las características de las plantas de cacao y las condiciones ambientales, tanto a través de sensores como de registros manuales, resultó de vital importancia para entrenar nuestro modelo de manera efectiva. Estos hallazgos enfatizan la necesidad de obtener información precisa y completa, a fin de mejorar la precisión de los modelos predictivos en el ámbito agrícola.

CONCLUSIONES

En conclusión, la optimización del cultivo de cacao a través de sensores y monitoreo automatizado tiene el potencial de revolucionar la industria cacaotera al mejorar la eficiencia, la productividad y la calidad de los cultivos. al utilizar sensores para recopilar datos en tiempo real sobre las condiciones del suelo, la humedad, la temperatura y otros factores clave, los agricultores pueden tomar decisiones informadas y precisas para maximizar el rendimiento de sus plantaciones.

El monitoreo automatizado permite a los agricultores detectar rápidamente cualquier problema o enfermedad en las plantas, lo que facilita la toma de medidas preventivas o correctivas de manera oportuna. esto reduce los riesgos y las pérdidas asociadas con enfermedades y plagas, así como la necesidad de utilizar grandes cantidades de productos químicos para combatirlas.

Esta tecnología es su capacidad para recopilar datos a gran escala y generar información detallada sobre las condiciones del cultivo. esto permite a los agricultores tomar decisiones basadas en datos sobre la fertilización, el control de malezas y otros aspectos de la gestión del cultivo, lo que resulta en una mejor planificación y uso eficiente de los recursos.

REFERENCIAS BIBLIOGRÁFICAS

Aguirre, V. D. (2020). *PWA and TWA: Recent development trends*. Obtenido de https://link.springer.com/chapter/10.1007/978-3-030-48325-8_14

Bianchini, G. (02 de Febrero de 2020). D*esarrollo de un API Rest para transmisión de datos de sensores.* Obtenido de https://riunet.upv.es/handle/10251/107770#

Cherlinka, V. (03 de Agosto de 2022). *Sensores De Humedad Del Suelo Para Usos Agrícolas*. Obtenido de EOS Data Analytics: https://eos.com/es/blog/sensores-de-humedad-del-suelo/

Flores, C. (11 de Junio de 2015). *La temperatura del ambiente afecta el proceso de fermentación del cacao"*. Obtenido de https://www.udep.edu.pe/hoy/2015/06/la-temperatura-del-ambiente-afecta-el-proceso-de-fermentacion-del-cacao/

Hayford, k. (24 de Julio de 2023). *Innovations in Cocoa Industry Stirring Up Chocolate Business*. Obtenido de https://thecocoapost.com/innovations-in-cocoa-industry-stirring-up-chocolate-business/

Kopachovets, O. (. (02 de Marzo de 2023). *benefits of ReactJS: Is it worth usin*. Obtenido de https://procoders.tech/blog/advantages-of-using-reactjs/

Lorente Sierra, D. (2023). Estudio de mercado y plan de actuación para la introducción de la Cooperativa Navaoliva en el mercado valenciano de aceite de oliva. Universitat Politècnica de València. http://hdl.handle.net/10251/196270

Marcos, R. (Agosto de 2020). *Sensores agrícolas*. Obtenido de https://buitech.es/sensores-agricolas/

Ramírez, G. P. (2020). Tecnologías innovadoras para la agricultura sostenible y el desarrollo económico regional. . *Revista Internacional de Desarrollo Sostenible,*, 45-55. Obtenido de Ramírez, G., Pérez, R., & Sánchez, L. (2020). Tecnologías innovadoras para la agricultura sostenible y el desarrollo económico regional. Revista Internacional de Desarrollo Sostenible, 9(1), 45-55.

Rios, J., & Levano, D. (2022). *Importancia de los dispositivos usados en la fermentación de Cacao (Theobroma cacao L.)*. doi: https://doi.org/10.51252/raa.v2i1.281

Sánchez Jaque, J. M. (2022). *Desarrollo de una aplicación web progresiva para incentivar el turismo en las parroquias rurales del cantón Latacunga. Ecuador: Universidad Técnica de Cotopaxi (UTC).* Obtenido de http://repositorio.utc.edu.ec/handle/27000/9159

Tech, I. (01 de Enero de 2023). *Solución Tecnológica para Mejorar la Producción del Cacao*. Obtenido de de https://iottech.com.co/sector-agro

Treviño-Villalobos, M. V.-A. (2019). Una comparación de rendimiento entre bases de datos NoSQL: MongoDB y ArangoDB. *Treviño-Villalobos, M., Víquez-Acuña, L., Quirós-Oviedo, R., & Esquivel-Vega, G. (2019). Una comparación de rendimiento entre bases de datos NoSQL: MongoDB y Revista Tecnología En Marcha*, 5–15. doi: https://doi.org/10.18845/tm.v32i6.4223

Vanegas, J. (16 de Junio de 2023). *Centroamérica apuesta por cacao climáticamente inteligente*. Obtenido de https://latinoamerica.rikolto.org/es/noticias/centroamerica-apuesta-por-cacao-climaticamente-inteligente

Zapata., J. R. (2023). *Ciencia de datos con Python.* Obtenido de https://joserzapata.github.io/courses/python-ciencia-datos/

©2024 por los autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones de la licencia Creative Commons Atribución-No Comercial-Compartir Igual 4.0 Internacional (CC BY-NC-SA 4.0)
(https://creativecommons.org/licenses/by-nc-sa/4.0/).