**Projeto final AAED**

Titulo:

Comparación de algoritmos de búsqueda de rutas en grafos urbanos: Un caso de estudio en Alto Selva Alegre, Arequipa, Perú

Abstract:

Las campañas de vacunación antirrábica en áreas urbanas densamente pobladas requieren una planificación precisa para maximizar la cobertura y minimizar los costos operativos. Este trabajo presenta un análisis y comparación de los algoritmos Dijkstra, Bellman-Ford y A\* aplicados al grafo vial del distrito de Alto Selva Alegre en Arequipa, Perú, para determinar distancias caminadas óptimas desde las viviendas hacia puntos de vacunación. Basados en el artículo: “Optimizar la ubicación de los sitios de vacunación para frenar una epidemia zoonótica”, replicamos y optimizamos parte de su metodología utilizando herramientas de código abierto, logrando reducir significativamente los tiempos de cálculo. Este estudio demuestra que el uso de algoritmos eficientes puede superar las limitaciones impuestas por plataformas de cálculo como Leaflet Routing Machine y Mapbox Directions API, ofreciendo una alternativa más ágil y económica para la planificación de campañas de salud pública.

1. Introducción

La rabia es una enfermedad zoonótica que afecta a miles de personas y animales cada año, particularmente en zonas urbanas densamente pobladas donde el control de la población canina es un desafío significativo. Las campañas anuales de vacunación antirrábica son fundamentales para prevenir brotes, pero su planificación presenta limitaciones logísticas y económicas.

Una de las estrategias clave en estas campañas consiste en seleccionar puntos de vacunación óptimos, tanto fijos como móviles. Los puntos fijos permanecen operativos durante todo el día en ubicaciones estratégicas, mientras que los móviles se desplazan por áreas de alta densidad canina donde los puntos de vacunacion fijos no tienen alcance. Sin embargo, establecer un número excesivo de puntos fijos incrementa significativamente los costos, lo que exige un enfoque metodológico para identificar un número óptimo de ubicaciones que maximicen la cobertura y minimicen las distancias caminadas por los residentes.

En este contexto, el artículo de Castillo-Neyra et al. ~\cite{Castillo-Neyra2024}, introduce un método basado en el cálculo de distancias caminadas desde cada vivienda hacia potenciales puntos de vacunación. Utilizando la API de Mapbox Directions, los autores calcularon estas distancias para un conjunto inicial de 70 puntos candidatos, seleccionando finalmente los 20 puntos más eficientes. Sin embargo, las restricciones de su herramienta, como el límite de 20 consultas por minuto, dificultaron la rapidez del análisis, llegando a requerir varios días para completar sus cálculos en ciertos escenarios.

En el presente estudio, replicamos y mejoramos parte de esta metodología mediante el uso de bibliotecas de código abierto como OSMnx y NetworkX, junto con técnicas de clustering empleando K-Means para simular los posibles puntos de vacunación ya que no contábamos con esa información. Además, evaluamos y comparamos el desempeño de tres algoritmos clásicos de rutas cortas: Dijkstra, Bellman-Ford y A\*. Nuestro objetivo principal es demostrar la eficiencia de estas herramientas en términos de tiempo de ejecución, ofreciendo una alternativa viable y más rápida para la planificación de campañas de vacunación.

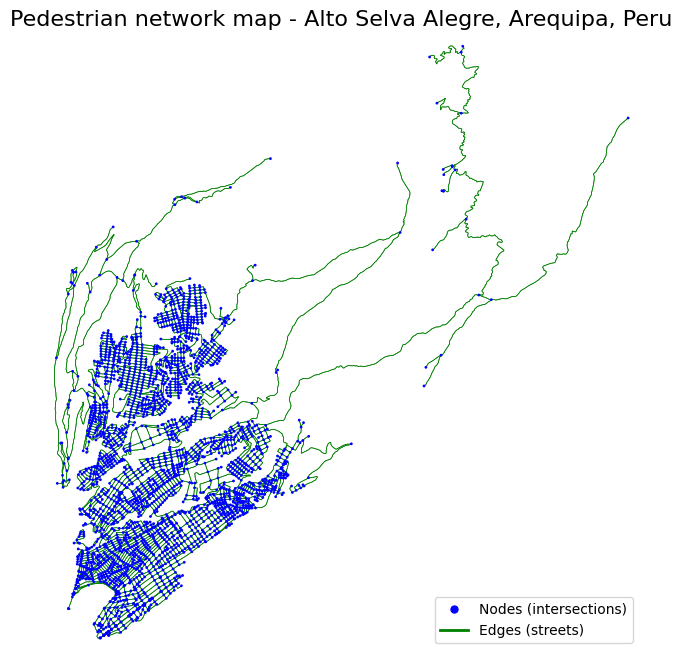
### **2. Metodología**

La metodología empleada en este estudio combina herramientas de código abierto para modelar grafos urbanos, técnicas de clustering para identificar puntos clave y algoritmos de rutas cortas para calcular distancias óptimas. El enfoque permite evaluar el desempeño de los algoritmos considerando el tiempo de ejecución de los resultados. A continuación, se detallan las etapas principales del proceso.

#### **2.1 Generación del Grafo Urbano**

El distrito de Alto Selva Alegre, ubicado en Arequipa, Perú, fue modelado utilizando la biblioteca **OSMnx**, que permite extraer datos viales desde OpenStreetMap (OSM). Esta herramienta es ampliamente reconocida por su capacidad de generar grafos urbanos de alta precisión, representando las intersecciones como nodos y las conexiones entre ellas como arcos ponderados por la distancia geográfica.

Dado que no se dispone de los datos de coordenadas GPS de las viviendas originales del distrito utilizados en el estudio base, se asumirá que cada nodo del grafo representa una vivienda. Aunque esta aproximación reduce la cantidad total de viviendas modeladas en comparación con el número real, resulta adecuada como referencia para los objetivos de este análisis. Con esta aproximación, será posible calcular el tiempo de ejecución del método propuesto por los autores del estudio original. Esto permitirá realizar una comparación directa entre dicho método y los tres algoritmos alternativos propuestos en este trabajo. En la figura 1 es mostrado el grafo con los nodos y vertices.



### **2.2 Identificación de Clustering**

En el artículo base, se seleccionaron 70 posibles puntos de vacunación que fueron determinados por personal del Ministerio de Salud, basándose en su experiencia y conocimiento previo en campañas de vacunación. A partir de estos 70 puntos iniciales, el artículo concluyó con la identificación de 20 puntos de vacunación fijos que optimizan el uso de recursos y maximizan la cobertura.

Para simular los 70 puntos iniciales mencionados en el artículo base, este estudio empleó el algoritmo de clustering **K-Means**, disponible en la biblioteca **Scikit-learn**, para dividir todos los nodos del grafo en 70 grupos. Este enfoque permitió realizar una agrupación basada en proximidad geográfica, proporcionando un método reproducible y objetivo para identificar posibles puntos estratégicos de vacunación.

El proceso se llevó a cabo de la siguiente manera:

1. **Determinación del Número de Clústeres**: Se seleccionó un total de 70 clústeres, en correspondencia con los puntos iniciales descritos en el artículo base. Esta decisión asegura consistencia en la comparación de resultados.
2. **Aplicación del Algoritmo K-Means**: Las coordenadas geográficas de los nodos del grafo se usaron como datos de entrada para el algoritmo. Esto permitió que K-Means formara clústeres que representan comunidades geográficamente coherentes dentro del distrito.
3. **Identificación de Centroides**: Los centroides de los clústeres generados se consideraron como los 70 posibles puntos de vacunación iniciales. Estos centroides se utilizaron posteriormente para calcular las distancias caminadas más cortas desde cada nodo hacia su punto correspondiente.

En la figura 2 se muestran los 70 centroides.

### **2.3 Implementación de Algoritmos de Rutas Cortas**

Para calcular las distancias caminadas desde cada casa (representada como nodo) hacia los posibles puntos de vacunación fijos (los centroides generados previamente), se evaluaron tres algoritmos clásicos de rutas cortas. Cada uno fue seleccionado por sus características específicas, permitiendo una comparación exhaustiva de su rendimiento en el contexto del grafo urbano modelado.

1. **Dijkstra**~\cite{Dijkstra1959}:  
   Este algoritmo calcula las rutas más cortas desde un nodo fuente hasta todos los demás nodos en un grafo ponderado.
   * **Ventajas**: Su simplicidad y robustez lo convierten en un enfoque estándar ampliamente utilizado en problemas de optimización de rutas.
   * **Limitaciones**: Aunque eficiente en grafos pequeños o moderados, su desempeño puede degradarse en grafos densos o grandes debido a la búsqueda exhaustiva en todas las direcciones posibles. Esto es particularmente relevante en el contexto del grafo modelado, que incluye numerosas conexiones viales.
2. **Bellman-Ford**~\cite{Ford1956, Bellman1958}:  
   Similar a Dijkstra, este algoritmo también encuentra rutas más cortas, pero se destaca por su capacidad para manejar arcos con pesos negativos, lo cual lo hace especialmente útil en escenarios donde estos puedan existir.
   * **Ventajas**: Su diseño iterativo garantiza la precisión de los resultados incluso en grafos complejos con arcos negativos.
   * **Limitaciones**: Aunque nuestro grafo no incluye pesos negativos, Bellman-Ford fue usado para evaluar su eficiencia comparativa. Su principal desventaja radica en su mayor costo computacional, ya que revisa todos los arcos del grafo en cada iteración, lo que puede resultar menos eficiente en estructuras densas.
3. **A**\*~\cite{Hart1968}:  
   Este algoritmo combina la búsqueda exhaustiva de Dijkstra con una función heurística que estima la distancia restante hacia el destino, lo que lo convierte en una opción más eficiente en grafos grandes y complejos.
   * **Ventajas**: La heurística aplicada en este caso fue la distancia euclidiana directa entre nodos, lo que permitió una aceleración significativa en los cálculos. Esto es especialmente útil en grafos urbanos densos como el modelado en este estudio.
   * **Limitaciones**: Aunque su rendimiento es superior en términos de velocidad, la calidad de los resultados puede depender en gran medida de la elección y precisión de la heurística.

### **2.4 Análisis y Evaluación de Desempeño**

Para evaluar la eficiencia y consistencia de los algoritmos evaluados, se desarrolló un análisis detallado que abarcó las siguientes etapas clave:

1. **Iteraciones Múltiples**:  
   Cada algoritmo fue ejecutado tres veces, garantizando que los resultados fueran consistentes y reproducibles. Durante cada iteración, se calcularon las distancias desde todos los nodos del grafo hacia los posibles puntos de vacunación (centroides). Este enfoque permitió identificar variaciones en los cálculos y asegurar que el rendimiento no estuviera influenciado por factores externos, como las condiciones del sistema o fluctuaciones de memoria.
2. **Medición de Tiempos de Ejecución**:  
   Se registraron los tiempos de ejecución totales para cada algoritmo al procesar la totalidad del grafo urbano. Estos datos fueron utilizados para realizar una comparación directa del desempeño entre los métodos implementados.

### **3. Resultados**

El análisis del caso de estudio se centró en un grafo con un total de **3031 nodos**, los cuales representan las viviendas del distrito de Alto Selva Alegre. Para cada vivienda, se calculó la distancia más corta hacia los **70 posibles puntos de vacunación fijos** simulados mediante el método descrito previamente. Esto resultó en un total de **212,170 consultas individuales** para determinar las rutas óptimas.

El enfoque del artículo base, que realiza un máximo de **20 consultas por minuto**, requeriría aproximadamente **176.81 horas (7.37 días)** para completar todas las consultas. Este tiempo fue calculado con base en los parámetros descritos en dicho estudio.

#### **3.1 Tiempo de Ejecución**

Los tiempos de ejecución obtenidos para los algoritmos implementados en este trabajo fueron los mostrados en la tabla 1 donde se muestra el resultado de cada iteracion realizada con cada uno de los algoritmos:

TABLA1: En esta tabla se muestran los resultados de las iteraciones realizadas con cada algoritmo.

* **A**\*: **2 horas**.
* **Dijkstra**: **2 horas y 30 minutos**.
* **Bellman-Ford**: **5 horas**.

En comparación con el método utilizado en el artículo base, que requiere más de **7 días**, los resultados obtenidos evidencian una **mejora drástica en el tiempo de ejecución**. Esto pone de manifiesto la ventaja de utilizar algoritmos optimizados y de código abierto para abordar problemas similares. En la tabla 2 se muestran los resultados promedios de los tres algoritmos utilizados junto al resultado que obtendrían los autores del paper base si hubieran realizado sus cálculos con nuestros datos.

TABLA2: En esta tabla se muestra los resultados promedio obtenidos de los tres algoritmos utilizados en este estudio y tambien el tiempo que se obtendría con el método utilizado en el articulo base.

### **4. Discusión**

La implementación de herramientas de código abierto como **OSMnx** y **NetworkX** en este estudio ha demostrado ser una alternativa económica y eficiente para abordar problemas de optimización en grafos urbanos. Estas bibliotecas no solo facilitan la extracción y modelado de datos espaciales de plataformas como **OpenStreetMap (OSM)**, sino que también ofrecen una flexibilidad que resulta especialmente valiosa en proyectos de investigación y aplicaciones prácticas de gran escala.

En contraste, herramientas comerciales como **Mapbox Directions API**, aunque dotadas de funcionalidades avanzadas y soporte técnico, presentan importantes limitaciones en escenarios con alta demanda computacional. Los costos asociados a su uso y las restricciones de cuotas pueden ser prohibitivos para estudios que requieren procesar un gran volumen de datos, como el análisis de rutas óptimas para miles de viviendas.

La evaluación del desempeño de los algoritmos revela que, aunque el algoritmo A\* presenta una destacada combinación de precisión y rapidez gracias a su enfoque heurístico, los análisis recientes indican que el algoritmo de Dijkstra es ligeramente superior en términos de eficiencia para el caso estudiado. Dijkstra demuestra una mayor consistencia al calcular rutas óptimas en grafos densos, lo que lo posiciona como una alternativa ligeramente más robusta para aplicaciones como la planificación de campañas de vacunación y otras intervenciones logísticas a nivel comunitario. Este hallazgo subraya la importancia de evaluar múltiples enfoques para identificar la solución más adecuada según las características específicas del problema y los recursos disponibles.

Por último, la combinación de enfoques reproducibles, herramientas de código abierto y algoritmos avanzados subraya la importancia de democratizar el acceso a soluciones tecnológicas. Estas metodologías pueden ser adaptadas fácilmente a otros contextos urbanos, ampliando su impacto en problemas críticos como la distribución de recursos en áreas vulnerables.

### **5. Conclusión**

Este estudio demuestra la efectividad de los algoritmos de rutas cortas aplicados a grafos urbanos como una estrategia para optimizar la selección de puntos de vacunación en contextos urbanos complejos. La integración de herramientas de código abierto, como OSMnx y NetworkX, junto con metodologías reproducibles, permitió no solo modelar la infraestructura vial con precisión, sino también reducir significativamente los tiempos de cálculo en comparación con otros métodos.

Aunque el algoritmo A\* se presenta teóricamente como una solución ideal debido a su combinación de precisión y rapidez mediante el uso de heurísticas, los resultados obtenidos en este estudio revelaron que el algoritmo de Dijkstra ofrece un desempeño ligeramente superior en el caso específico analizado. Este hallazgo destaca la importancia de evaluar rigurosamente diferentes enfoques para identificar la solución más adecuada según las características del problema y del entorno urbano.

Los resultados obtenidos confirman que el enfoque propuesto no solo es viable, sino también escalable, ofreciendo una solución práctica y económica para la planificación de campañas de salud pública. La flexibilidad de estas metodologías permite su adaptación a otras ciudades o regiones, ampliando su aplicabilidad a diferentes contextos geográficos y logísticos.

Además, el enfoque centrado en el uso de datos abiertos y algoritmos eficientes subraya el potencial de las tecnologías accesibles para abordar desafíos urbanos complejos. Este estudio representa un paso hacia el desarrollo de herramientas basadas en datos que pueden mejorar la toma de decisiones en políticas públicas, especialmente en áreas críticas como la salud, el transporte y la distribución de recursos esenciales.

### **6. Futuras Investigaciones**

El desarrollo de este trabajo abre múltiples oportunidades para investigaciones futuras. Una línea prometedora consiste en explorar modelos híbridos que integren algoritmos de rutas cortas con técnicas de optimización multiobjetivo, lo que permitiría incorporar factores adicionales como la densidad poblacional, la accesibilidad a servicios esenciales y las características socioeconómicas de las áreas estudiadas. Este enfoque podría ofrecer soluciones más integrales, especialmente en contextos donde las decisiones de ubicación deben equilibrar múltiples criterios.

Asimismo, se sugiere realizar un análisis más detallado sobre el impacto de diferentes heurísticas en el desempeño del algoritmo A\*, especialmente en grafos urbanos de mayor escala y complejidad. Evaluar heurísticas alternativas, como las basadas en datos históricos de tráfico o patrones de movilidad humana, podría mejorar la eficiencia del algoritmo y adaptarlo mejor a escenarios dinámicos.

Otra dirección relevante sería establecer un marco para comparar los resultados obtenidos con los reportados en el artículo base, lo que permitiría evaluar la precisión relativa de las soluciones propuestas. Aunque este análisis no fue posible en el presente estudio debido a la falta de acceso a los datos exactos y al tiempo considerable que requeriría dicho esfuerzo, su implementación futura proporcionaría una validación más rigurosa y fortalecería la confianza en los métodos desarrollados.

Finalmente, la implementación de estas metodologías en entornos reales, como ciudades con infraestructuras más complejas o en situaciones de emergencia sanitaria, proporcionaría datos valiosos para validar y refinar los modelos propuestos, ampliando su aplicabilidad y robustez.

### **Referencias**

1. Castillo-Neyra, R., et al. *Optimizing Fixed Vaccination Points in Urban Areas*.
2. Boeing, G. (2017). *OSMnx: A Python package for analyzing street networks*.
3. Dijkstra, E. W. (1959). *A note on two problems in connexion with graphs*.