



## Trabajo de Fin de Grado

---

Aplicación en Python para ayudar en la discriminación entre TEA y TDAH mediante el análisis del perfil de velocidad y trayectoria de la escritura

TITULACIÓN: Grado en Ingeniería Informática

AUTOR: Pablo Santana Susilla

---

TUTORIZADO POR:  
María Cristina Carmona Duarte

[Diciembre/2024]

DIRECTORA DE LA ESCUELA DE INGENIERÍA INFORMÁTICA

## Agradecimientos

*A mis padres, Amaya y Marcelo, por su apoyo incondicional durante estos seis años. Gracias por creer en mí, motivarme a seguir adelante, estar siempre a mi lado, incluso en los momentos en los que pensé en rendirme, y por entenderme cuando pasaba días estudiando en mi habitación o cuando apenas tenía tiempo para estar con ustedes.*

*A mi abuelo Javier, quien me vio iniciar esta andadura, pero ya no está. A quienes me apoyaron desde el principio y ya no están, a quienes han permanecido conmigo y a quienes se sumaron en el camino: este logro también es gracias a ustedes.*

*También quiero expresar mi agradecimiento a María Cristina, cuya orientación y compromiso durante la supervisión de este trabajo han sido fundamentales para alcanzar los objetivos propuestos. Su dedicación y profesionalismo han sido un pilar esencial en el desarrollo de este proyecto.*

*Este TFT ha sido realizado gracias al proyecto de investigación  
PID2021-122687OA-I00, financiado por  
MCIN/AEI/10.13039/501100011033/FEDER, UE.*

# Resumen

Este trabajo de fin de título tiene como objetivo desarrollar una aplicación en Python que contribuya a la discriminación entre el Trastorno del Espectro Autista (TEA) y el Trastorno por Déficit de Atención e Hiperactividad (TDAH), a través del análisis del perfil de velocidad y trayectoria de la escritura. La motivación principal del estudio radica en la necesidad de mejorar el diagnóstico diferencial entre ambos trastornos, que a menudo presentan síntomas solapados y cuya detección temprana puede facilitar intervenciones más eficaces.

La aplicación desarrollada analiza dibujos de cuadrados hechos por sujetos en una tablet, utilizando técnicas de visión por computador para detectar patrones geométricos y estudiar el perfil de velocidad de los trazos. Se realizaron análisis estadísticos para identificar diferencias significativas entre los diferentes grupos. Los resultados obtenidos mostraron que la aplicación es capaz de discriminar de manera efectiva entre los distintos diagnósticos, destacando la variabilidad en la velocidad de escritura de los niños con TDAH y la rigidez y consistencia en los trazos de los niños con TEA, características que los diferencian de los pacientes neurotípicos.

Este proyecto propone una herramienta complementaria para los profesionales de la salud, que puede mejorar la precisión del diagnóstico de estos trastornos del desarrollo y servir como base para futuras investigaciones en el campo de la neurociencia.

# Abstract

The aim of this final degree project is to develop a Python application that contributes to the discrimination between Autism Spectrum Disorder (ASD) and Attention Deficit Hyperactivity Disorder (ADHD), through the analysis of the writing speed and trajectory profile. The main motivation for the study lies in the need to improve the differential diagnosis between the two disorders, which often have overlapping symptoms and whose early detection can facilitate more effective interventions.

The developed application analyses drawings of squares made by subjects on a tablet, using computer vision techniques to detect geometric patterns and study the speed profile of the strokes. Statistical analyses were performed to identify significant differences between the different groups. The results obtained showed that the application is able to effectively discriminate between the different diagnoses, highlighting the variability in the writing speed of children with ADHD and the rigidity and consistency in the strokes of children with ASD, characteristics that differentiate them from neurotypical patients.

This project proposes a complementary tool for health professionals, which can improve the diagnostic accuracy of these developmental disorders and serve as a basis for future research in the field of neuroscience.

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación y ámbito de trabajo . . . . .	2
1.2. Recursos hardware y software utilizados . . . . .	2
1.2.1. Hardware . . . . .	2
1.2.2. Software . . . . .	3
<b>2. Estado actual y objetivos iniciales</b>	<b>5</b>
2.1. Estado del arte . . . . .	5
2.2. Objetivos . . . . .	7
<b>3. Competencias específicas y aportaciones del trabajo</b>	<b>8</b>
3.1. Competencias específicas . . . . .	8
3.2. Aportaciones del trabajo . . . . .	9
<b>4. Desarrollo</b>	<b>11</b>
4.1. Planificación y metodología . . . . .	11
4.1.1. Metodología . . . . .	11
4.1.2. Planificación inicial . . . . .	12
4.1.3. Planificación real . . . . .	13
4.1.4. Organización de los sprints . . . . .	14
4.2. Recopilación y análisis de datos . . . . .	16
4.2.1. Base de datos . . . . .	16
4.2.2. Proceso de recopilación de datos . . . . .	17
4.2.3. Conceptos y Herramientas de Análisis . . . . .	18
4.3. Códigos y algoritmos desarrollados . . . . .	22
4.3.1. Detección de cuadrados . . . . .	23
4.3.2. Cálculo y representación de porcentaje de cuadrados . . . . .	26
4.3.3. Cálculo de similitud del perfil de velocidad de la escritura . . . . .	33
4.3.4. Cálculo de diferencia entre dibujos de cuadrados . . . . .	35
<b>5. Discusión de los resultados obtenidos</b>	<b>40</b>
5.1. Distribución de porcentaje de cuadrados encontrados . . . . .	40
5.2. Cálculo de similitud del perfil de velocidad de la escritura . . . . .	42
5.3. Diferencia de Hausdorff entre cuadrados . . . . .	43

5.4. Comparación de resultados con trabajos previos . . . . .	44
5.5. Limitaciones del estudio . . . . .	45
<b>6. Conclusiones y trabajo futuro</b>	<b>46</b>
6.1. Conclusiones . . . . .	46
6.2. Propuestas para investigación futura . . . . .	47
<b>A. Anexo: Manual de usuario</b>	<b>48</b>
<b>B. Anexo: Presupuesto</b>	<b>51</b>
<b>C. Anexo: Normativa y Legislación</b>	<b>54</b>
C.1. Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales . . . . .	54
C.2. Reglamento General de Protección de Datos . . . . .	54
C.3. Normas Éticas en Investigación . . . . .	55
C.4. Consentimiento Informado para la Participación en el Estudio . . . . .	55
C.5. Legislación Adicional Relacionada . . . . .	56

# Índice de figuras

1.1. Logos de las herramientas utilizadas en el proyecto. . . . .	4
4.1. Esquema general de la metodología ágil . . . . .	11
4.2. Ejemplo de algunos commits del proyecto utilizando Git. . . . .	12
4.3. Estructura de ambas bases de datos . . . . .	17
4.4. Ejemplo de alineación de series temporales mediante DTW. . . . .	19
4.5. Ejemplo de comparación de conjuntos mediante la Diferencia de Hausdorff .	20
4.6. Estructura del sistema de ficheros . . . . .	23
4.7. Resultado del algoritmo de detección de cuadrados para un varón de 10 años diagnosticado con TEA . . . . .	25
4.8. Resultado del algoritmo de detección de cuadrados para una niña de 10 años diagnosticada con TEA . . . . .	25
4.9. Porcentajes de cuadrados encontrados en pacientes masculinos neurotípicos .	27
4.10. Porcentajes de cuadrados encontrados en pacientes masculinos con TDAH . .	28
4.11. Porcentajes de cuadrados encontrados en pacientes masculinos con TEA . . .	28
4.12. Porcentajes de cuadrados encontrados en pacientes femeninas neurotípicas .	29
4.13. Porcentajes de cuadrados encontrados en pacientes femeninas con TDAH . .	30
4.14. Porcentajes de cuadrados encontrados en pacientes femeninas con TEA . . .	31
4.15. Distribución de porcentaje de cuadrados encontrados para pacientes masculinos	31
4.16. Distribución de porcentaje de cuadrados encontrados para pacientes femeninas	32
4.17. Distribución de similitud del perfil de velocidad neurotípicos . . . . .	33
4.18. Distribución de similitud del perfil de velocidad neurotípicas . . . . .	34
4.19. Distribución de similitud del perfil de velocidad por grupos . . . . .	34
4.20. Cálculo de diferencia entre cuadrados de pacientes neurotípicos . . . . .	36
4.21. Cálculo de diferencia entre cuadrados de pacientes femeninas neurotípicas . .	37
4.22. Cálculo de diferencia entre cuadrados de pacientes masculinos con TEA . . .	37
4.23. Cálculo de diferencia entre cuadrados de pacientes femeninas con TEA . . .	38
4.24. Cálculo de diferencia entre cuadrados de pacientes masculinos con TDAH . .	38
4.25. Cálculo de diferencia entre cuadrados de pacientes femeninas con TDAH . .	39
A.1. Estructura de la carpeta automatic_versions . . . . .	49
A.2. Botón de selección de sexo . . . . .	50
A.3. Botón de selección para distance_dtw.py . . . . .	50

# Índice de cuadros

4.1. Planificación inicial del proyecto . . . . .	13
4.2. Planificación real del proyecto . . . . .	14
4.3. Planificación de los sprints en el Proyecto . . . . .	16
4.4. Organización interna de cada archivo ‘.csv’ utilizado en el proyecto. . . . .	16
4.5. Distribución de pacientes por diagnóstico y sexo . . . . .	16
4.6. Número de dibujos en cada base de datos . . . . .	17
4.7. Resultados de la prueba t de Student para pacientes masculinos neurotípicos	27
4.8. Resultados de la prueba ANOVA para pacientes masculinos neurotípicos . .	27
4.9. Resultados de la prueba t para pacientes femeninas neurotípicas . . . . .	29
4.10. Resultados de la prueba ANOVA para pacientes femeninas neurotípicas . . .	29
4.11. Diagnósticos disponibles por edad y diagnóstico de pacientes masculinos . . .	32
4.12. Diagnósticos disponibles por edad y diagnóstico de pacientes femeninas . . .	33
4.13. Prueba T para comparaciones entre diagnósticos (9-10 años) . . . . .	35
4.14. Resultado de la prueba ANOVA para los grupos de 9-10 años . . . . .	35

# Índice de Algoritmo

4.1. Ejemplo de <code>cv2.approxPolyDP()</code> para la detección de cuadrados . . . . .	22
4.2. Reescalado de dibujos realizados . . . . .	24
4.3. Detección de cuadrados en imágenes . . . . .	24
4.4. Detección de cuadrados y eliminación de los no cuadrados . . . . .	26
4.5. Reescalado de cuadrados detectados . . . . .	35
4.6. Cálculo de la diferencia de Hausdorff entre cuadrados . . . . .	36

# Capítulo 1

## Introducción

‘Los números no mienten, pero necesitan un intérprete.’

---

Stephen Few

El Trastorno del Espectro Autista (TEA)[34] y el Trastorno por Déficit de Atención e Hiperactividad (TDAH)[35] son condiciones neuropsiquiátricas que impactan significativamente la vida de las personas que los padecen, y aunque ambas patologías comparten algunas características, difieren significativamente en sus manifestaciones clínicas y efectos sobre la vida diaria.

En primer lugar, ambos trastornos pueden afectar a la ejecución de tareas que requieren coordinación motora fina y atención sostenida, como la escritura. Sin embargo, las diferencias en la naturaleza de estas dificultades son notables: mientras que el TDAH se caracteriza por dificultades para mantener la atención e impulsividad, el TEA está más relacionado con dificultades en la interacción social, la comunicación y conductas repetitivas o restrictivas.

Investigaciones recientes han señalado que tanto las personas con TEA como con TDAH muestran patrones únicos en la velocidad y la trayectoria de la escritura [11]. Así, por ejemplo, individuos con TDAH pueden exhibir variabilidad en la velocidad debido a la falta de atención y control impulsivo [8], mientras que aquellos con TEA podrían mostrar patrones más consistentes, pero menos flexibles debido a la rigidez motora y comportamientos repetitivos.

En este contexto, la propuesta de desarrollar una aplicación en Python para analizar el perfil de velocidad y trayectoria de la escritura cobra relevancia. El proyecto incluye el análisis de datos recogidos a partir de pruebas de escritura realizadas por individuos con diferentes diagnósticos (TEA, TDAH y neurotípicos), organizados por edad y sexo. Estos datos han sido procesados mediante algoritmos que permiten extraer métricas clave, como la velocidad y la trayectoria. Esta herramienta busca complementar los métodos de diagnóstico convencionales, facilitando una detección temprana y precisa que permita intervenciones más efectivas y mejora la calidad de vida de las personas afectadas y sus familias.

## 1.1. Motivación y ámbito de trabajo

La motivación detrás de este proyecto radica en la necesidad de proporcionar herramientas innovadoras y efectivas que complementen los métodos tradicionales de diagnóstico. El diagnóstico diferencial entre el TEA y el TDAH puede ser complejo debido a la superposición de síntomas y la variabilidad individual en la presentación de ambos trastornos. La escritura, como forma de expresión motora compleja, ofrece un terreno fértil para explorar patrones distintivos que pueden diferenciar estos trastornos.

Este trabajo no busca descubrir nuevas características, sino aprovechar las ya conocidas, como la velocidad y la trayectoria de la escritura, para una detección rápida y precisa. El diagnóstico precoz es crucial, ya que posibilita implementar tratamientos adecuados durante las etapas formativas del desarrollo de los pacientes.

En última instancia, este proyecto aspira a contribuir significativamente a la comprensión científica de estas patologías, abriendo nuevas vías para la investigación y el desarrollo de tecnologías que puedan transformar positivamente la práctica clínica y educativa en el ámbito del neurodesarrollo infantil.

## 1.2. Recursos hardware y software utilizados

Para el desarrollo de este trabajo de fin de título se han utilizado diversos recursos informáticos, tanto de hardware como de software, entre los que destacan:

### 1.2.1. Hardware

- **Ordenador portátil:** Lenovo IdeaPad S340
  - **Sistema Operativo:** Windows 10 de 64 bits.
  - **RAM instalada:** 8GB.
  - **Procesador:** Intel(R) Core(TM) i7-8565U CPU @ 1.80GHz 1.99 GHz.
  - **Núcleos:** 4 núcleos.
- **iPad (2018):** Utilizado para la toma de datos de los distintos pacientes.
- **Apple Pencil:** Utilizado para realizar las pruebas en el iPad.

## 1.2.2. Software

En lo que al software se refiere, se ha hecho uso de:

- **Visual Studio Code v1.90.2 (\*User set up\*)**: utilizado como editor de código fuente.
- **Python (versión 3.11.7)**: como lenguaje de alto nivel [1].
- **PDData**: Para la toma de datos se utilizó esta aplicación desarrollada en la Universidad de Las Palmas de Gran Canaria disponible en TestFlight [2].

### 1.2.2.1. Bibliotecas y herramientas utilizadas

Para llevar a cabo el análisis y desarrollo del proyecto, se utilizó un entorno virtual de Python, también conocido como ‘environment’ [26], con diversas bibliotecas y herramientas de software.

El uso de este entorno virtual permitió aislar las dependencias específicas del proyecto, evitando conflictos con otros proyectos y garantizando la reproducibilidad de los resultados. Además, facilitó la gestión de las versiones de las bibliotecas, asegurando que todas las herramientas utilizadas fueran compatibles entre sí y contribuyendo a un desarrollo eficiente y organizado.

A continuación, se detalla la lista de las bibliotecas más destacables utilizadas junto con una breve descripción de su propósito:

- **Fastdtw (versión 0.3.4)**: Paquete que permite una implementación eficiente y aproximada del cálculo de la distancia DTW para secuencias largas, optimizando el rendimiento del análisis de datos temporales [28].
- **GitHub**: Plataforma de desarrollo colaborativo que utiliza el sistema de control de versiones Git. Utilizada para gestionar el código fuente del proyecto, realizar un seguimiento de los cambios y colaborar con otros desarrolladores de manera eficiente [16].
- **GPT-4**: Utilizado para la corrección y mejora del código y del texto presente en esta memoria [24].
- **Matplotlib (versión 3.4.2)**: Utilizada para la visualización de datos y generación de gráficos, como los dibujos realizados, sus respectivas gráficas de velocidad y las figuras que se mostrarán en la sección códigos desarrollados [18].
- **NumPy (versión 1.20.3)**: Utilizada para operaciones matemáticas y manejo de matrices. Utilizada principalmente para realizar operaciones numéricas eficientes en arrays multidimensionales [17].
- **OpenCV (versión 4.5.1)**: Utilizada para el procesamiento de imágenes y visión por computadora. Utilizada para detectar y dibujar contornos de cuadrados en imágenes generadas a partir de datos CSV, facilitando la identificación automática de formas geométricas en un contexto de visión por computador [19].

- **Overleaf**: Plataforma en línea que permite a los usuarios escribir, editar y colaborar en documentos científicos y académicos utilizando LaTeX, un sistema de composición de textos especializado en la producción de documentos técnicos y científicos [25].
- **Pandas (versión 1.2.4)**: Utilizada para la manipulación y análisis de datos. En el caso de este trabajo, se utiliza específicamente para leer archivos CSV y manipular datos tabulares [21].
- **SciPy (versión 1.6.3)**: Utilizada para estadísticas y funciones científicas. En este caso, se utiliza específicamente para importar una función estadística necesaria para el análisis. En concreto, se importa `ttest_ind` desde `scipy.stats`. Esta función se utilizó para realizar la prueba t de Student para comparar las medias de dos muestras independientes [31].
- **tk (versión 8.6.12)**: Utilizada para la creación de interfaces gráficas de usuario. Ha sido usada para implementar una ventana de selección de directorio a través de un botón [14].

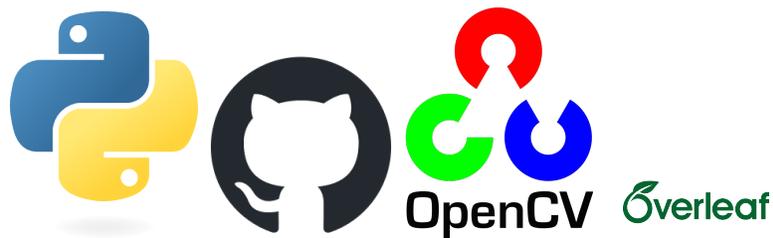


Figura 1.1: Logos de las herramientas utilizadas en el proyecto.

---

<sup>1</sup>Los logos fueron obtenidos de sus respectivas páginas oficiales: Python (<https://www.python.org>), OpenCV (<https://opencv.org>), GitHub (<https://github.com/logos>), y Overleaf (<https://www.overleaf.com>).

# Capítulo 2

## Estado actual y objetivos iniciales

### 2.1. Estado del arte

El Trastorno por Déficit de Atención e Hiperactividad (TDAH) y el Trastorno del Espectro Autista (TEA) son dos de los trastornos neurodesarrollados más comunes en la infancia, con una prevalencia significativa en la población infantil [20]. Según el manual “Diagnóstico y Estadístico de los Trastornos Mentales”(DSM-5)[3], el TDAH se caracteriza por *“síntomas de inatención, hiperactividad e impulsividad que pueden afectar el rendimiento académico y social del individuo”*, mientras que el TEA se caracteriza por *“el deterioro persistente de la comunicación social recíproca y la interacción social, y los patrones de conducta, intereses o actividades restrictivos y repetitivos”*.

A pesar de su prevalencia, la bibliografía científica presenta una carencia respecto a la representación de género en la investigación de estos trastornos, especialmente en relación a la participación del sexo femenino en estudios clínicos y la manifestación de síntomas en este grupo de población.

Estudios epidemiológicos han demostrado una predominancia de casos diagnosticados en niños en comparación con niñas, lo que sugiere un desequilibrio en las relaciones de género [20]. Dicha subrepresentación ha generado una carencia de datos específicos sobre la presentación y características del trastorno en pacientes femeninas, lo que dificulta la comprensión completa de su naturaleza y el desarrollo de intervenciones efectivas.

En la investigación de Mahendiran et al. [20], se examinó las diferencias de género en la función adaptativa social, proporcionando información valiosa sobre la presentación y características de estos trastornos en ambos sexos. Los hallazgos encontrados subrayan la importancia de considerar estas diferencias en la investigación y el tratamiento, destacando la necesidad de reducir la brecha en la representación de niñas.

Además, investigaciones recientes han identificado el fenómeno del ‘masking’ o camuflaje en el sexo femenino con TEA, en el que el individuo oculta o compensa los signos del trastorno para mantener o adaptarse a las expectativas y normas sociales [10]. Este camuflaje

hace que los síntomas de TEA en mujeres puedan pasar desapercibidos o ser subestimados, dificultando así su diagnóstico y, en consecuencia, aumentando la subrepresentación de niñas en estudios clínicos. A menudo, esta estrategia de adaptación genera diagnósticos tardíos o erróneos, afectando negativamente la intervención y el tratamiento temprano, y subrayando la necesidad de mejorar las herramientas de diagnóstico con una perspectiva de género. Esto concuerda con la investigación de Mahendiran et al. [20], que señala que los síntomas de las niñas con TEA tienden a ser menos evidentes y más internalizados que los niños. Esta situación puede llevar a una intervención inadecuada para las niñas, perpetuando las disparidades en el tratamiento y los resultados clínicos.

Otras investigaciones recientes también han identificado diferencias significativas en la presentación de síntomas entre niños y niñas con TDAH, mostrando síntomas más relacionados con la inatención en las niñas [22]. Esta falta de reconocimiento temprano agrava las brechas en la representación y el tratamiento adecuado para las niñas con TDAH.

En el estudio de Cohen et al. [8], sobre la escritura a mano en pacientes con TDAH, se reveló que presentan una escritura menos legible, más desorganizada y más lenta en comparación con aquellos sujetos sin este trastorno. Estos hallazgos subrayan la relevancia de las características motoras en estos trastornos como un aspecto importante a considerar. En el TDAH, los estudios han demostrado problemas en la coordinación temporal, lo que afecta a tareas que requieren ritmo y sincronización. Por otro lado, en el TEA, las dificultades se centran en la integración de movimientos coordinados, especialmente en tareas que exigen sincronización visual y motora [29]. Estas diferencias en las habilidades motoras y en la escritura podrían servir como biomarcadores diagnósticos, aunque es necesario realizar más investigaciones para profundizar en estas características y su impacto en ambos trastornos.

En este contexto, las técnicas de visión por computador y la detección de formas geométricas pueden ser útiles para identificar patrones específicos asociados con TDAH y TEA en tareas que impliquen habilidades motoras finas como la escritura. La detección de formas en imágenes es una herramienta potente que permite el análisis automático de características visuales. Por ejemplo, para detectar cuadrados utilizando OpenCV en Python, se sigue un enfoque de varios pasos: primero, la conversión de la imagen a escala de grises; después, la aplicación de un desenfoque gaussiano para reducir el ruido; luego, el uso del algoritmo de Canny para detectar bordes; y, finalmente, la identificación de contornos que correspondan a cuadrados mediante la aproximación de contornos y la verificación de propiedades geométricas [15]. Este método ha demostrado ser eficaz para detectar formas en entornos controlados y puede ser adaptado para analizar patrones de escritura, como la forma de las letras o los trazos, en busca de irregularidades relacionadas con TDAH o TEA.

Dada la escasez de datos específicos sobre niñas, resulta fundamental abordar esta brecha para mejorar la atención y los resultados clínicos en este grupo menos representado. Se propone desarrollar una aplicación que utilice la velocidad y trayectoria de la escritura para diferenciar entre ambas patologías, con especial atención a las diferencias de género. En el desarrollo de la aplicación separar los datos por género garantizará su precisión y efectividad, contribuyendo de esta manera a reducir la brecha en la representación de niñas en la investigación y mejorando su temprana intervención y los resultados para todos los afectados.

## 2.2. Objetivos

El objetivo principal de este trabajo de fin de título es desarrollar una aplicación que ayude en la discriminación entre TEA y TDAH. La herramienta propuesta se basa en el análisis del perfil de velocidad y la trayectoria de la escritura, proporcionando un enfoque cuantitativo para identificar características específicas asociadas a estos trastornos del neurodesarrollo.

A partir de este objetivo general, se establecen los siguientes objetivos específicos:

1. **Mejorar el diagnóstico temprano:** Facilitar y ayudar la detección precoz de TEA y TDAH mediante el análisis objetivo de características motoras específicas de la escritura.
2. **Reducir la brecha en la representación de pacientes femeninas:** Contribuir a reducir la brecha en la representación de pacientes femeninas diagnosticadas con ambas patologías, mejorando su temprana intervención y los resultados para todos los afectados.
3. **Incrementar la precisión y objetividad:** Proporcionar un enfoque cuantitativo para la evaluación de estos trastornos del desarrollo, complementando los métodos tradicionales que pueden llegar a ser subjetivos y depender del profesional encargado de evaluar al paciente (psiquiatra, psicólogo clínico, neurólogo o pediatra).
4. **Ofrecer soporte a profesionales de la salud:** Asistir a los profesionales de la salud en la toma de decisiones clínica ofreciendo datos y métricas concretas que respalden el diagnóstico.
5. **Promover la investigación:** Servir como herramienta para la investigación continua proporcionando datos que podrían generar nuevos métodos de intervención y diagnóstico.

# Capítulo 3

## Competencias específicas y aportaciones del trabajo

### 3.1. Competencias específicas

La realización de este trabajo ha abordado las siguientes competencias del Grado en Ingeniería Informática [13] impartido en la Universidad de Las Palmas de Gran Canaria:

- TI3: *“Capacidad para emplear metodologías centradas en el usuario y la organización para el desarrollo, evaluación y gestión de aplicaciones y sistemas basados en tecnologías de la información que aseguren la accesibilidad, ergonomía y usabilidad de los sistemas.”*
  - El diseño y desarrollo de la aplicación en Python incluyeron la implementación de una interfaz de usuario intuitiva y fácil de usar, especialmente para profesionales de la salud, educadores y pacientes.
  - La aplicación buscó ser una herramienta efectiva para complementar métodos de diagnóstico tradicionales, asegurando que fuera accesible para sus usuarios.
  - La validación continua durante el desarrollo del proyecto aseguró que la aplicación no solo fuera funcional, sino también práctica y útil, reforzando la usabilidad y accesibilidad.
- TI5: *“Capacidad para seleccionar, desplegar, integrar y gestionar sistemas de información que satisfagan las necesidades de la organización, con los criterios de coste y calidad identificados.”*
  - El proyecto incluyó la adquisición de datos de velocidad de escritura usando tablets y su procesamiento mediante diferentes algoritmos, seleccionando herramientas y tecnologías en función de su efectividad para los objetivos del proyecto.
  - El uso de Python, una herramienta ampliamente reconocida por su eficiencia y bajo coste, mostró una consideración cuidadosa de los costes asociados al desarrollo y mantenimiento del sistema.

## CAPÍTULO 3. COMPETENCIAS ESPECÍFICAS Y APORTACIONES DEL TRABAJO9

- CI01: *“Capacidad para diseñar, desarrollar, seleccionar y evaluar aplicaciones y sistemas informáticos, asegurando su fiabilidad, seguridad y calidad, conforme a principios éticos y a la legislación y normativa vigente.”*
  - Durante el desarrollo de la aplicación, se demostró la capacidad de diseñar y desarrollar una solución informática desde cero. Utilizando Python y diversas bibliotecas como NumPy y Pandas, se implementaron algoritmos para analizar el perfil de velocidad y trayectoria de la escritura.
  - Se utilizaron dispositivos precisos como el iPad y el Apple Pencil para la captura de datos, garantizando la calidad y precisión de la información obtenida.
  - Se respetaron estrictos principios éticos, asegurando el consentimiento informado de los participantes en todas las etapas del proyecto.
- G8: *“Conocimiento de las materias básicas y tecnologías que capaciten para el aprendizaje y desarrollo de nuevos métodos y tecnologías, así como las que les doten de una gran versatilidad para adaptarse a nuevas situaciones.”*
  - Durante el desarrollo de este trabajo, se ha hecho uso de diversas herramientas y tecnologías, tales como Python, OpenCV y algoritmos de aprendizaje automático para analizar patrones en la escritura. La capacidad para integrar diferentes bibliotecas como NumPy, Pandas y SciPy en la solución propuesta muestra la versatilidad adquirida.
- G9: *“Capacidad para resolver problemas con iniciativa, toma de decisiones, autonomía y creatividad. Capacidad para saber comunicar y transmitir los conocimientos, habilidades y destrezas de la profesión de Ingeniero Técnico en Informática.”*
  - Durante el desarrollo de este proyecto, se presentaron problemas complejos vinculados con la clasificación de datos y la identificación de formas geométricas que fueron resueltos mediante la implementación de soluciones creativas. La autonomía en la toma de decisiones para seleccionar tecnologías y ajustar los parámetros de los algoritmos resultó ser clave.
  - Se ha logrado mostrar los resultados obtenidos de manera efectiva a través de una interfaz gráfica sencilla y accesible, facilitando la interpretación de los datos.

### 3.2. Aportaciones del trabajo

Este proyecto aporta valor al campo de la Ingeniería Informática aplicada a la salud a través del desarrollo de una herramienta práctica e innovadora que aborda las necesidades del diagnóstico diferencial de TEA y TDAH. Las principales contribuciones son:

1. **Diseño e implementación de algoritmos especializados:** Se han desarrollado diferentes algoritmos con los que poder visualizar y estudiar diferentes parámetros, tanto de los dibujos realizados por los pacientes, como de sus respectivas gráfica de

velocidad y tiempo. Estos además integran una interfaz con la que los clínicos puedan visualizar y añadir nuevos datos.

2. **Integración de tecnología en el ámbito clínico:** La integración de técnicas de procesamiento de señales y análisis de datos en un contexto clínico, proporcionando un enfoque objetivo y cuantitativo para la evaluación de los trastornos estudiados.
3. **Impacto clínico directo:** La aplicación está diseñada para ser utilizada por los clínicos, permitiendo que los médicos visualicen los resultados a través de una interfaz accesible e intuitiva, lo que facilita el diagnóstico y seguimiento de pacientes con sospecha de TEA o TDAH. Esto supone una mejora en la atención y tratamiento de estos trastornos.
4. **Contribución al conocimiento:** Contribuciones teóricas y prácticas al campo de la psicología y la neurociencia computacional, al proporcionar nuevas perspectivas sobre cómo las características motoras pueden reflejar diferencias neuropsicológicas sutiles entre los trastornos estudiados.
5. **Implementación de una interfaz gráfica intuitiva:** Diseño de una interfaz que permite a usuarios no técnicos ejecutar análisis de forma sencilla, favoreciendo su adopción en distintos entornos clínicos y académicos.

# Capítulo 4

## Desarrollo

### 4.1. Planificación y metodología

#### 4.1.1. Metodología

Para organizar el desarrollo del proyecto, se adoptó una metodología ágil [27], que permitió la flexibilidad necesaria para adaptarse a los cambios y nuevas necesidades surgidas durante el trabajo. Este enfoque se basó en la entrega continua de valor a través de iteraciones cortas, denominadas sprints. Cada sprint incluyó objetivos claros, reuniones periódicas de seguimiento con María Cristina, tutora de este trabajo, y revisiones que fomentaron la mejora continua y la adaptación de las estrategias iniciales, tal y como se puede observar en el esquema presentado en la figura 4.1.

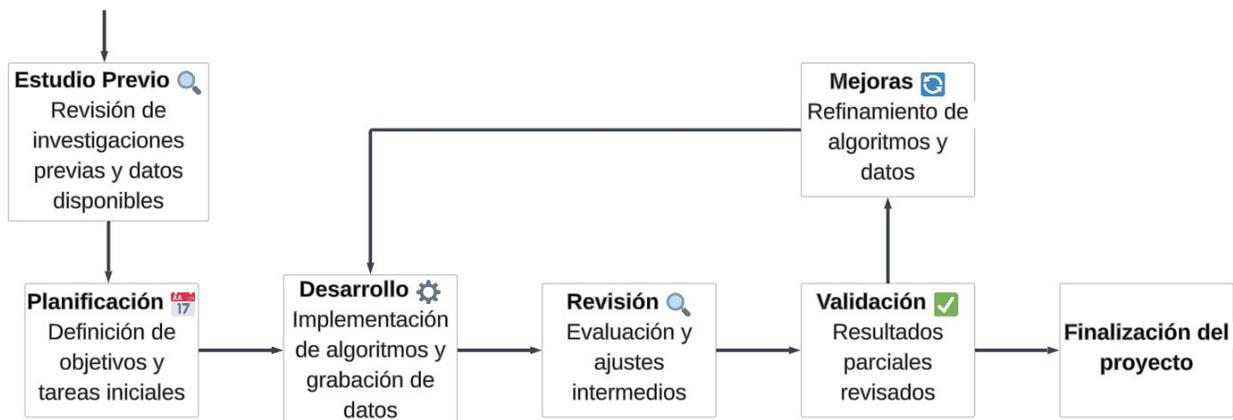
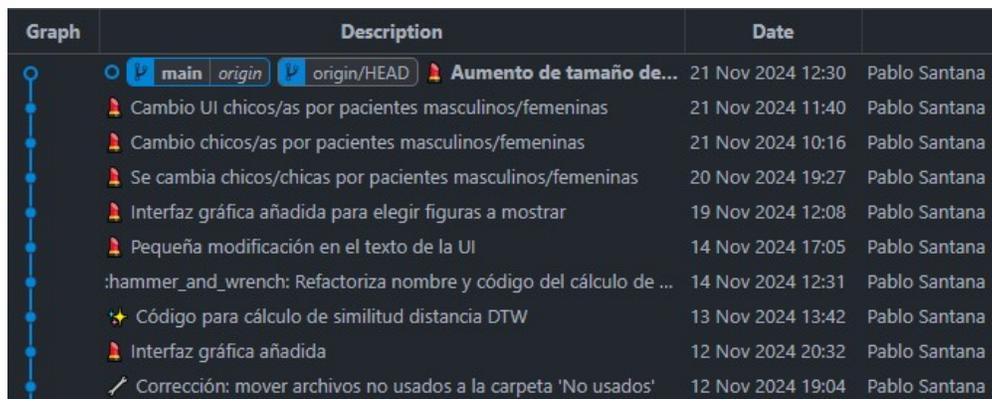


Figura 4.1: Esquema general de la metodología ágil

Aunque el trabajo se llevó a cabo en una única rama, esta herramienta permitió documentar el progreso de manera efectiva, asegurando la trazabilidad de las modificaciones realizadas y facilitando la recuperación de versiones anteriores en caso de ser necesario (ver figura 4.2).



Graph	Description	Date	
○	main origin origin/HEAD 🔥 Aumento de tamaño de...	21 Nov 2024 12:30	Pablo Santana
🔥	Cambio UI chicos/as por pacientes masculinos/femeninas	21 Nov 2024 11:40	Pablo Santana
🔥	Cambio chicos/as por pacientes masculinos/femeninas	21 Nov 2024 10:16	Pablo Santana
🔥	Se cambia chicos/chicas por pacientes masculinos/femeninas	20 Nov 2024 19:27	Pablo Santana
🔥	Interfaz gráfica añadida para elegir figuras a mostrar	19 Nov 2024 12:08	Pablo Santana
🔥	Pequeña modificación en el texto de la UI	14 Nov 2024 17:05	Pablo Santana
🔨	:hammer_and_wrench: Refactoriza nombre y código del cálculo de ...	14 Nov 2024 12:31	Pablo Santana
🌟	🌟 Código para cálculo de similitud distancia DTW	13 Nov 2024 13:42	Pablo Santana
🔥	Interfaz gráfica añadida	12 Nov 2024 20:32	Pablo Santana
✍️	✍️ Corrección: mover archivos no usados a la carpeta 'No usados'	12 Nov 2024 19:04	Pablo Santana

Figura 4.2: Ejemplo de algunos commits del proyecto utilizando Git.

A lo largo de cada iteración, se definieron tareas específicas que fueron evaluadas y ajustadas conforme avanzaba el proyecto, permitiendo priorizar actividades según la urgencia y relevancia. Este enfoque combinó la metodología ágil con el uso de herramientas técnicas como Git para ofrecer un desarrollo eficiente y bien documentado.

En la planificación inicial, se definieron las fases principales y sus respectivas tareas, estimando la duración de cada una. Posteriormente, con el avance del proyecto, estas fases se dividieron en sprints para gestionar mejor el trabajo y priorizar las actividades según la urgencia y relevancia del momento.

#### 4.1.2. Planificación inicial

La planificación inicial del proyecto consistió en la definición de las fases principales que compondrían el trabajo, así como una estimación preliminar de la duración de cada una de ellas. La tabla 4.1 muestra un resumen de esta planificación, desglosando cada fase en sus respectivas tareas y asignando una estimación de horas para su ejecución.

Fases	Duración Estimada (horas)	Tareas
Estudio previo / Análisis	60	Tarea 1.1: Estudio de las investigaciones previas realizadas en el campo. Tarea 1.2: Estudio de las características de la base de datos previa.

*Continúa en la siguiente página*

Fases	Duración Estimada (horas)	Tareas
		Tarea 1.3: Estudio de las herramientas que se van a utilizar, métodos de parametrización y de clasificación automática.
Diseño / Desarrollo / Implementación	100	Tarea 2.1: Grabación de nuevos datos. Tarea 2.2: Desarrollo de diferentes técnicas de parametrización y clasificación adecuadas al número de muestras. Tarea 2.3: Desarrollo de una aplicación en Python para la representación de los resultados y visualización de los datos.
Evaluación / Validación / Prueba	80	Tarea 3.1: Evaluación de los resultados. Tarea 3.2: Mejora de los algoritmos de detección y parametrización.
Documentación / Presentación	60	Tarea 4.1: Redacción y revisión de la memoria. Tarea 4.2: Redacción y preparación de la presentación para la defensa.

Tabla 4.1: Planificación inicial del proyecto

### 4.1.3. Planificación real

Tal y como se indicó anteriormente, el desarrollo del proyecto siguió una metodología ágil, organizada en sprints de entre dos a cuatro semanas, lo que permitió ajustar el plan conforme avanzaba el trabajo. Cada sprint incluía revisiones periódicas para evaluar el progreso y realizar las modificaciones necesarias.

En primer lugar, se realizó un análisis exhaustivo de investigaciones previas sobre TEA y TDAH en relación con los perfiles de velocidad de escritura, partiendo del trabajo de María Cristina Carmona Duarte, *‘Assessment of Autism and ADHD: A Comparative Analysis of Drawing Velocity Profiles and the NEPSY Test’* [11]. A partir de este punto de partida, se realizó una búsqueda más profunda de artículos que pudieran complementar y expandir la investigación existente.

En paralelo, se desarrollaron varios algoritmos y códigos para visualizar los datos recolectados, tanto los dibujos de cuadrados realizados por los pacientes, como las gráficas de velocidad correspondientes. Participé en el proceso de recolección de datos en dos ocasiones, lo que permitió refinar los algoritmos de análisis de manera iterativa.

A medida que avanzaban los sprints, los resultados obtenidos fueron revisados y validados periódicamente en reuniones con mi tutora, lo que permitió realizar ajustes continuos en los

códigos y mejorar la representación de los datos. La siguiente tabla 4.2 muestra con mayor detalle las fases y su duración.

Fases	Duración (horas)	Tareas
Revisión Bibliográfica	40 horas	Tarea 1.1: Revisión de investigaciones previas realizadas en el campo.
	10 horas	Tarea 1.2: Estudio de la base de datos y análisis de artículos clave.
	10 horas	Tarea 1.3: Selección de herramientas y algoritmos.
Diseño de la Aplicación	10 horas	Tarea 2.1: Grabación de nuevos datos
	60 horas	Tarea 2.2: Prototipado y refinamiento del diseño de la interfaz.
	30 horas	Tarea 2.3: Desarrollo de la aplicación en Python para la visualización de resultados.
Desarrollo y Prueba	30 horas	Tarea 3.1: Implementación de algoritmos y extracción de características.
	50 horas	Tarea 3.2: Mejora continua de algoritmos de análisis y visualización.
Validación y Documentación	45 horas	Tarea 4.1: Redacción y revisión de la memoria.
	15 horas	Tarea 4.2: Preparación de la presentación para la defensa.

Tabla 4.2: Planificación real del proyecto

#### 4.1.4. Organización de los sprints

El proyecto se organizó en sprints de entre dos a cuatro semanas, cada uno enfocado en un conjunto específico de tareas según la planificación real, permitiendo así una adaptación y revisión continua. La siguiente tabla 4.3 muestra la distribución de tareas en cada sprint.

Sprint	Duración (semanas)	Tareas y Objetivo
Sprint 1	Semanas 1 a 2	Tarea 1.1: Revisión exhaustiva de investigaciones previas. Objetivo: Establecer una base teórica sólida en relación con TEA y TDAH en perfiles de velocidad de escritura, analizando estudios previos relevantes.

*Continúa en la siguiente página*

<b>Sprint</b>	<b>Duración (semanas)</b>	<b>Tareas y Objetivo</b>
Sprint 2	Semana 3	Tarea 1.2: Estudio detallado de la base de datos y análisis de artículos clave. Objetivo: Comprender la estructura de los datos y seleccionar artículos clave para complementar la investigación teórica.
Sprint 3	Semana 4	Tarea 1.3: Selección de herramientas y algoritmos adecuados para el análisis de los datos. Objetivo: Identificar y seleccionar herramientas de software y métodos de análisis que se utilizarán en el desarrollo del proyecto.
Sprint 4	Semanas 5 a 7	Tarea 2.1: Grabación de nuevos datos en el Hospital Materno Infantil de Gran Canaria y en el CEIP Fernando de Guanarteme. Tarea 3.1: Implementación de algoritmos y extracción de características. Objetivo: Desarrollar diferentes algoritmos y códigos con los que poder obtener diferentes métricas de velocidad y trayectoria de la escritura.
Sprint 5	Semanas 8 a 10	Continuación de Tarea 3.1: Refinamiento de los algoritmos que obtuvieron métricas destacables. Tarea 2.2: Prototipado y refinamiento del diseño de la interfaz. Tarea 2.3: Desarrollo de una aplicación en Python para la representación de resultados y visualización de datos. Objetivo: Completar el prototipo de la aplicación en Python para procesar y visualizar los datos de manera efectiva.
Sprint 6	Semanas 11 a 17	Continuación de Tarea 3.2: Mejora continua de los algoritmos de análisis y optimización de visualización de los datos. Objetivo: Optimizar los algoritmos de análisis para mejorar la precisión y claridad en la representación gráfica de los datos..
Sprint 8	Semanas 18 a 20	Tarea 4.1: Redacción y revisión inicial de la memoria del proyecto. Objetivo: Documentar el trabajo realizado en el desarrollo de la aplicación y los resultados obtenidos en el análisis.

*Continúa en la siguiente página*

<b>Sprint</b>	<b>Duración (semanas)</b>	<b>Tareas y Objetivo</b>
Sprint 9	Semanas 21 a 24	Tarea 4.1: Redacción final y revisión exhaustiva de la memoria. Tarea 4.2: Preparación de la presentación para la defensa del proyecto. Objetivo: Finalizar la documentación del proyecto y preparar la presentación para la defensa, asegurando una explicación clara de los resultados y metodologías utilizadas.

Tabla 4.3: Planificación de los sprints en el Proyecto

## 4.2. Recopilación y análisis de datos

### 4.2.1. Base de datos

El desarrollo de este proyecto se basó en una base de datos preexistente que se complementó con muestras adicionales durante el transcurso de este trabajo. La base de datos está compuesta por archivos ‘.csv’ que contienen información detallada sobre los dibujos realizados por los pacientes y sus respectivas gráficas de velocidad. A cada dibujo le corresponde un archivo. Estas muestras fueron organizadas y estructuradas para facilitar el uso de los algoritmos desarrollados durante el proyecto.

Cada archivo ‘.csv’ registra la trayectoria de los dibujos mediante las siguientes columnas:

<b>Columna 1</b>	<b>Columna 2</b>	<b>Columna 3</b>
Posición en eje X	Posición en eje Y	Tiempo asociado a cada muestra (en segundos)

Tabla 4.4: Organización interna de cada archivo ‘.csv’ utilizado en el proyecto.

La organización de la base de datos se diseñó según el diagnóstico y el sexo de los pacientes, con el número de participantes registrados mostrado en la tabla 4.5.

<b>Diagnóstico</b>	<b>Sexo</b>	<b>Número de pacientes</b>
Neurotípico	Masculino	114
	Femenino	98
TEA	Masculino	23
	Femenino	12
TDAH	Masculino	10
	Femenino	6
<b>Total</b>		<b>263</b>

Tabla 4.5: Distribución de pacientes por diagnóstico y sexo

Estos datos representan la trayectoria de los dibujos en coordenadas (X, Y) y el momento temporal de cada muestra capturada, las cuales fueron obtenidas utilizando un iPad.

Durante el desarrollo de este trabajo, se utilizaron dos versiones de la base de datos mencionada previamente en la tabla 4.5. La primera versión incluye todos los dibujos realizados por los pacientes, mientras que la segunda versión contiene únicamente aquellos dibujos que han sido identificados y catalogados como cuadrado (ver tabla 4.6). El código utilizado para catalogar un dibujo como cuadrado se presenta en la *sección códigos desarrollados*. (ver el algoritmo 4.4)

Base de datos	Número de dibujos
V3	2435
V4	1461

Tabla 4.6: Número de dibujos en cada base de datos

Ambas versiones mantienen una estructura jerárquica, descrita en la Figura 4.3. Esta organización facilitó el procesamiento eficiente de los datos.

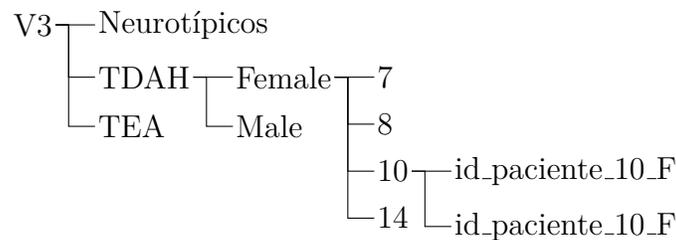


Figura 4.3: Estructura de ambas bases de datos

Siendo:

- **V3 o V4:** El nombre que se le ha dado a la base de datos.
- **Neurotípicos, TDAH y TEA:** El diagnóstico del paciente.
- **Female y Male:** El sexo del paciente.
- **7,8,10,14, etc:** La edad del paciente
- **id\_paciente\_10\_F:** ID del paciente concatenado de la edad y de la inicial del sexo (Female o Male)

#### 4.2.2. Proceso de recopilación de datos

Para la recopilación de datos de pacientes con diferentes diagnósticos realicé una visita al Hospital Materno Infantil de Gran Canaria. Antes de llevar a cabo las pruebas, se requirió que los padres de los menores firmaran una autorización que detallaba el propósito de la

recolección de dichos datos, cumpliendo así con los requisitos del comité ético. Además, se les explicó en qué consistían las pruebas que sus hijos iban a realizar.

El objetivo de la actividad consistió en observar y registrar meticulosamente los datos mientras los pacientes llevaban a cabo actividades de dibujo predefinidas en un iPad. El protocolo de recolección de datos incluyó una serie de actividades específicas de dibujo, durante las cuales se registraron tanto los trazos realizados como los datos relacionados con la velocidad y la trayectoria. La actividad en la que se centra este estudio consistió en la repetición de la tarea de dibujar un cuadrado aproximadamente en diez ocasiones consecutivas. Cada repetición generó un archivo ‘.csv’ que contenía los datos de velocidad y trayectoria del respectivo dibujo, siguiendo la estructura que se indicó anteriormente en la tabla 4.4. Lamentablemente, estas pruebas solo se pudieron llevar a cabo con una niña diagnosticada con TEA, debido a limitaciones en la recopilación de datos del segundo participante, un varón diagnosticado con TDAH.

Por otro lado, también visité el colegio Fernando Guanarteme, colaborando en la recogida de datos de un grupo de cinco alumnos neurotípicos. En esta ocasión, se logró realizar las pruebas y recoger datos de este grupo de alumnos, quienes mostraron una gran curiosidad por las pruebas. Estaban interesados en conocer el objetivo de las actividades y participaron activamente en el proceso.

Estas experiencias no solo me proporcionaron una valiosa visión del proceso de recopilación de datos en un entorno clínico, sino que también me permitió comprender mejor la estructura de los dibujos utilizados y el proceso de recopilación de los datos.

### 4.2.3. Conceptos y Herramientas de Análisis

Para analizar las diferencias en los patrones de escritura entre los grupos estudiados, se han empleado herramientas estadísticas y técnicas computacionales avanzadas. Estas permiten evaluar la significancia de las diferencias observadas y describir con precisión las trayectorias de escritura de los participantes. A continuación, se presentan los conceptos clave y su aplicación en este trabajo, mostrando cómo contribuyen a alcanzar los objetivos establecidos.

#### 4.2.3.1. Conceptos Estadísticos: p-value y F-statistic

El p-value [23] y el F-statistic [32] son herramientas fundamentales en este trabajo para evaluar la significancia estadística de las diferencias observadas en los datos obtenidos.

##### **P-value**

El p-value es una medida empleada en las pruebas de hipótesis para evaluar la importancia estadística de los resultados obtenidos. Indica la probabilidad de obtener un resultado tan extremo como, o más extremo que, el observado en el experimento, bajo la premisa de que la hipótesis nula es verdadera. Un p-value bajo (típicamente menor a 0.05) indica que las muestras son significativamente diferentes, y por lo tanto, se puede rechazar la hipótesis nula

a favor de la hipótesis alternativa. En este trabajo, el p-value se ha empleado en el análisis de varianza (ANOVA) para evaluar la significancia estadística de los datos.

### F-statistic

El F-statistic es una medida utilizada en el análisis de varianza (ANOVA) para comparar las variaciones entre los grupos de datos con la variación dentro de los grupos. Específicamente, el F-statistic es la razón entre la variabilidad explicada por el modelo y la variabilidad no explicada. Un valor alto del F-statistic indica que hay una diferencia significativa entre las medias de los grupos comparados. El p-value asociado al F-statistic se utiliza para determinar si estas diferencias son estadísticamente significativas.

#### 4.2.3.2. Comparación de Series Temporales: Distancia DTW

La Distancia DTW (Dynamic Time Warping) [33] es una medida utilizada para evaluar la similitud entre dos series temporales que pueden tener diferencias en duración o velocidad. A diferencia de otras métricas de distancia, la distancia DTW permite alinear no linealmente las series para identificar patrones comunes, aunque ocurran en momentos distintos.

El proceso de DTW consiste en “estirar” o “comprimir” partes de las series para hacer que puntos similares queden alineados. La distancia DTW es un número que indica qué tan parecidas son las dos series: una distancia DTW baja indica una alta similitud entre las series temporales analizadas, mientras que una distancia alta sugiere diferencias significativas en los patrones de velocidad.

En la siguiente imagen 4.4 las dos series (en rojo y azul) se alinean mediante el algoritmo DTW. Las líneas verticales muestran cómo se emparejan puntos similares entre las series, incluso si están en posiciones de tiempo diferentes.

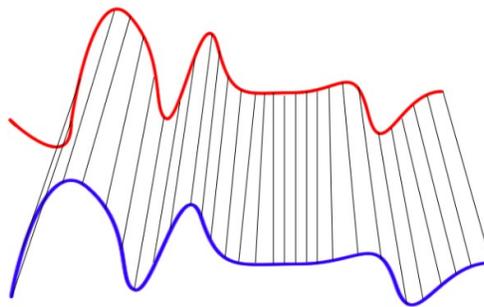


Figura 4.4: Ejemplo de alineación de series temporales mediante DTW.

1

La distancia DTW resulta especialmente útil en este trabajo para comparar los perfiles de velocidad de los trazos de escritura, permitiendo identificar si existen patrones distintivos en los movimientos entre diferentes grupos de pacientes. Una distancia DTW baja indica una

<sup>1</sup>La imagen de ejemplo fue obtenida de ResearchGate (<https://www.researchgate.net>).

alta similitud entre las series temporales analizadas, mientras que una distancia alta sugiere diferencias significativas en los patrones de velocidad.

Matemáticamente se define como:

$$DTW(A, B) = \min \sum_{i=1}^K |A(i) - B(j)| \quad (4.1)$$

Donde  $A$  y  $B$  son las series temporales a comparar, y  $K$  representa el número de puntos en el camino óptimo que minimiza la diferencia acumulada entre las series.

En este trabajo, la distancia DTW se utiliza para comparar los perfiles de velocidad de los trazos de escritura. Esto nos ayuda a ver si hay patrones de movimiento similares o distintos entre ellos mismos, y los diferentes grupos de pacientes.

#### 4.2.3.3. Comparación de Formas: Diferencia de Hausdorff

La diferencia de Hausdorff mide la similitud entre dos conjuntos de puntos, capturando la máxima discrepancia en sus posiciones. A diferencia de otras métricas, esta distancia considera la mayor de las distancias mínimas desde cada punto de un conjunto hasta el otro, reflejando la diferencia más significativa en su estructura espacial.

En la figura 4.5, los conjuntos  $A$  (azul) y  $B$  (rojo) se comparan, y la distancia de Hausdorff, marcada con una línea punteada verde, representa la distancia máxima mínima entre los conjuntos.

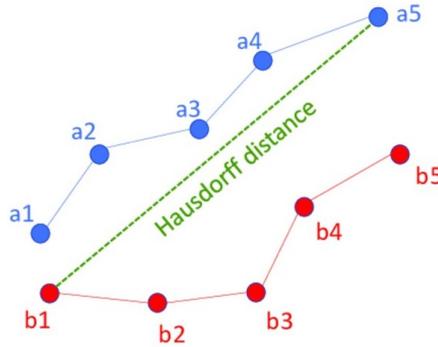


Figura 4.5: Ejemplo de comparación de conjuntos mediante la Diferencia de Hausdorff

2

Matemáticamente se define como:

$$d_H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|b - a\| \right\} \quad (4.2)$$

<sup>2</sup>La imagen de ejemplo fue obtenida de ResearchGate (<https://www.researchgate.net>).

En este trabajo, se emplea la diferencia de Hausdorff para evaluar si los patrones de movimiento son parecidos o distintos entre los mismos grupos de pacientes y entre diferentes grupos de pacientes.

#### 4.2.3.4. Detección de formas geométricas

La función `cv2.approxPolyDP()` de OpenCV es una herramienta fundamental para la simplificación de contornos detectados en una imagen, facilitando así la identificación de formas geométricas específicas. Esta función se utiliza para aproximar un contorno a una forma poligonal menos compleja, lo que resulta útil para detectar y analizar formas como cuadrados y triángulos.

##### Parámetros de `cv2.approxPolyDP()`:

- **Contorno** (`contour`): Es una secuencia de puntos que representa el contorno detectado en una imagen. Este contorno es generalmente obtenido mediante una operación de detección de bordes, como `cv2.findContours()`.
- **Epsilon**: Este parámetro controla la precisión del polígono aproximado. Se calcula como una fracción de la distancia total del contorno. Un valor menor de `epsilon` resulta en una aproximación más precisa, mientras que un valor mayor simplifica el contorno, reduciendo el número de vértices del polígono.
- **Cerrado** (`closed`): Un valor booleano que indica si el polígono debe ser cerrado. Se establece en `True` para asegurar que el polígono sea cerrado.

##### Funcionamiento:

1. **Detección del Contorno**: Primero, se detectan los contornos en la imagen utilizando `cv2.findContours()`, generando una lista de contornos que representan las formas presentes en la imagen.
2. **Aproximación del Contorno**: Se aplica `cv2.approxPolyDP()` a cada contorno detectado para simplificarlo a un polígono con menos vértices.
3. **Identificación de Formas**: Al analizar el número de vértices del polígono simplificado, se pueden identificar formas geométricas específicas, como cuadrados.

```

1 # Deteccion de bordes
2 bordes = cv2.Canny(gris , 50, 150)
3
4 # Encontrar contornos
5 contornos , _ = cv2.findContours(bordes , cv2.RETR_EXTERNAL, cv2.
    CHAIN_APPROX_SIMPLE)
6
7 for contorno in contornos:
8     # Aproximar el contorno
9     epsilon = 0.02 * cv2.arcLength(contorno , True)
10    aproximacion = cv2.approxPolyDP(contorno , epsilon , True)
11
12    # Verificar si el contorno es un cuadrado (4 vertices)
13    if len(aproximacion) == 4:
14        cv2.drawContours(imagen , [aproximacion] , -1, (0 , 255, 0) , 3)

```

Algoritmo 4.1: Ejemplo de `cv2.approxPolyDP()` para la detección de cuadrados

#### 4.2.3.5. Integración de conceptos en el análisis

Los conceptos y herramientas descritos se integraron para realizar un análisis detallado de las trayectorias de escritura:

1. **Validación estadística:** Las diferencias observadas en los resultados de DTW se contrastaron con pruebas ANOVA para determinar su significancia estadística mediante p-value y F-statistic.
2. **Detección de formas geométricas:** La identificación de formas geométricas se realizó utilizando la función `cv2.approxPolyDP()` de OpenCV.
3. **Comparación temporal:** Los perfiles de velocidad de escritura se analizaron utilizando DTW para alinear las variaciones entre series temporales de diferentes pacientes.
4. **Análisis espacial:** Las trayectorias espaciales de los cuadrados se evaluaron con la diferencia de Hausdorff, identificando patrones espaciales comunes o divergentes entre los grupos.

Estas herramientas y conceptos están integrados en los algoritmos desarrollados en este trabajo, como se detallará en la siguiente subsección. Los resultados obtenidos contribuyen directamente a discriminar patrones de escritura entre los diferentes grupos estudiados, facilitando la identificación de diferencias clave para el diagnóstico diferencial.

### 4.3. Códigos y algoritmos desarrollados

Antes de comenzar con esta sección donde se explicarán los códigos y algoritmos desarrollados para evaluar y analizar análisis el perfil de velocidad y trayectoria de la escritura de

los pacientes, se pretende dejar claro que la separación de los datos entre sexos en el análisis no fue una decisión arbitraria, sino que responde a diferencias clínicamente significativas en cómo se manifiestan los síntomas de TEA y TDAH en cada género. Tal y como se mencionó en anteriormente, estudios previos han mostrado que las niñas tienden a presentar síntomas de manera menos evidente, en muchos casos, más internalizados, lo que suele conllevar a diagnósticos tardíos o incluso a la falta de diagnóstico. En el caso del TEA, el fenómeno de ‘camuflaje’ o ‘masking’, del que se habló anteriormente en ‘Estado del arte’, en el cual las niñas aprenden a ocultar o minimizar sus síntomas para adaptarse socialmente, agrega un reto adicional a la detección temprana y precisa.

Teniendo en cuenta estas particularidades, los códigos y algoritmos se diseñaron para analizar los datos de manera diferenciada entre pacientes masculinos y femeninas. Esta separación no solo permite que la herramienta detecte patrones específicos en cada grupo, sino que también contribuye a un diagnóstico más preciso e inclusivo, asegurando que las características únicas de las niñas no queden diluidas en un análisis conjunto.

En el momento de la entrega de este trabajo, los códigos desarrollados generan las figuras calculadas basándose en los datos disponibles hasta la fecha o, alternativamente, considerando nuevos sujetos siempre que se incorporen siguiendo la estructura de ficheros descrita en la figura 4.3.

Para ejecutar los códigos y obtener las figuras que se muestran a continuación, el sistema de archivos debe estar organizado de la siguiente manera:

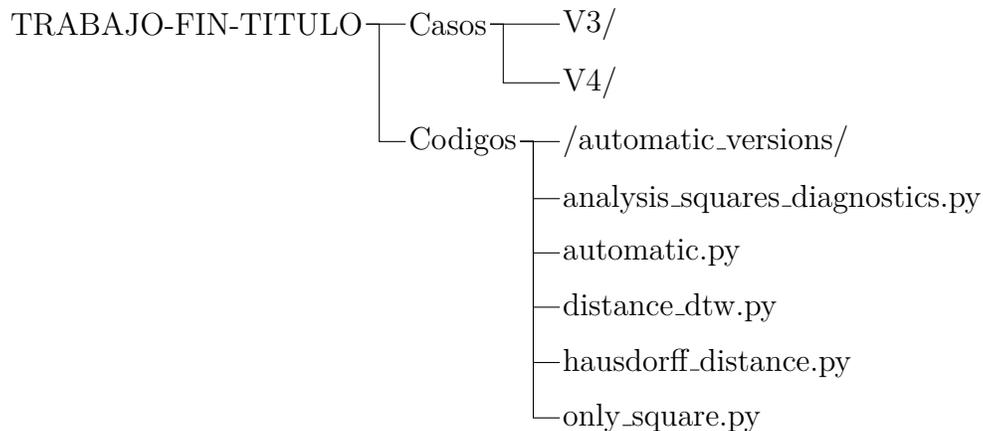


Figura 4.6: Estructura del sistema de ficheros

### 4.3.1. Detección de cuadrados

El código que se encuentra en el fichero **automatic.py**, contiene un código desarrollado para realizar un análisis automatizado de dibujos representados en los archivos CSV de la base de datos. Este análisis tiene como objetivo principal la detección de cuadrados a partir de los dibujos realizados en la tablet que se usó para la recopilación de los datos.

El archivo define dos funciones principales. La función `extraer_nombre(archivo)` se

encarga de extraer el nombre deseado del archivo CSV utilizando expresiones regulares. Esto permite identificar y extraer la parte relevante del nombre del archivo para su posterior análisis. La función `procesar_archivos(directorio)` procesa los archivos CSV dentro de un directorio dado. Esta función realiza la lectura de datos, normalización y suavizado de coordenadas, y busca la presencia de cuadrados en los dibujos.

Para asegurar que todos los dibujos se representen con el mismo tamaño en la visualización, el código realiza un reescalado de las coordenadas de cada dibujo. Este proceso implica una normalización de las coordenadas  $x$  e  $y$  para ajustar los valores originales a un rango uniforme de 0 a 255, de modo que cada dibujo ocupe el mismo tamaño visual dentro de una imagen de  $256 \times 256$  píxeles. Así, independientemente de las dimensiones originales de los dibujos, estos se escalan para ocupar el mismo espacio en el lienzo visual, permitiendo una comparación homogénea entre todos.

Este ajuste de escala se realiza mediante las siguientes transformaciones:

```

1 x_norm = ((x_data - x_min) / (x_max - x_min) * 255).astype(np.uint8)
2 y_norm = ((y_data - y_min) / (y_max - y_min) * 255).astype(np.uint8)

```

Algoritmo 4.2: Reescalado de dibujos realizados

Donde,  $x\_data$  y  $y\_data$  representan las coordenadas originales del dibujo, mientras que  $x\_min$ ,  $x\_max$ ,  $y\_min$  y  $y\_max$  son los valores mínimo y máximo en cada eje. Al aplicar esta normalización, se garantiza que el dibujo se ajuste a un rango uniforme, permitiendo una comparación visual directa entre todos los dibujos en los gráficos de las figuras 4.7 y 4.8.

El proceso de detección de cuadrados se realiza mediante un enfoque basado en procesamiento de imágenes utilizando la biblioteca OpenCV en Python. Primero, se genera una máscara de contorno en blanco y negro que representa los bordes del dibujo. Luego, se buscan los contornos en esta máscara y se aproximan a polígonos con la función `cv2.approxPolyDP()`, buscando aquellos con cuatro vértices. Estos cuadrados detectados se registran y se visualizan junto con los dibujos originales [15].

A continuación, se presenta el fragmento de código responsable de la detección de cuadrados:

```

1 # Encontrar contornos
2 contours, _ = cv2.findContours(canvas_blur, cv2.RETR_EXTERNAL, cv2.CHAIN_APPROX_SIMPLE)
3 for contour in contours:
4     approx = cv2.approxPolyDP(contour, 0.015 * cv2.arcLength(contour, True), True) # Se
        ajusta el umbral
5     if len(approx) == 4 and cv2.isContourConvex(approx): # Se verifica la convexidad
6         area = cv2.contourArea(approx)
7         if area > 1000: # Se ajusta el area minima del cuadrado
8             num_cuadrados_encontrados += 1
9             archivos_cuadrados.append(archivo_csv)
10            cv2.drawContours(canvas, [approx], -1, (127), 2) # Dibujar contorno detectado

```

Algoritmo 4.3: Detección de cuadrados en imágenes

Cabe destacar que con el objetivo de mejorar la detección de cuadrados, de forma empírica, se optó por reducir el umbral a 0.015 y ajustar el área mínima del cuadrado a 1000 unidades cuadradas.

En las figuras 4.7 y 4.8, se pueden observar dos ejemplos del resultado del procesamiento realizado por `automatic.py` para un varón y una niña, ambos de 10 años y diagnosticados con TEA.

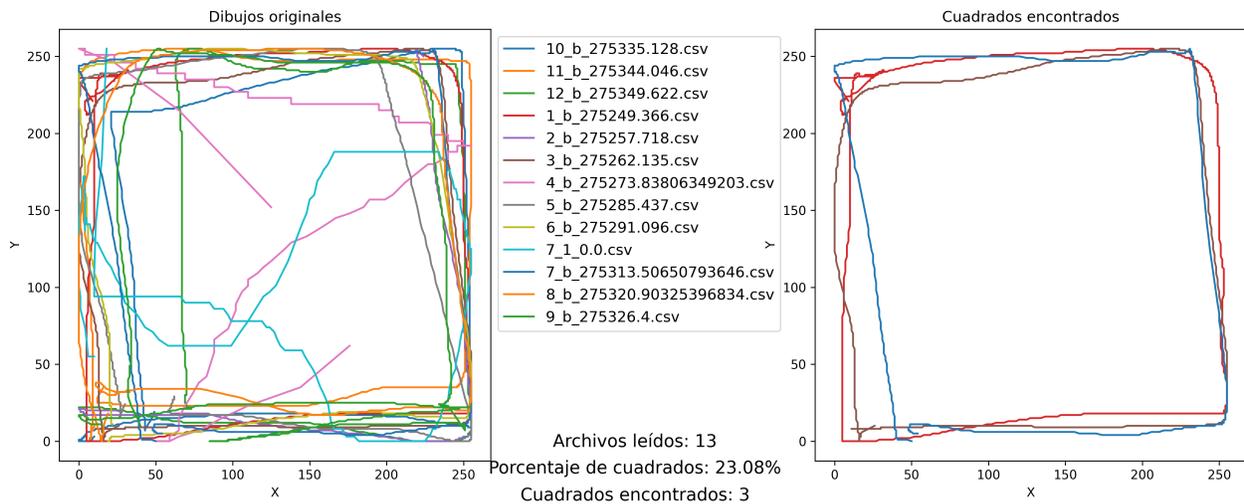


Figura 4.7: Resultado del algoritmo de detección de cuadrados para un varón de 10 años diagnosticado con TEA

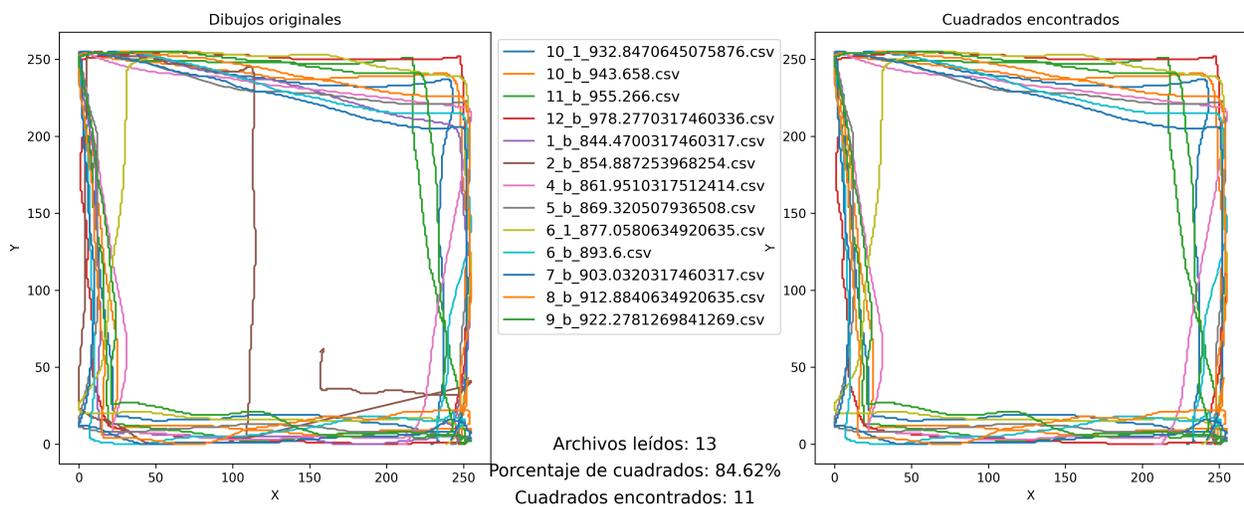


Figura 4.8: Resultado del algoritmo de detección de cuadrados para una niña de 10 años diagnosticada con TEA

Como se indicó anteriormente, se empleó una versión de este código para crear la segunda versión de la base de datos utilizada, eliminando aquellos archivos que guardaban la información de dibujos no catalogados como cuadrados. Dentro del sistema de ficheros, esta versión modificada puede verse en el archivo **only\_square.py**. A continuación, en el algoritmo 4.4, se presenta la parte más significativa del mismo, donde si el dibujo realizado no es detectado como cuadrado, se elimina el fichero que contiene su información.

```

1 # Encontrar contornos
2 contours, _ = cv2.findContours(canvas_blur, cv2.RETR_EXTERNAL, cv2.
   CHAIN_APPROX_SIMPLE)
3 cuadrado_encontrado = False
4 for contour in contours:
5     approx = cv2.approxPolyDP(contour, 0.02 * cv2.arcLength(contour, True)
6     , True) # Se ajusta el umbral
7     if len(approx) == 4 and cv2.isContourConvex(approx): # Se verifica la
8         convexidad
9         area = cv2.contourArea(approx)
10        if area > 1000: # Ajustar el area cuadrada minima
11            cuadrado_encontrado = True
12            break
13
14 if not cuadrado_encontrado: # Si no se encontro un cuadrado, eliminar
15     el archivo
16     os.remove(archivo_csv)

```

Algoritmo 4.4: Detección de cuadrados y eliminación de los no cuadrados

### 4.3.2. Cálculo y representación de porcentaje de cuadrados

El código que se encuentra en el fichero **automatic\_v2.py** extiende las funcionalidades de **automatic.py**, añadiendo la capacidad de calcular el porcentaje de cuadrados detectados y representar su distribución a través de diagramas de cajas, donde cada caja representa la variabilidad de los datos para una edad específica.

La figura 4.9, muestra el diagrama de cajas correspondiente a los porcentajes de cuadrados detectados en pacientes masculinos neurotípicos. En este gráfico, se muestra una tendencia ascendente en el porcentaje de cuadrados encontrados a medida que la edad aumenta, particularmente entre las edades de 4 a 7 años. Esto sugiere un desarrollo progresivo de las habilidades relevantes a medida que los pacientes masculinos neurotípicos crecen. Después de los 7 años, el porcentaje de cuadrados encontrados parece estabilizarse, aunque aún se observa variabilidad.

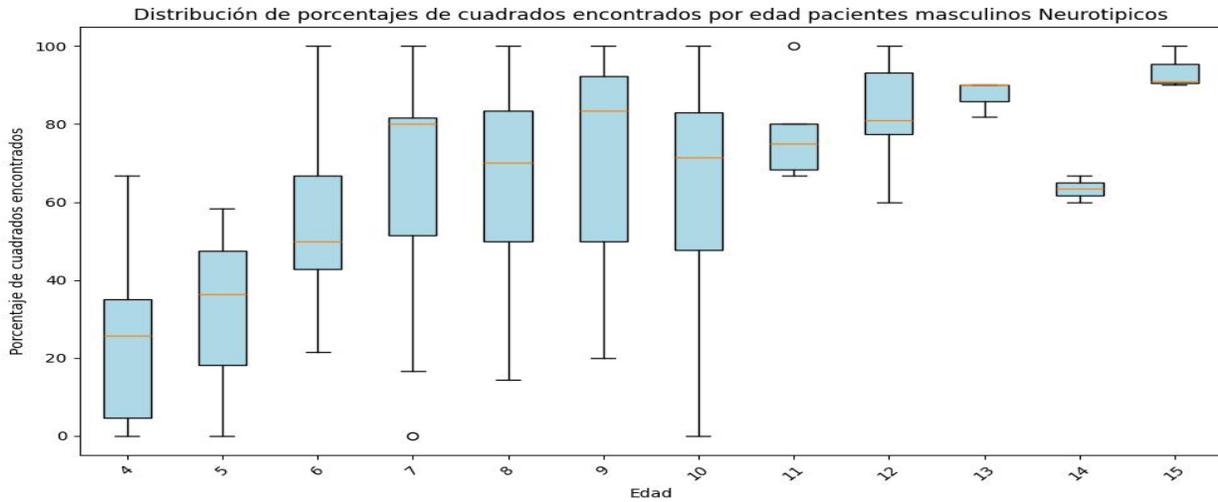


Figura 4.9: Porcentajes de cuadrados encontrados en pacientes masculinos neurotípicos

Con el objetivo de analizar si las diferencias observadas en la figura anterior son significativas, se realizó un análisis estadístico sobre los porcentajes de cuadrados detectados. Este análisis permite identificar patrones de desarrollo relacionados con la edad y comparar los distintos grupos para determinar si las diferencias son estadísticamente significativas. A continuación, se presentan los resultados obtenidos a partir de pruebas t de Student (tabla 4.7) y análisis de varianza (ANOVA) (tabla 4.8).

Comparación de Grupos de Edad	T-statistic	P-value
4-6 vs 7-10	-2.9005	0.0078
7-10 vs 11-13	-3.5069	0.0008
4-6 vs 11-13	-4.9278	6.0152e-05

Tabla 4.7: Resultados de la prueba t de Student para pacientes masculinos neurotípicos

ANOVA	F-statistic	P-value
Grupos de 4-6, 7-10, 11-13	9.6077	0.0001

Tabla 4.8: Resultados de la prueba ANOVA para pacientes masculinos neurotípicos

Los análisis estadísticos realizados indican diferencias significativas en el porcentaje de cuadrados encontrados entre los distintos grupos de edad, lo que sugiere una evolución en las habilidades cognitivas y perceptuales a medida que los niños crecen. Específicamente, se observaron diferencias significativas entre los grupos de edad 4-6 años y 7-10 años ( $p = 0,0078$ ), así como entre los grupos 7-10 años y 11-13 años ( $p = 0,0008$ ). Estas diferencias son aún más pronunciadas al comparar los grupos de edad 4-6 años y 11-13 años, con un p-value extremadamente pequeño ( $p = 6,0152 \times 10^{-5}$ ), lo que proporciona una fuerte evidencia de que las habilidades evaluadas mejoran significativamente con la edad.

Estos hallazgos subrayan la importancia de considerar el desarrollo relacionado con la edad. La mejora significativa en las habilidades evaluadas con el crecimiento sugiere que los programas optimizados para diferentes grupos de edad pueden apoyar de manera más efectiva el desarrollo cognitivo y perceptual de los niños.

En este contexto, resulta relevante explorar cómo estas distribuciones varían en condiciones específicas, como TDAH y TEA, incluso sin realizar un análisis estadístico detallado. Las siguientes figuras (figura 4.10 y figura 4.11) ilustran las distribuciones de cuadrados encontrados para pacientes masculinos con TDAH y TEA, respectivamente.

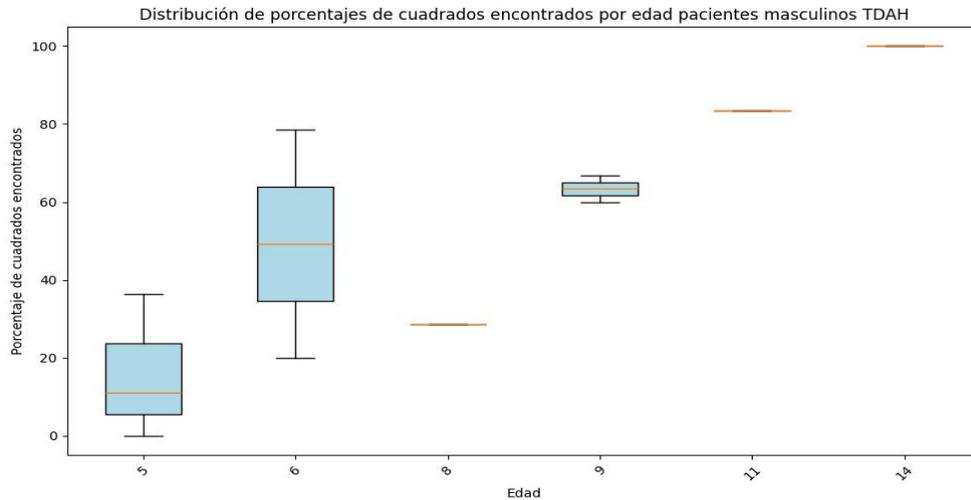


Figura 4.10: Porcentajes de cuadrados encontrados en pacientes masculinos con TDAH

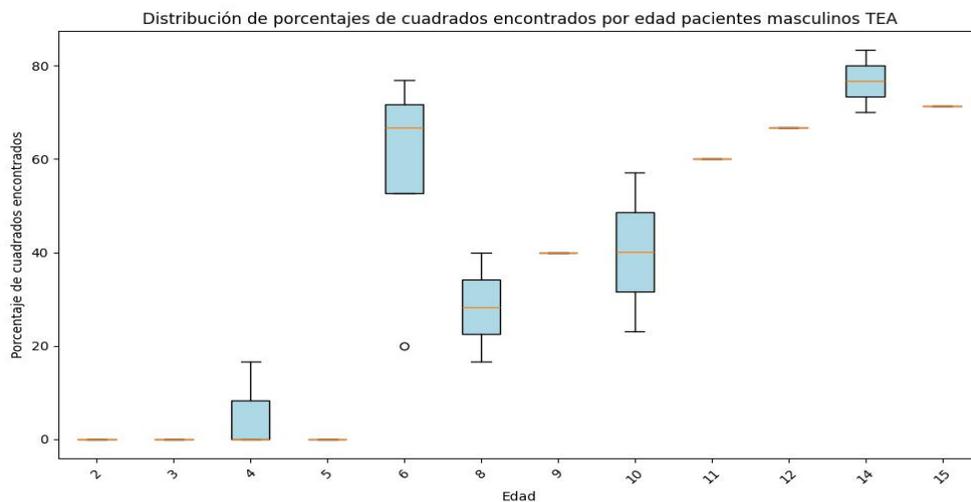


Figura 4.11: Porcentajes de cuadrados encontrados en pacientes masculinos con TEA

La figura 4.12 muestra el diagrama de cajas correspondiente a los porcentajes de cuadrados detectados en pacientes neurotípicas, evidenciando una tendencia ascendente en el número de cuadrados encontrados a medida que la edad aumenta, particularmente entre las edades de 4 a 7 años, lo que sugiere un desarrollo progresivo de las habilidades relevantes a medida que las niñas crecen. Después de los 7 años, el porcentaje de cuadrados encontrados parece estabilizarse, aunque aún se observa cierta variabilidad.

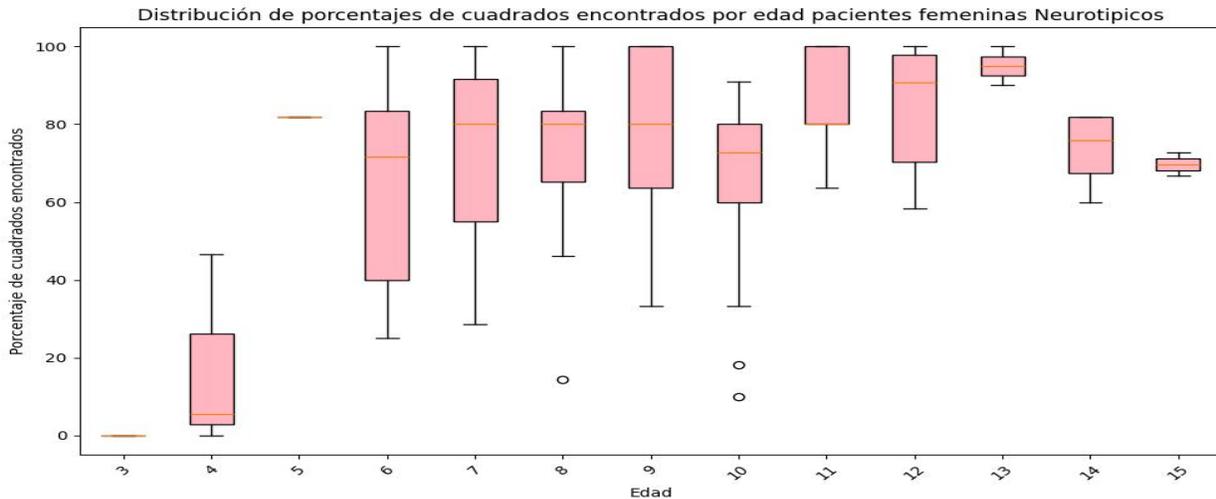


Figura 4.12: Porcentajes de cuadrados encontrados en pacientes femeninas neurotípicas

De la misma manera que se hizo para los pacientes masculinos neurotípicos, también se realizó un análisis estadístico para las pacientes femeninas neurotípicas con el fin de analizar si las diferencias observadas en la figura 4.12 son significativas. A continuación, se presentan los resultados obtenidos a partir de pruebas t de Student (tabla 4.9) y análisis de varianza (ANOVA) (tabla 4.10).

Comparación de Grupos de Edad	T-statistic	P-value
4 vs 6-7	-1.8516	0.1479
4 vs 8-10	-1.9861	0.1366
4 vs 11-14	-2.5346	0.0793
6-7 vs 8-10	-0.2659	0.7919
8-10 vs 11-14	-2.2703	0.0286
6-7 vs 11-14	-1.9128	0.0640

Tabla 4.9: Resultados de la prueba t para pacientes femeninas neurotípicas

ANOVA	F-statistic	P-value
Grupos de 4, 6-7, 8-10, 11-14	5.2280	0.0023

Tabla 4.10: Resultados de la prueba ANOVA para pacientes femeninas neurotípicas

Los análisis estadísticos realizados indican diferencias significativas en el porcentaje de cuadrados encontrados entre algunos grupos de edad en pacientes femeninas neurotípicas. En particular, se observó una diferencia significativa entre los grupos de edad 8-10 años y 11-14 años ( $p = 0,0286$ ), lo que sugiere que, aunque las habilidades relacionadas con la detección de cuadrados tienden a estabilizarse después de los 7 años, continúan mejorando ligeramente en edades mayores.

El análisis ANOVA también muestra diferencias significativas entre los grupos de edad ( $p = 0,0023$ ), indicando que al menos uno de los grupos difiere significativamente de los demás. Este hallazgo refuerza la idea de que la edad influye en el desarrollo de estas habilidades.

Estos resultados subrayan la importancia de considerar el desarrollo relacionado con la edad al diseñar programas dirigidos a mejorar habilidades cognitivas y perceptuales específicas, como las asociadas con el reconocimiento de formas geométricas.

A continuación, se presentan las distribuciones de cuadrados encontrados en pacientes con TDAH y TEA, ilustradas mediante diagramas de cajas. Estas figuras (figura 4.13 y figura 4.14) ofrecen una representación visual que permite observar posibles patrones o tendencias en la detección de cuadrados para pacientes masculinos con TDAH y TEA, respectivamente:

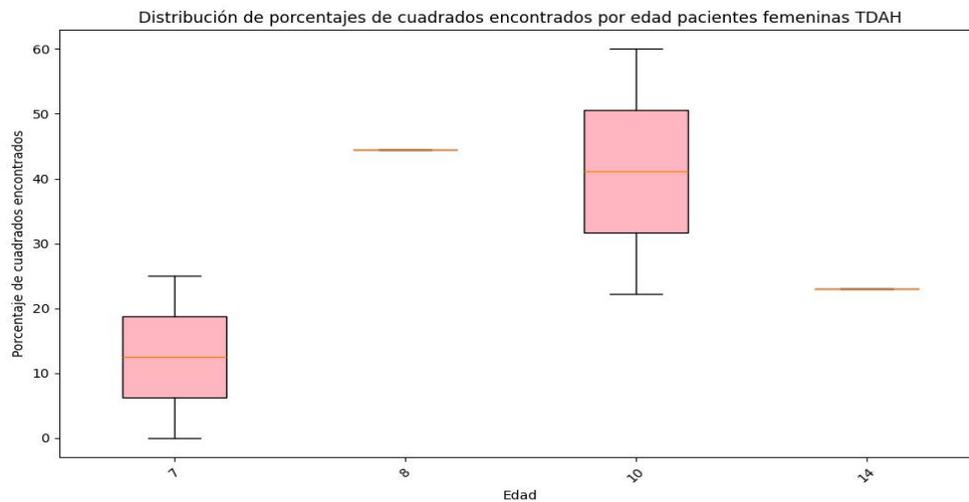


Figura 4.13: Porcentajes de cuadrados encontrados en pacientes femeninas con TDAH

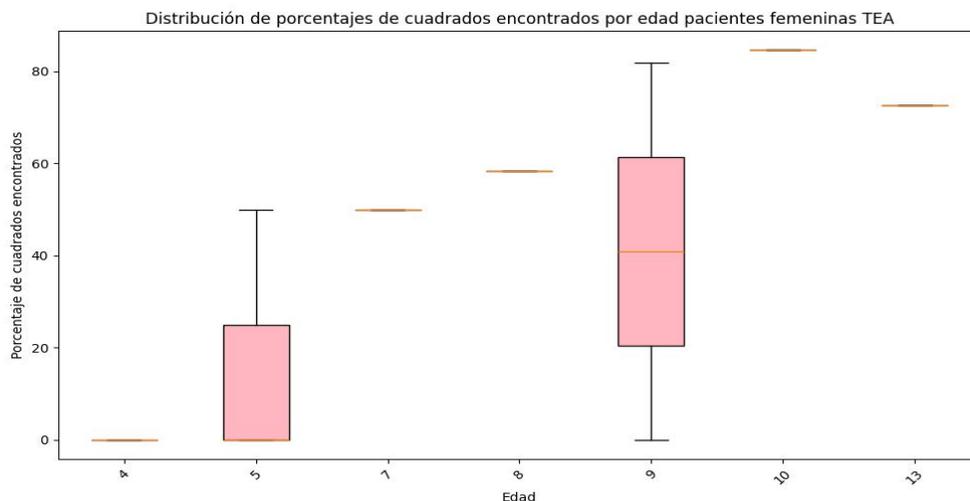


Figura 4.14: Porcentajes de cuadrados encontrados en pacientes femeninas con TEA

A continuación, se presenta en una misma figura la representación de los tres diagnósticos con los que trabaja este proyecto con el objetivo de facilitar la comparación de los valores obtenidos. Esta figura se obtiene ejecutando el código que se encuentra en el archivo **analysis\_squares\_diagnostics.py**, ejecutando los directorios del sexo que se desea graficar.

A continuación, en la figura 4.15, se presenta el diagrama de cajas correspondientes a la distribución de porcentajes de cuadrados encontrados para los sujetos masculinos con los tres diagnósticos trabajados:

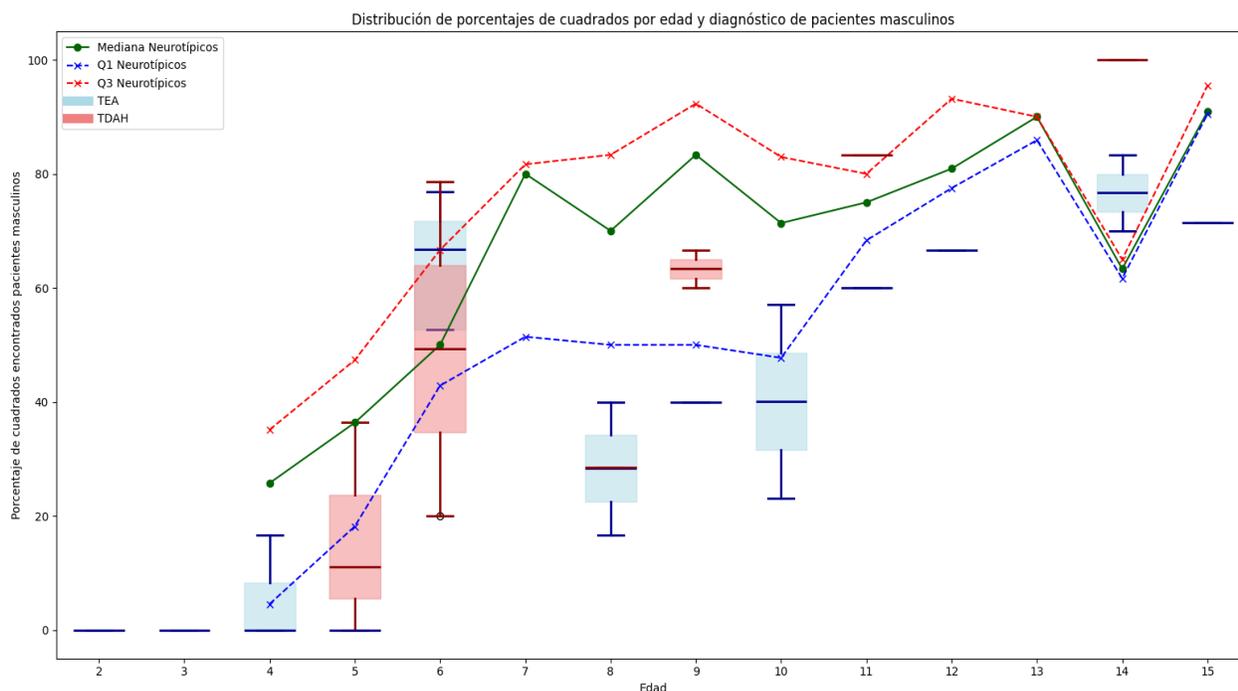


Figura 4.15: Distribución de porcentaje de cuadrados encontrados para pacientes masculinos

Además, con el objetivo de depurar el código y verificar qué datos se están leyendo al generar las figuras, se añadió un fragmento de código que imprime los diagnósticos disponibles para cada edad. La tabla 4.11 presenta un resumen de los diagnósticos registrados por edad para los pacientes masculinos:

Edad	Diagnósticos
2-3	TEA
4	Neurotípicos, TEA
5-6	Neurotípicos, TEA, TDAH
7	Neurotípicos
8-9	Neurotípicos, TEA, TDAH
10	Neurotípicos, TEA
11	Neurotípicos, TEA, TDAH
12	Neurotípicos, TEA
13	Neurotípicos
14	Neurotípicos, TEA, TDAH
15	Neurotípicos, TEA

Tabla 4.11: Diagnósticos disponibles por edad y diagnóstico de pacientes masculinos

De la misma manera que se hizo para los sujetos masculinos, a continuación, se presenta el diagrama de cajas (figura 4.16) correspondiente a la distribución de porcentajes de cuadrados encontrados para los sujetos femeninos con los tres diagnósticos trabajados:

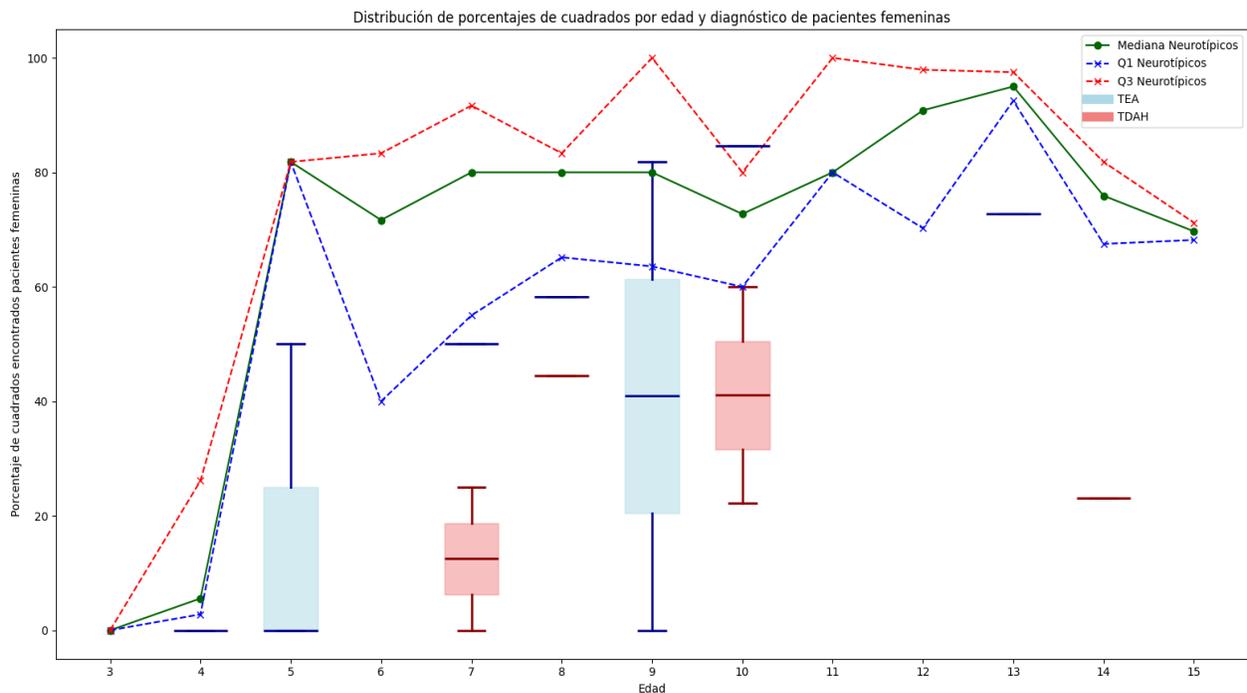


Figura 4.16: Distribución de porcentaje de cuadrados encontrados para pacientes femeninas

Para este caso, también se incluye una tabla, como la 4.12, que presenta los diagnósticos disponibles para cada edad, de manera similar a lo realizado anteriormente con los sujetos de sexo masculino:

Edad	Diagnósticos
3	Neurotípicos
4-5	Neurotípicos, TEA
6	Neurotípicos
7-8	Neurotípicos, TEA, TDAH
9	Neurotípicos, TEA
10	Neurotípicos, TEA, TDAH
11-12	Neurotípicos
13	Neurotípicos, TEA
14	Neurotípicos, TDAH
15	Neurotípicos

Tabla 4.12: Diagnósticos disponibles por edad y diagnóstico de pacientes femeninas

### 4.3.3. Cálculo de similitud del perfil de velocidad de la escritura

El código desarrollado en el archivo `distance_dtw.py` está diseñado para analizar las velocidades de escritura de pacientes en función del diagnóstico, sexo y edad para compararlas mediante el cálculo de la distancia de tiempo de deformación (*Dynamic Time Warping*, *DTW*), como se detalla en la subsección ‘Comparación de Series Temporales: Distancia DTW’.

Las figuras 4.17 y 4.18 representan los resultados obtenidos para los pacientes masculinos y femeninos neurotípicos, respectivamente:

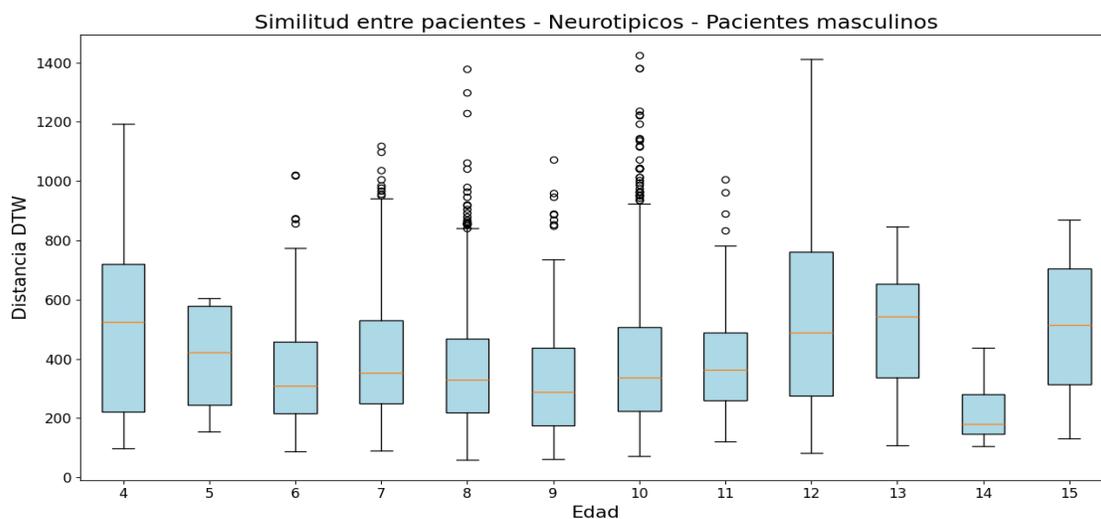


Figura 4.17: Distribución de similitud del perfil de velocidad neurotípicos

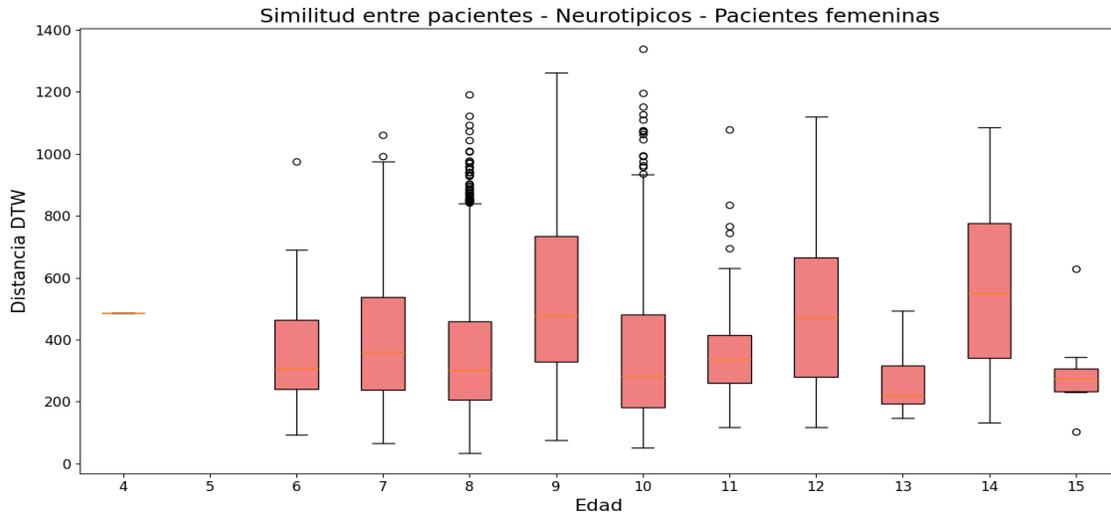


Figura 4.18: Distribución de similitud del perfil de velocidad neurotípicas

Con el objetivo de comprobar si existían diferencias significativas en los resultados obtenidos, se realizó, como en algoritmos anteriores, un análisis estadístico para los pacientes de edad 9 y 10 años de los tres diagnósticos tratados en este trabajo de fin de título. La figura 4.19 muestra la variabilidad de la distancia dtw para el caso mencionado:

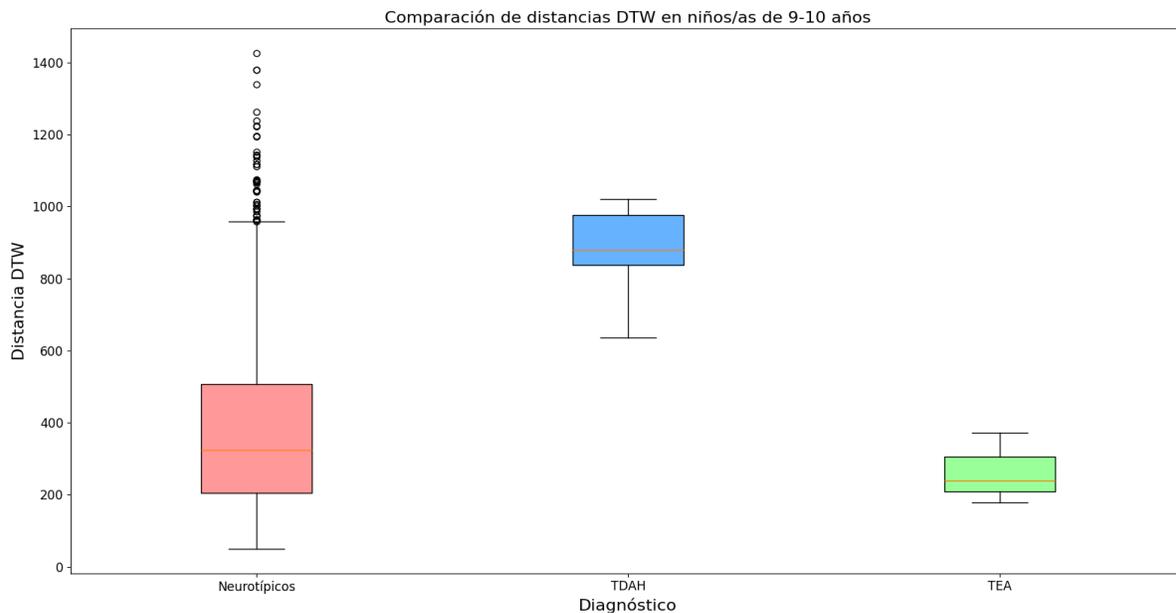


Figura 4.19: Distribución de similitud del perfil de velocidad por grupos

Las tablas 4.13 y 4.14 muestran los resultados obtenidos a partir de las pruebas t de Student y el análisis de varianza (ANOVA) respectivamente:

Comparación de diagnósticos	T-statistic	P-value
Neurotípicos vs. TDAH	-8.54570	0.00033
Neurotípicos vs. TEA	2.11943	0.16613
TDAH vs. TEA	7.59329	0.00032

Tabla 4.13: Prueba T para comparaciones entre diagnósticos (9-10 años)

ANOVA	P-value
Grupos de 9-10	9.14989e-07

Tabla 4.14: Resultado de la prueba ANOVA para los grupos de 9-10 años

En el análisis de las velocidades de escritura entre grupos de diagnóstico (Neurotípicos, TDAH y TEA) para niños de 9-10 años, se encontró una diferencia significativa entre al menos dos de los grupos mediante ANOVA (valor  $p = 9.15e-07$ ). Esto indica que las distancias DTW, que reflejan similitudes de velocidad, varían notablemente entre los diagnósticos.

Las comparaciones por pares muestran que las distancias DTW son significativamente diferentes entre los Neurotípicos y el grupo TDAH ( $p = 0.00033$ ), siendo las velocidades de los Neurotípicos más consistentes entre sí. Entre Neurotípicos y TEA, no se observan diferencias significativas ( $p = 0.16613$ ), lo que sugiere que ambos grupos presentan patrones de velocidad similares. Sin embargo, entre TDAH y TEA, vuelven a haber diferencias significativas ( $p = 0.00032$ ), con velocidades más consistentes en el grupo TEA.

#### 4.3.4. Cálculo de diferencia entre dibujos de cuadrados

El código desarrollado en el archivo `hausdorff_distance.py` calcula la diferencia de Hausdorff, una medida de similitud entre dos conjuntos de puntos, utilizada aquí para comparar las figuras cuadradas detectadas en los datos de cada paciente.

El código analiza los puntos normalizados de cada archivo `.csv`, detecta cuadrados y los reescala para estandarizar su tamaño, facilitando comparaciones consistentes, como se muestra en el algoritmo 4.5.

```

1 def reescalar_cuadrado(puntos):
2     # Calcula el centroide del cuadrado
3     centroide = np.mean(puntos, axis=0)
4     # Traslada los puntos para centrar el cuadrado en el origen
5     puntos -= centroide
6     # Calcula la distancia máxima desde el centroide a cualquier punto del
7     # cuadrado
8     max_distancia = np.max(np.linalg.norm(puntos, axis=1))
9     # Reescala los puntos dividiendo por la distancia máxima
10    puntos /= max_distancia
11    return puntos

```

Algoritmo 4.5: Reescalado de cuadrados detectados

Esta función calcula el centroide de los puntos y luego reescala el cuadrado dividiendo los puntos por la máxima distancia al centroide, asegurando que todos los cuadrados tengan un tamaño uniforme para facilitar las comparaciones.

Para cada par de figuras cuadradas detectadas, el código calcula dos distancias de Hausdorff dirigidas usando la función `directed_hausdorff(c1, c2)` de `scipy.spatial.distance`. Posteriormente, toma el valor máximo de estas distancias como la diferencia entre las figuras. Este valor representa cuán distintas son las figuras cuadradas entre sí en términos de su forma. A continuación se muestra la parte del código que realiza este cálculo:

```

1 def calcular_diferencia_entre_cuadrados(cuadrado1, cuadrado2):
2     # Calcula la distancia de Hausdorff desde cuadrado1 a cuadrado2
3     distancia1 = directed_hausdorff(cuadrado1, cuadrado2)[0]
4     # Calcula la distancia de Hausdorff desde cuadrado2 a cuadrado1
5     distancia2 = directed_hausdorff(cuadrado2, cuadrado1)[0]
6     # Retorna la mayor de las dos distancias como medida de disimilitud
7     return max(distancia1, distancia2)

```

Algoritmo 4.6: Cálculo de la diferencia de Hausdorff entre cuadrados

Esta función toma dos figuras cuadradas y calcula la diferencia de Hausdorff entre ellas, considerando ambas direcciones y seleccionando el valor máximo de las distancias para obtener una medida robusta de diferencia. Finalmente, las diferencias de Hausdorff se agrupan por sexo y edad y se representan en gráficos de cajas (boxplots) para facilitar el análisis comparativo.

Las figuras 4.20 y 4.21 se corresponden a las diferencias entre cuadrados de los pacientes masculinos y femeninos respectivamente:

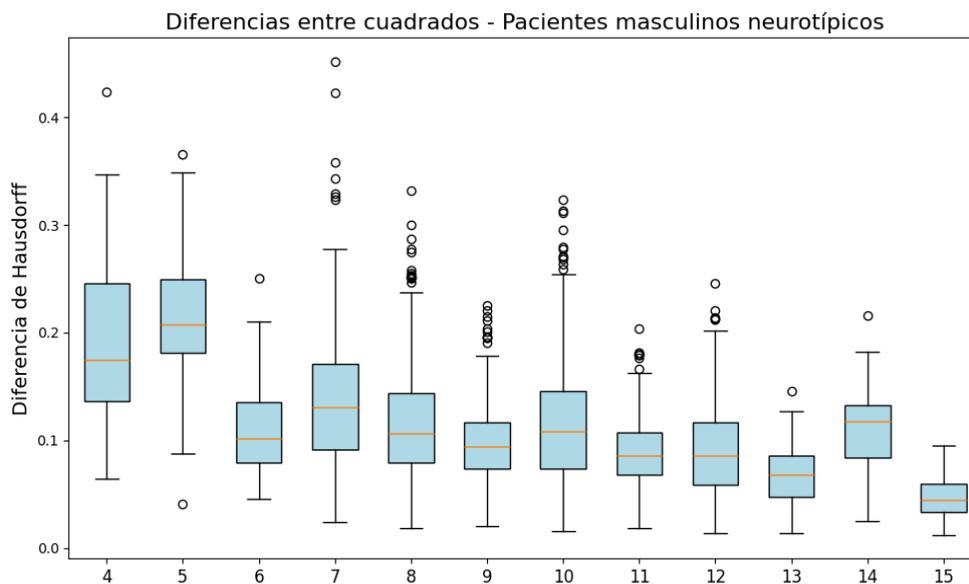


Figura 4.20: Cálculo de diferencia entre cuadrados de pacientes neurotípicos

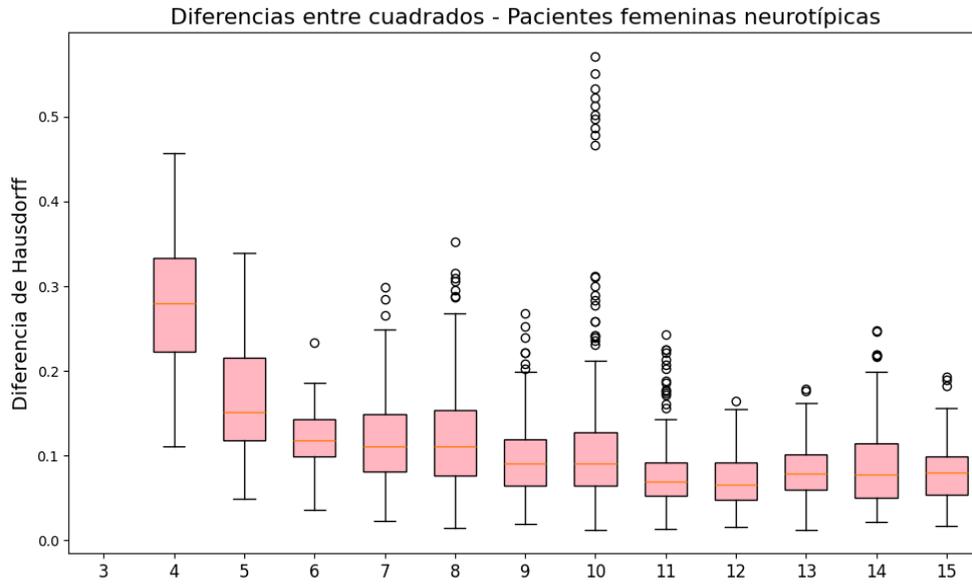


Figura 4.21: Cálculo de diferencia entre cuadrados de pacientes femeninas neurotípicas

A continuación, se presentan las figuras obtenidas para los pacientes masculinos y femeninos diagnosticados con TEA respectivamente 4.22 y 4.23:

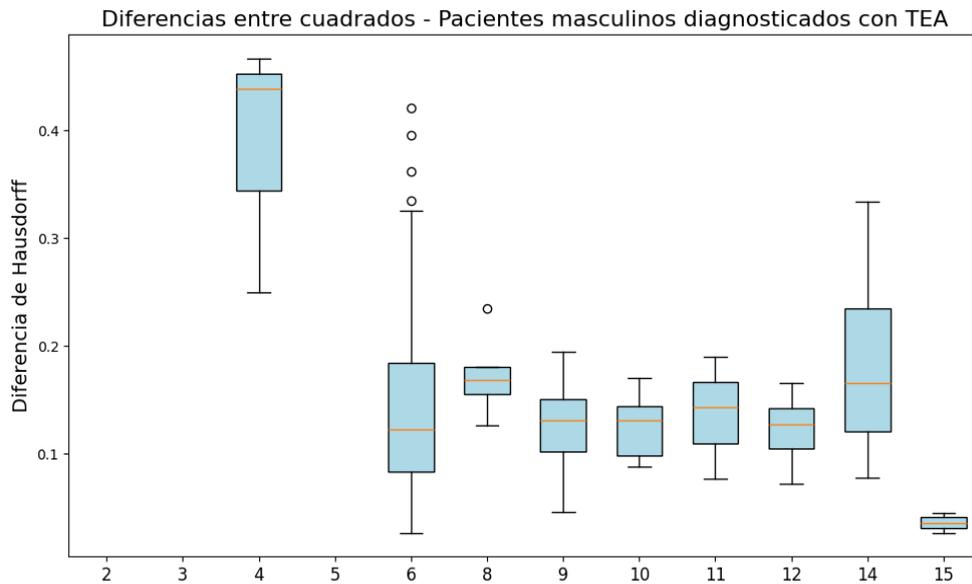


Figura 4.22: Cálculo de diferencia entre cuadrados de pacientes masculinos con TEA

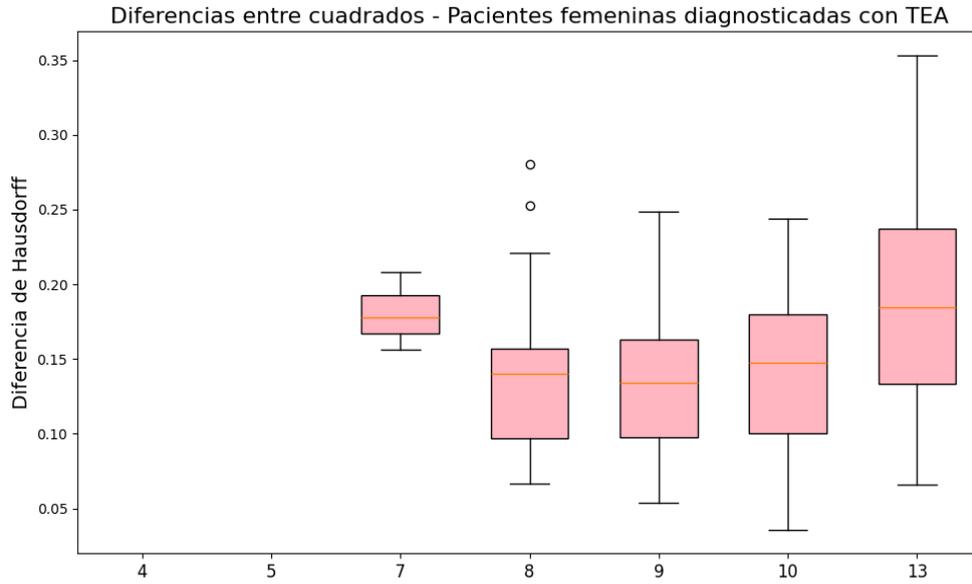


Figura 4.23: Cálculo de diferencia entre cuadrados de pacientes femeninas con TEA

Por último, se presentan las figuras obtenidas por la diferencia de Hausdorff para pacientes masculinos diagnosticados con TDAH 4.24 y femeninas 4.24:

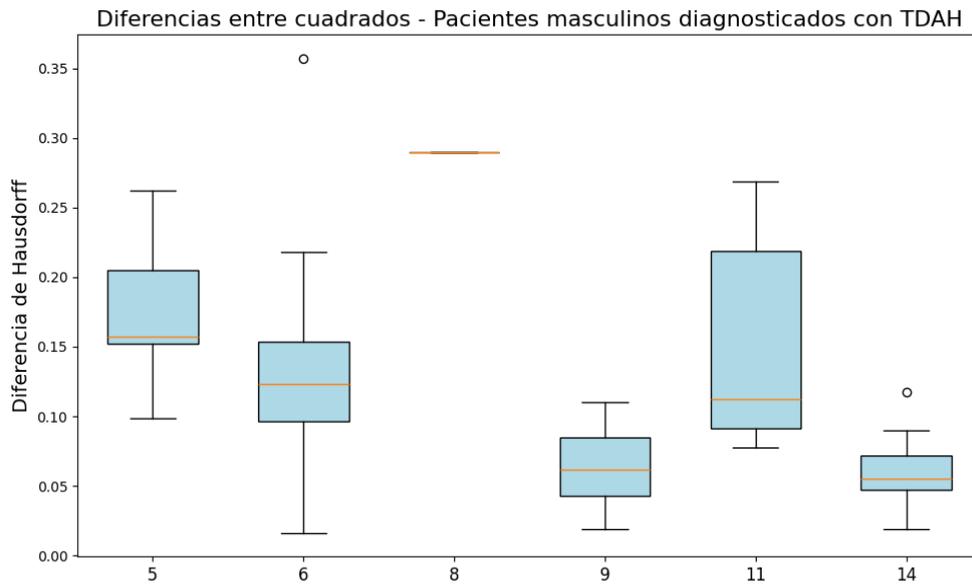


Figura 4.24: Cálculo de diferencia entre cuadrados de pacientes masculinos con TDAH

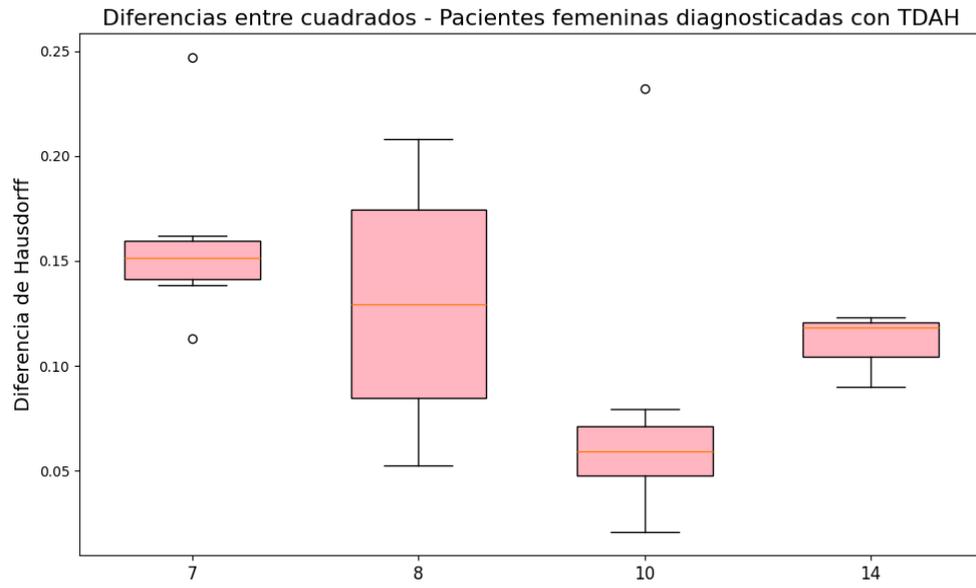


Figura 4.25: Cálculo de diferencia entre cuadrados de pacientes femeninas con TDAH

# Capítulo 5

## Discusión de los resultados obtenidos

Este capítulo presenta y analiza los resultados obtenidos mediante las distintas técnicas de procesamiento de datos utilizados en los perfiles de velocidad y escritura de pacientes diagnosticados con Trastorno del Espectro Autista, Déficit de Atención con Hiperactividad y pacientes neurotípicos. El objetivo principal de este análisis es evaluar y determinar si ciertas características en los patrones de escritura, como la precisión en la representación de figuras geométricas, en este caso, la representación de cuadrados, y la similitud en los perfiles de velocidad podrían contribuir a la discriminación entre estos diagnósticos.

### 5.1. Distribución de porcentaje de cuadrados encontrados

Los gráficos de distribución del porcentaje de cuadrados encontrados en las trayectorias de escritura presentan varias tendencias diferenciadas por diagnóstico y edad de los participantes.

En los pacientes neurotípicos (figura 4.9), se observa una variabilidad muy notable en el porcentaje de cuadrados encontrados a edades tempranas. Los datos son muy dispersos en las edades entre seis y ocho años, lo que sugiere que las habilidades de escritura pueden variar mucho en esas edades. A medida que los pacientes masculinos crecen, especialmente después de los diez años, la mediana del porcentaje de cuadrados encontrados tiende a ser mucho más estable y más alta; esto indica que el desarrollo de los sujetos se relaciona con la precisión de la escritura.

En contraste, en el conjunto de los pacientes masculinos con TDAH (Figura 4.10), los datos presentan menos variaciones que para los neurotípicos, especialmente en los grupos de ocho y catorce años donde el porcentaje de cuadrados parece mantenerse estable; esto implica que el grupo de pacientes con TDAH puede ser menos variable en cuanto a las habilidades de escritura para esas edades. Sin embargo, es importante considerar que los datos son limitados para ciertas edades y esto puede afectar a la precisión de los resultados.

En cuanto a los pacientes masculinos con TEA (Figura 4.11), los gráficos presentan una gran variabilidad en el porcentaje de cuadrados encontrados, especialmente en la edad de 10 años. Las edades más tempranas tienden a tener porcentajes más bajos de cuadrados, lo que puede indicar problemas con la precisión de la escritura a edades más tempranas. Al mismo tiempo, la alta variabilidad en las edades mayores puede indicar que el desarrollo puede estar influido por diferencias en el desarrollo de habilidades motoras finas.

La figura 4.12 correspondiente a las pacientes femeninas neurotípicas, revela un patrón similar al de los pacientes masculinos neurotípicos: una media y precisión crecientes con la edad, junto con una reducción en la variabilidad. Esto sugiere que, al igual que en los pacientes masculinos, existe un posible desarrollo en las habilidades de escritura a medida que las niñas crecen.

Para el caso de las pacientes femeninas con TDAH (figura 4.13), se observa una variabilidad en los porcentajes de cuadrados encontrados, especialmente para las pacientes de diez años, aunque el número de datos para otros grupos de edad es limitado, lo que dificulta obtener conclusiones firmes. Asimismo, en el grupo de pacientes femeninas con TEA (Figura 4.14), se observa una alta variabilidad a lo largo de las edades, especialmente a los 9 años, lo cual puede reflejar las diferencias individuales en habilidades motoras finas en este grupo.

Las figuras 4.15 y 4.16 muestran la distribución de porcentajes de cuadrados hallados según la edad y diagnóstico de los pacientes, resaltando las medianas y los cuartiles de los participantes neurotípicos frente a los valores observados en pacientes con TEA y TDAH. Estas figuras corroboran que los neurotípicos, en general, presentan una mayor consistencia en sus porcentajes de cuadrados encontrados, mientras que los pacientes con TEA y TDAH muestran una mayor variabilidad y un menor rendimiento en varias edades. Esto respalda las observaciones previas acerca del desarrollo diferencial de habilidades motoras en función del diagnóstico.

Los análisis estadísticos (tabla 4.7 y tabla 4.9) que incluyeron pruebas t y el análisis de varianza demostraron diferencias notables para ciertas comparaciones de edad dentro de los grupos neurotípicos de pacientes masculinos y pacientes femeninas. Esto sugiere que las habilidades de escritura de estos grupos se desarrollan en gran medida con la edad. Al comparar los diagnósticos, los resultados mostraron diferencias estadísticamente significativas, particularmente entre los grupos de neurotípicos y del espectro autista entre los 8 y 10 años de edad, lo que puede indicar diferencias en las habilidades de escritura relacionadas con el diagnóstico.

Los resultados obtenidos muestran que tanto los pacientes masculinos como las pacientes femeninas desarrollan sus habilidades de manera diferente y, conforme crecen, su precisión al dibujar mejora progresivamente. Tanto los pacientes masculinos como las pacientes femeninas con TDAH y TEA muestran una mayor variabilidad y, en general, un rendimiento inferior en comparación con los individuos neurotípicos, especialmente a edades tempranas. Estos resultados coinciden con la revisión de Takagi et al. [29], que concluye que las personas con TDAH y TEA tienden a mostrar dificultades en el desarrollo de habilidades motoras, reflejándose en un rendimiento inferior y una mayor variabilidad en comparación con sus pares neurotípicos.

## 5.2. Cálculo de similitud del perfil de velocidad de la escritura

En la figura 4.19 se muestra la comparación de distancias DTW entre los grupos masculinos y femeninos con distintos diagnósticos, correspondientes a edades de nueve y diez años. Los análisis estadísticos revelan diferencias significativas entre los grupos, destacando patrones más similares entre los neurotípicos y TEA, mientras que el grupo TDAH presenta mayor variabilidad.

En cuanto a los neurotípicos, las figuras 4.20 y 4.21 muestran las distancias DTW distribuidas por edades para pacientes masculinos y pacientes femeninas, respectivamente. En el caso de los pacientes masculinos (figura 4.20), se observa una mayor dispersión en edades más tempranas (4-8 años), con una tendencia hacia la estabilización en edades mayores. En las pacientes femeninas (4.21), la variabilidad es similar, aunque con menos datos en edades extremas. Este comportamiento refleja el desarrollo progresivo de habilidades motoras finas con la edad en pacientes neurotípicos, lo que reduce las diferencias en los patrones de escritura.

Para los pacientes con TEA, las figuras 4.22 y 4.23 presentan las distribuciones de distancias DTW. En los pacientes masculinos (figura 4.22), las distancias muestran cierta uniformidad, aunque con variabilidad mayor en las edades más tempranas (4-6 años). Por otro lado, en las niñas (figura 4.23) no se observan datos significativos, lo que podría indicar una falta de información en este grupo. La limitada cantidad de datos en pacientes femeninas con TEA representa una limitación para realizar un análisis concluyente sobre sus patrones de escritura.

Respecto a los pacientes con TDAH, las figuras 4.24 y 4.25 muestran los resultados para pacientes femeninas y masculinos, respectivamente. En los pacientes masculinos (figura 4.25), las distancias DTW son notablemente más elevadas y presentan mayor dispersión en comparación con los neurotípicos y TEA. Esto refleja posibles dificultades motoras o inconsistencias en los movimientos al escribir. En las niñas (figura 4.24), aunque los datos son menos abundantes, también se observa una tendencia similar de mayor variabilidad. Este comportamiento puede estar asociado con características propias del TDAH, como la falta de control motor fino, que impactan directamente en los patrones de velocidad de escritura.

En general, los resultados sugieren que los pacientes con TDAH presentan una mayor variabilidad en sus distancias DTW, diferenciándose claramente de los otros dos grupos. Por el contrario, los pacientes neurotípicos y con TEA muestran patrones más similares, especialmente en el rango de edades de 9-10 años. Esta estabilización podría estar relacionada con el desarrollo motor que se produce a medida que los pacientes masculinos y femeninas crecen, lo que se refleja en una menor dispersión en las distancias DTW para edades mayores. Esto sugiere que las características de la velocidad de escritura podrían utilizarse para distinguir entre los grupos de diagnóstico en futuros estudios.

Es importante destacar la limitación en la cantidad de datos disponibles, especialmente en pacientes femeninas con TEA, lo que podría sesgar los resultados. También se observa una

mayor variabilidad en las edades más tempranas en todos los grupos, lo que podría deberse a una mayor inestabilidad motora en estas etapas del desarrollo. Finalmente, se sugiere complementar este análisis con un estudio más profundo de las diferencias entre sexos dentro de cada diagnóstico para identificar posibles patrones específicos.

### 5.3. Diferencia de Hausdorff entre cuadrados

El análisis de diferencias de Hausdorff entre cuadrados proporciona información valiosa sobre la variación en la precisión de dibujo de cuadrados por cada grupo. Esta medida permite evaluar la exactitud en la ejecución de las figuras cuadradas, ofreciendo una visión detallada de las habilidades motoras de los pacientes en función de su edad y diagnóstico.

En las figuras correspondientes a pacientes neurotípicos, tanto masculinos como femeninas (figuras 4.20 y 4.21 respectivamente), se observa que las diferencias tienden a ser menores a medida que los pacientes avanzan en edad, especialmente a partir de la franja de los 6-8 años. Este patrón sugiere que, a medida que desarrollan sus habilidades motoras, los pacientes neurotípicos logran mejorar y estabilizar su capacidad de trazar figuras con mayor precisión. Esta tendencia hacia una menor variabilidad en las diferencias de Hausdorff refleja un constante progreso en la consistencia de los trazos, lo que podría atribuirse al desarrollo motor propio de la infancia y la adolescencia.

En contraste, las figuras correspondientes a los pacientes con TDAH y TEA muestran valores de diferencia más elevados y una mayor varianza. En el caso de los pacientes masculinos y las pacientes femeninas con TDAH (figura 4.24 y figura 4.25), se observa una notable dispersión en las diferencias de Hausdorff, lo cual sugiere que, aunque algunos pacientes logran mejorar con la edad, otros continúan presentando grandes variaciones en la precisión de sus trazos. Esta mayor dispersión podría estar relacionada con las dificultades en el control de la atención y el movimiento, propias del diagnóstico de TDAH, que impactan la capacidad de mantener una trayectoria estable y repetible en el dibujo.

En lo que respecta a los pacientes con TEA, tanto pacientes masculinos como pacientes femeninos (figura 4.22 y figura 4.23), la variabilidad también es alta, especialmente en edades tempranas. La dispersión de los valores de diferencia en las figuras sugiere que las dificultades en la percepción y control motor afectan su capacidad para trazar figuras con precisión. Aunque algunos pacientes muestran una tendencia hacia la estabilización en la precisión de los trazos con la edad, otros continúan experimentando variaciones significativas, lo cual podría estar relacionado con las características particulares del TEA en cuanto a habilidades motrices y perceptivas.

Estos resultados destacan cómo el diagnóstico y la edad influyen directamente en la capacidad de los pacientes para dibujar figuras geométricas de forma precisa. En este contexto, la diferencia de Hausdorff ha demostrado que puede ser una herramienta efectiva para evaluar la precisión y consistencia en el dibujo, evidenciando patrones de desarrollo diferenciados entre los diferentes grupos, proporcionando una base para analizar cómo los distintos diagnósticos afectan las habilidades motoras de cada paciente.

## 5.4. Comparación de resultados con trabajos previos

Los resultados obtenidos complementan los estudios previos en varios aspectos clave y aportan nuevas perspectivas metodológicas y prácticas al análisis del perfil de velocidad y trayectoria de la escritura en pacientes con TEA, TDAH y neurotípicos. En el estudio de Cohen et al. [8], por ejemplo, se señala que los niños con TDAH tienden a tener una escritura más desorganizada y muy variable, debido a los problemas de control motor y atención característicos de este trastorno. En este trabajo, estas observaciones se reflejan en las mayores distancias DTW obtenidas para los pacientes con TDAH, especialmente en edades tempranas. Además, este análisis también permite cuantificar dichas variaciones de manera precisa, hecho que no se había explorado en investigaciones previas.

Además, las diferencias de género observadas concuerdan los hallazgos de Mahendiran et al. [20], donde ya se destacaron variaciones adaptativas sociales y diferencias motoras entre pacientes con TEA y TDAH. Este trabajo amplía esa perspectiva al analizar específicamente los trazos de escritura en ambos géneros. Sin embargo, la menor representación de niñas y, en especial, las diagnosticadas con TEA, sigue siendo una limitación importante, como también se señaló en el estudio de Montagut et al [22]. La escasez de datos para este grupo refuerza la necesidad de futuros estudios que aborden esta brecha.

De igual manera, en el estudio de Takagi et al. [29], se identificaron características motoras distintivas en TEA y TDAH a través de una revisión sistemática. Este trabajo no solo confirma sus hallazgos, como la rigidez motora en pacientes con TEA y la mayor dispersión en TDAH, sino que también introduce un enfoque computacional y cuantitativo. El uso de métricas avanzadas como la distancia DTW y la diferencias de Hausdorff permite analizar las diferencias de forma cuantitativa, ofreciendo una herramienta robusta para discriminar entre patrones motores asociados a diferentes diagnósticos.

Sin embargo, este estudio presenta limitaciones, al igual que muchos otros. La limitada cantidad de datos, en especial los relacionados con pacientes femeninas diagnosticadas con TEA podría afectar a la generalización de los resultados. Además, las condiciones controladas bajo las que se recopilaron los datos, pueden no reflejar con precisión las condiciones naturales de escritura de los pacientes. Superar estas limitaciones requiere que futuras investigaciones aumenten el tamaño de la muestra y que incluyan diversos entornos de evaluación.

En definitiva, este trabajo no solo complementa el ‘estado del arte’, sino que también introduce nuevos enfoques metodológicos que permiten un análisis más profundo, riguroso y detallado. Estas contribuciones no solo refuerzan la utilidad de las métricas trabajadas para el diagnóstico diferencial entre TEA y TDAH, sino que también sientan las bases para futuras aplicaciones clínicas y académicas.

## 5.5. Limitaciones del estudio

A pesar de los resultados positivos, es importante destacar algunas limitaciones del estudio. En primer lugar, la muestra utilizada, aunque representativa, es relativamente pequeña, especialmente en los subgrupos de pacientes masculinos y femeninas con TDAH y TEA (Tabla 4.5). Esto puede limitar la generalización de los resultados a una población más amplia.

Además, la recogida de datos se realizó en entornos controlados, lo que podría no reflejar completamente las condiciones cotidianas en las que se manifiestan las dificultades de escritura en pacientes con estos trastornos.

Otra limitación importante es la dependencia del análisis en herramientas y algoritmos específicos. Aunque estos demostraron ser eficaces en la detección de patrones de escritura, pueden no capturar todas las sutilezas presentes en los trazos de los pacientes. Futuras investigaciones deberían considerar la integración de técnicas más avanzadas, como aprendizaje automático y redes neuronales, para mejorar la precisión y la robustez del diagnóstico.

Para mitigar estas limitaciones, es fundamental ampliar la muestra en futuros estudios, incluir condiciones más variadas para la recopilación de datos y analizar con mayor profundidad cómo otros factores pueden afectar los patrones observados.

# Capítulo 6

## Conclusiones y trabajo futuro

### 6.1. Conclusiones

Este trabajo de fin de título presenta una herramienta desarrollada en Python, diseñada como herramienta complementaria a los métodos convencionales para discriminar entre el Trastorno del Espectro Autista (TEA) y el Trastorno por Déficit de Atención e Hiperactividad (TDAH). El análisis se centra en el análisis del perfil de velocidad y trayectoria de la escritura, específicamente en la prueba de dibujo de formas geométricas simples.

A partir de los resultados obtenidos y de la discusión llevada a cabo en el capítulo anterior, se pueden extraer las siguientes conclusiones de la aplicación desarrollada:

1. **Eficiencia en la discriminación de diagnósticos:** La herramienta desarrollada ha demostrado ser efectiva al permitir diferenciar entre los patrones de escritura de los tres diagnósticos estudiados. Concretamente, el análisis mediante la distancia DTW reveló que los pacientes con TDAH presentan una mayor variabilidad en la velocidad de escritura, mientras que los patrones de los pacientes con TEA muestran una mayor consistencia y rigidez, asemejándose a los resultados obtenidos por los pacientes neurotípicos.
2. **Relevancia de las métricas utilizadas:**
  - **Discriminación de TDAH:** El cálculo de similitud del perfil de velocidad de la escritura mediante la **distancia DTW es una herramienta efectiva**. Los pacientes con TDAH muestran una mayor variabilidad en la velocidad de los trazos, evidenciando lapsos de pérdida de control motor y atención, características asociadas a este diagnóstico.
  - **Discriminación de TEA:** El cálculo de la **distancia de Hausdorff es una herramienta efectiva** para la discriminación de TEA ya que permite cuantificar el grado de precisión en la reproducción de figuras geométricas, en este caso, la representación de cuadrados, lo que es útil para capturar la rigidez motriz asociada a este diagnóstico.

3. **Evolución de patrones según la edad:** Los datos obtenidos indican que los pacientes neurotípicos tienden a reducir la variabilidad de sus trazos conforme crecen y maduran, evidenciando un desarrollo motor más estable. Por el contrario, los pacientes con TEA y TDAH mantienen una mayor dispersión, especialmente a edades tempranas.
4. **Muestra limitada:** A pesar de los resultados obtenidos, este trabajo presenta limitaciones como el reducido tamaño de la muestra en ciertos grupos, como las pacientes femeninas con TEA, o haberse realizado en entornos controlados, lo que podría limitar la generalización de los resultados a contextos cotidianos.

## 6.2. Propuestas para investigación futura

El presente trabajo ha demostrado que el análisis del perfil de velocidad y trayectoria de la escritura es una herramienta útil para la discriminación entre TEA, TDAH. Sin embargo, existen diversas áreas que podrían beneficiarse de un mayor desarrollo y profundización en futuras investigaciones:

- **Ampliación del tamaño de la muestra:** A pesar de que este trabajo ha proporcionado resultados relevantes, el tamaño de la muestra utilizada es limitado, especialmente en el caso de los pacientes diagnosticados con TEA y TDAH, pero sobre todo en el caso de las niñas con estos diagnósticos, por lo que es necesario aumentar el número de participantes femeninas para obtener conclusiones más robustas y generalizables.
- **Incorporación de nuevas métricas:** Aunque este trabajo se ha centrado en el análisis de la representación de dibujos de cuadrados, se deberían explorar otras métricas que podrían complementar este análisis; por ejemplo, el análisis de la presión del trazo, el tiempo que se tarda en trazar un cuadrado o la fluidez del movimiento podría aportar información adicional sobre las diferencias motoras entre individuos con TEA y TDAH.
- **Desarrollo de algoritmos de aprendizaje automático:** Una posible línea de investigación futura sería el desarrollo de modelos de aprendizaje automático que pudieran automatizar la clasificación entre TEA, TDAH y neurotípicos con mayor precisión. El uso de redes neuronales podría mejorar la capacidad predictiva de la herramienta.
- **Aplicación en entornos clínicos:** Si bien el presente estudio ha tenido un enfoque investigativo, futuras investigaciones podrían centrarse en la implementación y validación de esta herramienta en entornos clínicos reales. Un estudio longitudinal que examine su efectividad en el diagnóstico temprano y en la monitorización de progresos terapéuticos sería de gran valor para los profesionales de la salud.
- **Análisis del impacto de la intervención terapéutica:** Un área de futuro interés podría ser investigar cómo los patrones de escritura evolucionan con la intervención terapéutica, especialmente en individuos con TEA y TDAH. Esto permitiría evaluar el impacto de distintas terapias y ajustar los tratamientos basados en los cambios observados en los perfiles de escritura.

# Apéndice A

## Anexo: Manual de usuario

Dentro del repositorio GitHub de este trabajo, se pueden encontrar varios archivos que contienen códigos desarrollados para la extracción de datos y parámetros. Es importante destacar, que para poder ejecutarlos obteniendo los resultados mostrados en este documento, se ha de seguir la misma estructura de ficheros indicada en la figura 4.3.

### Ejecución de `automatic.py`

Para poder ejecutar el código que se encuentra en este fichero, en VS Code deberemos de pulsar las teclas de nuestro teclado **Ctrl + Alt + N**, o buscar en la parte superior derecha el símbolo  .

A continuación y a través de tkinter, aparecerá (normalmente en la mitad izquierda de la pantalla) un botón como el siguiente:

**Seleccionar directorio**

Pulsando sobre dicho botón, se nos abrirá las diferentes carpetas

Nombre	Fecha de modificación	Tipo
Neurotípicos	DD/MM/2024 HH:MM	Carpeta de archivos
TDAH	DD/MM/2024 HH:MM	Carpeta de archivos
TEA	DD/MM/2024 HH:MM	Carpeta de archivos

Si navegamos por los distintos directorios, llegaremos a las carpetas en las que se encuentran los datos de cada dibujo realizado por cada pacientes. Un ejemplo de ruta sería:

Ruta del archivo

`./V3/TEA/Male/10/id_paciente_10_M`

Para procesar un paciente, seleccionamos (no hacer doble ‘click’) una carpeta ‘id\_paciente\_10\_M’, y presionamos el botón en la parte inferior derecha:

**Seleccionar carpeta**

Tras esto, obtendremos los resultados de la ejecución del código con una figura como la 4.7

## Ejecución de `automatic_v2.py`

Antes de comenzar con la descripción de cómo ejecutar `automatic_v2.py`, se debe aclarar que para una mayor rapidez y facilitar la representación de las figuras mostradas en este documento, se ha decidido crear un fichero para cada sexo y diagnóstico quedando de la siguiente manera:

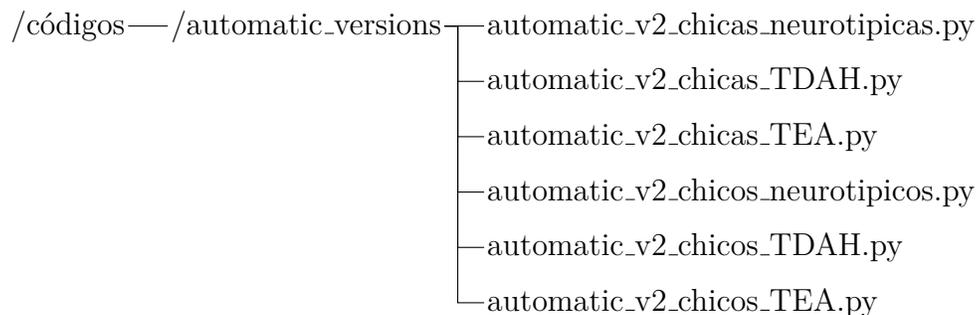


Figura A.1: Estructura de la carpeta `automatic_versions`

Para obtener las mismas figuras adjuntas en este trabajo, bastará con abrir alguno de los archivos y pulsar las teclas **Ctrl + Alt + N** o presionar el símbolo  , tal y como se indicó anteriormente.

## Ejecución de `analysis_squares_diagnostics.py`

Para poder ejecutar el código que se encuentra en este fichero y obtener las figuras 4.15 y 4.16, en VS Code deberemos de pulsar las teclas de nuestro teclado **Ctrl + Alt + N**, o buscar en la parte superior derecha el símbolo .

A continuación y a través de tkinter, aparecerá una ventana como la siguiente, en el que, dependiendo de qué género se quiera tratar, deberemos de seleccionar ‘masculino’ o ‘femenino’:

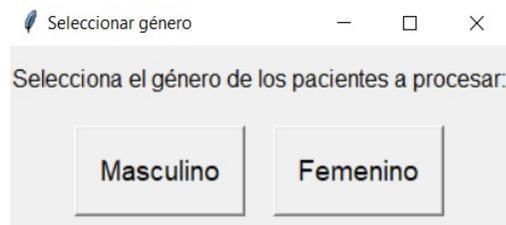
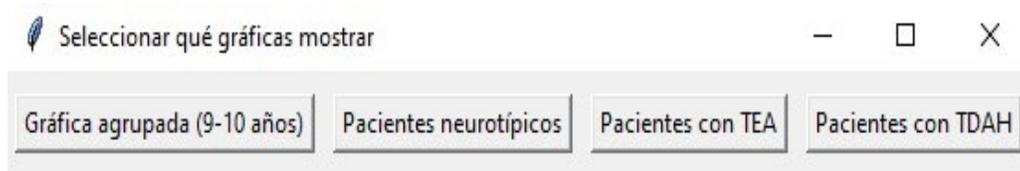


Figura A.2: Botón de selección de sexo

## Ejecución de `distance_dtw.py`

Para poder ejecutar el código que se encuentran en este fichero, al igual que los códigos anteriores, en VS Code deberemos de pulsar las teclas de nuestro teclado **Ctrl + Alt + N**, o buscar en la parte superior derecha el símbolo .

A continuación y a través de tkinter, aparecerá (normalmente en la cerca de la esquina superior izquierda de la pantalla) una interfaz gráfica como la siguiente

Figura A.3: Botón de selección para `distance_dtw.py`

Sobre el que tendremos que hacer 'click' en el apartado que queramos procesar y el código solo ejecutará los datos necesarios para obtener las figuras relacionadas con la opción seleccionada. Una vez finalice su ejecución, mostrará deseada y que se muestran en su conjunto en Cálculo de similitud del perfil de velocidad de la escritura.

# Apéndice B

## Anexo: Presupuesto

El desarrollo de esta memoria implicó la utilización de diversos recursos. A continuación se detalla el presupuesto estimado para la implementación de este proyecto:

- **Ordenador portátil:** Se utilizó un ordenador portátil con las especificaciones mencionadas en el apartado ‘Recursos hardware y software utilizados’.
- **Dispositivos de entrada:** Para la recogida de datos, se hizo uso de un iPad Pro 2018 y un Apple Pencil para que los pacientes realizaran los dibujos. [36] [9]
- **Un ingeniero informático:** Para el desarrollo de los códigos y la redacción de esta memoria.
- **Tiempo de desarrollo:** Tiempo empleado para el desarrollo de este trabajo.

### Costo del Portátil

El portátil utilizado para realizar el trabajo tiene las siguientes características:

- Costo inicial: €900
- Vida útil: 6 años
- Horas de uso diario: 10 horas

La vida útil del portátil en horas es:

$$6 \text{ años} \times 365 \text{ días/año} \times 10 \text{ horas/día} = 21.900 \text{ horas}$$

El costo por hora de uso del portátil es:

$$\frac{900 \text{ €}}{21.900 \text{ horas}} \approx 0,041 \text{ €/hora}$$

Para 300 horas de uso, el costo del portátil es:

$$300 \text{ horas} \times 0,041 \text{ €/hora} = 12,30 \text{ €}$$

## Costo del Salario del Ingeniero Informático

El salario promedio de un ingeniero informático en España es €27,000 al año, lo que equivale a €13.85 por hora según la Universidad Alfonso X el Sabio.[30]

Para 300 horas de trabajo, el costo es:

$$300 \text{ horas} \times 13,85 \text{ €/hora} = 4.155\text{€}$$

## Costo de la Electricidad

Asumimos un consumo de 50W (0.05 kW) por hora y un costo de 0,30€ por kWh.

El costo de la electricidad para 300 horas es:

$$300 \text{ horas} \times 0,05 \text{ kW} \times 0,30 \text{ €/kWh} = 4,5 \text{ €}$$

## Costo del iPad

El costo inicial del iPad fue de 879€. La vida útil estimada en términos de horas es de 2.000 horas. La amortización por hora se calcula:

$$\frac{879 \text{ €}}{2.000 \text{ horas}} = 0,4395 \text{ €/hora}$$

Con un uso de aproximadamente 15 horas, la amortización total es de:

$$0,4395 \text{ €/hora} \times 15 \text{ horas} = 6,59 \text{ €}$$

## Costo del Apple Pencil

El costo inicial del Apple Pencil fue de 100 EUR. La vida útil estimada en términos de horas es de 500 horas. La amortización por hora se

$$\frac{100 \text{ €}}{500 \text{ horas}} = 0,20 \text{ €/hora}$$

Con un uso de aproximadamente 15 horas, la amortización total es de:

$$0,20 \text{ €/hora} \times 15 \text{ horas} = 3,00 \text{ €}$$

## Costo Total

Sumando todos los costos, tenemos:

$$\begin{aligned} &12,30 \text{ € (portátil)} + 6,59 \text{ € (iPad)} + 3,00 \text{ € (Apple Pencil)} \\ &\quad + 4.155 \text{ € (salario)} + 4,50 \text{ € (electricidad)} \\ &= 4.181,39 \text{ €} \end{aligned}$$

El costo total estimado para la realización de este trabajo es de **4.181,39€**

# Apéndice C

## Anexo: Normativa y Legislación

Este trabajo de fin de título ha sido desarrollado conforme a la normativa vigente en materia de protección de datos personales, seguridad de la información y ética en la investigación. A continuación, se detallan las principales leyes y normativas que se han considerado, así como el procedimiento seguido para obtener el consentimiento informado de los participantes, asegurando el respeto a sus derechos y la correcta gestión de la información.

### C.1. Ley Orgánica de Protección de Datos Personales y Garantía de los Derechos Digitales

La “*Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales*” (LOPDyGDD)[7] adapta el derecho español al Reglamento General de Protección de Datos (RGPD)[12] de la Unión Europea. Esta ley establece los principios y derechos que deben respetarse al tratar datos personales, incluyendo el consentimiento explícito de los participantes, el derecho de acceso, rectificación, cancelación y oposición, así como la obligación de garantizar la seguridad y confidencialidad de los datos personales tratados.

En este trabajo, se han seguido las directrices establecidas por la LOPDyGDD para asegurar que los datos personales de los participantes, especialmente aquellos recogidos durante las sesiones de escritura en tablets, sean tratados con la máxima confidencialidad. Se obtuvo el consentimiento informado por escrito de los padres o tutores legales de todos los menores participantes, asegurando que comprendan el propósito del estudio y sus derechos.

### C.2. Reglamento General de Protección de Datos

El “*Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de*

*datos personales y a la libre circulación de estos datos*” (RGPD)[12] es aplicable en todos los estados miembros de la Unión Europea. El RGPD establece principios como la minimización de datos, la limitación de la finalidad y la necesidad de implementar medidas técnicas y organizativas apropiadas para proteger los datos personales.

En cumplimiento con el RGPD, se adoptó un enfoque de minimización de datos, recopilando únicamente la información estrictamente necesaria para los fines del estudio, garantizando la anonimización de los datos sensibles cuando fue requerido.

### C.3. Normas Éticas en Investigación

El estudio también ha cumplido con normas éticas para asegurar que los participantes no sean expuestos a riesgos físicos o psicológicos durante la recolección de datos. El consentimiento informado obtenido de los participantes y sus tutores legales incluyó una explicación detallada sobre la naturaleza del estudio, los procedimientos a seguir, los posibles riesgos y beneficios, y el derecho a retirarse en cualquier momento sin penalización alguna.

### C.4. Consentimiento Informado para la Participación en el Estudio

Previo a la realización de las pruebas, se obtuvo un consentimiento informado por escrito de los padres o tutores legales de los menores. Este documento tenía como propósito asegurar que los padres o tutores comprendieran completamente los detalles del estudio.

#### Aspectos Claves del Consentimiento

1. **Propósito del estudio:** Se explicó que el estudio tenía como objetivo analizar datos relacionados con los patrones de movimiento de los pacientes durante pruebas en tablets, contribuyendo a la investigación en el desarrollo cognitivo y neurológico.
2. **Descripción de las pruebas:** Las pruebas consistían en actividades en un iPad, registrando datos como velocidad y trayectoria de los trazos, sin implicar riesgos físicos o emocionales.
3. **Confidencialidad de los datos:** Se garantizó que todos los datos serían tratados de manera confidencial, y la identidad de los participantes no sería revelada en ningún momento.
4. **Voluntariedad de la participación:** Se subrayó que la participación era completamente voluntaria, y los padres o los pacientes podían retirarse en cualquier momento sin consecuencias negativas.

5. **Aprobación ética:** El estudio fue revisado y aprobado por el comité ético correspondiente, cumpliendo con las normativas vigentes de protección de menores en investigaciones científicas.

Este consentimiento fue un requisito indispensable para asegurar que los derechos de los participantes, en particular de los menores, fueran protegidos en todo momento.

## C.5. Legislación Adicional Relacionada

Además de las normativas mencionadas, se tuvieron en cuenta otras leyes relevantes:

- **Ley 14/2007, de 3 de julio, de Investigación Biomédica:** regula la investigación en seres humanos y protege los derechos de los participantes.[5]
- **Ley 41/2002, de 14 de noviembre, reguladora de la autonomía del paciente y derechos y obligaciones en materia de información y documentación clínica:** protege los derechos de los menores en estudios de investigación.[4]
- **Ley 9/2017, de 8 de noviembre, de Contratos del Sector Público:** relevante al haber utilizado fondos públicos, estableciendo obligaciones de transparencia.[6]

# Bibliografía

- [1] (2024). *Python Language Reference, version 3.11.7*. Python Software Foundation. Accedido: 30 de enero de 2024.
- [2] Apple Inc. (2024). Testflight overview. Accedido: 17 de octubre de 2024.
- [3] Association, A. P. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Publishing, 5th edition edition. Accedido: 23 de junio de 2024.
- [4] Boletín Oficial del Estado(BOE) (2002). Ley orgánica 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica. Accedido: 11 de septiembre de 2024.
- [5] Boletín Oficial del Estado(BOE) (2007). Ley orgánica 14/2007, de 3 de julio, de investigación biomédica. Accedido: 11 de septiembre de 2024.
- [6] Boletín Oficial del Estado(BOE) (2017). Ley de contratos del sector público, texto refundido aprobado por el real decreto legislativo 3/2011, de 14 de noviembre. Accedido: 11 de septiembre de 2024.
- [7] Boletín Oficial del Estado(BOE) (2018). Ley orgánica de protección de datos personales y garantía de los derechos digitales. Accedido: 11 de septiembre de 2024.
- [8] Cohen, R., Sasson, A., and Gal, E. (2016). Handwriting in children with adhd: The role of graphology. *Journal of Attention Disorders*, 20(3):209–217.
- [9] Componentes, P. (2024). Apple pencil. Accedido: 5 de septiembre de 2024.
- [10] Conversation, T. (2024). ¿qué es el ‘masking’ en el autismo? Accedido: 1 de noviembre de 2024.
- [11] Csibra, B., Bunford, N., and Gácsi, M. (2024). Development of a human-analogue, 3-symptom domain dog adhd and functionality rating scale (dafrs). *Scientific Reports*, 14(1).
- [12] Diario Oficial de la Unión Europea (2016). Reglamento general de protección de datos. Accedido: 11 de septiembre de 2024.
- [13] Escuela de Ingeniería Informática de la Universidad de Las Palmas de Gran Canaria (2024). Grado en ingeniería informática - rev03. Accedido: 24 de julio de 2024.

- [14] Foundation, P. S. (2024). *Tkinter: Python's standard GUI (Graphical User Interface) package, version 8.6.12*. Accedido: 14 de junio de 2024.
- [15] GeeksforGeeks (2024). How to detect shapes in images in python using opencv. Accedido: 25 de Marzo 2024.
- [16] GitHub, Inc. (2024). Github: Plataforma de desarrollo colaborativo utilizando git. Accedido: 26 de marzo de 2024.
- [17] Harris, C. R. et al. (2021). *NumPy: Array processing for numbers, strings, records, and objects, version 1.20.3*. Accedido: 30 de enero de 2024.
- [18] Hunter, J. D. (2021). *Matplotlib: Visualization with Python, version 3.4.2*. Accedido: 14 de junio de 2024.
- [19] Itseez (2021). Opencv documentation, version 4.5.1. *OpenCV*. Accedido: 25 de Marzo 2024.
- [20] Mahendiran, T., Dupuis, A., Crosbie, J., Georgiades, S., Kelley, E., Liu, X., Nicolson, R., Schachar, R., Anagnostou, E., and Brian, J. (2019). Sex differences in social adaptive function in autism spectrum disorder and attention-deficit hyperactivity disorder. *Frontiers in psychiatry*, 10:607.
- [21] McKinney, W. (2021). *Pandas: Powerful data structures for data analysis, version 1.2.4*. Accedido: 14 de junio de 2024.
- [22] Montagut Asunción, M., Mas Romero, R. M., Fernández Andrés, M. I., and Pastor Cerezuela, G. (2018). Influencia del sesgo de género en el diagnóstico de trastorno de espectro autista: una revisión. *Escritos de Psicología (Internet)*, 11(1):42–54.
- [23] Montgomery, D. C. (2019). *Design and Analysis of Experiments*. John Wiley & Sons.
- [24] OpenAI (2024). Chatgpt: Generative pre-trained transformer by openai. Accedido: Durante el desarrollo del trabajo.
- [25] Overleaf (2024). Overleaf, online latex editor. Accedido: 29 de enero de 2024.
- [26] Python Software Foundation (2024). Python virtual environments (venv). Accedido: 30 de enero de 2024.
- [27] Red Hat (2024). ¿qué es la metodología ágil? Accedido: 23 de noviembre de 2024.
- [28] Salvador, S. and Chan, P. (2024). *FastDTW: Fast Approximation of Dynamic Time Warping*. Accedido: 1 de noviembre de 2024.
- [29] Takagi, S., Hori, H., Yamaguchi, T., Ochi, S., Nishida, M., Maruo, T., and Takahashi, H. (2022). Motor functional characteristics in attention-deficit/hyperactivity disorder and autism spectrum disorders: A systematic review. *Neuropsychiatric Disease and Treatment*, 18:1679–1695.

- [30] Universidad Alfonso X el Sabio (2024). ¿cuánto gana un ingeniero informático en España? Accedido: 26 de julio de 2024.
- [31] Virtanen, P. et al. (2021). *SciPy: Scientific Computing Tools for Python, version 1.6.3*. Accedido: 14 de junio de 2024.
- [32] Wackerly, D., Mendenhall, W., and Scheaffer, R. (2008). *Mathematical Statistics with Applications*. Cengage Learning.
- [33] Wikipedia contributors (2024). Deformación dinámica del tiempo. Accedido: 5 de noviembre de 2024.
- [34] World Health Organization (2024a). Autism spectrum disorders. Accedido: 23 de junio de 2024.
- [35] World Health Organization (2024b). Mental disorders: Attention-deficit/hyperactivity disorder (adhd). Accedido: 23 de junio de 2024.
- [36] Xataka (2018). ipad pro 2018: características y ficha técnica. Accedido: 5 de septiembre de 2024.

