



Universidad del Desarrollo
Facultad de Ingeniería

Clasificador del nivel del estudiante, en el simulador de realidad virtual para procedimientos de laparoscopia.

POR: MIGUEL ANGEL DIAZ MOYA

Proyecto de grado presentado a la Facultad de Ingeniería de la Universidad del
Desarrollo para optar al grado académico de Magíster en Data Science

PROFESORES GUÍA:

Dr. JOSE IGNACIO GUZMAN MONTOTO
MSC. CAMILO RODRIGUEZ-BELTRAN

Santiago, Enero 2022

TABLA DE CONTENIDO

| | |
|---|-------------------------------|
| RESUMEN..... | 3 |
| 1. INTRODUCCIÓN | 4 |
| 2. HIPÓTESIS Y OBJETIVOS..... | 6 |
| 3. TRABAJOS RELACIONADOS..... | 7 |
| 4. DESCRIPCIÓN DE DATOS Y METODOLOGÍA | 9 |
| 4.1. CONTEXTO DEL ESCENARIO VIRTUAL | 9 |
| 4.2. CATEGORIZACIÓN DE LOS USUARIOS | 10 |
| 4.3. GENERACIÓN DEL DATA SET | 10 |
| 4.4. DATOS | 11 |
| 4.5. METODOLOGÍA | 16 |
| 5. ANÁLISIS Y MODELAMIENTO DE DATOS | 17 |
| 5.1. LIMPIEZA Y TRATAMIENTO DE DATOS | 17 |
| 5.2. ANÁLISIS ESTADÍSTICOS DE LOS DATOS | 18 |
| 5.3. MODELAMIENTO DE DATOS | 19 |
| 5.4. DESARROLLO DEL MODELO | 20 |
| 5.4.1. <i>Definición del Modelo</i> | 20 |
| 5.4.2. <i>Configuración del Modelo</i> | 21 |
| 5.4.3. <i>Función de Pérdida</i> | 21 |
| 5.4.4. <i>Optimizador</i> | 22 |
| 5.4.5. <i>Entrenamiento del Modelo</i> | 22 |
| 5.4.6. <i>Overfitting</i> | 23 |
| 5.4.7. <i>Matriz de Confusión</i> | 23 |
| 5.4.8. <i>Métricas</i> | 24 |
| 5.4.8.1. <i>Exactitud</i> | 25 |
| 5.4.8.2. <i>Precisión</i> | 25 |
| 5.4.8.3. <i>Sensibilidad, exhaustividad</i> | 26 |
| 5.4.8.4. <i>Especificidad</i> | 26 |
| 5.4.8.5. <i>F1 Score</i> | 26 |
| 5.4.9. <i>Evaluación del Modelo con los datos de Prueba</i> | 27 |
| 6. RESULTADOS | 28 |
| 7. CONCLUSIONES | 30 |
| 8. TRABAJO FUTURO Y RECOMENDACIONES | 31 |
| CATEGORIZACIÓN DE LOS USUARIOS | ¡ERROR! MARCADOR NO DEFINIDO. |
| BIBLIOGRAFÍA | 32 |

Te haces Magister una sola vez en la Vida...
José Ignacio Guzmán Montoto

Agradecimientos

Luego de recorrer este camino de esfuerzo y dedicación, con los costos personales, familiares y laborales que esto conlleva, dedicando tiempo y atención en cada una de las exigencias que mi casa de estudios me ha propuesto a lo largo de mi proceso de formación de post grado para optar al Título de Magister en Data Science. Quiero agradecer en primer lugar a Dios y a la vida por darme la oportunidad que no todos tienen de sumar una especialización a mi carrera profesional, me siento muy afortunado y agradecido por este hito en mi vida.

Sin lugar a duda agradecer a *Mi Esposa Evelin*, por ser siempre mi apoyo incondicional en cada uno de los desafíos que me planteo.

A *Mi Familia*, por la formación y dedicación que han puesto en mí, para llegar a ser quién soy.

A *José Ignacio Guzmán*, por su dedicación, calidad profesional y humana, por guiar mi investigación de Magister en Data Science, preocupándose de que cada detalle marque la diferencia para lograr un trabajo del mejor nivel.

A mis profesores del Programa de Magister de la Universidad del Desarrollo, por su dedicación y compromiso en la formación de Profesionales especialistas en Ciencia de Datos.

Santiago, 25 de enero 2022

M.D.

Resumen

La Dirección de Innovación de la Facultad de Ingeniería de la Universidad del Desarrollo, llevó a cabo el desenvolvimiento de un prototipo de simulador en realidad virtual inmersiva para cirugía de mínimo acceso portátil.

A través del cual, permite a los estudiantes realizar actividades motrices en las cuales se registra la información de orientación y posición de sus acciones a lo largo de la actividad, generando un informe automático en donde se retro-alimenta al alumno con cantidad de acciones, válidas e inválidas en distintos eventos durante cada sesión de entrenamiento.

Con el avance de la tecnología, los simuladores clínicos han demostrado ser una herramienta efectiva para el desarrollo de las habilidades quirúrgicas necesarias que debe poseer un médico especialista; A través de los entrenamientos en el simulador es posible almacenar y procesar los datos, lo que permite mejorar el proceso de aprendizaje del alumno en la técnica de laparoscopia inmersiva de mínimo acceso.

La clasificación es un problema de modelado que asigna una etiqueta de clase tipo discreta a cada fila de registro, en la practica la asignación de este identificador de clase se utilizará para la clasificación del practicante, pero es importante resaltar que este modelamiento de la información tiene la capacidad de producir una predicción de valores continuos.

En el aprendizaje supervisado los predictores y las variables de respuesta son conocidas para construir modelos matemáticos con la intención de predecir o clasificar observaciones que se obtengan posteriormente. Estos métodos se consideran supervisados debido a que se construyen con valores etiquetados, es decir, la máquina “aprende” de los datos con el propósito de predecir resultados futuros. El algoritmo de aprendizaje categorizará mediante la diferenciación con respecto al tipo de la variable de salida involucrada en nuestro problema.

A través del modelamiento de la data generada en una sesión de entrenamiento, es posible entregar al usuario, una calificación objetiva de su nivel de avance, en base a sus propios registros de entrenamiento en el simulador. Y así, aumentar la exactitud con la que se miden los avances de cada alumno, y no solo depender de la supervisión de un médico especialista. De esta forma, podemos mejorar la retroalimentación de los alumnos en la adquisición de las habilidades necesarias para mejorar su desempeño en técnicas de laparoscopia inmersiva.

1. Introducción

La cirugía de laparoscopia es un procedimiento quirúrgico de mínima invasión para el paciente, en la cual se ha logrado acortar el tiempo de recuperación de este, durante el post operatorio, disminuyendo tanto el tiempo de recuperación, como el riesgo de infección, ya que la inserción es menor que en una cirugía convencional. Por otra parte, el gran desafío se encuentra en alcanzar altos niveles de preparación para estudiantes en esta área, ya que por ser una cirugía indirecta y por la imagen que se proyecta en dos dimensiones en una pantalla, el especialista debe adquirir las destrezas necesarias para no perder el sentido de profundidad de la cavidad que se encuentra interviniendo quirúrgicamente.

En el método convencional, los médicos presencian una intervención quirúrgica de laparoscopia realizada por parte de un especialista, pero sin intervenir, ya que en los centros médicos no se está permitido, dado el alto riesgo para el paciente y la complejidad que esto conlleva. Por otra parte, las habilidades que posee el médico especialista en el procedimiento laparoscópico no son transferibles a un cirujano solo con la exposición de la teoría, sino que deben ser adquiridas a través de la práctica.

Los simuladores deben presentar escenarios que acerquen a los estudiantes a un ambiente lo más real posible, y permitan adquirir las habilidades necesarias en un ambiente cercano al del trabajo quirúrgico. Por esta razón es que varios han sido los esfuerzos por crear simuladores que aumenten la velocidad con que el alumno adquiere las destrezas necesarias en esta materia.

Dentro de algunos tipos de simuladores, se encuentran los modelos de banca, los cadavéricos, los animales y de los de realidad virtual. Los modelos de banca, también denominados Box Trainers, consiste en una caja opaca, y son muy útiles para desarrollar destrezas básicas y suturas. En el caso de los entrenamientos con cadáveres, estos tienen la dificultad de que los tejidos no tienen la elasticidad propia de un ser vivo, además de la falta de irrigación sanguínea dentro del área de trabajo, por lo que su comportamiento dista de un escenario como si fuese un paciente vivo.

Por otra parte, los entrenamientos con animales no se han fomentado, debido al dilema ético, al igual que el entrenamiento con simuladores cadavéricos.

Finalmente, la simulación de estos procedimientos quirúrgicos con realidad virtual permite interactuar en tiempo real con escenarios predefinidos, y adquirir habilidades que en otros casos podrían tomar meses e incluso años.

Los simuladores disponibles actualmente en Chile están divididos en dos grupos. Uno basado en una maqueta, donde el cirujano actúa con los instrumentos reales "agarrando" objetos y trasladándolos de un origen a un destino dentro de un recipiente quirúrgico. En este caso se utilizan cámaras para la observación de la escena. Si bien utiliza durante el entrenamiento los instrumentos reales y realiza un contacto con objetos,

tiene dificultad en lograr la repetitividad de los ejercicios, la falta de portabilidad del sistema y la ausencia de un almacenamiento y análisis de la actividad del alumno durante la simulación. Un segundo grupo son simuladores importados, que utiliza la realidad virtual.

Los simuladores virtuales existentes en el mercado por su parte resuelven este problema, sin embargo, carecen de la facilidad de poder ser trasladados ya que están diseñados fundamentalmente para laboratorios. La movilidad de este tipo de simuladores, permite realizar entrenamiento a distancia, pudiendo corregir y retroalimentar al usuario desde un lugar remoto acerca de los avances y correcciones en cada sesión de entrenamiento, sin tener que desplazarse a un laboratorio para entregar su feedback presencialmente.

No obstante, la supervisión y retroalimentación con la que cuentan los simuladores modulares y de realidad virtual se basa en la observación de cada una de las sesiones por un profesor especialista, que luego de concluir el entrenamiento, genera un informe con indicadores como conteo de acciones válidas y erradas, tiempo de la sesión, para así corregir al alumno en base a resultados absolutos que carecen de haber profundizado en los datos generados en su sesión de entrenamiento.

Se puede mejorar aún más la calidad de los reportes, realizando un análisis en profundidad de los datos recolectados por el simulador a través de la sesión de entrenamiento, ya que cada estudiante tiene un propio patrón de comportamiento en cada sesión, por lo que es posible entregar un reporte acabado de las acciones y puntos de mejoría totalmente personalizable.

El simulador de realidad virtual, con análisis de resultados de cada sesión personalizados, aumenta considerablemente el progreso del usuario, permitiendo encontrar errores o modificando patrones que visualmente es difícil de corregir, ya que la percepción del entrenador puede variar; Es así como incorporando análisis de datos bajo un programa desarrollado en Python, podemos establecer los lineamientos en base a los datos de entrada que se generan en cada una de las sesiones de entrenamiento, y emitir información gráfica e indicadores, que permitan corregir de manera efectiva las desviaciones en el proceso de formación, de los futuros especialistas en procedimientos de laparoscopia.

La necesidad de una preparación avanzada de los cirujanos es aplicable a cada estudiante de postgrado de la especialidad de cirugía en Chile. Además, el sistema basado en el simulador que se propone, puede ser utilizado por todos los cirujanos activos, que puedan adquirir el producto para mantener sus habilidades o pasar cursos de superación a distancia sin necesidad de trasladarse a las salas de capacitación o al menos reduciendo de esta forma de una manera considerable las clases presenciales.

2. Hipótesis y Objetivos

Hipótesis

Si se utilizan técnicas de análisis de datos para construir un modelo de clasificación, será posible obtener el nivel de avance de un estudiante que se prepara en el simulador con realidad virtual para la cirugía de mínimo acceso.

Objetivo general

Obtener un modelo basado en el análisis de datos que permita retroalimentar sobre el nivel de desempeño en una sesión de entrenamiento.

Objetivos Específicos

- Programar el procesamiento y limpieza de datos necesario para que el modelo implementado puedan recibir los datos adecuadamente desde los archivos del simulador.
- Describir el frame de datos que se utilizará en cada sesión de entrenamiento de acuerdo a las especificaciones realizadas por los médicos cirujanos de la facultad de medicina de la Universidad del Desarrollo, en conjunto con el equipo de innovación de la misma Institución.
- Identificar patrones de información de los usuarios del simulador de realidades aumentadas, que se puedan clasificar en categorías o clases.

3. Trabajos Relacionados

Evaluar el desempeño de los alumnos en el desarrollo de habilidades adquiridas durante la realización de ejercicios de simulación en laparoscopia, es el principal desafío que podemos encontrar en general en simuladores, sin importar si son físicos, virtuales, o robotizados.

Su curva de aprendizaje es uno de los principales factores restrictivos para su amplia utilización en la práctica clínica en nuestro país. A pesar de los evidentes beneficios que conlleva en la adquisición de habilidades en procedimientos de laparoscopia, la difusión de esta tecnología ha sido lenta, siendo el acceso a la tecnología una de sus mayores restricciones para ampliar su utilización.

La realización de esta técnica mínimamente invasiva requiere de habilidades avanzadas y específicas que no necesariamente se adquieren con el aprendizaje de la cirugía tradicional o abierta. Dichas habilidades, habitualmente se aprenden como un acto autodidacta por parte de cirujanos experimentados, a pesar de que existe evidencia que el entrenamiento inapropiado de estos procedimientos puede comprometer la seguridad de los pacientes.

Esta brecha entre el nivel de práctica exigida y la realidad es lo que ha llevado a desarrollar simuladores de entrenamiento avanzado con el objetivo de aumentar las oportunidades de práctica de los cirujanos en formación, tal como se indica en la Tabla 1.

| Características | Dificultades |
|--------------------------------------|--|
| Visión en 2 dimensiones | Percepción de profundidad reducida |
| Eje ojo-mano-objetivo perturbado | Disminuye la ergonomía y destreza |
| Instrumentos largos y rígidos | Temblor manual natural aumentado y disminución del rango de movimientos |
| Puertos de entrada abdominales fijos | Efecto fulcrum <i>* apoyo sobre el que pivota la palanca del instrumento quirúrgico</i> |
| Inestabilidad de la cámara | Aumento de fatiga |
| Capacidad táctil limitada | Menor destreza manual fina |

Tabla 1 - Características y dificultades que presenta la cirugía laparoscópica

Los simuladores de realidad virtual permiten a los sujetos en entrenamiento interactuar con un ambiente simulado por computador que reproduce partes específicas o un procedimiento quirúrgico completo. Además del entrenamiento, son una potente herramienta de evaluación de habilidades quirúrgicas, los cuales entregan retroalimentación pos entrenamiento, de variables como el tiempo del ejercicio, la eficiencia de los movimientos, e integridad en las suturas, entre otras.

Varios estudios han demostrado que el entrenamiento mediante simulación laparoscópica mejora las capacidades y permite desarrollar destrezas en los sujetos que realizan el entrenamiento durante, e incluso, posteriormente al período de formación.

En la Tabla 2, se comparan las principales características de cada tipo de simulador disponible, en donde se analizan características diferenciadoras desde aspectos como los costos, la retroalimentación durante la formación del alumno, entre otros.

| Tipo de simulador | Características |
|---|---|
| Caja laparoscopia o <i>endotrainers</i> | Bajo costo, portátiles, usos ilimitados. Habilidades básicas y avanzadas. Retroalimentación sensorial adecuada. Requiere supervisión de entrenador |
| Simuladores de realidad virtual | Registro y evaluación de variables métricas. Registro del desempeño y logro de objetivos. Alto costo inicial. Menos retroalimentación sensorial |
| Modelos de simulación híbridos | Muy bajo costo. Baja representatividad de la realidad. Valor discutible como método de aprendizaje |
| Modelos cadavéricos humanos y animales (<i>ex vivo</i>) | Óptima representatividad de la realidad. Cuestionamiento ético. Disponibilidad ilimitada. Alto costo. Requiere centro especializado, personal entrenado para su preparación y manejo. |

Tabla 2 - Características de los diferentes tipos de modelos de simulación

Los simuladores de realidad virtual permiten recolectar datos objetivos sobre el desempeño de cada usuario, sin embargo, los actuales simuladores presentes en el mercado deben elaborar una mejor metodología para la evaluación del desempeño de cada alumno, ya que actualmente no se considera una profundización de los datos que se generan durante una sesión de entrenamiento.

4. Descripción de datos y Metodología

4.1. Contexto del Escenario Virtual

Los escenarios desarrollados actualmente en el simulador son tres, en donde encontramos ejercicios de coordinación, traslado y agarre, siendo este último escenario en donde centraremos nuestro análisis.

El escenario de agarre consiste en realizar el traslado de seis pelotitas desde un lugar a otro, con una pinza que en su punta posee una tenaza con la cual tomaremos cada pelotita y la trasladaremos hacia el fondo del escenario para depositarla en un recipiente.

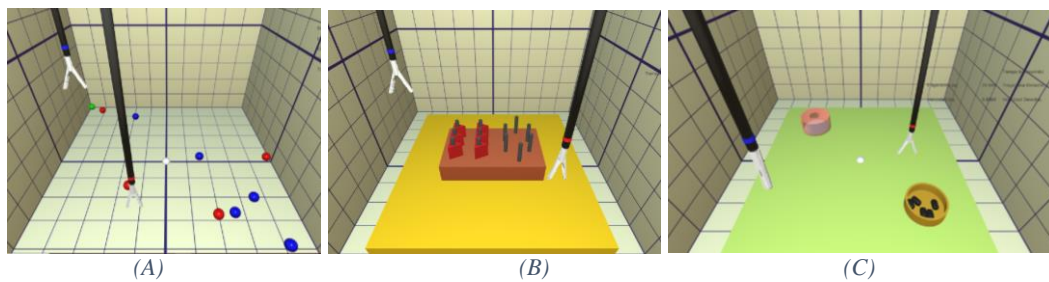


Imagen 1 - Escenarios de entrenamiento
(A) Escenario de Coordinación (B) Escenario de Traslado (C) Escenario de Agarre

Durante el trayecto de traslado se debe sostener la pelotita sin que esta se nos suelte, ya que el contexto de este escenario, es similar a trasladar algún material quirúrgico dentro de una cavidad abdominal humana en un procedimiento de laparoscopia inmersiva.

La habilidad que se busca desarrollar con el entrenamiento en este escenario es habituarse a trabajar en un ambiente que posee profundidad, la cual durante una cirugía se ve en dos dimensiones, a través de una pantalla, ver imagen 2.

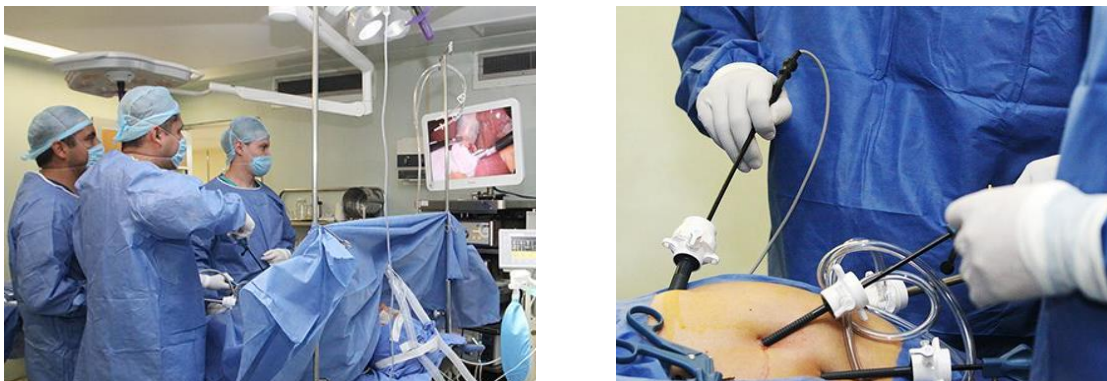


Imagen 2 - Retiro de vesícula en procedimiento de laparoscopia
Fuente: <http://www.imss.gob.mx/prensa/archivo/201711/348>

El sistema de referencia que se utilizó en el desarrollo del escenario de “agarre”, es el de sistema de coordenadas mano izquierda (ver imagen 3), el cual se utiliza principalmente

en desarrollos de videojuegos, debido al uso de la matriz de la cámara con la cual vemos el escenario dentro del ambiente virtual. Por otra parte, el sistema de coordenadas de mano derecha no se utiliza en este tipo de problemas, debido a que este generalmente se aplica en problemas matemáticos.

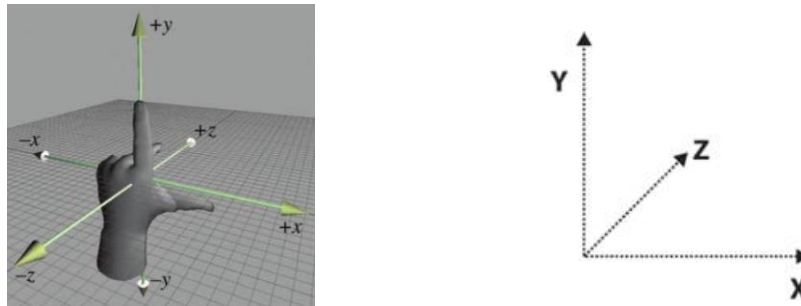


Imagen 2 - Coordenadas de Mano Izquierda
Fuente: <https://programmerclick.com/article/15612037685/>

4.2. Categorización de los usuarios

La categorización de cada uno de los usuarios se realizó analizando el perfil de cada uno, en donde se consideró la experiencia previa, el área profesional en el cual se desenvolvían, lo cual tuvo relación con los resultados preliminares de la investigación, que se verán en el capítulo 5 – Resultados.

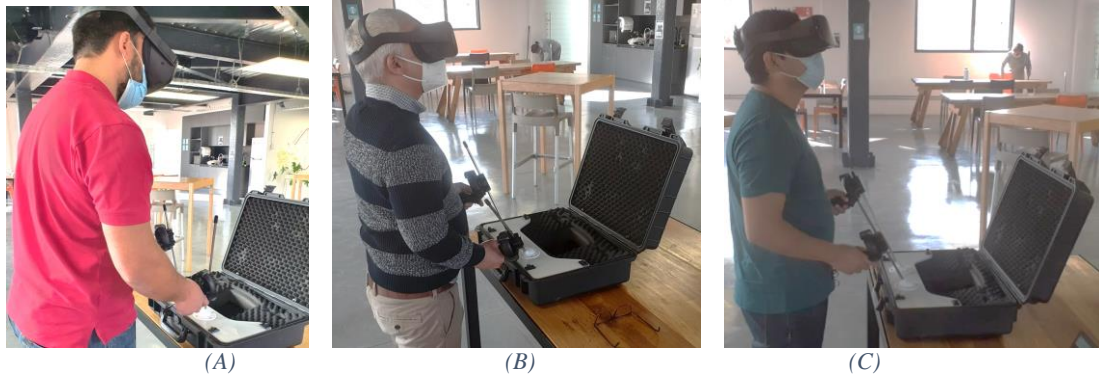
Luego de obtener el set de datos, y realizando análisis preliminares absolutos en base a tiempo de desempeño, la distancia recorrida y la cantidad de registros que generó cada usuario durante sus diez sesiones de entrenamiento (*ver tabla 6*) es que se etiquetaron como se indica en la *tabla 3*.

| Usuario | Categoría del usuario | Etiqueta |
|-----------|-----------------------|----------|
| Usuario 1 | Nivel bajo | 0 |
| Usuario 2 | Nivel medio | 1 |
| Usuario 3 | Nivel Avanzado | 1 |

Tabla 3 - Categorización de usuarios

4.3. Generación del Data Set

Este set de datos fue creado a partir de los entrenamientos de tres usuarios diferentes, que realizaron diez entrenamientos cada uno en un ejercicio denominado *de agarre* en un simulador de realidad virtual para entrenamientos en laparoscopia inmersiva, desarrollado por el área de innovación de la Universidad del Desarrollo, en conjunto con la facultad de Medicina de la misma Universidad.



(A) (B) (C)
 Imagen 3 - Usuarios realizando entrenamientos en Simulador de Realidad Virtual
 (A) Usuario novato, (B) Usuario Intermedio, (C) Usuario Avanzado

El ejercicio de entrenamiento, consiste en trasladar seis pelotas desde el pocillo amarillo hacia el pocillo marrón, los cuales se encuentran dentro de un escenario virtual y son trasladados con una pinza que ingresa desde la parte superior de la escena.

4.4. Datos

El desarrollo de este escenario se realizó en el motor de Realidad Virtual de Unity, en donde a través de la API *Tin Can Api*, se almacena cada registro a una frecuencia de 16 ms.

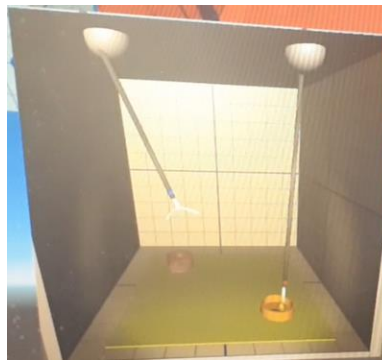


Imagen 4 - Escenario "De agarre" en Unity

Cada frame de datos registra dos grupos de coordenadas, tanto de posición, como de rotación, estas se exportan desde un servidor web en donde se almacenan los datos luego de cada sesión de entrenamiento en el simulador de realidad virtual.

En la imagen siguiente se muestra la estructura con la cual los datos son exportados desde el servidor hacia un disco local, esta data debe pasar por un proceso de limpieza y pre procesamiento para obtener un set de datos limpio y listo para comenzar a ser procesado.

Position =(-2.8, -8.1, 45.5)|Rotation = (0.0, 1.0, -0.1, 0.1)
 Position =(-2.8, -8.3, 45.5)|Rotation = (0.0, 1.0, -0.1, 0.1)
 Position =(-2.7, -8.4, 45.5)|Rotation = (0.0, 1.0, -0.1, 0.1)
 Position =(-2.7, -8.6, 45.5)|Rotation = (0.0, 1.0, -0.1, 0.1)
 Position =(-2.7, -8.7, 45.6)|Rotation = (0.0, 1.0, -0.1, 0.1)

Tabla 1 – Fragmento de datos proveniente de un entrenamiento

Este fragmento del set de datos, fue importado desde el servidor para aprendizaje de datos *ProProfs LRS*, descargados en ficheros.txt, a continuación, se muestra la arquitectura básica del servidor.

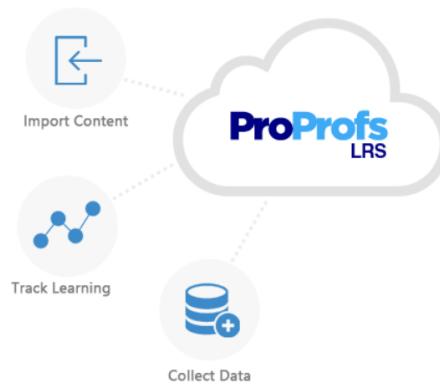


Imagen 5 - Arquitectura básica del servidor ProProfs LRS

Fuente: <https://www.proprofs.com/training/features/learning-record-store/>

Luego de diez sesiones de entrenamiento realizadas por tres usuarios diferentes, se obtuvo un set de datos con una dimensionalidad de ochenta y nueve mil doscientas once filas por catorce columnas, a continuación, se muestra un fragmento del set de datos luego de haber pasado por el proceso de limpieza de datos.

| | index | test | Px | Py | Pz | Rx | Ry | Rz | Rw | date | time | user | timestamp | nivel |
|---|-------|------|------|------|------|----|----|------|-----|--------|----------|--------|-----------|-------|
| 0 | 0 | 1 | -2.8 | -8.1 | 45.5 | 0 | 1 | -0.1 | 0.1 | oct-21 | 15:50:21 | miguel | 0 | 0 |
| 1 | 1 | 1 | -2.8 | -8.3 | 45.5 | 0 | 1 | -0.1 | 0.1 | oct-21 | 15:50:21 | miguel | 16 | 0 |
| 2 | 2 | 1 | -2.7 | -8.4 | 45.5 | 0 | 1 | -0.1 | 0.1 | oct-21 | 15:50:21 | miguel | 32 | 0 |
| 3 | 3 | 1 | -2.7 | -8.6 | 45.5 | 0 | 1 | -0.1 | 0.1 | oct-21 | 15:50:21 | miguel | 48 | 0 |
| 4 | 4 | 1 | -2.7 | -8.7 | 45.6 | 0 | 1 | -0.1 | 0.1 | oct-21 | 15:50:21 | miguel | 64 | 0 |

Tabla 4 – Fragmento del dataset, luego del proceso de limpieza de datos

Al ser un desarrollo de innovación, se debe realizar la documentación del significado de cada una de las columnas, para entender posteriormente el análisis de los datos en cada uno de los ejercicios realizados en el simulador, lo cual se muestra en la Tabla 5.

| Columna | Descripción |
|-----------|--|
| index | Index o índice de cada registro del set de datos |
| test | corresponde a cada uno de los test realizados |
| Px | dato de posición en su coordenada x en el espacio tridimensional |
| Py | dato de posición en su coordenada y en el espacio tridimensional |
| Pz | dato de posición en su coordenada z en el espacio tridimensional |
| Rx | dato de rotación en su coordenada x en el espacio tridimensional |
| Ry | dato de rotación en su coordenada y en el espacio tridimensional |
| Rz | dato de rotación en su coordenada z en el espacio tridimensional |
| Rw | dato de posición en su coordenada w en el espacio tridimensional |
| date | fecha del ejercicio |
| time | hora del ejercicio |
| user | identificación del usuario |
| timestamp | Registro de datos cada 16 ms |
| nivel | Etiqueta del usuario que realizó el ejercicio |

Tabla 5 - Documentación del set de datos

La siguiente tabla, muestra un resumen de algunos indicadores del set de datos, en donde se distinguen las dos clases etiquetadas, cero para usuarios novatos y uno, para los usuarios intermedio y avanzado.

| sesión | etiqueta | nivel | test | registros | % registros | distancia recorrida | tiempo total (s) |
|--------|----------|----------|------|-----------|-------------|---------------------|------------------|
| 1 | 0 | bajo | 1 | 5923 | 6.6% | 1.45 | 82.4 |
| 2 | 0 | bajo | 2 | 3460 | 3.9% | 0.26 | 48.1 |
| 3 | 0 | bajo | 3 | 4285 | 4.8% | 0.24 | 59.6 |
| 4 | 0 | bajo | 4 | 5767 | 6.5% | 1.33 | 80.2 |
| 5 | 0 | bajo | 5 | 3955 | 4.4% | 0.25 | 55.0 |
| 6 | 0 | bajo | 6 | 6394 | 7.2% | 0.42 | 89.0 |
| 7 | 0 | bajo | 7 | 4028 | 4.5% | 0.43 | 56.0 |
| 8 | 0 | bajo | 8 | 3832 | 4.3% | 0.29 | 53.3 |
| 9 | 0 | bajo | 9 | 5903 | 6.6% | 0.37 | 82.1 |
| 10 | 0 | bajo | 10 | 3596 | 4.0% | 0.22 | 50.0 |
| 11 | 1 | medio | 11 | 3047 | 3.4% | 0.69 | 42.3 |
| 12 | 1 | medio | 12 | 2758 | 3.1% | 0.49 | 38.3 |
| 13 | 1 | medio | 13 | 3252 | 3.6% | 0.50 | 45.2 |
| 14 | 1 | medio | 14 | 2400 | 2.7% | 0.25 | 33.3 |
| 15 | 1 | medio | 15 | 2090 | 2.3% | 0.24 | 29.0 |
| 16 | 1 | medio | 16 | 2742 | 3.1% | 0.33 | 38.1 |
| 17 | 1 | medio | 17 | 2232 | 2.5% | 0.32 | 31.0 |
| 18 | 1 | medio | 18 | 2135 | 2.4% | 0.53 | 29.6 |
| 19 | 1 | medio | 19 | 2011 | 2.3% | 0.37 | 27.9 |
| 20 | 1 | medio | 20 | 1909 | 2.1% | 0.25 | 26.5 |
| 21 | 1 | avanzado | 21 | 1063 | 1.2% | 0.08 | 14.7 |
| 22 | 1 | avanzado | 22 | 1706 | 1.9% | 0.12 | 23.7 |
| 23 | 1 | avanzado | 23 | 1413 | 1.6% | 0.09 | 19.6 |

| | | | | | | | |
|----|---|----------|----|------|------|------|------|
| 24 | 1 | avanzado | 24 | 1293 | 1.4% | 0.10 | 17.9 |
| 25 | 1 | avanzado | 25 | 2251 | 2.5% | 0.13 | 31.2 |
| 26 | 1 | avanzado | 26 | 2062 | 2.3% | 0.20 | 28.6 |
| 27 | 1 | avanzado | 27 | 1766 | 2.0% | 0.11 | 24.5 |
| 28 | 1 | avanzado | 28 | 1933 | 2.2% | 0.15 | 26.8 |
| 29 | 1 | avanzado | 29 | 1369 | 1.5% | 0.09 | 19.0 |
| 30 | 1 | avanzado | 30 | 2636 | 3.0% | 0.39 | 36.6 |

Tabla 6 - Resumen general del set de datos

| estadísticos | novato | | | intermedio | | | avanzado | | |
|----------------------------|--------|--------|-------|------------|--------|-------|----------|--------|-------|
| | Px | Py | Pz | Px | Py | Pz | Px | Py | Pz |
| Promedio | -9.13 | -9.76 | 47.06 | -8.51 | -9.76 | 47.24 | -8.57 | -10.42 | 47.47 |
| Mediana | -8.70 | -9.90 | 45.4 | -5.60 | -9.90 | 45.5 | -4.90 | -10.30 | 45.50 |
| Max | 2.20 | -1.90 | 55.9 | 0.10 | 1.50 | 53.9 | 1.60 | -2.00 | 55.30 |
| Min | -20.90 | -17.10 | 37.7 | -16.80 | -13.50 | 38.9 | -18.00 | -13.30 | 39.40 |
| Rango | 23.10 | 15.20 | 18.2 | 16.90 | 15.00 | 15.00 | 19.60 | 11.30 | 15.90 |
| Varianza | 37.93 | 2.55 | 24.31 | 32.85 | 4.53 | 19.49 | 35.57 | 2.38 | 20.96 |
| Desviación estándar | 6.16 | 1.6 | 4.93 | 5.73 | 2.13 | 4.41 | 5.96 | 1.54 | 4.58 |

Tabla 7 - Estadísticos de posición x y z, para usuarios novato, intermedio y avanzado

En la tabla anterior, se distinguen tres clases, novato, intermedio y avanzado, en donde las dos últimas clases distan sustancialmente de la primera clase en sus mediciones estadísticas, esto se ve con claridad por ejemplo en el promedio de las coordenadas del usuario intermedio y avanzado.

Si bien los datos son espaciales y poseen datos en sus tres ejes, se puede representar cada uno como una serie en donde se han registrado los casi noventa mil datos de las 30 sesiones de entrenamiento de los 3 usuarios en el simulador de realidad virtual.

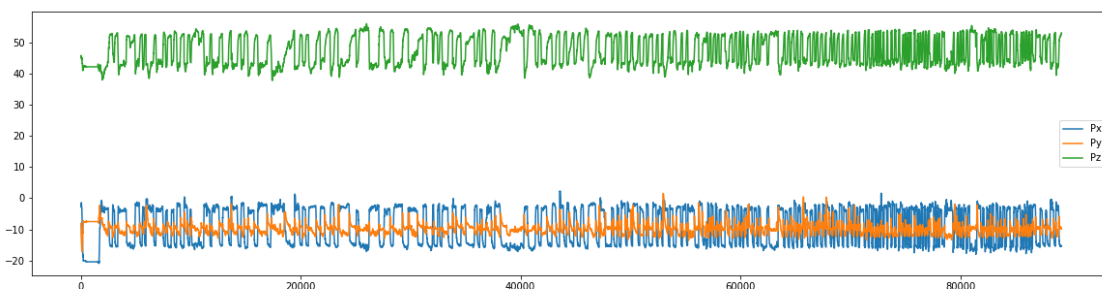


Gráfico 1 - Tendencia de Posición data set

Cada tendencia de coordenadas se puede representar por separado, pudiendo ver con claridad que cada una tiene una amplitud diferente y una escala de registro diferente.

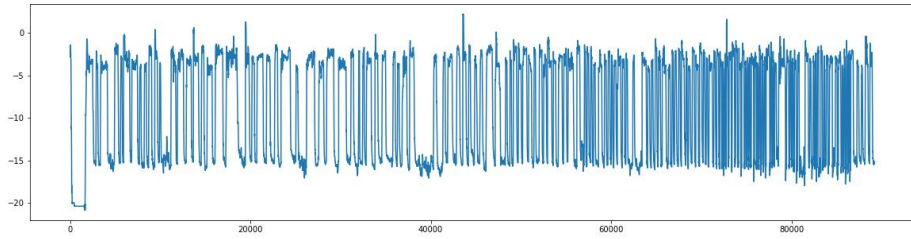


Gráfico 2 - Tendencia de posición coordenada Px

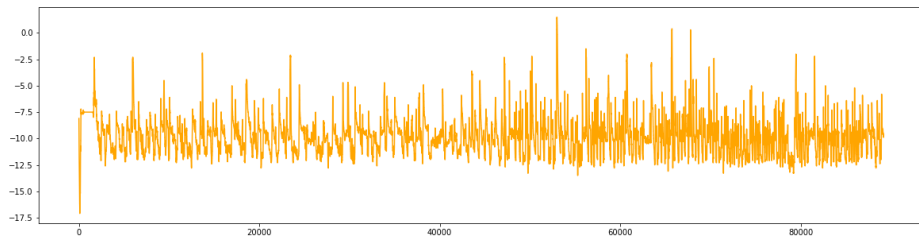


Gráfico 3 - Tendencia de posición coordenada Py

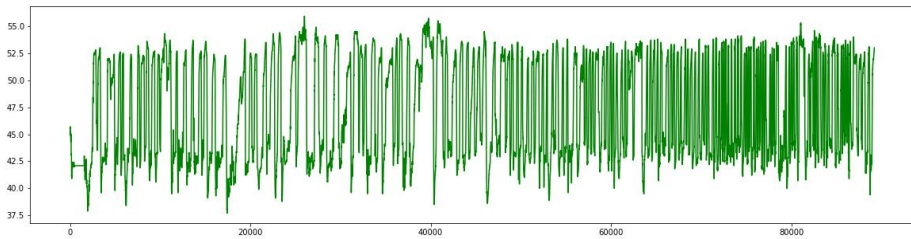


Gráfico 4 - Tendencia de posición coordenada Pz

Para los tres usuarios se registraron datos en cada una de las diez sesiones de entrenamiento realizadas, pero de acuerdo a los resultados obtenidos en la tabla 9, se han etiquetado solo con dos clases, novato, avanzado.

Esto debido a la similitud de las clases intermedio y avanzado, tanto en su cantidad de registros, en sus tiempos, como en sus distancias registradas.

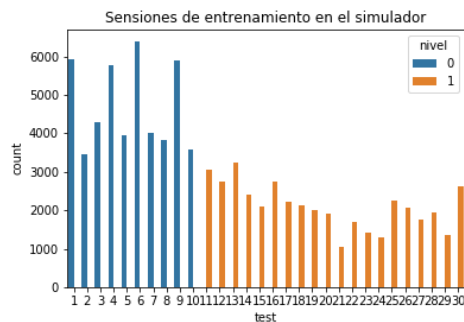


Gráfico 5 - sesiones de entrenamiento para 2 clases de usuarios

Además de esta forma se obtiene un set de datos balanceados, tal como se indica en la tabla a continuación

| registros | valor | % data set |
|-----------|-------|------------|
| Novato | 47143 | 52.84% |
| Avanzado | 42069 | 47.16% |
| Total | 89212 | 100% |

Tabla 8 - Registros del set de datos

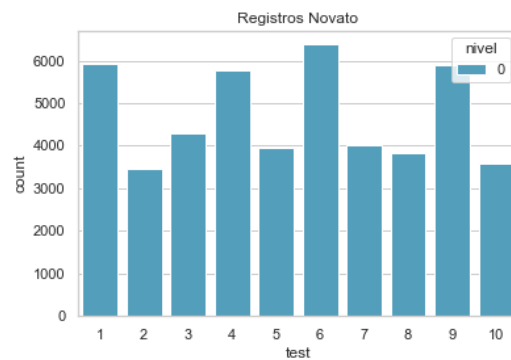


Gráfico 6 - clase 0 (Novato)

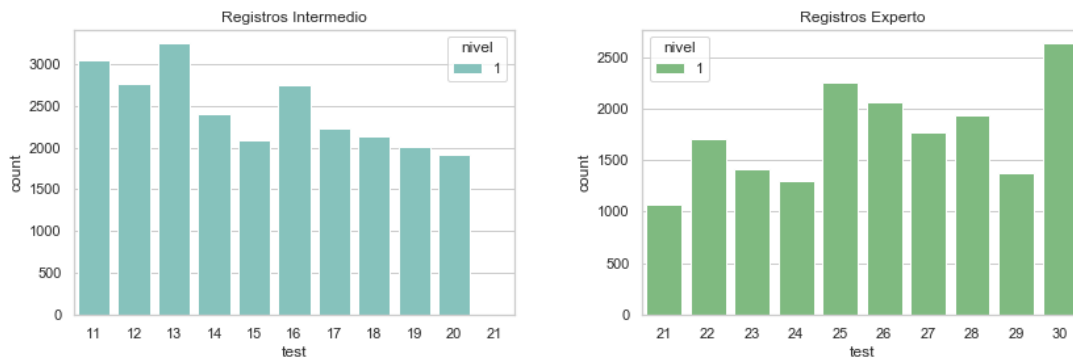


Gráfico 7 - clase 1 (Avanzado)

4.5. Metodología

La metodología de trabajo de la red neuronal artificial, comprende varias etapas, para el desarrollo y validación del modelo, seguiremos la siguiente metodología de trabajo:

- Limpieza y estructuración de datos
- Análisis estadísticos del set de datos
- Entrenamiento de la red neuronal
- Utilización de la red neuronal

Se ha comenzado con un modelo simple, escogido arbitrariamente, para comprobar la efectividad de la función de pérdida y el algoritmo de optimización escogido. El pre procesamiento de los datos no ha sido una tarea menos ya que se han realizado diversas combinaciones de parámetros para optimizar la capacidad computacional en el pre-procesamiento de la red.

5. Análisis y Modelamiento de datos

5.1. Limpieza y tratamiento de datos

Para realizar la limpieza y el tratamiento de los datos, se han utilizado módulos bastante conocidos de Python, los cuales son:

- *Keras*, biblioteca de redes neuronales de código abierto escrita en Python.
- *Matplotlib*, biblioteca para la generación de gráficos a partir de datos contenidos en listas o arrays en el lenguaje de programación Python y su extensión matemática NumPy
- *Numpy*, biblioteca para el lenguaje de programación Python que da soporte para crear vectores y matrices grandes multidimensionales, junto con una gran colección de funciones matemáticas de alto nivel para operar con ellas.
- *Pandas*, es una biblioteca de software escrita como extensión de NumPy para manipulación y análisis de datos para el lenguaje de programación Python.
- *Scipy*, biblioteca libre y de código abierto para Python. Se compone de herramientas y algoritmos matemáticos.
- *Seaborn*, es una biblioteca de visualización de datos de Python basada en matplotlib que proporciona una interfaz de alto nivel para dibujar gráficos estadísticos atractivos e informativos.
- *Sklearn*, es una biblioteca para aprendizaje automático de software libre para el lenguaje de programación Python.
- *Tensorflow*, es una biblioteca de código abierto para aprendizaje automático a través de un rango de tareas, y desarrollado por Google para satisfacer sus necesidades de sistemas capaces de construir y entrenar redes neuronales para detectar y descifrar patrones y correlaciones

5.2. Análisis estadísticos de los Datos

El análisis estadístico realizado arroja los primeros resultados que son indicadores de la distancia que existe entre las clases, ya que por ejemplo existe un 52.84% de los registros que son del usuario novato, y un 47.16% que son registros tanto para el usuario intermedio, como para el avanzado.

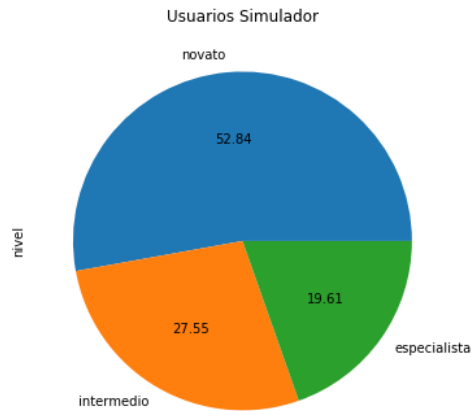


Gráfico 8 - Registros de usuarios en sesiones de entrenamiento

Por otra parte, en la imagen a continuación, podemos ver que, si bien existe una evidente mayor cantidad de registros para el usuario novato, producto de su poca experiencia en las sesiones de entrenamiento, esto nos ayuda a establecer un primer indicador de aprendizaje, *cantidad de registros*.

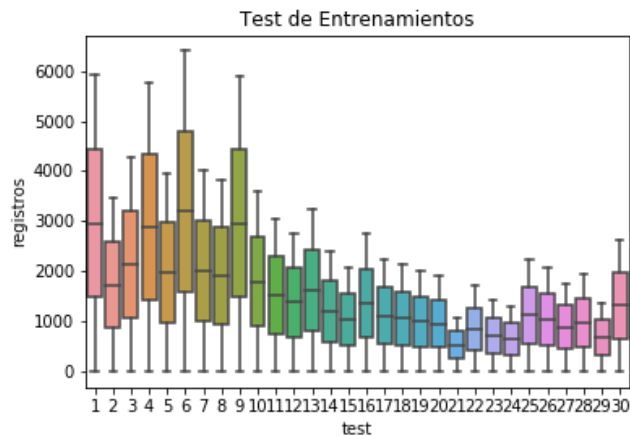


Gráfico 9 - Diagrama de caja y bigote para las sesiones de entrenamiento

En cuanto a los tiempos, se obtuvo un promedio de 41 segundos para el total de las sesiones de entrenamiento, siendo el tiempo más alto registrado con 89 segundos. En la gráfica siguiente podemos observar que los usuarios intermedio y avanzado, registran

tiempos para sus 20 sesiones de entrenamiento por debajo del promedio del total de las sesiones realizadas.

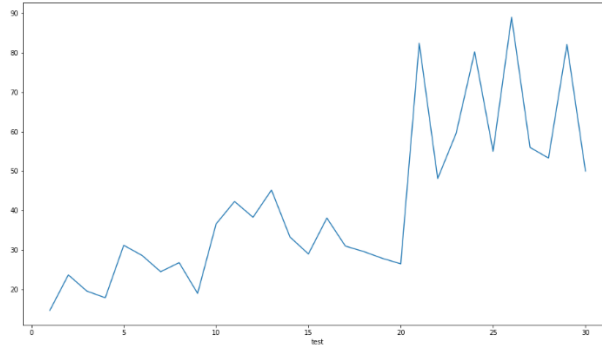


Gráfico 10 - Tiempo total para cada test realizado

| nivel | distancia (cms) |
|--------------|------------------------|
| bajo | 5.26 |
| medio | 3.98 |
| alto | 1.45 |

Tabla 9 – Distancia recorrida por cada clase

5.3. Modelamiento de Datos

5.3.1. Separar los datos para entrenar y evaluar el modelo

Si bien nos encontramos con un gran set de datos provenientes desde la base de datos del simulador de realidad virtual al empezar, haremos una separación de estos para utilizarlos para entrenamiento y evaluación del modelo.

5.3.2. Datos de entrenamiento, validación y prueba

Para la configuración y evaluación de nuestro modelo, habitualmente se dividen los datos disponibles en dos conjuntos de datos:

- Datos de entrenamiento (*training*)
- Datos de prueba (*test*)

Y a su vez, una parte de los datos de entrenamiento (*training*) se reservan como datos de validación (*validation*).

Los datos de entrenamiento que nos quedan después de sacar los datos de validación y prueba son los que usaremos para que el algoritmo de aprendizaje calcule los parámetros del modelo, mientras que los de validación los utilizaremos para tunear los hiperparámetros.

Cuando se busca mejorar el algoritmo ajustando los hiperparámetros con los datos de validación, estamos inevitablemente incidiendo en el modelo, el cual puede sesgar los resultados a favor del conjunto de validación. Es por esta razón que disponemos de un conjunto de datos reservados de prueba final, los cuales son datos que el modelo no ha visto anteriormente durante la etapa de entrenamiento (ni como datos de entrenamiento ni como datos de validación).

Esto permite obtener una medida de comportamiento del algoritmo más objetiva y evaluar si nuestro modelo generaliza correctamente. En la figura siguiente se esquematiza esta división y el propósito del uso de cada subconjunto de datos.

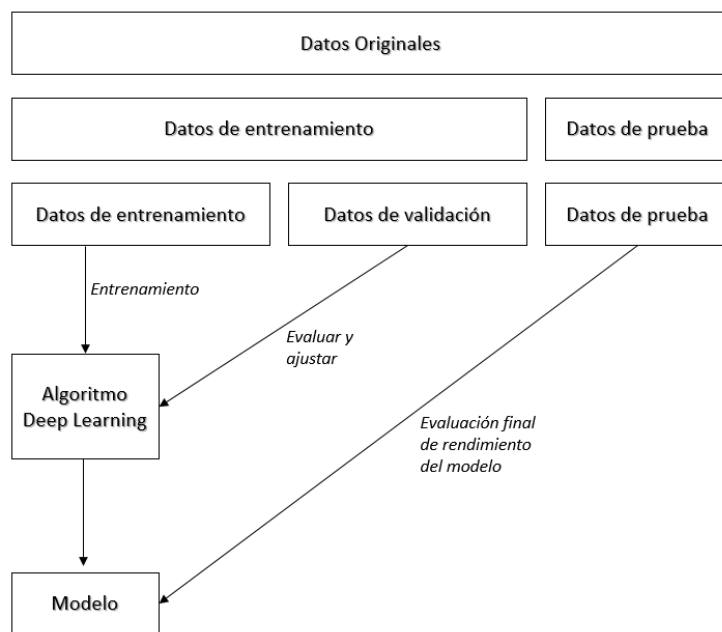


Gráfico 11 - Repartición de los datos en tres grupos (entrenamiento, validación y prueba) y el propósito de cada uno de ellos

5.4. Desarrollo del Modelo

5.4.1. Definición del Modelo

Para establecer las restricciones que necesitamos en la salida de la red neuronal, lo primero es establecer cuál es la última capa, qué tipo de activación usaremos y cuál será la relación de esta con la función de pérdida que definiremos para el modelo.

La arquitectura de nuestro modelo es una arquitectura simple de 2 capas densamente conectadas, de 64 neuronas.

Podemos inspeccionar la arquitectura de la red con el método `summary()`, en donde se obtuvo un total de 75970 parámetros entrenables, los cuales se desglosan en la columna *Param #*:

| Layer (type) | Output Shape | Param # |
|--------------------------|--------------|---------|
| lstm_1 (LSTM) | (None, 128) | 67584 |
| dropout_1 (Dropout) | (None, 128) | 0 |
| dense_1 (Dense) | (None, 64) | 8256 |
| dense_2 (Dense) | (None, 2) | 130 |
| Total params: 75,970 | | |
| Trainable params: 75,970 | | |
| Non-trainable params: 0 | | |

Imagen 7 - Arquitectura de la red Neuronal, método `summary()`

5.4.2. Configuración del Modelo

Algunas definiciones para la construcción de nuestro modelo son la función de pérdida y el optimizador.

5.4.3. Función de Pérdida

La función de pérdida está basada en una métrica de rendimiento. Una medida de rendimiento típica es el error cuadrático medio (*mean square error*) (MSE) que es la raíz cuadrada de la suma ponderada del cuadrado de las diferencias entre los valores previstos y los valores observados.

$$MSE = \sqrt{\frac{1}{m} \sum (Y_{estimada} - Y_{real})^2}$$

Aunque el MSE es generalmente la medida de rendimiento más utilizada para realizar regresión, además se puede utilizar otra función, como el error absoluto medio (*mean absolute error*) (MAE), también llamado promedio absoluto de la desviación.

$$MAE = \frac{1}{m} \sum |Y_{estimada} - Y_{real}|$$

Ambas métricas las utilizamos en modelos de regresión. MAE es más robusto para los valores atípicos, ya que no utiliza el cuadrado

5.4.4. Optimizador

En nuestro código, hemos implementado el optimizador de Adam, el cual, está diseñado para acelerar el proceso de optimización, por ejemplo, disminuir el número de evaluaciones de funciones requeridas para alcanzar los óptimos, o mejorar la capacidad del algoritmo de optimización, lo que puede entregar en un mejor resultado final.

5.4.5. Entrenamiento del Modelo

El valor del hiperparámetro *epoch* es muy importante considerar al momento de entrenar nuestro modelo, ya que es muy probable que a mayor cantidad de épocas nuestro modelo muestre una mayor precisión de la red para nuestros datos de entrenamiento. Pero por otra parte también se debe considerar que, si el número de épocas es demasiado alto, nuestra red podría tener problemas de sobreajuste.

Debido a esto, es que es muy importante encontrar el valor adecuado de este hiperparámetro. Y para esto tiene mucha importancia los datos de validación que se han definido en *validation_split*.

Se ha corrido el modelo en varias ocasiones, para distintas épocas de entrenamiento, en donde se han obtenido los resultados de la Tabla 10.

| epoch | Métricas de Validación | | | | Métricas de Entrenamiento | | | |
|-------|------------------------|----------|----------|----------|---------------------------|----------|----------|----------|
| | val_loss | val_acc | val_MSE | val_MAE | loss | accuracy | MSE | MAE |
| 10 | 0.499838 | 0.756833 | 0.165865 | 0.315847 | 0.509304 | 0.754032 | 0.168619 | 0.344167 |
| 30 | 0.221032 | 0.920112 | 0.063958 | 0.118724 | 0.220093 | 0.920407 | 0.061492 | 0.125166 |
| 50 | 0.400685 | 0.817099 | 0.129407 | 0.261045 | 0.41607 | 0.814341 | 0.13301 | 0.281212 |
| 100 | 0.315199 | 0.871759 | 0.090575 | 0.16333 | 0.318273 | 0.89446 | 0.084379 | 0.131454 |

Tabla 10 - Métricas para distintas Épocas de entrenamiento

De acuerdo a la tabla anterior, hemos obtenido distintas gráficas de Train Erros vs Val Error, en donde podemos comparar las distintas épocas de entrenamiento que se han probado, y con la variación y testeo de este hiperparámetro, se puede ver en la gráfica de 50 épocas, que el modelo de entrenamiento se sobreajuste al pasar de las 45 épocas de entrenamiento, por lo que un valor óptimo de este hiperparámetro son 40 épocas.

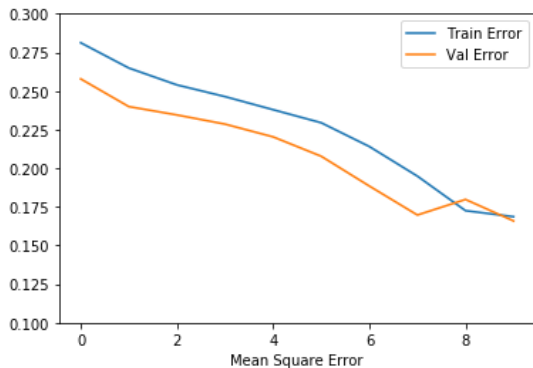


Gráfico 11 - 10 epoch de entrenamiento

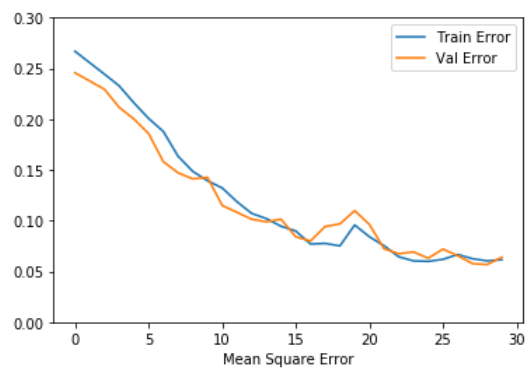


Gráfico 12 - 30 epoch de entrenamiento

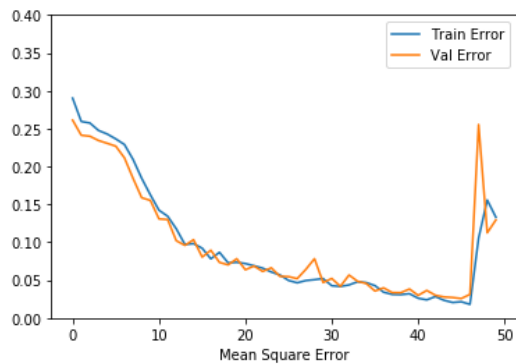


Gráfico 13 - 50 epoch de entrenamiento

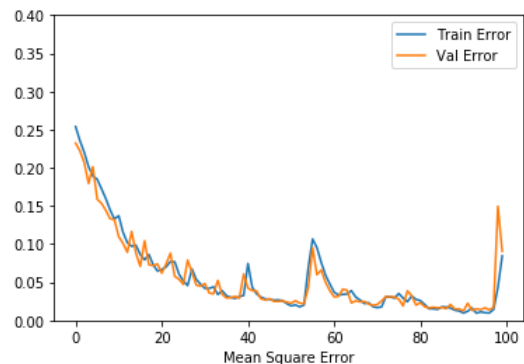


Gráfico 14 - 100 epoch de entrenamiento

5.4.6. Overfitting

Este concepto de *overfitting* (*sobreajuste*), se produce cuando el modelo obtenido se ajusta tanto a los ejemplos etiquetados de entrenamiento que no puede realizar las predicciones correctas en ejemplos de datos nuevos que nunca ha visto antes.

Por lo tanto, el sobreajuste sucede, cuando se modelan los datos de entrenamiento demasiado bien, aprendiendo detalles de estos que no son generales. Esto debido a que sobreentrenamos nuestro modelo y este estará considerando como válidos solo los datos idénticos a los que nuestro conjunto de entrenamiento, incluido sus defectos (*ruído* para el problema que estamos resolviendo).

5.4.7. Matriz de Confusión

La matriz de confusión sirve para mostrar cuándo una clase es confundida con otra, lo cual nos permite trabajar de forma separada con los distintos tipos de error asociados.

| | | | |
|---------------------|---|--|---|
| | | clase predicha | |
| | | + | - |
| clase actual | + | TP True Positives | FN False Negatives Type II error |
| | - | FP False Positives Type I error | TN True Negatives |

Gráfico 15 - Matriz de confusión

Donde:

- TP True Positives (verdaderos positivos)
- TN True Negatives (verdaderos negativos)
- FP False Positives (falsos positivos)
- FN False Negatives (falsos negativos)

5.4.8. Métricas

Las siguientes métricas se utilizan comúnmente para evaluar el rendimiento de los modelos de clasificación.

| Métrica | Fórmula | Interpretación |
|--|-------------------------------------|---|
| <u>Exactitud</u> <i>Accuracy</i> | $\frac{TP + TN}{TP + TN + FP + FN}$ | Rendimiento general del modelo |
| <u>Precisión</u> <i>Precision</i> | $\frac{TP}{TP + FP}$ | Qué tan precisas son las predicciones positivas |
| <u>Sensibilidad</u> <i>Recall</i> <i>Sensitivity</i> | $\frac{TP}{TP + FN}$ | Convergencia de muestra positiva real |
| <u>Especificidad</u> <i>Specificity</i> | $\frac{TN}{TN + FP}$ | Cobertura de muestra negativa actual |
| <u>Puntuación F1</u> <i>F1 score</i> | $\frac{2TP}{2TP + FP + FN}$ | Métrica híbrida útil para clases no balanceadas |

Tabla 11 - Principales métricas

Los valores de la diagonal principal **TP**=905 y **TN**=786 corresponden con los valores estimados de forma correcta por el modelo, tanto los *verdaderos positivos TP*, como los *verdaderos negativos TN*.

La otra diagonal, por tanto, representa los casos en los que el modelo *se ha equivocado*, **FN**=48 *falsos negativos*, **FP**=44 *falsos positivos*.

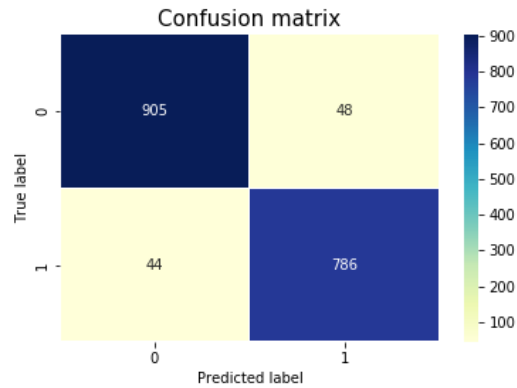


Gráfico 16 - Matriz de confusión

5.4.8.1. Exactitud

La exactitud (*accuracy*) es el porcentaje de predicciones correctas frente al total. Por tanto, es el cociente entre los casos bien clasificados por el modelo *verdaderos positivos (TP)* y *verdaderos negativos (TN)*, es decir, los valores en la diagonal de la matriz de confusión, y la suma de todos los casos.

El valor obtenido para este modelo es de un 73%. No es maravilloso, pero podemos considerarlo *aceptable*). Para calcularlo a mano, a partir de la matriz de confusión:

| <u>Fórmula</u> | <u>Ecuación</u> | <u>Exactitud del Modelo</u> |
|-------------------------------------|---|-----------------------------|
| $\frac{TP + TN}{TP + TN + FP + FN}$ | $\frac{905 + 786}{905 + 786 + 44 + 48}$ | $0.948 \approx 94.8\%$ |

5.4.8.2. Precisión

La precisión, (*precision*) se refiere a lo cerca que está el resultado de nuestra predicción del valor verdadero. Por tanto, es el cociente entre los casos positivos bien clasificados por nuestro modelo y el total de las predicciones positivas. Este cálculo, a partir de la matriz de confusión

| <u>Fórmula</u> | <u>Ecuación</u> | <u>Precisión del Modelo</u> |
|----------------------|------------------------|-----------------------------|
| $\frac{TP}{TP + FP}$ | $\frac{905}{905 + 44}$ | $0.953 \approx 95.3\%$ |

El valor obtenido para este modelo es de un 62%. Por tanto, nuestro modelo es más preciso que exacto.

5.4.8.3. Sensibilidad, exhaustividad

La sensibilidad (*recall*) representa la tasa de nuestros valores verdaderos positivos (*True Positive Rate*) ó TP. Indica la proporción entre los casos positivos bien clasificados por el modelo, respecto al total de positivos. Para calcularlo en este caso:

| <u>Fórmula</u> | <u>Ecuación</u> | <u>Sensibilidad del Modelo</u> |
|----------------------|------------------------|--------------------------------|
| $\frac{TP}{TP + FN}$ | $\frac{905}{905 + 48}$ | 0.9496 \approx 95.0% |

Este valor representa la habilidad del modelo de detectar los casos relevantes. Un 95% es claramente un valor muy bueno para una métrica. Por lo tanto, decimos que nuestro algoritmo de clasificación tiene una alta sensibilidad, considera muchos valores positivos.

5.4.8.4. Especificidad

La especificidad, por su parte, es la tasa de verdaderos negativos, (*true negative*). Corresponde a la proporción entre los casos negativos bien clasificados por el modelo, respecto al total de casos clasificados como negativos. Este cálculo es el siguiente:

| <u>Fórmula</u> | <u>Ecuación</u> | <u>Especificidad del Modelo</u> |
|----------------------|------------------------|---------------------------------|
| $\frac{TN}{TN + FP}$ | $\frac{905}{905 + 44}$ | 0.9469 \approx 95.0% |

En este caso, la especificidad tiene un valor muy bueno. Esto quiere decir que su capacidad de discriminar los casos negativos es muy buena. Por lo tanto, es difícil obtener falsos positivos.

5.4.8.5. F1 Score

La métrica F1 score, combina tanto la precisión como la sensibilidad en un solo valor de medida, esto nos permitirá comparar el rendimiento combinado entre la precisión y la sensibilidad entre varias soluciones.

| <u>Fórmula</u> | <u>Ecuación</u> | <u>F1 Score del Modelo</u> |
|-----------------------------|---------------------------------|----------------------------|
| $\frac{2TP}{2TP + FP + FN}$ | $\frac{2x905}{2x905 + 44 + 48}$ | 0.9516 \approx 95.1% |

5.4.9. Evaluación del Modelo con los datos de Prueba

Finalmente, se debe testear como generaliza el modelo al utilizar el conjunto de datos de prueba, que no hemos utilizado anteriormente durante el entrenamiento del modelo, esto es un método que nos servirá como indicador de los que nos podemos esperar del modelo, cuando este lo utilicemos en un entorno de producción, Para esto se utilizará el método *evaluate()* :

```
Test score:          0.16724725416117148
Test accuracy:      0.9484015629656584
Test MSE:           0.04405356670948674
Test MAE:           0.04405356670948674
```

Gráfico 17 – Evaluación del Modelo de Entrenamiento con datos de prueba

6. Resultados

Los resultados obtenidos en la elaboración de este modelo entregan información que permitirá al evaluador de las sesiones de entrenamiento, retroalimentar al estudiante durante su proceso de aprendizaje y adquisición de habilidades en el entrenamiento en el simulador de realidad virtual para procedimientos de laparoscopia.

Uno de los indicadores a considerar luego de haber analizado el universo de datos, es la distancia de la trayectoria de cada una de las sesiones de entrenamiento, en donde la muestra de las sesiones de un novato de etiqueta 0 recorrió un 35.19% más de su trayectoria con respecto al usuario especialista, realizando movimientos poco eficientes que aumentaron sus registros en el set de datos.

Por otra parte, los tiempos de desempeño de los usuarios en general, las sesiones fueron realizadas en 2.06 minutos, de los cuales el usuario especialista solo utilizó el 19.3% del tiempo total para realizar las 10 sesiones de entrenamiento en el ejercicio de agarre. Esto está relacionado con la utilización eficiente de las trayectorias de traslado desde el punto A hacia el punto B en el escenario virtual, con lo que se puede determinar la adquisición de habilidades para dimensionar profundidad en un espacio que se visualiza en 2 dimensiones, pero que en la virtualidad se encuentra en 3 dimensiones.

Una similitud muy estrecha con la realidad de los procedimientos de laparoscopia en donde un especialista opera en un espacio físico de 3 dimensiones, pero observando un ambiente de 2 dimensiones que puede observar solo a través de un monitor.

En cuanto a las métricas del modelo, tanto la sensibilidad como la especificidad dieron 95%. Por tanto, este modelo es balanceado en sus valores verdaderos positivos, como en sus verdaderos negativos. Por lo que nuestro modelo evitaría los casos falsos positivos y falsos negativos.

El proyecto del simulador de realidad virtual ha sido desarrollado como un prototipo como parte del área de innovación de la Universidad del Desarrollo, y como tal se consideraron detalles y variados aspectos que acercan al estudiante especialista a un ambiente realista que permite posteriormente obtener los datos de cada registro que se genera propio del ejercicio realizado. Pero, además, a lo largo del modelamiento de los datos encontramos oportunidades de mejora, que permitirán obtener métricas aún más precisas y agregar una mayor cantidad de columnas con información para el análisis, ya que el simulador registra datos tanto de posición en sus 3 coordenadas, como de rotación de las pinzas en coordenadas quaternions, es posible optimizar el código base para que el set de dato posea variables físicas asociadas a la velocidad y trayectoria, como la rapidez, velocidad media, velocidad instantánea, entre otras variables, sería posible analizar por ejemplo cuales son las trayectorias críticas e incorporar esto como un indicador más dentro del análisis de los datos, este indicador se asocia al tiempo de registro de cada frame de datos. Si bien se intentó incorporar una columna de timestamp arbitrariamente de 16 (ms),

los resultados no fueron reales, ya que se obtuvo una pendiente positiva constante, que no se pudo utilizar objetivamente para los análisis de esta tesis.

En cuanto a la red neuronal, se diseñó una arquitectura apropiada luego de variadas pruebas de laboratorio computacional en Python, en donde se modeló de tal forma que esta aprendiera del set de datos de posición (x, y, z), considerando como variable objetivo el nivel del usuario que realizó cada una de las 10 sesiones de entrenamiento, en donde se etiquetó a cada usuario con un 0 si su nivel era novato, y un 1 si su nivel era intermedio o avanzado, sin embargo luego de haber ajustado los parámetros se obtuvo un modelo que logra discriminar más que sobre si el nivel del estudiante es 0 ó 1, es que si su comportamiento en los registros de las sesiones de entrenamiento es más el de un novato o un avanzado.

Otro resultado que se pudo ver en esta tesis es que hay una muy alta variabilidad en la trayectoria de los registros para el set completo de datos, el cual es un set pequeño de datos, esto hace que la red no haya sido alimentada con un gran volumen de datos como es habitual en el modelamiento de las redes neuronales.

Si bien, es un argumento válido para haber seleccionado otro modelo de Redes neuronales, no es tan así, ya que se explica por la secuencialidad de la que se registran los datos en el espacio, muy similar a una serie de tiempo, es así como se optó por la implementación de este modelo LSTM asumiendo que el set de datos es pequeño, pero con la particularidad de la secuencialidad de los datos que se registraron en las sesiones de entrenamiento.

7. Conclusiones

La optimización de los procesos de entrenamiento para un profesional de la salud a través de los métodos convencionales, son extensos y en el caso de la capacitación de médicos especialistas en la cirugía de laparoscopia, carecen de herramientas digitales avanzadas, que permitan aumentar la productividad en cuanto al entrenamiento de este tipo de cirugías tan complejas.

Con el desarrollo de prototipos de entrenamiento en realidad aumentada como es el simulador de realidad virtual inmersivo para cirugía de mínimo acceso incorporando analítica avanzada en su proceso de personalización de reportes a cada alumno, es posible aumentar la eficiencia con que el alumno se entrena en la adquisición de la técnica laparoscópica.

Si bien este estudio contempla datos de un simulador como parte de un desarrollo de innovación, es un primer avance en la exploración de nuevas y mejores herramientas de aprendizaje para los especialistas del área de la salud.

Como parte del aprendizaje de este estudio, se encuentra la cantidad de datos que se utilizó, siendo este un set de datos pequeño, y variable en su cantidad de registros, permitió dar un pequeño avance en el análisis de estos datos.

Por otra parte, el hecho de obtener retroalimentación de los datos de cada sesión y un análisis avanzado de estos datos proporciona a los estudiantes, la oportunidad de conocer aún más, cuales son los aspectos que debe mejorar y en qué se debe concentrar en sus próximas sesiones de entrenamiento.

Ningún método de evaluación por si solo puede proporcionar todos los datos que se requieren para evaluar algo tan complejo como la prestación de servicios profesionales por parte de un médico exitoso, sin embargo, el desarrollo de una herramienta de evaluación para validar la capacitación de un estudiante merece ampliar la cantidad de sesiones de entrenamiento para mejorar la precisión del algoritmo y conseguir mejores resultados de desempeño.

8. Trabajo Futuro y Recomendaciones

- Definir indicadores y umbral para cada clase de usuario.
- Realizar nuevas sesiones de trabajo con el simulador de entrenamiento en laparoscopia para aumentar el volumen de datos y mejorar los resultados del Modelo.
- Al aumentar la cantidad de usuarios que utilicen el simulador, se recomienda estudiar el comportamiento entre individuos, para otorgar nuevas herramientas en el proceso de clasificación inicial de usuarios.
- Incorporar modelamiento de datos en una versión 2.0 del Simulador de VR, para probar el modelo sobre nuevos usuarios y nuevos datos.
- Incorporar una marca de tiempo en Unity para cada registro tridimensional para incorporar cálculos de variables físicas como trayectoria, desplazamiento, rapidez, etc. Que sirvan como nuevas entradas para el modelamiento de los nuevos datos y poder controlar de cierta manera la salida desde la capa de entrada de los datos.
- Registrar la penalización de caída del objeto que se está manipulando dentro de la sesión.
- Realizar seguimiento a la curva de aprendizaje de profesionales que se preparan en esta técnica con simulador vs quienes se preparan sin simulador.

Bibliografía

- [1] López, Sara Morales, Silvia Angélica Ávila Juárez, Alba Brenda Daniel, Fátima Molina Carrasco, Hugo Erick Olvera Cortés, Ana Gabriela Ortiz Sánchez, y Karla Strassburger Lona. «¿Cómo se construyen los escenarios para la enseñanza basada en simulación clínica?», s. f., 12.
- [2] Dunn, Fletcher, y Ian Parberry. *3D Math Primer for Graphics and Game Development*. Wordware Game Math Library. Plano, Tex: Wordware Pub, 2002.
- [3] Medina Echavarría, Ricardo Adolfo, John Edisson Díaz Figueroa, Edgar Eduardo Gonzalez Villamil, Jorge Andres Perez Oliva, y Alex Hernando Rangel Ospino. «Complemento al simulador de vuelo del helicóptero Bell 206, para las prácticas en la formación de pilotos de la ESAVI». *Revista EIA* 17, n.º 34 (17 de octubre de 2020): 1-14. <https://doi.org/10.24050/reia.v17i34.1327>.
- [4] Ramos, Jorge Luis Estepa, Jorge Luis Estepa Pérez, Tahiluma Santana Pedraza, y Arístides Sánchez. «Construcción de un simulador artesanal para la adquisición de habilidades en cirugía laparoscópica Construction of a craft simulator for the acquisition of skills in laparoscopic surgery», 2020, 6.
- [5] Murillo, Javier García, Mauricio Arias Correa, y Édison Valencia Díaz. «Diseño de prototipo de simulador para entrenamiento en cirugía laparoscópica». *Revista Ingeniería Biomédica*, s. f., 7.
- [6] Monserrat, Carlos, Óscar López, y Mariano Alcañíz. «Estado del Arte en Simulación Quirúrgica», s. f., 8.
- [7] Capua, Nelly Esther, Gerardo Flores, Rodolfo Calamandrei, Adriana Elías, y Ester Olga Graiff. «Evaluación de la adquisición de destrezas con simulador laparoscópico. Experiencia inicial.», s. f., 19.
- [8] Henao, Oscar, Jaime Escallón, Jessica Green, Mónica Farcas, Juan Manuel Sierra, William Sánchez, y Allan Okrainec. «Fundamentos de cirugía laparoscópica en Colombia con telesimulación: una herramienta adicional para la formación integral de cirujanos». *Biomédica* 33, n.º 1 (15 de agosto de 2012): 107-14. <https://doi.org/10.7705/biomedica.v33i1.1425>.
- [9] Guáqueta, Víctor Augusto Beltrán, y Dary Luz Hernández Perilla. «IMPACTO DEL ENTRENAMIENTO CON SIMULADORES DE LAPAROSCOPIA DENTRO DEL PROGRAMA DE CIRUGÍA GENERAL – UNIVERSIDAD DEL ROSARIO», 2013, 96.
- [10] Inzunza Agüero, Martín, José Quezada González, Cristian Jarry Trujillo, Alberto Torres Gueren, Rodrigo Tejos Sufán, Gabriel Escalona Vivas, Valeria Abiuso Baesler, et al. «IMPACTO DEL PRIMER “LEARNING CENTER” DE CIRUGÍA MÍNIMAMENTE INVASIVA EN CHILE». *Revista de Cirugía* 71, n.º 5 (13 de octubre de 2019). <https://doi.org/10.35687/s2452-45492019005365>.
- [11] García-Ceja, Enrique, y Ramon Brena. «Long-Term Activity Recognition from Accelerometer Data». *Procedia Technology* 7 (2013): 248-56. <https://doi.org/10.1016/j.protcy.2013.04.031>.

- [12] Ramachandran, Prabhu, y Gaël Varoquaux. «Mayavi: A Package for 3D Visualization of Scientific Data». *Computing in Science & Engineering* 13, n.º 2 (marzo de 2011): 40-51. <https://doi.org/10.1109/MCSE.2011.35>.
- [13] Liagkou, Vasiliki, Dimitrios Salmas, y Chrysostomos Stylios. «Realizing Virtual Reality Learning Environment for Industry 4.0». *Procedia CIRP* 79 (2019): 712-17. <https://doi.org/10.1016/j.procir.2019.02.025>.
- [14] Ramasubramani, Vyas, y Sharon C. Glotzer. «Rowan: A Python Package for Working with Quaternions». *Journal of Open Source Software* 3, n.º 27 (30 de julio de 2018): 787. <https://doi.org/10.21105/joss.00787>.
- [15] Kerrigan, Nicolas. «Simulación, ¿una Necesidad En El Entrenamiento Para La Cirugía Laparoscópica Colorrectal?» *Revista Chilena de Cirugía* 69, n.º 6 (noviembre de 2017): 508-12. <https://doi.org/10.1016/j.rchic.2017.06.004>.
- [16] Granados, Dr Mauro Soto, y Dr Javier Valencia Rosenberg. «Simulador para el dominio de procedimientos básicos en cirugía laparoscópica. Diseño de un modelo práctico y económico», s. f., 4.
- [17] Ledo, María J Vidal, Raidell Avello Martínez, Mabel A Rodríguez Monteagudo, y José Alberto Menéndez Bravo. «Simuladores como medios de enseñanza», s. f., 13.
- [18] Tovar, Daniel Ricardo Ramos, y Sergio Alexander Salinas. «Simuladores virtuales para entrenamiento de habilidades para laparoscopia». *REVISTA INGENIERÍA BIOMÉDICA*, s. f., 11.
- [19] Heroy, Andrew M, Zackary Gill, Samantha Sprague, David Stroud, y John Santerre. «Stationary Exercise Classification Using IMUs and Deep Learning» 3, n.º 1 (2020): 21.
- [20] Reyes Rocabado, Jimmy, Carlos Escobar Flores, Juan Duarte Vargas, y Pedro Ramirez Peradotto. «UNA APLICACIÓN DEL MODELO DE REGRESIÓN LOGÍSTICA EN LA PREDICCIÓN DEL RENDIMIENTO ESTUDIANTIL». *Estudios pedagógicos (Valdivia)* 33, n.º 2 (2007). <https://doi.org/10.4067/S0718-07052007000200006>.
- [21] Aguilar, E., A. Vivas, y J. Sabater. «Una aproximación a la realidad aumentada y sus aplicaciones quirúrgicas». *Entre ciencia e ingeniería* 12, n.º 24 (12 de diciembre de 2018): 15. <https://doi.org/10.31908/19098367.3811>.
- [22] Aguilar, E, y A Vivas. «Una aproximación a la realidad aumentada y sus aplicaciones quirúrgicas1 An Approach to Augmented Reality and The Surgical Applications Uma Abordagem ao Aumento da Realidade e Suas Aplicações Cirúrgicas» 12, n.º 24 (2018): 10.
- [23] Numfu, Manoch, Andreas Riel, y Frederic Noel. «Virtual Reality Based Digital Chain for Maintenance Training». *Procedia CIRP* 84 (2019): 1069-74. <https://doi.org/10.1016/j.procir.2019.04.268>.