



Desarrollo de un sistema automatizado basado en autómatas finitos para la limpieza y gestión de datos generados por paneles solares en bases de datos NoSQL

Development of an Automated System Based on Finite Automata for Cleaning and Managing Data Generated by Solar Panels in NoSQL Databases

José Daniel Arenas Betancur
jdarenas@poligran.edu.co

Jonatan Arenas Sepúlveda
jarenass@poligran.edu.co

Wilber Misas García
wmisas@poligran.edu.co
Institución Universitaria Politécnico Grancolombiano
Colombia

Recepción: 22/03/2024
Aceptación: 01/07/2025

DOI: <https://doi.org/10.15765/6g02c989>

Resumen

La expansión del uso de paneles solares ha generado grandes volúmenes de datos provenientes de sensores que registran variables como producción de energía, temperatura e irradiancia. Sin embargo, estos registros suelen presentar errores como valores faltantes, datos duplicados o inconsistencias que afectan la calidad del análisis y la precisión de los modelos predictivos.

Ante esta problemática, se plantea el desarrollo de un sistema automatizado basado en autómatas finitos para la detección, clasificación y corrección de errores en los datos generados por sistemas fotovoltaicos. Este sistema se integra con una base de datos NoSQL, específicamente MongoDB, lo que permite almacenar los registros corregidos de forma eficiente y estructurada.

La metodología implementada incluye el diseño del autómata, su integración con la base de datos y la validación mediante conjuntos de datos simulados. Los resultados demuestran una mejora significativa en



la calidad y procesamiento de los datos respecto a métodos manuales, reduciendo errores humanos y aumentando la eficiencia.

Este proyecto aporta una solución práctica y escalable para el tratamiento automatizado de datos en sistemas solares, mejorando la confiabilidad de la información y facilitando su análisis en aplicaciones energéticas.

Palabras clave

Autómatas finitos, limpieza de datos, sistemas fotovoltaicos, bases de datos NoSQL, automatización, calidad de datos

INTRODUCCIÓN

A nivel internacional, se han propuesto diversas estrategias para mejorar la integridad de los datos en entornos energéticos, entre ellas, el uso de técnicas de aprendizaje automático, lógica difusa y modelos estadísticos (Rahman et al., 2022). Sin embargo, estas soluciones suelen requerir grandes volúmenes de datos etiquetados o altos recursos computacionales. En este contexto, los autómatas finitos surgen como una alternativa eficiente, de bajo costo computacional y fácilmente integrable con sistemas de bases de datos modernas (Hopcroft & Ullman, 2006; Academia.edu, s.f.).

Esta investigación propone el desarrollo de un sistema automatizado basado en autómatas finitos para la clasificación, limpieza y validación de los datos generados por paneles solares, almacenados en bases de datos NoSQL como MongoDB. La pertinencia del estudio se fundamenta en la necesidad de garantizar la calidad de los datos para análisis confiables y eficientes, especialmente en contextos donde se manejan grandes volúmenes de información de forma continua (Xpert.Digital, 2020).

Además, el uso de bases de datos NoSQL permite una mayor flexibilidad y escalabilidad en el almacenamiento, alineándose con las demandas actuales de sistemas energéticos inteligentes (OpenSistemas, 2023). La investigación busca aportar una solución práctica y replicable para empresas del sector energético y académico, mejorando la eficiencia en la gestión de datos y promoviendo el uso responsable de la información en el ámbito de las energías renovables.

OBJETIVOS

Desarrollar un sistema automatizado basado en autómatas finitos para la gestión y limpieza de los datos generados por paneles solares, optimizando su almacenamiento en bases de datos NoSQL y mejorando la precisión y eficiencia en el procesamiento de la información.

3.1. Objetivos específicos

3.1.1. Diseñar un autómata finito que clasifique los registros de los paneles solares en “dato válido”, “dato faltante”, “dato duplicado” y “dato erróneo”, estableciendo las reglas de transición entre estos estados para garantizar la detección de inconsistencias.



3.1.2. Implementar un sistema automatizado que procese los registros provenientes de los paneles solares, integrando el autómata finito con una base de datos NoSQL como MongoDB para almacenar los datos clasificados y corregidos.

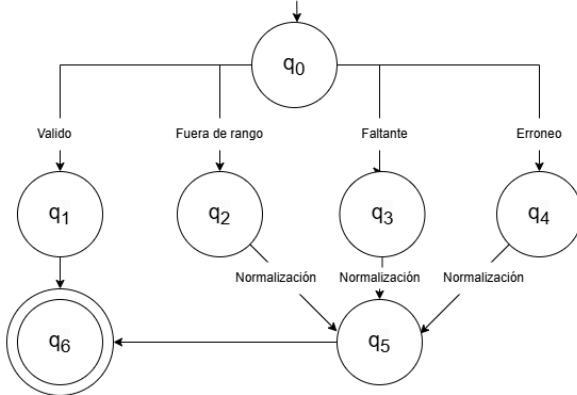
3.1.3. Validar la precisión y eficiencia del sistema automatizado mediante pruebas con conjuntos de datos reales y simulados, comparando los resultados obtenidos con los métodos manuales de limpieza de datos para evaluar su desempeño y fiabilidad.

MÉTODO

Este proyecto sigue un enfoque computacional-experimental, centrado en el desarrollo, implementación y validación de un sistema automatizado para la limpieza de datos.

Fase 1: Diseño del sistema

- Definir los estados del autómata para clasificar los registros en “dato válido”, “dato faltante”, “dato duplicado” y “dato erróneo”.
- Establecer las reglas de transición entre los estados del autómata.
- Seleccionar MongoDB como base de datos NoSQL para almacenar los registros procesados.



Fase 2: Implementación del sistema

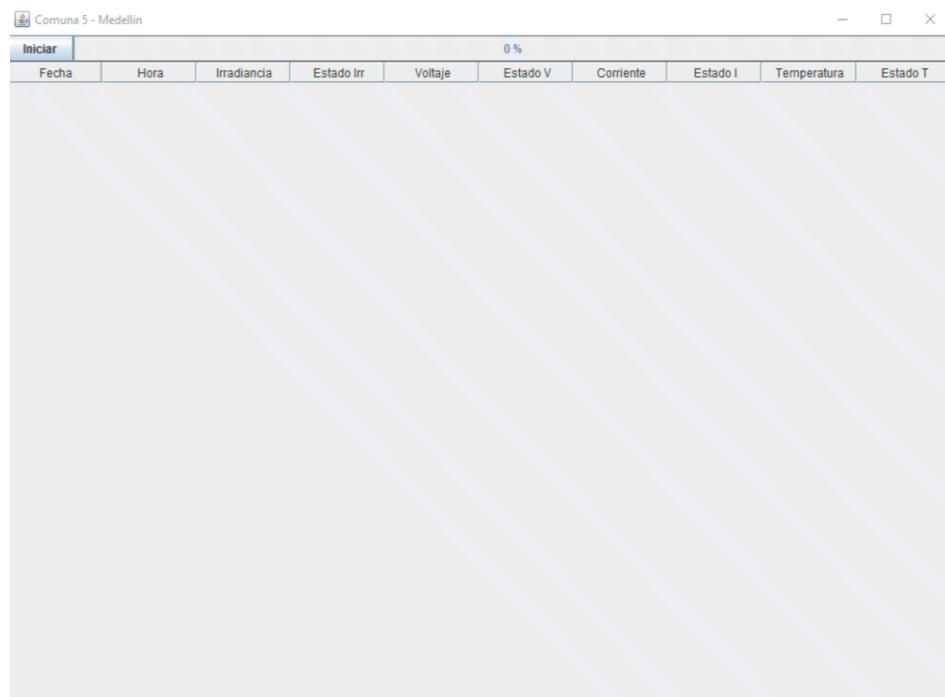
- Desarrollar el autómata finito que procese los registros y los clasifique según su estado.
- Integrar el autómata con la base de datos para almacenar los datos corregidos.
- Implementar funciones de corrección automática según el tipo de error identificado.

Fase 3: Validación del sistema

- Realizar pruebas con conjuntos de datos simulados evaluando la fiabilidad del autómata.



- Evaluar la precisión del sistema en la detección y corrección de errores.
- Comparar los resultados obtenidos con métodos manuales de limpieza para validar la eficiencia.



Este
garantiza

enfoque
un



procesamiento sistemático y automatizado de los datos, mejorando la calidad de la información y facilitando su gestión en grandes volúmenes.

RESULTADOS

Durante las pruebas realizadas con datos simulados, el sistema demostró alta precisión en la detección y corrección de errores, lo que generó una mejora significativa en comparación con los métodos manuales de limpieza de datos. Se observó una reducción en el tiempo de procesamiento y una mayor consistencia en la estructura de los datos almacenados.

La integración con MongoDB fue exitosa, permitiendo la trazabilidad de los datos corregidos y su recuperación eficiente. Esto evidenció que el sistema es capaz de gestionar grandes volúmenes de información con confiabilidad y eficiencia.

En general, los resultados validaron la viabilidad del enfoque basado en autómatas finitos para entornos de datos provenientes de sistemas fotovoltaicos, proporcionando una herramienta útil para el análisis y monitoreo de estos sistemas en tiempo real.

DISCUSIÓN Y CONCLUSIÓN

- Se logró el objetivo general al desarrollar un sistema automatizado basado en autómatas finitos que permite gestionar y limpiar los datos generados por paneles solares, optimizando su almacenamiento en bases de datos NoSQL y mejorando la eficiencia en el procesamiento.
- El autómata finito fue correctamente diseñado para clasificar los datos en cuatro categorías: válidos, faltantes, duplicados y erróneos, cumpliendo con el primer objetivo específico y asegurando una estructura de control clara para la limpieza.
- El sistema automatizado fue implementado con éxito e integrado con MongoDB. Los registros procesados se almacenan adecuadamente en la base de datos, permitiendo trazabilidad y recuperación eficiente de los datos corregidos, cumpliendo con el segundo objetivo específico.
- La validación del sistema se realizó mediante el uso de datos simulados, demostrando que el sistema ejecuta sus funciones de manera eficiente, precisa y confiable. Los resultados obtenidos mostraron mejoras significativas respecto a los métodos manuales, cumpliendo así el tercer objetivo específico.
- El sistema demuestra ser funcional y útil en contextos donde se manejan grandes volúmenes de información proveniente de sensores solares, brindando una solución efectiva para garantizar la calidad de los datos y facilitar su análisis posterior.

REFERENCIAS BIBLIOGRÁFICAS

Academia.edu. (s.f.). Aplicaciones de autómatas finitos deterministas. Recuperado de https://www.academia.edu/8425334/Aplicaciones_de_Aut%C3%B3matas_Finitos_Deterministas



Rahman, S., et al. (2022). Data-driven fault detection in PV systems. Renewable Energy.
<https://www.sciencedirect.com/science/article/pii/S2590123024000884>

Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2018). Metodología de la investigación (6^a ed.). McGraw-Hill.

Montgomery, D. C. (2020). Design and analysis of experiments (10th ed.). Wiley.

Hopcroft, J. E., & Ullman, J. D. (2006). Introduction to Automata Theory, Languages, and Computation.
<https://dl.acm.org/doi/abs/10.1145/568438.568455>

OpenSistemas. (2023). NoSQL: Soluciones prácticas para bases de datos (con ejemplos). Recuperado de
<https://opensistemas.com/soluciones-practicas-para-bases-de-datos-nosql>

Xpert.Digital. (2020). Detección temprana de errores en sistemas fotovoltaicos. Recuperado de
<https://xpert.digital/es/deteccion-temprana-de-errores-en-sistemas-fotovoltaicos>