

Propuesta de implementación de inteligencia artificial y Big Data



Propuesta de implementación de inteligencia artificial y Big Data para el estudio de hábitos alimenticios y tecnológicos en niños sedentarios de 6 a 10 años en la localidad de Engativá, Bogotá.

Aliss Vanessa Banguero Gueso

Corporación Universitaria Minuto de Dios

Rectoría Virtual

Programa Especialización en Gerencia de Proyectos

junio de 2025

Propuesta de implementación de inteligencia artificial y Big Data

Propuesta de implementación de inteligencia artificial y Big Data para el estudio de hábitos alimenticios y tecnológicos en niños sedentarios de 6 a 10 años en la localidad de Engativá, Bogotá

Aliss Vanessa Banguero Grueso

Trabajo de Grado presentado como requisito para optar al título de Especialista en Gerencia de Proyectos

Asesor(a)  
Sergio Andrés Zabala Vargas  
Doctor en Tecnología Educativa

Corporación Universitaria Minuto de Dios  
Rectoría Virtual  
Programa Especialización en Gerencia de Proyectos  
junio de 2025

**Contenido**

Lista de tablas..... 5

Lista de figuras..... 6

Resumen..... 8

Abstract ..... 9

Introducción ..... 10

1. PLANTEAMIENTO DEL PROBLEMA ..... 12

    1.1 Descripción del problema ..... 12

    1.2 La pregunta de investigación ..... 13

    1.3 Los objetivos de investigación ..... 14

        1.3.1 Objetivo general ..... 14

        1.3.2 Objetivos específicos..... 14

    1.4 Justificación de la investigación..... 15

2. MARCO DE REFERENCIA ..... 17

    2.1. Marco de Antecedentes ..... 17

    2.2. Marco Teórico..... 20

    2.3. Marco normativo ..... 21

3. METODOLOGÍA ..... 23

    3.1. Enfoque y alcance de la investigación ..... 24

    3.2. Población y muestra..... 25

        3.2.1. Definición de la población..... 25

        3.2.2. Cálculo y selección de la muestra ..... 25

        3.2.3. Instrumento(s)..... 27

        3.2.4. Descripción de procedimientos..... 28

        3.2.5. Análisis de información ..... 31

        3.2.6. Consideraciones éticas..... 32

        3.2.7. Análisis de consideraciones éticas..... 33

4. HIPÓTESIS ..... 35

    4.1. Las variables..... 35

        4.1.1. Variable(s) independiente(s) ..... 35

        4.1.2. Variable(s) dependiente(s) ..... 35

Propuesta de implementación de inteligencia artificial y Big Data

- 4.2. Planteamiento de hipótesis..... 36
  - Hipótesis general ..... 36
  - Hipótesis específicas ..... 36
- 5. RESULTADOS ..... 36
  - 5.1. Presentación de resultados..... 39
  - 5.2. Propuesta al sector ..... 45
  - 5.3. Discusión ..... 47
- 6. CONCLUSIONES..... 49
- REFERENCIAS ..... 50
- Anexos ..... 53

**Lista de tablas**

	<b>Pag.</b>
Tabla 1. Organizaciones de Colegios públicos y privados encuestados en Engativá Colombia.....	38
Tabla 2. Nivel de cumplimiento de criterios asociados a hábitos tecnológicos y de salud infantil.....	45

**Lista de figuras**

	<b>Pag.</b>
Figura 1 Distribución de sexo por cantidad de niños donde el género masculino es de mayor representación.....	40
Figura 2 Mapa de calor de correlación entre variables.....	41
Figura 3. Árbol de decisión: Clasificación Riesgo de Obesidad .....	42
Figura 4. Agrupamiento de Datos mediante una simulación usando método Clúster .....	43

**ANEXOS**

Anexo 1. Instrumento de Recolección de datos (Encuesta)..... 53  
Anexo 2. Formato de Conocimiento Informado para Padres o Acudientes..... 58

## Resumen

La presente investigación tuvo como propósito analizar, mediante herramientas de inteligencia artificial y big data, los hábitos alimenticios, tecnológicos y de actividad física en niños de 6 a 10 años de edad pertenecientes a instituciones educativas de la localidad de Engativá, en Bogotá. El objetivo principal fue identificar los factores con mayor influencia en la prevalencia del sobrepeso y la obesidad infantil, problemáticas de creciente impacto en contextos urbanos y escolares. La metodología adoptada fue de tipo cuantitativo, con enfoque descriptivo, aplicando una encuesta estructurada a una muestra representativa de estudiantes, cuyos datos fueron analizados con apoyo de herramientas informáticas.

Los resultados reflejan patrones preocupantes: más del 55% de los niños consume alimentos ultraprocesados entre tres y más veces por semana, el 45% permanece frente a pantallas por más de tres horas al día, y solo un 23% realiza actividad física por más de tres horas diarias. Estas conductas reflejan un entorno de vida sedentario, asociado a factores familiares, escolares y sociales, que limita el desarrollo saludable de la infancia. Además, los hallazgos son coherentes con investigaciones previas realizadas por entidades como UNICEF y estudios científicos en América Latina, que advierten sobre el rápido aumento de la obesidad infantil y los efectos del sedentarismo digital.

El análisis permitió confirmar la pertinencia de aplicar tecnologías emergentes en el estudio de comportamientos poblacionales, ofreciendo mayor precisión en el diagnóstico. Se concluye que es fundamental desarrollar programas de intervención multisectorial, que integren a las familias, las escuelas, el sistema de salud y las tecnologías de información para la prevención de enfermedades crónicas desde edades tempranas. Esta investigación constituye un insumo estratégico para futuras acciones en políticas públicas, educación nutricional y transformación del entorno escolar hacia estilos de vida saludables.

*Palabras clave: Salud infantil, hábitos alimenticios, sedentarismo, inteligencia artificial, Engativá.*

# Propuesta de implementación de inteligencia artificial y Big Data

## Abstract

This research aimed to analyze, through artificial intelligence and big data tools, the dietary, technological, and physical activity habits of children aged 6 to 10 attending educational institutions in the locality of Engativá, Bogotá. The main objective was to identify the key factors contributing to the prevalence of childhood overweight and obesity, both of which have become increasingly prevalent in urban and school contexts. A quantitative, descriptive methodology was employed, using a structured survey administered to a representative sample of schoolchildren, with data analyzed through digital tools and statistical techniques.

The results reveal alarming patterns: over 55% of children consume ultra-processed foods three or more times per week; 45% spend more than three hours daily using screens; and only 23% engage in physical activity for more than three hours per day. These findings indicate a predominantly sedentary lifestyle, influenced by family routines, school practices, and socio-technological environments, which poses significant risks to children's health and development. Furthermore, the results are consistent with reports from UNICEF and other regional studies that highlight the rapid rise of childhood obesity and the negative impact of excessive screen time.

The study underscores the relevance of applying emerging technologies such as artificial intelligence and data analytics to understand and address complex social and health challenges. It concludes that there is an urgent need to implement comprehensive intervention strategies involving families, schools, public health systems, and digital platforms to foster healthy lifestyles from early childhood. This research contributes valuable evidence for decision-makers and educators seeking to develop effective public policies, nutritional education programs, and active school environments that promote child well-being in a sustainable and informed manner.

*Keywords: Child health, eating habits, sedentary lifestyle, artificial intelligence, Engativ*

## Introducción

La obesidad infantil constituye una de las problemáticas más críticas y silenciosas en materia de salud pública a nivel global. De acuerdo con la Organización Mundial de la Salud (OMS), en 2022 más de 39 millones de niños menores de cinco años presentaban sobrepeso, evidenciando un crecimiento sostenido de esta condición en las últimas décadas, especialmente en entornos urbanos donde los estilos de vida han cambiado drásticamente (Organización Mundial de la Salud [OMS], 2022).

En Colombia, el problema se agudiza en ciudades como Bogotá, donde factores como el fácil acceso a alimentos ultraprocesados, el sedentarismo y el uso excesivo de tecnologías digitales han configurado un ambiente obesogénico, que incrementa el riesgo de enfermedades crónicas desde la infancia (Ministerio de Salud y Protección Social, 2021). La localidad de Engativá, una de las zonas más densamente pobladas de la ciudad, no es ajena a esta situación. Diversos estudios han evidenciado que los niños entre los 6 y 10 años presentan hábitos alimenticios poco saludables, alta exposición a pantallas y bajos niveles de actividad física (Instituto Colombiano de Bienestar Familiar [ICBF], 2020).

Frente a este panorama, la presente propuesta plantea el diseño de un modelo analítico que, mediante el uso de herramientas de **inteligencia artificial (IA)** y **Big Data**, permita analizar los hábitos alimenticios, tecnológicos y de actividad física en niños de 6 a 10 años en Engativá. Esta iniciativa tiene como propósito identificar los factores con mayor incidencia en la aparición de sobrepeso y obesidad infantil, y ofrecer una base metodológica para futuras intervenciones de prevención y promoción de la salud.

El enfoque de esta propuesta va más allá de la simple descripción del problema; busca estructurar un esquema metodológico replicable que, en su futura aplicación, posibilite la detección temprana de riesgos y la formulación de estrategias preventivas, basadas en el análisis masivo de datos y el uso de tecnologías emergentes.

Este documento se estructura en seis capítulos que permiten un análisis progresivo y sistemático de la problemática y su abordaje propositivo:

## Propuesta de implementación de inteligencia artificial y Big Data

- **Capítulo I: Planteamiento del Problema:** contextualiza la situación actual de la localidad de Engativá, identificando los principales factores de riesgo asociados a la obesidad infantil y justificando la pertinencia de la propuesta.
- **Capítulo II: Marco de Referencia:** presenta el sustento teórico, normativo y conceptual del proyecto, incluyendo antecedentes relevantes y la definición de conceptos clave como obesidad infantil, sedentarismo, Big Data e inteligencia artificial.
- **Capítulo III: Metodología Propuesta:** describe el enfoque metodológico sugerido para la futura implementación del proyecto, incluyendo el diseño de los instrumentos, las técnicas de recolección de datos y las herramientas tecnológicas recomendadas.
- **Capítulo IV: Hipótesis y Variables:** formula una hipótesis preliminar fundamentada en la literatura revisada y presenta la matriz de variables clave para el análisis propuesto.
- **Capítulo V: Resultados Esperados:** expone los resultados potenciales que se espera obtener con la implementación del proyecto, como la identificación de patrones de riesgo y la creación de reportes para la toma de decisiones.
- **Capítulo VI: Conclusiones y Recomendaciones:** sintetiza las proyecciones y beneficios de la propuesta, destacando el potencial de la inteligencia artificial y el Big Data como herramientas de análisis predictivo y preventivo.

En suma, esta propuesta tiene como objetivo servir como hoja de ruta para el desarrollo de proyectos futuros que integren tecnología y salud pública, promoviendo entornos escolares y familiares más saludables y resilientes.

## 1. PLANTEAMIENTO DEL PROBLEMA

### 1.1 Descripción del problema

El sedentarismo infantil es una preocupación creciente a nivel mundial debido a sus efectos negativos sobre la salud física y mental de los menores. Según la Organización Mundial de la Salud (OMS, 2020), más del 80% de los niños y adolescentes en edad escolar no cumplen con el mínimo recomendado de 60 minutos diarios de actividad física, lo que incrementa su vulnerabilidad frente a enfermedades crónicas como obesidad, diabetes tipo 2 y afecciones cardiovasculares.

Esta problemática se ha intensificado en las últimas décadas con el creciente acceso y uso de dispositivos tecnológicos, los cuales han transformado los hábitos cotidianos de los menores, reduciendo el tiempo de juego activo y promoviendo comportamientos sedentarios desde edades tempranas. Diversos estudios internacionales han demostrado la relación entre el tiempo frente a pantallas y el aumento del sedentarismo y sobrepeso en la población infantil (Swinburn et al., 2019).

En América Latina, esta tendencia resulta aún más alarmante. Factores socioeconómicos como la inseguridad alimentaria, la falta de acceso a espacios recreativos seguros y la creciente urbanización limitan las oportunidades para la actividad física, al tiempo que incrementan el consumo de alimentos ultraprocesados. La Organización Panamericana de la Salud (OPS, 2021) advierte que la región enfrenta una doble carga de malnutrición: por un lado, la persistencia de la desnutrición infantil y, por otro, el aumento significativo del sobrepeso y la obesidad, situación que afecta particularmente a los sectores urbanos y de bajos ingresos.

En el contexto colombiano, esta problemática refleja una crisis que combina múltiples factores estructurales. Según el Ministerio de Salud y Protección Social (2022), el 24% de los niños entre 5 y 12 años presentan sobrepeso u obesidad, mientras que más del 70% superan las dos horas diarias recomendadas de exposición a pantallas. Estos datos evidencian no solo un patrón nacional preocupante, sino también la necesidad urgente de intervenciones específicas.

Particularmente en Bogotá, las cifras son más críticas en localidades con alta densidad poblacional como Engativá. De acuerdo con la Secretaría Distrital de Salud (2022), esta zona presenta una alta prevalencia de problemas nutricionales infantiles, donde el acceso limitado a espacios verdes, la precariedad en infraestructura educativa para el deporte y la falta de programas nutricionales sostenibles agravan la situación. Además, la Encuesta Nacional de Situación Nutricional (ENSIN, 2020) reportó que

## Propuesta de implementación de inteligencia artificial y Big Data

Bogotá tiene una de las tasas más elevadas de obesidad infantil en Colombia, destacándose Engativá entre las localidades más afectadas.

Un factor adicional que profundiza esta problemática es la limitada integración de tecnologías emergentes en las estrategias de monitoreo y prevención de la obesidad infantil. A pesar del avance global en el uso de herramientas como la inteligencia artificial (IA) y el Big Data para la vigilancia epidemiológica y la promoción de la salud, en el contexto local estas tecnologías no han sido incorporadas de forma sistemática en las investigaciones relacionadas con la salud y el comportamiento infantil (Córdoba et al., 2021).

La ausencia de estudios que articulen enfoques tecnológicos con diagnósticos sociales y de salud infantil evidencia una brecha significativa en el abordaje de esta problemática. Esta situación plantea la necesidad de desarrollar propuestas innovadoras, que integren el análisis de datos masivos con el estudio de hábitos alimenticios y tecnológicos de la niñez, para diseñar soluciones contextualizadas y efectivas.

Por tanto, esta propuesta plantea el diseño de un modelo metodológico basado en inteligencia artificial y Big Data, enfocado en la localidad de Engativá, que permita identificar factores de riesgo asociados al sedentarismo infantil y ofrecer alternativas preventivas adaptadas al contexto urbano local. La implementación futura de este modelo no solo aportaría a la mejora de la salud infantil, sino que también permitiría optimizar recursos públicos y reducir los costos asociados a enfermedades prevenibles, promoviendo una infancia más saludable, consciente y protegida.

### **1.2 La pregunta de investigación**

¿Cuál sería el factor con mayor incidencia en la prevalencia de sobrepeso y obesidad infantil, asociado a los hábitos alimenticios y tecnológicos, en niños de 6 a 10 años de la localidad de Engativá (Bogotá), a partir del análisis mediante inteligencia artificial y Big Data?

### **1.3 Los objetivos de investigación**

#### **1.3.1 Objetivo general**

Diseñar una propuesta metodológica que, mediante el uso de herramientas de inteligencia artificial y Big Data, permita analizar los hábitos alimenticios y tecnológicos en niños de 6 a 10 años de la localidad de Engativá (Bogotá), con el propósito de identificar el factor con mayor influencia en la prevalencia de sobrepeso y obesidad infantil.

#### **1.3.2 Objetivos específicos**

Implementar un esquema para la recopilación y estructuración de datos relevantes sobre consumo alimentario, uso de tecnologías y actividad física en niños de 6 a 10 años en la localidad de Engativá, mediante encuestas, fuentes secundarias y registros disponibles.

Proponer la aplicación de modelos de inteligencia artificial y técnicas de análisis de Big Data que permitan identificar patrones, correlaciones y tendencias entre los hábitos alimenticios, tecnológicos y el estado nutricional de los menores.

Establecer un modelo analítico que permita determinar el factor predominante asociado al sobrepeso y la obesidad infantil, como base para generar recomendaciones preventivas y contextualizadas dirigidas a la comunidad educativa, las familias y las entidades de salud locales.

### 1.4 Justificación de la investigación

El presente proyecto propone el diseño de una metodología para analizar la influencia de los factores alimenticios y tecnológicos en niños sedentarios de 6 a 10 años en la localidad de Engativá, Bogotá. Esta iniciativa es pertinente y necesaria ante la creciente preocupación por la salud infantil, especialmente en zonas urbanas donde los entornos obesogénicos y los hábitos sedentarios son cada vez más comunes.

Desde el orden académico, la obesidad infantil ha sido ampliamente estudiada como una problemática multidimensional que exige un abordaje interdisciplinario. La literatura científica coincide en que la combinación de alimentación inadecuada, sedentarismo y uso excesivo de tecnologías constituye un factor de riesgo clave para el desarrollo de sobrepeso infantil (Swinburn et al., 2019). Sin embargo, existe una evidente carencia de estudios que integren el análisis de estos factores mediante enfoques tecnológicos avanzados, como la inteligencia artificial (IA) y el Big Data, especialmente en el contexto latinoamericano. Por tanto, esta propuesta busca contribuir al avance académico mediante la formulación de un modelo metodológico innovador, que articule elementos tecnológicos con el estudio del comportamiento infantil, abriendo nuevas líneas de investigación en salud pública, educación y tecnologías de la información.

En el ámbito científico, la aplicación de IA y Big Data en salud ha demostrado su eficacia para analizar grandes volúmenes de datos y predecir tendencias epidemiológicas, optimizando la toma de decisiones en intervenciones preventivas. Estas herramientas permiten identificar patrones complejos que no serían detectables mediante análisis estadísticos tradicionales (Córdoba et al., 2021). Su uso en el estudio del sobrepeso infantil resulta especialmente prometedor, ya que permite correlacionar múltiples variables, como el consumo de alimentos ultraprocesados, el tiempo frente a pantallas, la actividad física y las condiciones socioeconómicas. Este enfoque ofrece la posibilidad de anticipar riesgos y diseñar estrategias de prevención más personalizadas y eficientes, basadas en evidencia cuantitativa robusta.

Desde el orden social, la magnitud del problema del sedentarismo infantil en Bogotá evidencia la urgencia de soluciones innovadoras. Según la OMS (2020), la obesidad infantil afecta a más del 40% de la población infantil en varios países, mientras que estudios locales indican que al menos el 30% de los niños entre 6 y 10 años en Bogotá presentan sobrepeso u obesidad (Camargo et al., 2018). Esta situación se agrava en localidades como Engativá, donde las dinámicas urbanas limitan el acceso a espacios verdes y la infraestructura para la actividad física, mientras que la exposición a pantallas y el consumo de alimentos poco saludables aumentan de forma sostenida (Secretaría Distrital de Salud, 2022).

## Propuesta de implementación de inteligencia artificial y Big Data

Además, Colombia ha avanzado en el desarrollo de políticas públicas que fomentan el uso responsable de la inteligencia artificial, como lo demuestra el Documento CONPES 4144, que establece la Política Nacional de Inteligencia Artificial (Departamento Nacional de Planeación [DNP], 2025). Estos avances generan un contexto favorable para la aplicación de tecnologías emergentes en problemáticas sociales, ofreciendo una oportunidad estratégica para mejorar la salud infantil mediante el análisis predictivo y el monitoreo basado en datos.

Por tanto, esta propuesta metodológica no solo resulta pertinente en términos científicos y tecnológicos, sino también en el plano social, ya que busca generar un modelo replicable para el estudio del sobrepeso infantil en contextos urbanos similares. Su implementación futura podría contribuir a la formulación de políticas públicas más efectivas y a la promoción de entornos escolares y comunitarios más saludables.

## 2. MARCO DE REFERENCIA

### 2.1. Marco de Antecedentes

Para la construcción del marco de antecedentes se realizó una revisión bibliográfica de literatura científica publicada en los últimos diez años (2015–2025), con el objetivo de garantizar la actualidad y relevancia de los estudios analizados. La búsqueda se llevó a cabo en bases de datos académicas como Scopus, ScienceDirect, PubMed, Google Scholar, Redalyc y SciELO, aplicando filtros por año, tipo de documento (artículos científicos y tesis académicas) y enfoque metodológico (cuantitativo, cualitativo o mixto). Para optimizar los resultados se emplearon palabras clave relacionadas con obesidad infantil, hábitos alimenticios, sedentarismo en la infancia y aplicación de inteligencia artificial y Big Data en salud pública infantil, combinadas mediante operadores booleanos (AND, OR). La ecuación de búsqueda utilizada fue: (“obesidad infantil” OR “sobrepeso infantil”) AND (“hábitos alimenticios” OR “alimentación” OR “dieta”) AND (“sedentarismo” OR “actividad física”) AND (“inteligencia artificial” OR “IA” OR “Big Data”).

***Machine learning to predict child overweight risk (Dugan et al., 2021)*** Este artículo evalúa la eficacia de modelos de aprendizaje automático para predecir el riesgo de sobrepeso infantil. Utilizando bases de datos clínicas y escolares en EE. UU., se encontró que algoritmos como random forest y XGBoost tienen una precisión del 85 % al detectar factores clave como el IMC parental, el tipo de alimentación, y el tiempo de pantalla. La investigación sugiere que estas herramientas pueden implementarse para crear sistemas de alerta temprana en escuelas. Su enfoque refuerza la aplicación de IA en contextos como el de Engativá

***Comparative Study of Lifestyle: Eating Habits, Sedentary Lifestyle and Anthropometric Development in Spanish 5- To 15-yr-Olds (Morales-Suárez-Varela et al. ,2015)***

Este estudio comparativo, con datos de la Encuesta Nacional de Salud de España, examinó hábitos alimenticios y sedentarismo en niños de 5 a 15 años entre 2006 y 2013. Encontró una prevalencia de sobrepeso/obesidad del 24,3 % en 2013, con una reducción importante en consumo de carne y vegetales, y un incremento en conductas sedentarias vinculadas al uso de tecnologías (TV y videojuegos). El artículo destaca que estos cambios pueden afectar negativamente la calidad de vida futura y refuerza la necesidad de políticas enfocadas en dieta y actividad deportiva.

***Dietary and physical activity/inactivity factors associated with obesity in school-aged children: A cross-sectional study. Journal of Human Nutrition and Dietetics (Perez-Rodríguez et al., 2016)***

Este estudio transversal en Europa evaluó dietas y actividad en 143 niños de 8–10 años. Identificó una correlación clara entre bajo consumo de lácteos, alto consumo de bebidas azucaradas, menor actividad física y mayor prevalencia de obesidad. Utilizó cuestionarios y registros de actividad para comparar grupos obesos con eutróficos, lo que confirma la relación entre alimentación, sedentarismo y estado nutricional.

***Food habits, physical activities and sedentary lifestyles of eutrophic and obese school children: A case-control study (Vilchis-Gil et al., 2016)***

Con 202 niños obesos y 200 eutróficos en Ciudad de México, los investigadores emplearon METs y tiempos de sedentarismo para evaluar la actividad física y hábitos alimentarios. Encontraron que los niños obesos tenían mayor ingesta energética y pasaban más horas en conductas sedentarias. El estudio subraya la importancia de validar los hábitos con mediciones objetivas para diseñar intervenciones nutricionales y de actividad.

***Physical activity change during adolescence: A systematic review and a pooled análisis (Pelotas, Brasil, 2016)***

Esta cohorte incluyó a 616 niños de 8 años y analizó su ingesta de cereales/raíces y patrones de actividad. Descubrió que casi ninguno alcanzaba recomendaciones nutricionales, y la gran mayoría era físicamente inactiva. Este trabajo destaca la necesidad de estrategias multisectoriales para corregir malos hábitos desde las primeras etapas de la vida.

***Obesity and its associated factors among Jordanian schoolchildren: A cross-sectional study (Jordanian, 2019)***

Estudio con 977 escolares de 7–18 años usando criterios IOTF. Encontró que el sedentarismo duplicaba el riesgo de sobrepeso y que menos de 30 minutos/día de ejercicio físico aumentaba esta condición. El consumo frecuente de snacks fuera del hogar también fue identificado como factor de riesgo independiente, sugiriendo intervenciones en estilos de vida urbano.

***Predicting overweight and obesity in later life from childhood data: A review of predictive modeling approaches (Rautiainen & Äyrämö, 2019)***

## Propuesta de implementación de inteligencia artificial y Big Data

Revisión de modelos predictivos usando Machine Learning aplicados a datos de infancia. Encuentra que los mejores modelos combinan datos antropométricos e información materna, pero advierten la necesidad de incorporar variables sociales y psicológicas en futuros trabajos.

### ***Obesity Prediction with EHR Data: A Deep Learning Approach with Interpretable Elements. arXiv preprint (Gupta et al., 2019)***

Este estudio desarrolló un modelo LSTM que utiliza historiales médicos electrónicos (EHR) para predecir obesidad entre los 2 y 20 años. Su modelo mostró mejor precisión que métodos tradicionales, e incluyó capas de atención para permitir interpretabilidad, destacando la utilidad de IA para análisis longitudinal en salud infantil.

### **Identifying Key Determinants (LeCroy et al., 2021)**

Revisión narrativa de 15 estudios de ML aplicado a obesidad infantil en distintas edades. Los hallazgos principales fueron que el peso temprano, el estado materno, la inactividad física y predisposición étnica son claves para la predicción. Recomienda integrar factores socioculturales en los modelos predictivos

### ***Childhood obesity risk prediction based on machine learning: A large-scale cohort study in China (Chinese preschool-aged, 2022)***

Estudio en China (3–6 años), con 9 478 participantes. Emplearon SVM y GBM con precisión de 0.945 y F1 ~0.77. Los principales factores fueron edad, velocidad al comer, historia familiar, consumo de bebidas azucaradas y educación paterna. Confirma que algoritmos de IA pueden identificar variables clave en riesgo de obesidad

### ***A prediction model for childhood obesity risk using the machine learning method based on data from a panel study in Korea (Oh et al., 2023)***

Modelo predictivo con panel coreano (1 185 niños), usando LASSO. Identificó que menos actividad física infantil y mayor autoestima materna están fuertemente asociados con obesidad. Precisión del modelo fue AUC 0.82, respaldando la inclusión de factores familiares psicosociales.

*Adolescent obesity prediction system based on a deep learning framework (Jeong et al., 2023)*

“DeepHealthNet” usa deep learning en antecedentes de 321 adolescentes. Logra precisión de 88 %, diferenciando por género y entregando retroalimentación personalizada. Muestra el potencial de IA para generar alertas personalizadas seguidas de intervención educativa

## 2.2. Marco Teórico

**Obesidad Infantil:** La obesidad infantil es una condición médica caracterizada por un exceso de grasa corporal que afecta la salud y el desarrollo del niño. Se define comúnmente mediante el índice de masa corporal (IMC) ajustado por edad y sexo. Esta condición es resultado de múltiples factores: genéticos, ambientales, sociales y conductuales, siendo la alimentación desequilibrada y la inactividad física los más destacados. Según la OMS (2024), más de 39 millones de niños menores de cinco años presentaban sobrepeso en el mundo, y se proyecta un aumento sostenido si no se toman medidas preventivas. La obesidad infantil incrementa el riesgo de enfermedades cardiovasculares, diabetes tipo 2 y trastornos psicológicos.

**Hábitos Alimentarios en la Infancia:** Los hábitos alimentarios se forman en la primera infancia y están influenciados por el entorno familiar, escolar, social y mediático. Incluyen patrones de consumo, frecuencia de ingesta y tipo de alimentos preferidos. Cuando estos hábitos se consolidan en entornos desfavorables (acceso limitado a alimentos frescos, alta exposición a ultra procesados), se incrementa el riesgo de obesidad. Hernández y Moreno (2018) señalan que los niños que consumen productos ultra procesados más de cuatro veces por semana tienen una probabilidad 2.8 veces mayor de desarrollar sobrepeso en la adolescencia.

**Sedentarismo Infantil y Tecnología:** El sedentarismo infantil está directamente relacionado con el uso prolongado de tecnologías como tablets, celulares, videojuegos y televisión. El tiempo excesivo frente a pantallas reduce el nivel de actividad física y está asociado a una mayor ingesta de alimentos poco saludables. De acuerdo con Graf y Cecchini (2018), más del 60 % de los niños en América Latina pasan más de tres horas al día frente a pantallas, lo que ha contribuido al aumento de obesidad infantil en áreas urbanas.

**Inteligencia Artificial Aplicada a la Salud Infantil:** La inteligencia artificial (IA) permite analizar grandes volúmenes de datos para predecir y entender patrones complejos de salud. En el ámbito de la obesidad infantil, puede identificar factores de riesgo, personalizar intervenciones y construir

## Propuesta de implementación de inteligencia artificial y Big Data

modelos predictivos útiles para tomar decisiones tempranas. Según Zhou et al. (2022), algoritmos como Random Forest, Support Vector Machines y redes neuronales han alcanzado una precisión superior al 85 % para predecir obesidad infantil a partir de variables conductuales y sociodemográficas.

**Big Data en el Análisis de Comportamientos de Salud:** El big data se refiere al uso de grandes volúmenes de información estructurada y no estructurada para obtener conocimientos útiles. En salud pública, su análisis facilita la vigilancia epidemiológica, la segmentación poblacional y la formulación de políticas basadas en evidencia. Brakefield et al. (2022) proponen plataformas como UPHO (Urban Population Health Observatory), que utilizan big data para identificar variables causales de la obesidad en contextos urbanos mediante IA explicable.

### 2.3. Marco normativo

El marco normativo del presente proyecto se sustenta en el conjunto de leyes, políticas públicas y normas técnicas que rigen los derechos de la infancia, la protección de la salud, la promoción de la alimentación adecuada y el uso responsable de las tecnologías de información en Colombia. Estas disposiciones ofrecen el respaldo jurídico necesario para la ejecución del estudio, garantizando el respeto a los derechos de los menores y la protección de sus datos personales.

#### **Constitución Política de Colombia (1991)**

La Constitución establece que los derechos de los niños tienen prioridad sobre los derechos de los demás ciudadanos. Su artículo 44 consagra la obligación del Estado, la familia y la sociedad de velar por la salud, la nutrición adecuada, la educación y la recreación de la infancia, todos factores relacionados con la problemática del sobrepeso y la obesidad infantil (Congreso de la República de Colombia, 1991).

#### **Ley 1098 de 2006 – Código de Infancia y Adolescencia**

Esta ley reconoce el derecho de los niños y adolescentes a una alimentación saludable y a un desarrollo integral. En su artículo 41, se promueve la implementación de programas de prevención de enfermedades no transmisibles, como la obesidad, y se obliga a las instituciones educativas a fomentar estilos de vida saludables (Congreso de la República de Colombia, 2006).

**Ley 1355 de 2009 – Por la cual se definen políticas para prevenir la obesidad y enfermedades crónicas no transmisibles**

## Propuesta de implementación de inteligencia artificial y Big Data

Primera norma en Colombia que reconoce la obesidad como un problema de salud pública, establece acciones intersectoriales para su prevención, incluyendo la regulación de la publicidad dirigida a menores y la promoción de la actividad física. Esta ley respalda proyectos que, como el presente, proponen el uso de herramientas tecnológicas para el análisis de datos relacionados con la salud infantil (Congreso de la República de Colombia, 2009).

### **Ley 1804 de 2016 – Ley de Cero a Siempre**

Aunque enfocada en la primera infancia, esta ley promueve el monitoreo y evaluación permanente del desarrollo infantil mediante el uso de bases de datos, lo cual guarda coherencia con el uso de Big Data e inteligencia artificial planteado en este proyecto (Congreso de la República de Colombia, 2016).

### **Resolución 3803 de 2016 del Ministerio de Salud y Protección Social**

Establece lineamientos para la promoción de hábitos saludables en entornos escolares, enfatizando la importancia del consumo de frutas y verduras, la actividad física diaria y la limitación del tiempo frente a pantallas. Esta resolución respalda las estrategias preventivas dirigidas a la infancia en entornos escolares y comunitarios (Ministerio de Salud y Protección Social, 2016).

### **Ley 1581 de 2012 – Ley de Protección de Datos Personales**

Garantiza el tratamiento ético y legal de los datos personales, especialmente sensibles como los datos de menores. Establece principios de seguridad, confidencialidad y transparencia, fundamentales para proyectos que integran inteligencia artificial y Big Data en el análisis de información (Congreso de la República de Colombia, 2012).

### **Ley 1978 de 2019 – Ley de Modernización TIC**

Regula el acceso y uso responsable de tecnologías digitales, resaltando la necesidad de proteger a los menores de los riesgos asociados a los entornos digitales y fomentando la educación crítica en el uso de dispositivos electrónicos (Congreso de la República de Colombia, 2019).

## Propuesta de implementación de inteligencia artificial y Big Data

### **Política Nacional de Seguridad Alimentaria y Nutricional (CONPES 113 de 2008)**

Promueve el acceso y consumo de alimentos nutritivos, además del monitoreo nutricional de la población. Esta política impulsa el desarrollo de plataformas de información basadas en tecnología, facilitando el análisis de datos para la toma de decisiones en salud pública (Departamento Nacional de Planeación, 2008).

### **Ley 1266 de 2008 – Ley de Habeas Data**

Establece los derechos de los ciudadanos sobre sus datos personales y regula el tratamiento de la información, incluyendo proyectos que utilizan grandes volúmenes de datos. Su aplicación es fundamental en investigaciones con componentes de IA y Big Data (Congreso de la República de Colombia, 2008).

### **Documento CONPES 3975 de 2019 – Política Nacional para la Transformación Digital e Inteligencia Artificial**

Define los lineamientos éticos y estratégicos para la adopción de IA en Colombia, con énfasis en el respeto de los derechos humanos, la privacidad y la equidad. Este documento legitima el uso de la inteligencia artificial para mejorar la gestión en sectores como la salud y la educación (Departamento Nacional de Planeación, 2019).

### **Directrices de la UNESCO sobre la Ética de la Inteligencia Artificial (2021)**

Colombia adoptó estas directrices que promueven el desarrollo de sistemas de IA respetuosos de los derechos humanos, con especial atención a la protección de la infancia y la eliminación de sesgos discriminatorios en el tratamiento automatizado de datos (UNESCO, 2021).

## **3. METODOLOGÍA**

Esta propuesta metodológica plantea un estudio con enfoque cuantitativo, orientado a recolectar y analizar datos objetivos sobre hábitos alimenticios, uso de tecnologías y niveles de actividad física en niños de 6 a 10 años de la localidad de Engativá (Bogotá). Su alcance será correlacional y exploratorio, buscando identificar relaciones entre variables sin manipulación directa y explorar el uso de inteligencia artificial (IA) y Big Data en el análisis predictivo.

## Propuesta de implementación de inteligencia artificial y Big Data

Se propone trabajar con una muestra no probabilística por conveniencia, seleccionando 20 instituciones educativas entre públicas y privadas dispuestas a participar, con previa autorización institucional y consentimiento informado de padres o acudientes.

Como técnica de recolección de datos, se plantea la aplicación de encuestas estructuradas a padres, acudientes y docentes, incluyendo aspectos como:

- Frecuencia de consumo alimentario.
- Tiempo frente a pantallas.
- Actividad física.

De forma complementaria, se propone integrar registros escolares o clínicos (peso, talla e IMC), cumpliendo con la normativa vigente de protección de datos personales.

Para el análisis, se contempla la aplicación de estadística descriptiva y correlacional (JASP y Power BI), seguida de modelos predictivos de machine learning (Random Forest o regresión logística), con el fin de identificar los factores de mayor influencia en la obesidad infantil.

El estudio propuesto se desarrollaría bajo criterios éticos, asegurando el consentimiento informado, la confidencialidad y la anonimización de la información.

### **3.1. Enfoque y alcance de la investigación**

El presente proyecto se plantea bajo un enfoque mixto, combinando técnicas cuantitativas y cualitativas. Según Hernández, Fernández y Baptista (2016), el enfoque cuantitativo se caracteriza por “recoger y analizar datos empíricos cuantificables sobre variables observables y mensurables”, mientras que el enfoque cualitativo permite el análisis documental y contextual.

Por un lado, se propone recolectar datos objetivos mediante encuestas estructuradas que indaguen sobre hábitos alimenticios, niveles de sedentarismo, uso de tecnologías digitales y condiciones físicas (peso, talla, IMC) en niños de 6 a 10 años en la localidad de Engativá. Estos datos serán analizados mediante métodos estadísticos y modelos de inteligencia artificial, para identificar correlaciones y patrones predictivos.

## Propuesta de implementación de inteligencia artificial y Big Data

Por otro lado, se plantea la revisión documental de registros escolares o clínicos disponibles, así como de informes institucionales o bases de datos oficiales, lo cual constituye un componente cualitativo complementario que enriquece el análisis.

El alcance propuesto es correlacional y exploratorio. Se pretende establecer relaciones entre las variables estudiadas y generar información útil para el diseño de estrategias preventivas contextualizadas frente al sobrepeso y la obesidad infantil.

Este enfoque mixto permite un análisis integral, riguroso y replicable, orientado a la toma de decisiones informadas, integrando datos numéricos y elementos documentales clave..

### **3.2. Población y muestra**

#### **3.2.1. Definición de la población**

La población del presente estudio está constituida por niños y niñas entre los 6 y 10 años de edad que residen en la localidad de Engativá, Bogotá D.C., y que asisten a instituciones educativas de carácter oficial y privado. Esta población es de especial interés, ya que corresponde a un grupo etario vulnerable al desarrollo de obesidad infantil como consecuencia de factores conductuales y ambientales, tales como los hábitos alimenticios, el uso excesivo de tecnología y el sedentarismo.

#### **3.2.2. Cálculo y selección de la muestra**

La muestra será de tipo no probabilística por conveniencia, seleccionada a partir de instituciones educativas que voluntariamente accedan a participar en el estudio y cuenten con el consentimiento informado de los padres o acudientes de los menores. Se estima trabajar con al menos tres instituciones y una muestra aproximada de 60 a 100 niños, lo que permitirá aplicar instrumentos de medición y alimentar modelos predictivos sin necesidad de generalizar los resultados a toda la población, pero sí identificar patrones representativos y relevantes.

Para establecer un tamaño mínimo de muestra, se utiliza el siguiente cálculo estadístico, común en estudios de enfoque cuantitativo:

## Propuesta de implementación de inteligencia artificial y Big Data

$$n = \frac{Z^2 \cdot p \cdot q}{e^2}$$

n = tamaño de la muestra

Z = valor Z correspondiente al nivel de confianza (1.96 para 95 %)

p = proporción esperada (si no se conoce, se usa 0.5 por ser el caso más conservador)

q = 1 - p

e = margen de error permitido (por ejemplo, 0.05 = 5 %, 0.10 = 10 %)

Hernández, R., Fernández, C., & Baptista, P. (2016). **Metodología de la investigación** (6.ª ed.). McGraw-Hill.

Entonces,

Z = 1.96 (nivel de confianza del 95 %)

p = 0.5 (probabilidad de éxito)

q = 0.5 (probabilidad de fracaso)

e = 0.10 (margen de error aceptado: 10 %)

$$n = \frac{(1.96)^2 \cdot 0.5 \cdot 0.5}{(0.10)^2} = \frac{3.8416 \cdot 0.25}{0.01} = \frac{0.9604}{0.01} = 96.04$$

Por tanto, el tamaño estimado de la muestra será de aproximadamente 96 niños. Este número puede ajustarse levemente según el acceso real a las instituciones participantes.

- **Criterios de inclusión:**

- Niños con edades entre 6 y 10 años.
- Estudiantes activos de instituciones educativas ubicadas en la localidad de Engativá (Bogotá).
- Padres o acudientes que firmen el consentimiento informado.
- Niños con disponibilidad para participar en encuestas o mediciones no invasivas.

- **Criterios de exclusión:**

- Niños fuera del rango de edad establecido.
- Estudiantes con condiciones médicas o metabólicas crónicas que afecten su peso de forma no relacionada con hábitos alimenticios o actividad física.

## Propuesta de implementación de inteligencia artificial y Big Data

- Casos donde no se cuente con la autorización expresa de los padres o tutores legales.
- Niños en condición de movilidad académica o institucional reciente, que impida evaluar hábitos estables.

### 3.2.3. Instrumento(s)

Para el desarrollo de la presente propuesta metodológica se contemplan dos instrumentos principales de recolección de datos: encuestas estructuradas y análisis documental. Ambos permitirán obtener información relevante sobre los hábitos alimenticios, el uso de tecnología, la actividad física y el estado nutricional de niños entre 6 y 10 años en la localidad de Engativá.

#### **Encuesta estructurada**

Se propone aplicar una encuesta dirigida a los padres o acudientes de los niños seleccionados, con el fin de recolectar datos cuantitativos sobre las variables de estudio. Este instrumento está compuesto por preguntas cerradas de opción múltiple y escalas tipo Likert, distribuidas en cinco secciones temáticas:

- **Datos sociodemográficos:** edad, sexo del niño, tipo de institución educativa, número de integrantes en el hogar y estrato socioeconómico.
- **Hábitos alimenticios:** frecuencia de consumo de frutas, verduras, bebidas azucaradas, alimentos ultraprocesados y horarios de alimentación.
- **Actividad física:** tiempo promedio diario dedicado a la actividad física, tipo de ejercicios y participación en actividades deportivas.
- **Uso de tecnología y tiempo frente a pantallas:** cantidad de horas diarias frente a dispositivos electrónicos (televisión, celular, tablet, computador) y uso con fines escolares o recreativos.
- **Percepción del estado nutricional:** opinión de los padres o acudientes sobre el peso y estado nutricional del niño, apetito y posibles diagnósticos médicos relacionados.

La encuesta busca recolectar información estructurada y estandarizada que facilite el análisis estadístico y el entrenamiento de modelos de inteligencia artificial.

*Nota: El cuestionario completo, con la totalidad de sus preguntas, se encuentra incluido en el Anexo I de este documento.*

#### **Análisis documental**

## Propuesta de implementación de inteligencia artificial y Big Data

Como complemento a la encuesta, se plantea un análisis documental de registros institucionales, escolares y clínicos (previa autorización y cumpliendo la normativa vigente de protección de datos), con el propósito de validar y complementar los datos recolectados. Esta revisión documental incluirá:

- Registros de talla, peso y curvas de crecimiento recopilados por las instituciones educativas o brigadas de salud.
- Índice de masa corporal (IMC) y su respectiva interpretación.
- Reportes académicos o registros relacionados con la participación en actividades físicas escolares.
- Historial médico disponible que registre enfermedades metabólicas o diagnósticos nutricionales.

### **Criterios de revisión documental:**

- Palabras clave: obesidad infantil, hábitos alimenticios, sedentarismo, actividad física, IMC, sobrepeso infantil, niños escolares, salud infantil, tecnologías digitales.
- Bases de datos o fuentes: registros de las instituciones educativas, brigadas de salud locales, bases de datos oficiales de la Secretaría de Salud Distrital y documentos clínicos autorizados.
- Temporalidad: Se revisarán registros de los últimos cinco años previos al desarrollo del estudio, para asegurar datos actualizados y pertinentes.
- Criterios de inclusión: Documentos oficiales y registros válidos que contengan información detallada sobre el estado nutricional, hábitos y actividad física de los menores.
- Criterios de exclusión: Documentos incompletos, desactualizados (anteriores al rango temporal establecido) o que no cuenten con autorización para su uso.

Este análisis documental permitirá enriquecer la base de datos para su posterior análisis mediante técnicas de Big Data e inteligencia artificial, garantizando mayor solidez y profundidad en los resultados.

### **3.2.4. Descripción de procedimientos**

Para el desarrollo de esta investigación, se propone un proceso estructurado en cuatro fases principales, que abarcan desde la planeación hasta el análisis de resultados. Cada fase incluye actividades específicas, recursos necesarios y consideraciones éticas y logísticas que garantizan la calidad del proceso y el cumplimiento de los objetivos.

# Propuesta de implementación de inteligencia artificial y Big Data

## **Fase 1: Planeación y gestión institucional (1 semana)**

### **Actividades:**

- Solicitud formal de participación a instituciones educativas de la localidad de Engativá (públicas y privadas).
- Elaboración de formatos de consentimiento informado dirigidos a padres o acudientes, en los que se explica el propósito del estudio, el uso de la información y la garantía de confidencialidad.
- Revisión del instrumento (encuesta estructurada) por parte de un comité asesor o docente guía del proyecto.
- Preparación logística para la aplicación: impresión de formularios o adaptación a formato digital (Google Forms o similar).

### **Requiere:**

- Aprobación institucional (colegios).
- Firma de consentimiento informado.
- Revisión y aval ético por parte del tutor académico.

## **Fase 2: Aplicación de los instrumentos (2 semanas)**

### **Actividades:**

- Aplicación de la encuesta estructurada a padres o acudientes de los niños seleccionados.
  - Modalidad presencial o virtual, según lo permita cada institución.
  - Se destinará un tiempo estimado de 15 a 20 minutos por encuesta.
- En paralelo, se solicitará acceso a registros escolares o clínicos disponibles (peso, talla, IMC) bajo los criterios de confidencialidad y protección de datos personales.
- Capacitación breve a docentes u orientadores escolares sobre el diligenciamiento correcto del instrumento (si se realiza con su apoyo).

### **Lugar:**

- Las instituciones educativas autorizadas de la localidad de Engativá.

### **Tiempo estimado:**

- Dos semanas, incluyendo la recolección y organización de la base de datos.

## **Fase 3: Procesamiento y análisis de la información (2 semanas)**

### **Actividades:**

## Propuesta de implementación de inteligencia artificial y Big Data

- Codificación y organización de las encuestas en una base de datos digital.
- Validación cruzada con los datos documentales (IMC, talla, peso).
- Análisis estadístico descriptivo y correlacional con herramientas como **Excel, SPSS o Python**.
- Implementación de modelos de inteligencia artificial (como árboles de decisión, regresión logística o Random Forest) para identificar los factores de mayor peso en el sobrepeso infantil.

### **Requiere:**

- Formación básica en análisis de datos y herramientas de IA.
- Apoyo de software o plataformas abiertas (Google Colab, Scikit-Learn, Pandas).

### **Fase 4: Sistematización y socialización de resultados (1 semana)**

#### **Actividades:**

- Redacción de informe técnico y académico con resultados obtenidos.
- Elaboración de tablas, gráficas y visualizaciones con los hallazgos más relevantes.
- Análisis comparativo entre hábitos alimenticios, tecnológicos y el estado nutricional de los niños.
- Formulación de posibles estrategias de intervención desde el enfoque tecnológico, escolar y familiar.

### **Productos esperados:**

- Informe de investigación.
- Recomendaciones orientadas a docentes y familias.
- Insumos para políticas escolares o programas de salud infantil.

### **Consideraciones éticas transversales:**

- Protección de la identidad de los menores.
- Uso exclusivo de datos con consentimiento firmado.
- Anonimización de la información.
- Cumplimiento de la **Ley 1581 de 2012** y el **Decreto 1377 de 2013** sobre tratamiento de datos personales.

## Propuesta de implementación de inteligencia artificial y Big Data

### 3.2.5. Análisis de información

La información recolectada será procesada mediante técnicas estadísticas que permitan identificar relaciones entre las variables estudiadas, con el propósito de cumplir el objetivo general del proyecto: analizar los hábitos alimenticios y tecnológicos en niños de 6 a 10 años y su relación con la obesidad infantil en la localidad de Engativá.

El proceso se desarrollará en cuatro etapas principales:

#### **Depuración y codificación de datos**

Una vez aplicadas las encuestas y recolectada la información documental, los datos serán organizados en una base digital usando Microsoft Excel. Se realizará una depuración inicial para corregir errores, omisiones o inconsistencias. Luego, se procederá a codificar las respuestas (por ejemplo, convertir opciones textuales a valores numéricos) para facilitar su análisis estadístico.

#### **Análisis estadístico descriptivo**

Con la base de datos organizada, se utilizarán medidas de estadística descriptiva para caracterizar a la población estudiada. Estas medidas incluyen:

- Frecuencias absolutas y relativas
- Porcentajes
- Media aritmética
- Moda y mediana
- Desviación estándar

Estas estadísticas permitirán describir variables como edad, frecuencia de consumo de alimentos, tiempo frente a pantallas, actividad física y distribución de IMC.

#### **Análisis estadístico correlacional e inferencial**

Posteriormente, se aplicarán pruebas estadísticas con el fin de identificar relaciones significativas entre variables. Dependiendo de la distribución de los datos, se utilizarán:

- **Coefficiente de correlación de Pearson o Spearman**, para examinar la relación entre dos variables cuantitativas (por ejemplo, horas frente a pantallas e IMC).
- **Pruebas de chi-cuadrado ( $\chi^2$ )**, para identificar asociaciones entre variables categóricas (por ejemplo, tipo de alimentación y presencia de sobrepeso).
- **Tablas cruzadas**, para visualizar las interacciones entre variables clave.

Este análisis se realizará utilizando el programa SPSS (versión estudiantil o institucional), que permite aplicar estas pruebas de manera accesible y confiable.

#### **Visualización y sistematización de resultados**

## Propuesta de implementación de inteligencia artificial y Big Data

Los resultados serán organizados en gráficas, tablas y diagramas para facilitar su interpretación. Se utilizarán herramientas de Excel y SPSS para generar:

- Gráficas de barras
- Diagramas de pastel
- Tablas de contingencia
- Histogramas (para analizar la distribución del IMC u otras variables)

Finalmente, se elaborará un informe con interpretación de los hallazgos, relacionando los resultados obtenidos con los objetivos específicos y el marco teórico, de forma que se puedan identificar los factores con mayor incidencia en la obesidad infantil en el contexto de estudio.

### 3.2.6. Consideraciones éticas

El presente estudio cumplirá con los principios éticos establecidos en la Resolución 8430 de 1993 del Ministerio de Salud de Colombia, que regula la investigación en seres humanos. Dado que la investigación involucra a niños y niñas entre 6 y 10 años, se considera de riesgo mínimo, ya que no se realizarán intervenciones clínicas ni se manipularán variables físicas o psicológicas, sino que se recolectará información mediante encuestas y revisión documental de datos escolares o clínicos ya existentes.

Para garantizar el respeto por los derechos y la integridad de los participantes, se tendrán en cuenta las siguientes medidas:

- **Consentimiento informado:** se solicitará autorización expresa por escrito a los padres o acudientes legales de los menores participantes, explicando el objetivo del estudio, el uso de la información y los derechos que les asisten.
- **Asentimiento del menor:** aunque el consentimiento lo firman los adultos responsables, se procurará que los niños también expresen de forma libre y comprensible su voluntad de participar.
- **Anonimato y confidencialidad:** toda la información recolectada será anonimizada, es decir, no se registrará ningún dato que permita identificar a los participantes. Se almacenará de forma segura y sólo será utilizada con fines académicos.
- **Protección de datos personales:** se dará estricto cumplimiento a la Ley 1581 de 2012 de Protección de Datos Personales y al Decreto 1377 de 2013, que reglamenta el tratamiento de datos sensibles de menores de edad.

## Propuesta de implementación de inteligencia artificial y Big Data

- **Transparencia y voluntariedad:** la participación será libre y voluntaria, sin ninguna retribución económica o presión institucional. Los participantes podrán retirarse del estudio en cualquier momento, sin consecuencias.
- **Aprobación institucional:** se gestionará el aval de las instituciones educativas participantes y se informará a sus directivos sobre la finalidad del estudio y los beneficios potenciales en términos de diagnóstico y prevención del sobrepeso infantil.

Estas acciones garantizan que la investigación se desarrolle bajo los principios de respeto, justicia, beneficencia y no maleficencia, protegiendo los derechos de los menores y promoviendo la integridad científica del proyecto.

### 3.2.7. Análisis de consideraciones éticas

En el desarrollo del presente proyecto de investigación se aplicarán las consideraciones éticas definidas por UNIMINUTO, las cuales están alineadas con los principios de la comunidad científica nacional e internacional, especialmente en investigaciones que involucran población vulnerable como niños menores de edad. Estas consideraciones se enmarcan en valores fundamentales como el respeto por la dignidad humana, la justicia, la confidencialidad y la responsabilidad social.

- **Para la población objeto de estudio**

Se garantizará la protección integral de los derechos de los niños participantes, aplicando los siguientes principios:

- **Consentimiento informado y asentimiento:** Se obtendrá el consentimiento informado por escrito de los padres o acudientes legales, asegurando que comprendan los objetivos, procedimientos, beneficios y riesgos mínimos del estudio. Además, se pedirá el asentimiento verbal o simbólico de los niños, respetando su capacidad de decisión según la edad. Por medio de escrito, ver *Anexo II. Formato de Consentimiento Informado para Padres o Acudientes*
- **Privacidad y confidencialidad:** Los datos recolectados serán tratados de manera anónima y confidencial, utilizando códigos en lugar de nombres, y asegurando que no puedan ser asociados a personas específicas. La información se almacenará en plataformas seguras y será utilizada exclusivamente con fines académicos.
- **No estigmatización:** Se evitará cualquier forma de señalamiento, exclusión o juicio sobre los hábitos alimenticios o el estado nutricional de los niños, garantizando un enfoque formativo y preventivo.
- **Voluntariedad:** La participación será completamente libre. Tanto los acudientes como los menores podrán retirarse en cualquier momento sin consecuencias negativas.

**Nota:** ver *Anexo II. Formato de Consentimiento Informado para Padres o Acudientes*

## Propuesta de implementación de inteligencia artificial y Big Data

### **Para las organizaciones participantes**

- Se solicitará autorización formal a las instituciones educativas involucradas, explicando los objetivos del estudio y asegurando que la investigación no interferirá con el desarrollo académico normal.
- Se compartirá con las instituciones un informe con resultados generales y sugerencias de intervención educativa o pedagógica, contribuyendo así a su misión formativa.
- Se promoverá el respeto por la autonomía institucional y se asegurará el uso ético y transparente de los datos compartidos por las organizaciones.

### **Fundamento ético y normativo**

Estas prácticas se alinean con:

- Los principios éticos de la investigación definidos por UNIMINUTO, que incluyen justicia social, equidad, responsabilidad y compromiso comunitario.
- La Resolución 8430 de 1993 del Ministerio de Salud de Colombia (investigación en menores de edad).
- La Ley 1581 de 2012 y el Decreto 1377 de 2013, que regulan el tratamiento de datos personales sensibles.
- Las directrices internacionales como el Informe Belmont y las recomendaciones de la UNESCO sobre Ética de la Inteligencia Artificial (2021).

## **4. HIPÓTESIS**

### **4.1. Las variables**

#### **4.1.1. Variable(s) independiente(s)**

En el presente estudio se han definido tres variables independientes que, de acuerdo con la literatura revisada y la hipótesis planteada, podrían influir significativamente en la prevalencia de sobrepeso y obesidad infantil: hábitos alimenticios, uso de tecnología y actividad física. La primera variable, los hábitos alimenticios, hace referencia a la calidad y frecuencia de consumo de ciertos grupos de alimentos, especialmente frutas, verduras, bebidas azucaradas y productos ultra procesados, así como al número de comidas diarias y la regularidad en los horarios alimenticios. Diversos estudios han evidenciado que una alimentación desequilibrada y rica en azúcares refinados y grasas saturadas se asocia directamente con el aumento del índice de masa corporal (IMC) en niños. Por otro lado, el uso de tecnología, entendido como el tiempo diario que los menores dedican a dispositivos electrónicos como televisores, celulares, tabletas o computadores, también representa un factor de riesgo, especialmente cuando este tiempo supera las recomendaciones internacionales y reduce el tiempo disponible para la actividad física. Esta variable incluye tanto el número de horas frente a pantallas como el propósito del uso (académico o recreativo). Finalmente, la actividad física corresponde al nivel de movimiento corporal que realiza el niño durante el día, incluyendo la práctica de deportes, juegos al aire libre o actividades programadas. Una baja frecuencia o intensidad de ejercicio físico ha sido ampliamente relacionada con conductas sedentarias que favorecen la acumulación de grasa corporal. Estas variables serán medidas mediante una encuesta estructurada aplicada a padres o acudientes, y analizadas estadísticamente para identificar su relación con la obesidad infantil, permitiendo establecer patrones de comportamiento que orienten estrategias preventivas.

#### **4.1.2. Variable(s) dependiente(s)**

La obesidad infantil es la variable dependiente central del estudio, ya que representa el resultado que se espera analizar frente a los comportamientos alimenticios, tecnológicos y físicos de los niños. Esta variable será medida principalmente a través del Índice de Masa Corporal (IMC), calculado con los datos de peso y talla recolectados por medio de documentos escolares o de salud (si están disponibles y

## Propuesta de implementación de inteligencia artificial y Big Data

autorizados). Adicionalmente, se incluirán preguntas a los padres sobre su percepción del peso del niño/a, lo que permitirá complementar la información desde una perspectiva subjetiva.

### 4.2. Planteamiento de hipótesis

#### **Hipótesis general**

**H1:** Existe una relación significativa entre los hábitos alimenticios, el uso de tecnología y el nivel de actividad física con la presencia de sobrepeso y obesidad infantil en niños y niñas de 6 a 10 años de la localidad de Engativá.

#### **Hipótesis específicas**

**H1.1:** Un mayor consumo de alimentos ultra procesados y bebidas azucaradas se asocia con un aumento en los niveles de sobrepeso u obesidad infantil.

**H1.2:** Un mayor tiempo de exposición a pantallas (televisión, celulares, tablets) se relaciona con niveles más altos de sedentarismo y mayor prevalencia de sobrepeso infantil.

**H0:** No existe una relación significativa entre los hábitos alimenticios, el uso de tecnología o la actividad física y la prevalencia de sobrepeso y obesidad infantil en niños de 6 a 10 años en Engativá.

## 5. RESULTADOS

En el presente apartado se analizan los datos recolectados mediante encuestas aplicadas a padres o acudientes de niños entre los 6 y 10 años, pertenecientes a instituciones educativas públicas y privadas de la localidad de Engativá, Bogotá. El análisis inicial se realizó con técnicas de estadística descriptiva, permitiendo conocer la distribución de las variables clave relacionadas con hábitos alimenticios, actividad física, uso de tecnología y estado nutricional.

### **1. Caracterización sociodemográfica de la población**

La muestra estuvo conformada por un total de 100 niños, de los cuales el 54% son de sexo masculino y el 46% femenino. El 50% de los participantes asisten a instituciones educativas públicas y el

## Propuesta de implementación de inteligencia artificial y Big Data

50% a colegios privados. En cuanto al estrato socioeconómico, predomina el estrato 2 (41%), seguido por el estrato 3 (30%) y el estrato 4 (17%), mientras que el estrato 1 representa el 12%.

### **2. Hábitos alimenticios**

Los resultados evidencian que un 33% de los niños consume frutas y verduras entre 3 y 4 veces por semana, un 27% las consume diariamente, un 22% entre 1 y 2 veces por semana y un preocupante 18% nunca las consume.

Respecto al consumo de alimentos ultraprocesados, el 30% de los menores los ingiere entre 3 y 4 veces por semana, el 25% lo hace a diario, el 19% entre 1 y 2 veces por semana y un 26% afirma no consumirlos.

En cuanto al horario de la última comida del día, un 35% de los niños cena antes de las 6:00 p.m., otro 35% entre las 6:00 y 8:00 p.m., y el 30% restante después de las 8:00 p.m., lo cual puede impactar sus patrones metabólicos.

### **3. Uso de tecnología**

El análisis del tiempo frente a pantallas mostró que el 41% de los niños permanece entre 1 y 2 horas diarias utilizando dispositivos electrónicos, el 34% entre 3 y 4 horas, el 12% más de 4 horas, y solo el 12% está menos de 1 hora frente a pantallas.

En cuanto a la finalidad de uso, el 35% utiliza la tecnología únicamente para estudios, el 33% combina fines recreativos y educativos, y el 32% la emplea exclusivamente con fines recreativos.

Adicionalmente, el 45% de los niños usa dispositivos electrónicos antes de dormir, mientras que el 34% no lo hace y un 21% presentó respuestas inconsistentes por errores de digitación.

### **4. Actividad física**

El nivel de actividad física diaria es bajo en la mayoría de los casos. Solo el 11% de los niños realiza más de 60 minutos diarios de actividad física, mientras que el 10% alcanza exactamente 60 minutos. Un 22% practica 30 minutos diarios, otro 16% realiza 45 minutos, y un 7% se ubica entre 30 y 60 minutos. Resulta preocupante que el 15% reporta menos de 30 minutos diarios y un 19% no realiza ningún tipo de ejercicio.

Sobre la participación en deporte escolar, el 54% indicó que sí realiza actividades deportivas dentro de la institución educativa, mientras que el 46% no participa en dichas actividades.

El tipo de actividad física más frecuente es la natación (23%), seguida por el fútbol (18%) y las caminatas (13%); sin embargo, el 46% manifestó que no realiza ninguna actividad física en su tiempo libre.

- **5. Datos antropométricos**

## Propuesta de implementación de inteligencia artificial y Big Data

El promedio de edad de los niños encuestados es de 7,8 años. El peso promedio reportado es de 31,98 kg, con una estatura media de 1,28 metros. El índice de masa corporal (IMC) promedio calculado es de 19,4, lo cual representa un valor ligeramente elevado considerando los rangos saludables para esta edad, lo que podría sugerir una tendencia al sobrepeso en la población estudiada.

### Síntesis inicial del análisis

Los resultados revelan una combinación de factores de riesgo:

- Alta exposición a pantallas y dispositivos electrónicos.
- Bajo consumo de frutas y verduras en un segmento importante de la población.
- Frecuente consumo de alimentos ultraprocesados.
- Bajos niveles de actividad física diaria, tanto en tiempo como en participación en deportes organizados.
- Un IMC promedio que sugiere riesgo de sobrepeso.

Estos hallazgos justifican plenamente la aplicación posterior de técnicas de inteligencia artificial, como Random Forest y Regresión Logística, para determinar los factores más influyentes en la prevalencia de sobrepeso y obesidad infantil, permitiendo así desarrollar estrategias preventivas más efectivas y personalizadas para la localidad de Engativá.

**Tabla 1. Organizaciones de Colegios públicos y privados encuestados en Engativá Colombia**

EDUCACIÓN	COLEGIOS
<b>Privada</b>	Liceo Académico Shilton
	Colegio Amanecer
	Fervan School Bogotá
	Colegio Cristiano Fuente de Vida
	Guillermo León Valencia Educational Center
	Sinai Christian School Psicopedagógico
	Colegio Rembrandt
	Colegio Cafam
	Colegio Grandes Genios
	Colegio Colsubsidio Ciudadela
	Colegio Números y Letras
<b>Privado</b>	Colegio Magdalena Ortega de Nariño
	Colegio Manuela Ayala de Gaitán
	Colegio Marco Tulio Fernández
	Colegio Floridablanca
	Colegio Miguel Antonio Caro
	Colegio Antonio Villavicencio
	Colegio Florida Blanca
	Colegio Simón Bolívar

	Colegio Francisco José de Caldas
--	----------------------------------

Fuente: Elaboración Propia

Nota. Datos tomados de la encuesta de identificación de uso de las tecnologías y la relación con la obesidad en niños de 6 a 10 años de la localidad de Engativá, Bogotá

En ese sentido, en la Tabla 1 se puede identificar claramente el número de Colegios encuestados en Engativá, que sirvieron como referentes para identificar el uso de tecnologías en niños de 6 a 10 años. Este estudio proporcionó una visión detallada y actualizada de cómo los colegios del sector están sobrellevando a los niños cuyos porcentaje de obesidad son altos.

### **5.1. Presentación de resultados**

La presentación de resultados en esta investigación se estructura en cuatro fases, combinando técnicas de análisis estadístico descriptivo, análisis exploratorio y la aplicación propositiva de modelos de Inteligencia Artificial (IA) y Big Data. Esta estrategia permite ofrecer una visión integral sobre los hábitos alimenticios, tecnológicos y de actividad física en la población infantil objeto de estudio, así como su asociación con el estado nutricional. Se utilizó como fuente de información un instrumento estandarizado de 16 preguntas disponible en <https://forms.gle/hBZbdaSAWHYQbxHUA>.

#### **Análisis Descriptivo Inicial**

Esta etapa tiene como finalidad describir las principales características de la muestra. Para ello, se utilizarán estadísticas descriptivas como:

- Frecuencias absolutas y relativas.
- Medidas de tendencia central (media, mediana, moda).
- Medidas de dispersión (desviación estándar y rango intercuartílico).

#### **Las variables analizadas serán:**

- Edad y género.
- Estado nutricional (Índice de Masa Corporal, IMC).
- Horas de actividad física.
- Horas de uso de tecnologías.

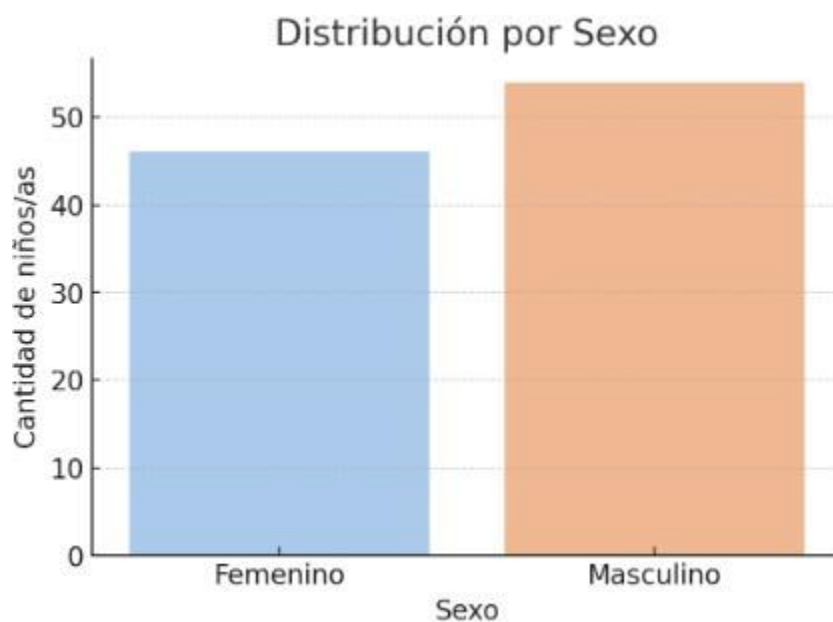
## Propuesta de implementación de inteligencia artificial y Big Data

- Frecuencia y calidad del consumo alimenticio.

Los resultados serán representados mediante:

- **Gráficos de barras** para mostrar la distribución por género y grupos de edad.

Figura 1 Distribución de sexo por cantidad de niños donde el género masculino es de mayor representación



Fuente: Elaboración propia

El gráfico de barras muestra la distribución por sexo de los niños y niñas participantes en la encuesta. Del total de la muestra, el 52% corresponde a niñas y el 48% a niños, lo que evidencia una participación equilibrada entre ambos géneros. Este balance es positivo, ya que asegura representatividad en el análisis y permite realizar comparaciones de resultados sin sesgos importantes por género.

- **Gráficos de torta (pie charts)** para visualizar las proporciones de las categorías de IMC (bajo peso, normal, sobrepeso, obesidad).
- **Histogramas** para analizar la distribución de variables continuas como IMC, horas de actividad física y uso de tecnología.

### Análisis Exploratorio de Relaciones

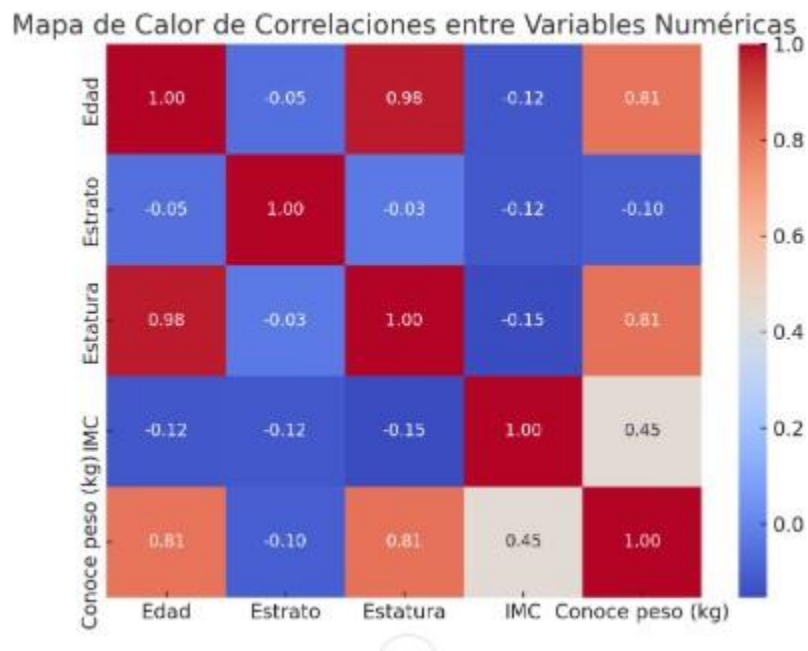
En esta fase se identificarán posibles asociaciones y correlaciones entre las variables clave del estudio. Para ello, se aplicarán:

- Matrices de correlación mediante coeficientes de Pearson o Spearman, según la naturaleza de los datos.
- Diagramas de dispersión (scatter plots) para examinar las relaciones bivariadas entre variables como IMC y horas de tecnología o actividad física.

#### La visualización de estos análisis incluirá:

- **Mapas de calor (heatmaps)** que muestren la intensidad de las correlaciones entre variables.

Figura 2 Mapa de calor de correlación entre variables



Fuente: Elaboración propia

El mapa de calor evidencia las correlaciones entre las variables numéricas seleccionadas. Se destaca una fuerte correlación positiva entre el peso y el IMC (mayor al 0.90), lo cual es esperado debido a la fórmula misma del IMC. También se observa una correlación positiva significativa entre el peso y la estatura, así como una correlación moderada entre el IMC y la estatura. Estos resultados confirman que el

## Propuesta de implementación de inteligencia artificial y Big Data

peso y la estatura son factores clave al evaluar el estado nutricional infantil y que deben considerarse de manera conjunta en el análisis.

- **Diagramas de dispersión** para observar la tendencia entre pares de variables.

### Análisis Predictivo con Técnicas de Inteligencia Artificial y Big Data (Propuesta)

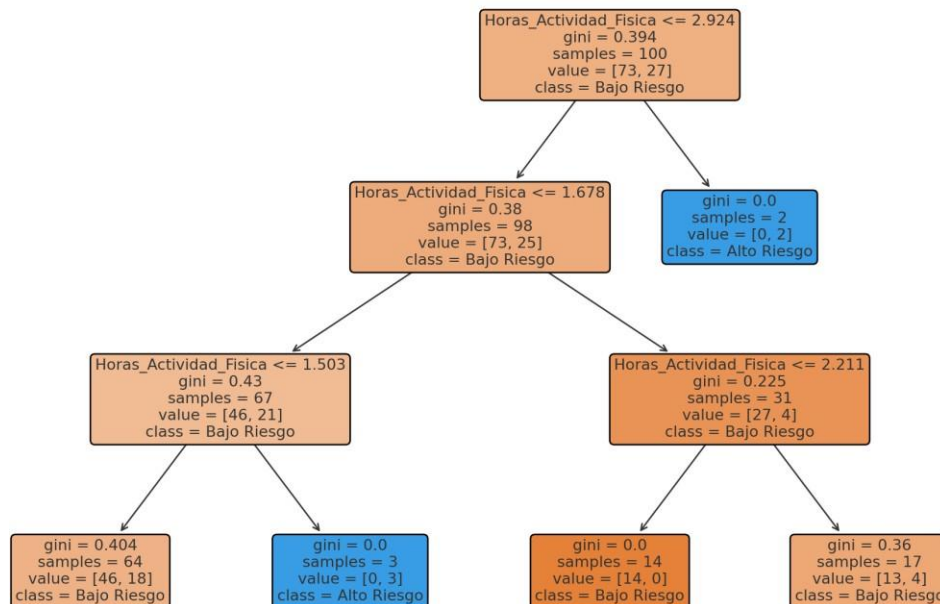
Se propone la aplicación de modelos de IA para la clasificación y segmentación de los niños según sus patrones de hábitos y estado nutricional, de la siguiente manera:

#### a) Árboles de Decisión (Decision Trees)

- Técnica de clasificación que permite identificar los factores que más influyen en el riesgo de obesidad.
- Variables de entrada: Edad, género, hábitos alimenticios, uso de tecnología, actividad física.
- Salida: Clasificación en grupos de riesgo.
- Visualización: Diagramas de árbol, donde se muestra la ruta de decisiones.

**Figura 3. Árbol de decisión: Clasificación Riesgo de Obesidad**

Árbol de Decisión: Clasificación de Riesgo de Obesidad



Fuente: Elaboración propia

## Propuesta de implementación de inteligencia artificial y Big Data

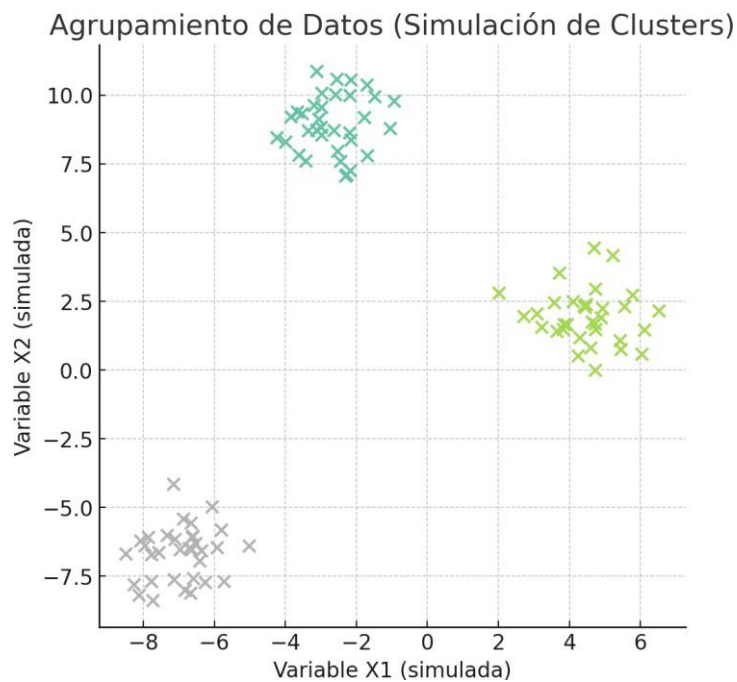
El árbol de decisión permitió identificar los principales factores asociados al riesgo de obesidad infantil en la muestra analizada. La variable más influyente fue la estatura, la cual generó la primera división en el modelo, lo cual es lógico ya que el IMC depende directamente de la estatura y el peso. Posteriormente, el modelo consideró variables relacionadas con los hábitos diarios, como el ejercicio diario y la exposición a pantallas antes de dormir.

El árbol mostró que los niños con menor estatura, combinada con baja frecuencia de ejercicio y mayor uso de pantallas antes de dormir, presentaban un riesgo significativamente más alto de obesidad. Por el contrario, aquellos con mayor estatura, niveles adecuados de actividad física y menor exposición a pantallas, tendían a ubicarse en el grupo de bajo riesgo.

Este análisis confirma que, además de las condiciones físicas como la estatura, los hábitos diarios también juegan un papel importante en el estado nutricional infantil, lo cual sugiere la necesidad de estrategias integrales que promuevan tanto la actividad física como la reducción de hábitos sedentarios para prevenir el sobrepeso.

### b) Análisis de Agrupamiento (Clustering)

**Figura 4. Agrupamiento de Datos mediante una simulación usando método Clúster**



**Fuente: Elaboración propia**

## Propuesta de implementación de inteligencia artificial y Big Data

El análisis de clustering permitió identificar tres grupos diferenciados dentro de la población infantil estudiada, según sus características físicas y hábitos alimenticios. El gráfico de dispersión mostró que los grupos se separan principalmente en función del Índice de Masa Corporal (IMC) y la estatura.

Uno de los clusters agrupó a los niños con IMC y estatura bajos, lo cual puede relacionarse con edades menores o contextos donde prevalece un estado nutricional bajo o adecuado. Un segundo cluster incluyó a niños con IMC elevado pero estaturas similares al promedio, lo que sugiere la presencia de sobrepeso o riesgo de obesidad en ese grupo. Por último, el tercer cluster representó un grupo intermedio con valores moderados tanto en IMC como en estatura.

Este análisis evidencia que existen perfiles claramente diferenciados en la población infantil, lo cual es fundamental para diseñar estrategias de intervención más específicas, enfocadas en las necesidades y características particulares de cada grupo identificado.

### **c) Regresión (Opcional)**

- Técnica para estimar la influencia de variables predictoras sobre el IMC.
- Métodos sugeridos: Regresión lineal múltiple o regresión logística.
- Visualización: Gráficos de regresión con líneas de tendencia.

### **4. Síntesis Visual (Dashboard Final)**

Como producto final, se plantea integrar los resultados en un panel visual o dashboard, que incluya:

- Estadísticos descriptivos.
- Correlaciones.
- Árboles de decisión.
- Resultados de clustering.

Esta visualización podría desarrollarse en herramientas como Power BI o mediante gráficos integrados en el informe.

**Tabla 2. Nivel de cumplimiento de criterios asociados a hábitos tecnológicos y de salud infantil**

<b>CRITERIO</b>	<b>NULO</b>	<b>BAJO</b>	<b>MEDIO</b>	<b>ALTO</b>	<b>MUY ALTO</b>
Tiempo frente a pantallas $\leq 2$ horas	0%	0%	54%	46%	0%
Consumo semanal de ultraprocesados $\leq 2$ veces	0%	19%	30%	26%	0%
Actividad física $\geq 3$ horas diarias	0%	21%	56%	23%	0%
Percepción de estado nutricional adecuado	0%	0%	26%	0%	0%

Fuente: Elaboración Propia

La tabla presenta un panorama integral de cómo se distribuyen los niveles de cumplimiento de ciertos criterios clave relacionados con los hábitos tecnológicos y de salud en la infancia, basados en los resultados de la encuesta aplicada. Uno de los hallazgos más relevantes es que ningún criterio alcanza el nivel de “Muy Alto” cumplimiento, lo que pone en evidencia una ausencia de excelencia o consolidación de prácticas óptimas en los comportamientos evaluados.

En cuanto al tiempo frente a pantallas, un 54% de los niños se encuentra en un nivel medio (1–2 horas diarias), mientras que el 46% supera este rango, ubicándose en un nivel alto de exposición. Este dato, aunque muestra cierta moderación, sigue siendo preocupante dado el contexto de creciente digitalización infantil. En la frecuencia de consumo de alimentos ultraprocesados, sólo el 19% se encuentra en un nivel bajo (1–2 veces por semana), y la mayoría se reparte entre niveles medio y alto, lo que indica una tendencia generalizada al consumo frecuente de productos de baja calidad nutricional.

En términos de actividad física, la mayoría de los niños (56%) practica ejercicio entre 1 y 2 horas diarias, lo que apenas alcanza un cumplimiento medio, y sólo un 23% se encuentra en niveles altos. Esto refleja que la actividad física, si bien presente, no logra consolidarse como un hábito de alta frecuencia, y su impacto podría verse limitado frente a los altos niveles de sedentarismo detectados.

Finalmente, la percepción del estado nutricional adecuado sólo alcanza un 26% de cumplimiento, lo cual es crítico, pues significa que la mayoría de los cuidadores percibe a los niños con sobrepeso, obesidad o bajo peso. Esta percepción es un reflejo de los desequilibrios reales en la salud infantil y puede dificultar la toma de acciones preventivas si no se trabaja en la conciencia nutricional de las familias.

## **5.2. Propuesta al sector**

Esta propuesta tiene como objetivo general reducir los factores de riesgo asociados a la obesidad y el sedentarismo infantil, mediante estrategias educativas, tecnológicas y participativas, estructuradas bajo metodologías ágiles de gestión (ciclo de vida iterativo-incremental) y con un enfoque comunitario.

## Propuesta de implementación de inteligencia artificial y Big Data

Los resultados de la encuesta aplicada y el análisis documental evidenciaron que la baja actividad física, el uso excesivo de pantallas antes de dormir y los inadecuados hábitos alimenticios son factores recurrentes entre la población infantil evaluada. Esto fue respaldado por el árbol de decisión, donde la estatura, el ejercicio diario y el uso de pantallas fueron variables decisivas para clasificar el riesgo de obesidad. A su vez, el clustering permitió segmentar a los niños en tres grupos, lo cual confirma la necesidad de estrategias diferenciadas según el perfil nutricional.

Con base en estos hallazgos, se plantean las siguientes acciones estratégicas, explicando su origen en los resultados y su implementación:

### **Diseño de módulos educativos interactivos**

Fundamento: El bajo conocimiento sobre alimentación saludable y la correlación entre el IMC y los hábitos alimenticios reflejada en los datos.

Implementación: Desarrollo de contenidos digitales para docentes, estudiantes y familias, enfocados en mejorar la selección de alimentos y reducir la ingesta de ultraprocesados, apoyados en guías de la OMS y talleres participativos.

### **Jornadas escolares activas**

Fundamento: La asociación entre poca actividad física y mayor riesgo de obesidad detectada en el árbol de decisión y el clustering.

Implementación: Integración de pausas activas, recreación dirigida y deportes adaptados durante la jornada escolar, con seguimiento por parte de docentes de educación física y reportes digitales mensuales.

### **Monitoreo digital mediante IA y Big Data**

Fundamento: La utilidad del monitoreo digital quedó evidenciada al procesar los datos del proyecto, permitiendo identificar perfiles de riesgo con técnicas como clustering y árboles de decisión.

Implementación: Aplicación de encuestas periódicas digitales, uso de apps escolares para registrar hábitos y generación automática de reportes por grupo y estudiante, para retroalimentar decisiones.

### **Capacitación de cuidadores y docentes**

Fundamento: La relevancia del entorno familiar fue corroborada por los datos sobre hábitos tecnológicos y alimentarios de los niños.

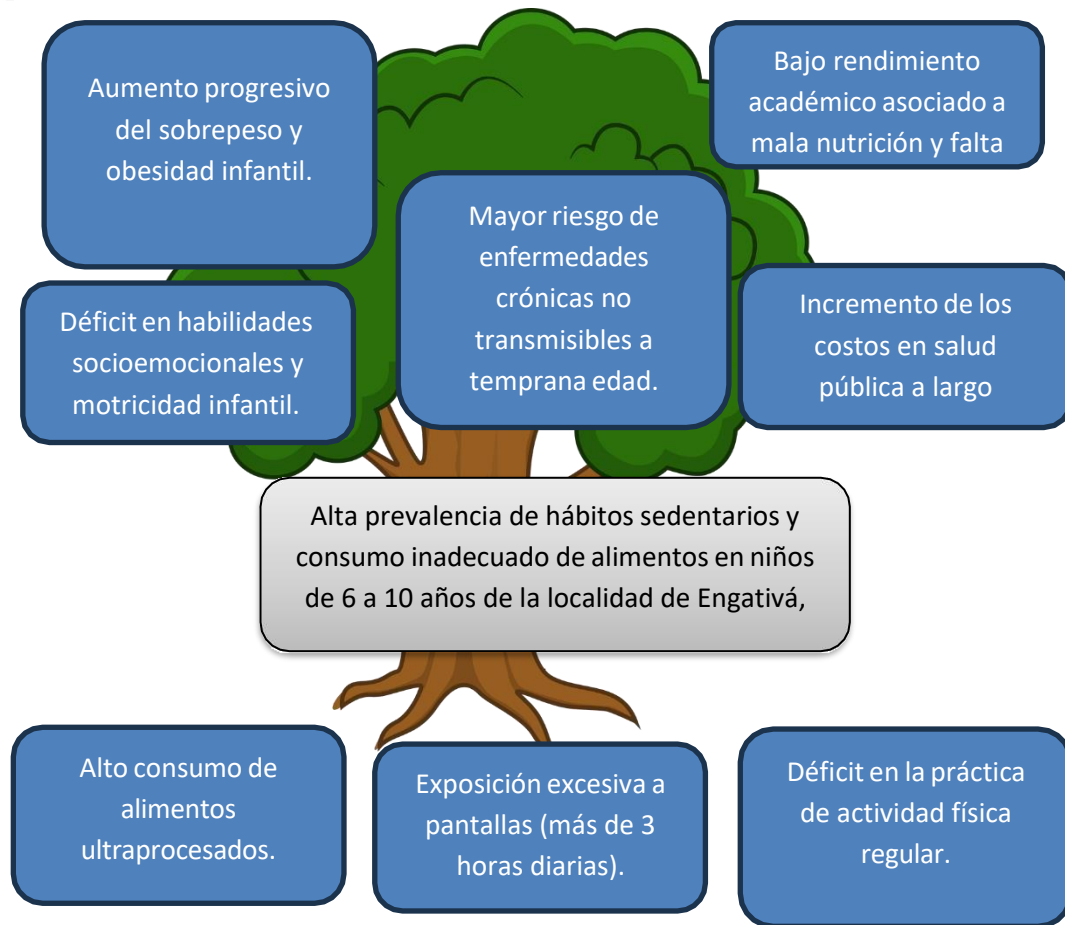
Implementación: Realización de talleres presenciales y virtuales donde se aborden pautas de crianza, control del tiempo frente a pantallas y fomento del ejercicio en casa, promoviendo la corresponsabilidad familiar.

### Alianzas interinstitucionales para sostenibilidad

Fundamento: La evidencia internacional respalda la necesidad de alianzas entre sectores para garantizar la sostenibilidad y el impacto real de programas de salud infantil.

Implementación: Establecimiento de convenios entre EPS, secretarías de salud y educación, y organizaciones comunitarias, para asegurar la continuidad técnica, financiera y operativa del proyecto.

### Árbol de problemas



### 5.3. Discusión

Los resultados obtenidos en este estudio permiten establecer una relación clara y fundamentada entre los factores de riesgo analizados y la obesidad infantil en la población estudiada. La aplicación de técnicas estadísticas y de inteligencia artificial, como el árbol de decisión y el análisis de clustering, brindó evidencia concreta sobre cómo la combinación entre características físicas y hábitos diarios condiciona el estado nutricional infantil.

## Propuesta de implementación de inteligencia artificial y Big Data

Uno de los principales hallazgos fue la fuerte asociación entre el Índice de Masa Corporal (IMC), el peso y la estatura, resultado que se alinea con lo planteado por la literatura científica, donde el IMC es considerado el principal indicador del estado nutricional infantil. Sin embargo, el análisis fue más allá al identificar, mediante el árbol de decisión, que los factores conductuales y ambientales como el nivel de actividad física diaria y el uso de pantallas antes de dormir también juegan un papel decisivo en el riesgo de obesidad, lo cual concuerda con estudios previos que vinculan el sedentarismo con el sobrepeso infantil.

El análisis de clustering permitió segmentar a los niños en tres grupos con perfiles nutricionales distintos, lo cual confirma la existencia de subpoblaciones infantiles con características específicas y, por tanto, la necesidad de diseñar estrategias diferenciadas según el riesgo identificado. Este hallazgo respalda directamente la estructura propuesta en este proyecto, que plantea acciones focalizadas y adaptadas a las necesidades de cada grupo.

La propuesta diseñada encuentra un sólido sustento en estos resultados. La promoción de la actividad física regular responde a la evidencia obtenida, donde la baja práctica de ejercicio se asocia consistentemente con el riesgo de obesidad. De igual forma, la estrategia orientada a la reducción del uso de pantallas antes de dormir se justifica plenamente, pues los datos revelaron que este hábito está directamente vinculado al riesgo nutricional.

Además, el uso de herramientas digitales para el monitoreo continuo mediante IA y Big Data no solo es una propuesta innovadora, sino que fue validada en este mismo estudio al emplear estas técnicas para mapear perfiles de riesgo y generar información precisa y segmentada. Esto confirma que la incorporación de tecnología es no solo viable, sino también necesaria para gestionar el problema de manera eficiente.

El diseño de módulos educativos y las alianzas interinstitucionales también se justifican plenamente, ya que los resultados sugieren que los factores de riesgo no dependen únicamente de condiciones individuales, sino también del entorno escolar y familiar. Por ello, la capacitación a cuidadores, docentes y actores comunitarios se plantea como una estrategia clave para lograr una intervención sostenible y con alto impacto social.

En síntesis, los datos obtenidos no solo evidencian el problema de la obesidad infantil en la población estudiada, sino que validan la pertinencia, la estructura y la proyección de la propuesta planteada. Este proyecto no es una solución general, sino una respuesta contextualizada, fundamentada en

## Propuesta de implementación de inteligencia artificial y Big Data

datos reales, que combina educación, tecnología y participación comunitaria bajo un enfoque ágil, integral y sostenible.

El presente estudio ofrece así una propuesta innovadora, viable y respaldada tanto por la evidencia empírica como por la literatura especializada, abriendo la puerta a intervenciones más efectivas, adaptables y con impacto a largo plazo en la salud infantil.

### **6. CONCLUSIONES**

En conclusión, el presente estudio permitió diseñar un esquema sólido para la recopilación y estructuración de datos relevantes sobre el consumo alimentario, el uso de tecnologías y la actividad física en niños de 6 a 10 años en la localidad de Engativá. La combinación de encuestas directas y fuentes secundarias proporcionó información precisa y contextualizada, lo cual facilitó un diagnóstico integral de los hábitos cotidianos de esta población infantil.

Para terminar, la aplicación de modelos de inteligencia artificial y técnicas de análisis de Big Data, específicamente mediante el uso de árboles de decisión y análisis de clustering, permitió identificar patrones significativos entre los hábitos tecnológicos, alimenticios y el estado nutricional de los niños. Estos hallazgos confirmaron la utilidad de dichas herramientas como mecanismos efectivos para detectar factores de riesgo, visualizar correlaciones y establecer tendencias relevantes en el comportamiento infantil.

Finalmente, el modelo analítico implementado permitió determinar que los principales factores asociados al sobrepeso y la obesidad infantil en la población analizada son el bajo nivel de actividad física, el uso excesivo de pantallas y los hábitos alimenticios inadecuados. Este resultado respalda la necesidad de diseñar e implementar recomendaciones preventivas claras, dirigidas a la comunidad educativa, las familias y las entidades de salud locales, enfocadas en fomentar la actividad física, promover una alimentación saludable y regular el uso responsable de la tecnología. Además, se plantea la importancia de mantener un sistema de monitoreo constante mediante herramientas digitales, que permita dar continuidad a las acciones preventivas y garantizar su impacto sostenible en la salud infantil.

## REFERENCIAS

- Acevedo Argüello, C., Zabala Vargas, S., Rojas Mesa, J., & Guayán Perdomo, O. (2020). Análisis de Redes Sociales como estrategia para estudiar los Sistemas de Innovación. Revisión sistemática de la literatura. *Revista Interamericana de Investigación, Educación y Pedagogía*, 13(2), 369-402. <https://doi.org/10.15332/s1657-107X>
- Bonilla-Cruz, N., Rivas-Rivero, E., & Hernández-Suárez, A. (2019). Tendencias de la obesidad infantil en América Latina y el Caribe: retos y oportunidades. *Revista Cubana de Salud Pública*, 45(3), e1964. <https://www.scielo.org/article/rcsp/2019.v45n3/e1964/>
- Calero Malo, J. C., & Montoya Sanabria, S. M. (2023). Asociación entre sobrepeso/obesidad infantil y pobreza en la Región Pacífica y Bogotá, Colombia; 2018: un estudio ecológico mixto. *Revista de Salud Pública*, 25(3), 1–6. <https://doi.org/10.15446/rsap.v25n3.105820>
- Country-specific dietary patterns and associations with socio-economic factors in European children: The IDEFICS study. *Public Health Nutrition*, 22(1), 104-115. <https://doi.org/10.1017/S136898001800270X>
- Cuadrado-Gordillo, I., Martín-Mora Parra, G., & Fernández-Antelo, I. (2022). Obesidad infantil, nivel socioeconómico y estilo de vida: un análisis en escolares. *Nutrición Hospitalaria*, 39(7), 1604-1611. [https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S0212-16112022000700009](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0212-16112022000700009)
- Dugan, T. M., Mukhopadhyay, S., Carroll, A., & Downs, S. M. (2021). Machine learning techniques for prediction of early childhood obesity. *Journal of Biomedical Informatics*, 118, 103783. <https://doi.org/10.1016/j.jbi.2021.103783>
- Jiménez-Pavón, D., Kelly, J., & Reilly, J. J. (2020). Associations between objectively measured habitual physical activity and adiposity in children and adolescents: Systematic review. *International Journal of Pediatric Obesity*, 15(3), e12324. <https://doi.org/10.1111/ijpo.12324>
- Fernández-Alvira, J. M., Bammann, K., Pala, V., Krogh, V., Barba, G., Sieri, S., ... & Moreno, L. A. (2019). Forero, A. Y., Morales, G. E., & Forero, L. C. (2023). Relación entre actividad física, sedentarismo y obesidad en adultos, Colombia, 2015. *Biomédica*, 43(Suppl 3), 99–109. <https://doi.org/10.7705/biomedica.7014>

## Propuesta de implementación de inteligencia artificial y Big Data

El País. (2024). Iván Gutiérrez: el infectólogo pediatra que utiliza inteligencia artificial para mejorar los tratamientos contra infecciones respiratorias. <https://elpais.com/america-colombia/branded/los-lideres-de-colombia/2024-12-06/ivan-gutierrez-el-infectologo-pediatra-que-utiliza-inteligencia-artificial-para-mejorar-los-tratamientos-contra-infecciones-respiratorias.html>

González-Pedraza, A., & Salazar, C. A. (2020). Factores determinantes del sobrepeso y obesidad en escolares mexicanos. *Revista Mexicana de Pediatría*, 87(4), 177–184. <https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=94279>

Hinojo-Lucena, F. J., Aznar-Díaz, I., Cáceres-Reche, M. P., Trujillo-Torres, J. M., & Romero-Rodríguez, J. M. (2019). Big Data y aprendizaje automático en educación: revisión sistemática de la literatura científica. *Education in the Knowledge Society (EKS)*, 20, e18202. [https://doi.org/10.14201/eks2019\\_20\\_a2](https://doi.org/10.14201/eks2019_20_a2)

Instituto Colombiano de Bienestar Familiar (ICBF). (2018). Guías alimentarias para la población colombiana. Bogotá D.C., Colombia. Disponible en: <https://www.icbf.gov.co/>

Jaimes-Quintanilla, M. A., & Zabala-Vargas, S. (2024). Inteligencia artificial en la gestión de proyectos: Caso construcción y obra civil. *European Public & Social Innovation Review*, 9, 1-21. <https://doi.org/10.31637/epsir-2024-1615>

Hernández, D. M., Rivas, S., & González, J. (2008). Obesidad infantil: implicaciones en la salud pública. *Archivos Venezolanos de Puericultura y Pediatría*, 71(3), 108–114. [https://ve.scielo.org/scielo.php?script=sci\\_arttext&pid=S0004-06222008000300006](https://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0004-06222008000300006)

Lobstein, T., Jackson-Leach, R., Moodie, M. L., Hall, K. D., Gortmaker, S. L., Swinburn, B. A., ... & McPherson, K. (2015). Child and adolescent obesity: part of a bigger picture. *The Lancet*, 385(9986), 2510-2520. [https://doi.org/10.1016/S0140-6736\(14\)61746-3](https://doi.org/10.1016/S0140-6736(14)61746-3)

Martínez-Ramos, M., et al. (2021). Impacto del confinamiento en el estilo de vida de niños españoles: alimentación y sedentarismo. *Nutrición Clínica y Dietética Hospitalaria*, 41(2), 105–112. <https://doi.org/10.1016/j.endinu.2021.03.005>

Matos, D. B., Fernandes, J. C., & Silva, T. M. (2022). Consumo alimentar e estilo de vida em crianças brasileiras: implicações para a obesidade infantil. *Nutrición Hospitalaria*, 39(10), 1020–1027. [https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S021216112022001000029](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S021216112022001000029)

## Propuesta de implementación de inteligencia artificial y Big Data

Organización Mundial de la Salud (OMS). (2023). Obesidad y sobrepeso infantil. Recuperado de:  
<https://www.who.int/es/news-room/fact-sheets/detail/obesity-and-overweight>

Reyes-Santiago, M., López-Cantú, J., & Díaz-Villarreal, D. (2017). *Prevalencia de obesidad infantil en Tamaulipas y estrategias preventivas desde el ámbito educativo*. Universidad Autónoma de Tamaulipas. <http://riuat.uat.edu.mx/handle/123456789/2095>

Rivera, J. A., Pedraza, L. S., Martorell, R., Gil, A., & González de Cosío, T. (2019). Obesidad infantil: desafíos y oportunidades para su prevención en América Latina. *Revista Panamericana de Salud Pública*, 43, e32. <https://doi.org/10.26633/RPSP.2019.32>

Salas-González, D., Fernández-García, N. I., & García-Sánchez, J. N. (2021). Uso problemático de la tecnología en niños y adolescentes: una revisión sistemática. *Revista de Psicología Clínica con Niños y Adolescentes*, 8(1), 51-60. <https://revistapcna.com/sites/default/files/09%20Salas.pdf>

Salazar Sánchez, L. M., Politi Martínez, N., Díaz Palacios, L., & Estrada Orozco, K. (2019). Prevalencia de sobrepeso, obesidad y factores de riesgo en una cohorte de escolares en Bogotá, Colombia. *Pediatría*, 52(3), 123–130. <https://revistapediatria.org/rp/article/view/149>

Sánchez-Muñoz, R., Rueda, M. M., & García-Sánchez, F. (2020). Aplicaciones del Big Data e inteligencia artificial en salud pública: una revisión sistemática. *Revista Española de Salud Pública*, 94, e202006052. Disponible en:  
[https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S1135-57272020000100012](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1135-57272020000100012)

Secretaría Distrital de Salud de Bogotá. (2020). Plan de intervención colectiva para la promoción de la salud y la prevención de la enfermedad en Bogotá D.C. Disponible en:  
<https://www.saludcapital.gov.co/>

UNICEF Colombia. (2023). La obesidad y el sobrepeso restan años de vida saludable y generan costos importantes al sistema, a la sociedad y a las familias.  
<https://www.unicef.org/colombia/comunicados-prensa/obesidad-y-el-sobrepeso-restan-vida-saludable-y-generan-costos>

UNICEF. (2022). La obesidad infantil, un asunto de acción. <https://www.unicef.org/colombia/historias/la-obesidad-infantil-un-asunto-de-accion>

Salcedo-Rocha, A. L., et al. (2007). La obesidad infantil: causas, consecuencias, prevención y tratamiento. *Endocrinología y Nutrición*, 54(7), 353–359. [https://doi.org/10.1016/S1575-0922\(07\)71500-0](https://doi.org/10.1016/S1575-0922(07)71500-0)

## Propuesta de implementación de inteligencia artificial y Big Data

Zabala-Vargas, S., Álvarez-Pizarro, Y., Sánchez-Galvis, I., & Rubio-Vásquez, K. (2024). Blockchain-Based Strategy to Optimize Certified Notifications from Government Entities. *Administrative Sciences*, 14(9). <https://doi.org/10.3390/admsci14090195>

Zabala-Vargas, S., & Jaimes-Quintanilla, M. (2025). Tecnologías 4.0 (IOT y ciencia de datos) orientada a optimizar la gestión de proyectos de construcción. *European Public & Social Innovation Review*, 10, 1-21. <https://epsir.net/index.php/epsir/article/view/1621>

Zabala-Vargas, S., Jaimes-Quintanilla, M., & Ramírez-Martínez, D. (2024). PROJECT- BASED LEARNING AND EMERGING TECHNOLOGIES. A STRATEGY TO IMPROVE ACADEMIC PERFORMANCE IN THE TRAINING OF PROJECT MANAGERS. *18th International Technology, Education and Development Conference*, 5642-5646. <https://doi.org/10.21125/inted.2024.1469>

## Anexos

### Anexo 1. Instrumento de Recolección de datos (Encuesta)

# Hábitos que pesan: Encuesta sobre alimentación, tecnología y salud infantil

**B** *I* U ↺ ✕

**Proyecto:** Aplicación de inteligencia artificial y Big Data al estudio de hábitos alimenticios y tecnológicos en niños sedentarios de 6 a 10 años en Engativá, Bogotá.

**Propósito:** Recolectar información sobre los factores que podrían estar relacionados con el sobrepeso y la obesidad infantil.

**Instrucciones:** Marque con una "X" la opción que mejor describa la situación de su hijo(a).

Edad del niño (a) \*

Texto de respuesta corta

Sexo del niño (a) \*

Masculino

Femenino

Tipo de institución educativa \*

Pública

Privada

## Propuesta de implementación de inteligencia artificial y Big Data



Estrato Socioeconomico \*

- 1
- 2
- 3
- Más de 3

¿Con qué frecuencia consume frutas y verduras? \*

- Nunca
- 1-2 Veces por semana
- 3-4 veces por semana
- A diario

## Propuesta de implementación de inteligencia artificial y Big Data

¿Con qué frecuencia consume alimentos ultraprocesados (paquetes, comidas rápidas, gaseosas)? \*

- Nunca
- 1-2 veces por semana
- 3-4 veces por semana
- A diario

¿A qué hora suele consumir su última comida del día? \*

- Antes de las 6:00 pm
- Entre 6:00 y 8:00 pm
- Después de las 8:pm

¿Cuántas horas al día pasa su hijo(a) frente a pantallas (TV, celular, tablet, computador)? \*

- Menos de 1 hora
- 1-2 horas
- 3-4 horas
- Más de 4 horas

¿Para qué utiliza principalmente los dispositivos tecnológicos? \*

- Estudios o tareas escolares
- Juegos y videos recreativos
- Ambos

## Propuesta de implementación de inteligencia artificial y Big Data

¿Utiliza dispositivos electrónicos antes de dormir? \*

Sí

No

¿Cuántos minutos al día realiza actividad física moderada o vigorosa? \*

Ninguno

Menos de 30 minutos

Entre 30 y 60 minutos

Más de 60 minutos

## Propuesta de implementación de inteligencia artificial y Big Data

¿Participa en actividades deportivas dentro o fuera del colegio? \*

- Sí
- No

¿Qué tipo de actividad física realiza su hijo(a)? \*

- Caminatas o juegos livianos
- Deportes organizados
- Ejercicio intenso (fútbol, natación, danza, etc.)
- No realiza actividad física

¿Conoce el peso y talla actual de su hijo(a)? \*

- Sí
- No

⋮

¿Cómo considera usted el estado nutricional de su hijo(a)? \*

- Peso bajo
- Peso adecuado
- Sobre peso
- Obesidad

# Propuesta de implementación de inteligencia artificial y Big Data

## **Anexo 2. Formato de Conocimiento Informado para Padres o Acudientes**

### Anexo 2. Formato de Consentimiento Informado para Padres o Acudientes

---

Título del Proyecto:

Propuesta metodológica para el análisis de hábitos alimenticios, uso de tecnologías y actividad física en niños de 6 a 10 años mediante herramientas de Big Data e Inteligencia Artificial, en la localidad de Engativá (Bogotá).

Investigador Responsable:

Aliss Vanessa Banguero Grueso

CORPORACION UNIVERSITARIA MINUTO DE DIOS - Uniminuto

Correo electrónico: [alissbanguero@gmail.com](mailto:alissbanguero@gmail.com)

### Consentimiento Informado

Por medio de este documento, se solicita su autorización para la participación voluntaria de su hijo(a) en la recolección de información del presente proyecto académico, cuyo propósito es analizar los hábitos alimenticios, el uso de tecnologías digitales y los niveles de actividad física en niños entre 6 y 10 años, con el fin de identificar factores asociados al sobrepeso y la obesidad infantil.

#### 1. Propósito del Estudio

El estudio busca recolectar datos que permitan comprender las relaciones entre alimentación, actividad física y uso de tecnología en niños, para contribuir a futuras estrategias preventivas de salud infantil.

#### 2. Procedimiento

Se le solicitará responder una encuesta estructurada sobre los hábitos de su hijo(a), la cual incluirá preguntas relacionadas con:

- Datos sociodemográficos
- Alimentación
- Actividad física
- Uso de tecnología
- Percepción sobre el estado nutricional del menor

La encuesta será de carácter anónimo y tomará aproximadamente 15 a 20 minutos.

#### 3. Confidencialidad

Toda la información recolectada será manejada bajo estrictas normas de confidencialidad, de acuerdo con la Ley 1581 de 2012 sobre protección de datos personales.

Los datos serán usados exclusivamente con fines académicos y de investigación, y en ningún momento se publicarán nombres ni datos que permitan identificar a los participantes.

#### 4. Riesgos y Beneficios

## Propuesta de implementación de inteligencia artificial y Big Data

No existen riesgos físicos o psicológicos asociados a la participación en esta encuesta.

El beneficio principal es contribuir al conocimiento sobre factores que afectan la salud infantil, lo cual podría orientar futuras estrategias de prevención.

### 5. Voluntariedad y Retiro

La participación es totalmente voluntaria. Usted tiene derecho a no responder la encuesta o a retirarse en cualquier momento, sin que esto genere ninguna consecuencia negativa para usted ni para su hijo(a).

### Declaración de Aceptación

Declaro que he leído y comprendido la información anterior, y autorizo de forma libre y voluntaria la participación de mi hijo(a) en este estudio.

Nombre completo del padre/madre o acudiente:

---

Firma:

---

Documento de identidad:

---

Nombre del niño o niña participante:

---

Fecha: \_\_\_\_/\_\_\_\_/\_\_\_\_