

Modelo predictivo para la acción anticipatoria y alerta temprana en eventos climáticos asociados a inundaciones (MONITOR)

Jeremías Pabón López¹, Jesús Alfredo Sanabria Mejía²

¹⁻²Facultad de Ingeniería y Ciencias Básicas (FICB), ³

Universidad Central de Colombia

Maestría en Analítica de Datos

Automatización e Integración de Datos IA

Bogotá, Colombia

¹jpabonl@ucentral.edu.co, ²jsanabriam4@ucentral.edu.co

April 13, 2024

Contents

1	Introducción (Max 250 Palabras) - (<i>Primera entrega</i>)	3
2	Características del proyecto de investigación que hace uso de Integración y Automatización de Datos para IA (Max 500 Palabras) - (<i>Primera entrega</i>)	4
2.1	Titulo del proyecto de investigación (Max 100 Palabras) - (<i>Primera entrega</i>)	4
2.2	Objetivo general (Max 100 Palabras) - (<i>Primera entrega</i>)	4
2.2.1	Objetivos especificos (Max 100 Palabras) - (<i>Primera entrega</i>)	4
2.3	Alcance (Max 200 Palabras) - (<i>Primera entrega</i>)	4
2.4	Pregunta de investigación (Max 100 Palabras) - (<i>Primera entrega</i>)	5
2.5	Hipotesis (Max 100 Palabras) - (<i>Primera entrega</i>)	5
3	Reflexiones sobre el origen de datos e información (Max 400 Palabras) - (<i>Primera entrega</i>)	6
3.1	¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (<i>Primera entrega</i>)	6
3.2	¿Cuales son las consideraciones legales o éticas del uso de la información? (Max 100 Palabras) - (<i>Primera entrega</i>)	6

3.3	¿Cuales son los retos de la información y los datos que utilizara en Integración y Automatización de Datos para IA? (Max 100 Palabras) - (<i>Primera entrega</i>)	7
3.4	¿Que espera de la utilización de Integración y Automatización de Datos para IA para su proyecto? (Max 100 Palabras) - (<i>Primera entrega</i>)	7
4	Diseño de integración y Automatización de Datos para IA (Diagrama) (<i>Primera entrega</i>)	8
5	integración de Datos (<i>Segunda entrega</i>)	9
6	Automatización de Datos (<i>Segunda entrega</i>)	9
7	IA (<i>Segunda entrega</i>)	10
8	Proximos pasos (<i>Tercera entrega</i>)	11
9	Lecciones aprendidas (<i>Tercera entrega</i>)	12
10	Bibliografía	13

1 Introducción (Max 250 Palabras) - (*Primera entrega*)

Colombia cuenta con diversos sistemas de alerta temprana para fenómenos de gran escala, gestionados por entidades nacionales como IDEAM, DIMAR, Corporación OSSO y SGC. A nivel regional, las Corporaciones Ambientales, como la CAR, también han implementado sistemas, al igual que algunos municipios, como el Área Metropolitana del Valle de Aburrá, o a nivel distrital las experiencias relacionadas con IDIGER en Bogotá; o el SIATA en Medellín.

Los sistemas existentes no están diseñados para generar alertas a nivel local que le permita a las comunidades prepararse de manera anticipada frente a los efectos del fenómeno de la niñez, otros eventos climáticos repentinos o alteraciones en los sistemas hídricos como taponamiento de ríos. Los datos y pronósticos existentes pueden estar dispersos en varias fuentes, lo que dificulta que las comunidades y autoridades accedan a información completa y actualizada. En muchas regiones alejadas del centro del país, las herramientas oficiales que ofrecen información son severamente limitadas.

Actualmente, la respuesta en zonas vulnerables tiende a ser reactiva más que anticipatoria. En lugar de prepararse proactivamente para los impactos previstos de La Niña, las comunidades y autoridades suelen esperar hasta que ocurra el desastre antes de tomar medidas. Este enfoque reactivo limita la eficacia de los esfuerzos de respuesta y exacerba los desafíos que enfrentan las comunidades afectadas, destacando la necesidad de medidas anticipadas para reducir la vulnerabilidad y mejorar la resiliencia.

En la actualidad, las comunidades no son alertadas oportunamente sobre eventos de inundaciones, estos eventos ocurren e impactan la vida de las comunidades desde el punto de vista de la salud; causando infecciones y otras enfermedades relacionadas con picaduras de mosquitos (dengue); en la educación, las inundaciones interrumpen el aprendizaje, dificultando la permanencia en la escuela; también muchas familias pierden sus pertenencias, reciben daños en sus viviendas e incluso sus medios de vida se ven severamente afectados, lo que impacta directamente en la economía de los hogares y en su seguridad alimentaria. Como consecuencia de lo anterior, las familias se ven afectadas emocionalmente, lo que requiere una atención adecuada para que las familias puedan aplicar estrategias de afrontamiento positivas, contribuir a la resiliencia y mitigar los riesgos de protección para los niños.

La propuesta de este proyecto busca incorporar fuentes de información oficiales disponibles e información generada por las comunidades para ofrecer una solución que permita generar información y alertas oportunamente a las comunidades, autoridades locales y nacionales, organizaciones de la sociedad civil y organismo de socorro en la preparación anticipada y respuesta efectiva sobre los eventos asociados a las inundaciones.

2 Características del proyecto de investigación que hace uso de Integración y Automatización de Datos para IA (Max 500 Palabras) - (*Primera entrega*)

2.1 Título del proyecto de investigación (Max 100 Palabras) - (*Primera entrega*)

Modelo predictivo para la acción anticipatoria y alerta temprana en eventos climáticos asociados a inundaciones (MONITOR)

2.2 Objetivo general (Max 100 Palabras) - (*Primera entrega*)

Desarrollar y pilotar un modelo predictivo basado en redes neuronales para la acción anticipatoria y alerta temprana en eventos climáticos asociados a inundaciones en Bogotá.

2.2.1 Objetivos específicos (Max 100 Palabras) - (*Primera entrega*)

- Identificar y recopilar datos de fuentes oficiales y comunitarias pertinentes para la predicción de inundaciones.
- Desarrollar un modelo basado en redes neuronales convolucionales (CNN) y redes neuronales recurrentes (RNN) para predecir las inundaciones en la zona de estudio.
- Pilotar el modelo en comunidades del distrito capital para evaluar su efectividad, aceptación y usabilidad por parte de las comunidades locales y las instituciones de gestión de desastres.

2.3 Alcance (Max 200 Palabras) - (*Primera entrega*)

Se propone desarrollar un sistema de alerta temprana que se basará en la recopilación de datos provenientes de plataformas institucionales de monitoreo del clima y de información comunitaria accesible. Utilizando esta información, se desarrollará un modelo predictivo que identificará variables clave para la predicción de fenómenos climáticos. Este modelo se implementará en un tablero de información público, disponible para ser insertado en plataformas web de instituciones y organizaciones relevantes.

Además, se establecerá un mecanismo de reporte accesible para la comunidad, permitiendo el envío de información que alimente el modelo a través de dispositivos tecnológicos y la red de internet. El sistema generará alertas de acuerdo con los protocolos y umbrales establecidos por las autoridades competentes, las cuales serán comunicadas mediante correos electrónicos a instituciones, organismos de socorro y organizaciones de la sociedad civil. Es importante destacar que

este modelo no pretende reemplazar los protocolos y alertas existentes, sino complementarlos para mejorar la toma de decisiones por parte de las partes interesadas.

2.4 Pregunta de investigación (Max 100 Palabras) - (*Primera entrega*)

¿Cómo puede la aplicación de un modelo predictivo basado en redes neuronales mejorar la capacidad de anticipación y respuesta ante eventos climáticos relacionados con inundaciones en Bogotá, involucrando a comunidades, instituciones gubernamentales y organizaciones de la sociedad civil?

2.5 Hipotesis (Max 100 Palabras) - (*Primera entrega*)

La implementación de redes neuronales convolucionales y recurrentes mejorará la anticipación de inundaciones en Bogotá.

La implementación exitosa de un modelo predictivo basado en redes neuronales convolucionales (CNN) y recurrentes (RNN) podría revolucionar la capacidad de anticipación y alerta temprana frente a eventos climáticos, específicamente inundaciones, en Bogotá. Al aprovechar la precisión y la capacidad de captura de patrones espaciales y temporales de datos de multisensores, el modelo podría ofrecer predicciones más precisas y detalladas sobre la ocurrencia y la intensidad de inundaciones en la ciudad. Además, la disponibilidad de datos precisos y oportunos, tanto de fuentes oficiales como comunitarias, será fundamental para la efectividad del modelo. La aceptación y adopción del modelo por parte de las comunidades locales y las instituciones de gestión de desastres también desempeñarán un papel crucial en su utilidad a largo plazo. Si se logra una implementación exitosa y una aceptación generalizada, este modelo podría no solo mejorar la capacidad de respuesta ante inundaciones en Bogotá, sino también servir como un modelo replicable y escalable para otras ciudades enfrentando desafíos similares en todo el mundo.

3 Reflexiones sobre el origen de datos e información

(Max 400 Palabras) - (*Primera entrega*)

Para el desarrollo del proyecto, se considerarán tres fuentes de información clave. En primer lugar, se aprovecharán los sistemas de información meteorológica proporcionados por entidades nacionales, regionales e internacionales, los cuales suministran datos y variables determinantes relacionados con las inundaciones. Además, se recurrirá a información de eventos históricos disponible en informes, noticias y otros documentos para establecer patrones y métricas sobre las probabilidades de ocurrencia de las inundaciones y su impacto. Por último, se contará con un canal que permitirá recopilar datos directamente desde las comunidades, incluyendo conocimientos ancestrales y reportes sobre eventos climáticos en tiempo real.

3.1 ¿Cual es el origen de los datos e información ? (Max 100 Palabras) - (*Primera entrega*)

El origen de los datos e información para el desarrollo del proyecto proviene de tres fuentes principales. En primer lugar, se aprovechan los sistemas de información meteorológica proporcionados por entidades nacionales, regionales e internacionales, que suministran datos y variables relevantes sobre las inundaciones. Además, se recurre a información de eventos históricos disponible en informes, noticias y otros documentos para establecer patrones y métricas sobre las probabilidades de ocurrencia de las inundaciones y su impacto. Por último, se cuenta con un canal que permite recopilar datos directamente desde las comunidades, incluyendo conocimientos ancestrales y reportes sobre eventos climáticos en tiempo real. Esta diversidad de fuentes garantiza una visión integral y actualizada para abordar eficazmente los desafíos relacionados con las inundaciones.

3.2 ¿Cuales son las consideraciones legales o éticas del uso de la información? (Max 100 Palabras) - (*Primera entrega*)

El uso de la información en un proyecto como este conlleva una serie de consideraciones legales y éticas importantes que deben ser tenidas en cuenta. En términos generales, es fundamental destacar el compromiso con la precisión y la transparencia en la comunicación de pronósticos y alertas. Esto es crucial para evitar consecuencias negativas. Además, un mal uso de esta información podría generar pánico, lo que a su vez podría resultar en accidentes o una falta de preparación real para atender emergencias. Por lo tanto, también es esencial garantizar un acceso equitativo a esta información para asegurar la seguridad y el bienestar de toda la sociedad,

3.3 ¿Cuales son los retos de la información y los datos que utilizara en Integración y Automatización de Datos para IA? (Max 100 Palabras) - (Primera entrega)

Los desafíos en la Integración y Automatización de Datos para IA implican superar barreras como el acceso a datos actualizados y fiables, esenciales para construir modelos predictivos efectivos. Además, se requiere capacidad de procesamiento de datos en tiempo real para manejar el volumen y la variedad de datos de múltiples fuentes. Es crucial evitar costos excesivos al integrar estas fuentes y asegurar eficiencia y rentabilidad en el proceso. Los resultados deben ser seguros y confiables, proporcionando información clara y útil para las autoridades y comunidades. Esto permitirá una toma de decisiones optimizada y una mejor preparación y respuesta frente a los fenómenos climáticos, específicamente las inundaciones.

3.4 ¿Que espera de la utilización de Integración y Automatización de Datos para IA para su proyecto? (Max 100 Palabras) - (Primera entrega)

Se espera que la utilización de diversas fuentes de información, junto con el conocimiento comunitario, alimente un modelo computacional autónomo para gestionar datos según parámetros técnicos establecidos. Este enfoque busca generar resultados eficientes que fortalezcan la preparación ante inundaciones, reduciendo sus impactos adversos en áreas como la salud, la educación, la economía y la salud mental. Además, se pretende abordar riesgos específicos de género, asegurando una gestión inclusiva y equitativa de desastres. Integrar esta perspectiva en la herramienta puede proporcionar información clave a instituciones para desarrollar planes de prevención y respuesta adaptados a cada comunidad, fortaleciendo la capacidad de anticipación y respuesta. Este enfoque proactivo busca no solo mitigar los efectos negativos de las inundaciones, sino también promover una mayor resiliencia comunitaria.

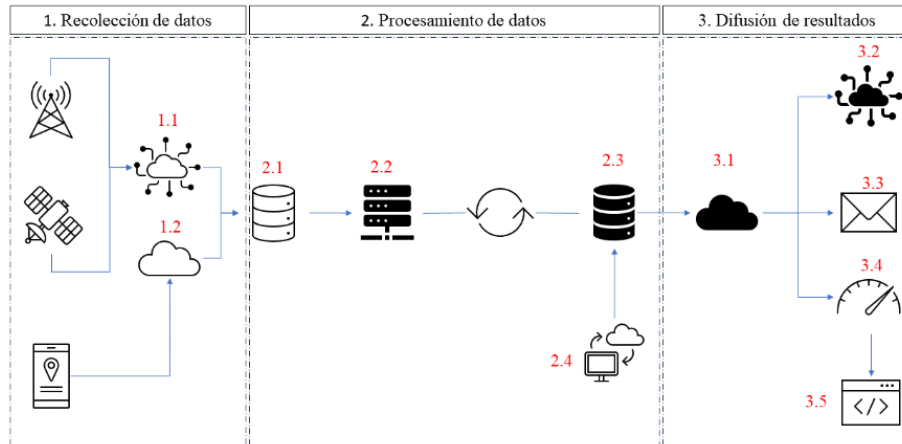


Figure 1: Diseño de integración y Automatización de Datos para IA - Fuente: elaboración propia

4 Diseño de integración y Automatización de Datos para IA (Diagrama) *(Primera entrega)*

1. Fuente de datos

(1.1) Datos institucionales API REST

1.1.1. Sensores (estaciones meteorológicas) → Disponible en: Datos abiertos

1.1.2. Imágenes de satélite → disponibles en:

1.1.3. Datos históricos → Disponibles en:

(1.2) Formulario comunitario CSV

2. Procesamiento de datos

(2.1) Se usará el servicio de big data Azure Synapse Analytics, que integra almacenamiento, análisis y unificación de datos de diferentes fuentes en una única plataforma.

(2.2) Para el desarrollo del modelo predictivo con redes neuronales, se utilizarán las bibliotecas y marcos de trabajo de Python, como TensorFlow, Keras y PyTorch, los cuales simplifican la implementación de redes neuronales convolucionales y recurrentes.

(2.3) Se usará Azure Data Lake Storage para el almacenamiento de los datos ya procesados por el modelo predictivo y que será accedido

para el desarrollo de informes y que retroalimenta el modelo para mejorar su predicción.

- (2.4) Administración de parámetros e información de notificaciones para las alertas / Análisis y generación de indicadores

3. Difusión de resultados

- (3.1) Se usará Azure Data Lake Storage para el almacenamiento de los datos ya procesados por el modelo predictivo y que será accedido para el desarrollo de informes y que retroalimenta el modelo para mejorar su predicción.
- (3.2) Se podrá disponible API para que sea accedida de forma externa, para esto se hará uso de Swagger Editor que es un editor de código abierto basado en navegador para diseñar y documentar API utilizando la especificación OpenAPI.
- (3.3) Para la activación del sistema de alerta para los actores interesados, se hará uso de Azure Communication Services que permite la comunicación multicanal para agregar voz, vídeo, chat, mensajería de texto/SMS y correo electrónico.
- (3.4) Se hará el desarrollo de un tablero de datos o dashboard en herramientas BI como Power BI y Tableau.
- (3.5) Código HTML para incrustar tablero de datos en sitios web externos con analizar automatizados claros y concisos que puedan ser usados y difundidos por diferentes canales (radios, redes sociales, etc)

5 integración de Datos (*Segunda entrega*)

Vulputate nec hac convallis rutrum eu ante volutpat aliquam ullamcorper pulvinar tristique velit nulla, cubilia felis tempor aptent vitae rhoncus parturient euismod mauris potenti dignissim magna. Nunc nec cum quisque quam tincidunt mauris nascetur conubia placerat fusce consequat eget erat vulputate, est rhoncus etiam dignissim luctus volutpat facilisi molestie torquent at convallis accumsan. Magnis morbi auctor sapien fusce turpis leo ad libero vivamus, sem enim ultrices elementum curae accumsan vel hendrerit. Etiam elementum dui a sodales auctor lacus proin interdum, porttitor netus tortor blandit sociis facilisi ullamcorper, mi aenean euismod diam placerat dignissim class.

6 Automatización de Datos (*Segunda entrega*)

Vulputate nec hac convallis rutrum eu ante volutpat aliquam ullamcorper pulvinar tristique velit nulla, cubilia felis tempor aptent vitae rhoncus parturient euismod mauris potenti dignissim magna. Nunc nec cum quisque quam tincidunt mauris nascetur conubia placerat fusce consequat eget erat vulputate, est rhoncus etiam dignissim luctus volutpat facilisi molestie torquent at convallis accumsan.

Magnis morbi auctor sapien fusce turpis leo ad libero vivamus, sem enim ultrices elementum curae accumsan vel hendrerit. Etiam elementum dui a sodales auctor lacus proin interdum, porttitor netus tortor blandit sociis facilisi ullamcorper, mi aenean euismod diam placerat dignissim class.

7 IA (*Segunda entrega*)

Vulputate nec hac convallis rutrum eu ante volutpat aliquam ullamcorper pulvinar tristique velit nulla, cubilia felis tempor aptent vitae rhoncus parturient euismod mauris potenti dignissim magna. Nunc nec cum quisque quam tincidunt mauris nascetur conubia placerat fusce consequat eget erat vulputate, est rhoncus etiam dignissim luctus volutpat facilisi molestie torquent at convallis accumsan. Magnis morbi auctor sapien fusce turpis leo ad libero vivamus, sem enim ultrices elementum curae accumsan vel hendrerit. Etiam elementum dui a sodales auctor lacus proin interdum, porttitor netus tortor blandit sociis facilisi ullamcorper, mi aenean euismod diam placerat dignissim class.

8 Proximos pasos (*Tercera entrega*)

9 Lecciones aprendidas (*Tercera entrega*)

10 Bibliografía