

Tabla de contenido

1. Introducción	1
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
2. Marco Conceptual	14
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
3.	Desarrollo del modelo de predicción	29
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
4.	Módulo de visualización	52
4.1.	Descripción desde el punto de vista del usuario	52
4.2.	Implementación	53
5.	Evaluación del impacto económico y social	58
5.1.	Impacto en el Instituto nacional del Torax	58
5.2.	Impacto potencial en Chile	62
6.	Conclusiones	65
6.1.	Conclusiones generales	65
6.2.	Trabajo Futuro	67
	Bibliografía	68

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