



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

SELECCIÓN, IMPLEMENTACIÓN Y VALIDACIÓN DE MODELOS DE PREDICCIÓN  
DE RIESGO PARA ENFERMEDADES CARDIOVASCULARES EN PACIENTES CON  
PROBLEMAS DE SUEÑO

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL E  
INGENIERO CIVIL EN COMPUTACIÓN

MOISÉS JAVIER HERNÁNDEZ ARANDA

PROFESOR GUÍA:  
JUAN DOMINGO VELÁSQUEZ SILVA

MIEMBROS DE LA COMISIÓN:  
AIDAN HOGAN  
ROCÍO BELÉN RUIZ MORENO  
ALEXANDRE BERGEL

SANTIAGO DE CHILE  
2017

RESUMEN DE LA MEMORIA PARA OPTAR AL  
TÍTULO DE: Ingeniero Civil Industrial e Ingeniero  
Civil en Computación  
POR: Moisés Javier Hernández Aranda  
FECHA: 21/08/2017  
PROFESOR GUÍA: Juan Domingo Velásquez Silva

## **SELECCIÓN, IMPLEMENTACIÓN Y VALIDACIÓN DE MODELOS DE PREDICCIÓN DE RIESGO PARA ENFERMEDADES CARDIOVASCULARES EN PACIENTES CON PROBLEMAS DE SUEÑO**

Según la Organización Mundial de la Salud (OMS), las Enfermedades Cardiovasculares (ECV) son la principal causa de muerte en todo el mundo, aún cuando la mayoría de ellas pueden prevenirse actuando sobre ciertos factores de riesgo. Para su prevención son fundamentales la detección precoz y el tratamiento temprano, por medio de cambios en hábitos de consumo y estilo de vida o por medio de la administración de fármacos.

Existen varias formas estándar de estimar el riesgo de enfermedad cardiovascular, siendo el índice de Framingham la más reconocida. Aunque en general estos métodos dan buenos resultados, en muchos casos no son suficientemente precisos en su predicción, lo cual podría derivar en eventos cardiovasculares graves que podrían haberse evitado. Por otro lado, existen factores de riesgo que no se consideran y que podrían implicar una predicción más precisa. En particular, en pacientes con problemas de sueño, una forma de mejorar la predicción de ECV podría ser en base al nivel de saturación de oxígeno en la sangre durante el sueño, que puede ser medido a través un examen médico llamado oximetría nocturna.

El objetivo de esta memoria es precisamente contribuir a la detección temprana de este tipo de enfermedades en este tipo de pacientes de forma más confiable, por lo que se diseña un modelo preventivo en base a factores de riesgo tradicionales y además información recogida de una oximetría nocturna. El modelo se entrena y valida a través de una base de datos que contiene 5.804 oximetrías y el historial clínico de los pacientes en cuanto a las enfermedades cardiovasculares sufridas durante los 10 años posteriores a la oximetría.

La versión final del modelo de predicción tiene un desempeño en términos de AUC de un 0,775 contra un 0,739 del índice de Framingham, por lo que efectivamente se da una mejora estadísticamente significativa. Sin embargo, el aporte de la oximetría nocturna al poder de predicción es mínimo y el alza se debe a otros factores, entre los que se cree estaría la especificidad del modelo en cuanto a su capacidad de realizar predicciones particularmente precisas en pacientes con problemas de sueño .

Se realiza además una evaluación económica y social para dar una idea aproximada del impacto que produciría la implantación del modelo en producción, resultando en un beneficio potencial para todo el país de entre 10 y 13 billones de pesos y 160.000 y 200.000 *Disability-Adjusted Life-Years* (DALYS) y en 10 años. Con el fin de facilitar la divulgación de los resultados, se desarrolla además un primer prototipo de un módulo de visualización del indicador de riesgo del modelo.



# Agradecimientos

Agradezco primero que todo a mis padres, por el apoyo constante que me han dado desde niño para poder estudiar en las mejores condiciones, lleno de oportunidades que me han permitido desarrollarme plenamente, haciendo las cosas que me hacen feliz.

Muchas gracias también a Bárbara Fuentes, que ha sido una pieza clave en los felices que han sido mis últimos cuatro años de universidad.

Agradezco también a mis amigos de la universidad, con quienes compartimos distintas etapas de este proceso: Especialmente a Sergio Orellana, Benjamín Bossi, Pedro Concha, Jaime Capponi, Paola Riadi y Martín Castillo. Con todos compartí muy lindos momentos.

Al mundo universitario del tenis, que me significó una actividad distinta a lo académico, haciendo que mi estadía completa por la universidad no fuera solo estudio, permitiéndome participar de competencias y distenderme en los entrenamientos de todas las semanas con algo que me apasiona. De todo eso me llevo muy gratos recuerdos.

Finalmente, a todos aquellos que me apoyaron constantemente en el desarrollo de mi memoria, incluyendo a todo el equipo del WIC y a Aidan Hogan. En ningún momento dejé de sentir el respaldo de todos y la disposición a ayudarme en lo que fuera que necesitara y además en un ambiente de trabajo particularmente grato.

# Tabla de contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Antecedentes generales . . . . .	1
1.2. Justificación y descripción del proyecto . . . . .	2
1.2.1. Análisis de las soluciones actuales . . . . .	2
1.2.2. Predicción de ECV en base a oximetrías . . . . .	6
1.2.3. Hipótesis de investigación . . . . .	6
1.3. Objetivos . . . . .	7
1.4. Apoyos Institucionales . . . . .	8
1.5. Metodología . . . . .	8
1.5.1. Selección de indicadores . . . . .	8
1.5.2. Implementación de un modelo inicial base . . . . .	8
1.5.3. Desarrollo del modelo final . . . . .	9
1.5.4. Módulo de visualización . . . . .	10
1.5.5. Evaluación del impacto económico . . . . .	10
1.6. Alcances . . . . .	10
1.7. Resultados Esperados . . . . .	11
1.8. Estructura del informe . . . . .	12
<b>2. Marco Conceptual</b>	<b>14</b>
2.1. Datos Utilizados . . . . .	14
2.2. Herramientas tecnológicas . . . . .	14
2.3. Índices de predicción de ECV tradicionales . . . . .	15
2.3.1. Índice de Framingham . . . . .	15
2.3.2. <i>HeartScore</i> o <i>SCORE</i> . . . . .	17
2.3.3. Índice <i>Prospective Cardiovascular Münster (PROCAM)</i> . . . . .	18
2.3.4. Matriz <i>ESH/ESC</i> . . . . .	20
2.4. Otros conceptos médicos relevantes . . . . .	20
2.4.1. Eventos Cardiovasculares . . . . .	20
2.4.2. Polisomnografía y oximetría . . . . .	22
2.4.3. <i>Disability adjusted life years (DALYs)</i> . . . . .	23
2.5. Modelos de <i>Machine Learning</i> . . . . .	23
2.5.1. Regresión logística . . . . .	23
2.5.2. <i>Support Vector Machine (SVM)</i> . . . . .	24
2.5.3. Red Neuronal . . . . .	24
2.5.4. <i>Naive Bayes</i> . . . . .	24
2.6. Conceptos relevantes de <i>Machine Learning</i> y desarrollo de software . . . . .	25
2.6.1. Matriz de confusión y métricas asociadas . . . . .	25

2.6.2.	<i>Cross-validation</i> o validación cruzada [41]	27
2.6.3.	Modelo Vista Controlador (MVC)	28
<b>3.</b>	<b>Desarrollo del modelo de predicción</b>	<b>29</b>
3.1.	Selección de indicadores	29
3.2.	Implementación de un modelo inicial base	31
3.2.1.	Implementación del índice inicial	31
3.2.2.	Implementación de la variable dependiente	32
3.2.3.	Extracción de un conjunto grande de oximetrías	33
3.2.4.	Regresión Logística o Logit	34
3.2.5.	Controlador	35
3.2.6.	<i>Testing</i>	36
3.2.7.	Resultados	36
3.3.	Desarrollo del modelo final	37
3.3.1.	Implementación y desempeño del modelo de Framingham	37
3.3.2.	Implementación de nuevas variables explicativas	39
3.3.3.	Inclusión de variables tradicionales al modelo	40
3.3.4.	Modificación del umbral de clasificación	41
3.3.5.	Exploración de modelos de <i>Machine Learning</i>	42
3.3.6.	<i>Testing</i>	43
3.3.7.	Medición de desempeño y resultados	44
3.4.	Perfeccionamiento del modelo	46
3.4.1.	Mejoras en la metodología de evaluación	46
3.4.2.	Mejoras del modelo	48
3.4.3.	Resultados	48
<b>4.</b>	<b>Módulo de visualización</b>	<b>52</b>
4.1.	Descripción desde el punto de vista del usuario	52
4.2.	Implementación	53
<b>5.</b>	<b>Evaluación del impacto económico y social</b>	<b>58</b>
5.1.	Impacto en el Instituto nacional del Torax	58
5.2.	Impacto potencial en Chile	62
<b>6.</b>	<b>Conclusiones</b>	<b>65</b>
6.1.	Conclusiones generales	65
6.2.	Trabajo Futuro	67
	<b>Bibliografía</b>	<b>68</b>

# Índice de tablas

3.1. Descripción de variables calculadas desde una oximetría consideradas en esta memoria . . . . .	31
3.2. Matriz de confusión del modelo base calculada con la técnica de validación cruzada . . . . .	37
3.3. Matriz de confusión del modelo base calculada con la técnica de separación de los datos utilizando un 30 % para <i>testing</i> . . . . .	37
3.4. Medidas de desempeño con umbral de 5 % . . . . .	44
3.5. Medidas de desempeño con umbral de 10 % . . . . .	44
3.6. Medidas de desempeño con umbral de 15 % . . . . .	44
3.7. AUC para índice de Framingham y modelos construidos . . . . .	46
3.8. Tiempo de ejecución en segundos de cada modelo con <i>cross-validation</i> de 20 <i>folds</i> . . . . .	47
3.9. AUC regresión logística para cada configuración posible de variables explicativas . . . . .	48
3.10. AUC <i>Support Vector Machine</i> para cada configuración posible de variables explicativas . . . . .	49
3.11. AUC Red Neuronal para cada configuración posible de variables explicativas . . . . .	49
3.12. AUC Naive Bayes para cada configuración posible de variables explicativas . . . . .	49
3.13. AUC modelo con únicamente las variables de Framingham como variables explicativas . . . . .	49
3.14. AUC Regresión Logística para cada configuración posible de variables explicativas . . . . .	50
3.15. AUC <i>Support Vector Machine</i> para cada configuración posible de variables explicativas . . . . .	50
3.16. AUC Red Neuronal para cada configuración posible de variables explicativas . . . . .	50
3.17. AUC Naive Bayes para cada configuración posible de variables explicativas . . . . .	50
3.18. Coeficientes modelo final . . . . .	51
5.1. Número de estudios de sueño en INT entre 2010 y 2015 . . . . .	58
5.2. <i>Recall</i> y número de personas que se atenderían preventivamente para cada corte definido . . . . .	59
5.3. Numero de personas que prevendrían una ECV producto de la implementación del modelo . . . . .	60
5.4. Estimación del costo por persona en DALYs y en pesos chilenos de tener una insuficiencia cardiaca o un infarto de miocardio . . . . .	60
5.5. Estimación de beneficios monetarios y en DALYs de la implementación del modelo en el hospital salvador para un corte al 5 % . . . . .	61

5.6. Estimación de beneficios monetarios y en DALYs de la implementación del modelo en el hospital salvador para un corte al 10 % . . . . .	61
5.7. Estimación de beneficios monetarios y en DALYs de la implementación del modelo en el hospital salvador para un corte al 20 % . . . . .	61
5.8. Mercado potencial estimado por año de 2015 a 2030 . . . . .	63
5.9. Estimación de beneficios potenciales monetarios y en DALYs de la implementación del modelo en la población chilena para un corte al 5 % . . . . .	63
5.10. Estimación de beneficios potenciales monetarios y en DALYs de la implementación del modelo en la población chilena para un corte al 10 % . . . . .	63
5.11. Estimación de beneficios potenciales monetarios y en DALYs de la implementación del modelo en la población chilena para un corte al 20 % . . . . .	64



# Índice de figuras

1.1.	Gráfico de una oximetría visualizada en <i>Sleepware</i> , la curva color calipso indica el nivel de saturación de oxígeno en la sangre . . . . .	3
1.2.	Gráfico de una oximetría junto a ciertas métricas calculadas automáticamente en <i>Stardust II</i> . La curva negra corresponde a la saturación de oxígeno	4
1.3.	Parámetros en base a los que <i>Somno Check Micro Cardio</i> calcula un indicador de riesgo cardiovascular . . . . .	5
2.1.	Matriz para la predicción de riesgo de ECV según índice de Framingham para hombres en la población chilena. [3] . . . . .	16
2.2.	Coefficientes para el cálculo del índice de Framingham . . . . .	17
2.3.	Matriz para la predicción de riesgo de ECV según índice SCORE [31] . . . .	19
2.4.	Matriz desarrollada en conjunto por ESH y ESC para la predicción de riesgo de ECV [34] . . . . .	20
2.5.	Factores de riesgo utilizados en la matriz ESH/ESC [34] . . . . .	21
2.6.	Matriz de confusión: mide desempeño de un modelo de clasificación . . . .	26
2.7.	Curva para medir el desempeño de modelos probabilísticos. La línea diagonal representa un modelo sin ningún poder de predicción y entre más se acerca la curva hacia la esquina superior izquierda mayor es la precisión en la predicción del modelo . . . . .	27
3.1.	Modelo auxiliar a Framingham utilizado para cuantificar su desempeño . . .	38
3.2.	Ilustración de una actualización en el <i>buffer</i> durante el algoritmo para el cálculo del indicador <i>SPO2-I</i> . . . . .	40
3.3.	Ilustración de una actualización en el <i>buffer</i> durante el algoritmo para el cálculo del indicador <i>PR-I</i> . . . . .	41
3.4.	Curvas ROC para el índice de Framingham y cada uno de los modelos construidos . . . . .	45
4.1.	Formulario prototipo funcional para el ingreso de variables de predicción de riesgo cardiovascular . . . . .	54
4.2.	Representación gráfica del nivel de riesgo cardiovascular . . . . .	55

# Capítulo 1

## Introducción

### 1.1 Antecedentes generales

Según la organización mundial de la salud [1] (OMS), las enfermedades cardiovasculares (ECV) son la principal causa de muerte a nivel mundial. Se calcula que en 2012 murieron por esta causa 17,5 millones de personas, lo cual representa un 31 % de todas las muertes registradas en el mundo. En 2014 en Chile, las enfermedades asociadas al sistema circulatorio correspondieron a un 27.5 % del total de las muertes [2].

La mayoría de las ECV pueden prevenirse actuando sobre factores de riesgo como el consumo de tabaco, las dietas malsanas y la obesidad, la inactividad física o el consumo nocivo de alcohol [3]. Para su prevención son fundamentales la detección precoz y el tratamiento temprano, por medio de servicios de orientación o la administración de fármacos, según corresponda.

Existen varias formas estándar de estimar el riesgo de enfermedad cardiovascular en base a factores como la edad, índice de colesterol, presión sanguínea, presencia de diabetes, consumo de tabaco, etc. El índice más popular se conoce como índice de Framingham [4] y se calcula en base a un modelo bien validado que lleva el mismo nombre (modelo de Framingham). Existen varios otros índices similares también validados y que se basan en factores similares.

Si bien los modelos de predicción de ECV en general dan buenos resultados, en muchos casos no son suficientemente precisos en su estimación [5]. Esto es un problema que podría ser mitigado, pues en índices tradicionales no se consideran ciertos factores de riesgo que, en algunos casos, podrían ser determinantes para una mejor predicción.

Por otra parte, existen una serie de problemas de sueño que son cada vez más comunes en Chile y el mundo. El insomnio, la hipersomnia y problemas respiratorios al dormir son solo algunos de ellos. En la encuesta de salud de 2016 cerca de un 10 % de las personas califica su calidad de sueño como mala o muy mala [6].

Recientemente ha surgido además investigación que muestra evidencias de una correla-

ción entre problemas de sueño y otro tipo de enfermedades, como cáncer [7, 8], problemas al cerebro [9, 10] y justamente enfermedades cardiovasculares [10-12]. Un posible buen indicador de riesgo vascular según la literatura es nivel de saturación de oxígeno en la sangre durante el sueño [10, 11], que puede ser medido en un examen médico conocido como oximetría nocturna. Por lo tanto, en pacientes con problemas de sueño se podría diseñar una forma alternativa de predecir el nivel de riesgo cardiovascular en base a dicho indicador.

## 1.2 Justificación y descripción del proyecto

La evidencia indica claramente que quienes tienen problemas de sueño tienen alta probabilidad de contraer enfermedades cardiovasculares. Así lo afirma en base a varios estudios, por ejemplo, *National Sleep Foundation* [13], reconocida organización sin fines de lucro dedicada a promover en Estados Unidos buenos hábitos de sueño. Se sabe, por ejemplo, que la Apnea Obstructiva Central incrementa significativamente el riesgo de un ataque al corazón [12]. Es muy común además que a pacientes con este tipo de problemas se les solicite realizarse oximetrías nocturnas, lo que abre una ventana a instituciones médicas para que utilicen dicha información para realizar predicciones de riesgo cardiovascular.

Dado lo anterior, se realizó una revisión de las alternativas de software que existen actualmente que pudieran servir para la predicción de ECV en pacientes con problemas de sueño y además una investigación del estado del arte en cuanto a formas de predicción de ECV basadas en oximetrías. El objetivo es tener una comprensión acabada del problema para poder determinar de que forma ser una contribución.

### 1.2.1 Análisis de las soluciones actuales

En primer lugar, se hace un análisis de las distintas alternativas de solución disponibles en el mercado, tomando en cuenta que actualmente existe *software* que se vende conjuntamente con el aparato que realiza las oximetrías en los pacientes. En base a la información que aportan, podría predecirse riesgo cardiovascular apenas el paciente se realiza el examen y en base a eso tomar acciones preventivas.

Se revisaron 3 programas disponibles actualmente en el mercado para analizar de que forma podrían servir para la predicción: *Stardust II* [14], *Sleepware* [15] y *Somno Check Micro Cardio* [16]

*Stardust* y *Sleepware* forman parte de esta revisión pues fueron facilitados por el Instituto Nacional del Tórax (INT) para este propósito. Ambos programas se utilizan a diario en muchas instituciones médicas y no están contruidos exclusivamente para visualizar nivel de saturación de oxígeno sino que tienen muchas otras funcionalidades. Si bien existen otras alternativas de software similares en el mercado, se recibió confirmación de parte del doctor Juan Carrillo, médico e investigador de la Unidad de Sueño del INT, de que son muy parecidas en cuanto a su funcionalidad, por lo que estos dos representan muy bien lo que el mercado ofrece.

*Somno Check Cardio*, por otra parte, fue escogido por ser el único software que se encontró en el mercado cuya principal función es calcular riesgo cardiovascular en base a desaturación de oxígeno. No se dispone de acceso al software pero sí al manual de uso y a información sobre la forma en que fueron validados los indicadores que utiliza.

Tanto *Stardust II* como *Sleepware* despliegan un gráfico que indica como varía a través del tiempo el nivel de saturación de oxígeno del paciente. Una forma de extraer información de los datos y de predecir a partir de eso el riesgo de enfermedades es a través del análisis visual de los médicos directamente desde el gráfico. Eso se hace actualmente pero sólo puede entregar una idea muy poco precisa del nivel de riesgo. En casos de problemas de sueño muy severos, por ejemplo, puede efectivamente determinarse que existe un riesgo alto si hay desaturaciones muy marcadas en el gráfico, pero si las desaturaciones son menos evidentes se vuelve infactible por la complejidad de la curva. La figura 1.1 muestra un caso de ejemplo de un gráfico para una oximetría.

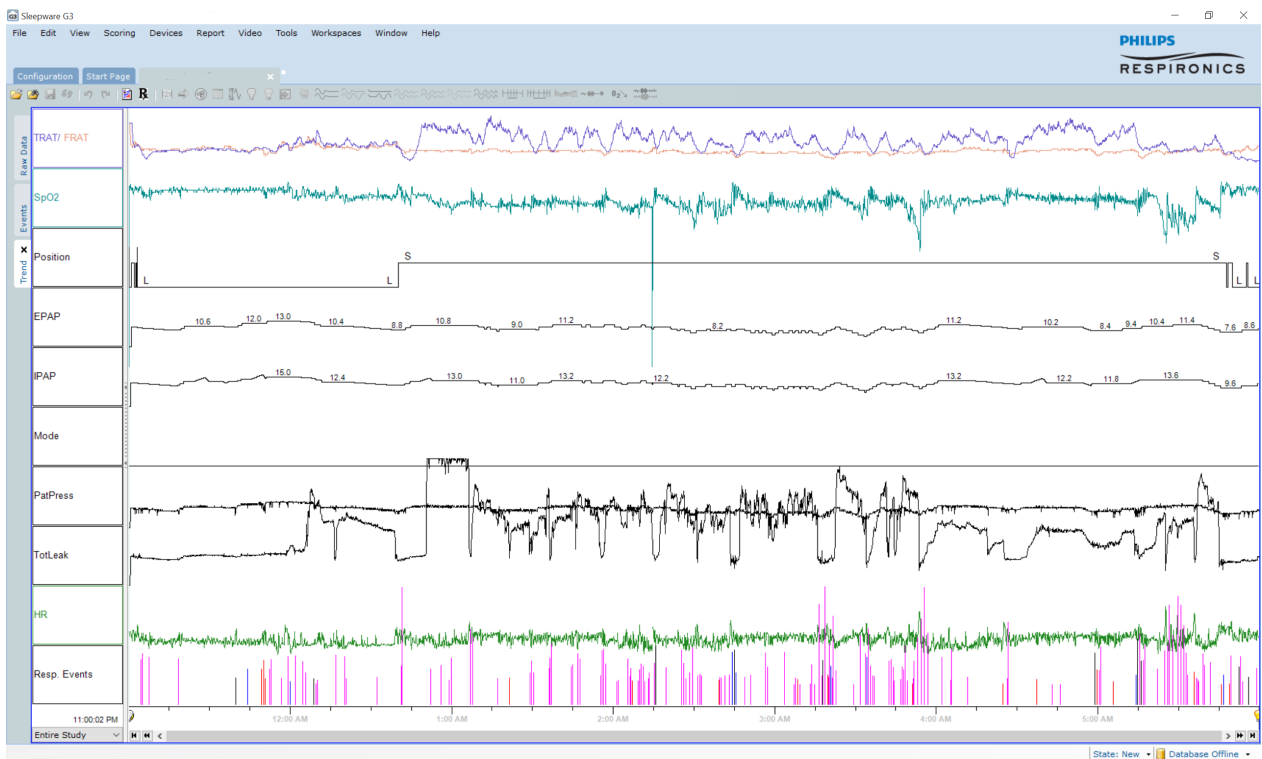
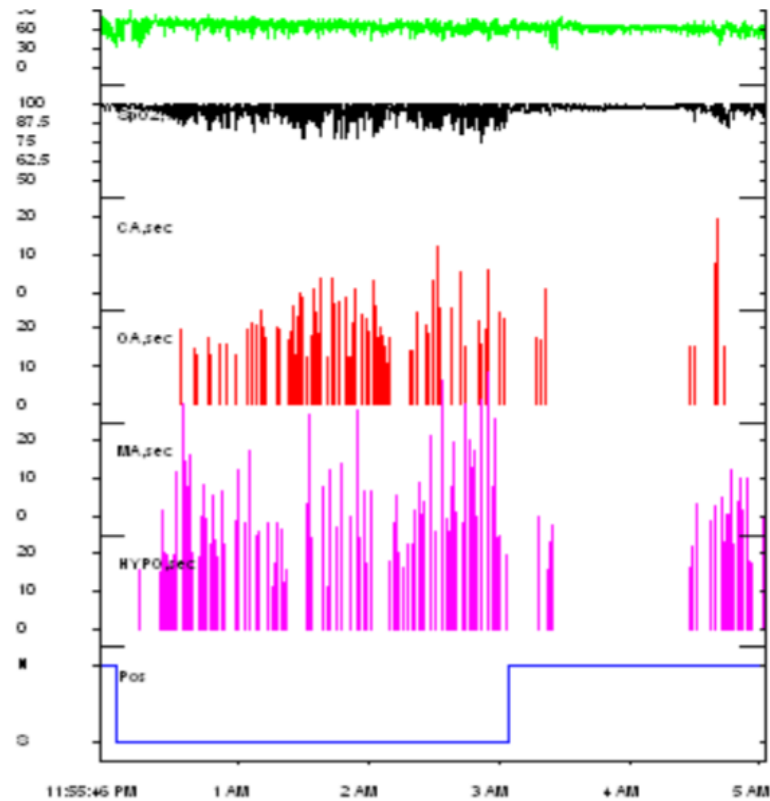


Figura 1.1: Gráfico de una oximetría visualizada en *Sleepware*, la curva color calipso indica el nivel de saturación de oxígeno en la sangre

Por otra parte, tanto *Stardust II* como *Sleepware* entregan métricas que ayudan a extraer información desde los datos. Por ejemplo, muestran el promedio de saturación y el tiempo total en que la saturación fue menos de 90 %. Sin embargo, existe evidencia que apunta a que se pueden calcular mejores predictores en base a la oximetría. Frecuentemente dos gráficos que a la vista de un médico representan a pacientes muy distintos y que eventualmente pudieran tener un riesgo de ECV muy diferente muestran métricas muy parecidas. Además, según conversaciones sostenidas con doctores en el Instituto Nacional del Tórax y el hospital del Salvador, incluso si alguno de estos índices fuera buen predictor, no es algo que sea de conocimiento general por lo que un médico normalmente no usará el dato si el software no se lo indica expresamente o no recibe una capacitación al respecto.

En la figura 1.2 se aprecia una oximetría junto con las métricas que el programa calcula de forma automática.



Oxime try

<95 % (minutes)	68	Average (%)	96
<90 % (minutes)	25	Desat Index (#/hour)	41.2
<85 % (minutes)	8.5	Desat Max (%)	29
<80 % (minutes)	2.5	Desat Max dur (sec)	56
<75 % (minutes)	0	Lowest SpO <sub>2</sub> (≥ 2 sec) (%)	72
<70 % (minutes)	0	# Episodes (≥ 5 min) ≤ 88%	0
<60 % (minutes)	0	Longest dur (min) SpO <sub>2</sub> ≤ 88%	0.6
<50 % (minutes)	0		
< 97 % (minutes)	117.5		

Figura 1.2: Gráfico de una oximetría junto a ciertas métricas calculadas automáticamente en *Stardust II*. La curva negra corresponde a la saturación de oxígeno

Por último, *Somno Check Micro Cardio* sí calcula un indicador pensado especialmente para predecir riesgo cardiovascular, por lo tanto se ajusta mucho mejor a las necesidades buscadas. La figura 1.3 ilustra los parámetros sobre a los cuales se basa el indicador.

A pesar de haber sido creado específicamente para la predicción de riesgo de ECV, tiene dos limitaciones principales. La primera de ellas es que el indicador que utiliza fue validado en base al nivel de correlación que tienen los parámetros con índices cardiovasculares tradicionales. Por lo tanto, no hay ninguna garantía de que en alguna circunstancia ten-

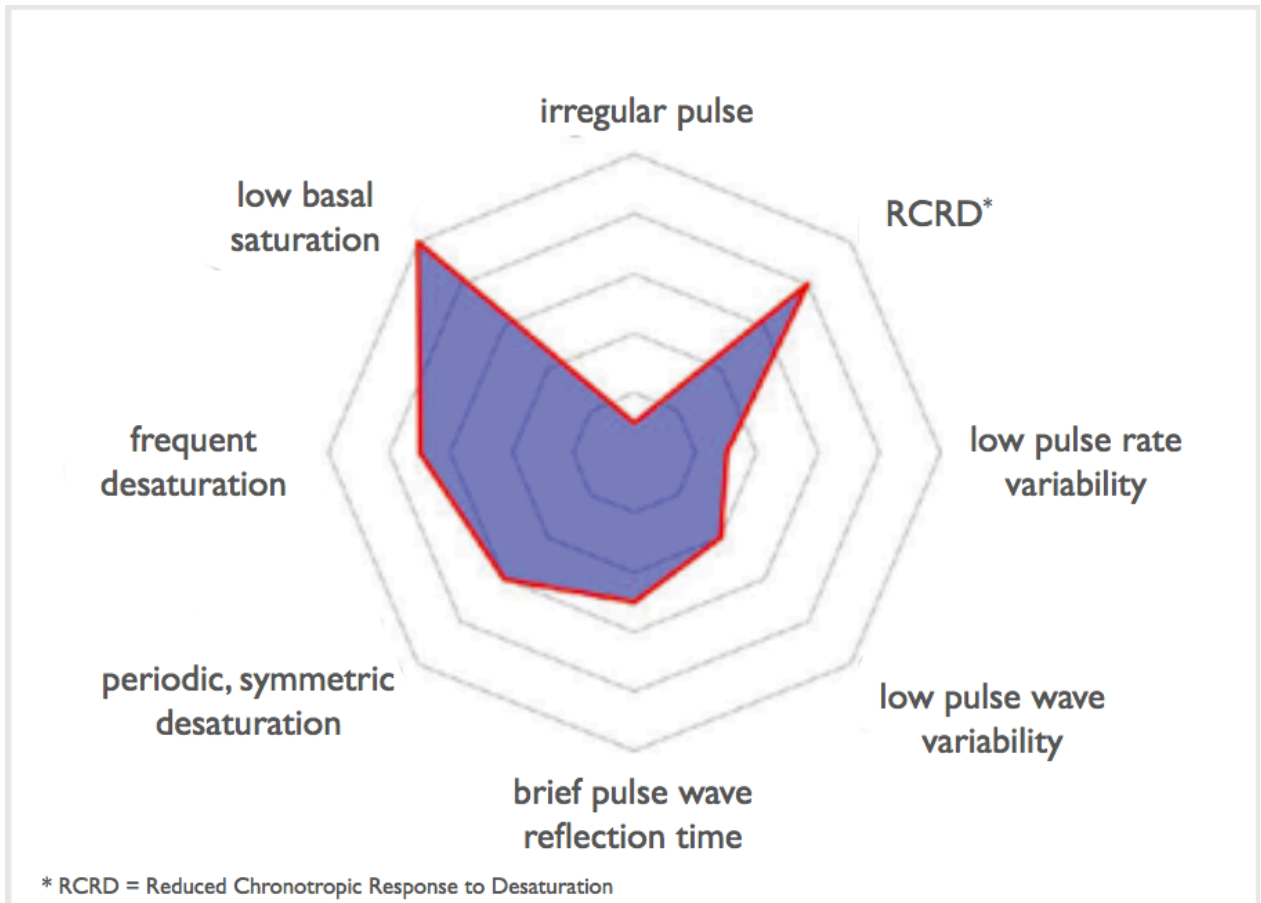


Figura 1.3: Parámetros en base a los que *Somno Check Micro Cardio* calcula un indicador de riesgo cardiovascular

ga mayor poder predictivo que modelos tradicionales y no puede suponerse que se haga cargo de la imprecisiones que ellos presentan. Podría ocurrir, por ejemplo, que el índice acertadamente indique un alto riesgo de ECV sólo en cierto porcentaje de entre todas las ocasiones en que el índice de Framingham hace la misma predicción, pero que el resto del tiempo ambos índices se comporten de forma equivalente. En ese caso, lo mejor en cuanto a la precisión de la predicción sería seguir utilizando el índice de Framingham sin ningún tipo de modificación. La segunda limitación es que el índice no fue validado en Chile y no puede suponerse que funcionará de la misma forma en nuestro país. Es sabido por ejemplo que el modelo de Framingham en algunos casos sobrestima y en otros subestima el riesgo de ECV [17-19]. De hecho, en Chile debió ser adaptado por la Universidad de Talca para nuestra población [20].

### **1.2.2 Predicción de ECV en base a oximetrías**

Por otro lado, existe investigación que puede tomarse como base para a futuro predecir de mejor forma el riesgo de enfermedades cardiovasculares en base a oximetrías. Tras una búsqueda bibliográfica se seleccionaron 2 artículos que muestran indicios claros de una correlación entre indicadores calculados en base a una oximetría y enfermedades cardiovasculares y que pueden utilizarse para el diseño del modelo de predicción de esta memoria:

- En el artículo de Sommermeyer, D.Zou, et al. [11] se revisa si existe correlación entre distintos indicadores recogidos desde una oximetría y los principales indicadores de riesgo de ECV (Índices de Framingham, PROCAM, SCORE y ESH/ESC) y en efecto se encuentra correlación significativa para cada uno de los indicadores recogidos.
- Según E. Watanabe, K. Kiyono et al. [21] se observa que en pacientes con Apnea Central, existe un índice que se correlaciona significativamente con la tasa de mortalidad asociada a enfermedades cardiovasculares.

### **1.2.3 Hipótesis de investigación**

Los antecedentes mencionados dan pie a la siguiente hipótesis para esta memoria:

- Se puede construir un indicador en base a oximetrías nocturnas y variables demográficas y clínicas que sirva como una alternativa más confiable al índice de Framingham en pacientes con problemas de sueño.

El objetivo de validar dicha hipótesis es contribuir a que en un futuro se implemente en centros de salud en Chile una herramienta que permita a los doctores detectar de forma temprana y con alta precisión enfermedades cardiovasculares en pacientes con problemas de sueño.

En caso de que la hipótesis propuesta se valide, y en base a lo desarrollado en esta memoria se podrá a futuro desarrollar un software que extienda las soluciones presentes ac-

tualmente en el mercado. De este modo, se podrá entregar una herramienta con la que médicos puedan sacar conclusiones más acertadas y de forma simple sobre el riesgo de ECV en pacientes con problemas de sueño.

### 1.3 Objetivos

El objetivo general se redacta de la siguiente forma:

- Seleccionar, implementar y validar un modelo para la predicción de riesgo cardiovascular a través de técnicas de *Machine Learning* que sirva como una alternativa más confiable a índices tradicionales en pacientes con problemas de sueño y sienten las bases para el desarrollo de una herramienta de apoyo al diagnóstico médico que permita tomar acciones preventivas más certeras.

Cabe mencionar que el cumplimiento del objetivo dependerá fuertemente de que se valide la hipótesis de investigación. En caso contrario se entregará una descripción completa de la investigación que contribuya a que futuros investigadores establezcan de forma más informada los pasos a seguir en el futuro.

Los objetivos específicos son los siguientes:

- Seleccionar en base a investigación previa los indicadores de riesgo de ECV derivados de una oximetría nocturna que permitirían de mejor manera contribuir al poder predictivo de un modelo de predicción de riesgo cardiovascular.
- Diseñar y construir potenciales modelos de predicción de ECV en base a indicadores calculados a partir de una oximetría nocturna.
- Validar y comparar potenciales modelos de predicción de ECV en base a indicadores calculados a partir de una oximetría nocturna, permitiendo seleccionar el de mejor desempeño.
- Implementar un módulo de visualización de un índice de riesgo cardiovascular calculado en base a un modelo validado de predicción de ECV.
- Cuantificar el potencial beneficio económico y social de implementar una herramienta de apoyo al diagnóstico médico basado en el modelo desarrollado

Sobre los últimos dos puntos cabe aclarar que sólo serán desarrollados en caso de que la hipótesis se valide.



## 1.4 Apoyos Institucionales

La investigación cuenta con el apoyo del INT y el centro de investigación *Web Intelligence Centre* (WIC). Ambas instituciones han conformado una alianza que pretende, una vez concluido el trabajo del alumno, continuar con el desarrollo del proyecto.

En el INT se cuenta con el apoyo de la Unidad del Sueño, que es el principal centro de derivación a nivel nacional de pacientes con problemas de sueño. En su representación se cuenta con el apoyo del doctor Juan Carrillo, médico del equipo de Medicina del Sueño. Él proporcionará guía, datos e información del funcionamiento del centro médico y de aspectos técnicos del área de la medicina que deben ser considerados en esta memoria. A su vez, el hospital ayudará con la validación del trabajo desde la opinión experta del área médica.

El Centro de Inteligencia Web (WIC) es una organización dedicada a hacer investigación aplicada buscando generar valor en la industria en base a herramientas de *Data Science*. De parte de ellos se recibirá constante apoyo y guía para el desarrollo de la tesis y eventualmente recursos económicos si fuera necesario.

## 1.5 Metodología

En resumen, la metodología consiste en la implementación y posterior validación de una serie de indicadores que luego se complementan para crear en conjunto un modelo de predicción de riesgo cardiovascular. Los indicadores son recogidos de literatura relacionada tras una búsqueda bibliográfica y la validación se hace principalmente en base a datos de *National Sleep Research Resource (NSRR)* [22].

El proceso consta de seis etapas, las cuales se detallan a continuación:

### 1.5.1 Selección de indicadores

El primer paso consiste en seleccionar un conjunto de indicadores que hayan tenido resultados alentadores en investigaciones previas y que, por lo tanto, se sospechen de utilidad para la validación de la hipótesis planteada en esta memoria. Dicha selección se realizará apoyada por profesionales del área en el Instituto Nacional del Tórax. Se concluye con la redacción del capítulo 3.1 en donde se señalan y describen los índices seleccionados.

### 1.5.2 Implementación de un modelo inicial base

Tras el paso anterior se crea una primera versión de un modelo de predicción de ECV. Para ello se realizará lo siguiente:

1. Se escoge y se implementa uno de los indicadores de entre la selección previa. Esto se realiza a partir de un archivo en formato EDF [23], que es el formato estándar de almacenamiento de este tipo de datos.
2. Se entrena una regresión logística para predecir riesgo de ECV. Para esto se utilizan datos de *NSRR*, en particular la base de datos *Sleep Heart Health Study (SHHS)* [24], que cuenta con oximetrías de 5804 pacientes mayores de 40 años tomadas entre los años 1995 y 1998 y además un seguimiento del historial clínico de estos pacientes en cuanto a las ECV padecidas posteriormente. Para la regresión se utiliza una variable dependiente que tiene dos posibles valores según si se contrajo o no alguna enfermedad cardiovascular en los siguientes 10 años.

Esta sección finaliza con la implementación de un primer indicador basado en una oximetría y con la implementación de un primer modelo de predicción que corresponde a una regresión logística de una única variable explicativa. Además, se obtiene información de primera mano de su poder predictivo. Todo esto se refleja en la redacción de la sección 3.2.

### 1.5.3 Desarrollo del modelo final

Se repite el proceso de implementación de la parte anterior para los indicadores faltantes. Con todos los indicadores ya implementados, se construyen una serie de modelos de predicción de ECV que deben ser entrenados y validados también a partir de la base de datos SHHS. El objetivo es poder comparar y escoger la mejor forma de predicción de entre una gama de modelos:

1. Red neuronal, *Support Vector Machine (SVM)*, clasificador bayesiano y regresión logística que consideren como variables explicativas cada uno de los indicadores
2. Red neuronal, *SVM*, clasificador bayesiano y regresión logística que consideren cada uno de los indicadores al mismo tiempo y que además consideren todas las variables explicativas clásicas del índice de Framingham.
3. Red neuronal, *SVM*, clasificador bayesiano y regresión logística que consideren únicamente las variables explicativas clásicas del índice de Framingham.

Además se implementa el índice de Framingham sin ningún tipo de modificación para cada uno de los pacientes.

Si los modelos del punto 2 dan resultados significativamente mejores que el índice de Framingham, se considera validada la hipótesis planteada. Si además, los resultados son significativamente mejores que los del punto 3, se confirma que las variables calculadas desde la oximetría son un aporte real al desempeño del modelo.

Para el entrenamiento y validación de estos modelos se utiliza la misma variable dependiente escogida para la validación del primer indicador, es decir, una variable de sólo dos posibles valores según si se contrajo o no alguna enfermedad cardiovascular en los

siguientes 10 años.

Se exhibe y valida el resultado obtenido a profesionales del área en el Instituto Nacional del Tórax, haciendo ajustes de ser pertinente.

Se concluye con varios modelos predictivos de mayor complejidad y que se espera tengan un mejor poder de predicción que el modelo inicial. Se obtiene además una validación de la implementación realizada por parte de profesionales del área. Todo esto se ve reflejado en la redacción de la sección 3.3.

#### **1.5.4 Módulo de visualización**

Este paso y los que vienen a continuación se realizan sólo si la hipótesis se encuentra validada. Consiste en la construcción de un primer prototipo funcional de un módulo de visualización que provea una interfaz para desplegar gráficamente la predicción del modelo en base a un archivo con datos de una oximetría en formato EDF, datos demográficos e historial médico del paciente.

Esta sección concluye con el prototipo funcional implementado y su documentación, reflejada en el capítulo 4.

#### **1.5.5 Evaluación del impacto económico**

Se realiza finalmente una estimación del impacto que conllevaría el uso del modelo encontrado como predictor de riesgo de ECV. Esto en cuanto a la mejora en la calidad de vida de los pacientes, el prolongamiento de su esperanza de vida y la mayor eficiencia en términos económicos en los tratamientos debido a una predicción temprana. Para cuantificar el aumento de la esperanza y la calidad de vida se utilizará la medida *Disability-Adjusted Life-Years* (DALYs) 2.4.3, que es aceptada internacionalmente por la comunidad científica para este propósito.

Se finalizara con la estimación justificada en el capítulo 5.

### **1.6 Alcances**

El foco de esta memoria está en la validación de la hipótesis propuesta. En el caso de que se valide, se espera obtener un modelo de predicción de riesgo de ECV cuya relevancia y validez esté bien justificada y un módulo funcional de visualización del indicador. Además, se entregará una estimación del potencial impacto económico y social que significaría la implementación de una herramienta de apoyo al diagnóstico médico basada en el modelo.

La validación de los indicadores tendrá la limitación de que será realizada con datos de Estados Unidos. Es fundamental que, una vez concluida la memoria, en caso de que la hi-

pótesis se valide, se replique el proceso de entrenamiento y validación con datos en Chile. Para este propósito, el Instituto Nacional del Tórax se encuentra actualmente construyendo una base de datos. Por lo tanto, el objetivo de esta memoria no es obtener un producto final sino pavimentar el camino para que en unos años se haya desarrollado un producto efectivo y validado que mejore la predicción de riesgo de ECV en pacientes con problemas de sueño. Por lo demás, si bien se espera que el producto sirva de apoyo al diagnóstico médico, no pretende que se logre prescindir del criterio médico al momento de hacer un diagnóstico. Tampoco pretende desplazar a otros índices de riesgo cardiovascular como el índice de Framingham, ya que su utilidad está acotada únicamente a pacientes con problemas de sueño que se realicen una oximetría.

El módulo de visualización consistirá en una interfaz donde un usuario podrá cargar una oximetría e ingresar cierta información demográfica y clínica del paciente. En base a ello podrá visualizar el riesgo cardiovascular estimado por el modelo. No pretende ser un producto final sino sólo una forma de guiar el proceso de validación e implementación que se realizará posteriormente con datos chilenos, además de ser una forma de motivar investigaciones futuras.

No se espera tampoco explotar todo el potencial de una oximetría nocturna para predicción de enfermedades. Existe investigación que apunta a predecir otro tipo de enfermedades a través de oximetrías que no será abordada en esta memoria [7-10]. Existen además diversas propuestas de parte del Instituto Nacional del Tórax sobre como seguir mejorando la capacidad de predicción del modelo que no están contempladas dentro del alcance de esta memoria.

En cuanto a la evaluación económica que se desarrolla en esta memoria, se estiman el beneficio monetario y en términos de mejora en la calidad de vida de la implementación del modelo. No se estiman los costos de implementación y de realización de tratamientos por lo que no se determina viabilidad o rentabilidad del proyecto.

Por último, una vez que culmine el desarrollo de un producto final, está contemplado realizar un proceso de capacitación para los médicos en los lugares que se utilice. Esto tampoco forma parte del alcance de esta memoria.

## **1.7 Resultados Esperados**

Se espera entregar un prototipo que consista en una interfaz que indique el riesgo cardiovascular de un paciente en base a un modelo de predicción seleccionado por el memorista que se calcule en base a datos obtenidos a partir de una oximetría nocturna. Junto con ello, se espera entregar una justificación de la validez y relevancia del indicador representado en el módulo.

Específicamente, se plantean los siguientes resultados esperados:

1. Selección en base a investigaciones previas de los indicadores derivados de la oximetría que mejores indicios den de aportar al poder predictivo de un modelo de predic-

ción de ECV. Se espera definir una lista de indicadores que reflejen el resultado de la selección junto a una justificación del proceso.

2. Diseño y construcción de modelos de predicción de ECV en base a indicadores seleccionados, con datos de NSRR. Concretamente se espera desarrollar el código en *Python* que represente el modelo junto a una descripción y justificación de lo realizado.
3. Validación y comparación de modelos. Se espera desarrollar en *Python* una metodología de evaluación de los distintos modelos, permitiendo una comparación entre ellos, escogiendo así el de mejor desempeño. Todo ello descrito y justificado en este informe.
4. Implementación de un módulo de visualización. Se espera obtener una herramienta que sirva de prototipo en la cual un usuario pueda ingresar los datos de un paciente con problemas de sueño y obtener de forma automática la predicción que realiza el modelo sobre el nivel de riesgo de ECV.
5. Cuantificación del beneficio económico que implicaría la implantación del modelo tanto exclusivamente en el INT como en todo el país. Se espera obtener una estimación del beneficio económico y social de aquello, estableciendo un rango, en pesos y en *Disability-Adjusted Life Years* (ver 2.4.3) que representaría el potencial beneficio y justificando el proceso.

## 1.8 Estructura del informe

Este informe se divide en seis capítulos, que se describen a continuación:

1. El primer capítulo introduce el tema de esta memoria, explica la necesidad planteada y justifica la relevancia del proyecto. Se especifican además los objetivos y el alcance de la memoria.
2. El segundo capítulo corresponde al Marco Conceptual del trabajo, en donde se describen aspectos fundamentales para la comprensión de la memoria, como conceptos médicos, herramientas matemáticas, características de los datos, etc.
3. El capítulo tres describe todo el desarrollo del modelo de predicción, desde la selección de indicadores de la oximetría a implementar hasta la obtención del modelo final.
4. El capítulo cuatro describe el módulo de visualización implementado, tanto del punto de vista del usuario como en cuanto a su implementación.
5. El quinto capítulo realiza una estimación del impacto económico y social que implicaría la implementación del modelo de predicción cardiovascular tanto de forma exclusiva en el Instituto Nacional del Tórax como a lo largo de todo el país.

6. El sexto y último capítulo resume el proceso y los resultados obtenidos y propone formas de seguir avanzando en la investigación que esta memoria inicia.

# Capítulo 2

## Marco Conceptual

### 2.1 Datos Utilizados

Se utiliza el juego de datos SHHS. Se trata de información recolectada a través de un estudio realizado entre 1995 y 2010 con 5804 personas en diferentes zonas de Estados Unidos. Para cada individuo se cuenta con una oximetría, datos clínicos y demográficos y seguimiento clínico de eventos cardiovasculares durante 10 años posteriores a la realización de la oximetría. La data se compone de 2765 hombres y 3039 mujeres de al menos 40 años.

### 2.2 Herramientas tecnológicas

Se describen las principales herramientas tecnológicas utilizadas durante el desarrollo de esta memoria:

- **Lenguaje de programación:** Se escoge Python 3.5.2 [25] como lenguaje de programación para la implementación de los modelos de *Machine Learning* y el módulo de visualización. Se escoge porque permite un desarrollo rápido con programación orientada a objetos, porque el alumno ya estaba familiarizado con el lenguaje al momento de comenzar su memoria y porque es uno de los lenguajes que se utilizan habitualmente en el WIC para el desarrollo de proyectos de investigación. Para la capa de presentación del módulo funcional se utiliza además html y librerías css y javascript [26].
- **Framework de desarrollo web:** Para el desarrollo del prototipo funcional del módulo de visualización se utiliza el *framework* de desarrollo web *Django* [27]. Se trata de un sistema *open source* que es ampliamente utilizado en la industria de desarrollo de software y destaca por permitir un desarrollo rápido con un diseño fácilmente escalable. Se escoge además porque se basa en el lenguaje de programación Python, que es el que se utiliza para el desarrollo del modelo en esta memoria y por lo tanto

la implementación del mismo en el módulo funcional se vuelve mucho más sencilla. Por último, el alumno ya estaba familiarizado con el *framework* lo que agiliza aún más el desarrollo.

- **Hardware:** Se utiliza un computador personal con las siguientes características:
  - **CPU:** 2,9 GHz Intel Core i5 .
  - **Memoria RAM:** 8 GB 1867 MHz DDR3.
  - **Disco Duro:** 499,06 GB GB SSD.
  - **Sistema Operativo:** OS X 10.10.5.
- **Sistema de versionamiento:** Se utiliza `git` [28], que es un sistema de versionamiento *open source* y gratuito que es ampliamente utilizado en la industria de desarrollo de software.

## 2.3 Índices de predicción de ECV tradicionales

Existen diversas formas de estimar el riesgo de ECV. La más conocida y la más comúnmente utilizada en Chile es el índice de Framingham. Sin embargo, existen varias otras. En esta memoria se utiliza el índice de Framingham como punto de comparación ya que, además de ser altamente reconocido, fue desarrollado en base a datos de Estados Unidos, que es precisamente de donde son los datos que se utilizan en esta memoria. En esta sección se describen los principales indicadores de riesgo identificados:

### 2.3.1 Índice de Framingham

El índice de Framingham [4] original surgió en 1948 tras una investigación encausada por el instituto *National Heart, Lung, and Blood Institute(NHLBI)*, en Estados Unidos.

Dicha organización ha trabajado desde entonces en este índice con datos de varios cientos de miles de pacientes, muchos de los cuales han recibido un seguimiento clínico durante varias décadas. Durante este tiempo han surgido diversas variantes del índice. Para esta memoria se utiliza la versión más reciente que no requiere de instrumentos de laboratorio. Esto pues la idea es que el cálculo del modelo no requiera de instrumentos adicionales una vez realizada la oximetría. Se estima el riesgo como una probabilidad de contraer una ECV en 10 años y se toma en consideración los siguientes factores:

- Edad
- Género
- Presencia de Diabetes



- Consumo de tabaco
- Presión sanguínea sistólica
- Utilización de tratamientos para regular la presión sanguínea
- Índice de masa corporal

El índice de Framingham ha sido adaptado a muchas poblaciones incluida la chilena (aunque en el caso de nuestro país, el índice utiliza el nivel de colesterol y no el índice de masa corporal para la predicción). Existe además una matriz (Ver Figura 2.1) que permite asociar cada probabilidad de riesgo a un nivel de riesgo en una población en particular, que puede ser bajo, moderado, alto o muy alto.

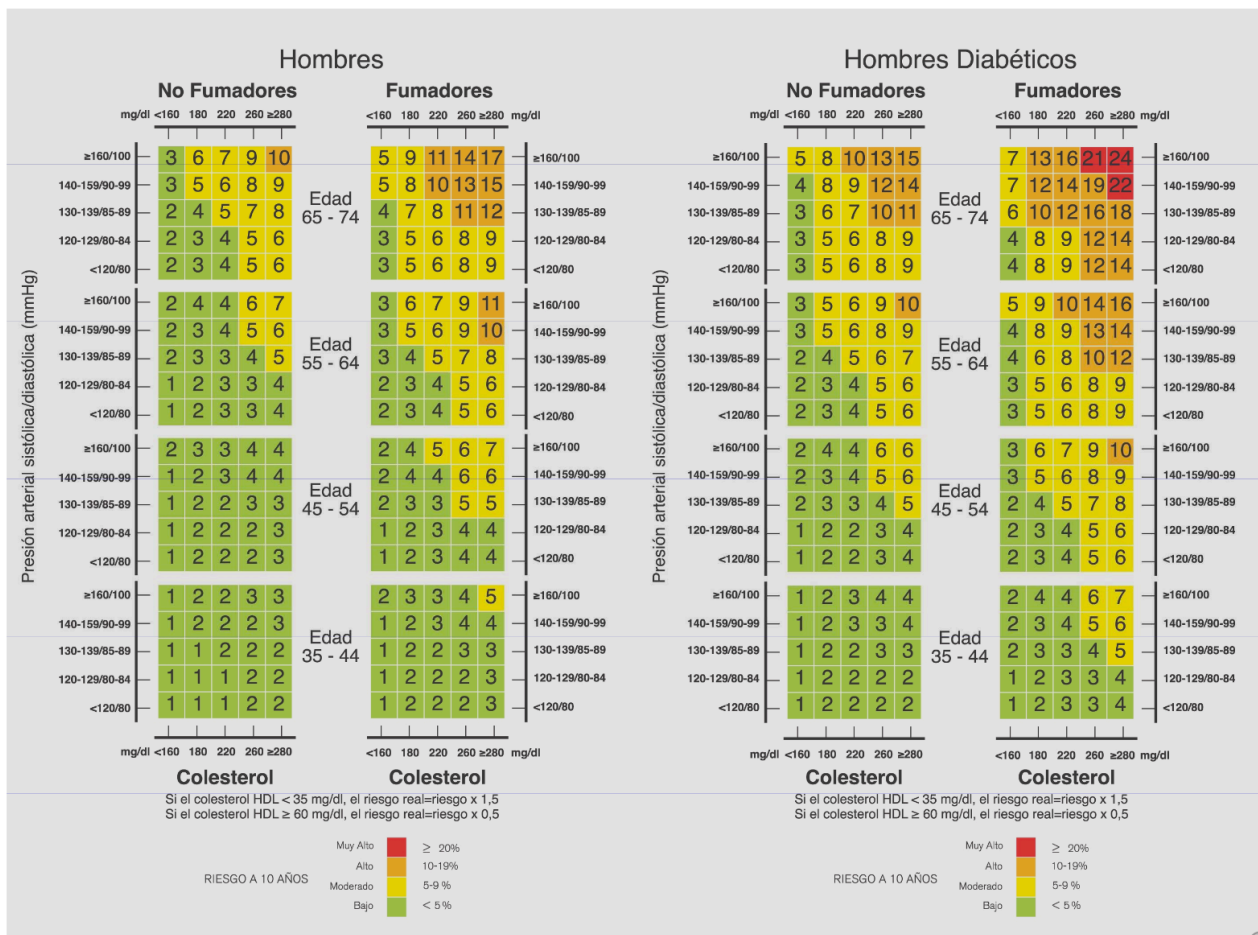


Figura 2.1: Matriz para la predicción de riesgo de ECV según índice de Framingham para hombres en la población chilena. [3]

Para el cálculo de la probabilidad se basa en las fórmulas 2.1 y 2.2:

$$1 - 0,94833^{exp(x\beta - 26,0145)} \quad (2.1)$$

$$1 - 0,88431^{exp(x\beta - 23,9388)} \quad (2.2)$$

La fórmula 2.1 se utiliza para pacientes mujeres y la fórmula 2.2 para pacientes hombres. Por otra parte, el término  $x\beta$  se refiere al producto punto entre un vector de coeficientes y un vector con las respectivas variables asociadas, que son edad, índice de masa corporal, presión sistólica, consumo de tabaco y presencia de diabetes. La figura 2.2 muestra el valor de cada uno de los coeficientes. Como se observa, estos coeficientes son diferentes en mujeres que en hombres. Además, el coeficiente asociado a la presión sistólica es diferente según si el paciente está recibiendo o no algún tratamiento para controlar la presión.

Men* (10-year Baseline Survival: So(10) = 0.88431)				
Variable	Beta**	p-value	Hazard Ratio	95% CI
Log of Age	3.11296	<.0001	22.49	(14.80, 34.16)
Log of Body Mass Index	0.79277	<.0066	2.21	(1.25, 3.91)
Log of SBP if not treated	1.85508	<.0001	6.39	(3.61, 11.33)
Log of SBP if treated	1.92672	<.0001	6.87	(3.90, 12.08)
Smoking	0.70953	<.0001	2.03	(1.75, 2.37)
Diabetes	0.53160	<.0001	1.70	(1.37, 2.11)

Women* (10-year Baseline Survival: So(10) = 0.94833)				
Variable	Beta**	p-value	Hazard Ratio	95% CI
Log of Age	2.72107	<.0001	15.20	(8.59, 26.87)
Log of Body Mass Index	0.51125	<.0609	1.67	(0.98, 2.85)
Log of SBP if not treated	2.81291	<.0001	16.66	(8.27, 33.54)
Log of SBP if treated	2.88267	<.0001	17.86	(8.97, 35.57)
Smoking	0.61868	<.0001	1.86	(1.53, 2.25)
Diabetes	0.77763	<.0001	2.18	(1.63, 2.91)

Figura 2.2: Coeficientes para el cálculo del índice de Framingham

### 2.3.2 *HeartScore* o *SCORE*

El índice *SCORE* o *HeartScore* [29] fue desarrollado por la Sociedad Europea de Cardiología (“European Society of Cardiology”) [30] y publicado en 2003. Es similar al de

Framingham pero fue desarrollado en base a datos de poblaciones europeas y por lo tanto pensado para aplicarse en poblaciones europeas. Se basa en datos recolectados de alrededor de 250.000 pacientes en 12 países de Europa. Existen 2 modelos y se ocupa uno u otro dependiendo del país de origen del paciente. Considera los siguientes factores de riesgo:

- Género
- Edad
- Consumo de Tabaco
- Presión sistólica
- Colesterol total

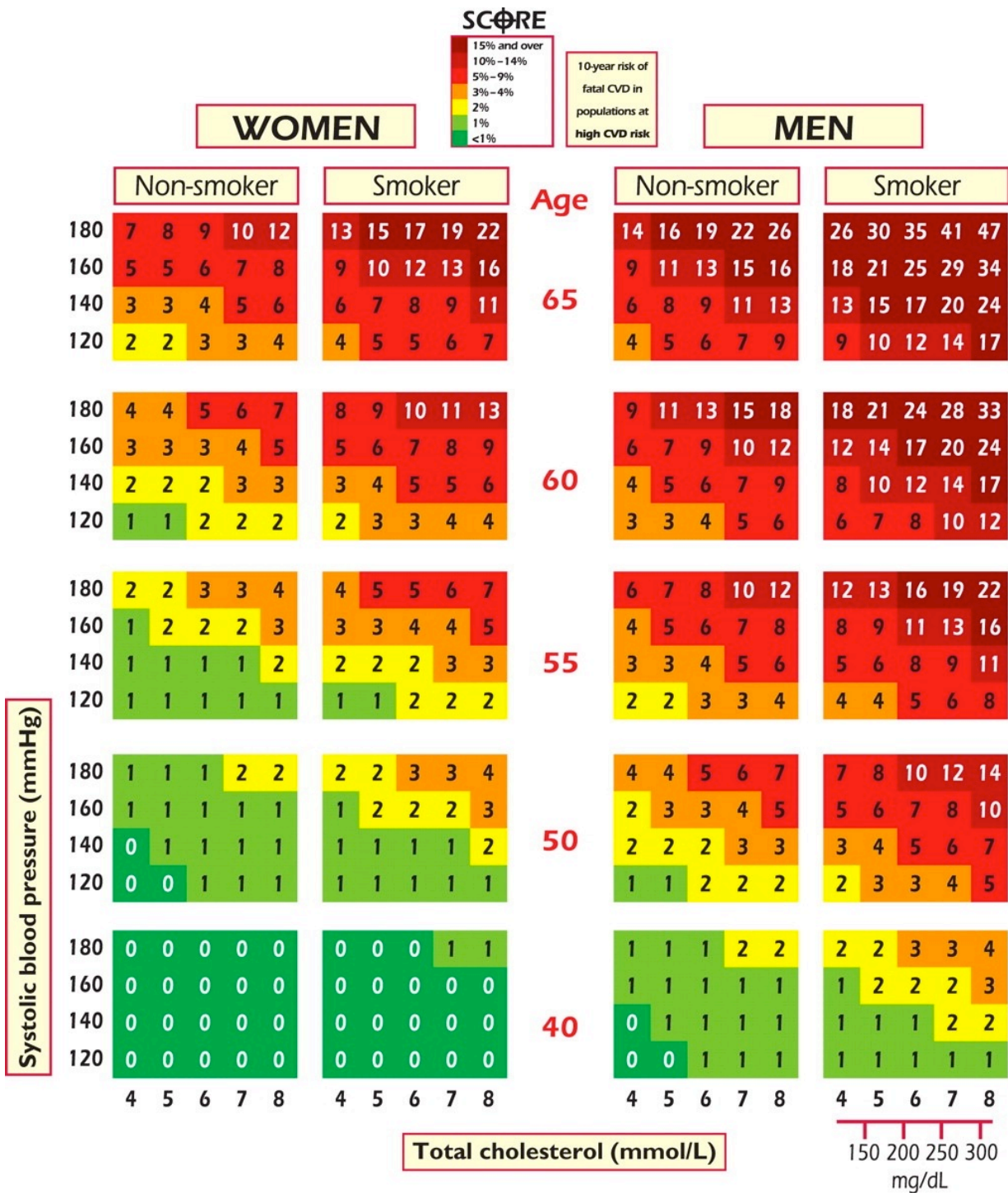
Al igual que el índice de Framingham, indica la probabilidad de contraer una ECV en 10 años desde el momento en que se calcula. Cuenta también con una matriz que asigna un nivel de riesgo asociado a cada probabilidad como muestra la figura 2.3 .

### **2.3.3 Índice *Prospective Cardiovascular Münster (PROCAM)***

El índice *PROCAM* [32] se basa en un estudio realizado en Alemania que comenzó en 1979 y que contó con alrededor de 20.000 participantes. Considera los siguientes factores de riesgo:

- Edad
- Presencia de Diabetes
- Colesterol HDL
- Lipoproteína de baja densidad (Colesterol LDL)
- Triglicéridos
- Consumo de tabaco
- Infartos en el historial de la familia

Este índice entrega cierto puntaje según el riesgo cardiovascular predicho y cada puntaje puede ser traducido en una probabilidad de riesgo de ECV a 10 años. Tal como está implementado está pensado para ser utilizado en hombres, sin embargo, se puede utilizar en mujeres dividiendo el puntaje en cuatro.



© 2007 ESC

High CVD risk countries are all those not listed under the low risk chart (Figure 4). Of these, some are at very high risk, and the high-risk chart may underestimate risk in these. These countries are Armenia, Azerbaijan, Belarus, Bulgaria, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Macedonia FYR, Moldova, Russia, Ukraine, and Uzbekistan.

Key: CVD = cardiovascular disease; SCORE = Systematic COronary Risk Evaluation

Reproduced with kind permission from Perk<sup>6</sup>

Figura 2.3: Matriz para la predicción de riesgo de ECV según índice SCORE [31]

## 2.3.4 Matriz ESH/ESC

La sociedad europea de hipertensión (*European Society of Hypertension (ESH)* [33]) en conjunto con ESC crearon una guía práctica [34] para tratar la hipertensión, que incluye una matriz de predicción de ECV. Dicha matriz no calcula probabilidad pero sí entrega un nivel de riesgo como muestra la figura 2.4. Existen 5 niveles de riesgo:

1. *Average risk*
2. *Low added risk*
3. *Moderate added risk*
4. *High added risk*
5. *Very high added risk*

Other risk factors, asymptomatic organ damage or disease	Blood pressure (mmHg)			
	High normal SBP 130–139 or DBP 85–89	Grade 1 HT SBP 140–159 or DBP 90–99	Grade 2 HT SBP 160–179 or DBP 100–109	Grade 3 HT SBP ≥180 or DBP ≥110
No other RF		Low risk	Moderate risk	High risk
1–2 RF	Low risk	Moderate risk	Moderate to high risk	High risk
≥3 RF	Low to moderate risk	Moderate to high risk	High risk	High risk
OD, CKD stage 3 or diabetes	Moderate to high risk	High risk	High risk	High to very high risk
Symptomatic CVD, CKD stage ≥ 4 or diabetes with OD/RFs	Very high risk	Very high risk	Very high risk	Very high risk

Figura 2.4: Matriz desarrollada en conjunto por ESH y ESC para la predicción de riesgo de ECV [34]

Según el número de factores de riesgo presentes en el paciente establece un nivel de riesgo. La figura 2.5 describe los factores de riesgo que utiliza.

## 2.4 Otros conceptos médicos relevantes

### 2.4.1 Eventos Cardiovasculares

En esta memoria se consideran eventos cardiovasculares los siguientes:

- **Angina:** Dolor o molestia en el pecho por falta de irrigación sanguínea
- **Angioplastia:** Procedimiento médico que se realiza con el fin de dilatar una arteria

Risk factors	Subclinical organ damage
<ul style="list-style-type: none"> <li>▪ Systolic and diastolic BP levels</li> <li>▪ Levels of pulse pressure (in the elderly)</li> <li>▪ Age (M &gt; 55 years; W &gt; 65 years)</li> <li>▪ Smoking</li> <li>▪ Dyslipidaemia               <ul style="list-style-type: none"> <li>- TC &gt; 5.0 mmol/l (190 mg/dl)</li> <li>or:</li> <li>- LDL/C &gt; 3.0 mmol/l (115 mg/dl)</li> <li>or:</li> <li>- HDL-C: M &lt; 1.0 mmol/l (40 mg/dl) W &lt; 1.2 mmol/l (46 mg/dl)</li> <li>or:</li> <li>- TG &gt; 1.7 mmol/l (150 mg/dl)</li> </ul> </li> <li>▪ Fasting plasma glucose 5.6–6.9 mmol/l (102–125 mg/dl)</li> <li>▪ Abnormal glucose tolerance test</li> <li>▪ Abdominal obesity (waist circumference &gt; 102 cm (M), &gt; 88 cm (W))</li> <li>▪ Family history of premature CV disease (M at age &lt; 55 years; W at age &lt; 65 years)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Electrocardiographic LVH (Sokolow-Lyon &gt; 38 mm; Cornell &gt; 2440 mm/ms)</li> <li>or:</li> <li>▪ Echocardiographic LVH♦ (LVMI M ≥ 125 g/m<sup>2</sup>, W ≥ 110 g/m<sup>2</sup>)</li> <li>▪ Carotid wall thickening (IMT &gt; 0.9 mm) or plaque</li> <li>▪ Carotid-femoral pulse wave velocity &gt; 12 m/sec</li> <li>▪ Ankle/Brachial BP index &lt; 0.9</li> <li>▪ Slight increase in plasma creatinine: M: 115–133 μmol/l (1.3–1.5 mg/dl) W: 107–124 μmol/l (1.2–1.4 mg/dl)</li> <li>▪ Low estimated glomerular filtration rate♦♦ (&lt; 60 ml/min/1.73 m<sup>2</sup>) or creatinine clearance♦♦♦ (&lt; 60 ml/min)</li> <li>▪ Microalbuminuria 30–300 mg/24 h or albumin-creatinine ratio: ≥ 22 (M); or ≥ 31 (W) mg/g creatinine</li> </ul>
Diabetes mellitus	Established CV or renal disease
<ul style="list-style-type: none"> <li>▪ Fasting plasma glucose ≥ 7.0 mmol/l (126 mg/dl) on repeated measurement</li> <li>or:</li> <li>▪ Postload plasma glucose &gt; 11.0 mmol/l (198 mg/dl)</li> </ul>	<ul style="list-style-type: none"> <li>▪ Cerebrovascular disease: ischaemic stroke; cerebral haemorrhage; transient ischaemic attack</li> <li>▪ Heart disease: myocardial infarction; angina; coronary revascularization; heart failure</li> <li>▪ Renal disease: diabetic nephropathy; renal impairment (serum creatinine M &gt; 133; W &gt; 124 μmol/l); proteinuria (&gt; 300 mg/24 h)</li> <li>▪ Peripheral artery disease</li> <li>▪ Advanced retinopathy: haemorrhages or exudates, papilloedema</li> </ul>
<p>Note: the cluster of three out of 5 risk factors among abdominal, obesity, altered fasting plasma glucose, BP ≥ 130/85 mmHg, low HDL cholesterol and high TG (as defined above) indicates the presence of metabolic syndrome.</p>	

M = men; W = women; CV = cardiovascular disease; IMT = intima-media thickness; BP = blood pressure; TG = triglycerides; C = cholesterol; ♦ = Risk maximal for concentric (wall thickness/radius ratio  $\geq 0.42$ ) LVH (left ventricular hypertrophy); ♦♦ = MDRD formula; ♦♦♦ = Cockcroft Gault formula.

Figura 2.5: Factores de riesgo utilizados en la matriz ESH/ESC [34]



o vena estrecha para restaurar un flujo sanguíneo obstruido

- **Derivación aortocoronaria con injerto (CABG por sus siglas en inglés):** Procedimiento médico para tratar una enfermedad de la arteria coronaria consiste en conectar un injerto que permita flujo sanguíneo por una ruta alternativa.
- **Insuficiencia cardíaca congestiva (CHF por sus siglas en inglés):** Enfermedad cardiovascular en la cual el corazón no bombea sangre adecuadamente
- **Infarto de miocardio (MI por sus siglas en inglés):** Necrosis (muerte de células) en el corazón por falta de riego sanguíneo.
- **MIP (Myocardial Infarction Procedure o procedimiento ante Infarto de miocardio):** Procedimiento médico efectuado para combatir un infarto de miocardio.
- **Angioplastia transluminal coronaria percutánea:** Procedimiento médico mínimamente invasivo para reestablecer el flujo sanguíneo en arterias coronarias bloqueadas.
- **Procedimiento de revascularización:** Procedimiento para restablecer el flujo sanguíneo hacia cierto órgano específico.
- **Stent:** Colocación de un tubo diminuto en un vaso sanguíneo con el objetivo de evitar que este se cierre.
- **Accidente cardiovascular:** Cese del flujo sanguíneo hacia el cerebro

Para el entrenamiento y validación de los modelos de *Machine Learning* elaborados en esta memoria si ocurre cualquiera de estos eventos se considera que el paciente padeció una enfermedad cardiovascular. Cabe mencionar que estos eventos no son excluyentes. Una enfermedad cardiovascular puede estar asociada a más de uno de los elementos descritos.

## 2.4.2 Polisomnografía y oximetría

La polisomnografía es un examen médico utilizado principalmente para detectar trastornos de sueño. Consiste en la medición de una serie de funciones corporales durante el tiempo en que un paciente duerme o intenta dormir (normalmente en la noche). Se registran una serie de señales, dentro de las cuales se encuentran el pulso cardíaco y el nivel de saturación de oxígeno. Estas dos señales se utilizan en esta memoria como indicadores para construir un modelo de predicción cardiovascular.

La oximetría nocturna, por otra parte, es un estudio muy similar a la polisomnografía, aunque más acotado. La principal diferencia es que mide únicamente dos señales, que son el nivel de saturación de oxígeno y el pulso cardíaco. Para otras aplicaciones evidentemente es menos útil, pero para efectos de esta memoria cumple igual de bien su propósito que la polisomnografía, ya que el modelo de predicción cardiovascular que se construye sólo se

basa en las dos señales descritas.

### 2.4.3 *Disability adjusted life years (DALYs)*

Los DALYs [35] son una forma estandarizada de cuantificar el impacto de acciones específicas relacionadas con el estado de salud de una población. Se utilizan frecuentemente en evaluaciones sociales en donde parte del beneficio o costo que se percibe no puede ser cuantificado de forma precisa en unidades monetarias pues tienen relación con la mejora o deterioro en la salud de las personas. Los DALYs ponderan el impacto de una iniciativa en términos de años de vida sumados o perdidos, así como también en términos a la mejora o deterioro de la calidad de vida por efectos de una discapacidad. Un DALY equivale a la ganancia de un año de vida en un estado cien por ciento saludable

## 2.5 Modelos de *Machine Learning*

En esta memoria se utilizan diversos modelos de *Machine Learning*, que permiten desarrollar a través de aprendizaje supervisado un modelo de predicción. Dicho de otra forma, permiten en base a herramientas computacionales el ajuste de un modelo utilizando datos recolectados previamente.

En el caso de esta memoria se predice, en base a una serie de variables tomadas de una oximetría, la probabilidad de un paciente de sufrir una ECV en los 10 años siguiente.

### 2.5.1 Regresión logística

La regresión logística [36] es un modelo de regresión en donde la variable dependiente es categórica. Es decir, puede tomar sólo 2 valores, que representan 2 estados posibles (cero o uno). En esta memoria estos estados son si hubo o no ECV en los 10 años siguientes a la realización de una oximetría.

Se estiman una serie de coeficientes de forma que se ajusten de la mejor forma posible a los datos. Una vez que los parámetros son estimados, el modelo permite predecir, en base al valor que toma cada variable explicativa, la probabilidad de que la variable dependiente tome cierto valor.

El modelo toma la siguiente forma, en donde  $x_1, x_2, \dots, x_n$  son las diferentes variables explicativas,  $\beta_0, \beta_1, \dots, \beta_n$  son los parámetros estimados y  $F(x_1, x_2, \dots, x_n)$  es la probabilidad estimada de que la variable dependiente tenga el valor uno :

$$F(x_1, x_2, \dots, x_n) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \quad (2.3)$$



## 2.5.2 *Support Vector Machine (SVM)*

SVM [37] se refiere a un método de clasificación que permite clasificar puntos en un espacio multidimensional en diferentes clases. Por ejemplo, en base al sueldo, la edad y el nivel de deuda de un sujeto podría clasificarse en alguien de bajo riesgo o de alto riesgo de no pago de un crédito. Para hacer esta clasificación, se construye en base a un set de datos de entrenamiento un hiperplano que separa las clases posibles buscando minimizar el error de clasificación asociado en dicho set. Una vez hecho esto, se obtiene un modelo que permite hacer la clasificación para cualquier dato fuera del set de entrenamiento. Existen varios algoritmos para la construcción de los hiperplanos. En esta memoria se utiliza *C-Support Vector Classification (CBC)* con la función *Radial basis de kernel*. El *kernel* es una función que se utiliza en este tipo de algoritmos y condiciona la precisión y complejidad en la clasificación.

Existe además una forma de estimar en base al modelo, una probabilidad de pertenecer a cada una de las clases. Esta técnica será utilizada en esta memoria para crear un modelo de predicción que precisamente estime la probabilidad de riesgo de ECV.

## 2.5.3 *Red Neuronal*

Una red neuronal [38] es un sistema muy complejo compuesto por una gran cantidad de unidades neuronales interconectadas entre ellas, de forma análoga a como lo hace un cerebro humano. Cada unidad recibe un input y retorna un output que a su vez una o más unidades recibirán de input. Se dividen en muchas capas de unidades neuronales y la capa final representa el output del modelo. La función que determina estos *outputs* es parametrizada y sus parámetros se ajustan en base a un set de datos. Aunque para muchas aplicaciones permite hacer predicciones muy precisas, tiene la desventaja de que el modelo resultante no es fácilmente interpretable.

## 2.5.4 *Naive Bayes*

*Naive Bayes* [39] es un método de *Machine Learning* que se basa en la ecuación 2.4 que representa al teorema de Bayes:

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)} \quad (2.4)$$

Donde  $A$  y  $B$  son ciertas ocurrencias que tienen asociada una probabilidad. Por ejemplo,  $A$  podría ser un evento cardiovascular y  $B$  el hecho de que un individuo cualquiera sea hombre, tenga 45 años y sea diabético. En base a esta ecuación, se puede estimar la probabilidad de ECV dada una serie de características de un paciente. Para hacerlo, se debe asumir que estas características son independientes, es decir, que el valor de una no influye en el de la otra y viceversa. Este es un supuesto muy fuerte, sin embargo a pesar de que no se cumple en la mayoría de los casos, el modelo en muchas ocasiones da resultados muy

buenos en situaciones reales, incluso con pocos datos. Diversos algoritmos se utilizan para hacer la clasificación y en esta memoria se utiliza el de *Gaussian Naive Bayes*.

## 2.6 Conceptos relevantes de *Machine Learning* y desarrollo de software

### 2.6.1 Matriz de confusión y métricas asociadas

Una forma estándar de medir el desempeño de un modelo de clasificación es a través de la matriz de confusión [40]. Se construye en base a las predicciones que éste realiza para un set de datos particular que contiene las variables explicativas del modelo así como también los resultados reales observados. Para cada observación de este set se compara los resultados reales con la predicción del modelo. En un modelo de clasificación de 2 clases donde una clase se considera positiva y la otra negativa, la matriz tiene 4 casillas, que representan lo siguiente:

- *True Positives*: El número de ocasiones en que el modelo predijo positivamente de forma acertada.
- *True Negatives*: El número de ocasiones en que el modelo predijo negativamente de forma acertada
- *False Positives*: El número de veces en que el modelo predijo positivamente de forma errónea.
- *False Negatives*: El número de veces en que el modelo predijo negativamente de forma errónea.

La figura 2.6 ilustra con un ejemplo la construcción de una matriz de confusión. Las casillas exteriores representan las sumas por columna o por fila de la matriz.

Sobre la base de esta matriz, pueden construirse una serie de métricas. En esta memoria se utilizarán las siguientes:

- **Accuracy**: Porcentaje de veces que el modelo predice acertadamente. Se calcula como se indica en la ecuación 2.6.1

$$(True\ Positives + True\ Negatives)/N \quad (2.5)$$

, donde  $N$  es el total de observaciones contabilizadas en la matriz.

- **Precision**: Porcentaje de veces en que se predijo acertadamente del total de veces que se predijo positivamente. Se calcula como indica la ecuación 2.6.1:

$$True\ Positives/(True\ Positives + False\ Positive) \quad (2.6)$$

		<b>Predicted: NO</b>	<b>Predicted: YES</b>	
n=165				
<b>Actual: NO</b>		TN = 50	FP = 10	60
<b>Actual: YES</b>		FN = 5	TP = 100	105
		55	110	

Figura 2.6: Matriz de confusión: mide desempeño de un modelo de clasificación

- **Negative Predictive Value:** Porcentaje de veces en que se predijo acertadamente del total de veces que se predijo negativamente. Se calcula como indica la ecuación 2.6.1:

$$\text{True Negatives} / (\text{False Negatives} + \text{True Negatives}) \quad (2.7)$$

- **Recall:** Porcentaje de veces en que se predijo acertadamente del total de veces en que el resultado real fue positivo. Se calcula como indica la ecuación 2.6.1:

$$\text{True Positives} / (\text{True Positives} + \text{False Negatives}) \quad (2.8)$$

- **Specificity:** Porcentaje de veces en que se predijo acertadamente del total de veces en que el resultado real fue negativo. Se calcula como indica la ecuación 2.6.1:

$$\text{True Negatives} / (\text{False Positives} + \text{True Negatives}) \quad (2.9)$$

- **False Positive Rate:** Porcentaje de veces en que se predijo erróneamente del total de veces que el resultado real fue negativo. Se calcula como indica la ecuación 2.6.1:

$$\text{False Positives} / (\text{False Positives} + \text{True Negatives}) \quad (2.10)$$

Este tipo de métricas se utilizan para medir el desempeño de un modelo de clasificación en dos clases, sin embargo, aunque pueden ayudar a realizar un primer análisis exploratorio, no sirven por sí solas para medir el desempeño de un modelo cuyo output es una probabilidad. En el caso de esta memoria, se busca precisamente construir modelos que arrojen una probabilidad de ECV y por lo tanto es necesaria una forma alternativa de medir

el desempeño. Una forma de hacerlo es a través de la curva *ROC* (*Receiver Operating Characteristic*). El *ROC* es un gráfico bidimensional en donde para un umbral de probabilidad dado se calculan las métricas *Recall* (también llamado *True Positive Rate*) y *False Positive Rate* y posteriormente se dibujan como un punto en un gráfico. Así, en base a muchos puntos, se dibuja una curva en el gráfico, tal como muestra la figura 2.7.

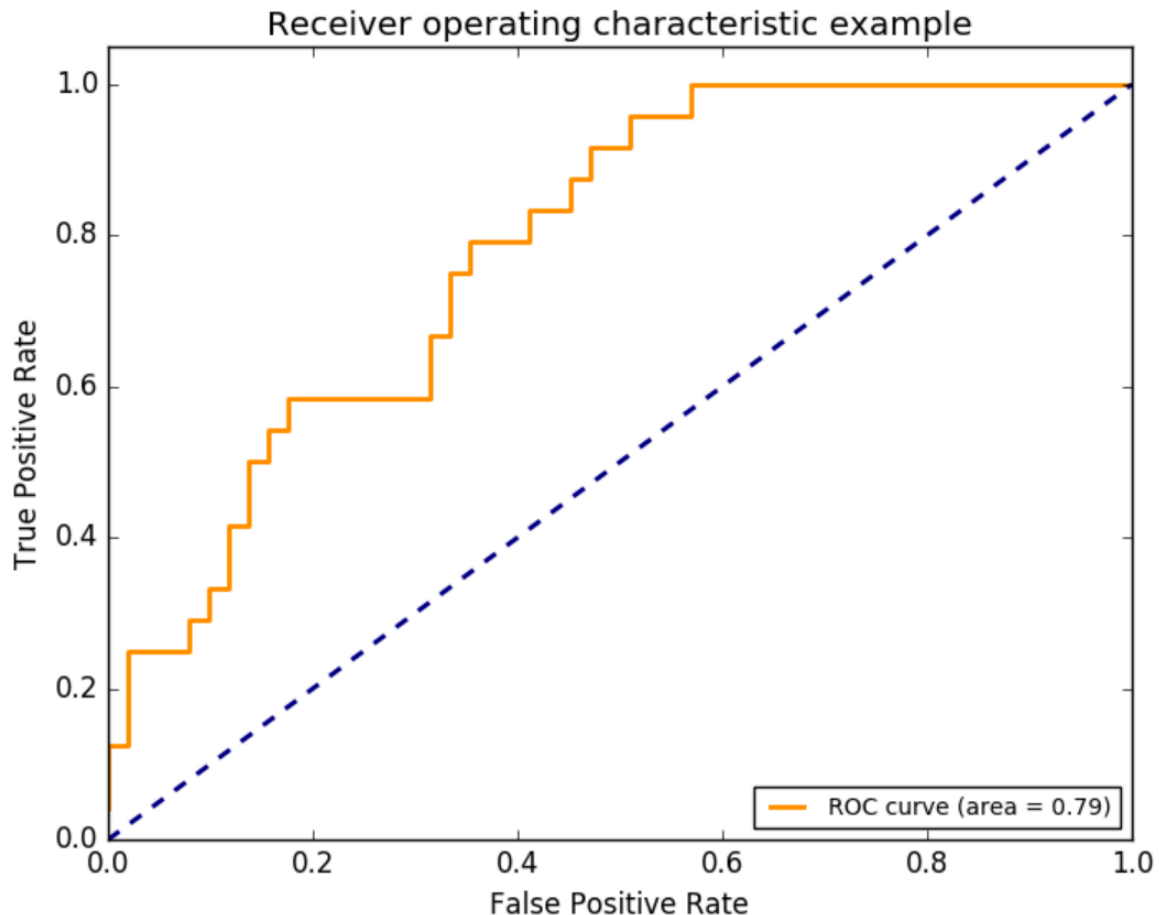


Figura 2.7: Curva para medir el desempeño de modelos probabilísticos. La línea diagonal representa un modelo sin ningún poder de predicción y entre más se acerca la curva hacia la esquina superior izquierda mayor es la precisión en la predicción del modelo

Los puntos del gráfico con un alto *Recall* y un bajo *False Positive Rate* representan buen desempeño del modelo, por lo tanto, entre más corrida esté la curva hacia la esquina superior izquierda del gráfico, mayor es el desempeño del modelo. Para cuantificar esto, se utiliza una medida de desempeño llamada *AUC* (*Area under curve*), que mide el área bajo la curva *ROC* en el gráfico.

### 2.6.2 *Cross-validation* o validación cruzada [41]

El término *cross-validation* se refiere a una técnica de validación de modelos de *Machine Learning* que busca predecir como será el desempeño de un modelo en la práctica, evitando que este se sobreestime debido a un sobreajuste en los datos. Para eso se divide

los de datos en  $k$  porciones todas del mismo tamaño, llamados *fold*s. Luego de eso, se entrena un modelo con  $k - 1$  *fold*s y se mide su desempeño en el *fold* restante. Se repite este proceso  $k$  veces dejando cada vez un *fold* distinto fuera de la fase de entrenamiento. Se utiliza en esta memoria para revisar el desempeño de los distintos modelos implementados y poder compararlos posteriormente.

### 2.6.3 Modelo Vista Controlador (MVC)

El concepto modelo-vista-controlador o *model-view-controller* [42] se refiere a una patrón de arquitectura de software en donde una aplicación se separa en tres componentes: modelo, vista y controlador.

El modelo es la parte del *software* que, como dice su nombre, se encarga de modelar el sistema, la vista se encarga de representar el modelo visualmente a través de alguna interfaz y el controlador a su vez se encarga de manipular la vista y el modelo.

Una de las ventajas de hacer esta separación es crear código que pueda extenderse fácilmente incluso por varios programadores al mismo tiempo. Se utiliza en esta memoria para facilitar la investigación una vez terminado el trabajo del alumno en cuanto al desarrollo de código que se requiera.

# Capítulo 3

## Desarrollo del modelo de predicción

### 3.1 Selección de indicadores

Tras una búsqueda bibliográfica se seleccionaron dos artículos que son utilizados como base para la investigación. Ambos artículos describen indicadores implementados a partir de una oximetría y que muestran indicios de que podrían ser útiles en la construcción de un modelo de predicción de ECV.

1. D. Sommermeyer, D. Zou, et al [11] encuentran una correlación entre varios índices calculados en base a una oximetría nocturna e índices clásicos de predicción de riesgo de ECV. Los índices son los siguientes:
  - (a) *Pulse wave attenuation index(PWA-I)*
  - (b) *Mean pulse propagation time(PPT)*
  - (c) *Mean respiratory-related pulse oscillation(RRPO)*
  - (d) *Pulse rate acceleration index(PR-I)*
  - (e) *Hypoxia Index(SPO<sub>2</sub>-I)*
  - (f) *Periodic and symmetric desaturation(PASD)*
  - (g) *Total time during which SpO<sub>2</sub> was below 90 % (TimeBelow90)*
  - (h) *Difference between PR-I and SpO<sub>2</sub>-I (PR\_SpO<sub>2</sub>-I)*

Cada uno de estos parámetros se correlaciona independientemente con el índice de riesgo cardiovascular ESH/ESC en una regresión logística. Además, en el artículo se crea un modelo que predice riesgo cardiovascular en base a los 7 primeros índices combinados (el último no se utiliza pues no tiene sentido incluir en un modelo multivariable una variable que es la resta de otras dos variables), que es validado con

cada uno de los siguientes índices tradicionales de predicción de riesgo:

- (a) Índice *ESH/ESC*
- (b) Índice de Framingham
- (c) Índice *PROCAM* [43]
- (d) Índice *SCORE* [44]

La principal limitación de este estudio es que la validación está hecha en base a otros indicadores de riesgo de ECV, lo cual no permite asegurar buenos resultados pues no se toma en cuenta las limitaciones de los indicadores tradicionales. Además, no entrega garantías de que índices derivados de la oximetría, en conjunto con variables tradicionales de predicción de riesgo cardiovascular, den mejores resultados que los modelos tradicionales.

2. E. Watanabe, K. Kiyono et al. [21] encuentran una correlación estadísticamente significativa entre la proporción del tiempo en que ocurrieron eventos de desaturación al 4 % (4 %POD) y la probabilidad de muerte en pacientes con Apnea Central y problemas cardiovasculares

Cabe mencionar que en este estudio la correlación fue estudiada para pacientes con problemas cardiovasculares y que además presentaban Apnea Central, para esta memoria, se buscará predictores de riesgo cardiovascular en un universo de pacientes más amplio. Por otra parte, la correlación que se encuentra es entre el índice 4 %POD y la mortalidad de los pacientes. En esta memoria se busca predecir el riesgo cardiovascular, es decir, la probabilidad de padecer un ECV, y no la probabilidad de muerte.

Los indicadores encontrados se revisan en conjunto con el doctor Juan Carrillo, médico de la Unidad del Sueño del INT, con el objetivo de seleccionar los que se implementarán en el marco de esta memoria.

En primer lugar, se descarta la implementación de los índices PWA-I y PPT debido a que se basan en señales que muchos de los oxímetros de pulso y polisomnógrafos no recogen. Por lo tanto, un modelo que las recoja en muchos casos no podría ser utilizado en la práctica. Adicionalmente, la base de datos SHHS en particular no cuenta con estas mediciones y por lo tanto el entrenamiento de un modelo que contemple dichos indicadores no es factible con la metodología diseñada para esta memoria.

Por otra parte, RRPO y PASD son dos índices que podrían tener poder de predicción y pueden ser calculados en base a cualquier oxímetro de pulso. Sin embargo, su descripción en los artículos académicos en donde se presenta es muy vaga y por lo tanto una implementación sin un estudio de mayor profundidad podría derivar en imprecisiones que afecten negativamente los resultados. Se decidió entonces buscar una colaboración de parte de investigadores que ya han implementado el índice para lograr resultados más precisos. Sin embargo, se resuelve dejar este aspecto fuera del alcance de esta memoria pues hay una alta incertidumbre en cuanto a los tiempos que el proceso pueda tardar.

Se decide entonces implementar los índices *PR-I*, *SPO<sub>2</sub>-I* y *TimeBelow90* tal como se presentan en D. Sommermeyer, D. Zou, et al. [11] y probar además ciertas variantes de los últimos dos. A continuación se presenta una breve descripción de cada uno de ellos, que se ilustra además en la tabla 3.1:

1. *PR-I*: Indica el número de aceleraciones de pulso mayores a un 10 % por hora.
2. *SPO<sub>2</sub>-I*: Indica el número de desaturaciones de oxígeno por hora, es decir, el número de caídas abruptas en la curva de SPO<sub>2</sub>. En estudios del sueño, se utilizan para el diagnóstico de problemas de sueño diferentes niveles de intensidad para definir las desaturaciones. En E. Watanabe, K. Kiyono et al. [21] se muestran indicios de poder de predicción de ECV si se considera como desaturación una caída al 4 % o más. En D. Sommermeyer, D. Zou, et al. [11] a su vez, se evidencian buenos resultados predictivos considerando como desaturación caídas al 2 % o más. Se decidió por lo tanto implementar tres variantes en total de este índice, con caídas al 2 %, 3 % y 4 %, seleccionando posteriormente la de mejores resultados.
3. *TimeBelow90*: Tiempo total de la oximetría durante el cual el porcentaje de saturación de oxígeno es menor a 90 %. Se implementan además dos variantes que son análogas pero se calculan con respecto a porcentajes diferentes. Específicamente, a un 85 % y 95 %. Se seleccionará finalmente la que entregue el mejor desempeño.

<b>PR-I</b>	Número de aceleraciones de pulso mayores a un 10 % por hora
<b>SPO<sub>2</sub>-I</b>	Número de desaturaciones de oxígeno por hora
<b>ProportionBelow90</b>	Porcentaje del tiempo durante el cual el nivel de saturación de oxígeno es menor a 90 %

Tabla 3.1: Descripción de variables calculadas desde una oximetría consideradas en esta memoria

## 3.2 Implementación de un modelo inicial base

### 3.2.1 Implementación del índice inicial

El primer índice que se decidió implementar fue *TimeBelow90*, que corresponde al total del tiempo durante el cual la saturación de oxígeno (SpO<sub>2</sub>) es de menos de 90 %. Cabe mencionar que se hizo un cambio menor en su implementación con respecto a la que plantea el artículo D. Sommermeyer, D. Zou, et al. [11]. En vez de calcular el tiempo total transcurrido se calculó la proporción del tiempo de medición en la que SpO<sub>2</sub> fue menor de 90 %. De esta forma, se evita asignar un valor más alto al índice sólo a causa a una duración mayor de la oximetría.

Se probó además una segunda implementación del índice en donde se calcula el porcentaje del tiempo de sueño con SpO<sub>2</sub> menor a 90 % (la diferencia entre tiempo de medición y tiempo de sueño existe pues no necesariamente el paciente está durmiendo durante toda



la medición). Finalmente se decidió no incluirla en el modelo pues podría haber dificultades prácticas en una futura implantación en producción. Esto pues requiere información adicional que no es necesariamente fácil de obtener. Sin embargo, su implementación no fue eliminada por si es útil en investigaciones venideras.

La implementación se hizo en base a un archivo EDF escogido entre los que incluye la base de datos de *NSRR*. Para la lectura desde *Python* se utilizó la librería `pyedflib` [45]. El archivo EDF es representado como un conjunto de arreglos en *Python* donde cada arreglo representa cada una de las señales de una polisomnografía. Para esta investigación la señal relevante es la que indica el nivel de saturación de oxígeno en la sangre. Por lo tanto, se almacena en una variable el arreglo correspondiente. Este arreglo contiene cada medición realizada de SpO<sub>2</sub> que representa el porcentaje de saturación de oxígeno en un instante dado. Se hace una medición por segundo durante el tiempo que dura la oximetría.

Para el cálculo de los diferentes índices se crea una clase llamada `IndexCalculator` que es la encargada de calcular los distintos índices. *IndexCalculator* tiene 2 variables de instancia `oximetry` y `sleep_time`. La primera representa la oximetría para la que se quiere calcular el indicador y la segunda el tiempo total de medición durante el cual el paciente estuvo dormido. Se crea entonces un primer método llamado `proportion_below_90`, que recibe un parámetro que puede tomar dos valores según el tipo de implementación deseado. En cada caso, se calcula el índice dividiendo el número de datos bajo 90 % en el número total de datos o en el tiempo durante el cual el paciente estuvo dormido. En caso de que se ingrese un argumento que no corresponda a alguna de las implementaciones realizadas en esta función se arroja una excepción.

### 3.2.2 Implementación de la variable dependiente

Por otra parte, se crea un archivo `xlsx` llamado `Eventos Cardiovasculares.xlsx` a partir de la base de datos SHHS, que contiene en su primera columna el *id* de cada paciente que sufrió un ECV y en una siguiente columna el tiempo que transcurrió desde que fue realizada la oximetría hasta que ocurrió el primer evento cardiovascular (un paciente puede sufrir más de un evento cardiovascular). La base de datos contiene mucha más información sobre cada evento, pero se hace un filtro para dejar sólo los datos más relevantes.

Para construir la variable en base al archivo, se crea la clase `DependentVariableBuilder`. Esta clase tiene un constructor que recibe un archivo como parámetro (que debe corresponder al archivo `Eventos Cardiovasculares.xlsx`), lo lee y va construyendo un diccionario en donde se asocia cada paciente que tuvo algún evento cardiovascular con el tiempo que transcurrió desde que se tomó la oximetría hasta que esto ocurrió y que se guarda como variable de clase. Además, se crea una variable de clase que contiene los identificadores de todos los pacientes para los que se encuentra disponible una oximetría en la base de datos.

Por otra parte, la clase posee un método (`ECV_to_5_or_10_years`) que en base al diccionario crea la variable dependiente tal como será utilizada para el primer modelo de predicción, es decir, una variable que puede tomar los valores 0, 1 o 2, que representan lo siguiente:

- 0: No hay ECV en los 10 años siguientes
- 1: Hay ECV en los 5 años posteriores a la oximetría
- 2: Hay ECV en los 10 años posteriores a la oximetría (pero no en los 5 primeros años)

Cabe mencionar que en aquellos pacientes para los que no se cuenta con oximetría la variable guarda el valor `None` de *Python*, que representa una entidad que no tiene valor asignado.

### 3.2.3 Extracción de un conjunto grande de oximetrías

El siguiente paso consiste en extraer cada una de las oximetrías disponibles y calcular, para cada una de ellas, el indicador ya implementado. Para esto se crea una clase llamada `OximetryReader` que tiene los siguientes métodos:

- `load_test_data`: Este método lee un conjunto de oximetrías de prueba y las retorna en un arreglo. Las oximetrías que lee son las de los pacientes correspondientes a los identificadores 200011 a 200015, por lo tanto, requiere que dichas oximetrías se encuentren almacenadas en la carpeta del proyecto (`prediccion_ecv`). Este método es creado principalmente para revisar el buen funcionamiento del código, sin embargo, se crean más adelante en esta memoria *tests* unitarios para cumplir este objetivo de forma más eficiente y que por lo tanto le restan utilidad práctica al método. Pese a lo anterior, el código no se elimina pues el memorista considera que puede ayudar a investigadores futuros a entender mejor el proceso de lectura de oximetrías, dada la simplicidad de su implementación.
- `load_data`: La base de datos SHHS cuenta con oximetrías tomadas en dos periodos, uno entre 1995 y 1998 y otro entre 2001 y 2003. En el primer período se tomaron oximetrías a todos los pacientes del estudio y en el segundo período se tomaron oximetrías adicionales a algunos de los pacientes. Este método está pensado para la lectura de oximetrías desde una carpeta que contenga todas las oximetrías de la base de datos SHHS tomadas en el primer período. Al igual que el método anterior, este método lee las oximetrías en formato EDF y las retorna todas en un arreglo. Cada elemento del arreglo corresponde a una oximetría. Se entrega como parámetro al método el número de oximetrías que se desea leer, y se leen las oximetrías desde la del paciente número 200001 (identificador del primer paciente) en orden numérico. Se puede indicar también la ruta en donde se encuentran los archivos EDF y además, opcionalmente se puede indicar a través de otro parámetro si se desea omitir alguna lectura por si un investigador en el futuro deseara excluir ciertos datos. Por defecto existe una serie de oximetrías que se omiten debido a que por razones que el memorista desconoce no se encuentran disponibles en la base de datos SHHS. Dichos pacientes no se consideraron en el estudio como parte de los datos. Por lo demás, cuando una oximetría se omite, en el arreglo retornado se asigna el valor `None` en vez de la señal de la oximetría.
- `read_oximetry`: Este método lee una única oximetría y se implementa para facilitar

el *testing* de la clase que se describe más adelante en esta memoria.

### 3.2.4 Regresión Logística o Logit

Con lo anterior, se pueden obtener todos los datos necesarios para el entrenamiento y validación de un primer modelo. Se trata de una regresión logística que tiene como única variable explicativa el indicador escogido (*ProportionBelow90*). El modelo predice si se presentará o no una ECV en los siguientes 10 años. Para eso, calcula una probabilidad de ECV y decide en base a qué es más probable que ocurra. En otras palabras, si la probabilidad de ocurrencia de ECV es mayor a 0,5, decide que habrá una ECV y si es menor, entonces decide que no habrá.

Se crea entonces la clase `ModelBuilder` que pretende construir un modelo Logit y calcular su desempeño, de modo de poder compararlo con otros modelos que se construyen más adelante en esta memoria. La clase tiene un constructor que recibe tres variables:

1. La variable dependiente como una lista de *Python*. Esta debe estar construida tal como se indica en la sección 3.2.2, es decir, debe tener valores 0, 1 o 2 según el tiempo hasta la primera ECV.
2. Las variables explicativas como una lista de *items*, en donde cada *item* es a su vez una lista con todos los valores de cada una de las variables de un paciente.
3. Una lista de los nombres de las variables explicativas

El constructor transforma las dos primeras listas en un arreglo de la librería `numpy` [46], ya que así es como son ingresados más adelante los datos en la librería que construye el modelo Logit. Se crean 2 métodos de la clase, ambos sirven para medir el desempeño del modelo pero lo hacen de formas distintas. A continuación se describen ambos:

- `cross_validation_scores`: Este método utiliza la técnica de *cross-validation* ( 2.6.2) o validación cruzada para revisar el desempeño del modelo. Recibe dos parámetros que corresponden al número de *folds* que se desea utilizar para la validación y el número de años a los que se desea hacer la predicción. Lo primero que se hace es construir una variable dependiente dicotómica en base a la variable dependiente que se ingresó en el constructor del objeto. Según a cuantos años se desee hacer la predicción se decide cuando tiene que valer uno y cuando debe tomar el valor cero. Entonces, por ejemplo, si la predicción se desea hacer a 10 años, se crea una variable que vale uno en caso de que haya habido una ECV en los 10 años siguientes a la oximetría y cero en caso contrario. Luego de eso, se utiliza la librería `Sklearn` para ajustar los parámetros de una regresión logística en base a la variable dependiente recién calculadas y a las variables explicativas de la clase. Luego de ello se imprimen los siguientes resultados del modelo:
  - Matriz de confusión
  - *Accuracy*

- *Precision*
- *Recall*
- `split_data_validation`: Este método difiere del anterior en la forma en que entrena y valida los datos. En este caso se entrena con una proporción de los datos y se valida el modelo con la proporción restante. Recibe como *input* el número de años a los que se desea hacer la predicción y la proporción del total de datos que se dejará para *testing*. A diferencia del método anterior, esta técnica de validación permite obtener de forma directa un modelo de predicción. En base a dicho modelo se imprimen, además de las métricas del método anterior, el intersección y los coeficientes ajustados del modelo y las probabilidades que arroja en base a las que hace la predicción.

### 3.2.5 Controlador

Se creó una clase que hará de controlador. Su objetivo es encargarse del manejo de las demás clases, según el modelo MVC (2.6.3). Se crean los siguientes métodos:

- `get_oximetries`: Este método utiliza la clase `OximetryReader` para leer las oximetrías. Si bien podría utilizarse directamente esta clase para la lectura es una buena práctica que lo haga el controlador.
- `calculate_indexes_from_oximetries_to_csv`: Este método, como dice su nombre, se encarga de leer todas las oximetrías, calcular los índices y guardarlos en un archivo *csv*. La idea de hacer esto es que al hacer cualquier cambio en el modelo no sea necesario leer cada vez todas las oximetrías (a menos que se quiera utilizar un nuevo índice). Recibe los siguientes argumentos:
  - `oximetries`: oximetrías a leer.
  - `indicator_names`: nombres de los indicadores que se desea calcular. A medida que se implementan nuevos indicadores el método debe ser extendido para que permita su cálculo.
  - `csv_file_name`: nombre del archivo donde se guardarán los índices calculados.

Con esta información, el método calcula los índices y los guarda en el archivo *csv*.

- `get_independent_variables_from_csv`: Este método se encarga de leer los índices de un archivo *csv* de modo de que estos sean almacenados en un arreglo y posteriormente puedan ser utilizados para la construcción de modelos de *Machine Learning*. Como argumentos recibe el número de oximetrías de las cuales se desea obtener el índice y el nombre del archivo *csv* desde el que se hará la lectura (por defecto es `indicadores.csv`).
- `build_dependent_variable`: Este método se encarga de construir la variable dependiente en base al archivo `Eventos cardiovasculares.xlsx`. Hace uso de la clase

DependentVariableBuilder, pero filtra la variable de modo que no haya valores None.

### 3.2.6 *Testing*

Se hace un proceso de *testing* básico con el objetivo de asegurar en cierta medida que lo implementado esté correcto y de tener una herramienta para revisar además que a lo largo de la memoria, con las modificaciones que el código va sufriendo, todo siga funcionando según lo esperado. Además, esto es de mucha utilidad para investigadores futuros que sigan trabajando sobre el código.

Se crea entonces una primera clase para *testing* llamada `IndexCalculatorTests`, que contiene *tests* unitarios para verificar el buen funcionamiento de los métodos de la clase `IndexCalculator`. Para esto extenderá la clase `TestCase`, de la librería `unittest` [47], con lo cual se habilita la realización de *tests* unitarios de forma simple. Esta clase posee dos variables globales que representan una oximetría de solo 10 mediciones y el tiempo de sueño de pacientes ficticios a las cuales se le asignan valores de forma arbitraria. La razón por la que se asigna una oximetría tan corta es que el investigador futuro pueda verificar fácilmente que el cálculo del indicador es el correcto. Una vez creada la clase se crea un método llamado `test` en donde se incluirán los *tests* unitarios. Los *tests* revisan que efectivamente el cálculo del índice sea correcto en base al resultado esperado en la oximetría de prueba. También revisa que la función arroje una excepción cuando corresponde.

Se creó además la clase `ControllerTest` que, tal como su nombre lo indica, realiza *testing* sobre los métodos de la clase `Controller` para verificar que estén funcionando correctamente. Revisa los siguientes puntos:

- Que una oximetría pueda leerse correctamente, incluso si se omite cierto archivo
- Que los arreglos se lean correctamente independientemente de si fueron leídos directamente desde la oximetría o desde el archivo CSV
- Que la regresión logística funcione sin arrojar errores.
- Que la variable dependiente sea construida correctamente

### 3.2.7 **Resultados**

Se calcula la matriz de confusión de este modelo, tanto con la técnica de validación cruzada como con la técnica de separación de los datos en dos (datos de entrenamiento y datos de *testing*). Para la última técnica se reservó un 30 % de los datos para *testing* y el resto para entrenamiento. Los resultados se despliegan en las tablas 3.2 y 3.3

Como puede observarse, en los datos de entrenamiento el modelo predice siempre que no habrá ECV. Esto se traduce en un muy mal rendimiento pues alrededor de un 23 % de las predicciones son casos en donde en realidad sí hubo ECV. Cabe mencionar que era espera-

		¿Modelo predijo EVC?	
		NO	SI
¿Hubo EVC?	NO	4520	0
	SI	1273	0

Tabla 3.2: Matriz de confusión del modelo base calculada con la técnica de validación cruzada

		¿Modelo predijo EVC?	
		NO	SI
¿Hubo EVC?	NO	1340	0
	SI	398	0

Tabla 3.3: Matriz de confusión del modelo base calculada con la técnica de separación de los datos utilizando un 30 % para *testing*

ble que este modelo no tuviera un buen desempeño principalmente pues utiliza una única variable explicativa. Y es que el objetivo de su construcción fue únicamente desarrollar una base para la implementación de modelos más complejos.

### 3.3 Desarrollo del modelo final

Teniendo un primer modelo base, el siguiente objetivo es volverlo más robusto. Para ello, en esta etapa se implementan dos índices adicionales derivados de la oximetría, se incorporan variables explicativas clásicas y se utilizan nuevos modelos de *Machine Learning*. Teniendo en perspectiva que el objetivo es lograr una predicción más precisa que la de un índice tradicional, se implementa también el índice de Framingham para tener un punto de comparación. Se decide utilizar el índice de Framingham y no algún otro índice tradicional, pues es el único que ha sido construido en base a datos de Estados Unidos y por lo tanto es el más razonable a ser comparado con modelos construidos con datos de Estados Unidos.

#### 3.3.1 Implementación y desempeño del modelo de Framingham

El primer paso consiste entonces en implementar el índice de Framingham. Para ello se crea la clase `FraminghamCalculator`. Tiene un único método (`get_10_years_risk`) que recibe las variables explicativas del índice de Framingham y en base a ello calcula una probabilidad de riesgo de ECV a 10 años. Cabe mencionar que en caso de que alguna de las variables explicativas sea `None`, el método arroja un error.

En base a la clase descrita, se puede calcular el índice de Framingham para cada uno de los pacientes de la base de datos SHHS. El controlador es quien se encarga de hacerlo. Para ello, se implementa el método `calculate_framingham_indexes`, que lee para cada paciente en SHHS las variables requeridas y calcula el índice. Finalmente retorna un arreglo

con el índice de Framingham calculado para cada paciente. Cabe mencionar que hay 323 pacientes para los que no está disponible alguna información en la base de datos que es necesaria para el cálculo del índice (por ejemplo, puede no haber información del índice de masa corporal de alguno de los pacientes). Para estos casos, el arreglo se completa con el valor `None`. Por otra parte, hay ciertas variables que pueden tomar sólo dos valores y que están codificadas con números enteros en la base de datos. La variable correspondiente al género, por ejemplo, está codificada con un 1 en caso de que el paciente sea hombre y un 2 en caso contrario. En estos casos se hizo una transformación, de modo que la variable final sea un booleano y así se tengan valores más legibles. En el caso anterior, por ejemplo, la variable género pasa a llamarse *male* y vale *True* si el paciente es hombre y *False* en caso contrario. Para el caso particular de la variable *smokstat\_s1* se hizo un filtro especial. Dicha variable puede tener tres valores posibles según si el paciente es fumador al momento de tomarse la oximetría, si lo fue en algún momento pero dejó de serlo o si nunca lo fue. Ya que el índice de Framingham se basa sólo en el estado del paciente en un instante dado en cuanto a su consumo de tabaco, la variable original fue transformada en una variable que simplemente toma el valor *True* o *False* según si el paciente es fumador o no al momento de tomarse la oximetría.

Finalmente, se debe medir el desempeño del índice. Para eso, se revisa que tan precisa es la predicción de acuerdo a la base de datos SHHS. Esto no es directo pues el output del índice es una probabilidad de ECV y la base de datos tiene información de si ocurrió o no una ECV. Con el objetivo de tener una primera medida de desempeño se define un modelo auxiliar, que en base a un umbral y según la probabilidad de riesgo de ECV arrojada por el índice de Framingham define si se arroja o no una alerta de riesgo cardiovascular (en la sección 3.3.7 se elabora un mecanismo más robusto de medición de desempeño). En la figura 3.1 se puede apreciar un diagrama que ilustra el funcionamiento del modelo. Este modelo sí es directamente comparable con los datos reales y por lo tanto puede medirse directamente su desempeño a través de una matriz de confusión.

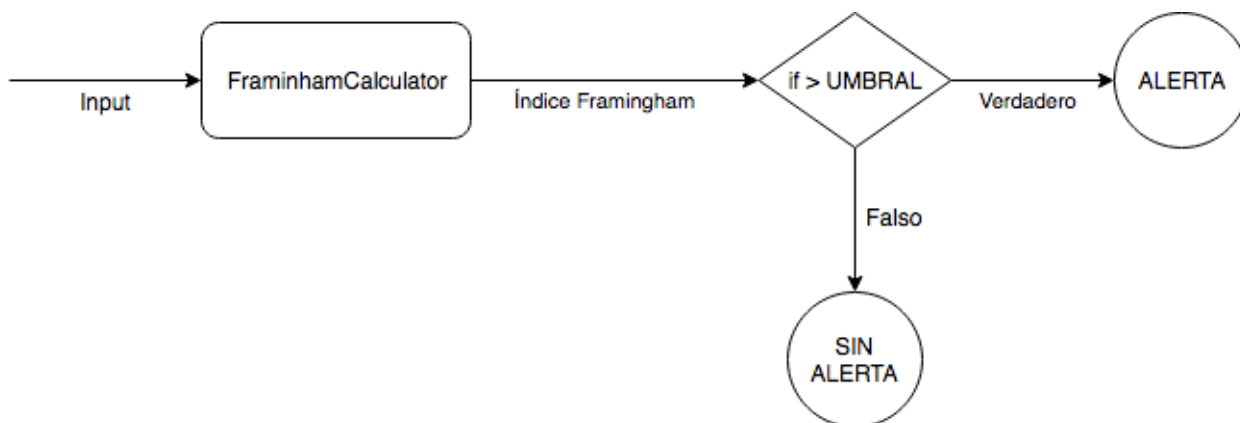


Figura 3.1: Modelo auxiliar a Framingham utilizado para cuantificar su desempeño

Para la implementación del modelo auxiliar se crea un método auxiliar privado en la clase `Controller` (`_get_framingham_indexes_score`) que recibe los siguientes argumentos:

- `true_values`: Desenlace real en la historia clínica del paciente. En otras palabras, es un arreglo de ceros y unos, en donde un uno significa que hubo una ECV en los 10

años posteriores a la oximetría y un cero que esto último no ocurrió.

- `framingham_index_array`: El arreglo con el cálculo del índice de Framingham para cada paciente.
- `threshold`: Umbral en base al cual se define el modelo.

El método calcula los resultados del modelo auxiliar para cada paciente y en base a ello imprime la matriz de confusión asociada. Retorna además dicha matriz en forma de un diccionario, en donde cada elemento corresponde a una casilla de la matriz.

Finalmente, para tener más de una medida de desempeño, se crea un método llamado `get_all_framingham_indexes_score` que llama al método anterior con distintos umbrales. Para ello, recibe como argumento un arreglo con los diferentes umbrales y calcula, usando el método anterior, la matriz de confusión para cada uno de los umbrales.

### 3.3.2 Implementación de nuevas variables explicativas

En primer lugar, para mejorar el modelo base creado en la sección 3.2, se implementaron dos nuevas variables explicativas desde la oximetría. Se trata de las dos variables seleccionadas en la sección 3.1 que habían quedado pendientes. A continuación se describe su implementación:

- *SPO2-I*: Este indicador pretende cuantificar qué tan seguido se presentan caídas abruptas o desaturaciones en el nivel de saturación de oxígeno. Para eso, se define primero una desaturación de oxígeno o como una baja de dos o más puntos porcentuales en una ventana de tiempo de 90 segundos. Luego, el promedio de caídas por hora corresponde al indicador.

Para su implementación, se crea el método `SP02_i`, en la clase `IndexCalculator`. En primer lugar, se crea un *buffer* que tiene 91 mediciones recogidas de la oximetría, partiendo por los primeros 91 datos (la oximetría recoge una medición cada un segundo). Luego, este se recorre de izquierda a derecha y se revisa si existe algún elemento en el *buffer*, cuyo valor sea dos puntos porcentuales menos que el primer elemento en el *buffer*. De ser así, se contabiliza la desaturación encontrada y se actualiza el *buffer*, de modo que su primer elemento pase a ser el dato en donde se encontró la desaturación. En caso de no encontrarse ninguna desaturación, también se actualiza el *buffer* corriéndose una posición a la derecha. La figura 3.2 ilustra este proceso. Esto se repite hasta haber revisado todos los datos de la oximetría. Finalmente, se retorna el número de desaturaciones encontradas dividido en el número de horas que dura la oximetría.

- *PR-I*: Este índice se basa en el ritmo cardíaco de un paciente y pretende cuantificar qué tan seguido se presentan incrementos abruptos durante el sueño. Se define un incremento abrupto como un aumento en al menos 10 puntos porcentuales con respecto a una mediana móvil calculada en base a un ventana de 20 elementos. Finalmente se retorna el promedio de incrementos abruptos por hora.



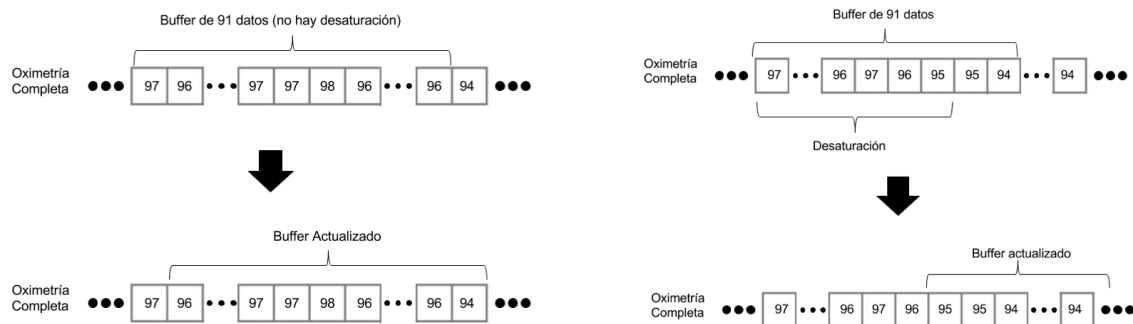


Figura 3.2: Ilustración de una actualización en el *buffer* durante el algoritmo para el cálculo del indicador *SPO2-I*

El indicador se encuentra implementado en el método `pr_i` y se basa en un *buffer* de 21 elementos, que en un comienzo corresponde a las primeras 21 mediciones de la oximetría. Lo primero que se hace es calcular la mediana del *buffer*. Luego, este se recorre de izquierda a derecha, revisando si alguno de los primeros 10 elementos es al menos un 10 % mayor a la mediana. De existir un incremento abrupto el *buffer* se actualiza de tal forma que su primer elemento corresponda a la medición donde se detectó el incremento. En caso de no existir un incremento se corre el *buffer* en una posición a la derecha. El objetivo de esta parte es detectar si al comienzo de la medición hubo una caída brusca en el ritmo cardíaco, lo cual indicaría que con alta probabilidad hubo un incremento abrupto justo antes de comenzar la medición que debe ser contabilizado.

Luego de esto, con el *buffer* actualizado, se recalcula la mediana y se revisa si el elemento en la posición central del *buffer* es o no un 10 % mayor que la mediana. De ser así, se corre el *buffer* en 10 posiciones hacia la derecha y en caso contrario se corre una posición. En cualquier caso se repite el proceso hasta que el *buffer* contenga los últimos 21 elementos de la señal de ritmo cardíaco. La figura 3.3 ilustra una actualización del *buffer*.

Finalmente, se hace un proceso análogo al de la primera parte para revisar si existe un incremento abrupto hacia el final de la revisión. Luego se retorna el número de incrementos abruptos promedio encontrados por hora.

### 3.3.3 Inclusión de variables tradicionales al modelo

Un siguiente paso para aumentar la robustez del modelo es incluir variables explicativas usadas tradicionalmente en índices de riesgo de ECV. Se decide utilizar las mismas varia-

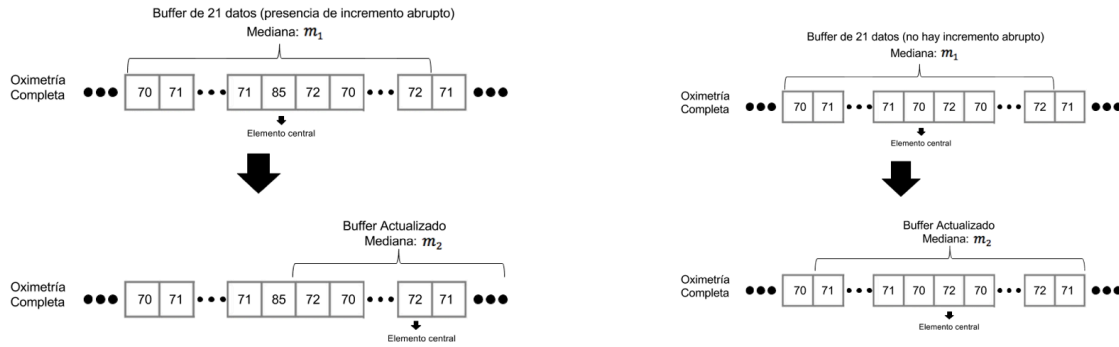


Figura 3.3: Ilustración de una actualización en el *buffer* durante el algoritmo para el cálculo del indicador PR-I

bles del índice de Framingham (2.3.1) por ser variables muy estudiadas de las que se ha validado su capacidad de predicción.

Para la lectura de dichas variables desde la base de datos SHHS se crea en el controlador el método `get_tradicional_independent_variables` que recibe como parámetro un archivo *csv* que corresponde a la base de datos de shhs (el nombre original del archivo es “shhs1-dataset-0.11.0.csv”). Desde el archivo lee cada una de las variables para cada paciente y las retorna en un arreglo en donde cada elemento corresponde a todas las variables para cierto paciente. A las variables originales se les aplican los mismos filtros que en el método `calculate_framingham_indexes`.

### 3.3.4 Modificación del umbral de clasificación

Una de las principales razones del mal desempeño del modelo inicial es que decide si existe o no riesgo de ECV en base a si la probabilidad de riesgo es o no mayor a un umbral de 0,5.

En primer lugar, esto presenta el inconveniente de que se está midiendo el desempeño de un tipo de modelo que no es el que requiere el problema. El objetivo final es que el modelo de predicción arroje una probabilidad de ECV y no una alerta basada en un umbral, pues una probabilidad es una medida mucho más precisa de riesgo que una clasificación en dos posibles clases (lo cual no quita que más adelante pueden definirse alertas en base a uno o más umbrales). Sin embargo, un modelo que arroja una probabilidad como output no permite una medición directa de su desempeño pues los datos observados son por naturaleza binarios, es decir, en la práctica se observa o no una ECV, pero no se observa una probabilidad asociada. Es por esta razón que, únicamente con fines de medir el desempeño del modelo vale la pena construir un modelo auxiliar que defina un umbral y clasifique en base a esto.

En segundo lugar, habiendo aceptado que es necesaria la construcción de un modelo auxiliar, un umbral de un 0,5 evidentemente conduce a un muy mal desempeño del modelo. Primero, porque los dato están desbalanceados, lo que hace que la probabilidad a priori

de acertar sea mucho mayor si se predice que no habrá ECV. Esto provoca que el modelo quede con cierta tendencia a predecir ausencia de ECV. Segundo, un umbral de 0,5 no considera que existen 2 tipos de errores que tienen un nivel de importancia diferente. Por un lado, predecir riesgo de ECV de forma errónea implica que se invertirán innecesariamente recursos en prevención de una enfermedad cuando el riesgo de contraerla es en realidad muy bajo. Por otro lado, predecir un riesgo bajo de ECV cuando en realidad el riesgo es considerable puede significar en el peor de los casos la muerte de un paciente o un decaimiento significativo en su estado de salud, habiendo podido tomar medidas de prevención. Este segundo caso es más grave que el primero y por lo tanto vale la pena bajar el umbral, de modo de poner el foco en evitar el segundo tipo de error.

Por todo lo anterior, se construye un modelo que clasifica de forma similar al original pero en base a a cualquier umbral que se defina. La implementación de este modelo se encuentra en la clase `ModelBuilder`, en el método `modified_model_scores`. El método recibe como parámetros la cantidad de años a los cuales se desea hacer la predicción, un umbral de clasificación y la proporción de los datos que desea utilizarse para *testing* (por defecto 0,30). Primero, se construye la variable dependiente en base al número de años de la predicción, luego se crea la matriz de confusión asociada al umbral de clasificación y finalmente se imprimen las siguientes medidas de desempeño:

- *Precision*
- *Negative Predicted Value*
- *Recall*
- *Specificity*

### 3.3.5 Exploración de modelos de *Machine Learning*

El siguiente paso consiste en entrenar diferentes modelos de *Machine Learning*. Esto es necesario pues, los diversos modelos tienen desempeños diferentes dependiendo de las características de los datos. Entonces, adicional a la regresión logística, se decide implementar varios para posteriormente medir el desempeño de cada uno y seleccionar el mejor. Los modelos fueron escogidos por ser ampliamente utilizados en distintas aplicaciones de *Machine Learning* y son los siguientes: *Support Vector Machine*, Red Neuronal y *Naive Bayes*

Para la implementación de dichos modelos se utilizó, al igual que para la regresión logística, la librería *Sklearn* [48]. Se crea el método `_build_confusion_matrix`, que en base a un objeto de *Sklearn* que representa alguno de los modelos descritos, un umbral, y los datos de entrenamiento y *testing*, construye un modelo y retorna su matriz de confusión. Este es un método privado de la clase `ModelBuilder` que es ejecutado por el método público `modified_model_scores`. Para el cálculo e impresión de las medidas de desempeño se utiliza un segundo método privado llamado `_print_confusion_matrix_scores` que precisamente calcula e imprime las medidas de desempeño descritas. El modelo basado en *Support Vector Machine* utiliza específicamente el método de clasificación *C-Support*

*Vector Classification* (CVC) o tipo uno y con *Radial basis function* (RBF) como *kernel*. La red neuronal, por su parte, utiliza una única capa oculta de 100 neuronas. Por último, la clasificación en el modelo de *Naive Bayes* se realiza a través del algoritmo *Gaussian Naive Bayes*. Junto al equipo técnico del WIC se escogieron dichas configuraciones por ser ampliamente utilizadas y de buenos resultados en aplicaciones previas de diversa índole, por lo que constituyen una buena base para iniciar el proceso de investigación que continuará en etapas posteriores a la realización de esta memoria.

### 3.3.6 Testing

Se realizaron una serie de *tests* unitarios para verificar el buen funcionamiento del modelo que se describen a continuación

#### Índice de Framingham

Para cerciorarse de que el cálculo del índice de Framingham esté funcionando correctamente, se creó la clase `FraminghamCalculatorTest`. Esta clase calcula a través del método `get_10_years_risk` de la clase `FraminghamCalculator` el índice de Framingham para un paciente ficticio con ciertos atributos. Luego, se mide el índice de Framingham a través de una herramienta oficial provista en el sitio web de los creadores del índice <sup>1</sup>. Finalmente, se compara que el índice calculado con ambos métodos tenga el mismo valor. Cabe mencionar, que para la comparación se utiliza el método `assertAlmostEqual` y no `assertEqual` debido a que el primer método permite que exista una diferencia infinitesimal entre ambos resultados, y en este caso, aquello puede darse ya que se trata de un número no entero de muchos dígitos y el computador puede redondear algunas cifras durante el cálculo.

A su vez, en la clase `ControllerTest`, se revisa que la matriz de confusión asociada a un conjunto de observaciones del índice de Framingham se construya correctamente. Esto se hace en base a un arreglo de 10 valores que representan 10 observaciones del índice y sobre el cual se calcula la matriz utilizando el método `get_framingham_indexes_score` y comprobando que el resultado sea el esperado.

#### Indicadores de la oximetría

Por otra parte, se revisó que el cálculo de los indicadores derivados de la oximetría se realice según lo esperable. Para eso, se agregan dos *tests* a la clase `IndexCalculatorTests`. Primero, se calcula el índice `spo2_i` y se revisa que entregue el resultado esperado en base a una oximetría de prueba. Luego, se crea una señal de pulso cardíaco de prueba que posee tres desaturaciones muy evidentes y se revisa que el resultado de utilizar el método `pr_i` coincida también con el número esperado.

---

<sup>1</sup><https://www.framinghamheartstudy.org/risk-functions/cardiovascular-disease/10-year-risk.php>

### 3.3.7 Medición de desempeño y resultados

Finalmente se mide el desempeño del modelo y se compara con el del índice de Framingham. Esta medición representa un desafío ya que el output del modelo es una probabilidad y por lo tanto no es directa la medición de desempeño.

Como primera aproximación al problema, se decidió construir tres modelos auxiliares, en base a los umbrales 5 %, 10 % y 15 %. Y medir el desempeño en cada caso, de forma de tener una primera idea del comportamiento de cada modelo.

Se calcularon las métricas *precision*, *NPV*, *recall*, y *specificity*. La tablas 3.4, 3.5 y 3.6 muestran los resultados obtenidos.

<b>Umbral: 5 %</b>				
	<i>Precision</i>	<i>NPV</i>	<i>Recall</i>	<i>Specificity</i>
<b>Framingham</b>	0,254	0,973	0,986	0,147
<b>Logit</b>	0,237	0,984	0,994	0,095
<b>SVM</b>	0,220	0	1	0
<b>Red Neuronal</b>	0,248	0,948	0,994	0,171
<b>Naive Bayes</b>	0,305	0,950	0,925	0,405

Tabla 3.4: Medidas de desempeño con umbral de 5 %

<b>Umbral: 10 %</b>				
	<i>Precision</i>	<i>NPV</i>	<i>Recall</i>	<i>Specificity</i>
<b>Framingham</b>	0,298	0,938	0,919	0,361
<b>Logit</b>	0,284	0,966	0,958	0,315
<b>SVM</b>	0,220	0	1	0
<b>Red Neuronal</b>	0,252	0,935	0,950	0,201
<b>Naive Bayes</b>	0,355	0,926	0,839	0,569

Tabla 3.5: Medidas de desempeño con umbral de 10 %

<b>Umbral: 15 %</b>				
	<i>Precision</i>	<i>NPV</i>	<i>Recall</i>	<i>Specificity</i>
<b>Framingham</b>	0,335	0,904	0,810	0,526
<b>Logit</b>	0,330	0,945	0,900	0,484
<b>SVM</b>	0,221	0,800	0,967	0,038
<b>Red Neuronal</b>	0,291	0,910	0,859	0,408
<b>Naive Bayes</b>	0,393	0,898	0,720	0,685

Tabla 3.6: Medidas de desempeño con umbral de 15 %

En primer lugar, se observa que a medida que aumenta el umbral aumentan las medidas *precision* y *specificity* pero disminuyen las medidas *NPV* y *recall*. Esto era de esperarse pues a medida que aumenta el umbral se es más flexible en términos de no equivocarse con una predicción negativa.

Por otra parte, dependiendo del umbral y la métrica que se definan uno u otro modelo puede ser mejor. En general, los modelos tienen mejor *negative predicted value* y *recall* que el índice de Framingham pero peor *precision* y *specificity*. La excepción es *Naive Bayes*, en cuyo caso ocurre justo al revés.

Considerando lo anterior, y a pesar de que las métricas entregan información valiosa del comportamiento de los modelos, es muy difícil decidir cual es el mejor. Además, incluso si la decisión fuera clara, hay información relevante que no se está considerando ya que solo se están calculando métricas en base a tres umbrales. Se decide entonces construir una curva ROC para cada modelo, en donde para un umbral dado se dibuja en un gráfico un punto que representa las métricas *True Positive Rate* y *False Positive Rate* (ver 2.6.1). El resultado se muestra en la figura 3.4:

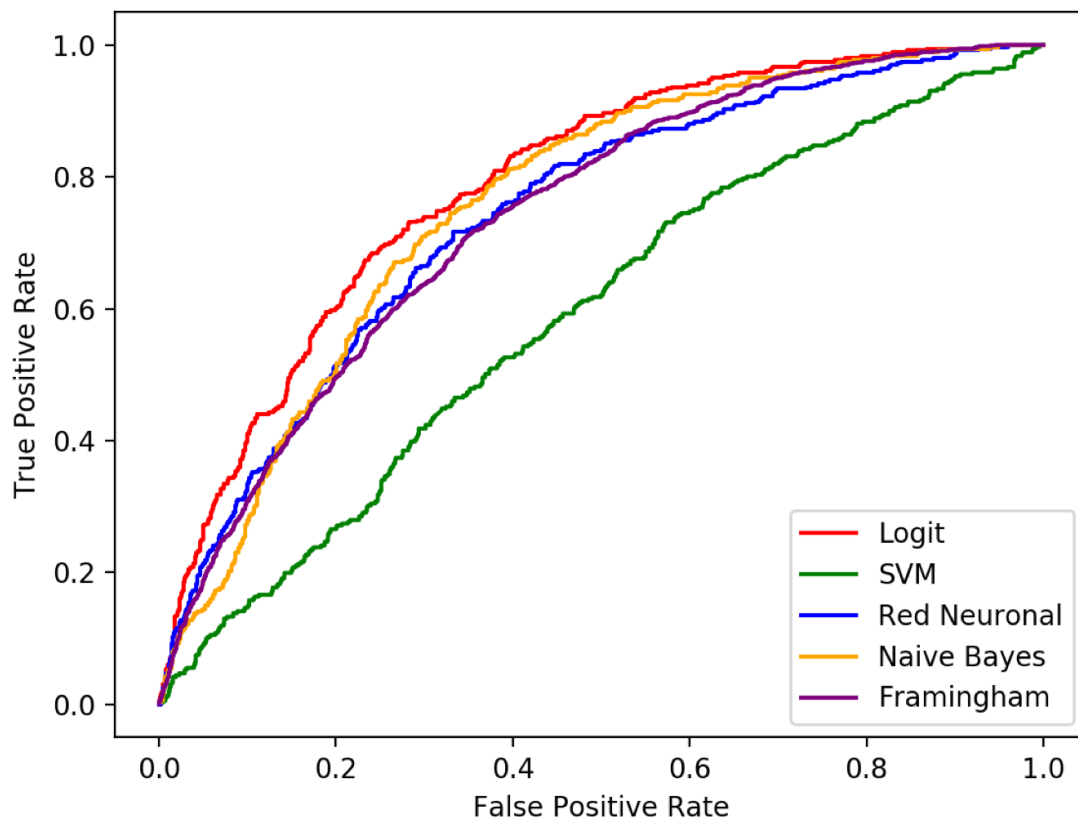


Figura 3.4: Curvas ROC para el índice de Framingham y cada uno de los modelos construidos

Los puntos más cercanos a la esquina superior izquierda del gráfico son los de mejor desempeño, ya que la medida *True Positive Rate* es mayor y la medida *False Positive Rate* es menor. Por lo tanto, el área bajo la curva en el gráfico es una excelente medida del desempeño de cada modelo. A esta área se le denomina AUC y para el índice de Framingham y los modelos construidos el resultado se muestra en la figura 3.7:

Como se aprecia en la tabla, el mejor modelo es el Logit (regresión logística), con un AUC

	<b>Area Under Curve (AUC)</b>
<b>Framingham</b>	0,739
<b>Logit</b>	0,790
<b>SVM</b>	0,589
<b>Red Neuronal</b>	0,741
<b>Naive Bayes</b>	0,758

Tabla 3.7: AUC para índice de Framingham y modelos construidos

de 0,790, que es mejor que el del índice de Framingham (0,739) de forma estadísticamente significativa. Con este resultado se da por validada la hipótesis de investigación.

## 3.4 Perfeccionamiento del modelo

Aunque el modelo en su esencia está completo, vale la pena afinar ciertos aspectos con el fin de mejorar la metodología de evaluación y avanzar en su robustez, de modo de lograr una mejor capacidad de predicción.

### 3.4.1 Mejoras en la metodología de evaluación

En primer lugar, se realizó una normalización a todas las variables del modelo, a excepción de las variables dicotómicas. El objetivo principal es asegurar que las diferencias de orden de magnitud entre variables no influya en el resultado del modelo. Para la normalización se utilizó el estadístico *z-score*, que se calcula como se indica en 3.1:

$$z_i = \frac{x_i - \bar{x}}{\sigma} \quad (3.1)$$

$x_i$  representa una observación particular,  $\bar{x}$  es el promedio de todas las observaciones y  $\sigma$  es la desviación estándar. En el caso particular de la edad, previo a la normalización, se calculó el logaritmo de la variable. La implementación se hizo en dos métodos, llamados `_normalize_traditional_independent_variables` y `_normalize_oximetry_variables`, que realizan el proceso de normalización de las variables tradicionales y de las de la oximetría respectivamente. Ambas reciben como *input* un arreglo de *items*, en donde cada *item* es a su vez un arreglo con cada una de las variables no normalizadas de un paciente particular. Se retorna en ambos casos la misma estructura, pero con las variables normalizadas.

Por otra parte, si bien ya había sido creado un método que permitía medir el desempeño de los datos en base a *cross-validation*, este sólo permitía hacerlo en base a métricas como *Accuracy* o *Recall*, que están pensadas para modelos de clasificación y no para modelos cuyo *output* es una probabilidad. Con este propósito es que se implementó el método `auc_scores_cross_validation`, que calcula el desempeño en términos del AUC para cada uno de los *folds*. Por defecto el método divide los datos en 20 *folds*, pero permite indicar un número distinto de *folds* a través de un parámetro opcional. Cabe mencionar que, previo a

la separación de *folds*, se cambia aleatoriamente el orden de los datos, de modo de que, si los pacientes en la base de datos original seguían cierto orden, se evite provocar un sesgo en el modelo. Para la separación de los datos en *folds*, se utiliza la clase `StratifiedKFold`, de la librería *Sklearn*, que permite hacerlo no sólo cambiando el orden de los datos, sino que además asegurando que se mantenga en cada *fold* una proporción de pacientes con eventos cardiovasculares similar a la del total de los datos. Con respecto a este punto, surge como interrogante el tiempo de ejecución que requiere la construcción de cada uno de los modelos, considerando que la aplicación de *cross-validation* ralentiza considerablemente el proceso. En este sentido, se aprecia que existen diferencias muy grandes entre los distintos modelos, siendo *Naive Bayes* el que menos tiempo toma y *SVM* el que más. Aún así, se considera que los resultados son satisfactorios ya que incluso los modelos que más demoran no toman demasiado tiempo. La tabla 3.8 muestra el tiempo total de ejecución en segundos para cada modelo utilizando *cross-validation* con 20 *folds*. El hecho de que los tiempos de ejecución sean bajos es una buena noticia pues permite que a futuro pueda complejizarse mucho más el modelo y puedan probarse distintas configuraciones sin que esto conlleve un costo demasiado alto en cuanto a recursos computacionales.

	<b>Tiempo de ejecución</b>
<b>Logit</b>	0.44
<b>SVM</b>	82.19
<b>Red Neuronal</b>	32.75
<b>Naive Bayes</b>	0.11

Tabla 3.8: Tiempo de ejecución en segundos de cada modelo con *cross-validation* de 20 *folds*

Adicionalmente, se corrieron los siguientes modelos, que no pretenden superar el desempeño del modelo ya construido, sino ser un punto de comparación frente al modelo construido:

- Red neuronal, *Support Vector Machine (SVM)*, clasificador bayesiano y regresión logística que consideren como variables explicativas únicamente los indicadores derivados de la oximetría.
- Red neuronal, *SVM*, clasificador bayesiano y regresión logística que consideren únicamente las variables explicativas clásicas

El primer tipo de modelo pretende revisar si la mejora de desempeño en relación al índice de Framingham de los modelos encontrados se debe al aporte de los indicadores de la oximetría por sí solas o a un ajuste particularmente bueno del modelo a la población estudiada.

El segundo tipo de modelo pretende establecer el poder de predicción que por si solos tienen los indicadores de la oximetría.



### 3.4.2 Mejoras del modelo

Una vez realizadas las implementaciones mencionadas, se prueba tres versiones distintas de dos de las variables implementadas desde la oximetría con el fin de poder comparar el desempeño de los modelos resultantes y escoger el mejor. Las variantes implementadas son las siguientes:

- **ProportionBelow85, ProportionBelow90 y ProportionBelow95:** Se refieren a la proporción del tiempo de medición en que el nivel de saturación de oxígeno estuvo bajo 85 %, 90 % y 95 % respectivamente. Si bien la variable *ProportionBelow90* es la que se probó originalmente y para la cual en [11] se encontraron indicios de capacidad predictiva, tanto *ProportionBelow85* como *ProportionBelow95* se utilizan habitualmente en estudios de sueño como indicadores de calidad de sueño y por lo tanto se decide revisar su potencial utilidad para el modelo.
- **SPO2-I al 2 %, 3 % y al 4 %:** Originalmente se implementó el índice SPO2-I de tal forma de contabilizar como desaturaciones cualquier caída en el nivel de saturación de oxígeno de al menos un 2 %. Dado que en [21] se muestran indicios de que el nivel de desaturaciones a un 4 % podría ser un buen indicador de riesgo cardiovascular y considerando que el nivel de desaturaciones al 3 % se usa habitualmente como un indicador de calidad de sueño, se decide revisar las tres variantes.

Por lo tanto, para cada uno de los modelos descritos en la parte 3.4.1, se prueban nueve alternativas diferentes correspondientes a las nueve combinaciones posibles de variantes.

### 3.4.3 Resultados

A continuación se exponen y analizan los resultados de cada uno de los modelos descritos en términos del AUC promedio calculado en base a *cross-validation* con 20 *folds*.

En primer lugar, se exhiben los resultados del modelo completo, incluyendo las variables de Framingham y las de la oximetría luego de las mejoras realizadas y con la metodología de evaluación actualizada. El desempeño en términos de AUC para cada configuración posible de las variables SPO2-I y *ProportionBelow* se muestra en las tablas 3.9, 3.10, 3.11 y 3.12:

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,773	0,774	0,775
	90 %	0,775	0,774	0,774
	95 %	0,773	0,773	0,774

Tabla 3.9: AUC regresión logística para cada configuración posible de variables explicativas

Se observa que la regresión logística sigue siendo la de mejor desempeño con un AUC de 0,775 seguida de cerca por la red neuronal. Sin embargo, el desempeño se reduce respecto al modelo previo, probablemente debido a que existía un sobreajuste que se redujo con las

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,713	0,685	0,713
	90 %	0,715	0,682	0,710
	95 %	0,685	0,717	0,678

Tabla 3.10: AUC *Support Vector Machine* para cada configuración posible de variables explicativas

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,768	0,769	0,769
	90 %	0,772	0,769	0,770
	95 %	0,767	0,771	0,772

Tabla 3.11: AUC Red Neuronal para cada configuración posible de variables explicativas

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,746	0,747	0,749
	90 %	0,744	0,743	0,745
	95 %	0,747	0,747	0,747

Tabla 3.12: AUC Naive Bayes para cada configuración posible de variables explicativas

mejoras realizadas a la metodología de evaluación de desempeño. A pesar de ello, sigue existiendo una diferencia significativa con respecto al índice de Framingham. La variación de los porcentajes en las variables de la oximetría no alteran mayormente el desempeño final, lo cual probablemente se debe a que dichas variables no tienen un peso demasiado alto en el modelo.

El modelo que únicamente contempla las variables de Framingham, por otra parte, arroja los resultados desplegados la tabla 3.13 :

	Logit	SVM	Red Neuronal	Naive Bayes
AUC	0,774	0,721	0,774	0,753

Tabla 3.13: AUC modelo con únicamente las variables de Framingham como variables explicativas

De aquí destaca principalmente el hecho de que los resultados son muy similares en cuanto a desempeño a los del modelo anterior, que incluye todas las variables. Las diferencias que se aprecian no son siquiera estadísticamente significativas. Esto implica que el peso de las variables de la oximetría es muy bajo en el modelo en relación al de otras variables. Este hallazgo es de suma importancia pues revela que el aporte de la oximetría al modelo es poco significativo y que la mejora del modelo versus el índice de Framingham se debe seguramente a otros factores. Aunque encontrar dichos factores queda fuera del alcance de esta memoria, se postula que lo más seguro es que se deba a que la población estudiada en la base de datos SHHS tiene ciertas características especiales que hacen que un

modelo entrenado y testeado en esos datos tenga un desempeño particularmente bueno. Una hipótesis es que esa mejora se debe a que la población estudiada presenta mayores problemas de sueño que lo normal, lo cual implicaría que valdría la pena desarrollar un método de predicción específico para ese tipo de pacientes.

Finalmente, en las tablas 3.14, 3.15, 3.16 y 3.17 se exhiben los resultados de los modelos cuyas variables explicativas son únicamente los indicadores derivados de la oximetría.

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,584	0,586	0,579
	90 %	0,584	0,585	0,570
	95 %	0,595	0,596	0,590

Tabla 3.14: AUC Regresión Logística para cada configuración posible de variables explicativas

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,496	0,472	0,480
	90 %	0,484	0,481	0,503
	95 %	0,528	0,534	0,509

Tabla 3.15: AUC *Support Vector Machine* para cada configuración posible de variables explicativas

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,605	0,607	0,593
	90 %	0,607	0,607	0,594
	95 %	0,609	0,604	0,602

Tabla 3.16: AUC Red Neuronal para cada configuración posible de variables explicativas

		SPO2-I		
		2 %	3 %	4 %
ProportionBelow	85 %	0,586	0,586	0,569
	90 %	0,581	0,583	0,571
	95 %	0,603	0,602	0,591

Tabla 3.17: AUC Naive Bayes para cada configuración posible de variables explicativas

Se comprueba que los indicadores de la oximetría tienen poder de predicción de riesgo cardiovascular como sugería la literatura. Un modelo sin ningún poder de predicción tiene un AUC de alrededor de 0,5 mientras que tres de los modelos en este caso tienen un AUC cercano a 0,6. La excepción es el *Support Vector Machine* que tiene un AUC marcadamente más bajo. Aunque se desconoce la razón se cree que puede deberse a un sobreajuste en los datos. Se deduce entonces que la baja relevancia encontrada de las variables de la oximetría en el modelo final no se deben a que estas no tengan poder de predicción, sino

que seguramente al alto nivel de correlación que existe entre estas con variables tradicionales de predicción de riesgo cardiovascular. Por otra parte, se aprecia claramente que la variable *ProportionBelow95* es la de mejor poder predictivo. *ProportionBelow90* como *ProportionBelow85* en tanto tienen un poder de predicción menor y similar entre sí. En cuanto a la variable SPO2-I, el rendimiento más bajo se da si se considera como desaturaciones caídas a un 4 %. Con 2 % o 3 % el desempeño es similar. Por último, el modelo de mejor desempeño en general, de entre los que únicamente utilizan las variables derivadas de la oximetría, es la red neuronal. Y dentro de las redes neuronales, el mejor es el que utiliza las variables *ProportionBelow95* y SPO2-I al 2 %.

El modelo que se decide utilizar como referencia para la creación del módulo funcional y la evaluación económica y social es la regresión logística que incluye variables tradicionales y también variables derivadas de la oximetría, con el índice SPO2-I al 2 % y con la variable *ProportionBelow90*. Esto ya que esta configuración de variables para este tipo de modelo es la que mejor desempeño tiene en términos de AUC, con un valor de 0,775. Los coeficientes del modelo se despliegan en la tabla 3.18.

<b>Intercepto</b>	-2,189
<b>ProportionBelow90</b>	0,053
<b>SpO2-I (2 %)</b>	-0,034
<b>PR-I</b>	0,009
<b>Sexo</b>	0,571
<b>Log(edad)</b>	1,039
<b>Presión sistólica</b>	-0,026
<b>IMC</b>	0,189
<b>Tratamiento Presión</b>	0,505
<b>Fumador</b>	0,799
<b>Diabético</b>	0,612

Tabla 3.18: Coeficientes modelo final

Se aprecia con claridad que el peso de las variables derivadas de la oximetría es efectivamente mucho menor al de otras variables. Las variables que más peso tienen en el modelo son el logaritmo de la edad, seguida por si el paciente es o no fumador.

# Capítulo 4

## Módulo de visualización

Una vez construido el modelo de predicción, se crea un prototipo funcional de un módulo de visualización del índice. Consiste en una aplicación web en donde un usuario ingresa una oximetría en formato EDF, junto a otros indicadores de riesgo cardiovascular y en base a aquello se calcula la probabilidad de contraer una enfermedad cardiovascular. El principal objetivo del módulo es mostrar de forma muy simple y clara el resultado final de la memoria, entendiendo que es un trabajo técnicamente complejo y por lo tanto es difícil de comunicar a quien no está familiarizado con técnicas de *Data Science*. Con esto se espera que sea más fácil motivar a futuros investigadores a seguir trabajando en el tema.

### 4.1 Descripción desde el punto de vista del usuario

El prototipo consiste en una aplicación web con una primera vista en donde se despliega un formulario con una serie de campos que permiten al usuario ingresar las variables del modelo. Los campos son los siguientes:

- **Oximetría:** Campo para ingresar un archivo en formato EDF. El archivo debe corresponder a una polisomnografía (ver 2.4.2). A pesar de que sólo se utilizan las señales correspondientes al nivel de saturación de oxígeno y al pulso cardíaco, se decidió utilizar una polisomnografía en vez de una oximetría principalmente ya que el primer tipo de examen es más común en el Instituto Nacional del Tórax que el segundo. De todas formas esto puede adaptarse fácilmente en el futuro para que acepte ambos tipos de exámenes.
- **Sexo:** Campo en donde el usuario selecciona una opción entre dos alternativas, indicando si el sexo del paciente es masculino o femenino.
- **Edad:** Campo en donde el usuario indica la edad del paciente.
- **Presión sistólica:** Campo en donde el usuario ingresa la presión sistólica del paciente en milímetros de mercurio (mmHg).

- **Índice de masa corporal:** Campo en donde el usuario ingresa el índice de masa corporal del paciente.
- **Tratamiento presión sanguínea:** Se despliega la pregunta: “¿Se encuentra actualmente con algún tratamiento para controlar la presión sanguínea?”. El usuario debe seleccionar entre dos alternativas, indicando si la respuesta es positiva o negativa en el paciente.
- **Consumo de tabaco:** Se despliega la pregunta: “¿Es fumador?”. El usuario debe seleccionar entre dos alternativas, indicando si la respuesta es positiva o negativa en el paciente. Se considera fumador aquel que durante el último mes ha fumado al menos un cigarrillo diario (definición entregada por la OMS).
- **Presencia de diabetes:** Se despliega la pregunta: “¿Es diabético?”. El usuario debe seleccionar entre dos alternativas, indicando si la respuesta es positiva o negativa en el paciente.

La figura 4.1 ilustra la vista descrita.

Al final del formulario, se muestra un botón que dice “Calcular Índice” y que precisamente permite al usuario calcular el riesgo de ECV en base al modelo desarrollado en esta memoria. Al presionar el botón, la aplicación calcula en base al archivo EDF las variables derivadas de la oximetría que el modelo utiliza y luego, utilizando además el resto de las variables ingresadas, calcula la probabilidad de ECV. Finalmente, la aplicación despliega una segunda vista que muestra el porcentaje calculado junto a una representación gráfica del mismo, ilustrando de forma simple el nivel de riesgo del paciente. La figura 4.2 muestra un ejemplo de la vista descrita.

## 4.2 Implementación

Para la implementación del módulo se crea un proyecto de Django (*framework* de desarrollo web) llamado `modulo_funcional` y una aplicación llamada `modulo_funcional_app`. Se crea además una carpeta llamada `templates` que contiene archivos `html` asociados a las vistas del proyecto. La documentación sobre la estructura y funcionamiento de un proyecto en Django se encuentra disponible en [27].

Para la creación del módulo, en primer lugar, debe diseñarse un sistema que maneje la lógica interna del módulo, correspondiente al cálculo del riesgo cardiovascular propiamente tal. Se crea entonces en primer lugar una clase llamada `IndexCalculator`, que se encarga de calcular los indicadores de la oximetría. Recibe como input dos señales: pulso cardíaco y de nivel de saturación de oxígeno. En base a ello calcula los indicadores. Su implementación es idéntica a la de la clase del mismo nombre utilizada en la construcción del modelo.

Luego de eso, se crea otra clase llamada `RiskCalculator` que se encarga de calcular la probabilidad de ECV en base al nivel de saturación de oxígeno, el pulso cardíaco y las va-

The image shows a web browser window with the title "Datos paciente" and the URL "127.0.0.1:8000/modulo\_funcional\_app/". The page content is a form titled "Índice de riesgo cardiovascular". The form includes the following fields and options:

- Oximetría:** A file upload button labeled "Browse..." with the text "No file selected."
- Sexo:** Radio buttons for "Masculino" and "Femenino".
- Edad:** A text input field with a clear button.
- Presión sistólica:** A text input field with a clear button.
- Índice de masa corporal:** A text input field with a clear button.
- ¿Se encuentra actualmente con algún tratamiento para controlar la presión sanguínea?:** Radio buttons for "Si" and "No".
- ¿Es fumador?:** Radio buttons for "Si" and "No".
- ¿Es diabético?:** Radio buttons for "Si" and "No".
- Calcular Índice:** A blue button at the bottom of the form.

Figura 4.1: Formulario prototipo funcional para el ingreso de variables de predicción de riesgo cardiovascular

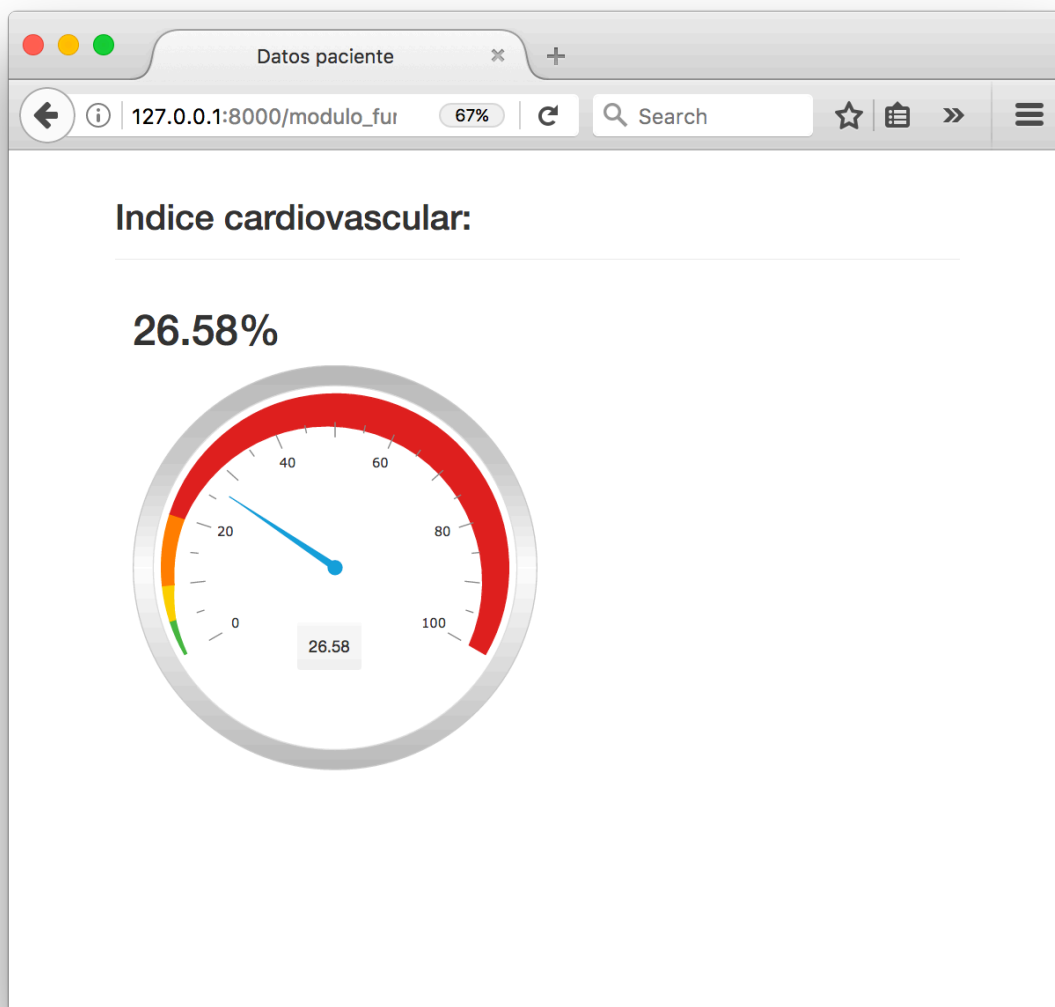


Figura 4.2: Representación gráfica del nivel de riesgo cardiovascular



riables tradicionales que se utilizan para cálculo del índice de Framingham. Tiene un único método llamado `calculate_risk` que utiliza la fórmula de la regresión logística con los parámetros estimados durante el entrenamiento del modelo. Como *input* recibe las señales de pulso y nivel de saturación de oxígeno y un diccionario que representa las variables del índice de Framingham. El diccionario debe tener las siguientes llaves:

- **male**: Indicando el sexo del paciente. `True` si es hombre y `false` si es mujer
- **age**: Indicando la edad del paciente.
- **systolic\_blood\_pressure**: Indicando la presión sistólica del paciente
- **body\_mass\_index**: Indicando el índice de masa corporal del paciente
- **blood\_pressure\_med\_treatment**: Indicando si el paciente está bajo tratamiento para la regularización de la presión sanguínea. `True` en caso afirmativo y `False` en caso contrario
- **smoker**: Indicando si el paciente es fumador. `True` en caso afirmativo y `False` en caso contrario.
- **diabetic**: Indicando si el paciente sufre de diabetes. `True` en caso afirmativo y `False` en caso contrario.

Por otra parte, debe crearse la vista en donde el usuario ingresa los datos del paciente. Para eso se extiende la clase `Form`, que provee *Django* y que se utiliza precisamente para el diseño de formularios. En esta clase se configura la estructura que tendrá el formulario en cuanto a los textos que acompañan cada uno de los elementos del mismo y en cuanto al tipo de información que el usuario podrá ingresar y la forma en que lo podrá hacer. Por ejemplo, ante la pregunta “¿Es diabético?”, el usuario puede seleccionar las alternativas “Si” o “No”, mientras que para señalar la edad del paciente el usuario puede ingresar cualquier número entero dentro de cierto rango. Existen además las siguientes restricciones para el ingreso de datos:

- **Rango Edad**: La edad debe ser un número entero entre 40 y 89 (Edades de los pacientes estudiados en la base de datos SHHS)
- **Rango Presión Sistólica**: La presión sistólica debe ser un número entero entre 40 y 260. Se considera un rango muy amplio. El objetivo principal de esta restricción es evitar errores de tipeo de parte del usuario.
- **Rango Índice de masa corporal**: El IMC debe ser un número entre 10 y 80. No necesariamente entero. También se considera un rango muy amplio.
- **Restricción de completitud**: Todos los datos deben ser completados para poder realizar el envío del formulario.

Finalmente, se crea un módulo llamado *Views* en donde se diseñan las 2 vistas del modelo. Tiene 2 métodos públicos:

1. `index`: Se encarga simplemente de desplegar el formulario para que lo rellene el usuario. Para eso, utiliza un *template* llamado `patientform.html` que despliega el formulario creado con su respectivo título. Con el objetivo de cuidar la estética del módulo se utiliza la librería *bootstrap3* [49], que estandariza los elementos desplegados en un sitio web permitiendo de forma simple un diseño bien cuidado desde el punto de vista estético.
2. `get_variables`: Se encarga de verificar que el formulario haya sido completado correctamente. De ser así, crea las variables de la oximetría y calcula el riesgo cardiovascular del paciente. Para esto, crea un diccionario con las variables del índice de Framingham con la estructura apropiada para poder utilizar la clase `RiskCalculator`. Para el cálculo de las variables de la oximetría crea un archivo temporal en la carpeta del proyecto que luego es recogido por la librería *pyedflib* para crear las señales correspondientes al pulso y nivel de saturación de oxígeno del paciente. Una vez calculado el riesgo, lo representa utilizando un *template* llamado `indicator_display`. Este utiliza la librería *jqwidgets* [50] para desplegar gráficamente el índice.

El módulo se exhibe al doctor Juan Carrillo con buena acogida. Se sugiere que en un producto final se ponga énfasis en la usabilidad del modelo. Particularmente, debe considerarse que el médico usualmente no sabe como acceder al archivo EDF que representa la oximetría y que por lo tanto, el sistema debiera recogerla de forma automática o a través de algún mecanismo muy simple de utilizar. Se recomienda además desplegar el valor de los indicadores calculados de la oximetría.

# Capítulo 5

## Evaluación del impacto económico y social

Esta sección tiene como objetivo estimar el beneficio económico y social que significaría la implementación del modelo en el Instituto Nacional del Torax y potencialmente a lo largo de todo Chile. La implementación consistiría, a grandes rasgos, en medir la probabilidad de riesgo de ECV en cada paciente que se realiza un estudio de sueño y tomar medidas preventivas en caso de existir riesgo. Es muy importante recalcar que lo que busca este análisis no es determinar la viabilidad o rentabilidad del proyecto sino meramente ilustrar la magnitud de su potencial impacto. Cabe mencionar también, que el análisis se centra en los dos eventos cardiovasculares más recurrentes según SHHS, que son infartos de miocardio e insuficiencia cardiaca.

### 5.1 Impacto en el Instituto nacional del Torax

Primero, se estima el impacto económico y social, asumiendo que el modelo se implementa únicamente en el Instituto Nacional del Torax. En una primera etapa, el modelo se probaría de forma exclusiva en esta institución y es por eso que es relevante hacer la estimación considerando únicamente este recinto.

Según la Unidad de Medicina del Sueño del INT, se realizaron 517 estudios de sueño (oximetrías y polisomnografías) en 2015 y en general, si se observa el mismo dato en años previos, se aprecia que existe una tendencia al alza en la cantidad de estudios realizados. La tabla 5.1 ilustra esta tendencia.

	Año					
	2010	2011	2012	2013	2014	2015
Estudios de Sueño	278	267	226	296	375	517

Tabla 5.1: Número de estudios de sueño en INT entre 2010 y 2015

La tasa de crecimiento promedio según los datos es de un 15 %. Suponiendo que esta

tasa se mantiene, y que la puesta en marcha del modelo comienza en 2018, en 10 años, el hospital habrá realizado 16.258 estudios de sueño.

Por otro lado, según los datos de SHHS, cerca de un 22 % de los pacientes con problemas de sueño tienen alguna enfermedad cardiovascular durante los 10 años siguientes a la realización del estudio de sueño.

Basado en el desempeño del modelo en los datos puede estimarse el porcentaje de personas en las que el modelo sería de ayuda para una prevención temprana. Esto presenta cierta dificultad pues dependiendo del criterio del doctor y de las características del paciente, se hacen recomendaciones distintas en cuanto al tratamiento a seguir para un mismo nivel de riesgo cardiovascular. Por otra parte, se hacen distintos tratamientos según el nivel de riesgo del paciente. Con el objetivo de simplificar el problema para efectos de este análisis, se define cierto nivel de riesgo cardiovascular y luego se asume que a todos los pacientes pasado dicho nivel se les realiza el mismo tratamiento. Y que a los pacientes bajo dicho nivel, por otra parte, no se les realiza ningún tipo de tratamiento.

Dado el supuesto anterior, el resultado varía según el nivel de riesgo escogido para el corte. Se decide entonces realizar el análisis para 3 cortes :

1. Corte en una probabilidad de ECV de 5 %
2. Corte en una probabilidad de ECV de 10 %
3. Corte en una probabilidad de ECV de 20 %

Se escogieron estos 3 cortes ya que ilustran los niveles de riesgo representados en la matriz de Framingham, en donde bajo un 5 % el riesgo se considera bajo, entre un 5 % y un 9 % se considera moderado, entre un 10 % y 19 % se considera alto y sobre 20 % se considera muy alto.

Para cada uno de estos cortes se calcula entonces el *recall* del modelo. El número de pacientes atendidos preventivamente puede estimarse simplemente como el porcentaje de pacientes que se hicieron un estudio de sueño de forma preventiva y que obtuvieron a través del modelo un nivel de riesgo superior al corte definido. Por ejemplo, si se define un corte en un 10 % y dado que, para éste umbral, el modelo tiene un *recall* de un 95 % (ver sección 3.4.3 ), el número de personas con riesgo real que reciben un tratamiento preventivo es de  $16,258 * 0,22 * 0,95 = 3,398$ . En la tabla 5.2 se muestran los resultados para cada uno de los cortes definidos

	<b>Corte</b>		
	5 %	10 %	20 %
<b>Recall</b>	0,986	0,950	0,812
<b>Pacientes con riesgo de ECV atendidas preventivamente</b>	3.526	3.398	2.904

Tabla 5.2: *Recall* y número de personas que se atenderían preventivamente para cada corte definido

Evidentemente, no todos los pacientes que se atienden preventivamente consiguen evitar el padecimiento de un ECV. Según N. J. Wald y M. R. Law [51], se estima que un tratamiento basado en fármacos puede reducir en un 88 % la probabilidad de padecer enfermedad cardíaca isquémica (enfermedades asociadas a problemas de circulación de la sangre hacia el corazón, como infartos de miocardio, angina, arritmia, insuficiencia cardíaca, etc). Suponiendo que un 88 % de los pacientes atendidos preventivamente evita que ocurra un problema cardiovascular, la tabla muestra el número de pacientes que se verían beneficiados en este sentido.

	<b>Corte</b>		
	<b>5 %</b>	<b>10 %</b>	<b>20 %</b>
<b>Personas beneficiadas</b>	3.103	2.990	2.556

Tabla 5.3: Numero de personas que prevendrían una ECV producto de la implementación del modelo

Por otro lado, el costo para la sociedad de una enfermedad puede clasificarse en 2 tipos:

1. El costo en dinero
2. El costo en cuanto a calidad de vida

Según B. Stevens, L. Pezzull et al. [52] el costo económico total del país completo por efecto de las insuficiencias cardíacas que se produjeron en Chile en 2015 es de 251.125 millones de pesos al año y por efecto de infartos agudos de miocardio (ataque al corazón) es de 333.766 millones de pesos al año. Esto incluye costos asociados a gastos del sistema de salud y costos en términos de pérdida de productividad. Según la misma fuente, el costo anual en DALYs (ver 2.4.3) es de 37.268 por efecto de insuficiencias cardíacas y 51.267 por efecto de infartos agudos de miocardio.

Dado que un 2 % de la población chilena adulta sufrió en 2015 una insuficiencia cardíaca y un 0,1 % sufrió un infarto agudo de miocardio [52] y tomando en cuenta además que el porcentaje de la población chilena mayor a 25 años en 2015 fue de cerca de 64 % [53] y que la población total en Chile en 2015 fue de cerca de 18 millones [54], se puede estimar fácilmente el costo por persona para el país de sufrir alguna de estas enfermedades. El resultado de la estimación se muestra en la tabla 5.4

	<b><i>Insuficiencia Cardíaca</i></b>	<b><i>Infarto de Miocardio</i></b>
Número de personas afectadas	230.400	11.520
Beneficio monetario total	251.125.000.000	333.766.000.000
Beneficio en DALYs total	37.268	51.267
<b>Beneficio monetario por persona</b>	1.089.952	28.972.743
<b>Beneficio DALYs por persona</b>	0,16	4.45

Tabla 5.4: Estimación del costo por persona en DALYs y en pesos chilenos de tener una insuficiencia cardíaca o un infarto de miocardio

Según los datos de SHHS, del total de pacientes que padecieron alguna ECV, un 26 % tuvo una insuficiencia cardíaca y un 16 % tuvo un infarto agudo de miocardio. El análisis

considera el impacto económico del modelo en términos de estas dos enfermedades que, según los datos en SHHS, son los dos eventos cardiovasculares más frecuentes y que además se encuentran bien estudiadas en cuanto a su impacto en la población chilena. Dado lo anterior, la estimación del impacto total para el INT de implementar el modelo de predicción desarrollado en esta memoria se ilustra en las tablas 5.5, 5.6 y 5.7.

	<b>Corte al 5 %</b>	
	<b><i>Insuficiencia Cardíaca</i></b>	<b><i>Infarto de Miocardio</i></b>
Numero de personas	807	496
<b>Beneficio monetario total</b>	880 Millones	14.371 Millones
<b>Beneficio DALYs total</b>	129	2.207
<b>Beneficio monetario agregado</b>	15.251 Millones	
<b>Beneficio DALYs agregado</b>	2.336	

Tabla 5.5: Estimación de beneficios monetarios y en DALYs de la implementación del modelo en el hospital salvador para un corte al 5 %

	<b>Corte al 10 %</b>	
	<b><i>Insuficiencia Cardíaca</i></b>	<b><i>Infarto de Miocardio</i></b>
Numero de personas	777	478
<b>Beneficio monetario total</b>	847 Millones	13.849 Millones
<b>Beneficio DALYs total</b>	124	2.127
<b>Beneficio monetario agregado</b>	14.696 Millones	
<b>Beneficio DALYs agregado</b>	2.251	

Tabla 5.6: Estimación de beneficios monetarios y en DALYs de la implementación del modelo en el hospital salvador para un corte al 10 %

	<b>Corte al 20 %</b>	
	<b><i>Insuficiencia Cardíaca</i></b>	<b><i>Infarto de Miocardio</i></b>
Numero de personas	665	409
<b>Beneficio monetario total</b>	725 Millones	11.850 Millones
<b>Beneficio DALYs total</b>	106	1.820
<b>Beneficio monetario agregado</b>	12.575 Millones	
<b>Beneficio DALYs agregado</b>	1.926	

Tabla 5.7: Estimación de beneficios monetarios y en DALYs de la implementación del modelo en el hospital salvador para un corte al 20 %

Como puede observarse, el beneficio monetario asociado a ambas causas descritas se estima entre 1.928 y 2.338 millones de pesos y entre 1.945 y 2.359 DALYs. Cabe mencionar que el beneficio estimado asume la utilización del modelo durante 10 años desde el año 2018. En promedio, el beneficio anual se estima entre 192,8 y 233,8 millones de pesos y entre 194 y 235 DALYs.

Cabe mencionar que este análisis estima los beneficios potenciales de la implementación del modelo pero no hace mención a los costos, tanto monetarios, como en términos de potencial empeoramiento en la calidad de vida de pacientes por causa de someterse a un tratamiento. Claramente, es necesaria su estimación para una evaluación de viabilidad y rentabilidad social del proyecto. Más aún, si se hiciera una propuesta basada meramente en este análisis evidentemente la mejor opción sería tratar a la mayor cantidad posible de personas, pero esto se debe en realidad a que no se está considerando el costo de los tratamientos. La razón por la cual no se realizó una estimación de costos es que para ello se requiere una definición más acabada de la forma de implementación del modelo en el sistema, lo cual a su vez requiere un entendimiento profundo de los procesos internos del INT o el lugar en donde se desee realizar la implantación. Y ambos aspectos están fuera de los alcances de esta memoria.

## 5.2 Impacto potencial en Chile

En Chile existen actualmente múltiples instituciones tanto en el sector público como el sector privado que atienden a pacientes con problemas de sueño utilizando oximetrías nocturnas o polisomnografías como herramienta de diagnóstico. Entre ellas se encuentran, además del INT, el Hospital Clínico de la Universidad Católica, la Clínica Alemana, la Clínica la Condes y el Hospital Militar por mencionar sólo algunas. Además, existe un número muy elevado de personas que tienen problemas de sueño. Es por esto que es interesante estimar el impacto potencial que la implantación del modelo podría significar a nivel nacional.

Según la Encuesta Nacional de Salud y Nutrición [6] realizada en 2016, en donde se hicieron preguntas relacionadas con diversos problemas de sueño a 8.649 chilenos, un 10 % califica la calidad de su sueño como mala o muy mala.

Por otra parte, en *World Population Data Sheet* [55], se proyecta que la población chilena crezca hasta 20,3 millones de personas en 2030, lo cual implica que la tasa de crecimiento promedio es de alrededor de 153.000 personas al año (asumiendo que en 2015 la población fue de 18 millones de habitantes).

Asumiendo esta tasa de crecimiento y que el 10 % de la población adulta tiene problemas de sueño (y suponiendo además que un 64 % de la población es adulta), el número de personas al año con problemas de sueño y que podrían potencialmente realizarse un estudio de sueño se muestra en la tabla 5.8.

En total, entre los años 2018 y 2027 el mercado potencial, es decir, el total de personas que podrían ser sujetas de una evaluación de riesgo cardiovascular en base al modelo planteado en esta memoria es de 12.254.400 personas. Utilizando los mismos supuestos de la sección 5.1, se deducen en las tabla 5.9, 5.10 y 5.11 los beneficios potenciales monetarios y en DALYs para la población Chilena.

En conclusión, se estima que en 10 años, el beneficio total sería de entre 10 y 13 billones de pesos y de entre 1.6 y 2 millones de DALYs aproximadamente. El beneficio anual

	<b>Población Total</b>	<b>Población Adulta</b>	<b>Población con problemas de Sueño</b>
2015	18.000.000	11.520.000	1.152.000
2016	18.153.000	11.617.920	1.161.792
2017	18.306.000	11.715.840	1.171.584
2018	18.459.000	11.813.760	1.181.376
2019	18.612.000	11.911.680	1.191.168
2020	18.765.000	12.009.600	1.200.960
2021	18.918.000	12.107.520	1.210.752
2022	19.071.000	12.205.440	1.220.544
2023	19.224.000	12.303.360	1.230.336
2024	19.377.000	12.401.280	1.240.128
2025	19.530.000	12.499.200	1.249.920
2026	19.683.000	12.597.120	1.259.712
2027	19.836.000	12.695.040	1.269.504
2028	19.989.000	12.792.960	1.279.296
2029	20.142.000	12.890.880	1.289.088
2030	20.300.000	12.992.000	1.299.200

Tabla 5.8: Mercado potencial estimado por año de 2015 a 2030

	<b>Corte al 5 %</b>	
	<b><i>Insuficiencia Cardíaca</i></b>	<b><i>Infarto de Miocardio</i></b>
Numero de personas	691.138	425.316
<b>Beneficio monetario total</b>	753.308 Millones	12.332.569 Millones
<b>Beneficio DALYs total</b>	110.582	1.892.656
<b>Beneficio monetario agregado</b>	13.075.876 Millones	
<b>Beneficio DALYs agregado</b>	2.003.238	

Tabla 5.9: Estimación de beneficios potenciales monetarios y en DALYs de la implementación del modelo en la población chilena para un corte al 5 %

	<b>Corte al 10 %</b>	
	<b><i>Insuficiencia Cardíaca</i></b>	<b><i>Infarto de Miocardio</i></b>
Numero de personas	665.904	409.787
<b>Beneficio monetario total</b>	725.804 Millones	11.872.657 Millones
<b>Beneficio DALYs total</b>	106.544	1.823.553
<b>Beneficio monetario agregado</b>	12.598.461 Millones	
<b>Beneficio DALYs agregado</b>	1.930.097	

Tabla 5.10: Estimación de beneficios potenciales monetarios y en DALYs de la implementación del modelo en la población chilena para un corte al 10 %



	<b>Corte al 20 %</b>	
	<b><i>Insuficiencia Cardíaca</i></b>	<b><i>Infarto de Miocardio</i></b>
Numero de personas	569.173	350.260
<b>Beneficio monetario total</b>	620.371 Millones	10.147.998 Millones
<b>Beneficio DALYs total</b>	91.068	1.558.658
<b>Beneficio monetario agregado</b>	10.768.369 Millones	
<b>Beneficio DALYs agregado</b>	1.649.725	

Tabla 5.11: Estimación de beneficios potenciales monetarios y en DALYs de la implementación del modelo en la población chilena para un corte al 20 %

promedio sería de entre 1 y 1.3 billones de pesos y entre 160.000 y 200.000 DALYs. Cabe mencionar que esto se refiere al beneficio que habría en caso de que todos los chilenos con problemas de sueño hicieran uso del modelo. Se aclara además que este análisis está hecho sólo para Chile, pudiéndose aumentar el beneficio si la solución se exporta hacia otros países.

# Capítulo 6

## Conclusiones

### 6.1 Conclusiones generales

La memoria consiste en el desarrollo de un modelo de predicción de riesgo cardiovascular en pacientes con problemas de sueño. El modelo se basa en factores de riesgo cardiovascular tradicionales, pero además en indicadores derivados de un examen médico llamado oximetría nocturna, en el que se mide el nivel de saturación de oxígeno en la sangre y el pulso cardíaco durante toda la noche. Este examen, que suele realizarse en pacientes con problemas de sueño, muestra indicios de poder predictivo de riesgo cardiovascular en diversas investigaciones recientes. Dado aquello, se plantea la siguiente hipótesis de investigación:

- Se puede construir un indicador en base a oximetrías nocturnas y variables demográficas y clínicas que sirvan como una alternativa más confiable al índice de Framingham en pacientes con problemas de sueño.

Es decir, el punto de comparación para el modelo que se construye es el índice de Framingham, que es una herramienta de predicción de riesgo cardiovascular muy reconocida y utilizada a nivel mundial.

La comparación entre modelos se hace a través de la métrica de desempeño AUC, que fue desarrollada justamente para la evaluación de modelos cuyo *output* es una probabilidad. En este caso, la probabilidad de padecer una enfermedad cardiovascular en los 10 años siguientes a la predicción del modelo.

El modelo final tiene un AUC de 0.775 contra un 0.739 que tiene el índice de Framingham, siendo la diferencia estadísticamente significativa. Con ello se valida la hipótesis de investigación postulada. Sin embargo, se encuentra que las variables derivadas de la oximetría tienen muy poco peso en el modelo en comparación a las variables tradicionales, lo cual indicaría que el alza en el desempeño se debe a ciertas características específicas de la población estudiada (probablemente la prevalencia de pacientes con problemas de sueño) más que a la relevancia de las variables de la oximetría.

Se realiza además un módulo de visualización del modelo, con el fin de facilitar la divulgación de los resultados a futuros investigadores. Por último, se estima los potenciales beneficios económicos y sociales de la implementación del modelo como herramienta de ayuda al diagnóstico médico exclusivamente en el Instituto Nacional del Tórax y también a lo largo de todo el país. Esto es considerando el efecto de la prevención las dos principales enfermedades cardiovasculares, que son Insuficiencia Cardíaca e Infarto de Miocardio.

El beneficio económico para el Instituto Nacional del Tórax es de entre 1.928 y 2.338 millones de pesos en 10 años. El costo social se midió en términos de DALYs, que permiten cuantificar el impacto en términos de calidad de vida. Se estimó un beneficio de entre 1.945 y 2.359 DALYs en 10 años. El beneficio potencial de la implementación en todo el país es de entre 10 y 13 billones de pesos y entre 1,6 y 2 millones de DALYs en 10 años.

En términos generales, se cumple el objetivo de la memoria, que consiste en la selección, implementación y validación de un modelo para la predicción de riesgo cardiovascular más confiable que índices tradicionales, sentando las bases para el desarrollo de una herramienta de apoyo al diagnóstico médicos.

Se hace primero una revisión de varios modelos con potencial poder de predicción, basado en investigaciones previas recogidas de la literatura. Aunque existen modelos de *Machine Learning* que no son probados e índices potencialmente útiles que no son implementados, se da un primer paso importante en cuanto a que se abren caminos para que continúe el curso de la investigación.

Sobre la implementación de los modelos escogidos, esta fue exitosa, logrando satisfactoriamente la implementación de los índices y el entrenamiento de los distintos modelos, incluyendo además, una sección de *testing*, que facilita el desarrollo de *software* a investigadores futuros.

Acerca del desempeño de los modelos, se logra una diferencia significativa con el índice de Framingham, lo cual es clave en la validación de la hipótesis de esta memoria y en el cumplimiento del objetivo general. Más aún, proyecta una línea de investigación muy esperanzadora para el WIC y el INT.

El módulo de visualización fue implementado sin inconvenientes, permitiendo con ello transmitir concretamente una idea de como funcionaría el modelo en un producto final, fomentando así el trabajo de investigación e implantación que debe realizarse tras esta memoria.

Finalmente, el beneficio de la implantación en términos tanto económicos como sociales fue cuantificado, lo cual se cree ayudará a impulsar la investigación futura, pues las cifras estimadas son muy prometedoras.

## 6.2 Trabajo Futuro

El objetivo de esta sección es servir de guía a futuros investigadores, permitiéndoles hacerse una idea de los aspectos de esta memoria sobre los que hay espacio para aportar a través de distintos caminos a la investigación.

En primer lugar, es muy importante recalcar que el modelo fue entrenado en base a datos de Estados Unidos y por lo tanto no puede garantizarse su buen funcionamiento en Chile. Dado esto, es imperativa la recolección y entrenamiento de datos en Chile para que el modelo tenga validez en nuestro país.

Por otra parte, el modelo en sí es perfectible de varias formas. En primer lugar, existen dos índices (RRPO y PASD) ya identificados que podrían calcularse a partir de una oximetría común y potencialmente ayudar a mejorar el desempeño del modelo según investigaciones previas. Además, existen señales adicionales que también podrían tener cierto poder predictivo de ECV que no son medidas por oxímetros de pulso tradicionales, pero sí podrían ser capturadas por algún otro dispositivo. Según el doctor Juan Carrillo, la polisomnografía y la actimetría son exámenes que podrían potencialmente contribuir a una mejor predicción. Adicionalmente, pueden hacerse optimizaciones en los modelos, por ejemplo, variando los hiperparámetros de la SVM o cambiando la configuración de la red neuronal. Existen además otros modelos de *Machine Learning* que podrían probarse, como por ejemplo árboles de decisión.

Por último, hay un tarea importante pendiente en cuanto a la implantación del modelo en producción. Para ello, en primer lugar debe realizarse un estudio de la dinámica interna de la institución en donde comenzaría la implementación, seguramente en el Instituto Nacional del Tórax, de forma de diseñar una estrategia de transferencia tecnológica acorde y poder estimar su costo. La estrategia debe incluir la implementación de un software listo para producción, capacitaciones a los doctores y la creación de un plan de prevención cardiovascular basado en el modelo, que incluya mediciones periódicas y distintos tratamientos según se requiera.

# Bibliografía

- [1] OMS Información general de Enfermedades cardiovasculares, <http://www.who.int/mediacentre/factsheets/fs317/es/>.
- [2] Ministerio de Salud Información desagregada sobre causas de muerte en Chile, <http://www.deis.cl/defunciones-y-mortalidad-general-y-por-grupos-de-edad/>.
- [3] L. Núñez, J. Marrugat, V. Mujica, M. C. Escobar, A. L. Jiménez, P. Pérez, I. Palomo y col., «Estimación de riesgo de enfermedad coronaria mediante la función de Framingham adaptada para la población chilena», *Revista médica de Chile*, vol. 137, n.º 10, págs. 1273-1282, 2009.
- [4] S. M. Grundy, R. Pasternak, P. Greenland, S. Smith y V. Fuster, «Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: a statement for healthcare professionals from the American Heart Association and the American College of Cardiology», *Journal of the American College of Cardiology*, vol. 34, n.º 4, págs. 1348-1359, 1999.
- [5] M. Lavente, «Predictive Value of the Framingham Risk Score in Identifying High Cardiovascular Risk», *Internacional Journal of Cardiovascular Sciences*,
- [6] *Encuesta Nacional de Salud y Nutrición Encuesta realizada en 2016 en donde se recopila información sobre los hábitos de sueño de la población chilena*, [http://ensanut.insp.mx/ensanut2016/descarga\\_bases.php#.WSxr-sm1tdi](http://ensanut.insp.mx/ensanut2016/descarga_bases.php#.WSxr-sm1tdi).
- [7] F. Campos-Rodriguez, M. A. Martinez-Garcia, M. Martinez, J. Duran-Cantolla, M. d. l. Peña, M. J. Masdeu, M. Gonzalez, F. d. Campo, I. Gallego, J. M. Marin y col., «Association between obstructive sleep apnea and cancer incidence in a large multicenter Spanish cohort», *American journal of respiratory and critical care medicine*, vol. 187, n.º 1, págs. 99-105, 2013.
- [8] F. J. Nieto, P. E. Peppard, T. Young, L. Finn, K. M. Hla y R. Farré, «Sleep-disordered breathing and cancer mortality: results from the Wisconsin Sleep Cohort Study», *American journal of respiratory and critical care medicine*, vol. 186, n.º 2, págs. 190-194, 2012.

- [9] I. Rosenzweig, S. C. Williams y M. J. Morrell, «The impact of sleep and hypoxia on the brain: potential mechanisms for the effects of obstructive sleep apnea», *Current opinion in pulmonary medicine*, vol. 20, n.º 6, págs. 565-571, 2014.
- [10] L. Lavie, «Oxidative stress in obstructive sleep apnea and intermittent hypoxia—revisited—the bad ugly and good: implications to the heart and brain», *Sleep medicine reviews*, vol. 20, págs. 27-45, 2015.
- [11] D. Sommermeyer, D. Zou, D. N. Eder, J. Hedner, J. H. Ficker, W. Randerath, C. Priegnitz, T. Penzel, I. Fietze, B. Sanner y col., «The use of overnight pulse wave analysis for recognition of cardiovascular risk factors and risk: a multicentric evaluation», *Journal of hypertension*, vol. 32, n.º 2, págs. 276-285, 2014.
- [12] J.-Y. Dong, Y.-H. Zhang y L.-Q. Qin, «Obstructive sleep apnea and cardiovascular risk: meta-analysis of prospective cohort studies», *Atherosclerosis*, vol. 229, n.º 2, págs. 489-495, 2013.
- [13] *National Sleep Foundation Influencia de los problemas de sueño en el riesgo de contraer enfermedades cardiovasculares*, <https://sleepfoundation.org/sleep-news/how-sleep-deprivation-affects-your-heart>.
- [14] *Sturdust II Aparato para medir el nivel de saturación de oxígeno en la sangre*, <http://www.usa.philips.com/healthcare/product/HC1025720/stardust-ii-portable-sleep-recorder>.
- [15] *Sleepware Medidor de saturación de oxígeno*, <http://www.usa.philips.com/healthcare/product/HC1082462/sleepware-g3-sleep-diagnostic-software>.
- [16] *Somno Check Cardio Medidor de saturación de oxígeno*, <http://www.weinmann-medical.com/en/>.
- [17] P. Brindle, E. Jonathan, F. Lampe, M. Walker, P. Whincup, T. Fahey y S. Ebrahim, «Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study», *Bmj*, vol. 327, n.º 7426, pág. 1267, 2003.
- [18] J. P. Empana, P. Ducimetière, D. Arveiler, J. Ferrières, A. Evans, J. B. Ruidavets, B. Haas, J. Yarnell, A. Bingham, P. Amouyel y J. Dallongeville, «Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations?The PRIME Study», *European Heart Journal*, vol. 24, n.º 21, págs. 1903-1911, nov. de 2003, ISSN: 0195-668X. DOI: 10.1016/j.ehj.2003.09.002. dirección: <https://academic.oup.com/eurheartj/article/24/21/1903/450072/Are-the-Framingham-and-PROCAM-coronary-heart> (visitado 15 de mayo de 2017).
- [19] D. Laurier, N. P. Chau, B. Cazelles, P. Segond y P.-M. Group12, «Estimation of CHD risk in a French working population using a modified Framingham model», *Journal of clinical epidemiology*, vol. 47, n.º 12, págs. 1353-1364, 1994.

- [20] M. G. I. Noguera y col., «TABLAS DE RIESGO CARDIOVASCULAR PARA LA POBLACION CHILENA», 2006.
- [21] E. Watanabe, K. Kiyono, S. Matsui, V. K. Somers, K. Sano, J. Hayano, T. Ichikawa, M. Kawai, M. Harada e Y. Ozaki, «Prognostic Importance of Novel Oxygen Desaturation Metrics in Patients With Heart Failure and Central Sleep Apnea», *Journal of Cardiac Failure*, 2016.
- [22] *National Sleep Research Resource Repositorio de bases de datos de investigaciones relacionadas con problemas de sueño*, <https://sleepdata.org>.
- [23] *European Data Format Formato digital en que se almacena una oximetría nocturna*, <http://www.edfplus.info/specs/edf.html>.
- [24] *National Sleep Research Resource Base de datos de pacientes enfocada en problemas respiratorios al dormir y sus efectos en enfermedades cardiovasculares*, <https://sleepdata.org/datasets/shhs>.
- [25] *Python Lenguaje de programación utilizado para la implementación de modelos de machine learning en esta memoria*, <https://www.python.org/>.
- [26] I. S. Graham, *The HTML sourcebook*. John Wiley & Sons, Inc., 1995.
- [27] *Django Framework de desarrollo web*, <https://www.djangoproject.com/>.
- [28] *Git Sistema de versionamiento utilizado en esta memoria*, <https://git-scm.com/>.
- [29] *HeartScore Medelo de predicción de riesgo cardiovascular*, <http://www.heartscore.org/Pages/background.aspx>.
- [30] *European Society of Cardiology Organización Europea dedicada al estudio de formas de prevenir enfermedades cardiovasculares*, <https://www.escardio.org/>.
- [31] *SCORE Índice europeo de predicción de riesgo cardiovascular*, [https://www.escardio.org/static\\_file/Escardio/Subspecialty/EACPR/Documents/score-charts.pdf](https://www.escardio.org/static_file/Escardio/Subspecialty/EACPR/Documents/score-charts.pdf).
- [32] G. Assmann, P. Cullen y H. Schulte, «Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study», *Circulation*, vol. 105, n.º 3, págs. 310-315, 2002.
- [33] *European Society of Hypertension Organización Europea dedicada al estudio de la hipertensión*, <http://www.eshonline.org/>.
- [34] G. Mancia, G. De Backer, A. Dominiczak, R. Cifkova, R. Fagard, G. Germano, G. Grassi, A. M. Heagerty, S. E. Kjeldsen, S. Laurent y col., «2007 ESH-ESC practice guidelines for the management of arterial hypertension: ESH-ESC task force on

- the management of arterial hypertension», *Journal of hypertension*, vol. 25, n.º 9, págs. 1751-1762, 2007.
- [35] N. Alvis y M. T. Valenzuela, «Los QALYs y DALYs como indicadores sintéticos de salud», *Revista médica de Chile*, vol. 138, págs. 83-87, 2010.
- [36] D. W. Hosmer Jr, S. Lemeshow y R. X. Sturdivant, *Applied logistic regression*. John Wiley & Sons, 2013, vol. 398.
- [37] *Support Vector Machine Explicación de la técnica de SVM*, <http://scikit-learn.org/stable/modules/svm.html>.
- [38] S. Haykin y N. Network, «A comprehensive foundation», *Neural Networks*, vol. 2, n.º 2004, pág. 41, 2004.
- [39] D. D. Lewis, «Naive (Bayes) at forty: The independence assumption in information retrieval», en *European conference on machine learning*, Springer, 1998, págs. 4-15.
- [40] S. Marsland, *Machine learning: an algorithmic perspective*. CRC press, 2015.
- [41] *Cross Validation Explicación de la técnica de Cross Validation*, [http://scikit-learn.org/stable/modules/cross\\_validation.html](http://scikit-learn.org/stable/modules/cross_validation.html).
- [42] J. Deacon, «Model-view-controller (mvc) architecture», *Online* [Citado em: 10 de março de 2006.] <http://www.jdl.co.uk/briefings/MVC.pdf>, 2009.
- [43] G. Assmann, P. Cullen y H. Schulte, «Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study», *Circulation*, vol. 105, n.º 3, págs. 310-315, 2002.
- [44] R. Conroy, K Pyörälä, A. e. Fitzgerald, S Sans, A Menotti, G. De Backer, D. De Bacquer, P Ducimetiere, P Jousilahti, U Keil y col., «Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project», *European heart journal*, vol. 24, n.º 11, págs. 987-1003, 2003.
- [45] *PyEDFlib Librería de python para lectura de archivos EDF*, <http://pyedflib.readthedocs.io/en/latest/>.
- [46] *Libría Numpy Librería de Python para el manejo de arreglos multidimensionales*, <http://www.numpy.org/>.
- [47] *Unittest Módulo de python para realización de tests unitarios*, <https://docs.python.org/2/library/unittest.html>.
- [48] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau,



M. Brucher, M. Perrot y E. Duchesnay, «Scikit-learn: Machine Learning in Python», *Journal of Machine Learning Research*, vol. 12, págs. 2825-2830, 2011.

- [49] *Bootstrap3 Librería de diseño web*, <https://django-bootstrap3.readthedocs.io/en/latest/>.
- [50] *jqwidgets Librería utilizada para crear la representación gráfica del riesgo cardiovascular en el módulo funcional*, <http://www.jqwidgets.com/>.
- [51] N. J. Wald y M. R. Law, «A strategy to reduce cardiovascular disease by more than 80 %», *Bmj*, vol. 326, n.º 7404, pág. 1419, 2003.
- [52] B. Stevens, L. Pezzullo, L. Verdian, J. Tomlinson y S. Zegenhagen, «The Economic Burden of Heart Diseases in Latin America»,
- [53] C. Factbook, «The World Factbook; 2010», *See also: <http://www.cia.gov/library/publications/the-worldfactbook>*, accessed January, vol. 30, 2015.
- [54] *Población chilena Estimación de la población chilena según el banco mundial*, <http://data.worldbank.org/indicator/SP.POP.TOTL>.
- [55] *Word Population Data Sheet Estudio que estima la evolución de la población mundial, incluyendo de forma desagregada la población de Chile*, <http://www.prb.org/pdf16/prb-wpds2016-web-2016.pdf>.