

(1) Presentación: Buenos días profesores, familia y amigos que me acompañan. Esta es la presentación de mi trabajo de título denominado: “Evaluación de modelos de aprendizaje automático para posicionamiento indoor utilizando Bluetooth low energy”.

(2) Tabla de contenidos

(3) Tabla de contenidos (Introducción)

(4) Introducción

- a) Desde la edad antigua, múltiples formas de localización han sido inventadas para ayudar a referenciar al ser humano, lo cual ha permitido facilitar así el comercio, la navegación y otros aspectos tan básicos y relevantes que hasta el día de hoy son requisitos para el correcto funcionamiento de la civilización.
- b) Dentro de los avances más importantes en este ámbito, es el desarrollo de la teoría científica y técnica denominada georreferenciación, la cual permite el posicionamiento espacial de una entidad en una localización geográfica y única, definida según un sistema de coordenadas y datum específicos.
- c) Gracias a GPS, el crecimiento y acceso de la georreferenciación y navegación está en progresivo aumento, a tal punto que cualquier persona con un smartphone puede saber su posición exacta, con un error de apenas unos pocos centímetros
- d) La búsqueda de una solución en esta memoria nace a partir de la problemática de georreferenciar dentro de una explotación minera o mina. No se puede georreferenciar por los métodos comunes, lo cual establece la necesidad de buscar otras alternativas que presenten resultados favorables en este tipo de entornos y recintos.

(5) Descripción del problema

- a) Algunos ejemplos de los problemas que resuelven estos sistemas son por ejemplo ayudar a encontrar tiendas dentro de un centro comercial, guiar personas discapacitadas visualmente en instalaciones o edificios, ayudar al transporte dentro de minas subterráneas para mover materiales. También se ha planteado sistemas de rescate.
- b) Cuando se usa tecnología GPS dentro de edificios o bajo tierra, existen muchos obstáculos e interferencia que empeoran significativamente la señal y exactitud, o simplemente las señales no pueden ser alcanzadas en el dispositivo.

- c) Las técnicas actuales de posicionamiento \textit{indoor} cuentan con problemas como la reflexión en múltiples objetos como paredes, muebles o el mismo cuerpo humano. Este problema se denomina “multi-path propagation”, y es uno de los principales causantes de error en la localización en interiores. Los métodos que confían plenamente en el indicador RSSI, el cual mide la fuerza de la señal, están sujetos a errores inherentes a la variación de las señales y al ruido.
- d) Problema: Mejorar exactitud de sistemas de posicionamiento en interiores mediante modelos que aprendan de las señales

(6) Objetivos:

- a) Diseñar un método de mapeo para un área mediante señales RSSI (\textit{fingerprint}).
- b) Comparar métodos de aprendizaje automático sobre mediciones RSSI para determinar cuál posee menor error y es más exacto en localización indoor.
- c) Determinar que tanto afectan los métodos de reducción de dimensionalidad tanto en precisión, error y tiempo de procesamiento para los algoritmos de máquinas de aprendizaje estudiados.

(7) Tabla de contenidos – Estado del arte

- a. Ahora corresponde hablar sobre los avances recientes, tecnologías y técnicas matemáticas que permiten la localización en interiores.

(8) Tecnologías para posicionamiento \textit{indoor}

- a) Esta aproximación está basada en el uso de datos provistos mediante cámaras, en formato de vídeo o fotografías. Habitualmente se utilizan dos técnicas, Sistemas de cámara fijos (objeto se mueve y tiene características) y sistemas de cámaras en movimiento ( el ambiente presenta características distinguibles).
- b) Una forma posible de utilizar la radiación infrarroja en sistemas de posicionamiento \textit{indoor} es poner en el objeto o persona a localizar, un transmisor infrarrojo con un identificador único. Luego, los receptores son colocados en lugares dentro del recinto, los cuales pueden detectar este identificador único y comunicar a un software especializado, el cual se encarga de calcular la posición mediante la distancia entre el transmisor y receptor. Costoso y complejo. Ventaja. No se afecta por interferencia electromagnética.
- c) El método consiste en un mapeo \textit{fingerprint} o huella, la cual registra el sonido ambiente de un determinado lugar, y luego construye una base de datos de todas las

mediciones realizadas. Posteriormente para estimar la posición del usuario, realiza una comparativa con su actual fingerprint y lo contrasta con la base de datos, eligiendo de esta manera la posición más “cercana” según el algoritmo de estimación seleccionado. Problemas: velocidad del sonido se ve afectado por temperatura, presión entre otros. Ventajas: bajos costos de implementación.

- d) La localización mediante RFID puede categorizarse en dos tipos, los cuales son localización del lector y localización de tags. Usuario porta un lector y los tags están dispuestos o usuario porta un tag y hay múltiples lectores estratégicamente puestos. Costoso y no escalable. Tags tienen información relativa a la posición. Poco alcance. No necesita LoS.

- (9) RSSI es una escala de referencia para medir el nivel de potencia de la fuerza de la señal recibida por el receptor. Se mide en dBm donde 0 RSSI indica señal ideal y valores más negativos indican mayor pérdida. Tx Power Es la potencia de salida o fuerza de la señal que el emisor produce durante el tiempo de transmisión. A mayor Tx Power, más estable es la señal, pero más energía se consume. Mejores: WiFi no requiere infraestructura donde ya hay, en otros lugares ( mina) costoso, mas popular. Bluetooth estándar WPAN, rango corto- medio, presente en todos los smartphones. Desarrollo de Beacons. Nombrar ZigBee, UWB, FM, GSM.

Ninguna tecnología presentada anteriormente es completamente efectiva en todos los escenarios, por lo que la elección de estas se debe hacer según las características del lugar, como su dinámica de cambio, obstáculos presentes, tamaño, entre otros.

- (10) Técnicas matemáticas Wireless para localización indoor: las que presentan mejores resultados. Por lo mismo es necesario establecer modelos matemáticos que reduzcan el error generado.

- a) Es el método más simple, y se basa en determinar una posición simbólica y aproximada de la posición del usuario.
- b) Antenas o emisores de ondas de radio. Según la señal más fuerte detectada por el usuario, es donde se localiza en el sistema.
- c) Ampliamente usado en redes celulares, ya que permite determinar la posición de un dispositivo con una precisión de 50-200 m, sin embargo, no es buena en espacios reducidos. GSM, Infrarrojo, Cell-ID.

- (11) Triangulación: La triangulación se divide en 2 puntos: lateración y angulación. Por una parte, la lateración estima la posición midiendo las distancias hacia múltiples puntos de referencia también llamados beacons. Angulación por su parte, no utiliza las distancias, sino que los ángulos relativos a cada beacon o punto de referencia. El principal problema es establecer la distancia para lateración, y los ángulos para angulación. Para ello se han diseñado muchos acercamientos que permiten obtener esos valores.

- (a) Técnicas de lateralización: ToA (Time of arrival), TDoA (Time difference of arrive). Ambos sufren de multicaminos y línea de visión. Método basado en RSS: modelos teóricos y empíricos, los cuales intentar predecir la distancia utilizando la diferencia entre la fuerza de la señal entre el transmisor y el receptor.
- (b) Técnicas de angulación: AoA ( Angle of arrival): es complejo y difícil de implementar, además de requerir software especializado, ya que para una localización muy precisa el ángulo debe ser medido con muy poco error y esto es difícil de lograr.

(12) Fingerprint: Fingerprint presenta dos etapas, la etapa offline y la etapa online. Durante la etapa offline se realiza un reconocimiento de las características del lugar, en donde se mide las coordenadas de cada estación base y la fuerza de la señal en determinados lugares, obteniendo así un mapa de señales del entorno. Durante la etapa online se utiliza la señal actual recibida de cada estación base y un modelo generado en la etapa offline, generando así una estimación de la posición del usuario. Para la estimación de la posición se han realizado variados acercamientos como es modelos de probabilidad, estimación de máxima verosimilitud, estimación bayesiana. medir cada punto del lugar en donde se realiza la localización. Mientras más puntos se registran, mejor será el radio-map ya que tiene un mayor número de datos de entrenamiento. Acercamiento determinista, probabilista y reconocimiento de patrones.

- (a) Motivación elección: Wifi y otras ya han sido muy exploradas, Bluetooth más fácil de implementar y en constante desarrollo. Fingerprint se selecciona porque no se influencia por el ruido, no depende de estimación de distancias, además algoritmos de machine learning son muy avanzados y pueden ayudar reconocer patrones en las señales, comparando algoritmos contemporáneos (Deep learning).

(13) Propuesta de solución:

(14) Propuesta:

Establecer un marco de trabajo para la recolección, entrenamiento y clasificación de algoritmos de machine learning utilizando Bluetooth Low Energy.

Comparación de diferentes clasificadores.

Utilizar técnicas de reducción de dimensionalidad.

Utilizar modelos sin necesidad de conexión a internet.

(15) Consideraciones Previas: Beacons

- a. Hablar y explicar Beacons, específicamente protocolos. Transmiten en 2.4 GHz, low energy. WiFi consume mas de 40 veces más energía que bluetooth.

(16) Leer

(17) Cabe destacar que los algoritmos de máquinas de aprendizaje a pesar que pueden resolver para problemas multivariados, es decir, predecir para dos variables, en este caso  $\$(x,y)\$$  como un punto; es mucho más sencillo separar esto en dos problemas, vale decir, un algoritmo de máquina de aprendizaje para  $\$x\$$  y otro para  $\$y\$$ .

(18) Entrenamiento

- (a) Entrenar técnicas de máquinas de aprendizaje muy conocidos y que han presentado buenos resultados a lo largo de muchos problemas.
- (b) Posteriormente, se seleccionan los mejores algoritmos, es decir, que presenten el mejor desempeño y luego son implementados.
- (c) Reducción de dimensionalidad.
  - a. Esto no ha sido mayormente explorado en la literatura.
  - b. Existe correlación espacial lineal de las señales adyacentes. PCA ayuda a eliminar esta correlación.
  - c. Los métodos de extracción de características pueden ayudar a agilizar la fase de entrenamiento, ya que este proceso es lento. Además, al ser menos componentes, en la fase online, las técnicas tardaran mucho menos tiempo en determinar la posición en tiempo real.
- (d) ¿Implementación en cliente o servidor? Resulta natural en cliente, ya que no siempre existe conexión en este tipo de lugares.

(19) Fase Online:

- (a) Para la fase online se reconocen dos etapas principales.
- (b) Colectar un vector de señales RSSI en la posición actual del usuario, es decir, el vector de intensidad de la señal.
- (c) Proveer este vector de entrada a los algoritmos de aprendizaje supervisado
- (d) Una vez que los algoritmos de clasificación proveen el resultado de la posición física, entonces la misma aplicación de la fase offline, es utilizada para mostrar en un mapa de tiempo real la localización actual de usuario.
- (e) Para realizar esta tarea se deben tener en cuenta las normalizaciones realizadas y aplicar correctamente la transformación PCA.

(20) Proceso de desarrollo: los datos son normalizados previamente antes de ser utilizados. Además, la razón de la conexión entre la normalización y el entrenamiento de los clasificadores se debe a que se utilizaran los algoritmos de machine learning utilizando

PCA y no utilizándolo, a modo de comparación en términos de tiempo y accuracy. Además, se ha omitido el ítem en donde los clasificadores ya entrenados son portados al dispositivo móvil. Finalmente, hay que notar que, en la fase online, el dispositivo móvil genera un nuevo vector de señales RSSI, luego se transforman los datos y finalmente se clasifican, lo que genera una posición estimada que se refleja en el dispositivo móvil. Este ciclo es constante, ya que la posición se actualiza continuamente según los parámetros definidos y frecuencia de actualización.

## (21) Experimentación

(22) Los parámetros más relevantes en este caso serían el TX Power y el Beacon Interval. En un despliegue real, lo más relevante es disminuir los costos, que en este caso están asociados a las baterías o más bien al cambio de estas a lo largo del tiempo. El valor de TX Power afecta principalmente a tres factores claves, estos son el rango de la señal, es decir, la distancia máxima alcanzada; el segundo corresponde a la estabilidad de la señal y finalmente la batería. Mientras menor es el intervalo, mayor es la estabilidad de la señal, además el intervalo afecta significativamente a la posición en interiores, debido a que mientras más rápido se mueve un usuario, si el intervalo es demasiado alto, su posición sufrirá saltos o cortes que no muestran realmente la ruta de desplazamiento. TX Power correspondiente a 7, y un valor del intervalo igual a 100ms.

(23) Lugar de experimentación: Estacionamiento subterráneo de la universidad Técnica Federico Santa María, Campus San Joaquín. Este lugar es ideal para realizar pruebas, ya que presenta interferencia debido a los vehículos en tránsito y la disposición de los objetos es cambiante. Las dimensiones constan de \$144.75m\$ de largo y \$36m\$ de ancho.

(24) Android: explicar pasos. Sklearn : Para la elaboración de modelos de clasificación, una de las librerías más utilizadas es scikit-learn o también conocida como SKlearn scikit-learn, la cual es una completa API para elaborar y realizar maquetas de machine learning. A pesar de que SKlearn presenta tantas ventajas para el desarrollo ágil de modelos de máquinas de aprendizaje, no es tan bueno en redes neuronales, y no implementa la mayor parte de los algoritmos modernos en este campo. Además, por la forma en que Tensorflow representa los datos, es decir, tensores o arreglos multidimensionales, y las operaciones referentes a grafos sin estado que transforman estos tensores, lo vuelve increíblemente rápido para entrenar y desplegar, incluso presenta API para dispositivos móviles en Java, C++ y GO .Core en c++.

(25) Leer la diapo, explicación posicionamiento Beacons y puntos de medición. Explicar mediciones. Un punto importante a definir es que valor asignar a los Beacons que no han sido detectados, en este caso y como el valor RSSI no puede ser mayor a cero, se decide asignar un valor de 100 dBm como un identificador de que no se ha detectado señal. Mediciones en diferentes días, 150 por posición.

(26) Entrenamiento de clasificadores: Para la gran mayoría de los algoritmos es necesario una normalización previa. Los clasificadores seleccionados en para este trabajo son Naive Bayes, SVM con Kernel de tipo Radial Basis Function (\textbf{RBF}), SVM con Kernel lineal, K-NN, Arboles de decisión, Random Forest, Adaboost, Neural network y QDA. Los parámetros de cada algoritmo son configurados según una búsqueda en grilla. 5-fold cross-validation con 20% de los datos. Se debe señalar que los gráficos son para la clasificación sobre la coordenada  $XX$ , sin embargo para  $YY$  los resultados son similares. Se usa similaridad de Jaccard o índice de Jaccard. Se debe notar que existen dos redes paralelas, esto se debe a que las redes neuronales para predecir la clase de X y la clase de Y se entrenan simultáneamente. Para red neuronal se utiliza test y train set del 20%.

(27) Hablar de red neuronal

(28) Con toda la información presentada, es claro que los mejores resultados se presentan principalmente en 3 algoritmos, estos son KNN, SVM con kernel RBF y Redes Neuronales profundas. Lo anterior demuestra que los datos efectivamente tienen estructuras no lineales, ya que K-NN no depende específicamente de la estructura de los datos, ya que busca un radio o vecindario de clasificación. SVM RBF define áreas de clasificación que precisamente son no lineales, por lo que puede buscar mejores representaciones en este tipo de estructuras que los clasificadores lineales. La forma en que modela estas fronteras no lineales es debido a la elección del kernel RBF el cual por su función de radio basal es flexible en torno a este tipo de regiones. Por último, las redes neuronales pueden aprender de datos no lineales debido a sus funciones de activación que precisamente son no lineales, además de poder aprender patrones más detallados en cada capa, por su capacidad de reducir la complejidad y dimensionalidad de los datos. costos que se deben considerar, en esta suerte de trade-off debido a la pérdida de precisión en los clasificadores por malos hiperparámetros, versus el costo de configuración previa de ellos en la fase de entrenamiento.

(29) Tres métodos: varianza o valores propios mayor a 1, varianza acumulada o información retenida mayor a 80%, y clasificadores.

(30) Según la comparativa, es claro que la mayoría de los clasificadores presentan una gran mejoría hasta la quinta componente, y posteriormente la ganancia no es significativa respecto a la información aportada a los respectivos algoritmos. Esto indica al igual que los otros métodos de selección, que a partir de la quinta componente la varianza aportada no es realmente significativa, por lo que puede reducirse la dimensionalidad de los datos a 5

(31) Los valores obtenidos entonces demuestran que a pesar de que la accuracy se reduce solo un poco, los valores de error medio decrecientan significativamente en el contexto del problema abordado, ya que para el posicionamiento en interiores se espera un error lo más pequeño posible, ya que esto significa una mejor localización. Por lo anterior, se debe tener en cuenta estos factores al momento de seleccionar las componentes principales, ya que, al perder información, se pierde exactitud en las mediciones, lo que repercute significativamente al momento de realizar pruebas reales. El comportamiento en términos

generales sigue el mismo patrón que al no aplicar técnicas de reducción de dimensionalidad, es decir, los clasificadores mantienen el orden relativo en cuanto a accuracy y nuevamente los clasificadores capaces de distinguir patrones no lineales en los datos son los dominantes, lo cual prueba estas características de los datos. También se debe notar que los tres primeros lugares corresponden nuevamente a redes neuronales profundas, máquina de vectores de soporte con kernel RBF y finalmente vecinos más cercanos con  $k=2$ . Por lo anterior, se decide utilizar los tres clasificadores NN, SVM y KNN para realizar la experimentación en el estacionamiento descrito anteriormente.

(32) Explicar método estático y dinámico.

(33) Como se observa en la imagen \autoref{fig:fase\_online}, en primer lugar se debe elegir el algoritmo a utilizar en el posicionamiento, luego seleccionar el patrón para elegir el punto en donde se realizarán las pruebas y finalmente iniciar, lo cual comienza a posicionar al usuario en tiempo real. A pesar de que se selecciona un algoritmo para la visualización del posicionamiento, internamente la aplicación ejecuta todos los algoritmos mostrados, es decir NN, SVM y KNN con sus respectivas versiones utilizando PCA. El fin de esto es ahorrar tiempo, ya que de esta forma se pueden obtener los resultados para todos los algoritmos concurrentemente sin necesidad de ejecutarlos de manera secuencial. Luego la aplicación tiene la capacidad de crear archivos de \textit{log}, es decir, guardar los resultados obtenidos, y el tiempo de ejecución.

(34) Sección resultados.

(35) Los resultados obtenidos en el método dinámico sin utilizar PCA. Se debe considerar que todos los valores están en metros. Dinámico sin pca: los mejores resultados en términos generales son obtenidos por el método KNN, ya que los promedios en general son bajos, así como el RMSE. Dinámico PCA: En este caso, los resultados son ligeramente mejores en términos de RMSE, lo cual indica que PCA si puede eliminar parte de la correlación espacial. En este caso nuevamente KNN con PCA obtiene el mejor resultado, seguido por Neural Network y finalmente SVM. Las varianzas de cada eje igualmente son relativamente altas, por lo que esto puede afectar mucho el posicionamiento en general, ya que, a mayor varianza, más error durante el tiempo se presenta, y más alejados están los valores desde el promedio.

(36) Método estático sin pca: Para este caso, la \autoref{tabla-estatica}, muestra que ahora los mejores resultados son obtenidos por el método Neural Networks con un RMSE de 4.4857 en promedio. Para todos los algoritmos, las varianzas se ven reducidas sustancialmente, lo que es esperable debido a que, al estar estático en una posición, las ondas electromagnéticas no se ven alteradas o cambiantes a través del tiempo, lo cual estabiliza la señal recibida por los Beacons. Estático con PCA: La \autoref{estatico-pca} muestra mejores resultados para todos los clasificadores, disminuyendo así los valores totales. Esto muestra nuevamente que mediante la técnica PCA es posible reducir el error. Por otra



parte, las varianzas permanecen pequeñas a excepción de Neural Network. El mejor resultado es obtenido por Neural Network con PCA con un error promedio de 3.9341 metros, seguido por KNN PCA y finalmente SVM PCA.

(37) CDF dinámico: el cual está definido como el valor de una variable aleatoria  $X$ , o su función de distribución, que al ser evaluada en  $x$ , es la probabilidad que  $X$  tome valores menores o iguales a  $x$ . **Explicar graficos. Para 100%**, Como se observa en la \autoref{cambio-cdf-dinamico}, todos los valores son mejorados utilizando PCA, además los errores menores se obtienen con NN y KNN, aunque KNN es más estable ya que con y sin PCA presenta valores semejantes.

(38) 100%: Los cambios para KNN y SVM no son significativos, y además los errores mínimos son muy similares, del orden de los 16 metros, lo cual indica que PCA funciona efectivamente para reducir la información redundante. Con respecto a NN, se presenta una anomalía al utilizar o no PCA, esto se debe a que las redes neuronales ya hacen una reducción de la dimensionalidad en sus capas escondidas, y retienen lo mejor de la información en cada paso, por lo que aplicar PCA no ayuda demasiado

(39) Análisis de distribución: NN se comporta de forma anómala, los valores no cambian significativamente solo outliers.

(40) Tiempos de ejecución: La tendencia es sumamente clara, PCA ayuda a disminuir el tiempo de procesamiento significativamente. KNN aumenta significativamente su tiempo de computo, ya que itera sobre cada registro del dataset en una búsqueda exhaustiva, por lo que utilizar PCA con KNN es un requisito. SVM por su parte reduce su tiempo de computo en un 53.0966\% lo cual es una ganancia muy significativa. Esto se debe a que el tiempo de entrenamiento de SVM kernelizado es cuadrático, por lo mismo demora mucho en entrenarse, sin embargo su tiempo de ejecución es lineal según el número de vectores de soporte, y lineal en el número de características. NN La mejora es de un 24.0867\%, lo cual es bastante, pero si se observan los valores, el cambio no es drástico para el sistema de posicionamiento

(41) \item Resultados estáticos son mucho mejores que los resultados dinámicos, sin embargo, el escenario de que el usuario este estático en un punto no es para nada realista  
\item Primero, los mejores valores de error medio son obtenidos por KNN y NN, en ambos métodos (estático y dinámico).

\item KNN es mucho menos disperso en ambos métodos y sus errores están más centrados en valores bajos, mientras NN presenta mucho mayor dispersión en el método dinámico, pero casi nada en el método estático, sobre todo al no utilizar PCA.

\item Mejor algoritmo es redes neuronales, a pesar de su distribución, mantiene valores bajos de error y tiempos de procesamiento.

(42) Conclusiones

(43)

- a. Las señales Bluetooth se ven afectada profundamente por cualquier objeto que se interponga, en este caso se ha probado particularmente con una persona, y como

es bien sabido, gran parte del cuerpo humano es agua, con lo cual la señal percibida en el receptor decae, por lo que esto afecta negativamente a los algoritmos y métodos matemáticos, ya que la información porta ruido e interferencia. Esto se nota cuando existen mucho flujo de personas.

- b. Es claro que los datos recolectados presentan estructuras no lineales, pero con correlaciones lineales en sus vecindarios, por lo que en general, los mejores algoritmos de máquinas de aprendizaje son aquellos capaces de reconocer estas estructuras.
- c. A partir de la experimentación en la fase online, se determina que el mejor algoritmo para el posicionamiento corresponde a Redes Neuronales artificiales, mediante un análisis de CDF, Boxplot y medidas de error medio como RMSE.
- d. Con respecto al uso de PCA, es claro que utilizarlo es definitivamente la mejor opción, sobre todo en algoritmos como KNN y SVM, debido a que reduce sustancialmente el tiempo de computo, y además mejora la precisión, disminuyendo el error presente en los algoritmos. A pesar de lo anterior, y de que PCA disminuye el error, al utilizarlo en redes neuronales se puede observar que su rango de dispersión aumenta significativamente, e incluso en el método estático presenta peores valores de error, por lo que para redes neuronales artificiales, aunque eventualmente pueda mejorar el tiempo de computo, es mejor no utilizar PCA.
- e. Los análisis indican entonces que utilizar Bluetooth con técnicas de máquinas de aprendizaje, pueden reducir el error a unos pocos metros, particularmente el mejor valor encontrado lo obtienen las redes neuronales artificiales con 3.9341 metros, lo cual es relativamente alto si se considera un posicionamiento en tiempo real, preciso y sin grandes errores, como es de esperar. Esto indica entonces que el sistema presentado en este trabajo puede ser la base para estimar la posición asociada a una región o zona geográfica de un recinto interior, pero no para determinar efectivamente la localización en tiempo real.
- f. Quedan muchas incógnitas abiertas, principalmente análisis de la distribución y densidad del posicionamiento de los Beacons Bluetooth, es decir, estudios sobre cuantos Beacons utilizar sobre una determinada región. Por otra parte, el análisis de la grilla.
- g. Con respecto a el modo de recolección de fingerprints, esto es aún un problema para esta técnica, ya que en espacios reducidos, en la fase offline la recolección es mucho más rápida, sobre todo con grillas de tamaño grande, sin embargo, si el recinto contiene muchos puntos de referencia, la recolección es lenta y requiere mucho esfuerzo. Actualización de fingerprints. clasificadores no supervisados, es decir, reconocimiento de patrones sin necesidad de indicar la posición, esto reduce considerablemente el esfuerzo para la recolección de fingerprint, ya que en la fase de entrenamiento no es necesario pasar al algoritmo la etiqueta de cada punto, es decir, la posición, sino que es el mismo algoritmo el que se encargaría de reconocer patrones en las señales

- h. A partir de los estudios realizados, se puede determinar que las máquinas de aprendizaje en conjunto con las señales Bluetooth Low Energy pueden ser un punto de partida para mejorar la precisión y disminuir el error del posicionamiento en interiores, en particular el Deep Learning, con el framework Tensorflow, el cual puede ser portado fácilmente a cualquier dispositivo móvil, con lo que es factible utilizar, además de ser sumamente rápido una vez que las redes están entrenadas. Fusión de sensores.
- i. Como conclusión final, se debe notar que en este caso solo se utiliza un lugar de experimentación, en particular, un estacionamiento, con lo que estos modelos entrenados no funcionarían en otros recintos. Este es el mayor problema del posicionamiento en interiores, lograr un modelo estándar, que funcione relativamente bien en gran parte de los escenarios y no dependa específicamente del lugar de experimentación.

Valores propios, Varianza explicada y Clasificadores