



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

MEDICIÓN DE CARGAS COGNITIVAS DURANTE ACTIVIDADES DE  
INTERACCIÓN HUMANO COMPUTADOR EN AMBIENTE MÓVIL  
USANDO SENSORES PSICO-FISIOLÓGICOS

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL  
INDUSTRIAL

HERNÁN FELIPE LIRA LÓPEZ

PROFESOR GUÍA:

SR. ÁNGEL JIMÉNEZ MOLINA

MIEMBROS DE LA COMISIÓN:

SR. IGNACIO CALISTO LEIVA

SR. ALBERTO CABEZAS BULLEMORE

Este trabajo ha sido financiado por el proyecto Fondecyt 11130252.

SANTIAGO DE CHILE

DICIEMBRE 2015

# Resumen Ejecutivo

RESUMEN DE LA MEMORIA  
PARA OPTAR AL TITULO DE  
INGENIERO CIVIL INDUSTRIAL  
POR: HERNÁN LIRA LÓPEZ  
FECHA: 07/12/2015  
PROF. GUÍA: SR. ÁNGEL JIMÉNEZ

El presente trabajo tiene como objetivo la medición cuantitativa de cargas cognitivas durante ciertas actividades en ambientes móviles mediante sensores psico-fisiológicos. Los sensores psico-fisiológicos entregan una medida de alguna característica fisiológica con lo que es posible determinar factores psicológicos como la carga cognitiva, el estrés, la fatiga, entre otros. El método de medición más utilizado es un cuestionario elaborado por NASA, que se aplica luego de realizadas las actividades por lo que el resultado final es subjetivo. El enfoque de este proyecto, en cambio, es medir en base a los datos objetivos que entregan las señales de los sensores. Para esto se utiliza un eyetracker, electrocardiograma, pulsoxímetro y temperatura.

Este tema de memoria está enmarcado en el proyecto Fondecyt “*A Cognitive Resource-Aware Mobile Service Framework to Support Human-Computer-Interactions in Ubiquitous Computing Environments*” liderado por el Profesor Ángel Jiménez. El problema tipo de estudio es cuando un usuario debe dividir su atención entre una actividad primaria física – por ejemplo, caminar, hablar – e interacciones del usuario con el computador u otro dispositivo – por ejemplo, leer o escribir desde el Smartphone.

Para la medición de carga cognitiva se desarrolla un diseño experimental que consiste en realizar tareas con un Smartphone en 3 escenarios con distinto nivel de demanda cognitiva, estos son, sin estímulos externos, en conversación y caminando. Los resultados esperados son principalmente dos: comprobar (o no) que las señales pueden discriminar los distintos niveles de demanda cognitiva y obtener clasificadores de esos niveles. Se realizó el experimento con 20 participantes validando a su vez el diseño con la aplicación del test NASA que afirmó que el diseño es coherente a lo que se busca.

Se aplica el proceso Knowledge Discovery in Databases con el fin de realizar todos los pasos adecuados para una correcta aplicación de un algoritmo de minería de datos.

Los resultados indican que todos los sensores son capaces de discriminar estadísticamente las actividades de distinto nivel de demanda cognitiva planteadas en este estudio. Se aplicó un test de hipótesis t y análisis de varianza de medidas repetidas para obtener lo anterior.

Se aplican también los algoritmos de Support Vector Machine, Naive Bayes y Regresión Logística. Los resultados indican que el modelo Naive Bayes fue el que mejor resultados obtuvo. En cuanto a los sensores, al modelar por cada tarea incluyendo a todos los participantes el sensor de BPM junto con el sensor de SO2 son los que mejor predicen. Incluso cuando se combinan su poder de predicción crece llegando al 81.42% de accuracy en algunas tareas.

*A mis padres por su amor y confianza en mí*

# Agradecimientos

En primer lugar quiero agradecer a mis padres. A mi Mamá, Verónica, por ser el mayor ejemplo de amor incondicional y entrega absoluta. Por siempre estar conmigo y levantarme en los momentos difíciles. No estaría finalizando esta etapa sino fuera gracias a ella. A mi Papá, Julio, porque siempre se preocupó de darme las mejores oportunidades y porque mi vida hoy sería muy distinta si él no hubiera sido parte de ella.

Agradecer también a mis tías, por su cariño y apoyo. A la Oga, en especial, por siempre preocuparse por mí y ayudarme cada vez que la necesité.

A Karina, por su amor, apoyo y especialmente por su ayuda durante la evaluación del diseño experimental en este trabajo.

También quiero agradecer a mis abuelos, Mami y Tata, que aunque ya no estén conmigo, cada día están presentes en mis recuerdos. Ellos son unas de las influencias más grandes en mi vida. Sé que les gustaría verme hoy, espero que lo estén haciendo desde algún lugar.

Agradecer a mi Profesor Guía, Ángel Jiménez, por darme la oportunidad y la confianza de trabajar con él y conocer un poco más cerca el mundo de la investigación que siempre fue un anhelo para mí.

También a Manuel Carrasco y José Zorrilla por su apoyo técnico y su disposición para resolver mis dudas durante esta memoria.

Finalmente, quiero agradecer a todas las personas que a lo largo de estos años han sido importantes para mí y me han permitido crecer como persona y formarme como profesional.

Hernán Lira López

# Tabla de Contenido

Resumen Ejecutivo .....	i
Agradecimientos .....	iii
Tabla de Contenido .....	iv
Índice de Tablas.....	vii
Índice de Figuras .....	viii
1. Introducción .....	1
1.1 Antecedentes generales.....	1
1.2 Contexto institucional.....	2
1.3 Descripción del proyecto.....	3
1.3.1 Proyecto .....	3
1.3.2 Oportunidad de investigación .....	3
1.4 Objetivos .....	4
1.4.1 Objetivo General.....	4
1.4.2 Objetivos Específicos .....	4
1.4.3 Hipótesis de investigación .....	5
1.5 Resultados Esperados y Alcances .....	5
1.6 Metodología.....	6
2. Marco Teórico .....	7
2.1 Recursos Cognitivos y Modelo de Wickens .....	7
2.2 Métodos de medición de carga cognitiva .....	12
2.2.1 Métodos subjetivos .....	12
2.2.2 Métodos basados en el desempeño de la tarea .....	12
2.2.3 Medidas Psico-Fisiológicas.....	13
2.3 Conceptualización del término Psico-Fisiología .....	13
2.4 Medidas fisiológicas en relación al sistema nervioso [30].....	14
2.4.1 El sistema nervioso autónomo (ANS) .....	15
2.4.2 El sistema nervioso Simpático .....	16

2.4.3	El sistema nervioso Parasimpático .....	16
2.5	El sistema Visual.....	16
2.5.1	Pupila.....	17
2.5.2	Medición .....	18
2.5.3	La Pupila y la carga cognitiva .....	18
2.6	El Sistema Cardiovascular .....	19
2.6.1	El corazón.....	19
2.6.2	El Ciclo Cardíaco .....	20
2.6.3	Medición .....	21
2.6.4	Sistema cardiovascular y carga cognitiva .....	22
2.7	Trabajos anteriores.....	23
2.8	Proceso Knowledge Discovery in Databases .....	26
2.8.1	Técnicas de minería de datos .....	28
2.8.2	Algoritmos de minería de datos .....	29
3.	Experimento .....	33
3.1	Instrumentación .....	33
3.1.1	Dispositivo de captura de datos .....	33
3.1.2	Sensores Psicofisiológicos .....	36
3.2	Grupo Experimental .....	39
3.3	Diseño del experimento .....	39
3.3.1	Consideraciones iniciales .....	39
3.3.2	Diseño.....	40
3.3.3	Escenarios .....	44
3.3.4	Consideraciones para la ejecución del experimento .....	48
3.4	Implementación.....	49
3.4.1	Grupo experimental.....	49
3.4.2	Procedimiento Experimental.....	50
3.4.3	Formato de Resultados.....	51
4.	Validación del diseño experimental .....	53
4.1	Validación previa.....	53

4.2	Validación posterior .....	54
5.	Análisis y resultados .....	56
5.1	Definición de observaciones .....	56
5.1.1	Etiquetado de Subtareas .....	57
5.2	Procesamiento de datos .....	58
5.2.1	Dilatación pupilar .....	58
5.2.2	Electrocardiograma (ECG) .....	60
5.2.3	Pulsoxímetro y temperatura .....	62
5.3	Análisis Estadístico .....	63
5.3.1	Análisis exploratorio de los datos .....	63
5.3.2	Test de Hipótesis .....	66
5.4	Minería de datos .....	69
5.4.1	Features .....	70
5.4.2	Selección de features .....	70
5.4.3	Modelamiento .....	71
5.5	Comparación con NASA TLX .....	80
5.6	Análisis por sensor .....	80
6.	Discusión .....	82
7.	Conclusiones y trabajo futuro .....	86
	Bibliografía .....	87
	Anexos .....	90
	A Detalle resultados SVM por sensor y subtarea .....	90
	B Detalle resultados Naive Bayes por sensor y subtarea .....	93

# Índice de Tablas

1	Recursos cognitivos	8
2	Control de las respuestas fisiológicas por parte del sistema nervioso	15
3	Medidas fisiológicas asociadas a procesos psicológicos	23
4	Indica recursos aplicados en Oficina + Actividades HCI	45
5	Recursos aplicados a Conversar + Actividades HCI	46
6	Recursos aplicados a Caminar + Actividades HCI	47
7	Detalle grupo experimental por género, edad y ocupación	49
8	Se indica el valor de carga cognitiva según medida entregada por NASA TLX para los tres escenarios del experimento	53
9	Resultados test t	54
10	Resultados test t	55
11	Resultados test t	55
12	Media y varianza primeros 10 participantes	64
13	Varianza entre participantes por señal y dificultad	65
14	Features iniciales por señal	70
15	Features finales por señal	71
16	Resultados modelo SVM	73
17	Accuracy por sensor	74
18	Accuracy por pareja de sensores	74
19	Resultados modelamiento con Naive Bayes.	75
20	Accuracy por sensor	76
21	Accuracy por pareja de sensores	76
22	Resultados modelo Regresión Logística	78
23	Los mejores 3 sensores para cada participante usando SVM	79
24	Los mejores 3 sensores para cada participante usando Naive Bayes	79
25	Accuracy promedio por participante y por total de tareas	83



# Índice de Figuras

1	Estructura de modelo de múltiples recursos	9
2	Sistema de procesamiento humano	11
3	Estructura del ojo humano	17
4	Estructura del corazón	20
5	Ciclo cardíaco ideal	21
6	Fases del proceso KDD	26
7	Técnicas de minería de datos	28
8	Hiperplano SVM	31
9	Raspberry Pi	34
10	Tarjeta aLaMode	34
11	Shield e-health	35
12	Tarjeta MoPi	35
13	Sensor ECG	36
14	Posición de electrodos ECG en el cuerpo	36
15	Sensor de Temperatura	37
16	Sensor de pulso y saturación de Oxígeno	37
17	Eyetracker by pupil labs	38
18	Imagen de la detección pupilar	38
19	Ejemplo de correo electrónico simple	41
20	Ejemplo de correo electrónico simple	41
21	Capturas de pantalla de app Despegar	42
22	Ejemplo de correo electrónico para app despegar	43
23	Pantalla inicio app Tripadvisor	43
24	Ejemplo correo para app Tripadvisor	43
25	Ejemplo correo electrónico para Crediautos	44
26	Participante utilizando Eyetracker	50
27	Participante realizando actividades	51
28	Gráfico que muestra las diferencias en los valores del test NASA	54
29	Captura de pantalla de video de Eyetracker	57
30	Interpolación de parpadeos	59
31	Filtrado de frecuencias altas	59
32	(A) Señal original; (B) Señal luego de aplicación T. de Fourier	61

# 1. Introducción

En este capítulo se introduce este estudio. Se comienza con los antecedentes generales y el contexto institucional del trabajo para entender el ambiente en el que se desarrolla. Luego se describe el proyecto a nivel general, explicando cómo surge esta oportunidad. Se continúa con el objetivo general del proyecto y los objetivos específicos que lo sustentan, además de la hipótesis de investigación. Se indican los resultados esperados, el alcance del proyecto y finalmente, se detalla la metodología del trabajo, en donde se muestra el procedimiento para el cumplimiento de cada una de sus etapas.

## 1.1 Antecedentes generales

En 1991 Mark Weiser dijo que “las tecnologías más profundas son las que desaparecen. Se tejen a sí mismos en la fábrica de la vida hasta hacerse indistinguible de esta” [1]. Con estas palabras comenzó con el desarrollo de la computación ubicua, rama de las tecnologías de información que tiene entre sus objetivos la disponibilidad no intrusiva de los sistemas y servicios computacionales en cualquier lugar y que sus procesos afecten lo menos posible al usuario, invisibilizando gran parte de las decisiones [2].

Esto toma gran relevancia si consideramos que actualmente las personas están expuestas a múltiples interacciones con distintos dispositivos computacionales. Si se quiere enviar o leer un correo electrónico, interactuar con una red social o buscar una dirección en Google Maps el usuario debe desarrollar múltiples tareas, denominadas tareas HCI (interacción humano computador por su sigla en inglés), para el cumplimiento de su objetivo. Para la realización de las tareas HCI se deben utilizar distintos recursos cognitivos<sup>1</sup>, que son los principales activos utilizados por la cognición para pensar, recordar, tomar decisiones, entre otras [7]. Desde la psicología se sabe que la cantidad de recursos cognitivos que se enfocan en determinada tarea HCI es limitada [3] luego, si la tarea HCI es muy demandante cognitivamente, tanto que sobrepasa la capacidad máxima, se produce lo que en psicología se conoce como agotamiento de recursos cognitivos [4] provocando errores en la ejecución de las tareas.

Ante este escenario, el desafío de la computación ubicua es entregar al usuario servicios y, en definitiva, la lógica del negocio, tomando en cuenta los factores ambientales,

---

<sup>1</sup> Atención, percepción, memoria de corto y largo plazo, control motor, entre otros.

temporales, físicos, sociales, psicológicos que entregan información sobre la situación en que está inmersa una persona. Así, al conocer el contexto de cada situación se puede entregar la información oportuna en el momento indicado con lo que se evita el agotamiento de recursos cognitivos y se interrumpe lo menos posible a la persona. Entonces si se quiere ofrecer un servicio, pero la persona está siendo muy demandada cognitivamente es preferible esperar y ofrecerlo en otra ocasión [5].

Con el fin de obtener mayor y mejor información de la situación en que está inmersa una persona es que en los últimos años se han desarrollado sensores psico-fisiológicos capaces de obtener métricas fisiológicas y asociarlas a distintos factores psicológicos, entre éstos la cognición, pudiendo así obtener la carga cognitiva en tiempo real. A partir de eso podrían existir muchos beneficios y aplicaciones para el desarrollo de sistemas computacionales y nuevas oportunidades de negocio.

## 1.2 Contexto institucional

Este trabajo de título se desarrolla en el marco del proyecto Fondecyt otorgado en el año 2013 titulado *A Cognitive Resource-Aware Mobile Service Framework to Support Human-Computer-Interactions in Ubiquitous Computing Environments* a cargo del profesor del Departamento de Ingeniería Industrial Ángel Jiménez Molina. En este proyecto se pretende desarrollar un mecanismo de ingeniería cognitiva que seleccione, componga y desarrolle funcionalidades, durante el tiempo en que se ejecute, tomando en consideración el contexto situacional y los recursos cognitivos empleados por el usuario de acuerdo a las tareas HCI realizadas [6].

El objetivo general del proyecto Fondecyt es reducir el agotamiento de recursos cognitivos mientras se proveen servicios. Para esto los objetivos específicos que lo sustentan son: 1) Modelar relaciones entre recursos mentales, tareas HCI y actividades; 2) Definir métricas para evaluar agotamiento de recursos cognitivos; 3) Crear algoritmos de Binding<sup>2</sup>; 4) Diseñar un modelo computacional para analizar servicios según demandas cognitivas e interferencias; 5) Desarrollar un framework para dispositivos móviles que englobe todo lo anterior.

---

<sup>2</sup> Algoritmo capaz de seleccionar el servicio idóneo a entregar según el contexto

## 1.3 Descripción del proyecto

### 1.3.1 Proyecto

Se tiene la siguiente situación, una persona realiza dos actividades. La actividad primaria es una actividad física como caminar, beber una bebida, esperar un bus o conversar. La otra actividad es una actividad HCI que realiza con un Smartphone [6]. Al realizar dos actividades simultáneamente la persona debe dividir su atención, con lo que asigna recursos cognitivos a cada una de éstas [7]. Como la cantidad de recursos cognitivos en una persona es limitada, eventualmente se saturará su capacidad con lo que se provocarán errores, stress y se reducirá la capacidad del individuo de tomar decisiones correctas [4].

Lo que pretende esta memoria es medir la carga cognitiva de personas expuestas a situaciones como la descrita en el párrafo anterior. La carga cognitiva se define como el nivel percibido de esfuerzo que se requiere para la realización de una tarea como una medida del trabajo mental producido durante su ejecución [8]. En trabajos futuros se podría nutrir con este nuevo input a los algoritmos de Binding y ofrecer servicios cuando la carga cognitiva sea la idónea. Así, este trabajo está inmerso en el objetivo específico número dos del Fondecyt, el cual busca definir métricas para evaluar el agotamiento de los recursos cognitivos.

### 1.3.2 Oportunidad de investigación

Para desarrollar este trabajo se sigue la metodología realizada por Haapalainen et al. [9] en donde se mide la carga cognitiva mediante sensores psico-fisiológicos que entregan medidas fisiológicas de cada individuo como la dilatación pupilar, pulso cardíaco, presión sanguínea, entre otras. Está bien estudiado en la literatura que estas medidas fisiológicas tienen directa relación con variables psicológicas como la carga cognitiva. Luego, al utilizar estos sensores durante el tipo de actividades descritas en el punto anterior se podría obtener la carga cognitiva utilizada en tiempo real.

La manera en que se ha medido la carga cognitiva en los estudios anteriores es proponiendo tareas HCI con distinto grado de dificultad, esto es, que requieren distintos niveles de demanda cognitiva. Se analizan los datos provenientes desde los sensores psico-fisiológicos concluyendo si es que los sensores son capaces de discriminar esos grados de dificultad o no. Luego, se aplican algoritmos de minería de datos para clasificar los datos

de cada señal según nivel de dificultad. Finalmente, los clasificadores resultantes son una medida de carga cognitiva.

Es importante señalar que los aportes de esta memoria al estudio de la aplicación de los sensores psico-fisiológicos para medir carga cognitiva son tres. Primero, no se ha considerado anteriormente la realización de experimentos en un ambiente móvil, esto es, la medición de carga cognitiva mientras el usuario utiliza una aplicación de un Smartphone por ejemplo. Segundo, no se ha considerado el contexto situacional en que está inmerso el usuario, esto es, si está caminando, sentado o escuchando música mientras realiza una tarea HCI. Tercero, la posibilidad de tener una medición de carga cognitiva en tiempo real y objetiva, a diferencia del método de medición más usado en la actualidad, NASA TLX (task load index) [10], que es un cuestionario que se le realiza al individuo luego de realizado el experimento, por lo tanto es subjetivo y no mide en tiempo real.

En esta memoria los sensores psico-fisiológicos que se utilizan son: Eyetracker, Electrocardiograma, Temperatura y Pulsoxímetro.

## 1.4 Objetivos

### 1.4.1 Objetivo General

Medición cuantitativa de cargas cognitivas durante actividades de interacción humano computador en un ambiente móvil usando sensores psico-fisiológicos para futuras aplicaciones de Computación Ubicua.

### 1.4.2 Objetivos Específicos

- Investigar sobre la teoría cognitiva, sensores psico-fisiológicos y la relación entre medidas fisiológicas con la carga cognitiva
- Diseñar e implementar experimentos fundamentando con la teoría cognitiva de Wickens y el uso de los sensores psico-fisiológicos
- Determinar si los sensores psico-fisiológicos son capaces de discriminar entre actividades con distinto nivel de demanda cognitiva mediante un análisis estadístico
- Evaluar modelos de minería de datos que indiquen el poder de clasificación de cada una de las señales según el nivel de demanda cognitiva comparando los resultados

obtenidos con los resultados de trabajos anteriores y con el cuestionario NASA TLX

- Discutir los resultados y entregar posibles vías de desarrollo del trabajo futuro

### 1.4.3 Hipótesis de investigación

En la literatura existe evidencia de que es posible la medición de cargas cognitivas mediante sensores psico-fisiológicos para actividades estáticas.

La hipótesis de esta memoria es que es posible medir la carga cognitiva para actividades en que el usuario debe dividir su atención en un entorno móvil mediante medidas psico-fisiológicas tomando en cuenta el contexto situacional.

## 1.5 Resultados Esperados y Alcances

Para este trabajo se esperan los siguientes resultados:

- 1) Estado del arte de la medición de carga cognitiva con sensores psico-fisiológicos
- 2) Diseño experimental
- 3) Aceptación o rechazo de la hipótesis de investigación.
  - Para cada sensor psico-fisiológico se debe obtener si es que son capaces de diferenciar a grupos con actividades de distinto nivel de dificultad o no. Esto se realizará aplicando test estadísticos y análisis de varianza según sea el caso.
- 4) Nivel de clasificación del algoritmo de minería de datos utilizado comparándolo con resultados de trabajos anteriores y con cuestionario NASA TLX.

En cuanto a los alcances no se pretende que se incluyan estrategias para mejorar el rendimiento de modelos o la implementación de nuevos modelos que no hayan sido utilizados en trabajos previos.

## 1.6 Metodología

El desarrollo del tema de memoria consta de las siguientes etapas basadas en Haapalainen et al. [9]:

- 1) Estudio del estado del arte: Se revisa la literatura del uso de recursos cognitivos aplicados a la computación ubicua con el fin de tener los fundamentos necesarios para entender el problema. Además se revisa la literatura existente con respecto al uso de sensores psico-fisiológicos y sus fundamentos científicos. Se sistematiza la información y se filtra lo relevante. Además se estudia el proceso KDD en profundidad.
- 2) Estudio de los métodos actuales de medición de cargas cognitivas: Se da énfasis al cuestionario NASA TLX.
- 3) Selección de aplicaciones y herramientas informáticas a utilizar
- 4) Diseño Experimental: Con la información recopilada se procede a diseñar las actividades a evaluar acorde a los recursos cognitivos que se quieran medir. Las actividades se deben construir a modo de tener al menos dos grados de dificultad. Para esto el diseño se basa en el modelo de Múltiples recursos de Wickens [7] en donde se exponen las interferencias producidas por compartir recursos cognitivos entre dos tareas y que al provocar más interferencias aumenta la carga cognitiva asociada.  
Además se determinan los escenarios de los experimentos, la cantidad y grupo experimental.  
Como parte del diseño experimental se realiza un desarrollo experimental “piloto” a un conjunto reducido de personas. Se miden actividades sencillas a modo de entrenamiento con los sensores y de análisis de los experimentos. Esto permite completar el diseño experimental en base a lo visto en la experiencia.
- 5) Realización del experimento: Aplicación de las pruebas creadas, validación de la dificultad de las pruebas y aplicación del cuestionario NASA TLX para la posterior evaluación.
- 6) Análisis de los datos: Aplicación del Proceso KDD para utilizar técnicas efectivas de minería de datos. La idea es entrenar clasificadores que a partir de los datos obtenidos desde los sensores sean capaces de predecir si la dificultad de la tarea es alta o baja con lo que se puede inferir la carga cognitiva.
- 7) Resultados: Tabla de resultados que muestra la efectividad de cada sensor para cada actividad.
- 8) Discusión: Comparación con resultados de trabajos anteriores y con el cuestionario NASA TLX
- 9) Conclusiones

## 2. Marco Teórico

En este capítulo se profundiza en conceptos vistos en el capítulo anterior y en otros que se verán en capítulos siguientes. Se comienza describiendo a los recursos cognitivos y el Modelo de Wickens. Luego se detallan los métodos de medición de carga cognitiva existentes y se entrega una conceptualización del término psico-fisiología. Se describen los sistemas fisiológicos atinentes a esta memoria y se entrega un resumen de trabajos anteriores relacionados con el tema. Se finaliza el capítulo con una descripción del proceso Knowledge Discovery in Databases que se utilizará para construir modelos de evaluación de carga cognitiva.

### 2.1 Recursos Cognitivos y Modelo de Wickens

Actualmente el Smartphone forma parte de la vida de las personas. Es común ver que mientras se realizan actividades como conversar, caminar o esperar por la llegada del metro se esté utilizando paralelamente este dispositivo. La teoría señala que al interactuar con un dispositivo móvil mientras se ejecutan otras tareas, se entra en una competencia por los mismos recursos cognitivos limitados que se necesitan para realizar las labores cotidianas [29].

Para sentar las bases de la teoría cognitiva a la que se hace referencia en esta memoria, se deben señalar los siguientes puntos:

- Los recursos cognitivos se pueden utilizar en paralelo por múltiples tareas,
- Los recursos cognitivos son limitados,
- Las tareas pueden utilizar varios recursos a la vez y diferentes tareas demandan diversos niveles de respuesta cognitiva con lo que es posible priorizar, posponer o terminar tareas que no tienen atención o un grupo de éstas por sobrecarga o agotamiento de recursos cognitivos [29].

Existen variados estímulos psicológicos estudiados en la literatura [30] como estrés, excitación, frustración, fatiga, trabajo mental, entre otros. Esta memoria solo se enfoca en el estudio del trabajo mental o carga cognitiva.

En la tabla 1 se listan los principales recursos cognitivos utilizados cuando se realiza algún tipo de trabajo mental.



#	Recurso cognitivo	Detalle
1	Control Motor	Sincronización, control, timing y finalización de las acciones motoras
	(manos)	Manipulación de objetos físicos
	(piernas)	Posicionamiento y movimiento de sí mismo
2	Sensación	Consumo de estímulos externos
3	Percepción	Organización de la data de los sentidos
4	Atención (visual) (auditiva) (motor)	Buscar, seleccionar e integrar
5	Control central	Control meta-nivel de las operaciones cognitivas – selección, inhibición, actualización, desplazamiento
6	Memoria de trabajo	Retención de información previa
7	Memoria Prospectiva	Control proactivo y sincronización de acciones
8	Memoria Episódica	“Tiempo de viaje” mental
9	Memoria Semántica	Conocimiento adquirido
10	Pensamiento consciente	Manipulación Consciente de los materiales de la memoria – abstracción, inferencia, razonamiento, resolución de problemas

Tabla 1: Recursos cognitivos. Extraída de [29]

Si una persona, por ejemplo, está escribiendo en su Smartphone, necesita usar el Control Motor para accionar sus manos, sensación, percepción, atención visual, control central y memoria de trabajo [29]. Y así, si está navegando en el browser o está esperando por la carga de alguna aplicación se están utilizando distintos tipos de recursos cognitivos a la vez.

Lo anterior se complica si se piensa que una persona además de utilizar su Smartphone, está realizando otra actividad, por ejemplo caminando en una calle congestionada. Para lograr cumplir con esa tarea deberá planear la ruta, caminar mientras

evita colisiones y se preocupa de su seguridad. Todas esas subtareas requieren el uso de control motor, atención, percepción, memoria.

Luego es probable que al realizar ambas actividades al mismo tiempo se produzcan interferencias.

Wickens [7], para estudiar este problema creó el Modelo de recursos múltiples cuya importancia radica en su habilidad para predecir niveles de interferencias en las tareas que se ejecutan al mismo tiempo en donde es posible que exista una sobrecarga del uso de recursos.

Este modelo nace a partir de la evidencia de que el desempeño dispar en la resolución de tareas no estaba únicamente relacionado con la dificultad de la misma, además resultan importantes ciertas características cualitativas y cómo la información es procesada en una estructura adecuada [7]. Por ejemplo, una tarea dual tendrá un desempeño pobre si es que la información de ambas es provista mediante el canal visual a diferencia de lo que ocurre si una tarea es provista por el canal visual y la otra por el canal auditivo.

Wickens identificó las dimensiones estructurales del procesamiento de información humano que permiten explicar los cambios en el desempeño de las tareas.

El modelo consta de 4 categorías principales: Etapas de procesamiento, Modalidades de percepción, Canales visuales y Códigos de procesamiento [7].

En la figura 1 se muestra un esquema de la estructura del modelo de múltiples recursos.

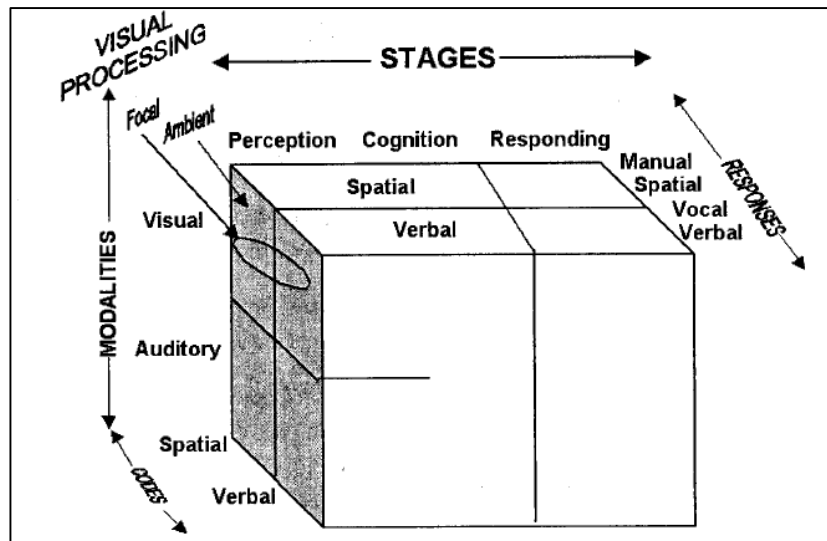


Figura 1: Estructura de modelo de múltiples recursos. Extraída de [7]

- Etapas de procesamiento

Son los recursos utilizados para realizar las actividades de percepción, las actividades cognitivas (por ejemplo, el uso de la memoria), y respuesta. Si una tarea requiere solo control motor no afecta el desempeño de tareas que requieran solo percepción y así sucesivamente [7].

- Modalidades de percepción

Es fácil darse cuenta de que es más fácil dividir la atención entre la vista y la audición que entre dos estímulos que lleguen por el mismo canal. Existen numerosos estudios que demuestran lo anterior y demuestran la dificultad de procesar dos señales que lleguen ya sea visual- visual o auditivo-auditivo [7].

- Canales visuales

Wickens separa los canales visuales en dos: Focal y ambiental. El canal focal es el que se utiliza al fijar la mirada en un objeto en particular percibiendo sus detalles y características, mientras que el canal ambiental es una visión periférica para tener sentido de orientación y de movimiento [7].

- Códigos de procesamiento

Esta dimensión define la distinción entre procesos espaciales y procesos lingüísticos o verbales. Esta separación radica en la eficiencia en el desempeño de las respuestas espaciales y verbales con lo que como en el caso de las modalidades de percepción si se ocupa el código espacial y verbal a la vez se disminuyen las interferencias entre las tareas [7].

En [31] se realizó una adaptación para explicitar de manera gráfica las interferencias en el modelo de Wickens y el sistema de procesamiento humano. La figura extraída de ese paper se muestra a continuación:

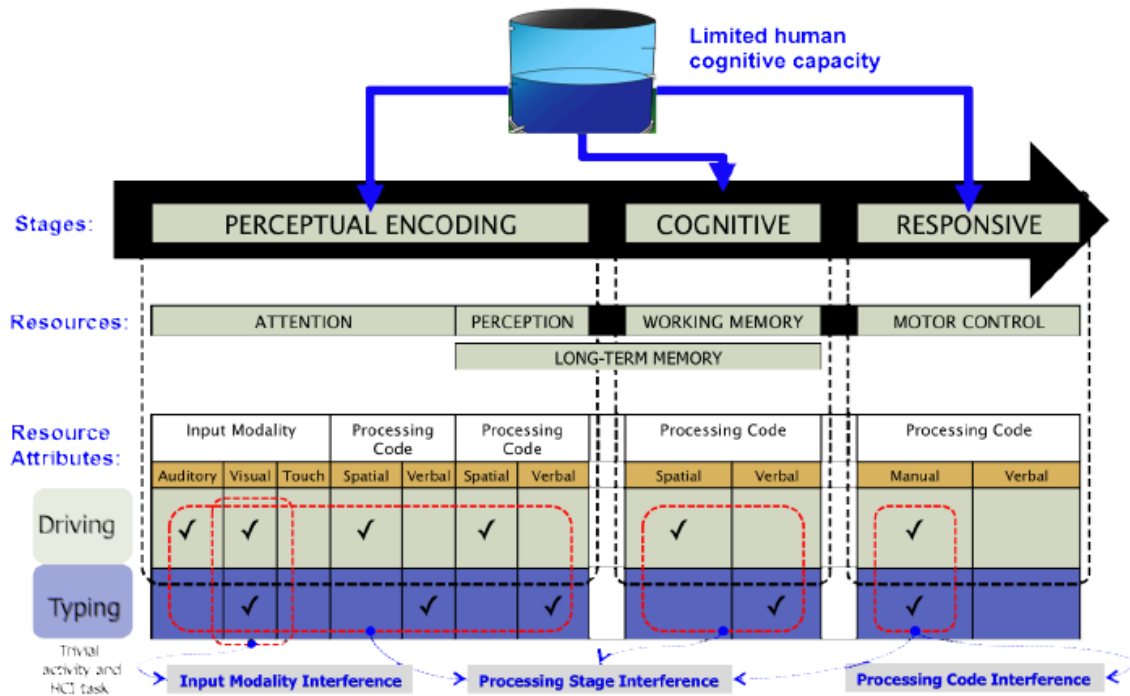


Figura 2: Sistema de procesamiento humano. Extraída de [31].

En la figura anterior se muestran los recursos utilizados mientras se maneja un automóvil y se escribe en el Smartphone.

De la figura también se observan los tipos de interferencia que se producen al realizar dos actividades:

- Interferencia por el canal: Esto tiene lugar cuando tanto la actividad física como la tarea HCI comparten el mismo canal. Lo que se produce es la detención del procesamiento y la interrupción de la tarea HCI. Esta interferencia es completa. [31]
- Interferencia en la etapa de procesamiento: Esta interferencia tiene lugar cuando se está utilizando la misma etapa de procesamiento para ambas tareas simultáneamente ya sea la etapa perceptual, cognitiva o de respuesta [31].
- Interferencia en el código de procesamiento: Se produce cuando se comparte el código de procesamiento ya sea espacial o verbal tanto en la misma etapa como en etapas distintas [31].

Al combinar interferencias se aumenta el grado de dificultad de una actividad [7]. Este resultado va a ser directamente aplicado en el diseño del experimento que se valida y se lleva a cabo.

## 2.2 Métodos de medición de carga cognitiva

### 2.2.1 Métodos subjetivos

Los métodos subjetivos, como los cuestionarios post actividad, permiten indicar la dificultad de la tarea en donde se les pregunta abiertamente a los individuos su percepción sobre ésta [11]. En general, poseen buena precisión, pero no permiten medir la carga cognitiva continuamente en tiempo real [9].

#### 2.2.1.1 NASA TLX (Task Load Index)

NASA TLX [10] es un ejemplo de test subjetivo. Consiste en evaluar la carga cognitiva de usuarios en múltiples dimensiones con una escala Likert. Se evalúa el desempeño mental, físico, demandas según el tiempo, esfuerzo y frustración [Hart et al. Development of NASA-TLX, 1988].

Consiste en un conjunto de ítems presentado en forma de afirmaciones o juicios, ante los cuales se pide la reacción de los participantes evaluando en una escala de 5 puntos. Se obtiene una puntuación por pregunta y al final una puntuación total dependiendo del peso relativo de cada variable.

Para ambientes ubicuos no es el test más aplicable, desde el punto de vista práctico, dado que en algunos escenarios del mundo real, por ejemplo cuando se quiere experimentar con tareas urgentes, no es posible aplicarlos inmediatamente después de las tareas con lo que se pierde capacidad de predicción [12].

### 2.2.2 Métodos basados en el desempeño de la tarea

Son métodos más objetivos que los anteriores, en donde el investigador debe medir el desempeño del sujeto durante la actividad, esto es, midiendo el tiempo que demora en realizarla y los resultados obtenidos. Estos métodos tienen la restricción de que el sujeto debe estar enfocado en una actividad, con lo que realizar dos o multi-tareas no es un escenario ideal. [9]

## 2.2.3 Medidas Psico-Fisiológicas

Las medidas psico-fisiológicas permiten detectar instantáneamente cambios en la carga cognitiva, por lo que es ideal si se quiere medir en tiempo real tareas duales o múltiples. Las dificultades que presenta son la incomodidad de utilizar los sensores y el complejo procesamiento de los datos [12].

Las comparaciones entre métodos han sido muy poco exploradas en la literatura. En esta memoria el enfoque será comparar el NASA TLX con las medidas psico-fisiológicas.

### 2.2.3.1 Sensores

Existe un amplio número de sensores capaces de proveer medidas directas o indirectas de la actividad cerebral. Se incluyen sensores que toman medidas directamente del cerebro como Electroencefalograma (EEG), Magnetoencefalograma (MEG), Resonancia Magnética Funcional (fMRI), y también medidas indirectas como la actividad cardiorrespiratoria, el pulso (HR), la variabilidad del pulso (HRV) así como también medidas de la actividad electrodermal y pupilometría [32].

Las ventajas del uso de estas medidas son [33]:

- Objetividad: Las medidas no dependen de la percepción del usuario ni están bajo el control de éste.
- Multidimensionalidad: Las medidas psicofisiológicas son sensibles a diferentes procesos cognitivos
- No intrusivos: Si bien algunos sensores requieren la colocación de electrodos, no se impide la ejecución de tareas
- Continuo: Se puede usar en tiempo real

## 2.3 Conceptualización del término Psico-Fisiología

La Psico-fisiología es una ciencia nueva, pero una idea antigua. Se basa en la creencia de que los síntomas experimentados por los seres humanos ante determinadas situaciones pueden ser medidos como reflejos propios de sus órganos vitales. Esto es, cómo relacionar la respuesta fisiológica del cuerpo con acciones psicológicas. [13].

Esta ciencia está íntimamente relacionada con la anatomía, fisiología y también con los fenómenos psicológicos, como la experiencia y comportamiento de organismos en un ambiente social y físico. Comparte muchas características con la neurociencia, pero la psico-fisiología busca integrar los procesos cognitivos con los procesos del sistema nervioso central y periférico. Entre la complejidad que se adhiere al moverse desde la fisiología a la psico-fisiología está la capacidad de los sistemas simbólicos de representación (por ejemplo, lenguaje, matemáticas) para comunicar y reflejar la experiencia y las influencias sociales y culturales del comportamiento y respuesta psicológica. Psicología y Psico-fisiología comparten el objetivo de explicar la experiencia humana y el comportamiento [13].

Los procesos psicológicos estudiados en Psico-fisiología corresponden a un amplio rango de actividades humanas incluyendo: aprendizaje, resolución de problemas, toma de decisiones, percepción, atención, dormir, sentir y respuesta emocional. Las respuestas fisiológicas incluyen señales captadas desde el cerebro, corazón, músculos, piel y ojos, entre otras [30].

Para el tema de esta memoria, la Psico-fisiología cognitiva se preocupa de las relaciones entre el sistema de procesamiento de información, como lo explicado en el capítulo 2.1, con eventos fisiológicos humanos [13].

La psico-fisiología está compuesta por variables dependientes e independientes. Las variables dependientes se refieren a lo que está siendo medido en la investigación, mientras que las variables independientes a lo que está siendo manipulado. Así, las variables dependientes son fisiológicas (ej, la dilatación pupilar) y las variables independientes son psicológicas (ej, resolución de un problema) [30].

Cuando se pretende usar medidas fisiológicas para evaluar estados psicológicos se debe tener en cuenta que cada medida fisiológica no es independiente de la otra. Son parte de un conjunto de respuestas complejas provenientes desde el sistema nervioso autónomo [13].

## 2.4 Medidas fisiológicas en relación al sistema nervioso [30]

Las medidas tomadas por los psicofisiólogos incluyen una o más de las siguientes: Electroencefalograma (EEG), Potencial Cerebral relacionado con eventos (ERP), Electromiograma (EMG, medida de la actividad muscular), Pupíloetría (medición de cambios del tamaño de la pupila), Electrooculografía (EOG, medida del movimiento del

ojo), Actividad Electrodermal (EDA o GSR, cambios en la actividad eléctrica de la superficie de la piel), Respuestas del corazón, Volumen sanguíneo, Presión sanguínea, Respiración, Consumo de oxígeno, Salivación, Temperatura de la piel, Motilidad gástrica.

Estas medidas están todas bajo el control del Sistema Nervioso. El sistema nervioso puede ser dividido en dos ramas. El sistema nervioso central (CNS, por sus siglas en inglés) y el sistema nervioso periférico (PNS). El CNS incluye al cerebro y a la médula espinal.

El sistema nervioso periférico se refiere a los nervios que están fuera de estas dos grandes estructuras. Este sistema está dividido entre el sistema somático (actividad muscular) y el sistema nervioso autónomo (ANS), el cual controla las estructuras viscerales (glándulas y órganos del cuerpo). Finalmente el ANS está subdividido en el sistema nervioso Parasimpático (PNS), que incluye a los mecanismos de inervación cuando el individuo está descansando, y el Sistema Nervioso Simpático (SNS), el que es dominante en situaciones que requieren movilización de energía.

La tabla 2 indica qué parte(s) del sistema nervioso controlan a las medidas fisiológicas.

Sistema nervioso Central	Sistema somático	Sistema nervioso autónomo
EEG ERP	EMG EOG	Pulso cardíaco (PNS, SNS) Presión sanguínea (PNS, SNS) Actividad electrodermal (SNS) Respuesta pupilar (PNS, SNS) Volumen sanguíneo (PNS, SNS)

Tabla 2: Control de las respuestas fisiológicas por parte del sistema nervioso.

### 2.4.1 El sistema nervioso autónomo (ANS)

Muchas de las respuestas fisiológicas de interés para esta ciencia, y en particular para esta memoria, están controladas por el sistema nervioso autónomo. Este sistema regula y coordina importantes actividades corporales como digestión, temperatura, presión sanguínea y muchos aspectos de comportamiento emocional. Estas actividades son realizadas independiente del control consciente del individuo. El objetivo principal del sistema es mantener los equilibrios en organismo, controlando aspectos tanto internos como externos. Inerva a tres tipos de células: Músculo liso, músculo cardíaco y glándulas.



Las glándulas sudoríparas son inervadas solo por el sistema nervioso simpático, lo que es una excepción dado que la mayoría de los órganos son inervados tanto por el sistema nervioso simpático como por el parasimpático.

## 2.4.2 El sistema nervioso Simpático

Controla las actividades en que se requiere movilidad, stress o situaciones de emergencia. Puede provocar la aceleración del corazón, dilatación de la pupila, incrementar el volumen y el flujo sanguíneo a los músculos, incrementar el sudor, la tensión muscular, entre muchas otras.

## 2.4.3 El sistema nervioso Parasimpático

Controla las funciones relacionadas con el descanso, reparación y relajación del cuerpo. Entre las reacciones que provoca este sistema está decremento del pulso cardíaco y de la presión sanguínea, estimulación del sistema digestivo, contracción pupilar, descanso y sueño.

Ambos sistemas se complementan integrando sus funciones permitiendo un flujo suave de las actividades y el comportamiento.

## 2.5 El sistema Visual

La ciencia ha descubierto algunos aspectos interesantes de los cambios en el tamaño de la pupila y la actividad mental. Pupilometría se refiere a la medición de las variaciones en el diámetro de la apertura pupilar del ojo [30].

Las dilataciones y contracciones de la pupila están gobernadas por el sistema nervioso autónomo. Su cambio de tamaño ocurre durante el procesamiento de información, percepción, memoria de corto plazo, aprendizaje, entre otros [30].

Para introducir el estudio de la pupila, primero se verán las estructuras principales del sistema visual.

Este está compuesto por las siguientes estructuras [18][19]:

- **Globo ocular:** Estructura mayor que da forma al ojo, de forma esferoidal.
- **Córnea:** Estructura transparente que permite la refracción de la luz con el fin de que se pueda enfocar la imagen en la retina.
- **Iris:** Tiene la forma de un disco perforado en su centro por un orificio circular, la pupila.
- **Pupila:** Orificio de tamaño variable. Su diámetro normal se encuentra comprendido entre 2 y 4 mm, pudiendo alcanzar los 8 mm en dilatación máxima y 0,5 mm en contracción máxima.
- **Retina:** Capa interna del globo ocular. En esta se inicia el proceso de visión con múltiples terminaciones nerviosas que recogen, elaboran y transmiten sensaciones visuales.
- Otras estructuras que se pueden nombrar son la esclerótica, cuerpo ciliar, coroides, humor vítreo, humor acuoso y cristalino.

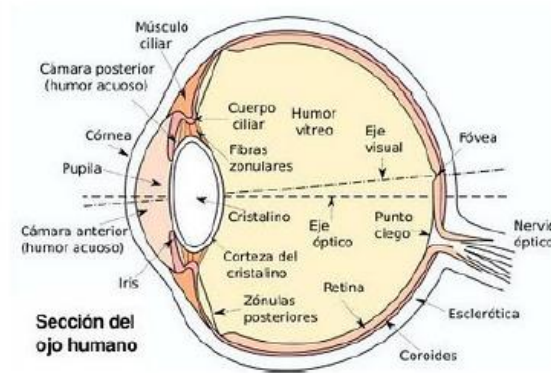


Figura 3 Estructura del ojo humano [36]

## 2.5.1 Pupila

La Pupila es un orificio al centro del iris en el ojo a través de la cual la luz pasa. Luego, la pupila no es más que un orificio rodeado por el músculo del iris. Este músculo, comandado por el sistema nervioso parasimpático se encarga de contraerla y dilatarla teniendo como función primaria ajustar los cambios de luminosidad dependiendo del ambiente [30].

En la pupila existen pequeños movimientos visualmente imperceptibles que están cognitivamente relacionados. Estos pequeños movimientos forman parte de la base de la pupilometría cognitiva proveyendo un único índice psico-fisiológico de actividad dinámica cerebral en la cognición humana [20].

## 2.5.2 Medición

La respuesta pupilar se ha mostrado que es una función de cuán difícil es una tarea. Los primeros estudios datan del año 1964 en donde se aumentaba gradualmente la dificultad de una multiplicación. El resultado fue que el tamaño pupilar aumentó en un rango de 4% a 30% en comparación con la situación previa a la pregunta. Otros estudios observaron relaciones con la memoria de corto plazo y excitación [30].

Históricamente la forma de medir el diámetro pupilar ha sido tomar múltiples fotografías al ojo mientras se le presentan los estímulos correspondientes. Actualmente se utiliza un dispositivo llamado Eyetracker que graba y mide la trayectoria del ojo en todo instante [21]. Posteriormente los datos son analizados ya sea manual o automáticamente.

Con respecto a la pupila, específicamente, el Eyetracker usa la metodología de la reflexión infrarroja de la córnea, tiene una cámara que mide la distancia y el ángulo de reflexión de la luz infrarroja del centro de la pupila a un determinado punto de fijación de una persona [37][19].

## 2.5.3 La Pupila y la carga cognitiva

Empíricamente, los cambios en el tamaño de la pupila han reflejado la variación de carga cognitiva con extraordinaria precisión. La respuesta pupilar es una correlación del uso de recursos cognitivos, luego, aunque no se sabe la relación directa entre ambas, se presume que debe ser una relación profunda en el sistema nervioso central [20].

Se ha comprobado que a medida que la tarea se hace más demandante en trabajo y tiempo el tamaño de la pupila aumenta y si es que los requerimientos son continuos la dilatación se mantendrá [38].

En [22] se muestra que la pupila es una medida significativa de la respuesta fisiológica en distintas tareas. En [20] se muestra la correlación existente entre la respuesta pupilar y tareas de distinta dificultad.

Además se conoce que existen otros factores que influyen en la respuesta pupilar y que degradan la relación con la cognición como por ejemplo la luminosidad. Es por esto que Xu, J. et al. [23] comprobó que aún existe la correlación entre la carga cognitiva y la pupila al modificar la dificultad de las tareas cambiando la luminosidad de la habitación en un ambiente controlado con cuatro niveles de luminosidad.

## 2.6 El Sistema Cardiovascular

El sistema cardiovascular es especialmente interesante para la psico-fisiología puesto que es sistema complejo con terminaciones nerviosas que regulan su acción por lo que es altamente sensible a los procesos neurológicos y a factores psicológicos como el estrés [13].

Está regulado por los sistemas simpático y parasimpático del sistema nervioso autónomo produciéndose patrones de inervación eléctrica que son fundamentales para las mediciones psico-fisiológicas [13].

El sistema cardiovascular está compuesto por el corazón y los vasos sanguíneos. Juntos se encargan de llevar la sangre oxigenada a todos los rincones del cuerpo humano y de llevar hacia los pulmones la sangre contaminada [13].

Existen varios estudios que asocian cambios en la actividad cardíaca con fenómenos psicológicos como trabajo mental, percepción, atención, resolución de problemas, detección de señales, entre otros [30].

### 2.6.1 El corazón

Es la bomba del sistema. Está formado por músculo cardíaco auricular, ventricular y fibras nerviosas encargadas de conducir los impulsos eléctricos del corazón.

La aurícula derecha recibe sangre de todo el cuerpo excepto de los pulmones. Las venas le llevan la sangre contaminada desde todo el cuerpo (Vena cava superior e inferior). La sangre fluye desde la aurícula derecha al ventrículo derecho y desde ahí a los pulmones vía la arteria pulmonar. En los pulmones la sangre se oxigena y vuelve a la aurícula izquierda por la vena pulmonar. Desde ahí va hacia el ventrículo izquierdo el cual bombea la sangre oxigenada a todo el organismo a través de la arteria Aorta [30].

Las aurículas y los ventrículos están conectados por un sistema eléctrico que permite una corta contracción ventricular después de la contracción auricular, además las fibras eléctricas permiten el bombeo de ambas partes del corazón. La despolarización de los dos nodos de tejido eléctricamente activo, el nodo sinoauricular (SA) y el nodo auricoventricular (AV) provee el impulso eléctrico que gatilla la contracción del corazón. El nodo SA es el marcapasos del corazón puesto que controla la tasa de bombeo al ser más rápido que el nodo AV [13].

El músculo cardíaco es diferente a los demás músculos del cuerpo debido a que su período de repolarización tiene una meseta de 0.2-0.3 segundos con lo que la sangre tiene

el tiempo necesario para viajar desde las aurículas a los ventrículos antes de que éstos se contraigan. [13]

La tasa normal de contracción también es influenciada por los nervios del Sistema nervioso Autónomo. El sistema nervioso Parasimpático influencia a los nodos SA y AV vía el nervio Vago disminuyendo los latidos en condiciones de reposo y descanso. El sistema nervioso Simpático incrementa los latidos en situaciones de stress, ejercicio físico o en situaciones de trabajo mental [13].

El latido del corazón representa la contracción que el corazón realiza al bombear sangre al cuerpo. El corazón humano normalmente se contrae alrededor de 72 veces por minuto en reposo. El control del latido es un mecanismo tanto interno del corazón como externo [30].

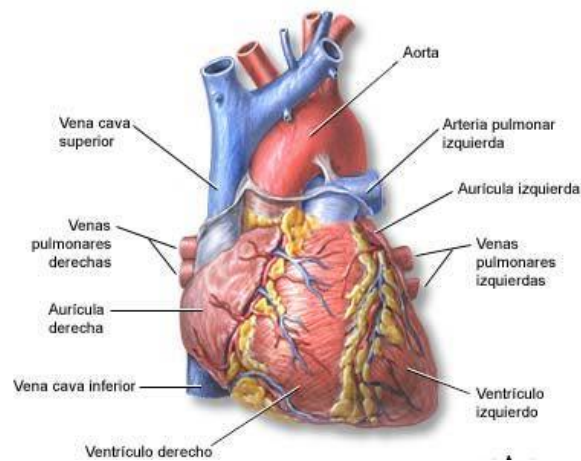


Figura 4 Estructura del corazón [39]

## 2.6.2 El Ciclo Cardíaco

Se denomina así a los eventos que ocurren en el corazón desde un latido hasta el siguiente. El ciclo está compuesto por dos etapas principales, *diástole* durante la cual el corazón se llena de sangre y *sístole* cuando el corazón bombea toda esa sangre al cuerpo [13].

El ciclo comienza cuando el nodo SA se despolariza en la aurícula derecha durante la última parte de la diástole. Esta primera onda de despolarización corresponde a la onda P en la señal eléctrica generada por el corazón. La onda P es seguida por una contracción auricular en donde aparece la onda QRS reflejando el comienzo de la sístole. Se produce la contracción ventricular en donde la presión ventricular es lo suficientemente alta para cerrar las válvulas auriculoventriculares entre las aurículas y los ventrículos. Luego, la

presión ventricular disminuye a causa del aumento de presión auricular con lo que se abren las válvulas y los ventrículos se llenan de sangre. Cuando la presión ventricular es mayor que la presión de la arteria Aorta se abre la válvula y la sangre se dirige a la circulación general. Ocurrido esto los ventrículos se repolarizan y se observa la onda T con lo que vuelve a comenzar la diástole. [13]

El tiempo entre el comienzo de la onda P y el comienzo de la onda QRS es alrededor de 160 mseg. El intervalo Q-T es alrededor de 300 mseg. Un ciclo cardíaco basado en una tasa de 72 latidos por minuto dura 830 mseg por lo que hay aproximadamente 370 mseg entre el fin de la onda T y el comienzo de la siguiente contracción auricular [13].

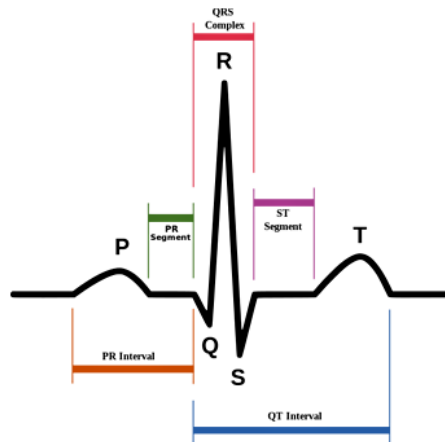


Figura 5: Ciclo cardíaco ideal [40]

### 2.6.3 Medición

- Electrocardiograma (ECG)

Es un dispositivo desarrollado para obtener la actividad eléctrica del corazón. Consta de tres electrodos que se pegan en el torso del sujeto y que recopilan los datos necesarios para obtener las ondas descritas en el ciclo cardíaco.

En un ECG normal se producen ondas R cuyo rango de amplitud es aproximadamente 2 mV. El peak de 2 mV debe ser amplificado por 2500 para poder trabajar [30].

- Pulso Cardíaco/Período cardíaco (HR)

Se denomina período cardíaco al tiempo en milisegundos entre dos latidos consecutivos. Típicamente se mide entre dos sucesivas ondas R en un ECG.

Habitualmente el período cardíaco se convierte a pulso cardíaco en donde se mide la cantidad de latidos por minuto [13].

- Variabilidad del pulso cardíaco (HRV)

Es un análisis basado en las amplitudes de las señales del intervalo cardíaco a varias frecuencias. El espectro HRV puede ser clasificado en tres bandas de frecuencia asociadas al tiempo del intervalo R-R. [27]

- Frecuencia Baja (0.02-0.06 Hz) asociado con la regulación de la temperatura corporal
- Frecuencia Media (0.07-0.14 Hz) asociado con la regulación de la presión arterial
- Frecuencia Alta (0.15-0.50 Hz) que refleja los efectos de la respiración en el intervalo cardíaco.

## 2.6.4 Sistema cardiovascular y carga cognitiva

Numerosos estudios existen comprobando la correlación entre el pulso cardíaco y la carga cognitiva [24].

En [30] se explicitan estudios que asocian el pulso cardíaco con el desempeño en una tarea mental de aprendizaje. Cuando aumenta la dificultad del aprendizaje, la persona está menos comprometida con la tarea y se produce una baja en el ritmo cardíaco.

En términos de resolución de problemas en [30] se explicitan correlaciones entre dificultad de problemas y aumento del ritmo cardíaco.

En [25] encontró que el pulso aumenta entre 10-15% al aumentar la dificultad de la tarea. También [26] dice que el pulso aumenta 15% conforme avanza el experimento. A su vez [24] obtuvo que el pulso cardíaco aumenta entre 9.7-19% mientras aumenta la dificultad de la tarea.

Además muchos estudios muestran que un aumento de la carga cognitiva está típicamente relacionado a una reducción del poder asociado a la Frecuencia Media implicando una supresión temporal de la presión arterial [27].

## 2.7 Trabajos anteriores

Para hacer un compendio de las medidas fisiológicas que se asocian a procesos psicológicos se presenta la tabla 3 en donde se indican los sensores que se utilizarán en esta memoria. [33]

Medida Fisiológica	Proceso Psicológico
Medidas cardiovasculares	HR es sensible a demandas cognitivas, restricciones de tiempo y también a la excitación.
Diámetro Pupilar	Está correlacionada con trabajo mental. Es un índice de cambios en el procesamiento de información. También responde a valencia emocional.
Oxímetro	Estrés [15]
Temperatura	Emocionalidad, relajación [15]

Tabla 3: Medidas fisiológicas asociadas a procesos psicológicos. [33]

El foco de esta memoria es medir carga cognitiva con un grupo de sensores. Para esto se han estudiado trabajos que han abordado el problema en escenarios distintos al ambiente móvil, pero que sirven para conocer la metodología asociada a este tipo de investigación.

Como se explicitó en la sección 1.6 la idea básica es crear actividades de distinta dificultad para entrenar clasificadores que utilizarán la data de los sensores con el fin de predecir si la carga asociada a una tarea es alta o baja.

A continuación se muestran estudios que siguen esta metodología, los sensores que utilizaron y los resultados que obtuvieron.

- Haapalainen et al. [9]

Objetivo	Medición de cargas cognitivas en tareas cognitivas básicas
Sensor(es) psicofisiológico	Eyetracker, EEG, ECG, flujo de calor, pulso
Experimento	Problemas desktop-based de percepción visual y velocidad cognitiva
Resultados	ECG y flujo de calor juntos distinguen con un 80% de precisión tareas de alta y baja demanda cognitiva



- Fritz et al [28]

Objetivo	Comprobar si a través de sensores biométricos se puede percibir la dificultad de una tarea mientras se está trabajando en ésta.
Sensor(es) psicofisiológico	Eyetracker, EEG
Experimento	8 tareas de comprensión de código variando la dificultad
Resultados	85% precisión la dificultad de una tarea usando Machine Learning

- Shi et al [14]

Objetivo	Evaluar los niveles de estrés y excitación al incrementar niveles de dificultad
Sensor(es) psicofisiológico	GSR (Respuesta galvánica de la piel)
Experimento	Interfaz de control de tráfico con 3 escenarios. 1) Usando gestos y hablando, 2) Solo hablando y 3) Solo usando gestos. Además se varía la dificultad en 5 formas: <ul style="list-style-type: none"> <li>- Complejidad visual</li> <li>- Número de entidades</li> <li>- Número de distractores</li> <li>- Tiempo limite</li> <li>- Número de acciones para finalizar</li> </ul>
Resultados	Resultados del GSR indican que existe un incremento significativo cuando aumento la dificultad de la tarea cognitiva

- Nourbakhsh et al [17]

Objetivo	Evaluar eficacia en la medición de GSR ante tareas de distinta dificultad
Sensor(es) psicofisiológico	GSR
Experimento	8 tareas aritméticas con 4 niveles de dificultad (binario, 1 dígito, 2 y 3 debiendo juntar los números y responder el número de dígitos existente)
Resultados	Se aprueba la hipótesis nula

- Nourbakhsh et al [41]

Objetivo	Medición de carga cognitiva
Sensor(es) psicofisiológico	GSR, eye blink
Experimento	8 tareas aritméticas con 4 niveles de dificultad (binario, 1 dígito, 2 y 3 debiendo juntar los números y responder el número de dígitos existente)
Resultados	Combinando los dos sensores se llega a una precisión de 75% para binario y 50% para 3 dígitos.

- Xu et al [23]

Objetivo	Verificar si se puede medir carga cognitiva mediante dilatación pupilar si es que la luminosidad cambia
Sensor(es) psicofisiológico	Eyetracker
Experimento	Tareas aritméticas: sumar 4 números que aparecen en pantalla. Varía en nivel de dificultad según el número de dígitos (4 grados de dificultad desde binario a 3 dígitos). Además para objeto de la hipótesis se varía la luminosidad de la habitación.
Resultados	Se aprueba la hipótesis nula

- Ikehara et al [15]

Objetivo	Evaluar carga cognitiva
Sensor(es) psicofisiológico	Eyetracker, presión del mouse, GSR, pulso oxímetro
Experimento	Seleccionar en la pantalla las fracciones cuyo valor es mayor a $1/3$ . Dos grados de dificultad. La dificultad baja son fracciones de rápida resolución. La dificultad alta se logra mediante un aumento en la dificultad de las tareas perceptivas (aumento del número de fracciones en pantalla y aumentando la velocidad en que son presentadas en pantalla) y aumento en la dificultad de tareas cognitivas (reducción de tiempo y aumento del rango de valores)
Resultados	GSR y eyetracker tienen la mayor significancia para detectar la dificultad de la tarea

## 2.8 Proceso Knowledge Discovery in Databases

Parte de la metodología de este estudio descrita en el capítulo 1 se basa en la aplicación del proceso Knowledge Discovery in databases (KDD) con el objetivo de procesar los datos, transformarlos, analizarlos y aplicar un algoritmo de predicción de minería de datos. El Proceso KDD está compuesto por nueve fases desde plantear los objetivos hasta aplicar el conocimiento descubierto [45].



Figura 6 Fases del proceso KDD

Las fases enunciadas en la figura 6 se detallan a continuación [45]:

- 1) **Comprensión del problema y definición de objetivos:** En esta fase inicial la idea es desarrollar un entendimiento acabado del problema a fin de establecer los objetivos del mismo tomando en cuenta que lo que se decida en esta fase tiene repercusiones para las fases futuras.
- 2) **Selección:** Esta es una fase muy importante dentro del proceso puesto que se debe saber qué tipo de datos se tienen disponibles y la calidad de los mismos. Se debe evaluar si los datos existentes con sus atributos son suficientes para explicar el problema planteado o si es el caso contrario, evaluar la posibilidad de recolección de nuevos datos. Los datos que se seleccionen en esta fase representarán la materia prima del modelo a desarrollar por lo que su calidad es imprescindible. Una vez

seleccionados es momento para organizarlos y planear las configuraciones óptimas que puedan facilitar el proceso y/o mejorar los resultados.

- 3) **Preprocesamiento:** En esta fase se trabaja con los datos disponibles con el fin de solucionar ciertos problemas que existen en toda base de datos. Problemas típicos son: Tratamiento de valores perdidos, datos fuera de rango, eliminación del ruido, entre otras. Hay casos más complejos en que deben utilizar algoritmos de minería de datos, aplicar ciertas funciones para eliminar interferencias o aplicar algún método estadístico.
- 4) **Transformación:** Consiste en agregarle valor a los datos transformándolos en medidas acordes al problema de estudio de forma tal que se puedan ajustar mejor al problema. Se puede aplicar alguna función apropiada, extraer características desde los datos, trabajar la dimensionalidad y los atributos, etc.
- 5) **Selección de la tarea de minería de datos:** En este punto se debe seleccionar el tipo de minería de datos con el que se trabajará. Esto tiene relación directa con el planteamiento de objetivos del punto uno. Se puede querer realizar una clasificación, clusterización, regresión. Se debe decidir si se quiere ocupar un algoritmo de predicción o de descripción. Para esto se debe saber que los algoritmos de predicción son parte de los métodos supervisados en donde se entrega data real y entrenan modelos con el fin de que se construya el conocimiento necesario para que se puedan predecir casos futuros. En el caso de la descripción se tienen como objetivos clasificar la data, encontrar relaciones, visualizar parámetros.
- 6) **Selección de algoritmo de minería de datos:** En este punto se debe especificar qué algoritmo se usará tomando en consideración los pro y contra de cada algoritmo pensando en el problema en sí y en cómo se ajusta el algoritmo. Se deben tomar en cuenta los parámetros, el nivel de detalle, de respuesta, etc.
- 7) **Implementación:** Se aplica el algoritmo de minería de datos teniendo especial cuidado en correr el algoritmo varias veces con el objetivo de encontrar los parámetros adecuados para los datos que se disponen.
- 8) **Evaluación e interpretación:** En esta etapa se evalúan e interpretan los resultados obtenidos con la referencia de los objetivos planeados en un comienzo. Esto se hace con el fin de concluir la fase y pasar a la siguiente o volver a fases anteriores para cambiar ciertas condiciones y re-evaluar.

- 9) Conocimiento descubierto: Utilizar el conocimiento descubierto para otros sistemas con el fin de ver los efectos que se pueden generar. Si resulta bien se concluye que el proceso KDD fue un éxito.

## 2.8.1 Técnicas de minería de datos

A gran escala, las técnicas de minería de datos se dividen en las que intentan verificar alguna hipótesis y las que intentan descubrir ciertos patrones o comportamientos presentes. A continuación se presenta una taxonomía de técnicas de minería de datos:

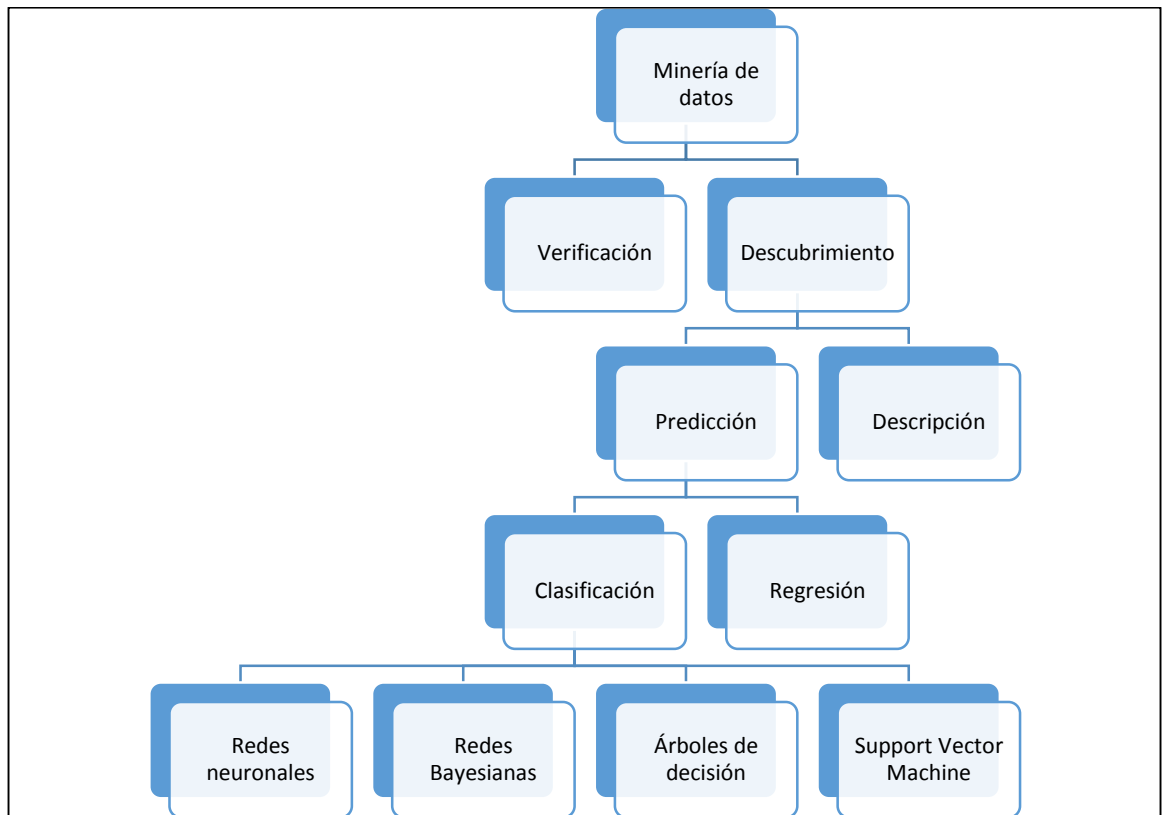


Figura 7 Técnicas de minería de datos [45]

Las técnicas de descubrimiento buscan crear modelos de comportamiento a partir de datos y modelos entrenados. Como se ve en la figura 7 las técnicas de descubrimiento se dividen en dos categorías. La primera es la de descripción, que intenta explicar e interpretar los datos desde el enfoque de la búsqueda de relaciones entre los datos. Los métodos predictivos en cambio pretenden construir modelos que sean capaces de tomar conjuntos de datos de entrenamiento y con esa información ser capaz de predecir para nuevos registros los valores de sus variables. Luego, los modelos predictivos tienen a su

vez dos clases más. Los algoritmos de clasificación y los algoritmos de regresión. Los algoritmos de clasificación pretenden predecir si un registro será incluido dentro de un determinado grupo o no, es decir, toman nuevos datos y devuelven el porcentaje de clasificación que obtienen. Los algoritmos de regresión se basan en generar pronósticos a partir de determinadas variables en un modelo. Así, un algoritmo de regresión puede predecir el precio de la acción de determinada compañía en base ciertas variables del mercado [45].

También es importante señalar la diferencia entre algoritmos supervisados versus no supervisados. Los primeros son algoritmos que intentan descubrir una relación entre las variables dependientes e independientes estudiadas modelándola utilizando patrones o modelos de comportamiento. Los segundos son algoritmos que no utilizan patrones y que buscan relaciones de acuerdo a la calidad de los datos [19].

## 2.8.2 Algoritmos de minería de datos

Como se explicó en el capítulo uno, en esta memoria se utilizan algoritmos de minería de datos. Se pretende saber si las señales de los sensores pueden clasificar la dificultad de las actividades con lo que los algoritmos a estudiar serán los de la categoría de Predicción.

- Naive Bayes [46]

Este algoritmo ha sido uno de los más utilizados en el ámbito de la Psicofisiología debido principalmente a sus buenos resultados y a su relativa facilidad de implementar.

Se tiene la conocida fórmula de Bayes:

$$P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^n P(E|F_i)P(F_i)}$$

Se considera a  $E$ , la evidencia de  $F$  y se calcula la probabilidad de que  $F$  ocurra dada la evidencia  $P(E|F)$ .

Si  $E = E_1, E_2, \dots, E_m$  teniendo evidencia de múltiples fuentes, la ecuación anterior se transforma a

$$P(F_j|E_1E_2 \dots E_m) = \left( \frac{P(E_1E_2 \dots E_m|F_j)P(F_j)}{P(E_1E_2 \dots E_m)} \right)$$

Lo que dificulta el cálculo por la dificultad de obtener el término  $P(E_1E_2 \dots E_m|F_j)$

Se supone entonces que los  $E_i$  son independientes dado  $F_j$  lo que permite escribir:

$$P(F_j|E_1E_2 \dots E_m) = \left( \frac{P(E_1|F_j)P(E_2|F_j) \dots P(E_m|F_j)}{P(E_1E_2 \dots E_m)} \right)$$

Que es el clasificador Naive Bayes para determinar la probabilidad de presencia en determinadas clases.

- Regresión Logística [45][19]

La regresión logística se utiliza para clasificar y derivan de las regresiones lineales. La diferencia está en que las regresiones logísticas predicen con cierta probabilidad la pertenencia o no a cierto grupo, luego su valor se ubica entre 0 y 1.

Definiendo el modelo  $f_\theta$ :

$$f_\theta(x) = h(\theta^T x)$$

$$h(y) = \frac{1}{1 + \exp^{-y}}$$

$f_\theta$  es una regresión logística,  $x$  es el caso que se quiere clasificar y  $\theta$  son los parámetros de la regresión.

$h$  es la función logística que sus valores varían entre 0 y 1.

Como el general de modelos de minería de datos, para entrenar o fijar los parámetros se establece una función de costo que al minimizarla encuentre los parámetros que se ajustan a los datos de entrenamiento. La función de costo se define así:

$$J(\theta) = -\frac{1}{n} \left[ \sum_{i=1}^n z^i \log f_\theta(x^i) + (1 - z^i) \log(1 - f_\theta(x^i)) \right]$$

Con  $z$  la etiqueta de la clase y  $x$  los datos.

Para evitar el sobreajuste de los datos se puede aplicar la técnica de regularización estableciendo un nuevo parámetro  $\gamma$  que se debe fijar mediante inspección:

$$J(\theta) = -\frac{1}{n} \left[ \sum_{i=1}^n z^i \log f_{\theta}(x^i) + (1 - z^i) \log(1 - f_{\theta}(x^i)) \right] + \frac{\gamma}{2n} \sum_{j=1}^t \theta_j^2$$

- Support Vector Machine (SVM) [44][19]

Algoritmo que actualmente está siendo muy usado por las investigaciones debido a sus excelentes resultados y a que dentro de su complejidad teórica existen muchas características que se pueden ajustar para un mejor pronóstico.

En la mayoría de los casos SVM es un clasificador binario.

Se supone un set de entrenamiento  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  donde  $\mathbf{x}$  son los vectores de datos e  $\mathbf{y}$  es la etiqueta de la clase que toma el valor de 1 o -1 si pertenece o no a la clase en cuestión.

Se busca la función lineal de la forma

$$f(\mathbf{x}) = \langle \mathbf{w} \cdot \mathbf{x} \rangle + b$$

De tal forma que si la clase de  $\mathbf{x}$  es positiva  $f(\mathbf{x}) \geq 0$ .  $w$  es el vector de pesos y  $b$  es denominado bias.

$$y_i = \begin{cases} 1 & \text{si } \langle \mathbf{w} \cdot \mathbf{x} \rangle + b \geq 0, \\ -1 & \text{si } \langle \mathbf{w} \cdot \mathbf{x} \rangle + b \leq 0. \end{cases}$$

SVM busca el hiperplano que separe a los datos de entrenamiento entre las clases positivas y las negativas.

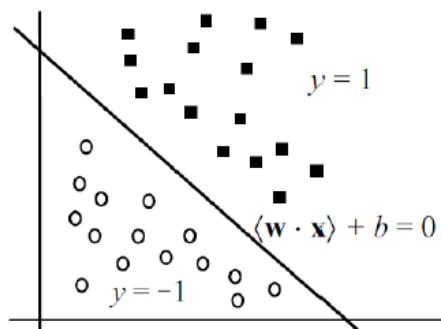


Figura 8 Hiperplano SVM [45]



Para encontrar este hiperplano se maximiza el margen entre los datos de distinta clase. Se expresa entonces como un problema de optimización

$$\begin{aligned} \text{Min} : & \frac{\langle \mathbf{w} \cdot \mathbf{w} \rangle}{2} \\ \text{s.a.} : & y_i(\langle \mathbf{w} \cdot \mathbf{x} \rangle + b) \geq 1 \end{aligned}$$

Este sin embargo es un caso ideal, por lo que para relajar el problema se agregan variables de relajación

$$\begin{aligned} \text{Min} : & \frac{\langle \mathbf{w} \cdot \mathbf{w} \rangle}{2} + C \sum_{i=1}^n \xi_i \\ \text{s.a.} : & y_i(\langle \mathbf{w} \cdot \mathbf{x} \rangle + b) \geq 1 - \xi_i \\ & \xi_i \geq 0. \end{aligned}$$

Por último, si el problema no es lineal se deben aplicar kernels que modifiquen el espacio en el que se encuentran los datos.

## 3. Experimento

En este capítulo se detalla el diseño experimental elaborado junto con la instrumentación utilizada. Además se indican aspectos relevantes de la implementación.

### 3.1 Instrumentación

Para el experimento se utilizan principalmente dos dispositivos. El primero es una arquitectura de hardware que es la responsable de obtener los registros desde los sensores de pulsoxímetro, temperatura y ECG y enviarlos a un servidor en la nube. El segundo es el software del Eyetracker instalado en una Tablet Microsoft Surface.

A continuación se presentan mayores detalles de los dispositivos.

#### 3.1.1 Dispositivo de captura de datos

Este dispositivo es capaz de obtener los valores de los sensores conectados al sujeto y guardarlos en una tarjeta de memoria SD para luego ser transferidos a un servidor remoto. Este dispositivo es el encargado de recibir los datos provenientes de los sensores, transformarlos para que puedan ser leídos digitalmente, enviarlos a una base de datos MongoDB y ajustar los timestamp para coordinar los tiempos de medición.

El dispositivo está compuesto por:

- Tarjeta Raspberry Pi modelo B+: Es un pequeño computador basado en Linux que se conecta vía serial a la tarjeta aLaMode. Obtiene los valores de los sensores y los escribe en un archivo utilizando formato JSON (Java Script Object Notation). Estos datos son enviados a un servidor Amazon. Tiene una tarjeta WiFi que le permite conectarse a la red interna.



Figura 9: Raspberry Pi [40]

- Tarjeta Alamode: Es una tarjeta Arduino basada en un chip Atmel 328 en formato Shield. La tarjeta se encarga de convertir las variables analógicas entregadas por los sensores en valores digitales para que puedan ser procesados por la Raspberry Pi.



Figura 10: Tarjeta aLaMode [40]

- Shield de e-health: Encargado de proveer la interfaz para conectar los sensores al dispositivo de captura de datos. Además tiene la función de convertir los valores eléctricos de las señales en valores que pueden ser procesados por la tarjeta aLaMode.



Figura 11: Shield e-health [40]

- Tarjeta MoPi: Encargada de regular y monitorear la fuente de alimentación del dispositivo que es un pack de 8 baterías tipo AA.



Figura 12: Tarjeta MoPi [40]

- Router WiFi: Encargado de proveer la conectividad a internet al dispositivo para enviar los datos a la nube, donde son procesados y guardados en una base de datos. El router WiFi se conecta a Internet usando un Dongle 3G.
- Servidor de procesamiento: El servidor principal del sistema está ubicado en la nube. Es un servidor virtual basado en Ubuntu Server.
- Todo el dispositivo se encuentra en un bolso de cámara fotográfica profesional para facilitar el desplazamiento.

### 3.1.2 Sensores Psicofisiológicos

Como se mencionó anteriormente, los sensores que se utilizan en este trabajo son los siguientes:

- Electrocardiograma: Herramienta de diagnóstico usada para evaluar las respuestas eléctricas y musculares del corazón.



Figura 13: Sensor ECG [40]

El sensor consta de tres electrodos que se conectan a la Shield e-health y a la persona de la siguiente manera:

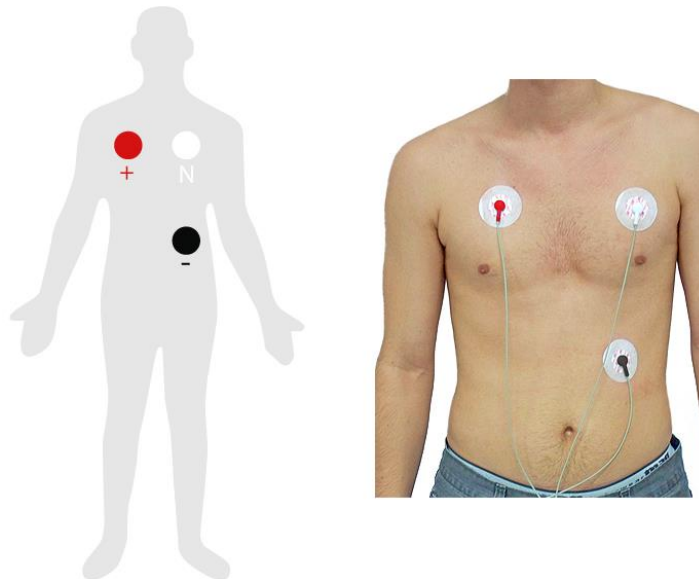


Figura 14: Posición de electrodos ECG en el cuerpo [40]

Este sensor entrega como dato la respuesta eléctrica del corazón en mV.

- Temperatura: Sensor que mide la temperatura superficial de la piel del individuo. Se pone en un dedo de la mano.



Figura 15 Sensor de Temperatura [40]

Este sensor entrega como dato la temperatura en grados Celsius.

- Pulsooxímetro: Mide la cantidad de latidos por minuto y la Saturación de Oxígeno en la sangre.



Figura 16: Sensor de pulso y saturación de Oxígeno [40]

Entrega como dato la cantidad de latidos por minuto (BPM por sus siglas en inglés) y el porcentaje de saturación de oxígeno en la sangre.

- Eyetracker: Sensor diseñado digitalmente e impreso en 3D. Tiene una precisión mínima de 60%, su frecuencia de captura de datos es de 120 hz, cámara HD y 5.7ms de latencia. Además la cámara que graba el ojo tiene una resolución máxima de 640x360 @30fps y por una cámara que graba la ubicación de la mirada con una resolución máxima de 1280x720 @30fps. [42]



Figura 17. Eyetracker by pupil labs.

Como se mencionó anteriormente, este sensor tiene un Hardware y Software independiente a los demás. El software denominado “Pupil Capture” se instala en una Tablet Microsoft Surface acondicionada con un sistema operativo Linux Ubuntu 15.04. A través de ese Software se ejecutan los comandos para comenzar a capturar datos.

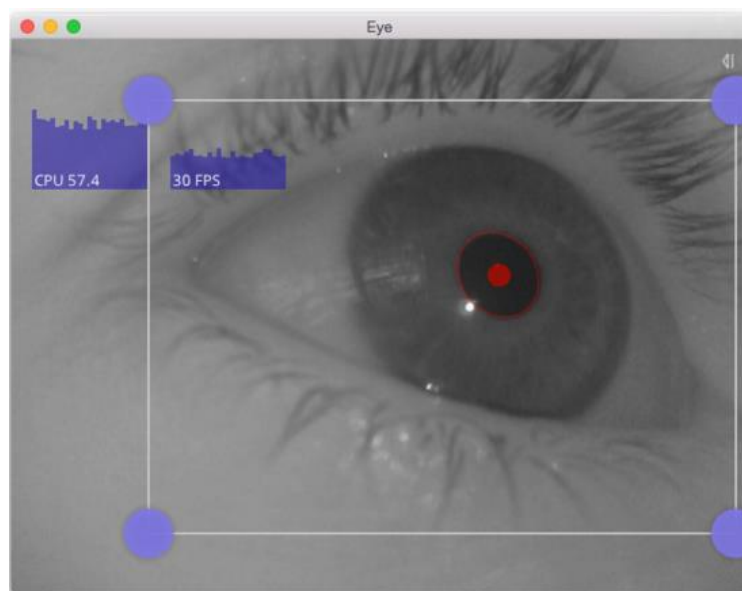


Figura 18. Imagen de la detección pupilar [42]

Además, es importante mencionar los otros elementos claves del experimento, que son:

- Smartphone Samsung Galaxy S5

- Sala de registro

## 3.2 Grupo Experimental

La definición del grupo experimental es un factor crítico para este trabajo. Siguiendo con las recomendaciones dadas en Haapalainen et al [9] y en [28] se tiene lo siguiente:

- Personas saludables. Esto es, que no tengan alguna enfermedad que pueda perjudicar los resultados del experimento
- Además no deben tener antecedentes de enfermedades neurológicas o psiquiátricas tanto ellos como su familia directa
- Sujetos entre 15 y 35 años
- Es importante que sea un grupo homogéneo en cuanto al uso de la tecnología, evitando así resultados que se deban a falta de práctica.
- Sin uso de medicamentos o drogas que puedan alterar el estudio
- Visión correcta, independiente a si usan lentes
- Deben firmar el consentimiento informado declarando todo lo anterior y permitiendo el trabajo con los datos recolectados.

## 3.3 Diseño del experimento

### 3.3.1 Consideraciones iniciales

La hipótesis de esta investigación se basa en considerar dos aspectos que no han sido abarcados por los trabajos anteriores. El primero es examinar escenarios en donde las actividades de interacción humano computador (HCI) sean en un ambiente móvil, y el segundo es que el contexto situacional en que está inmerso el participante debe ser relevante, esto es, importa si el participante está sentado en una oficina versus si está caminando.

Para el diseño de los experimentos se consideran escenarios en donde se ejecuten dos actividades simultáneamente. Una actividad es física (primaria) y la otra es de interacción con un dispositivo móvil (HCI).



Los escenarios están diseñados con el objetivo de que las actividades que los formen sean cotidianamente realizadas por las personas, esto con el fin de acercar esta investigación a posibles desarrollos futuros de aplicaciones u otros.

Además es importante señalar que las tareas están pensadas para estimular el trabajo mental de los participantes y así medir la carga cognitiva que se produce en la resolución de éstas. No se pretende someter a los participantes a estrés, frustración u otros posibles estímulos psicológicos.

Como se vio en el marco teórico, la metodología que siguen los papers estudiados se basa en exponer a los participantes a escenarios de distinta dificultad con el fin de verificar si es que los sensores que utilizan son capaces de discriminarlos, y así concluir sobre qué sensor o grupo de sensores entregan una mayor capacidad de predicción de la dificultad en ese escenario.

Se tiene entonces que la variable independiente de este estudio es la dificultad de la tarea, que se manipulará intencionalmente con el fin de ver los efectos producidos en las variables dependientes que son las señales de cada sensor.

En el caso de esta investigación, los escenarios de distinta dificultad son diseñados en base al modelo de Múltiples Recursos de Wickens [7] variando la cantidad de interferencias que se producen entre las distintas interfaces con lo que a su vez varía la carga cognitiva necesaria para la ejecución de la tarea.

### 3.3.2 Diseño

Este diseño está basado en los papers estudiados y expuestos en el marco teórico, en el trabajo de Ángel Jiménez et al [31], y en el paper [43] donde crean un framework para el correcto diseño de escenarios.

El experimento consiste en la aplicación de tres sets de actividades que estarán compuestos por tareas de dos niveles de dificultad. La idea fundamental es tener escenarios con baja y alta demanda de carga cognitiva con el fin de comprobar si es que las señales de los sensores son estadísticamente distintas y así poder predecir esa dificultad.

A continuación se detallan tanto las actividades físicas como las actividades HCI que se realizarán para luego plantear los sets de actividades y el flujo del proceso.

Las actividades físicas (primarias) a desarrollar son las siguientes:

- En la Oficina: Ambiente cerrado con baja demanda cognitiva. El lugar será la oficina del Profesor Ángel Jiménez en el Departamento de Industrias. La persona

se sentará en el escritorio y comenzará el experimento sin estímulos externos ni interrupciones. La persona solamente debe controlar su espacio personal.

- **Conversación:** En el mismo lugar, la persona debe responder preguntas del examinador y además escuchar música por un audífono en su idioma nativo. En este caso tampoco se requiere movimiento corporal. Sí se requiere poner atención, examinar, analizar y hacer sentido a las preguntas para responder. Además requiere recursos auditivos, visuales y de memoria.
- **Caminar:** La persona deberá salir de la oficina y ejecutar las tareas en el pasillo contiguo a la oficina. Se eligió este espacio debido a que la luminosidad se mantiene constante evitando así interferencias para el Eyetracker. El usuario camina por el pasillo hasta que completa las tareas HCI. Este escenario requiere planear la ruta, tener una noción visual del ambiente, percepción del espacio, control motor.

Las actividades HCI que se desarrollan están pensadas con el fin de acercar el experimento a la cotidianidad de las personas explicitando actividades del diario vivir. Estas actividades son las siguientes:

- **Correo electrónico simple:** El participante debe leer correos electrónicos simples desde la aplicación Gmail del smartphone en una cuenta previamente creada. La persona debe acceder a una aplicación, leer y responder los correos correspondientes. Esta actividad involucra decidir el camino para abrir la aplicación, esperar por la carga, seleccionar el correo, leer, calcular y responder. La idea es replicar correos con preguntas cotidianas. Un ejemplo de correo es:

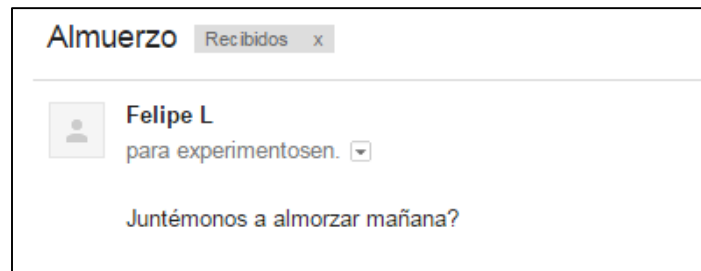


Figura 19 Ejemplo de correo electrónico simple

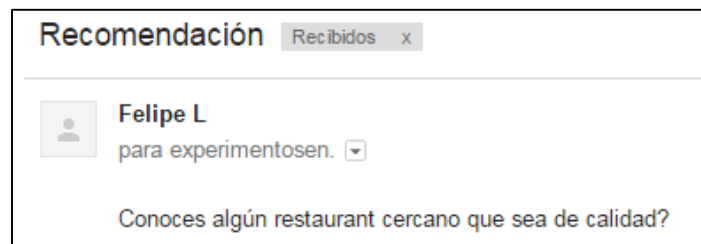


Figura 20 Ejemplo de correo electrónico simple

- Mensajería Instantánea en Hangout de Google: El participante debe responder mensajes del experimentador en la aplicación Hangout de Google. Las preguntas que el examinador realiza tienen como objetivo estimular la memoria de la persona. Ejemplos pueden ser, ¿Qué comió en el almuerzo de ayer? ¿Cuál era su actividad favorita de la infancia?. Esta actividad tiene como subtareas recibir el mensaje, decidir el camino para abrir la aplicación, esperar que cargue, leer el mensaje, analizar, recordar, escribir y esperar.
- Correos de búsqueda de información: El participante debe leer y responder correos electrónicos que le solicitan la búsqueda de precios de hoteles, viajes, autos y nombres de restaurantes. Las aplicaciones a las que debe acceder para encontrar esa información son Despegar, Crediautos y Tripadvisor. En esta actividad la persona debe recibir el mensaje, retener el requerimiento, decidir el camino para abrir el navegador o la aplicación en el móvil, abrir la página, ingresar el requerimiento y volver a escribir la respuesta.

En Despegar se le solicita al participante que busque el precio de un viaje o un hotel para una fecha y un destino indicados. Debe memorizar el requerimiento, acceder a la aplicación, ingresar el requerimiento, esperar la respuesta y responder el correo. No está permitido que el participante vuelva al correo por la información correcta si es que la olvidó.



Figura 21 Capturas de pantalla de app Despegar

A continuación se muestra un ejemplo del correo electrónico tipo.

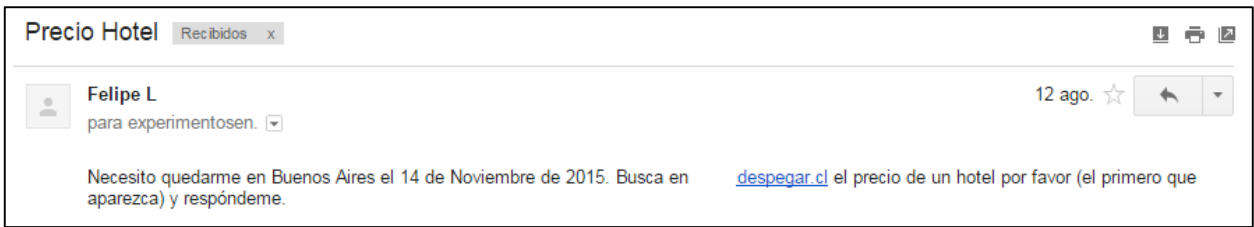


Figura 22 Ejemplo de correo electrónico para app despegar

Para Tripadvisor lo que se solicita es el nombre de restaurantes de cierta especialidad en una determinada ciudad del mundo. El participante debe seleccionar restaurantes, ingresar la ciudad, filtrar el tipo de restaurant y responder con la información adecuada.

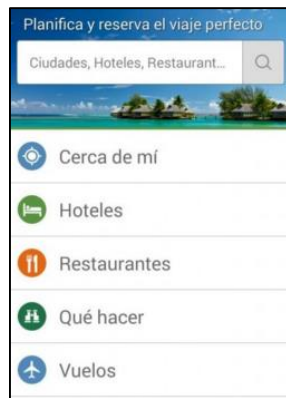


Figura 23 Pantalla inicio app Tripadvisor

A continuación se muestra un ejemplo del correo electrónico tipo.

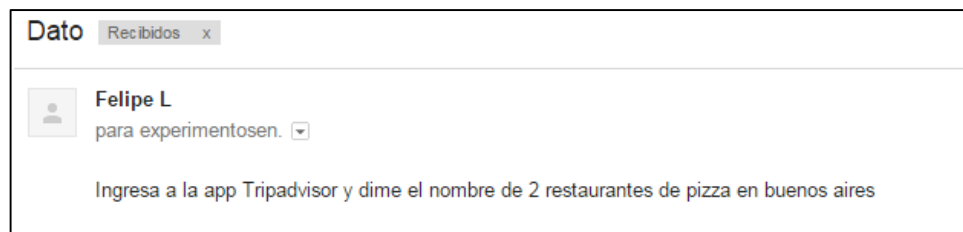


Figura 24 Ejemplo correo para app Tripadvisor

Por último, en Crediautos se le solicita al participante el precio de un vehiculo en que se le otorga la marca y el modelo del mismo. Debe ingresar estos datos en la aplicación y responder con la respuesta adecuada.

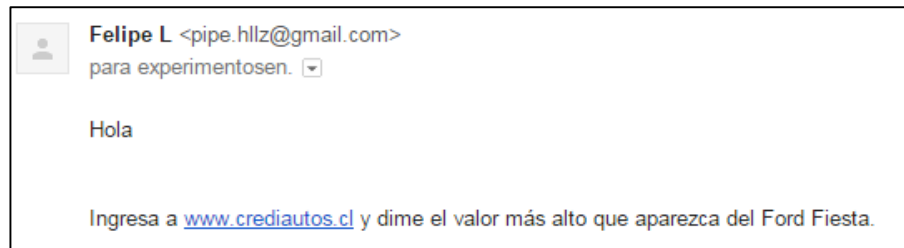


Figura 25 Ejemplo correo electrónico para Crediautos

### 3.3.3 Escenarios

Para cada actividad primaria el participante debe realizar las tres actividades HCI descritas anteriormente.

El experimento consiste en la aplicación de tres sets de actividades en donde para cada set existe un escenario de baja demanda cognitiva, esto es, las interferencias según el modelo de Wickens son mínimas o nulas, y un escenario con alta demanda cognitiva en donde el número de interferencias es mayor.

A continuación se describen las subtareas de cada actividad y las interferencias producidas para cada escenario [7][31]. En rojo se señala la alta demanda de recursos cognitivos, mientras que el color amarillo indica baja demanda.

- Oficina + actividades HCI:

Al estar sentado en un escritorio sin estímulos externos, la actividad física que la persona realiza es mínima y solo debe preocuparse de controlar su espacio personal [29]. Este escenario es el de baja demanda cognitiva. Si a eso le añadimos las subtareas que se requieren para las actividades HCI, según [7] y [31] el requerimiento de recursos es el siguiente (con rojo se denota una alta demanda de recursos y con amarillo una baja demanda):

HCI tasks & processing	Codificación perceptual								Proceso central		Respuesta	
	Atención selectiva				Percepción				Memoria		Responsivo	
	Modalidad de entrada				Codigo procesamiento		Codigo procesamiento		Codigo procesamiento		Codigo procesamiento	
	Audición	Visión focal	Visión ambiental	Tacto	Espacial	Verbal	Espacial	Verbal	Espacial	Verbal	Espacial	Verbal
Escribir		Yellow				Yellow		Yellow		Yellow	Red	
Decidir un camino		Yellow								Red		
Leer		Yellow				Yellow		Yellow		Red		
Esperar por la carga		Red	Red									
Mirar		Red				Yellow		Yellow		Yellow		
Analizar										Red		
Escuchar	Red					Yellow		Yellow		Red		
Sentado (controlar espacio)	Yellow		Yellow		Yellow		Yellow				Red	

Tabla 4. Indica recursos aplicados en Oficina + Actividades HCI

Cada una de estas subtarefas HCI se realizan una tras otra por lo que no entran en conflicto entre ellas. Entre las actividades HCI y la actividad física se ve que tanto Escribir como Controlar espacio ocupan el recurso de respuesta espacial al mismo tiempo por lo que se produce una interferencia de código de procesamiento.

- Conversación + actividades HCI:

Este es el escenario de alta demanda cognitiva. La persona continúa sentada, pero debe realizar las tareas HCI, poner atención a la conversación con el examinador y además tiene como estímulo externo música en su idioma nativo. El esquema de recursos es el siguiente para la actividad física y la actividad HCI:

HCI tasks & processing	Codificación perceptual								Proceso central		Respuesta	
	Atención selectiva				Percepción		Memoria		Responsivo			
	Modalidad de entrada				Codigo procesamiento		Codigo procesamiento		Codigo procesamiento		Codigo procesamiento	
	Audición	Visión focal	Visión ambiental	Tacto	Esp.	Verbal	Esp.	Verbal	Esp.	Verbal	Esp.	Verbal
Escribir												
Buscar en la pantalla												
Decidir un camino												
Leer												
Esperar por la carga												
Mirar												
Analizar												
Escuchar												
Sentado (controlar espacio)												
Hablar												
Analizar												
Escuchar												

Tabla 5 Recursos aplicados a Conversar + Actividades HCI

Se ve que se repite la tarea de analizar y escuchar dado que se tiene que analizar las preguntas en voz alta y además el requerimiento de la actividad HCI, también se debe escuchar las preguntas y escuchar música al mismo tiempo.

En este caso se producen los tres tipos de interferencias planteadas en el modelo de Wickens. Interferencia por el canal al estar escuchando música y conversar en donde la audición es fundamental. Interferencia de código de procesamiento al utilizar el código verbal de la memoria de trabajo para hablar, analizar y escuchar en la conversación además de leer, analizar y calcular los requerimientos de las tareas HCI. Por último se ve que hay interferencia de etapas de procesamiento al existir interferencia entre el uso del control motor y el uso de distintos canales input.

- Caminar + actividades HCI:

En este caso las tareas se realizan mientras la persona camina. Aquí no se tiene el estímulo de la música ni se está sosteniendo una conversación. El esquema de recursos es el siguiente para la actividad física:

HCI tasks & processing	Codificación perceptual								Proceso central		Respuesta	
	Atención selectiva				Percepción				Memoria		Responsivo	
	Modalidad de entrada				Codigo procesamiento		Codigo procesamiento		Codigo procesamiento		Codigo procesamiento	
	Audición	Visión focal	Visión ambiental	Tacto	Esp.	Verbal	Esp.	Verbal	Esp.	Verbal	Esp.	Verbal
Escribir		Yellow				Yellow		Yellow		Yellow	Red	
Buscar en la pantalla		Yellow			Yellow	Yellow		Yellow		Red		
Decidir un camino		Yellow								Red		
Leer		Yellow				Yellow		Yellow		Red		
Esperar por la carga		Red	Red									
Mirar		Red				Yellow		Yellow		Yellow		
Analizar										Red		
Caminar	Yellow		Yellow				Yellow				Red	
Hablar	Yellow	Yellow	Yellow			Yellow		Yellow		Red		Red
Encontrar camino			Yellow				Yellow		Red			
Evitar colisiones			Yellow				Yellow				Red	

Tabla 6 Recursos aplicados a Caminar + Actividades HCI

En este caso vemos que no se producen interferencias por canal. Se produce interferencia de código de procesamiento entre Escribir, evitar colisiones y caminar al utilizar el control motor espacial. Además se producen otras interferencias de etapas de procesamiento. A partir de lo anterior se puede inferir que este escenario posee una demanda cognitiva media entre los otros dos y se verá en el análisis de los resultados si es que es posible comprobar lo anterior.

Luego, las actividades que se aplican son las siguientes

- Actividad 1: Oficina + Correos simples + Correos búsqueda de información (baja demanda cognitiva)



- Actividad 2: Oficina + Hangout (baja demanda cognitiva)
- Actividad 3: Conversación + Correos simples + Correos búsqueda de información (alta demanda cognitiva)
- Actividad 4: Conversación + Hangout (alta demanda cognitiva)
- Actividad 5: Caminar + Correos simples + Correos búsqueda de información (baja y alta demanda cognitiva)

Es importante señalar que se quiere verificar si es que en la actividad de correos de búsqueda de información las señales son capaces de discriminar entre dificultad alta y baja dado que según la teoría es la actividad más difícil de todas debido a que el participante debe, además de las acciones normales a las demás actividades, memorizar requerimientos y responder a estos sin olvidar y sin perder la idea de la conversación.

### 3.3.4 Consideraciones para la ejecución del experimento

En cualquier experimento es clave mantener el control de los escenarios y la validez interna, esto es,

- El participante debe ejecutar las tareas tal como se lo indica el experimentador. Para esto se realiza, previo a la toma de datos real, una actividad de práctica que le enseña a los participantes a usar las aplicaciones y los caminos que deben seguir para cada actividad HCI. Además sirve para que se familiaricen con el Smartphone y evitar datos que no estarían relacionados con la dificultad de la tarea en sí.
- Debe existir una equivalencia entre todos los participantes con el fin de poder compararlos
- El espacio físico debe ser el mismo para todos los participantes y en las mismas condiciones
- El instrumento de medición debe mantener una estabilidad igual para todos los participantes. En este caso, es importante siempre calibrar los sensores antes de cada medición
- Para evitar el factor de aprendizaje de los participantes, al comienzo de cada sesión se debe elegir al azar el orden de las actividades con el fin de que al sumarizar a los participantes se atenúe este factor

- Por último, es fundamental que el comportamiento del experimentador sea equivalente para todos los participantes con el fin de no afectar los resultados.

## 3.4 Implementación

### 3.4.1 Grupo experimental

De acuerdo a los requisitos que se plantearon en el diseño experimental el grupo experimental definitivo fue compuesto de la siguiente manera:

P	G	E	Ocupación
1	M	23	Estudiante
2	M	21	Estudiante
3	F	24	Estudiante
4	M	22	Estudiante
5	F	26	Estudiante
6	F	27	Memorista
7	F	23	Estudiante
8	M	22	Estudiante
9	F	24	Estudiante
10	M	25	Estudiante
11	F	19	Estudiante
12	M	22	Estudiante
13	M	24	Memorista
14	M	20	Estudiante
15	M	21	Estudiante
16	M	22	Estudiante
17	M	21	Estudiante
18	M	25	Memorista
19	F	30	Estudiante
20	F	20	Estudiante

Tabla 7. Detalle grupo experimental por género, edad y ocupación.

La cantidad total de participantes en el estudio es mayor, pero por errores de software o hardware se consideran los 20 que completaron exitosamente el experimento.

### 3.4.2 Procedimiento Experimental

- 1) Al momento de que cada sujeto ingresa a la sala por primera vez se le informa el contexto general del proyecto junto con los objetivos que pretende el experimento
- 2) Además se le pide que complete el documento de Consentimiento informado y que complete un formulario con datos personales básicos
- 3) Se conectan sensores a la tarjeta ehealth, se enciende el dispositivo y se conecta a la raspberry desde terminal
- 4) Conectarse a eyetracker desde Tablet
- 5) Dar indicaciones para poner correctamente el sensor de Electrocardiograma. Poner el sensor de temperatura, y el sensor de oxígeno y pulso cardíaco.
- 6) Poner el bolso con los sensores junto al brazo no dominante
- 7) Verificar que se están transmitiendo datos desde el dispositivo mediante ejecución de comando específico desde la terminal
- 8) Poner el eyetracker en el sujeto y calibrarlo.



Figura 26 Participante utilizando Eyetracker

- 9) Realizar la actividad de prueba
- 10) Pedir al participante que cierre sus ojos durante 90 segundos para que se relaje [5]
- 11) Al comienzo de cada tarea se le entregan las instrucciones correspondientes
- 12) Sacar al azar el orden de las actividades
- 13) Realizar actividad 1<sup>3</sup>

---

<sup>3</sup> Se hace referencia a la primera actividad en realizarse según el orden obtenido. No necesariamente a la actividad 1 expuesta en la sección anterior.



Figura 27 Participante realizando actividades

- 14) Luego de cada actividad los sujetos deben responder un cuestionario NASA que se les suministra con el fin de conocer sus impresiones en términos de: demanda mental, demanda física, demanda temporal, desempeño, esfuerzo y frustración. Este cuestionario es una aplicación web de libre disposición en [www.nasatlx.com](http://www.nasatlx.com)
- 15) Pedir al participante que se relaje por 30 segundos cerrando sus ojos
- 16) Realizar actividad 2
- 17) Cuestionario NASA
- 18) Relajación
- 19) Realizar actividad 3
- 20) Cuestionario NASA
- 21) Relajación
- 22) Realizar actividad 4
- 23) Cuestionario NASA
- 24) Relajación
- 25) Realizar actividad 5
- 26) Cuestionario NASA
- 27) Cuando el sujeto complete el último cuestionario NASA se le retiran todos los sensores y se terminan las mediciones.

### 3.4.3 Formato de Resultados

El Eyetracker entrega los siguientes resultados cada vez que se graba la captura pupilar [42]:

- world.mkv: Video con la dirección de la mirada
- world\_timestamps.npy: Arreglo 1d con los timestamps para cada frame del video
- pupil\_positions.npy: arreglo con la posición de la pupila

- gaze\_positions.npy: arreglo con la posición de la mirada
- info.csv: archivo con meta data

Mediante un procesamiento que se detalla en el capítulo 5 se obtiene la medida de dilatación pupilar.

Para el caso del dispositivo de captura de datos con los otros sensores el formato de resultados es un archivo .csv con:

timestamp	BPM	SO2	TMP	ECG
1439493290	80	96	29.54	1.83714827
1439493291	87	96	28.68	1.84400158
1439493292	88	96	29.48	1.84327632

Donde ECG es el valor analógico de captura de electrocardiograma, BPM es el pulso por minuto, SO2 es la saturación de oxígeno, TMP es la temperatura en grados Celsius.

Timestamp se refiere al timestamp UNIX, que es el número de segundos transcurridos desde el 1 de Enero de 1970.

Los registros del Eyetracker también están etiquetados por timestamp por lo que simplemente se hace un match entre las dos tablas para obtener la base de datos con que se trabajará.

## 4. Validación del diseño experimental

El objetivo de la validación es asegurar que el diseño experimental es apropiado, esto es, que el control de la dificultad de las tareas está teniendo el resultado esperado. Así, se espera que entre las actividades del experimento exista evidencia estadística de que los resultados son distintos.

Para comparar actividades se aplica test t y ANOVA. La hipótesis nula del test t señala que las medias de los grupos analizados son iguales, por lo que si se quiere demostrar que efectivamente los grupos difieren en sus medias se debe poder rechazar la hipótesis nula. A su vez, la hipótesis nula de ANOVA señala que las varianzas de los grupos analizados son iguales, por lo que se debe rechazar esta hipótesis para demostrar que los grupos difieren en sus varianzas. Todos los análisis se realizan con un intervalo de un 95% de confianza.

Esta validación se realiza en dos etapas. La primera fue previa a la ejecución del experimento y la segunda fue con datos anexos recolectados durante las sesiones experimentales.

### 4.1 Validación previa

Se realizaron las actividades con 5 personas que no participaron del experimento final. Durante las sesiones se siguió el mismo protocolo experimental y se aplicó el cuestionario NASA TLX a cada participante. Los resultados se muestran en la tabla siguiente:

	Oficina	Conversar	Caminar
1	45.33	77.79	32.07
2	25.34	38.73	14.73
3	28.47	69.93	23.39
4	34.54	40.27	40.4
5	23.27	69.67	47.46

Tabla 8: Se indica el valor de carga cognitiva según medida entregada por NASA TLX para los tres escenarios del experimento.

En el siguiente gráfico se aprecia la diferencia de carga entre el escenario Oficina versus el escenario Conversar.

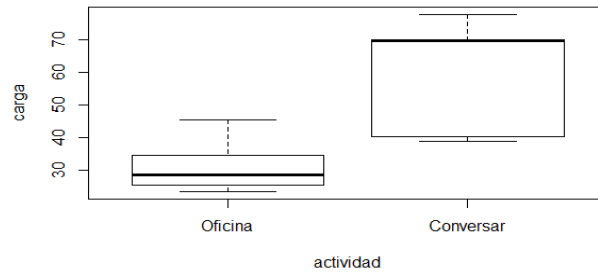


Figura 28 Gráfico que muestra las diferencias en los valores del test NASA

Al aplicar ANOVA y test t a los resultados se obtiene que con el 95% de confianza se rechazan las hipótesis nulas con lo que los grupos se diferencian entre sí en sus medias y varianzas.

F	p-value	t	p-value
9.354	0.0156	-3.529	0.024

## 4.2 Validación posterior

Al igual que en [9] se realiza la validación del experimento mediante tres medidas: NASA TLX, Tiempo de ejecución y Performance de las tareas. Para esto se toman datos recolectados de los 20 participantes en el experimento.

Cada actividad está compuesta por 10 secciones, ya sea correos o preguntas por Hangout. Se tienen los resultados por actividad y se procede a aplicar test t para compararlos. Se va a comparar la actividad 1 con la actividad 3 y la actividad 2 con la actividad 4 y la 1 con la 5.

- Performance: Se considera la performance al porcentaje de éxitos obtenidos en la actividad. Aplicando ANOVA y test t se tiene lo siguiente:

		1 vs 3	2 vs 4	1 vs 5
t	p-value (t)	0.0013	0.009	0.009

Tabla 9 Resultados test t

Por lo tanto, en todos los casos se rechaza la hipótesis nula y se tiene que las actividades son distintas en sus medias.

- Tiempo de ejecución: Se registra el tiempo que demoran en completar las actividades cada uno de los participantes. Aplicando ANOVA y test t se tiene lo siguiente:

		1 vs 3	2 vs 4	1 vs 5
t	p-value (t)	0.0416	0.048	<2.2E-16

Tabla 10 Resultados test t

Por lo tanto, en todos los casos se rechaza la hipótesis nula y se tiene que las actividades son distintas en sus medias.

- NASA TLX: Una vez aplicado el cuestionario NASA a los 20 participantes se tiene lo siguiente con respecto a la dificultad de las actividades:

		1 vs 3	2 vs 4	1 vs 5
t	p-value (t)	9.95E-09	9.84E-09	2.42E-07

Tabla 11 Resultados test t

Por lo tanto, en todos los casos se rechaza la hipótesis nula y se tiene que las actividades son distintas en sus medias.

Se concluye entonces que el diseño experimental es apropiado dado que las actividades difieren estadísticamente de acuerdo al grado de dificultad que se implementó.



# 5. Análisis y resultados

En este capítulo se presentan los resultados obtenidos del experimento. Para esto se siguen las etapas del modelo KDD con el fin de detallar tanto la preparación de los datos como la aplicación de modelos estadísticos y de minería de datos.

## 5.1 Definición de observaciones

Para analizar los datos obtenidos es necesario tener claro cómo se definirá una observación en este trabajo previo al procesamiento de los mismos. Durante las tareas HCI se deben realizar variadas subtareas en pos de cumplir el objetivo planteado. Ejemplos de subtareas son: Desbloquear el Smartphone, Buscar la aplicación, Acceder, Ingresar al Correo, Leer Correo, etc. De todas las subtareas no se está interesado en describirlas una por una, sino que interesa estudiar a las que según la teoría tienen más recursos cognitivos involucrados tal como se explicó en la sección del diseño experimental. Bailey e Iqbal [44], señalan que la carga cognitiva baja en determinados cambios entre subtareas y en subtareas de transición como por ejemplo acceder a alguna aplicación. Es por esta razón que se definieron las siguientes subtareas que interesan medir:

- 1) Leer correo electrónico simple
- 2) Responder correo electrónico simple
- 3) Leer correo electrónico de búsqueda de información
- 4) Buscar la aplicación correspondiente
- 5) Ingresar la información requerida
- 6) Leer información obtenida desde aplicación
- 7) Responder correo electrónico de búsqueda de información
- 8) Leer pregunta en Hangout
- 9) Responder pregunta en Hangout
- 10) Leer correo electrónico durante caminata

De esta forma se acotan los subconjuntos de datos y se hacen comparables las subtareas para los distintos escenarios. Por ejemplo, se pretende comparar la subtarea 1 en el escenario sin estímulos externos con la misma tarea 1 en el escenario de conversación y música. Así se realiza para las demás subtareas sucesivamente.

Además es importante señalar que no se comparará la actividad de caminar con las otras debido a que la actividad física va a ocasionar una variación en las señales fisiológicas

que no se puede interpretar como cambio en la carga cognitiva. Luego, esta actividad se comparará con si misma interactuando con correos fáciles y con correos difíciles.

### 5.1.1 Etiquetado de Subtareas

En la práctica, el proceso de etiquetado de los registros por subtarea es bastante lento. No se encontró una aplicación que realizara un tracking del comportamiento del usuario en el Smartphone al nivel de detalle que se necesita por lo que todas las etiquetas se hicieron de forma manual. Para esto se tiene un archivo obtenido desde el software Pupil Capture llamado timestamps.npy (la extensión .npy hace referencia a una librería de Python creada especialmente para el análisis de datos). Este archivo contiene los timestamps por cada frame del video por lo que se necesita entonces, conocer el frame de inicio y fin de las subtareas. Entonces, mediante un programa de edición de video se corre el video generado por Pupil Capture de la grabación hecha por el Eyetracker en donde se ve qué es lo que está mirando el participante y en particular se ve el Smartphone, obteniendo así para cada video los frame de inicio y fin de cada subtarea. Se obtienen los timestamps y se etiquetan en la base de datos usando R.

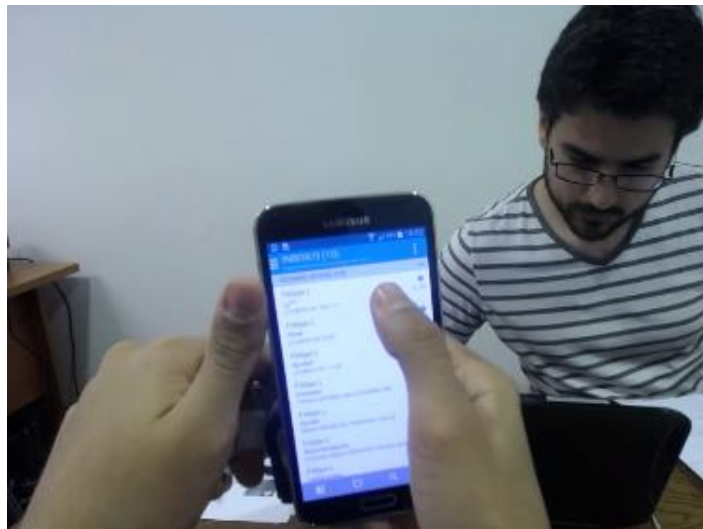


Figura 29 Captura de pantalla de video de Eyetracker

## 5.2 Procesamiento de datos

Como se señaló en el capítulo 2, la dilatación pupilar, electrocardiograma, pulso y temperatura han demostrado en trabajos anteriores ser buenos predictores del nivel de carga cognitiva. A continuación se detalla el procedimiento para procesar cada señal.

### 5.2.1 Dilatación pupilar

#### 5.2.1.1 Selección

Mediante el software Pupil Capture del Eyetracker se obtienen archivos de extensión .npy para las tareas de cada participante. La extensión .npy refiere a Numpy Python, una librería en lenguaje Python diseñada especialmente para el procesamiento de datos.

Se utiliza un script de Python para juntar todos los archivos de un mismo individuo y obtener la dilatación pupilar sin procesar. Este nuevo archivo contiene la dilatación pupilar medida en píxeles, el timestamp de la medición y la confiabilidad de la medida. Se eliminan registros que no corresponden a una actividad, así como errores que se hayan producido durante la sesión experimental.

#### 5.2.1.2 Pre-procesamiento

Un siguiente paso para trabajar con los datos es eliminar los artefactos que distorsionan las mediciones.

- Sacadas. Las sacadas son movimientos repentinos del ojo. Es un fenómeno normal, pero que distorsiona el ángulo en que el eyetracker debe medir el diámetro pupilar con lo que se pierde confiabilidad en la medida. Como el software del eyetracker está programado en código libre se tiene un script incorporado que filtra las sacadas con lo que se deben eliminar los registros si no cumplen ciertos criterios de confiabilidad.
- Pestañeos. Al pestañar, el eyetracker pierde los registros debido a que el párpado bloquea la luz infrarroja que detecta a la pupila. Como el intervalo de tiempo en

que ocurre este evento es de pocos milisegundos, se recomienda realizar una interpolación lineal entre los dos valores adyacentes al pestañeo.

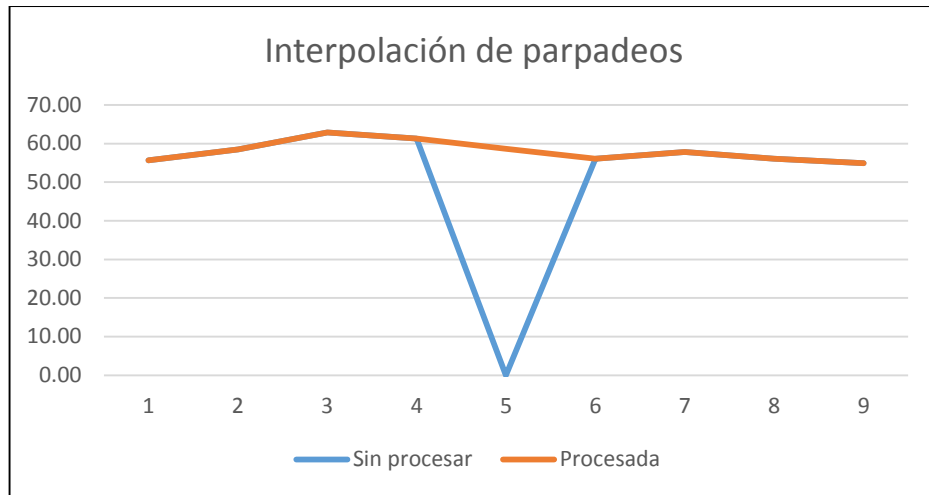


Figura 30 Interpolación de parpadeos

- Filtro de Frecuencias altas. Se utiliza un script en Python que filtran frecuencias que escapan de la media de la muestra y que por lo tanto se asume que no tienen relación con la actividad en sí.

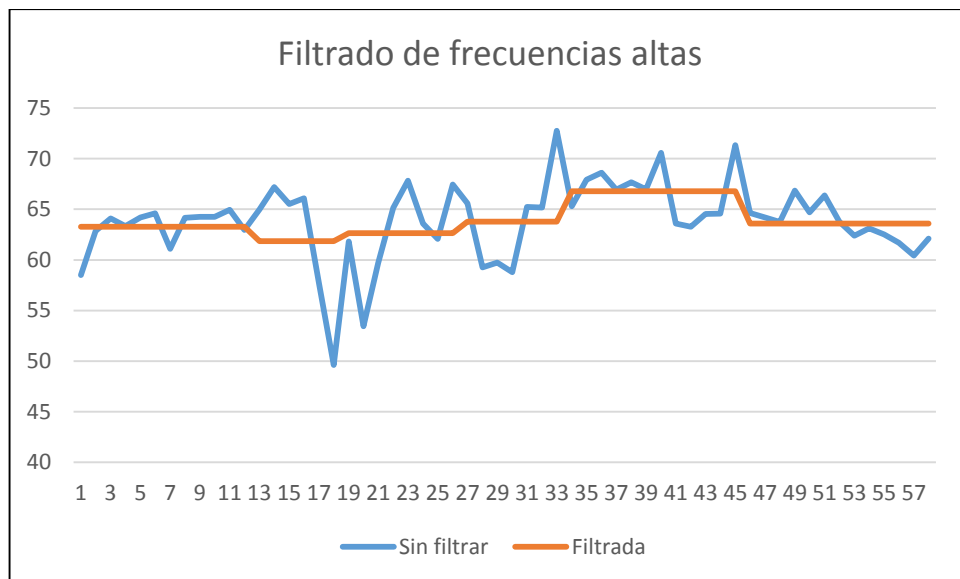


Figura 31 Filtrado de frecuencias altas

- Eliminación de registros. Por la naturaleza móvil de este experimento existen otros factores que provocan que caiga la confiabilidad de las mediciones entre los que se cuentan el tamaño de las pestañas y el tamaño de los párpados que varía

entre cada persona. Debido a eso se eliminan todos los registros que tengan una confiabilidad menor al 60%.

Luego de eliminados los ruidos, se debe etiquetar cada registro según la actividad y la subtarea a la que corresponde tal como se explicó en el capítulo 3.

### 5.2.1.3 Transformación

La medida que interesa es la variación del tamaño de la pupila por lo que se define una línea base tal como se realiza en [28]. Esta base es el tamaño pupilar de 3 segundos luego de que el participante termina el proceso de relajación previo a cada actividad. A partir de eso se calculan las demás medidas.

También para hacer comparables a todos los participantes se normalizan los registros. Se utiliza para esto una transformación Z-score que calcula la media y la varianza de cada una de las observaciones y se transforman los valores en un conjunto de datos que posee una distribución normal estándar de promedio 0 y varianza 1 [37].

## 5.2.2 Electrocardiograma (ECG)

### 5.2.2.1 Selección

Como se mencionó en el capítulo 3, el ECG forma parte del Dispositivo de Captura de Datos por lo que los registros se almacenan en un servidor virtual. Se debe acceder a la máquina virtual y descargar los registros. Está configurado para obtener archivos .csv. La señal en bruto presenta interferencias y ruidos que son necesario atenuar.

### 5.2.2.2 Pre-procesamiento

En primer lugar es necesario aplicar a la señal la transformada rápida de Fourier. Esto se hace para cambiar el espacio de las muestras desde el espacio tiempo a frecuencias y así poder eliminar frecuencias altas que no corresponden a lo que se está midiendo. Además, resulta particularmente útil en para este experimento dado que mientras se

camina se producen muchas interferencias en este sensor. Este proceso se realiza en bloques de 500 registros a la vez.

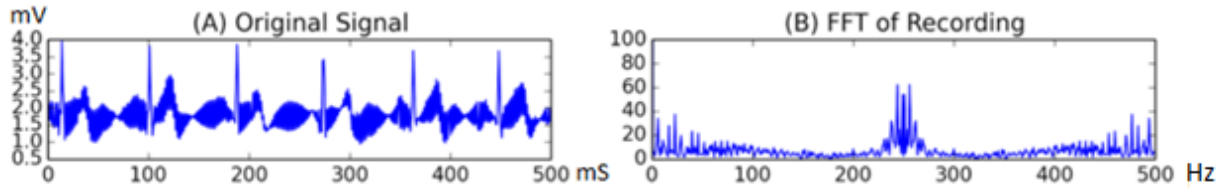


Figura 32 : (A) Señal original; (B) Señal luego de aplicación de transformada de Fourier.

En (B) se puede apreciar que existe un ruido en torno a los 250 Hz. Para eliminar las frecuencias altas se aplica un Filtro pasa bajo de 110 Hz como se muestra en (C). Una vez hecho el filtro se aplica la transformada inversa de Fourier con lo que se obtiene una señal limpia y apta para continuar con el trabajo.

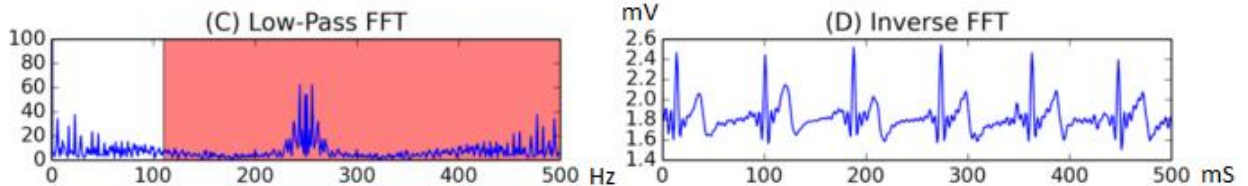


Figura 33 : (C) Aplicación de filtro pasa bajo. (D) Señal luego de la transformada inversa de Fourier.

Luego se procede a etiquetar cada registro de acuerdo al participante, a la actividad y a la subtarea correspondiente usando los registros del software del eyetracker como se explicó anteriormente. Cabe señalar que en ciertos períodos de tiempo para el participante 4 y el participante 12, el ECG entregó un valor constante. Estos registros se eliminaron, pero se continúa analizando a esos participantes dado que corresponden solo al 8% de la cantidad de datos para una actividad en particular.

### 5.2.2.3 Transformación

Para trabajar con la data de ECG, Haapalainen et al. [9] recomienda calcular la desviación absoluta media del ECG (ECGMAD). Señala que en su experimento el ECGMAD junto con el flujo de calor predijeron en un 81.1% la carga cognitiva. Para esto a cada registro se le aplica la fórmula:

$$ECGMAD = |ECG - 1/n(\sum_{i=1}^n ECG)|$$

Con n el total de registros para una actividad.

## 5.2.3 Pulsoxímetro y temperatura

### 5.2.3.1 Selección

Al igual que el sensor ECG, los sensores Pulsoxímetro y temperatura pertenecen al Dispositivo de captura de datos. Se descargan los archivos .csv individuales de cada señal desde el servidor virtual y se separan las porciones de datos que corresponden a cada participante. De estos sensores se tiene los latidos por minuto (BPM, por sus siglas en inglés), saturación de oxígeno (SO<sub>2</sub>) y temperatura en grados celsius.

A diferencia de los sensores anteriores, aquí hubo mayor incidencia de valores perdidos ya sea por desconexión del sensor pulsoxímetro durante algún instante del experimento o errores del sensor de temperatura. Aun así, la completitud de la data del sensor de temperatura y pulsoxímetro es del 96% y 98% respectivamente.

### 5.2.3.2 Pre-procesamiento

En el caso del sensor de pulsoxímetro, cuando los BPM superan la barrera de 100 se copia al servidor un registro de 700. Es decir, si el pulso es 101 latidos por minuto, en el servidor está el registro de 707 latidos por minuto. El equipo de soporte técnico intentó arreglar este problema, pero no fue posible por lo que se hicieron todas las pruebas necesarias para llegar a la conclusión de que lo mejor es procesar esos registros una vez realizada la muestra. A raíz de aquello se creó un script simple en Python que soluciona el problema.

No es necesario ningún otro tratamiento especial para estos dos sensores a diferencia de los anteriores. Las labores de preprocesamiento son las habituales en todo proceso KDD, esto es: Eliminación de Outliers, Sustitución de valores perdidos mediante interpolación lineal o por la media según sea el caso. En el caso del sensor de temperatura se encontraron bloques de datos fuera de rango que se sustituyeron por la media del participante en la subtarea correspondiente.

Al igual que los sensores anteriores, luego se procedió a realizar un consolidado con los registros etiquetados por timestamps, actividad y subtarea a la que pertenecen usando los registros del software del Eyetracker.

### 5.2.3.3 Transformación

En este caso la transformación utilizada también es la normalización que se aplicó para los registros de dilatación pupilar con el fin de eliminar las diferencias entre usuarios ya que existe un amplio rango en que las medidas fisiológicas pueden considerarse normales.

## 5.3 Análisis Estadístico

El propósito de este apartado es presentar estadísticos descriptivos de las señales obtenidas durante el experimento y además, comprobar si los sensores son capaces de diferenciar grupos de actividades con distinto nivel de dificultad mediante test de hipótesis.

### 5.3.1 Análisis exploratorio de los datos

La tabla n° 12 contiene la media, desviación estándar y varianza de cada uno de los participantes en las actividades ejecutadas. El tipo de actividades se encuentra agregado por nivel de dificultad.



		BPM		SO2		TMP		Dilatación		ECGMAD	
		Baja	Alta	Baja	Alta	Baja	Alta	Baja	Alta	Baja	Alta
1	M	91.00	94.50	96.10	97.50	29.40	27.70	60.50	68.30	0.03	0.02
	SD	4.30	3.70	0.82	0.60	0.83	0.05	5.22	6.50	0.04	0.02
2	M	65.50	69.10	95.80	96.70	28.40	27.90	52.40	66.20	0.04	0.03
	SD	3.90	4.70	0.40	0.60	1.01	1.03	4.04	4.30	0.03	0.03
3	M	83.94	85.56	97.34	98.06	30.87	28.9	72.09	75.55	0.02	0.01
	SD	5.15	5.49	0.60	0.32	7.62	4.58	6.66	0.02	0.01	0.01
4	M	70.78	75.89	96.33	98.05	27.76	27.72	65.91	74.39	0.04	0.04
	SD	3.15	4.26	0.55	0.43	0.05	0.06	4.64	5.33	0.05	0.04
5	M	60.41	65.40	94.17	98.61	27.75	27.71	66.40	68.85	0.08	0.11
	SD	1.75	3.91	10.29	0.49	0.05	0.05	4.27	7.75	0.04	0.01
6	M	83.69	84.89	98.50	98.63	27.66	27.75	61.23	62.18	0.03	0.02
	SD	3.55	5.19	0.49	0.80	0.06	0.05	3.52	4.43	0.02	0.03
7	M	75.11	78.15	98.82	98.85	27.74	27.73	77.94	74.53	0.09	0.02
	SD	4.54	5.83	0.39	0.41	0.05	0.05	8.50	6.65	0.17	0.01
8	M	87.77	89.06	95.74	96.18	27.73	27.73	80.07	73.26	0.02	0.04
	SD	3.45	4.48	1.12	1.89	0.05	0.05	8.63	13.59	0.04	0.04
9	M	74.52	76.38	95.56	97.02	27.77	27.74	70.93	68.86	0.20	0.13
	SD	3.59	4.29	1.31	2.11	0.23	0.05	8.45	6.82	0.24	0.20
10	M	86.27	88.06	92.19	95.47	27.73	27.72	90.85	98.69	0.02	0.02
	SD	3.79	6.28	2.43	4.32	0.05	0.05	4.78	6.12	0.01	0.01

Tabla 12 Media y varianza primeros 10 participantes

En la mayoría de los casos existe tendencia a que la media de los grupos varíe, ya sea aumentando o disminuyendo su valor.

Tanto la señal de BPM, de dilatación pupilar, de SO2 y temperatura mantienen un comportamiento estable entre ambos grupos, esto es, siempre en las actividades de mayor dificultad la media de BPM, SO2 y dilatación es mayor mientras que la de temperatura es menor.

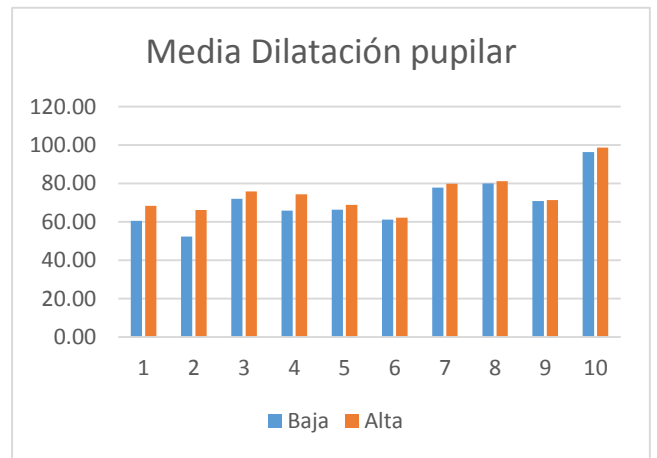
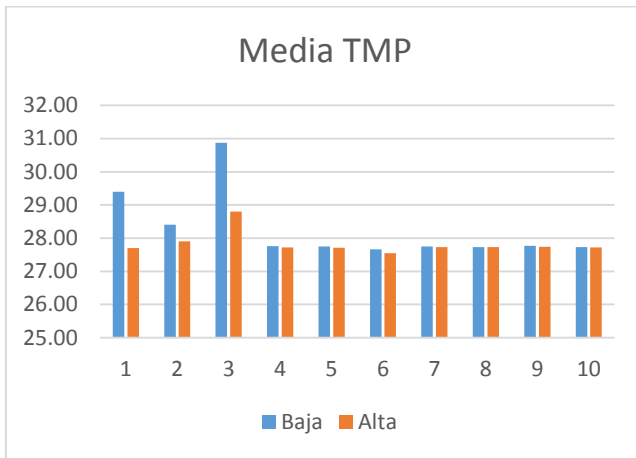
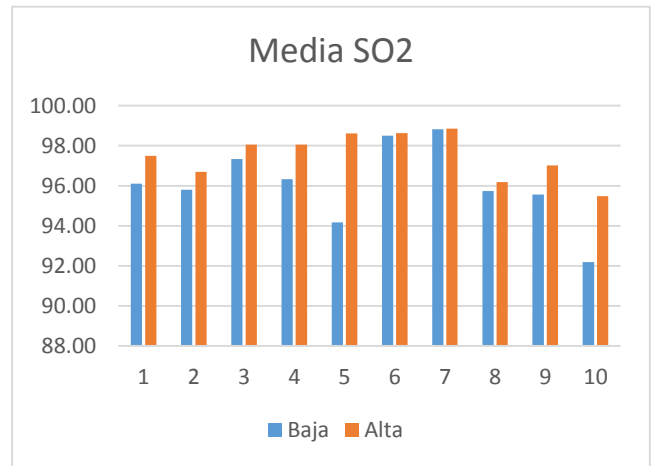
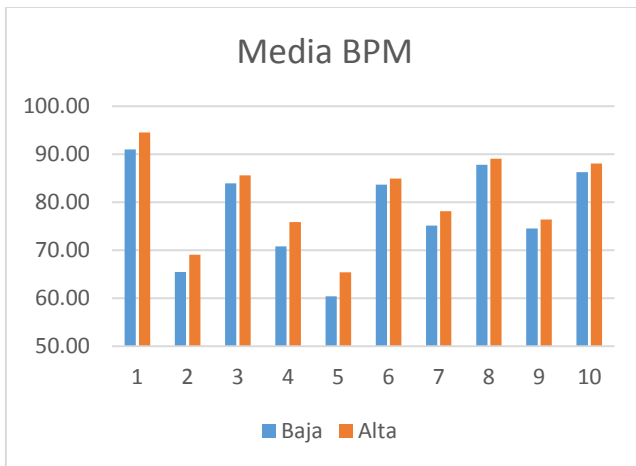


Figura 34 al 37. Medias para 4 señales de 10 participantes de acuerdo a la dificultad de la tarea.

Para el caso de ECGMAD no existe un comportamiento claro. En la mayoría de los casos disminuyen sus medias, pero hay otros en que se mantienen o aumentan.

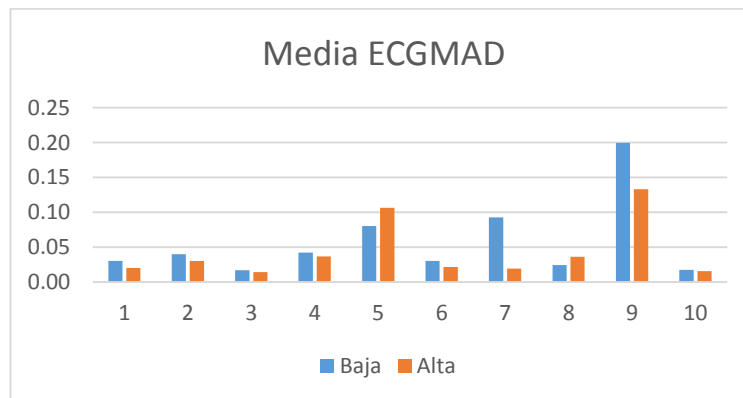


Figura 38 Media ECGMAD para 10 participantes

A continuación se muestra una tabla con la varianza entre participantes para cada sensor.

	BPM		SO2		TMP		Dilatación		ECGMAD	
	Baja	Alta	Baja	Alta	Baja	Alta	Baja	Alta	Baja	Alta
VAR	111.0	123.2	13.8	39.0	13.9	6.9	193.4	1	0.01	0.01

Tabla 13 Varianza entre participantes por señal y dificultad

Se observa que la varianza entre participantes es alta para las señales de BPM, SO2, TMP y Dilatación. Es claro que las respuestas fisiológicas varían de persona a persona, en donde los niveles normales cubren un rango amplio. El nivel de latidos por minuto que para una persona puede representar un nivel de carga cognitiva alto para otras puede ser su nivel base. Es por esto que para analizar a todos los participantes es necesario normalizar las señales como se mencionó en el apartado anterior.

### 5.3.2 Test de Hipótesis

El objetivo de esta sección es comprobar estadísticamente si las señales son capaces de discriminar entre actividades con distinta dificultad.

En primer lugar se debe señalar que el total de datos tienen distribución normal por lo que se puede aplicar un test de hipótesis t y ANOVA. Esto se comprobó aplicando una función en Rstudio, suite de programación del lenguaje R, que se utiliza de ahora en adelante.

Se debe señalar que como a todos los participantes se les aplican las mismas actividades, sin diferenciarlas por grupos, es necesario aplicar ANOVA de medidas repetidas con el fin de analizar la muestra intra-sujetos. La hipótesis nula es la misma de un ANOVA de una vía, pero en este caso se incorpora el error asociado a la variación entre los participantes.

Para realizar el análisis, la data se segmentó por subtareas con el fin de acotarla lo más posible y que las comparaciones entre grupos sean atingentes. Así, por cada señal se comparan las subtareas 1 de dificultad baja y alta, las subtareas 2 y así sucesivamente.

Para recordar lo visto en la sección 5.1, la numeración de las subtareas es la siguiente:

1. Leer correo simple
2. Responder correo simple
3. Leer correo de búsqueda de información
4. Buscar aplicación correspondiente
5. Ingresar requerimiento

6. Leer y memorizar información
7. Responder correo de búsqueda de información
8. Leer pregunta Hangout
9. Responder en Hangout

Para la señal de BPM se tiene lo siguiente con respecto a ANOVA de medidas repetidas:

BPM	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
F	209.2	1470	93.18	77.18	912.8	652.3	727.2	85.02	99.01
p-value	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16

Por lo tanto, con un intervalo de confianza del 95% se rechaza la hipótesis nula para todas las subtareas utilizando la señal de BPM

Los resultados del test t son los siguientes:

BPM	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
p-value (t)	6.31e-16	<2e-16	2.58e-09	0.05089	<2e-16	<2e-16	<2e-16	1.31e-07	2.50e-07

También se rechaza la hipótesis nula para todas las subtareas.

Para la señal de SO2 se tiene lo siguiente con respecto a ANOVA de medidas repetidas:

SO2	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
F	51.02	459.2	261.7	134.8	727.8	291	446.8	364	431.9
p-value	1.82E-12	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16

En este caso se ve que se rechaza la hipótesis nula para todas las subtareas.

Los resultados del test t son los siguientes:

SO2	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
p-value (t)	<2e-16	<2e-16	0.05758	0.939	0.01131	0.0499	0.01951	<2e-16	<2e-16

También se rechaza la hipótesis nula en todas las subtareas salvo en la número cuatro.

Para la señal de temperatura de tiene lo siguiente con respecto a ANOVA de medidas repetidas:

TMP	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
F	48.71	407.1	39.46	16.16	157.1	24.38	27.81	364.6	401.2
p-value	5.58E-12	<2e-16	5.00E-10	6.45E-05	<2e-16	8.88E-07	1.48E-07	<2e-16	<2e-16

En este caso se rechaza la hipótesis nula para todas las subareas.

Para el test t se tiene lo siguiente:

TMP	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
p-value (t)	5.16E-06	<2e-16	1.06E-05	0.004099	<2e-16	0.0005591	3.15E-09	<2e-16	<2e-16

Se rechaza la hipótesis nula de todas las subareas.

Para la dilatación pupilar se tiene lo siguiente con respecto a ANOVA de medidas repetidas:

EYE	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
F	0.876	36.49	137.2	84.96	377.7	163.9	105.2	23.8	50.89
p-value	0.349	1.71E-09	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	1.13E-06	1.77E-12

También se rechaza la hipótesis nula para todas las subareas.

Los resultados del test t son los siguientes:

EYE	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
p-value (t)	0.021	4.65E-13	<2e-16	<2e-16	<2e-16	<2e-16	<2e-16	1.61E-05	1.40E-07

Por lo tanto se rechaza la hipótesis nula para todas las subareas.

Por último, los resultados del ANOVA de medidas repetidas de la señal de ECGMAD:

ECG	1 vs 1	2 vs 2	3 vs 3	4 vs 4	5 vs 5	6 vs 6	7 vs 7	8 vs 8	9 vs 9
F	30.25	110.6	2.319	2.493	1.598	3.48	29.52	12.29	12.78
p-value	4.90E-08	<2e-16	0.128	0.115	0.206	0.0623	6.19E-08	0.000461	0.000326

En este caso no hay evidencia para rechazar la hipótesis nula en las subtareas 3, 4, 5 y 6. En el resto se rechaza la hipótesis nula.

Luego, de los resultados mostrados se puede señalar que los sensores usados en esta investigación son capaces de diferenciar subtareas de distinta dificultad ya que las señales capturadas de esas subtareas varían significativamente.

Las señales de pulso, saturación de oxígeno, temperatura y dilatación pupilar son capaces de discriminar entre todas las tareas de acuerdo al grado de dificultad. Para la señal de ECG, el test de hipótesis concuerda con el análisis exploratorio dado que en varias actividades no se observa un comportamiento concluyente.

También se realizó un test de hipótesis por participante. En este caso al reducirse la cantidad de registros la distribución que sigue no es normal por lo que se debe realizar un test de Wilcoxon. Este test tiene la misma hipótesis nula que el test t ajustándose a otro tipo de distribuciones.

Los resultados obtenidos por el test también indican que se puede diferenciar entre ambos niveles de dificultad para cada subtarea de cada participante. A modo de ejemplo se detallan los resultados de un participante para la subtarea 3:

3 vs 3	BPM	SO2	TMP	Dilatación	ECG
p-value	0.003337	0.0003025	9.37E-05	0.007184	0.02162

## 5.4 Minería de datos

Este apartado tiene como objetivo mostrar el proceso que se debe seguir para aplicar un algoritmo de minería de datos. Para eso se detallan las características (features en inglés), su selección y el modelamiento que se lleva a cabo con el objetivo de predecir la dificultad de las tareas en cada actividad.

### 5.4.1 Features

En la literatura existen variadas formas de extracción de features. En los primeros estudios se limitaban a estudiar solamente las señales recibidas directamente del sensor. A medida que se fue profundizando más se probaron diversas features que tuvieron resultados dispares en cuanto a su capacidad de predicción. Para este trabajo, se siguen las features propuestas por Haapalainen et al.[9] dado que forma parte de la metodología y este modelamiento se basa en ese trabajo.

En la tabla 14 se muestran las features que se analizarán en el siguiente apartado.

Señal	Features
BPM	Media, Mediana, Varianza, Mínimo, Máximo, Raíz cuadrada de dos pulsos sucesivos
SO2	Media, Mediana, Varianza
TMP	Media, Mediana, Varianza
Dilatación pupilar	Media, Mediana, Varianza, Max. y Min. Tamaño pupilar Diferencia respecto de la línea base
ECGMAD	Media, Mediana, Varianza

Tabla 14 Features iniciales por señal

### 5.4.2 Selección de features

La idea principal es que de las 21 features presentadas algunas no explicarían de buena forma el modelo por lo que se deben descartar. Se debe comprobar si la idea es cierta aplicando algoritmos de filtrado. Por lo tanto se intenta obtener el subconjunto de las features mostradas en la sección anterior que estén más correlacionadas con la variable independiente.

La metodología es la siguiente [45]:

- En primer lugar se realiza un Análisis de Componentes Principales cuyo objetivo es agregar al modelo las features que más varianza expliquen.
- Los primeros cinco componentes explican aproximadamente el 90% de la varianza. El sexto componente solo explica el 3% por lo que se elimina.

- Se extraen las variables eigenvalues mayores a 0.05 para los componentes principales que precisamente coinciden con las señales del estudio. En este punto se tiene un total de 15 features desde BPM, SO2, TMP, Dilatación pupilar y ECG.
- Se crea la matriz de correlación a esas 15 features
- Los resultados son pares de features con su correlación correspondiente. Si la correlación entre un par es mayor a 0.6 se mantiene solo una de las dos features.
- Finalmente se obtiene el conjunto de features a modelar. Se muestran en la tabla 15.

Señal	Feature
BPM	Media
SO2	Media
TMP	Media
Dilatación pupilar	Media, Varianza
ECG	Media, Mediana, Varianza

Tabla 15 Features finales por señal

### 5.4.3 Modelamiento

Se utilizan 3 algoritmos de Minería de datos. Support Vector Machine, Naive Bayes y Regresión Logística. Estos son los algoritmos que más se han utilizado para analizar los modelos de carga cognitiva.

SVM es el modelo más utilizado en la actualidad, tiene una fuerte base teórica, además en este caso los parámetros no son tantos por lo que es posible ir ajustándolos por inspección. El kernel a utilizar es lineal.

Regresión también ajusta bien a este tipo de problema. Es de fácil implementación y permite predecir funciones discretas, suficiente para este modelo.

Naive Bayes ha sido el modelo más utilizado y en particular lo utilizan en [9] por lo que es interesante implementarlo para poder realizar comparaciones.

Para poder hacer comparable a los 3 modelos, todos los algoritmos son entrenados y testeados con las mismas features, el mismo número de registros y la misma proporción de datos para entrenar y para testear. Se crean archivos con los valores de las features etiquetadas según el nivel de dificultad. Si el número de registros de una etiqueta es mayor



a los de la otra, se crea un archivo compuesto aleatoriamente por un subconjunto de esos registros para que cada etiqueta contenga la misma cantidad de datos. Se pretende saber cuán preciso es cada sensor al predecir la dificultad de la tarea.

Se desarrollan los modelos siguiendo dos enfoques obtenidos desde la literatura, el primero desde Fritz et al[28] en donde se enfocan cada tarea para todos los participantes al mismo tiempo y el segundo de Haapalainen et al [9] en donde aplican los algoritmos a cada participante por separado.

### 5.3.3.1 Por tarea

Para este caso se crean los archivos correspondientes a las subtareas del experimento y se proceden a entrenar los modelos. La proporción de datos es 70% de entrenamiento y 30% de testeo. La idea es comparar una misma subtarea en el escenario de baja dificultad y en el de alta dificultad y observar el poder de predicción cada modelo. Por ejemplo, interesa comparar la subtarea 1 en el escenario en que no se tienen estímulos externos con la misma subtarea 1 con estímulos externos (conversación y música). Tal como se explicó anteriormente los datos de los participantes se normalizaron minimizando así los efectos de la varianza entre participantes.

- Support Vector Machine

En la tabla 16 se detallan los resultados del modelo para todos los participantes por subtarea. Se presentan los ratios de evaluación del modelo, accuracy, precisión y recall. Se incluyen las dos mejores señales que pueden discriminar el nivel de dificultad con gran eficacia y la combinación de las dos mejores señales individuales junto con la siguiente mejor pareja de señales.

	Sensor	Accuracy	Precision	Recall
Subtarea 1	BPM	74.49%	71.72%	80.87%
	SO2	68.12%	61.68%	95.65%
	BPM+SO2	76.94%	69.53%	81.79%
	BPM+EYE	75.76%	82.65%	74.86%
Subtarea 2	BPM	76.00%	75.46%	77.10%
	TMP	71.02%	65.98%	86.84%
	BPM+TMP	78.16%	76.40%	81.52%
	BPM+SO2	78.15%	77.71%	78.98%
Subtarea 3	BPM	62.41%	60.82%	70.08%
	SO2	65.66%	62.24%	82.21%
	BPM+SO2	69.85%	70.00%	85.08%
	BPM+TMP	63.84%	61.25%	75.34%
Subtarea 4	SO2	63.15%	60.13%	78.02%
	EYE	64.87%	60.95%	82.76%
	SO2+EYE	71.88%	70.45%	86.86%
	TMP+EYE	67.23%	62.89%	84.03%
Subtarea 5	BPM	72.07%	72.47%	71.24%
	SO2	67.76%	63.97%	81.43%
	BPM+SO2	72.40%	73.33%	83.97%
	BPM+TMP	72.11%	71.98%	73.41%
Subtarea 6	BPM	74.11%	79.88%	64.44%
	SO2	70.17%	67.42%	78.04%
	BPM+SO2	78.01%	79.61%	70.09%
	BPM+EYE	78.88%	82.40%	73.40%
Subtarea 7	BPM	72.03%	79.38%	59.53%
	SO2	67.19%	65.28%	73.44%
	BPM+SO2	80.98%	85.33%	69.52%
	BPM+EYE	74.65%	81.09%	64.23%
Subtarea 8	SO2	68.53%	62.81%	89.14%
	TMP	65.36%	60.63%	85.38%
	SO2+TMP	75.83%	74.73%	77.33%
	SO2+ECG	71.13%	66.56%	85.03%
Subtarea 9	SO2	67.84%	63.96%	80.08%
	TMP	65.50%	61.73%	79.53%
	SO2+TMP	72.36%	69.31%	77.21%
	SO2+ECG	71.82%	66.87%	86.53%
Subtarea 10	BPM	56.73%	52.08%	66.89%
	SO2	56.46%	52.44%	67.11%
	BPM+SO2	57.65%	55.90%	61.24%
	BPM+TMP	56.70%	55.43%	68.68%

Tabla 16 Resultados modelo SVM

Se aprecia la predominancia del sensor BPM y SO2 en el modelo. En cada uno de los casos el accuracy de predicción es mayor con los dos sensores juntos que cada uno por separado, llegando incluso al 80.98% en la subtask 7. En la mayoría de los casos la mejor pareja de sensores está compuesta por los mejores sensores individuales salvo en la subtask 6 en donde la combinación BPM+EYE alcanza un 78.88% de eficacia, superando a BPM+SO2. Además los Recall de los modelos son en gran parte mayores que 70% lo que indica que el número de falsos positivos es minoría con respecto al total.

La subtask con menor coeficiente de predicción es la número 10 en donde se responden correos electrónicos mientras se camina. Solo se alcanza el 57.65% de aciertos cuando se combina el sensor BPM con el sensor SO2. Este resultado es consistente con las dificultades de procesamiento de la data que se discutirá con mayor detalle en el capítulo 6. Después de la 10, la que continúa en menor poder de predicción es la subtask 4 con un 63.15% y 64.87% de aciertos para los sensores SO2 y Eyetracker respectivamente. En anexos se muestran los resultados obtenidos de todos los sensores.

En la tabla 17 se presenta el promedio de accuracy para todas las subtasks de cada sensor en el modelo SVM.

Sensor	Accuracy
BPM	70%
SO2	69%
TMP	59%
EYE	62%
ECG	57%

Tabla 17 Accuracy por sensor

En la tabla 18 se presenta el promedio de accuracy para todas las subtasks por pareja de sensores usando SVM en donde se aprecia la supremacía de la pareja BPM+SO2.

Sensores	Accuracy
BPM+SO2	71.34%
BPM+TMP	69.62%
BPM+EYE	67.59%
BPM+ECG	66.94%
SO2+TMP	69.26%
SO2+EYE	67.67%
SO2+ECG	67.28%
TMP+EYE	64.35%
TMP+ECG	60.01%
EYE+ECG	60.06%

Tabla 18 Accuracy por pareja de sensores

- Naive Bayes

	Sensor	Accuracy	Precision	Recall
Subtarea 1	BPM	74.64%	71.91%	80.87%
	SO2	70.43%	64.13%	92.75%
	BPM+SO2	76.52%	69.93%	82.66%
	BPM+EYE	76.27%	82.01%	76.86%
Subtarea 2	BPM	76.00%	75.46%	77.10%
	TMP	68.69%	63.54%	87.77%
	BPM+TMP	78.82%	77.73%	80.81%
	BPM+SO2	81.42%	80.43%	83.06%
Subtarea 3	BPM	64.93%	64.16%	66.58%
	EYE	61.31%	59.36%	70.08%
	BPM+EYE	72.93%	70.70%	91.99%
	BPM+SO2	68.22%	62.43%	91.51%
Subtarea 4	SO2	63.79%	60.26%	81.03%
	EYE	65.30%	62.28%	77.59%
	SO2+EYE	74.04%	70.92%	91.95%
	BPM+EYE	67.23%	64.34%	77.31%
Subtarea 5	BPM	72.07%	72.47%	71.24%
	EYE	66.85%	63.06%	81.43%
	BPM+EYE	73.17%	72.97%	80.08%
	BPM+SO2	76.35%	75.40%	78.20%
Subtarea 6	BPM	74.11%	80.79%	63.25%
	SO2	72.79%	71.08%	76.85%
	BPM+SO2	80.76%	80.72%	76.35%
	BPM+EYE	78.41%	81.04%	74.11%
Subtarea 7	BPM	71.33%	79.35%	57.66%
	SO2	67.03%	65.10%	73.44%
	BPM+SO2	80.89%	86.00%	68.53%
	BPM+EYE	74.23%	80.38%	64.06%
Subtarea 8	SO2	67.56%	63.74%	79.81%
	TMP	66.32%	62.33%	80.64%
	SO2+TMP	74.17%	71.50%	79.55%
	SO2+EYE	69.42%	65.92%	80.49%
Subtarea 9	SO2	67.84%	63.96%	80.08%
	TMP	65.50%	61.73%	79.53%
	SO2+TMP	72.76%	70.32%	76.76%
	SO2+EYE	69.80%	64.43%	88.59%
Subtarea 10	BPM	57.95%	53.13%	68.23%
	SO2	56.23%	51.78%	66.02%
	BPM+SO2	58.32%	57.04%	62.57%
	BPM+TMP	56.65%	56.86%	63.62%

Tabla 19 Resultados modelamiento con Naive Bayes.

El formato de resultados que se presenta en la tabla 19 es el mismo que en el caso de SVM.

Los resultados son similares a los del modelo SVM. En este caso también los dos mejores sensores juntos predicen la dificultad de la tarea mejor que cada uno por separado superando el 80% de accuracy en las subtareas 6 y 7. La subtarea 4 nuevamente tiene bajo poder de predicción por separado, junto con la número 3, pero cuando se combinan ambos sensores el poder de predicción aumenta a 74.04% y 72.93% respectivamente. A su vez se vuelve a confirmar la baja predicción en la subtarea 10. Los mejores sensores son los mismos que en SVM salvo para la subtarea 3 y la 5 en que en vez del SO2 se agrega la Dilatación pupilar. En la mayoría de los casos la mejor pareja de sensores está compuesta por los mejores sensores individuales salvo en las subtareas 2 y 5 en que BPM+SO2 supera a BPM+TMP y a BPM+EYE respectivamente. En anexos se muestran los resultados obtenidos de todos los sensores.

En la tabla 20 se presenta el promedio de accuracy para cada uno de los sensores en todas las subtareas usando Naive Bayes.

Sensor	Accuracy
BPM	71%
SO2	69%
TMP	61%
EYE	60%
ECG	58%

Tabla 20 Accuracy por sensor

En la tabla 21 se presenta el promedio de accuracy para cada pareja de sensores en todas las subtareas usando Naive Bayes.

Sensores	Accuracy
BPM+SO2	71.70%
BPM+TMP	69.53%
BPM+EYE	68.90%
BPM+ECG	66.55%
SO2+TMP	67.93%
SO2+EYE	68.10%
SO2+ECG	66.33%
TMP+EYE	63.27%
TMP+ECG	61.12%
EYE+ECG	59.47%

Tabla 21 Accuracy por pareja de sensores

Se aprecia nuevamente la supremacia de la pareja BPM+SO2 con un 71.7% seguido de BPM+TMP con un 69.53%. Para Naive Bayes se obtienen buenos resultados con la pareja BPM+EYE a diferencia de lo que ocurre con SVM.

- Regresión Logística

En la tabla 22 se muestran los resultados para el modelo aplicando Regresión Logística.

	Sensor	Accuracy	Precision	Recall
Subtarea 1	BPM	70.92%	66.82%	83.12%
	SO2	68.30%	61.92%	95.04%
	BPM + SO2	71.65%	69.70%	76.16%
Subtarea 2	BPM	75.46%	75.65%	75.65%
	TMP	70.09%	65.50%	84.97%
	BPM + TMP	76.58%	74.91%	78.33%
Subtarea 3	BPM	62.00%	60.53%	69.25%
	EYE	60.13%	63.44%	79.51%
	BPM + EYE	63.91%	63.52%	67.83%
Subtarea 4	SO2	59.70%	57.89%	71.12%
	EYE	64.01%	60.45%	81.03%
	SO2 + EYE	66.55%	68.15%	72.15%
Subtarea 5	BPM	71.04%	71.88%	69.17%
	EYE	67.46%	63.41%	82.65%
	BPM + EYE	71.92%	72.87%	76.49%
Subtarea 6	BPM	72.32%	78.95%	60.86%
	SO2	69.33%	66.95%	76.37%
	BPM + SO2	73.43%	73.93%	79.17%
Subtarea 7	BPM	70.78%	78.66%	57.03%
	SO2	67.58%	65.52%	74.22%
	BPM + SO2	72.16%	71.73%	77.27%
Subtarea 8	SO2	65.77%	61.29%	83.57%
	TMP	63.57%	59.11%	85.38%
	SO2 + TMP	69.41%	71.73%	71.83%

Subtarea 9	SO2	64.83%	67.04%	67.79%
	TMP	63.54%	66.23%	65.38%
	SO2+TMP	70.49%	71.73%	73.91%
Subtarea 10	BPM	56.46%	52.44%	67.11%
	SO2	54.78%	51.13%	64.97%
	BPM+SO2	58.63%	57.67%	67.21%

Tabla 22. Resultados modelo Regresión Logística

Se observa que los resultados de este modelo son peores a los anteriores. Se mantiene la tendencia de la mejor predicción de los dos mejores sensores, pero los resultados rondan el 70% de precisión. Los mejores sensores son los mismos que los obtenidos con Naive Bayes.

### 5.3.3.2 Por participante

Como se mencionó anteriormente en esta sección se verá los resultados del modelamiento por participantes, esto es, sin normalizar los datos dado que no afecta la varianza inter participantes. Para esto cada feature fue modelada de la misma forma que en la sección anterior comparando la dificultad de una subtarea con estímulos externos y sin estos, evidentemente con la diferencia de que la porción de datos es menor y se debe correr el algoritmo muchas más veces.

En la tabla 23 se presentan los 3 mejores sensores, con su respectivo accuracy, que se obtuvieron para 10 participantes del estudio utilizando SVM. Estos accuracy representan el promedio de todas las subtasks de la actividad leer correos simples para cada participante.

	1er sensor	2do sensor	3er sensor
1	TMP= 98.7%	SO2= 92.01%	EYE= 82.7%
2	TMP= 100%	EYE= 95.9%	SO2= 77.3%
3	TMP= 100%	SO2= 87.4%	EYE= 78.8%
4	SO2= 97.6%	EYE= 90.8%	BPM= 86.1%
5	SO2= 98.2%	BPM= 91.3%	EYE= 68.7%
6	TMP= 92.7%	BPM= 59.8%	SO2= 57.8%
7	ECG= 78.8%	EYE= 78.1%	BPM= 70.8%
8	ECG= 88.1%	EYE=77.9%	BPM=76.7%
9	BPM=94.3%	SO2=91.5%	TMP= 88.3%
10	SO2= 92.8%	BPM=89.4%	TMP=85.6%

Tabla 23 Los mejores 3 sensores para cada participante usando SVM

A continuación se presentan los resultados una vez aplicado Naive Bayes.

	1er sensor	2do sensor	3er sensor
1	TMP=99.1%	SO2=92%	EYE=84%
2	TMP=100%	EYE= 99.2%	SO2= 74.2%
3	TMP=100%	SO2=87.4%	EYE=78.1%
4	SO2=97.6%	EYE=91.5%	BPM=86.1%
5	SO2=97.3%	EYE=91.5%	BPM=91.3%
6	TMP=93.1%	ECG=70.1%	EYE=67.8%
7	ECG=96.1%	EYE=78.1%	BPM=75.5%
8	ECG=96%	BPM=76.7%	TMP=73.6%
9	BPM=96.3%	SO2=90.8%	TMP=88.4%
10	SO2=93.9%	BPM=89,9%	EYE=80.2%

Tabla 24 Los mejores 3 sensores para cada participante usando Naive Bayes.

Se puede apreciar que los resultados de modelar por participante son bastante buenos para ambos modelos. En todos los participantes hay por lo menos un sensor que tiene un porcentaje de aciertos mayor al 90% en el caso de Naive Bayes. Incluso para el participante 2 y 3 la predicción es del 100% con el sensor de temperatura.



## 5.5 Comparación con NASA TLX

Como se mencionó en la sección 4.2 se tienen los resultados del test NASA aplicado a los 20 participantes. El valor que se entrega en este test es una unidad de medida de carga cognitiva.

Para comparar los resultados del test se sigue lo realizado por Chen et al[12] en que aplicó los resultados del test al modelo de minería de datos utilizado para comprobar su capacidad de predicción. Se normalizan los datos de nasa para cada participante para evitar los efectos de la varianza. Se debe tener en consideración que la cantidad de datos es mínima con respecto a lo que se tiene desde los sensores por lo que las conclusiones pueden no ser concluyentes.

Los resultados son los siguientes:

NASA TLX	
Accuracy	100%
Precisión	100%
Recall	100%

Al normalizar los datos al modelo le resulta muy fácil acertar las distintas clases debido a que prácticamente todos los registros correspondientes a la tarea más fácil quedan con valor negativo o muy pequeño. Luego, como se señaló anteriormente este resultado no es concluyente dado que no se puede comparar la cantidad de datos de ambas medidas.

## 5.6 Análisis por sensor

- BPM: Es el sensor que resultó con el mayor poder predictivo, con un 71% por si solo y alcanzando un 81.42% combinado con el sensor SO2. Es uno de los dos mejores sensores en 7 de las 10 subtareas estudiadas. En todos los participantes mostró una tendencia a aumentar a medida que aumenta la dificultad de la tarea por lo que se comprobó lo expuesto en la literatura.
- SO2: El resultado de este sensor resulta un tanto inesperado dado que no existe amplia evidencia en la literatura sobre el aumento de la saturación de oxígeno bajo el aumento de dificultades en tareas cognitivas. Es el segundo sensor que mejor predice con un 69%.

- **Temperatura:** Este sensor resulta particularmente atractivo debido a que la evidencia en la literatura indica que a medida que aumenta la dificultad cognitiva aumenta la temperatura superficial de la piel, pero en este caso la temperatura media de la piel bajaba cuando se aumentaba la dificultad. Este comportamiento se mantuvo para todos los participantes con énfasis en los primeros 3 participantes, luego las diferencias son menores. Por la naturaleza del sensor es posible que puedan existir ciertas interferencias provocadas por factores ambientales.
- **Eyetracker:** Este sensor tiene un accuracy de 60%. Es un resultado más pobre que los trabajos anteriores estudiados cuya capacidad de predicción ronda el 70%. La primera razón es instrumental, este sensor es móvil a diferencia de la mayoría de los utilizados anteriormente en que una estructura fija mantiene la posición y mirada de la persona apuntando al objetivo de mira. En este caso existen muchas posibilidades de que el participante mueva la cabeza provocando que la confianza en la medición de la pupila disminuya teniendo que eliminar esa data. Además, como se debe estar mirando el Smartphone, la posición natural es a unos 45 grados con respecto a la horizontal provocando otros problemas de confianza en la medición al tener el párpado un tanto caído. Otro inconveniente que se existió fue participantes con las pestañas muy largas que provocaban una pérdida breve de la señal. Otro punto a considerar son los cambios de luminosidad, a pesar de que éstos se intentaron controlar, pueden existir ciertas variaciones que provoquen ruido en la muestra por lo que a futuro se pueden aplicar filtros que corrijan el problema.
- **Electrocardiograma:** Es el sensor con peor poder predictivo de este estudio con un 58% de accuracy. Es un resultado lejano al 71.4% que obtuvo Haapalainen et al [9] y al 81.1% que obtuvieron en el mismo estudio al combinarlo con el flujo de calor.

Para trabajar con los registros del ECG se siguieron todos los pasos del procesamiento estándar, se eliminaron los ruidos e interferencias mediante la Transformada de Fourier, filtro pasa bajo y Transformada Inversa de Fourier. También se extrajeron las features recomendadas por la literatura.

Del análisis exploratorio se pudo apreciar un comportamiento bastante aleatorio de la señal, sin seguir tendencias iguales para todos los participantes. Además el ECG fue el único sensor que no fue capaz de rechazar la hipótesis nula de ANOVA y t para cuatro subtareas. Una de las razones de este resultado puede ser la recomendación de algunos autores por utilizar gel conductor en los electrodos. En la literatura existe evidencia de buenos resultados con gel y sin gel por lo que se optó por no utilizarlo con el fin de asemejar lo más posible el experimento a condiciones cotidianas. Queda propuesto entonces para futuros estudios la utilización de gel conductor y su comparación con el escenario en que no se use.

## 6. Discusión

En este estudio se tenía como objetivo realizar una evaluación de la carga cognitiva necesaria para realizar actividades HCI usando sensores psicofisiológicos siguiendo la metodología planteada en [9]. En la experimentación se hizo énfasis en actividades cotidianas utilizando el Smartphone más la presencia o ausencia de otros estímulos que le otorgan dificultad a las tareas agregando más complejidad al contexto. Esta dificultad se validó empíricamente utilizando las medidas de NASA TLX, % de Performance y el tiempo de realización de las tareas. Las actividades desarrolladas son 5 y se acotó la data a las subtareas que de acuerdo a la teoría requieren la mayor cantidad de recursos cognitivos.

El primer resultado que se tiene, luego del análisis estadístico, fue la comprobación de que las señales fisiológicas estudiadas varían significativamente en sus medias entre actividades de distinta dificultad con un 95% de confianza.

Luego del procesamiento de la data se aplicaron modelos de predicción de minería de datos cuyo objetivo es discriminar entre las señales que corresponden a actividades de dificultad alta versus dificultad baja.

Se aplicaron los modelos por tarea normalizando a todos los participantes y separando a cada participante. Los resultados obtenidos del modelo por tarea señalan que los sensores que fueron mejores predictores son el sensor BPM y el sensor SO2 con un 71% y 69% de éxitos. Cuando se aplica el modelo con ambos sensores a la vez llegan a accuracy de 81.42% para una tarea en particular. Las señales que juntas predicen mejor son BPM junto a SO2 y BPM junto a TMP logrando un 71% y 69% de accuracy en promedio respectivamente.

Cuando se aplica el modelo por participante las señales demuestran un poder de predicción mayor llegando a predecir en algunos casos con una certeza de un 100% (sensor de temperatura).

Se podría pensar que el camino está en desarrollar modelos particulares a cada individuo, pero si miramos detenidamente los resultados obtenidos la respuesta no es tan clara. Es cierto que en prácticamente todos los participantes se llega a por lo menos una señal que predice en un porcentaje mayor a 90% la dificultad de una tarea, pero la varianza de cada sensor es muy grande. Por ejemplo, viendo el caso de la temperatura en que para los participantes 2 y 3 se tiene un accuracy de 100%, para los participantes 7 y 9 su accuracy es de 52% y 26% respectivamente. Esto se hace latente si vemos las accuracy medias de los sensores:

Sensor	Accuracy	Accuracy
BPM	67,5%	71%
SO2	73%	69%
TMP	61%	61%
EYE	60%	60%
ECG	56%	58%

Tabla 25 Accuracy promedio por participante y por total de tareas.

Luego, es claro que no hay un sensor que predomine por sobre otro para todos los participantes por igual por lo que un enfoque de trabajo futuro sería, más que preocuparse por algún sensor en sí, preocuparse por el total de predicción que un grupo de sensores realiza.

Lo anterior es concordante con la literatura estudiada en donde los resultados son dispares en los sensores que mejor predicen la dificultad de las tareas

Para poner estos resultados en contexto, existen trabajos anteriores que aplican la metodología que se aplicó aquí, enfocados en actividades desktop-based sin agregar mayor complejidad al contexto. Haapalainen et al [9] predice la dificultad de una tarea con un 81.1% utilizando ECGMAD y flujo de calor. Cifra no alejada a lo obtenido en este trabajo con otros sensores. Los mismos autores aplicaron nuevamente la metodología obteniendo cifras similares, pero con el sensor de Electroencefalograma. Chen et al [12] obtiene un accuracy de 70% utilizando el diámetro pupilar. Fritz et al [28] predicen con un 84.3% la dificultad de una tarea usando Eyetracker, Conductividad galvánica de la piel y Electroencefalograma. Así, existen variados estudios que logran predecir del orden de un 75% tareas difíciles versus tareas fáciles.

La respuesta al por qué las cifras de predicción son parecidas, pero el sensor que lo logra es distinto se puede deber a variadas razones. En primer lugar los instrumentos de medición son distintos. Se han configurado dispositivos con diversas arquitecturas que pueden provocar ciertas tendencias disimiles. Una homogenización en la instrumentación podría generar resultados mucho más comparables con otros estudios.

En cuanto a este estudio, existen diferencias claras con respecto a los resultados del eyetracker y sobre todo del ECG. Aparte de la diferencia de base entre el tipo de actividades que se realizaron, en ambos casos, es probable que el diseño móvil del experimento haya afectado el desempeño de esos sensores. Por un lado el eyetracker es portátil, a diferencia de los más utilizados en los últimos años, lo que facilita el movimiento de la cabeza con la correspondiente pérdida de confianza en las mediciones, además, la posición en que naturalmente se acciona un Smartphone genera un ángulo que a la cámara ocular del eyetracker le dificulta la percepción de la pupila. En el caso del ECG, para cumplir con las características de movilidad y cierta cotidianeidad del experimento se optó

por no utilizar gel de conducción en los electrodos lo que pudo provocar medidas con menos precisión y por consiguiente pobres resultados en los modelos de predicción.

Con respecto a los resultados por actividad y escenario, lo primero a comentar es la actividad de correo electrónico simple. Compuesto por las subtareas 1 y 2, según la teoría de Wickens la dificultad de esta actividad es menor dado que solo se ocupa la dimensión verbal al tratarse de mensajes cortos, precisos y que exigen una respuesta de la misma manera. Debido a esto, esta actividad resultaba clave para averiguar si es que las señales eran significativamente distintas ante dificultades disimiles dado que, al regirse por la teoría, si eso se cumplía la probabilidad de que con las demás actividades ocurriera lo mismo era mayor. Lo mismo pasaba para el modelo de predicción. Los resultados indican que el clasificador SVM tiene un accuracy de 76% y 78% para las subtareas 1 y 2 respectivamente. En el caso de Naive Bayes, al combinar las señales de BPM y SO2 en la subtarea 2 se obtiene un accuracy de 81.42%, que es la mejor efectividad obtenida en este estudio y un resultado bueno considerando el contexto de trabajos anteriores.

La actividad de correos de búsqueda de información es según la teoría la más complicada dado que combina recursos cognitivos de memoria, verbales, visuales sumado a la conversación y la música en la dificultad alta. Los resultados del modelo de predicción reafirman la teoría al ser las subtareas 6 y 7 en las que mejor se predice la dificultad de las mismas. El clasificador Naive Bayes tiene un accuracy de 80.76% para la subtarea 6 y 80.89% para la subtarea 7. Por el contrario la subtarea 3 debería ser a lo menos similar a la subtarea 1 debido a que ambas son leer un correo electrónico, pero los resultados indican que la subtarea 3 tiene menor efectividad que ésta y además es la subtarea con menor accuracy de esta actividad si se ven los resultados de los sensores combinados.

Hangout es una actividad simple que si bien tiene el componente agregado de estar expectante esperando por que llegue el mensaje existe una familiaridad del grupo experimental con este tipo de aplicaciones. El accuracy del clasificador de Naive Bayes para esta actividad es 74.17% combinando el sensor de SO2 más Temperatura, resultado también bueno considerando el contexto.

La actividad de leer correos electrónicos mientras se camina es la que peor resultados tuvo con una accuracy del 58.32% con Naive Bayes y 57.65% con SVM. Esto se debe a varios factores. En primer lugar se tuvo que eliminar casi el 50% de la data de los participantes 5, 7, 9, 13 y 14 por errores en registros de los sensores con valores excesivamente altos o bajos. Se apreciaron ciertas tendencias a aumentar la temperatura muy por sobre lo normal, errores en los registros de ECG en donde por largos períodos se tenía un valor constante y menor frecuencia de registro del eyetracker debido a pequeños cambios en la luminosidad y a la imposibilidad de mantener un ángulo de mirada óptimo para la cámara mientras la persona camina.

Para finalizar se debe señalar que se comprobó la hipótesis de esta memoria y para este estudio efectivamente es posible medir la carga cognitiva para actividades en que el usuario debe dividir su atención en un entorno móvil mediante medidas psico-fisiológicas tomando en cuenta el contexto situacional.

## 7. Conclusiones y trabajo futuro

La medición de carga cognitiva en tiempo real representa un desafío para la Computación Ubicua con el fin de adquirir mayor información para tomas de decisiones que representen avances en el desarrollo de aplicaciones o de todo tipo de estrategias de negocio.

Midiendo la carga cognitiva con sensores psicofisiológicos se habían obtenido buenos resultados de clasificación con respecto a actividades desktop-based.

En este trabajo se comprobó que también es posible evaluar la carga cognitiva para actividades HCI en que están presentes otro tipo de estímulos como música o una conversación, situación que acerca los escenarios a las situaciones cotidianas.

También se comprobó que el diseño experimental basado en el modelo de múltiples recursos cognitivos de Wickens era atinente al problema teniendo como resultado que las subtareas que se suponían más demandantes cognitivamente fueron las que mejor se predijeron en los modelos.

Utilizando el sensor BPM y el sensor SO2 se llegó a clasificar con 81.42% de accuracy la dificultad de una subtask en particular, resultado comparable con los obtenidos por trabajos anteriores.

Posibles caminos de trabajo futuro son: Aumentar la frecuencia de registro de datos para el dispositivo de captura de manera de poder tener un mayor detalle de tareas que duren pocos segundos y estudiar cómo varía la carga cognitiva en esos períodos. Desarrollar modelos que se comporten como una caja negra con respecto a qué sensor está siendo el más efectivo, sino que se preocupe del total agregado de la predicción evitando así las constantes variaciones de efectividad de los sensores. También se puede profundizar en la investigación de esta área desarrollando modelos que sean capaces de medir carga cognitiva en tiempo real usando la metodología planteada en este trabajo, de esa forma se podrían desarrollar aplicaciones que solo ofrezcan un servicio cuando los sensores predigan que la carga cognitiva es la adecuada.

# Bibliografía

- [1] M. Weiser. The computer for the 21st century. 1991.
- [2] M. Weiser. Ubiquitous Computing. 1993.
- [3] Simon H.A. Designing organizations for an information rich world, 1971
- [4] Miettinen and Oulasvirta. Predicting time-sharing in mobile interaction, 2007
- [5] Ferreira et al. Assessing real-time cognitive load based on psychophysiological measures for younger and older adults, 2013
- [6] Documento Fondecyt. Sr. Ángel Jiménez-Molina
- [7] C. Wickens. Multiple resources and performance prediction. Theoretical issues in ergonomics science, 3(2):159–177, 2002.
- [8] Yin, B. et al. Speech-based cognitive load monitoring system, 2008
- [9] Haapalainen, Kim, Forlizzi, Dey. Psycho-Physiological measures for assessing cognitive load
- [10] Hart, S.G. and Staveland, L.E. (1988) Development of NASATLX (Task Load Index): results of empirical and theoretical research. In Hancock, P.A. and Meshkati, N. (Eds.), Human Mental Workload. Amsterdam, North-Holland, 139-183.
- [11] Wilson et al. Psychophysiological assessment of workload in multi-task environments, 1991
- [12] Chen et al., A comparison of four methods for cognitive load measurement, 2011
- [13] Cacioppo, Tassinari, Berntson. The Handbook of Psychophysiology, Third Edition, 2007
- [14] Shi et al. Galvanic Skin Response (GSR) as an Index of Cognitive Load. 2007
- [15] Assessing Cognitive Load with Physiological Sensors. 2005
- [16] Wilson, G. F. An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures. 2009
- [17] Nourbakhsh et al. Using Galvanic Skin Response for Cognitive Load Measurement in Arithmetic and Reading Tasks. 2012
- [18] F. J. A. Fernández, J. F. Pérez, and S. F. López, Oftalmología en Atención Primaria. Alcalá, 2002
- [19] Aracena, C. Estudio de la relación entre neurodatos, dilatación pupilar y emocionalidad basado en técnicas de minería de datos. 2014
- [20] Beatty, Lucero-Wagoner. The Pupillary System.
- [21] J. Velásquez, "Combining eye-tracking technologies with web usage mining for identifying website keyobjects". 2013
- [22] Beatty, J.: Task-evoked pupillary responses, processing load, and the structure of processing resources.



- [23] Xu, J. et al. Pupillary Response Based Cognitive Workload Measurement under Luminance Changes
- [24] Fredericks et al. An investigation of myocardial aerobic capacity as a measure of both physical and cognitive workloads
- [25] Ettema, J.H., Zielhuis, R.L., 1971. Physiological parameters of mental load. *Ergonomics* 14 (1), 137–144
- [26] Hitchen, M., Brodie, D.A., Harness, J.B., 1980. Cardiac responses to demanding mental load.
- [27] Ryu, K. Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic
- [28] Fritz, Begel, Müller, Yigit-Elliot, Züger. Using Psycho-Physiological Measures to Assess Task Difficulty in Software Development
- [29] Oulasvirta, Tamminen, Roto, Kuorelahti. Interaction in 4-second bursts: The fragmented nature of attentional resources in mobile HCI, 2005.
- [30] Andreassi. *Psychophysiology: Human Behaviour and Physiological response*, 2000.
- [31] Jimenez-Molina et al. Cognitive Resource-Aware Unobtrusive Service Provisioning in Ambient Intelligence Environments. 2014
- [32] Coyne, Baldwin, Cole, Sibley, Roberts. Applying Real Time Physiological Measures of cognitive load to improve training.
- [33] Dirican, Göktürk. *Psychophysiological Measures of Human Cognitive States Applied in HCI*. 2011
- [34] Setz, Arnrich, Schumm, La Marca. Discriminating stress from cognitive load using a wearable EDA device. 2010
- [35] <http://es.dreamstime.com/foto-de-archivo-libre-de-regal%C3%ADas-secci%C3%B3n-representativa-de-la-piel-image9845525> visitada en Mayo 2015
- [36] S. U. F. de Ingeniería Biométrica Informática, “clasificación de los sistemas biométricos.” <http://redyseguridad.fi-p.unam.mx/proyectos/biometria/clasificacionsistemas/recoirisyretina.html>
- [37] Jadue. Incidencia de la dilatación pupilar como variable predictiva del comportamiento de los usuarios en una página web antes de tomar una decisión. 2014
- [38] Kahneman, Beatty. Pupil diameter and load memory. 1966.
- [39] <https://anapaoar.files.wordpress.com/2013/02/3.jpg> visitada en Mayo 2015
- [40] [www.cooking-hacks.com](http://www.cooking-hacks.com) visitada en Mayo 2015
- [41] Nourbakhsh, Wwang, Chen. GSR and Blink Features for cognitive load classification. 2013.
- [42] [www.pupil-labs.com](http://www.pupil-labs.com) visitada en Octubre 2015
- [43] Hughes, Volante, Stowers, Leyva, Oglesby, Busbey, Salas. Cognition and Physiological response: Towards a model of validated physiological measurement. 2014
- [44] Iqbal, Bailey. Understanding Changes in Mental Workload during Execution of Goal-Directed Tasks and Its Application for Interruption Management

[45] O. Maimon and L. R. (Eds.), The Data Mining and Knowledge Discovery Handbook.  
Springer, 2005.

[46] Pacheco S. El clasificador Naive Bayes en la extracción de conocimiento de bases de datos

# Anexos

## A Detalle resultados SVM por sensor y subtarea

	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 1</b>	BPM	74.49%	71.72%	80.87%	<b>Subtarea 2</b>	BPM	76.00%	75.46%	77.10%
	SO2	68.12%	61.68%	95.65%		SO2	68.88%	64.13%	85.59%
	TMP	65.36%	60.04%	91.88%		TMP	71.02%	65.98%	86.84%
	EYE	54.35%	52.88%	79.71%		EYE	51.27%	50.70%	93.26%
	ECG	58.55%	54.75%	98.55%		ECG	58.52%	54.77%	98.13%
	BPM+SO2	76.94%	69.53%	81.79%		BPM+SO2	78.15%	77.71%	78.98%
	BPM+TMP	75.08%	79.76%	77.70%		BPM+TMP	78.16%	76.40%	81.52%
	BPM+EYE	75.76%	82.65%	74.86%		BPM+EYE	76.31%	76.22%	76.53%
	BPM+ECG	76.78%	84.47%	74.57%		BPM+ECG	76.98%	76.53%	77.86%
	SO2+TMP	75.59%	72.89%	93.71%		SO2+TMP	74.63%	70.48%	84.80%
	SO2+EYE	71.69%	70.65%	89.43%		SO2+EYE	67.84%	63.69%	83.06%
	SO2+ECG	72.03%	70.98%	89.43%		SO2+ECG	70.19%	65.57%	85.10%
	TMP+EYE	67.29%	66.74%	89.43%		TMP+EYE	69.07%	64.34%	85.61%
	TMP+ECG	67.29%	66.60%	90.00%		TMP+ECG	70.75%	65.69%	86.94%
EYE+ECG	61.53%	60.73%	99.43%	EYE+ECG	60.64%	56.12%	97.76%		
	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 3</b>	BPM	62.41%	60.82%	70.08%	<b>Subtarea 4</b>	BPM	62.93%	60.56%	74.13%
	SO2	65.66%	62.24%	82.21%		SO2	63.15%	60.13%	78.02%
	TMP	48.98%	49.42%	80.86%		TMP	48.27%	49.01%	86.20%
	EYE	59.91%	56.87%	82.47%		EYE	64.87%	60.95%	82.76%
	ECG	48.85%	49.34%	80.86%		ECG	46.12%	47.73%	81.89%
	BPM+SO2	69.85%	70.00%	85.08%		BPM+SO2	62.61%	59.68%	77.73%
	BPM+TMP	63.84%	61.25%	75.34%		BPM+TMP	65.55%	62.50%	77.73%
	BPM+EYE	63.70%	61.36%	73.97%		BPM+EYE	65.97%	63.38%	75.63%
	BPM+ECG	62.47%	60.76%	70.41%		BPM+ECG	60.71%	59.85%	65.13%
	SO2+TMP	61.37%	57.61%	86.03%		SO2+TMP	66.18%	62.07%	83.19%
	SO2+EYE	65.34%	61.62%	81.37%		SO2+EYE	71.88%	70.45%	86.86%
	SO2+ECG	61.23%	57.62%	84.93%		SO2+ECG	65.76%	61.68%	83.19%
	TMP+EYE	62.74%	59.55%	79.45%		TMP+EYE	67.23%	62.89%	84.03%
	TMP+ECG	48.36%	49.02%	82.19%		TMP+ECG	59.03%	55.47%	91.60%
EYE+ECG	58.90%	56.26%	80.00%	EYE+ECG	65.76%	62.30%	79.83%		

	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 5</b>	BPM	72.07%	72.47%	71.24%	<b>Subtarea 6</b>	BPM	74.11%	79.88%	64.44%
	SO2	67.76%	63.97%	81.43%		SO2	70.17%	67.42%	78.04%
	TMP	27.50%	28.63%	30.09%		TMP	54.89%	72.04%	15.99%
	EYE	66.84%	63.06%	81.43%		EYE	63.24%	66.46%	53.46%
	ECG	50.33%	50.18%	97.08%		ECG	52.14%	66.07%	8.83%
	BPM+SO2	72.40%	73.33%	83.97%		BPM+SO2	78.01%	79.61%	70.09%
	BPM+TMP	72.11%	71.98%	73.41%		BPM+TMP	76.04%	77.44%	73.40%
	BPM+EYE	71.07%	71.38%	71.46%		BPM+EYE	78.88%	82.40%	73.40%
	BPM+ECG	72.34%	72.28%	72.46%		BPM+ECG	77.22%	80.21%	72.21%
	SO2+TMP	67.84%	64.58%	79.04%		SO2+TMP	72.84%	71.72%	75.30%
	SO2+EYE	69.28%	66.10%	79.16%		SO2+EYE	68.68%	70.82%	63.42%
	SO2+ECG	68.56%	64.85%	81.08%		SO2+ECG	70.23%	67.56%	77.67%
	TMP+EYE	67.31%	64.79%	75.81%		TMP+EYE	64.06%	66.48%	56.53%
	TMP+ECG	62.57%	59.83%	76.53%		TMP+ECG	56.58%	76.70%	18.76%
EYE+ECG	64.01%	61.96%	72.57%	EYE+ECG	61.92%	65.43%	50.36%		
	<b>Sensor</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>		<b>Sensor</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>
<b>Subtarea 7</b>	BPM	72.03%	79.38%	59.53%	<b>Subtarea 8</b>	BPM	56.06%	57.37%	47.13%
	SO2	67.19%	65.28%	73.44%		SO2	68.53%	62.81%	89.14%
	TMP	55.07%	54.73%	58.75%		TMP	65.36%	60.63%	85.38%
	EYE	61.40%	65.33%	48.59%		EYE	54.83%	53.14%	81.74%
	ECG	53.12%	66.66%	12.50%		ECG	53.47%	51.85%	97.00%
	BPM+SO2	80.98%	85.33%	69.52%		BPM+SO2	69.07%	65.11%	82.28%
	BPM+TMP	73.15%	76.83%	66.22%		BPM+TMP	67.70%	63.64%	82.69%
	BPM+EYE	74.65%	81.09%	64.23%		BPM+EYE	56.15%	56.54%	53.43%
	BPM+ECG	72.40%	79.56%	60.23%		BPM+ECG	53.68%	54.04%	49.59%
	SO2+TMP	71.16%	71.53%	70.22%		SO2+TMP	75.83%	74.73%	77.33%
	SO2+EYE	68.00%	70.38%	62.06%		SO2+EYE	69.14%	63.94%	87.91%
	SO2+ECG	69.16%	68.03%	72.21%		SO2+ECG	71.13%	66.56%	85.03%
	TMP+EYE	60.60%	67.12%	41.43%		TMP+EYE	64.95%	61.38%	80.77%
	TMP+ECG	55.36%	80.19%	14.14%		TMP+ECG	65.77%	61.59%	83.93%
EYE+ECG	61.51%	66.43%	46.42%	EYE+ECG	54.78%	52.65%	95.47%		

	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 9</b>	BPM	58.71%	59.40%	44.24%	<b>Subtarea 10</b>	BPM	56.73%	52.08%	66.89%
	SO2	67.84%	63.96%	80.08%		SO2	56.46%	52.44%	67.11%
	TMP	65.50%	61.73%	79.53%		TMP	51.33%	52.08%	71.22%
	EYE	55.71%	54.55%	62.53%		EYE	54.98%	53.51%	76.23%
	ECG	54.85%	52.61%	86.90%		ECG	53.60%	52.58%	74.17%
	BPM+SO2	67.69%	68.16%	66.48%		BPM+SO2	57.65%	55.90%	61.24%
	BPM+TMP	67.87%	63.66%	78.70%		BPM+TMP	56.70%	55.43%	68.68%
	BPM+EYE	57.04%	60.36%	41.20%		BPM+EYE	56.35%	55.07%	69.36%
	BPM+ECG	61.30%	57.42%	87.63%		BPM+ECG	55.53%	54.42%	68.40%
	SO2+TMP	72.36%	69.31%	77.21%		SO2+TMP	54.79%	53.62%	68.55%
	SO2+EYE	70.10%	66.51%	81.04%		SO2+EYE	54.79%	54.07%	61.63%
	SO2+ECG	71.82%	66.87%	86.53%		SO2+ECG	52.72%	52.22%	60.24%
	TMP+EYE	66.18%	62.36%	81.73%		TMP+EYE	54.10%	53.42%	61.63%
	TMP+ECG	62.45%	58.82%	83.56%		TMP+ECG	51.91%	52.74%	58.50%
EYE+ECG	58.41%	55.13%	90.63%	EYE+ECG	53.11%	53.81%	59.47%		

## B Detalle resultados Naive Bayes por sensor y subtarea

	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 1</b>	BPM	74.64%	71.91%	80.87%	<b>Subtarea 2</b>	BPM	76.00%	75.46%	77.10%
	SO2	70.43%	64.13%	92.75%		SO2	68.42%	64.49%	82.07%
	TMP	63.91%	59.52%	86.96%		TMP	68.69%	63.54%	87.77%
	EYE	55.07%	53.33%	81.16%		EYE	55.98%	54.58%	71.50%
	ECG	64.64%	59.10%	95.07%		ECG	61.63%	56.92%	95.85%
	BPM+SO2	76.52%	69.93%	82.66%		BPM+SO2	81.42%	80.43%	83.06%
	BPM+TMP	74.40%	79.50%	76.50%		BPM+TMP	78.82%	77.73%	80.81%
	BPM+EYE	76.27%	82.01%	76.86%		BPM+EYE	76.31%	76.06%	76.84%
	BPM+ECG	72.37%	73.43%	83.71%		BPM+ECG	78.41%	76.65%	81.73%
	SO2+TMP	74.92%	73.82%	89.43%		SO2+TMP	74.07%	68.24%	90.10%
	SO2+EYE	71.69%	70.65%	89.43%		SO2+EYE	70.90%	68.54%	77.35%
	SO2+ECG	70.17%	68.28%	92.86%		SO2+ECG	71.01%	64.47%	93.67%
	TMP+EYE	67.63%	67.10%	89.14%		TMP+EYE	68.71%	63.48%	88.16%
	TMP+ECG	68.98%	67.29%	92.86%		TMP+ECG	69.98%	63.96%	91.63%
EYE+ECG	65.42%	63.67%	97.14%	EYE+ECG	62.33%	57.32%	96.73%		
	<b>Sensor</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>		<b>Sensor</b>	<b>Accuracy</b>	<b>Precision</b>	<b>Recall</b>
<b>Subtarea 3</b>	BPM	64.93%	64.16%	66.58%	<b>Subtarea 4</b>	BPM	60.77%	58.44%	74.56%
	SO2	61.26%	57.92%	82.74%		SO2	63.79%	60.26%	81.03%
	TMP	58.83%	55.63%	87.87%		TMP	59.05%	55.73%	87.93%
	EYE	61.31%	59.36%	70.08%		EYE	65.30%	62.28%	77.59%
	ECG	53.26%	54.57%	83.90%		ECG	53.87%	52.16%	93.53%
	BPM+SO2	68.22%	62.43%	91.51%		BPM+SO2	62.18%	59.42%	76.89%
	BPM+TMP	65.75%	63.04%	76.16%		BPM+TMP	63.87%	60.06%	82.77%
	BPM+EYE	72.93%	70.70%	91.99%		BPM+EYE	67.23%	64.34%	77.31%
	BPM+ECG	64.66%	64.04%	66.85%		BPM+ECG	59.03%	58.57%	61.76%
	SO2+TMP	59.04%	56.37%	80.00%		SO2+TMP	59.66%	56.50%	84.03%
	SO2+EYE	65.89%	61.65%	84.11%		SO2+EYE	74.04%	70.92%	91.95%
	SO2+ECG	60.27%	56.86%	85.21%		SO2+ECG	63.87%	60.51%	79.83%
	TMP+EYE	62.33%	59.57%	76.71%		TMP+EYE	60.50%	57.02%	85.29%
	TMP+ECG	60.27%	56.66%	87.40%		TMP+ECG	57.35%	54.87%	82.77%
EYE+ECG	58.63%	56.80%	72.05%	EYE+ECG	61.55%	61.90%	60.08%		

	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 5</b>	BPM	72.07%	72.47%	71.24%	<b>Subtarea 6</b>	BPM	74.11%	80.79%	63.25%
	SO2	66.84%	63.06%	81.43%		SO2	72.79%	71.08%	76.85%
	TMP	38.91%	42.20%	59.83%		TMP	58.35%	68.04%	31.50%
	EYE	66.85%	63.06%	81.43%		EYE	62.41%	67.21%	48.44%
	ECG	49.14%	50.18%	92.59%		ECG	53.69%	76.27%	10.73%
	BPM+SO2	76.35%	75.40%	78.20%		BPM+SO2	80.76%	80.72%	76.35%
	BPM+TMP	73.29%	73.27%	73.53%		BPM+TMP	76.51%	79.58%	71.26%
	BPM+EYE	73.17%	72.97%	80.08%		BPM+EYE	78.41%	81.04%	74.11%
	BPM+ECG	72.22%	72.54%	71.50%		BPM+ECG	73.78%	78.57%	65.32%
	SO2+TMP	68.02%	65.16%	77.49%		SO2+TMP	71.17%	73.06%	66.98%
	SO2+EYE	68.68%	65.82%	77.72%		SO2+EYE	68.45%	72.08%	60.10%
	SO2+ECG	68.32%	64.60%	81.08%		SO2+ECG	69.40%	71.39%	64.61%
	TMP+EYE	66.29%	63.52%	76.53%		TMP+EYE	63.23%	67.96%	49.88%
	TMP+ECG	63.11%	59.48%	82.28%		TMP+ECG	54.33%	59.38%	27.08%
EYE+ECG	63.83%	61.58%	73.53%	EYE+ECG	58.96%	63.07%	42.99%		
	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 7</b>	BPM	71.33%	79.35%	57.66%	<b>Subtarea 8</b>	BPM	57.35%	60.03%	44.01%
	SO2	67.03%	65.10%	73.44%		SO2	67.56%	63.74%	79.81%
	TMP	54.92%	54.59%	58.43%		TMP	66.32%	62.33%	80.64%
	EYE	62.10%	68.85%	44.21%		EYE	54.83%	54.31%	60.89%
	ECG	51.32%	50.84%	80.00%		ECG	58.92%	55.22%	94.27%
	BPM+SO2	80.89%	86.00%	68.53%		BPM+SO2	69.07%	69.04%	69.23%
	BPM+TMP	72.82%	81.71%	58.74%		BPM+TMP	67.77%	64.60%	78.71%
	BPM+EYE	74.23%	80.38%	64.06%		BPM+EYE	57.73%	60.97%	43.13%
	BPM+ECG	74.56%	81.18%	63.89%		BPM+ECG	60.62%	57.04%	86.26%
	SO2+TMP	69.58%	68.92%	71.21%		SO2+TMP	74.17%	71.50%	79.55%
	SO2+EYE	68.41%	69.35%	65.89%		SO2+EYE	69.42%	65.92%	80.49%
	SO2+ECG	66.67%	64.12%	75.54%		SO2+ECG	72.71%	67.48%	87.77%
	TMP+EYE	60.68%	67.39%	41.26%		TMP+EYE	65.36%	61.64%	81.46%
	TMP+ECG	59.19%	59.52%	57.24%		TMP+ECG	63.09%	59.17%	84.62%
EYE+ECG	63.84%	67.89%	52.41%	EYE+ECG	58.97%	55.47%	91.21%		

	Sensor	Accuracy	Precision	Recall		Sensor	Accuracy	Precision	Recall
<b>Subtarea 9</b>	BPM	61.85%	60.45%	68.68%	<b>Subtarea 10</b>	BPM	57.95%	53.13%	68.23%
	SO2	67.84%	63.96%	80.08%		SO2	56.23%	51.78%	66.02%
	TMP	65.50%	61.73%	79.53%		TMP	56.01%	56.54%	52.19%
	EYE	62.19%	59.44%	76.92%		EYE	54.98%	55.47%	50.82%
	ECG	54.08%	52.88%	75.54%		ECG	53.60%	54.09%	48.07%
	BPM+SO2	63.23%	61.66%	70.05%		BPM+SO2	58.32%	57.04%	62.57%
	BPM+TMP	65.42%	61.98%	79.94%		BPM+TMP	56.65%	56.86%	63.62%
	BPM+EYE	57.52%	57.99%	54.80%		BPM+EYE	55.24%	55.69%	61.54%
	BPM+ECG	56.76%	57.24%	53.70%		BPM+ECG	53.11%	53.86%	58.78%
	SO2+TMP	72.76%	70.32%	76.76%		SO2+TMP	55.94%	56.40%	61.54%
	SO2+EYE	69.80%	64.43%	88.59%		SO2+EYE	53.68%	54.27%	60.58%
	SO2+ECG	68.25%	64.72%	81.30%		SO2+ECG	52.62%	53.32%	59.88%
	TMP+EYE	63.23%	60.19%	78.29%		TMP+EYE	54.74%	55.17%	61.96%
	TMP+ECG	62.54%	59.36%	79.67%		TMP+ECG	52.33%	53.23%	57.93%
EYE+ECG	50.65%	50.42%	82.41%	EYE+ECG	50.56%	51.67%	55.17%		