

Robustez de las métricas de clasificación de cadencia de tecleo frente a variaciones emocionales

DOCTORADO EN CIENCIAS INFORMÁTICAS, FACULTAD DE INFORMÁTICA, UNLP

Tesista: Enrique P. Calot

Jorge S. Ierache (director), Waldo Hasperué (codirector)

Defendida el 18 de junio de 2019.

Laboratorio de Sistemas de Información Avanzados,
Facultad de Ingeniería, Universidad de Buenos Aires, Argentina
ecalot@lsia.fi.uba.ar

Resumen La dinámica de tecleo permite la identificación de una persona por la forma en que escribe. La tesis aquí resumida se enfoca en analizar la robustez de los algoritmos de análisis de cadencia de tecleo frente a variaciones en los registros biométricos mediante electroencefalografía y cuestionario de autoinforme, utilizando el enfoque dimensional para modelar estados afectivos. Se realizó un experimento para capturar patrones de tecleo en diferentes estados afectivos. Los resultados sugirieron que la tasa de aciertos para ciertas distancias de clasificación, como las métricas A y R, la distancia de Camberra, Manhattan y una distancia basada en Minkowski se ven influenciadas negativamente por los cambios en las respuestas de excitación y valencia. La distancia euclídea fue la menos afectada de las seis evaluadas.

Keywords: Keystroke dynamics, Biometrics, EEG, Sentiment analysis

1. Introducción

El ritmo de escritura es único para cada persona, y las variaciones entre las personas son tan profundas que permiten la identificación del mecanógrafo. En el teclado, el estudio de los patrones de tecleo así como los ritmos en sí se denominan *keystroke dynamics*. En las computadoras, el inicio de esta área particular de investigación se remonta a Gaines et al. [1], quien en 1980 realizó el primer análisis estadístico con textos libres y palabras fijas. Cada estilo de escritura es exclusivo de su propietario y no puede ser robado, perdido u olvidado [2].

El análisis de pulsación de teclas, como característica biométrica, se puede aplicar sobre un texto estático o un texto libre continuo. Los textos estáticos, como las contraseñas, con frecuencia se escriben utilizando la misma secuencia; mientras que los textos libres continuos comparan estilos de escritura usando secuencias de teclas comunes entre dos textos para determinar si pertenecen a la misma persona. La longitud de estos textos puede diferir considerablemente. Los textos libres requieren cadenas de teclas más largas para lograr la misma precisión que en textos estáticos [3].

2. Motivación

La necesidad de brindar una mayor capacidad de autenticación de usuarios con el empleo de biometría fue en aumento en los últimos años. La cadencia de tecleo permite generar firmas de seguridad sin tener que usar dispositivos específicos, siendo completamente transparente y no invasiva para el usuario. Teniendo en cuenta la hipótesis de que la variabilidad emocional influiría en la calidad de la autenticación, es esencial construir lotes de datos representativos de la variabilidad emocional; además, determinar qué métrica es la más adecuada para la cadencia de tecleo, cuál adoptar, qué mejor técnica usar y qué tan robusta es cada una.

3. Temas de investigación

La tesis [4] se centró en los cambios emocionales a corto plazo, principalmente la relajación y la emoción. Sin pretender incursionar en el área de la neurociencia, solo se enfocó en analizar la robustez de los algoritmos de *keystroke dynamics* frente a las variaciones en los registros biométricos a través de la electroencefalografía (EEG), en particular las ondas alfa, beta o cerebrales que, según la literatura [5-15], están asociados con estados de relajación y excitación.

El desarrollo abarca tres preguntas, las dos primeras son secundarias y necesarias para avanzar sobre la tercera, que es la central; a saber

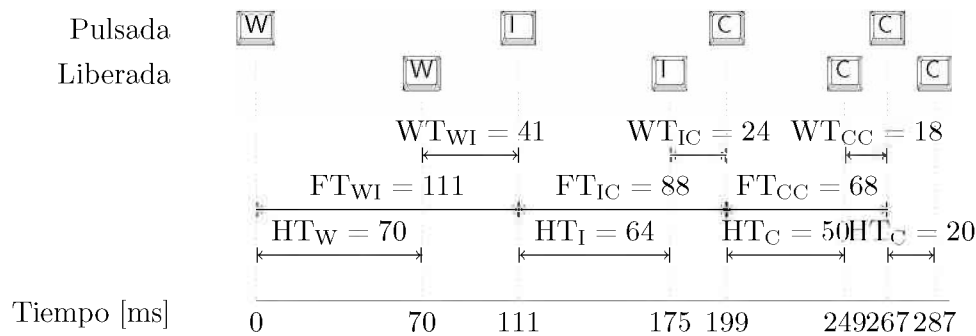


Figura 1: Tiempo interteclas.

1. ¿qué métricas se deben utilizar en el análisis de los resultados?
2. ¿qué cantidad de keystrokes es necesaria para que las pruebas sean estadísticamente significativas?
3. pregunta principal: ¿pueden los estados emocionales afectar la calidad de los algoritmos de clasificación mediante métricas utilizando cadencia de tecleo?

En clasificación de cadencia de tecleo las métricas son formas de medir la distancia entre dos patrones para sí decidir si pertenecen o no a una misma clase (por ejemplo, la misma persona).

3.1. Keystroke dynamics

La Figura 1 presenta un ejemplo de la cadencia de para la palabra "wicc". Hold time (HT) y wait time (WT) son el tiempo que una tecla está presionada y el tiempo interteclas respectivamente. Si se presiona una tecla dos veces, generará un solo evento HT, que es la media de ambos tiempos; lo mismo ocurrirá cuando ocurra un evento WT entre el mismo par de teclas.

Combinando los valores compartidos entre dos muestras nos permite formar un vector de números y calcular su distancia implica encontrar una medida de cuán distintos son los patrones de ambas muestras. Los vectores pueden ser contrastados con varias métricas, como por ejemplo Manhattan (\mathcal{L}^1) y Euclídea (\mathcal{L}^2), aunque la tesis presenta otras varias.

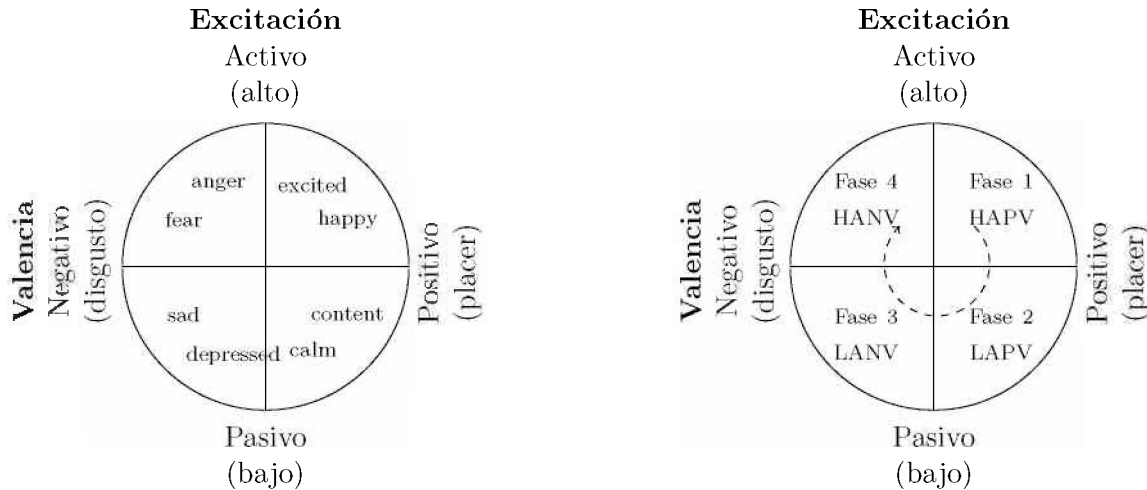
3.2. Modelos emocionales

Se utiliza el *Circumplex of Affectivity* de James Russell [16, 17], también conocido como *circumplejo* o modelo de Excitación-Valencia. Este modelo es bidimensional y cada eje es excitación (relajado *vs.* exaltado) y valencia (placer *vs.* disgusto). La Figura 2a ejemplifica la ubicación de 8 emociones utilizando este modelo [18]. La Figura 2b muestra las fases (con sus nombres en inglés) empleadas en la experimentación de esta tesis y como las personas fueron estimuladas: fase 1, high arousal positive valence (HAPV); fase 2, low arousal positive valence (LAPV); fase 3, low arousal negative valence (LANV); y fase 4, high arousal negative valence (HANV).

3.3. Electroencefalografía (EEG) como interfaz cerebro-máquina (BCI)

Una *interfaz cerebro-computadora* (BCI) permite a una persona controlar o comunicarse con el mundo externo sin la necesidad de utilizar las rutas de salida estándar del cerebro, los nervios o los músculos. Los sistemas BCI proporcionan una nueva opción de comunicación y control esencial para las personas con discapacidad que deterioran los canales normales de comunicación y control. También podrían proporcionar a las personas sin discapacidades un canal de control suplementario o un canal de control en circunstancias especiales [19-21].

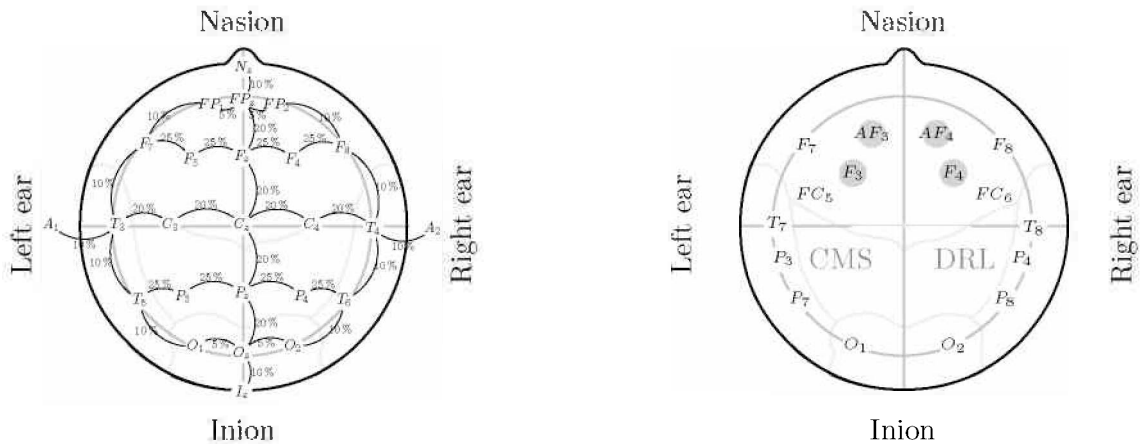
La experimentación de esta tesis utiliza una de las formas posibles de leer la actividad bioeléctrica del cerebro (EEG): los headsets *Emotiv EPOC*, un dispositivo de bajo costo compuesto por 14 electrodos más 2 electrodos de referencia (también utilizados para reducir el ruido), construidos para poder colocarlos en la cabeza. Los electrodos se disponen en el cuero cabelludo en posiciones ya especificadas según el *Sistema internacional 10-20* (ver Figura 3a) [22]. En este sistema, las distancias entre los electrodos suelen ser del 10% o 20% del tamaño de la cabeza (de ahí el nombre 10-20). La configuración del electrodo usando EPOC incluye electrodos intermedios (ver Figura 3b).



(a) Ubicación de ocho emociones. Se conserva el nombre de las mismas en el inglés, idioma original.

(b) Orden de las fases de adquisición de la experimentación.

Figura 2: Modelo Excitación-Valencia.



(a) Posicionamiento de electrodos en el sistema internacional 10-20.

(b) Ubicación de los electrodos del Emotiv EPOC. AF y F, 3 y 4 están en la zona de emociones.

Figura 3: Ubicación de los electrodos EEG.

3.4. Adquisición

La Figura 2b de la Sec. 3.2 describe el camino por el cual el participante es guiado a través del circulejo de Russell. Para esto, el individuo fue estimulado con videos e imágenes obtenidas de IAPS —de los cuales se conocen sus valores aproximados en el circulejo gracias al trabajo de Lang et al. [23]— y luego, el usuario ingresó textos libres, capturando sus patrones de keystroke dynamics. A lo largo del proceso, se midieron las ondas EEG. La Figura 4 describe la arquitectura del experimento.

OpenViBE se utilizó para recibir y almacenar, filtrar y procesar posteriormente las muestras en bruto. Se puede usar para calcular la posición del estado emocional en el modelo de Excitación-Valencia [24, 6].

3.5. Hipótesis secundarias

Se eligieron 10 oraciones de aproximadamente 50 caracteres cada una, con un total de 533 caracteres. Esto induce a los usuarios a generar un promedio de 739 pulsaciones de teclas (oscilando entre 666 y 823 con una desviación estándar de 37.7) y conjuntos de eventos de dimensión promedio 282 (oscilando entre 220 y 365 con una desviación estándar de 46.3) gracias a la alta repetición de pulsaciones de teclas. Aunque las oraciones se introducen sin clasificar, las cadencias son las mismas

en todas las fases y temas. Solicitar más pulsaciones de teclas pone en peligro todo el experimento a medida que el usuario comienza a aburrirse.

Las preguntas *¿qué métricas se deben utilizar en el análisis de los resultados?* y *¿qué cantidad de keystrokes es necesaria para que las pruebas sean estadísticamente significativas?* fueron resueltas tras analizarse cuatro conjuntos de datos (ver Figura 5): uno propio (LSIA, que contiene 68808 muestras de más de 300 keystrokes) y tres pertenecientes a PROSODY [25] (una base de datos de 4,6 millones de keystrokes tipeados por personas con contenido de opinión: revisión de restaurantes (REVIEW), matrimonio gay (GAY) y control de armas (GUN); diseñados para producir diferentes estados emocionales en los participantes, aunque no se han adquirido mediciones biométricas).

Las mejores métricas resultaron ser las Métricas R [3] y A [3], Manhattan (\mathcal{L}^1) y Minkowski con $p = 0,4$ ($\mathcal{L}^{0,4}$) [26]; y se decidió por un valor de 350 keystrokes por muestra ya que todos los casos fueron asintóticos no se observan diferencias significativas al superar este umbral [27].

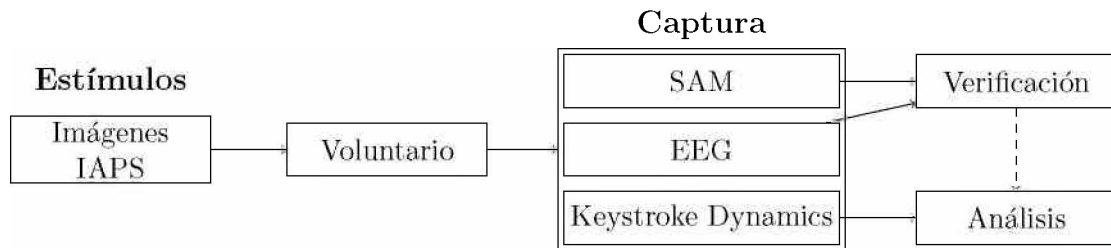


Figura 4: Descripción experimental.

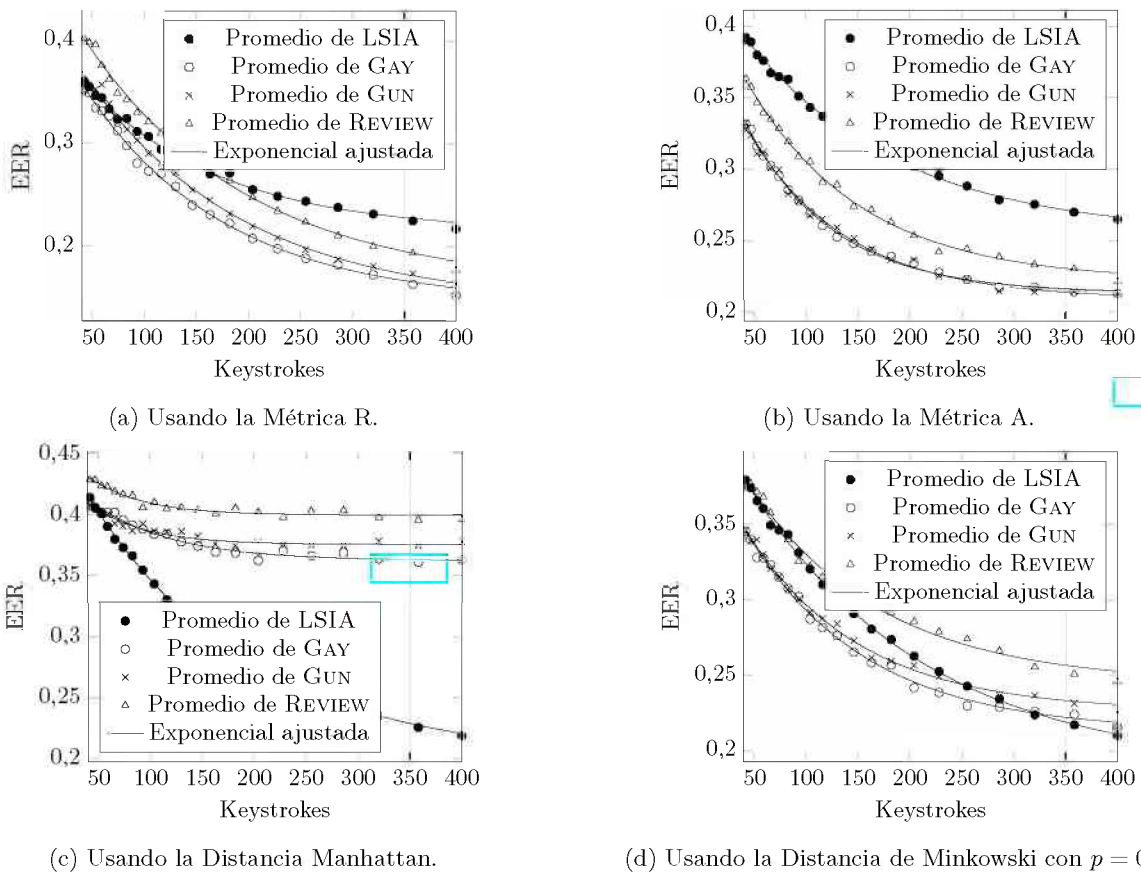


Figura 5: Curvas exponenciales producidas por el EER para pulsaciones de teclado que varían en el rango 40–400 para diferentes conjuntos de datos y métricas.

Cuadro 1: Resultados del experimento para las métricas evaluadas.

		Métrica A		Métrica R		Camberra		Minkowski ($\mathcal{L}^{0,4}$)		Manhattan (\mathcal{L}^1)		Euclídea (\mathcal{L}^2)	
		Usuario	EER	máx/mín	EER	máx/mín	EER	máx/mín	EER	máx/mín	EER	máx/mín	EER
Emociones no combinadas	1	29.1±5	37/23	22.3±2	23/19	31±3	34/27	33.4±3	38/29	35.3±2	38/34	39.3±5	46/31
	2	29.2±4	35/23	30.5±4	35/23	31.2±4	35/25	9.1±6	19/4	11.3±6	19/4	13±3	19/11
	3	30.8±0	31/31	23.7±2	27/20	28.5±2	31/27	32.4±2	35/31	34.7±2	38/31	36.1±2	38/34
	4	26.5±3	31/22	18.9±3	22/15	22.8±3	27/19	27.7±2	31/27	30.6±3	35/27	32.9±3	39/31
	5	34.7±2	38/31	25.8±5	35/20	35.8±2	38/34	33.5±3	37/27	33.5±4	37/27	39.2±3	43/35
	6	31.2±3	37/29	27.8±4	35/23	33.8±3	38/31	37±2	40/35	39.2±3	42/34	41.4±5	50/37
	7	30.9±0	31/31	24.1±3	27/19	28.8±2	31/27	24.6±2	27/23	23.4±4	27/17	32.4±6	42/23
	8	29.7±4	35/23	20.9±3	26/17	29.5±3	31/23	32.3±2	35/31	33.8±4	38/27	39±3	42/35
	9	31.5±1	34/31	25.1±2	27/23	33.8±3	37/29	28.5±2	31/27	31.5±4	35/23	38.9±5	46/31
	10	30.8±0	31/31	26.4±3	31/23	32.6±3	38/29	27.3±3	31/23	28.8±2	31/27	31.9±4	37/27
	11	34.1±1	35/33	32.5±8	38/19	35.7±2	38/35	27.7±4	35/23	31.8±4	39/27	36.2±2	38/33
	12	30.3±4	38/27	21.9±4	29/19	26.5±3	31/23	30±6	35/23	31.8±3	35/27	36.7±4	42/31
Total		30.7	38/22	25.0	38/15	30.8	38/19	28.6	40/4	30.5	42/4	34.7	50/11
SWT	p -value	4.302E-01		9.722E-01		7.190E-01		8.938E-03		1.077E-02		8.870E-04	
	Statistic H_0	0.9345		0.9776		0.9556		0.7982		0.8053		0.7023	
		No Rechazada		No Rechazada		No Rechazada		Rechazada		Rechazada		Rechazada	
		Métrica A		Métrica R		Camberra		Minkowski ($\mathcal{L}^{0,4}$)		Manhattan (\mathcal{L}^1)		Euclídea (\mathcal{L}^2)	
		Usuario	EER	máx/mín	EER	máx/mín	EER	máx/mín	EER	máx/mín	EER	máx/mín	EER
Emociones combinadas	1	28.5±5	37/22	18.6±1	20/17	27.3±2	31/25	32.6±2	35/29	34.8±3	40/31	38.7±3	43/35
	2	28±3	31/26	30.3±3	34/26	32±2	34/29	8.1±3	15/6	9.7±4	14/3	11.9±3	17/9
	3	29.4±1	31/29	22.3±1	23/20	26.3±2	29/23	30.9±2	34/29	34.2±3	40/31	38.2±2	40/34
	4	26.5±4	31/20	16.9±2	20/14	21.2±2	23/19	26.3±1	29/26	29±3	33/26	32.6±2	35/29
	5	33.7±1	34/31	23.7±5	31/20	33±1	34/31	30.7±3	33/26	33.9±3	38/29	39.3±2	43/37
	6	31.4±4	37/25	26.7±3	31/23	32.8±3	37/29	34.3±2	37/31	38.4±2	40/34	41.7±4	48/37
	7	29.1±1	31/29	23.3±2	26/20	27.7±2	31/26	23.6±2	26/21	22.5±3	26/17	31.8±5	40/25
	8	27.6±3	31/23	17.8±3	23/14	27.4±3	31/23	30.9±3	37/29	33.5±3	37/29	38.2±3	44/34
	9	29.1±1	31/29	22.9±2	26/20	31.6±3	35/26	28±2	31/26	30.3±4	34/23	36.1±4	43/29
	10	28.6±0	29/29	26.3±3	31/23	30.5±2	34/29	26.4±2	29/23	28.3±1	29/27	33.2±3	38/29
	11	34.4±2	37/32	32.8±7	40/23	35.4±1	37/34	27±3	31/23	30.9±4	37/26	36.5±2	38/33
	12	29.2±5	37/23	18.2±2	23/17	24.8±5	33/20	28.8±4	33/23	32±3	37/29	36.6±3	42/31
Total		29.6	37/20	23.3	40/14	29.2	37/19	27.3	37/6	29.8	40/3	34.6	48/9
SWT	p -value	6.20E-02		4.89E-01		8.32E-01		3.14E-03		1.16E-02		6.62E-04	
	Estad. H_0	0.8682		0.9393		0.9635		0.7567		0.8081		0.6890	
		No Rechazada		No Rechazada		No Rechazada		Rechazada		Rechazada		Rechazada	
PrT	p -value	1.658E-03		1.384E-03		5.779E-04		2.526E-03		8.590E-03		5.563E-01	
	Estad. H_0	4.13		4.24		4.77		78		73		47	
		Rechazada		Rechazada		Rechazada		Rechazada		Rechazada		Not Rechazada	
$\Delta\mu$ [%]		11		17		16		13		7		2	

4. Resultados experimentales

El cuadro 1 muestra los resultados obtenidos para seis métricas diferentes, donde el p -value muestra la probabilidad de que la hipótesis sea rechazada. Todos los valores de equal error rate (EER, el punto donde el error de aciertos es igual al error de rechazos) [26] se expresan en % a menos que se especifique lo contrario. $\Delta\mu$ es el desvío obtenido entre dos emociones distintas: la diferencia entre los errores de evaluar muestras con emociones combinadas (muestras de personas con varias emociones) y no combinadas (muestras de personas con la misma emoción).

El desvío es visible, pero se debe probar que es estadísticamente significativo. Para ello fueron realizados tests de normalidad (Shapiro–Wilk, o SWT) [28]. Las muestras de las métricas A, R y Camberra no rechazaron la hipótesis de normalidad, por lo que fue posible realizar un t-test pareado [29]. Para las demás métricas se utilizó la prueba de los rangos con signo de Wilcoxon (conocida en inglés como *Wilcoxon signed rank test* o SRT) [30] que es válida para estos casos.

En las primeras cinco métricas, se observa un cambio cerca de $\Delta\mu = 1\%$. Para la norma euclídea, en cambio, se ve un valor de $\Delta\mu = 2\%$.

5. Conclusiones

Se examinaron y revisaron las métricas de clasificación que prevalecen actualmente. Además, se contribuye a la verificación de los algoritmos basados en distancias: la métrica A y su valor empírico del umbral $t = 1,25$ encontrado por Gunetti y Picardi [3]. Adicionalmente, se propone el uso de la métrica de Minkowski con $p = 0,4$, y se muestra que Minkowski se comporta de manera adecuada frente a un amplio conjunto de datos tomados de un entorno de producción. Tras haber recopilado información en un entorno hostil y no idealizado, la distancia de Minkowski puede disminuir el EER medio de 21,9% en la métrica A o 17,4% en la métrica R a una tasa media de 15,6%, lo cual es una mejora sustancial. Esto es importante por su aplicabilidad en el área de la seguridad, exponiendo a posibles intrusos. Además, se hace una diferenciación de los tres tipos de agregación en las métricas de Minkowski especificando cuándo dividir por el número de eventos; en la literatura revisada, la mayoría de los artículos y tesis suelen pasar por alto esta aclaración.

Además, se confirmó empíricamente en dos set de datos, que para los métodos considerados basados en distancias —A, R, Manhattan, Minkowski y euclídea—, el EER asintótico más bajo se alcanza, en promedio, de manera exponencialmente decreciente desde el EER máximo en función del número de keystrokes de la sesión. El ajuste demostró ser preciso, con un error cuadrático medio inferior al 3,5% en el subconjunto de ajuste más desfavorable.

Con respecto a la pregunta hipótesis central, como prueba preliminar, el experimento, predice resultados futuros prometedores en conjuntos de datos más grandes. El EER entre una muestra que incluye un estado emocional varió alrededor del 1% en cinco métricas. De esto, se puede concluir que hay indicios de que los estados emocionales influyen en la calidad de la clasificación de los individuos y que, si un patrón de la persona capturado durante ese estado emocional no está disponible, la precisión del experimento puede empeorar. La métrica euclídea, por otro lado, fue notablemente resistente al cambio emocional y es digna de más investigación, especialmente con muestras más grandes.

Para el desarrollo de la tesis y su adecuada experimentación fue necesaria la realización de un framework de postprocesamiento de datos biométricos en Python compatible con hojas de notas de Jupyter y, para cuando se requiere procesar datos de manera intensiva, en el contexto de esta tesis se portó el framework al lenguaje C y se agregó soporte de paralelización. Para la captura de keystrokes fueron diseñados cuatro *keystroke loggers* (keylogger que además registra los tiempos de las pulsaciones), los primero tres, desarrollados en ANSI C: uno para Mac OS X, uno para Microsoft Windows y uno para GNU/Linux. El cuarto, desarrollado en JavaScript, es capaz de adquirir cadencia de tecleo en un entorno web. Los cuatro capturadores son compatibles entre sí produciendo información en un formato estandarizado y creado en el contexto de esta tesis. Este formato resultó de gran aporte por permitir compatibilidad con los frameworks de postprocesamiento de los datos.

5.1. Contexto

La tesis para el doctorado en Ciencias Informáticas de la Facultad de Informática de la UNLP se radicó en el Laboratorio de Sistemas de Información Avanzados (LSIA) de la Facultad de Ingeniería

de la UBA, fue desarrollada en el marco del proyecto UBACyT 20020130200140BA y actualmente se extiende en el PDE-44-2019 realizándose transferencia tecnológica a empresa de comercio electrónico más grande de la región. El proyecto se desarrolló en colaboración con el ISIER (FICCTE-SECYT UM) que asistió en asuntos de BCI, bajo el PID 01-001/12/14. La dirección estuvo a cargo de Jorge Ierache (LSIA, UBA) y la co-dirección de Waldo Hasperué (LIDI, UNLP).

5.2. Producción científica relacionada con la tesis

Durante el desarrollo de esta tesis se han comunicado resultados parciales a través de diversas publicaciones que a continuación se detallan

Trabajos previos en congresos

- Enrique P. Calot, Juan Manuel Rodríguez, y Jorge S. Ierache. Improving versatility in keystroke dynamic systems. En *Proceedings del XIX Congreso Argentino de Ciencias de la Computación*, número 5606, 2013. ISBN 978-987-23963-1-2. URL <http://hdl.handle.net/10915/32428>
- Enrique P. Calot, Juan Manuel Rodríguez, y Jorge S. Ierache. Improving versatility in keystroke dynamic systems. En Jorge Raúl Finochietto y Patricia Mabel Pesado, editores, *Computer Science & Technology Series. XIX Argentine Congress of Computer Science, Selected papers*, págs. 289–298. Editorial de la Universidad Nacional de La Plata (EDULP), 2014. ISBN 978-987-1985-49-4. URL <http://lsia.fi.uba.ar/papers/calot14b.pdf>
- Enrique P. Calot, Francisco Pirra, Juan Manuel Rodríguez, Gustavo Pereira, Juan Iribarren, y Jorge S. Ierache. Métodos adaptativos de educación de dinámica de tecleo centrado en el contexto emocional de un individuo aplicando interfaz cerebro computadora. En *XVI Workshop de Investigadores en Ciencias de la Computación*, mar. 2014. ISBN 978-950-34-1084-4. URL <http://lsia.fi.uba.ar/papers/calot14.pdf>

Series internacionales y capítulos de libro

- Enrique P. Calot, Jorge S. Ierache, y Waldo Hasperué. Robustness of keystroke dynamics identification algorithms against brain-wave variations associated with emotional variations. En *Intelligent Systems and Applications*, vol. 1037 de *Advances in Intelligent Systems and Computing*. Springer, 2019. ISBN 978-3-030-29515-8. doi: 10.1007/978-3-030-29516-5_15
- Nahuel González, Enrique P. Calot, y Jorge S. Ierache. A replication of two free text keystroke dynamics experiments under harsher conditions. En *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, págs. 1–6, sep. 2016. doi: 10.1109/BIOSIG.2016.7736905
- Nahuel González y Enrique P. Calot. Finite context modeling of keystroke dynamics in free text. En *Biometrics Special Interest Group (BIOSIG), 2015 International Conference of the*, págs. 1–5, sep. 2015. ISBN 978-3-88579-639-8. doi: 10.1109/BIOSIG.2015.7314606

Congresos internacionales

- Enrique P. Calot, Jorge S. Ierache, y Waldo Hasperué. Document typist identification by classification metrics applying keystroke dynamics under unidealised conditions. En *Second International Workshop on Computational Document Forensics, IWCDF at ICDAR 2019, Sydney, Australia, September 22–25, 2019*, págs. 19–24. IEEE, 2019. ISBN 978-1-7281-5054-3. doi: 10.1109/ICDARW.2019.70136
- Enrique P. Calot, Nahuel González, Jorge S. Ierache, y Waldo Hasperué. A convergence model for distance-based keystroke dynamics classification methods, 2019. Comunicado el 7 de febrero de 2019

Congresos nacionales

- Enrique P. Calot y Jorge S. Ierache. Multimodal biometric recording architecture for the exploitation of applications in the context of affective computing. En *Proceedings del XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017)*, número 10529, págs. 1030–1039, 2017. ISBN 978-950-34-1539-9. URL <http://hdl.handle.net/10915/63866>

Además de siete comunicaciones a WICC [\[37–43\]](#).

Esta tesis doctoral

- Enrique P. Calot. *Robustez de las métricas de clasificación de cadencia de tecleo frente a variaciones emocionales*. Tesis doctoral, Universidad Nacional de La Plata, 2019. URL <http://sedici.unlp.edu.ar/handle/10915/76652>

5.3. Futuros trabajos

Como futuras líneas de investigación, se requiere un trabajo adicional sobre la correlación entre los patrones emocionales y las ondas cerebrales. Además, experimentar con un mayor volumen de personas fortalecería los resultados registrados por este trabajo. La diferencia en el comportamiento de la norma euclidiana merece un estudio más profundo, al igual que las distancias de Minkowski con valores de p cercanos a dos.

La incorporación de otros registros biométricos como la frecuencia cardíaca (HR) o el electrocardiograma (ECG) aplicando el EEG como un marcador robusto inicial se considerará como una futura línea de investigación. El conjunto de datos EEG actuará como ya clasificado; por lo tanto, para el análisis de recursos humanos, se explorará la predicción del conjunto de datos con audio y video (como marcadores del estado emocional) utilizando clasificadores previamente entrenados.

Bibliografía

- [1] R. Stockton Gaines, William Lisowski, S. James Press, y Norman Shapiro. Authentication by keystroke timing: Some preliminary results. Technical report, DTIC Document, 1980.
- [2] Despina Polemi. Biometric techniques: review and evaluation of biometric techniques for identification and authentication, including an appraisal of the areas where they are most applicable. *Reported prepared for the European Commission DG XIII*, 4, 1997.
- [3] Daniele Gunetti y Claudia Picardi. Keystroke analysis of free text. *ACM Transactions on Information and System Security (TISSEC)*, 8(3):312–347, ago. 2005. ISSN 1094-9224. doi: 10.1145/1085126.1085129.
- [4] Enrique P. Calot. *Robustez de las métricas de clasificación de cadencia de tecleo frente a variaciones emocionales*. Tesis doctoral, Universidad Nacional de La Plata, 2019. URL <http://sedici.unlp.edu.ar/handle/10915/76652>.
- [5] Lin Shu, Jinyan Xie, Mingyue Yang, Ziyi Li, Zhenqi Li, Dan Liao, Xiangmin Xu, y Xinyi Yang. A review of emotion recognition using physiological signals. *Sensors*, 18(7), 2018. ISSN 1424-8220. doi: 10.3390/s18072074.
- [6] Danny Oude Bos. EEG-based emotion recognition-The Influence of Visual and Auditory Stimuli. *Capita Selecta (MSc course)*, 2006.
- [7] Panagiotis C. Petrantonakis y Leontios J. Hadjileontiadis. Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis. *IEEE Transactions on Affective Computing*, 1(2):81–97, jul. 2010. ISSN 1949-3045. doi: 10.1109/t-affc.2010.7.
- [8] Gyanendra K. Verma y Uma Shanker Tiwary. Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals. *NeuroImage*, 102: 162–172, nov. 2014. doi: 10.1016/j.neuroimage.2013.11.007.
- [9] Dongmin Shin, Dongil Shin, y Dongkyoo Shin. Development of emotion recognition interface using complex EEG/ECG bio-signal for interactive contents. *Multimedia Tools and Applications*, 76(9):11449–11470, dic. 2016. ISSN 1380-7501. doi: 10.1007/s11042-016-4203-7.
- [10] Wei-Long Zheng, Jia-Yi Zhu, y Bao-Liang Lu. Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, 2017. ISSN 1949-3045. doi: 10.1109/taffc.2017.2712143.
- [11] Anas Samara, Maria Luiza Recena Menezes, y Leo Galway. Feature extraction for emotion recognition and modelling using neurophysiological data. En *2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on CyberSpace and Security (IUCC-CSS)*. Institute of Electrical and Electronics Engineers (IEEE), dic. 2016. ISBN 978-1-5090-5566-1. doi: 10.1109/iucc-css.2016.027.

- [12] Xiang Li, Dawei Song, Peng Zhang, Guangliang Yu, Yuexian Hou, y Bin Hu. Emotion recognition from multi-channel EEG data through convolutional recurrent neural network. En *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Institute of Electrical and Electronics Engineers (IEEE), dic. 2016. ISBN 978-1-5090-1611-2. doi: 10.1109/bibm.2016.7822545.
- [13] Jianhai Zhang, Ming Chen, Sanqing Hu, Yu Cao, y Robert Kozma. PNN for EEG-based emotion recognition. En *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. Institute of Electrical and Electronics Engineers (IEEE), oct. 2016. ISBN 978-1-5090-1897-0. doi: 10.1109/smc.2016.7844584.
- [14] Hernan F. Garcia, Mauricio A. Alvarez, y Alvaro A. Orozco. Gaussian process dynamical models for multimodal affect recognition. En *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Institute of Electrical and Electronics Engineers (IEEE), ago. 2016. ISBN 978-1-4577-0220-4. doi: 10.1109/embc.2016.7590834.
- [15] Jingxin Liu, Hongying Meng, Asoke Nandi, y Maozhen Li. Emotion detection from eeg recordings. En *12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, págs. 2173–2178, 2016. URL <http://bura.brunel.ac.uk/handle/2438/13138>.
- [16] James A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980. ISSN 0022-3514. doi: 10.1037/h0077714.
- [17] Jonathan Posner, James A. Russell, y Bradley S. Peterson. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, 17(03), sep. 2005. doi: 10.1017/s0954579405050340.
- [18] Hatice Gunes, Björn Schuller, Maja Pantic, y Roddy Cowie. Emotion representation, analysis and synthesis in continuous space: A survey. En *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, págs. 827–834. IEEE, mar. 2011. doi: 10.1109/FG.2011.5771357.
- [19] Jonathan R. Wolpaw, Dennis J. McFarland, Gregory W. Neat, y Catherine A. Forneris. An EEG-based brain-computer interface for cursor control. *Electroencephalography and clinical neurophysiology*, 78(3):252–259, mar. 1991. ISSN 0013-4694. URL <http://view.ncbi.nlm.nih.gov/pubmed/1707798>.
- [20] Jonathan R. Wolpaw, Niels Birbaumer, Dennis J. McFarland, Gert Pfurtscheller, y Theresa M. Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002. doi: 10.1016/S1388-2457(02)00057-3.
- [21] Jonathan R. Wolpaw y Dennis J. McFarland. Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. En *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, págs. 17849–17854. National Academy of Sciences, dic. 2004. doi: 10.1073/pnas.0403504101.
- [22] Erim Yurci. Emotion detection from eeg signals: Correlating cerebral cortex activity with music evoked emotion. *Universitat Pompeu Fabra, Barcelona*, 2014.
- [23] Peter J. Lang, Margaret M. Bradley, y Bruce N. Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical report A-8*, 2008.
- [24] Lisa A. Feldman. Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of personality and social psychology*, 69(1):153, 1995.
- [25] Ritwik Banerjee, Song Feng, Jun Seok Kang, y Yejin Choi. Keystroke patterns as prosody in digital writings: A case study with deceptive reviews and essays. En *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, págs. 1469–1473, Doha, Qatar, oct. 2014. Association for Computational Linguistics.
- [26] Enrique P. Calot, Jorge S. Ierache, y Waldo Hasperué. Document typist identification by classification metrics applying keystroke dynamics under unidealised conditions. En *Second International Workshop on Computational Document Forensics, IWCDF at ICDAR 2019, Sydney, Australia, September 22–25, 2019*, págs. 19–24. IEEE, 2019. ISBN 978-1-7281-5054-3. doi: 10.1109/ICDARW.2019.70136.
- [27] Enrique P. Calot, Nahuel González, Jorge S. Ierache, y Waldo Hasperué. A convergence model for distance-based keystroke dynamics classification methods, 2019. Comunicado el 7 de febrero de 2019.
- [28] Samuel Sanford Shapiro y Martin B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965. doi: 10.1093/biomet/52.3-4.591.

- [29] Cyril H. Goulden. Methods of statistical analysis. chapter 4, págs. 50–55. John Wiley & Sons, Inc., New York, 2 edition, mayo 1956.
- [30] Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, 1(6):80–83, 1945. doi: 10.2307/3001968.
- [31] Enrique P. Calot, Juan Manuel Rodríguez, y Jorge S. Ierache. Improving versatility in keystroke dynamic systems. En *Proceedings del XIX Congreso Argentino de Ciencias de la Computación*, número 5606, 2013. ISBN 978-987-23963-1-2. URL <http://hdl.handle.net/10915/32428>.
- [32] Enrique P. Calot, Juan Manuel Rodríguez, y Jorge S. Ierache. Improving versatility in keystroke dynamic systems. En Jorge Raúl Finochietto y Patricia Mabel Pesado, editores, *Computer Science & Technology Series. XIX Argentine Congress of Computer Science, Selected papers*, págs. 289–298. Editorial de la Universidad Nacional de La Plata (EDULP), 2014. ISBN 978-987-1985-49-4. URL <http://lsia.fi.uba.ar/papers/calot14b.pdf>.
- [33] Enrique P. Calot, Francisco Pirra, Juan Manuel Rodríguez, Gustavo Pereira, Juan Iribarren, y Jorge S. Ierache. Métodos adaptativos de educación de dinámica de tecleo centrado en el contexto emocional de un individuo aplicando interfaz cerebro computadora. En *XVI Workshop de Investigadores en Ciencias de la Computación*, mar. 2014. ISBN 978-950-34-1084-4. URL <http://lsia.fi.uba.ar/papers/calot14.pdf>.
- [34] Enrique P. Calot, Jorge S. Ierache, y Waldo Hasperué. Robustness of keystroke dynamics identification algorithms against brain-wave variations associated with emotional variations. En *Intelligent Systems and Applications*, vol. 1037 de *Advances in Intelligent Systems and Computing*. Springer, 2019. ISBN 978-3-030-29515-8. doi: 10.1007/978-3-030-29516-5_15.
- [35] Nahuel González, Enrique P. Calot, y Jorge S. Ierache. A replication of two free text keystroke dynamics experiments under harsher conditions. En *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, págs. 1–6, sep. 2016. doi: 10.1109/BIOSIG.2016.7736905.
- [36] Nahuel González y Enrique P. Calot. Finite context modeling of keystroke dynamics in free text. En *Biometrics Special Interest Group (BIOSIG), 2015 International Conference of the*, págs. 1–5, sep. 2015. ISBN 978-3-88579-639-8. doi: 10.1109/BIOSIG.2015.7314606.
- [37] Enrique P. Calot y Jorge S. Ierache. Multimodal biometric recording architecture for the exploitation of applications in the context of affective computing. En *Proceedings del XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017)*, número 10529, págs. 1030–1039, 2017. ISBN 978-950-34-1539-9. URL <http://hdl.handle.net/10915/63866>.
- [38] Germán M. Concilio, Enrique P. Calot, Jorge S. Ierache, y Hernán D. Merlino. Application of Keystroke Dynamics Modelling Techniques to Strengthen the User Identification in the Context of E-commerce. En Armando Eduardo De Giusti, editor, *VII Workshop de Seguridad Informática (CACIC 2018, Tandil, Buenos Aires, Argentina)*, 2018. In press.
- [39] Gustavo Pereira, Facundo Nervo, Juan Iribarren, Enrique P. Calot, Jorge S. Ierache, y Norberto Mazza. Líneas de investigación aplicada al empleo de interfaces cerebro-maquina. En *Proceedings del XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015)*, pág. 5, jun. 2015. URL <http://hdl.handle.net/10915/46146>.
- [40] Nahuel González, Enrique P. Calot, Ariel Liguori, Francisco Pirra, Juan Manuel Rodríguez, Gustavo Pereira, Facundo Nervo, Juan Iribarren, y Jorge S. Ierache. Educación de dinámica de tecleo centrado en el contexto emocional de un individuo. En *Proceedings del XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015)*, pág. 5, jun. 2015. URL <http://hdl.handle.net/10915/46260>.
- [41] Enrique P. Calot, Ezequiel L. Aceto, Juan Manuel Rodríguez, Ariel Liguori, María Alejandra Ochoa, Hernán D. Merlino, Enrique Fernández, Nahuel González, Francisco Pirra, y Jorge S. Ierache. Líneas de investigación del laboratorio de sistemas de información avanzados. En *Proceedings del XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015)*, pág. 4, jun. 2015. ISBN 978-987-633-134-0. URL <http://hdl.handle.net/10915/46107>.
- [42] Enrique P. Calot, Federico M. Rossi, Nahuel Gonzalez, Waldo Hasperué, y Jorge S. Ierache. Avances en educación de dinámica de tecleo y el contexto emocional de un individuo aplicando interfaz cerebro computadora. En *XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina)*, págs. 872–876, jun. 2016. ISBN 978-950-698-377-2. URL <http://hdl.handle.net/10915/53247>.
- [43] Federico M. Rossi, Enrique P. Calot, y Jorge S. Ierache. Educación emocional de un individuo en contextos multimodales en computación afectiva. En *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, Buenos Aires, Argentina)*, jul. 2017. ISBN 978-987-42-5143-5.