



Universidad
de La Laguna

Escuela Superior de
Ingeniería y Tecnología
Sección de Ingeniería Informática

Trabajo de Fin de Grado

Redes Neuronales Artificiales aplicadas a medicina

Artificial Neural Networks applied to medicine

Óscar Trujillo Acosta

La Laguna, 4 de marzo de 2015

D. **Patricio García Báez**, con N.I.F. 43.356.987-D profesor Titular de Universidad adscrito al Departamento de Ingeniería Informática y de Sistemas de la Universidad de La Laguna, como tutor

D. **Carmen Paz Suárez Araujo**, con N.I.F. 43.640.373-N profesor Titular de Universidad adscrito al Departamento de Informática y Sistemas de la Universidad de Las Palmas de Gran Canarias, como cotutor

C E R T I F I C A (N)

Que la presente memoria titulada:

"Redes Neuronales Artificiales aplicadas a medicina."

ha sido realizada bajo su dirección por D. **Óscar Trujillo Acosta**, con N.I.F. 12.345.678-X.

Y para que así conste, en cumplimiento de la legislación vigente y a los efectos oportunos firman la presente en La Laguna a 20 de Junio de 2016.

Agradecimientos

FALTA

Licencia



© Esta obra está bajo una licencia de
Creative Commons Reconocimiento 4.0
Internacional.

Artificial Neural Networks applied to medicine

Resumen

La computación tradicional es un campo que ha ido ligado a la historia de la especie humana desde sus inicios. Siglos y siglos de avance técnico han dado sus frutos y hoy en día disponemos de máquinas capaces de realizar cálculos computacionales muy superiores a lo que un ser humano es capaz de hacer, no obstante no todo es una derrota y es que los animales seguimos poseyendo algo único: la capacidad de razonar. La naturaleza nos ha dotado de un sistema nervioso formado por una red neuronal muy compleja que nos permite aprender, razonar, deducir, etc.

En el último siglo el ser humano ha aunado esfuerzos en pos del desarrollo de sistemas que tratan de recrear esta forma de procesamiento humana, sistemas que han evolucionado hasta lo que hoy conocemos como: "Redes Neuronales Artificiales" o "RNA".

Se propone con este proyecto el desarrollo de una red neuronal artificial aplicada a la medicina, en concreto su funcionalidad será la de detectar tres tipos de cáncer: Cáncer de mama, cáncer de piel y cáncer de pulmón.

Palabras clave: Redes Neuronales Artificiales, medicina, detección de cáncer, sistemas inteligentes

Abstract

FALTA

Keywords: *Keyword1, Keyword2, Keyword3, ...*

Índice General

Capítulo 1. Introducción	1
1.1 Antecedentes	1
1.2 Objetivos	1
1.3 Programa de Apoyo a Trabajos Finales Libres	1
1.4 Alcance	2
 Capítulo 2. Redes Neuronales Artificiales, Datos y Herramientas	 3
2.1 Redes Neuronales Artificiales	3
2.1.1 Historia	3
2.1.2 Funcionamiento	3
2.1.3 Estado actual	3
2.2 Base de Datos	4
2.2.1 Datos en RNA	4
2.2.2 Heuristic and Evolutionary Algorithms Laboratory (HEAL)	4
2.2.3 Resultados de obtenidos por HEAL .	4
2.3 Herramientas	6
2.3.1 Herramientas de gestión de RNA ...	6
2.3.2 Microsoft Azure Machine Learning .	8
 Capítulo 3. Estudio	 15
3.1 Principios de los estudios	15
3.2 Resultados para cada tipo de cáncer analizado	18
3.2.1 Melanoma	18
3.2.2 Cáncer de pulmón	24
3.2.3 Cáncer de mama	31
3.3 Estudio de los sistemas bajo el subconjunto de test	38
3.3.1 Melanoma	39
3.3.2 Cáncer de pulmón	40
3.3.3 Cáncer de mama	41
3.4 Comparación de resultados con HEAL	42
3.4.1 Melanoma	42
3.4.2 Cáncer de pulmón	42

3.4.3 Cáncer de mama	43
Capítulo 4. Conclusiones y líneas futuras	44
Capítulo 5. Summary and Conclusions	45
Capítulo 6. Presupuesto	46
6.1 Conceptos	46
Bibliografía	47

Índice de figuras

Figura 2.1. Resultados HEAL cáncer de mama	5
Figura 2.2. Resultados HEAL melanoma	5
Figura 2.3. Resultados HEAL cáncer de pulmón	5
Figura 2.4. Algoritmos Azure	8
Figura 2.5. Árbol de decisión	10
Figura 2.6. Grafo acíclico dirigido	10
Figura 2.7. Grafo acíclico dirigido	11
Figura 2.8. Ejemplo resultados Azure	11
Figura 2.9. Ejemplo grafica ROC Azure	13
Figura 2.10. Ejemplo grafica P/R Azure	14
Figura 3.1. Esquema de las pruebas	15
Figura 3.2. Grafico bloque prueba	16
Figura 3.1. Melanoma bloque 1. (a)ROC. (b)P/R.	19
Figura 3.2. Melanoma bloque 2. (a)ROC. (b)P/R.	20
Figura 3.3. Melanoma bloque 3. (a)ROC. (b)P/R.	21
Figura 3.4. Melanoma bloque 4. (a)ROC. (b)P/R.	23
Figura 3.5. Resultados finales melanoma	24
Figura 3.6. Pulmón bloque 1. (a)ROC. (b)P/R.	25
Figura 3.7. Pulmón bloque 2. (a)ROC. (b)P/R.	26
Figura 3.8. Pulmón bloque 3. (a)ROC. (b)P/R.	27
Figura 3.9. Pulmón bloque 4. (a)ROC. (b)P/R.	29
Figura 3.10. Resultado final pulmón	30
Figura 3.11. Mama bloque 1. (a)ROC. (b)P/R.	32
Figura 3.12. Mama bloque 2. (a)ROC. (b)P/R.	33
Figura 3.13. Mama bloque 3. (a)ROC. (b)P/R.	34
Figura 3.14. Mama bloque 4. (a)ROC. (b)P/R.	36
Figura 3.15. Resultado final mama	37
Figura 3.16. Esquema de módulos subconjunto de test.	38
Figura 3.17. Melanoma test. (a)ROC. (b)P/R.	39
Figura 3.18. Melanoma test, índices	39
Figura 3.19. Mama test. (a)ROC. (b)P/R.	41

Figura 3.20. Mama test. (a)ROC. (b) P/R.....41

Figura 3.21. Mama test. (a)ROC. (b) P/R.....40

Figura 3.22. Mama test índices.....40

Índice de tablas

Tabla 3.1. Configuración redes melanoma bloque 1.....	18
Tabla 3.2. Resultados redes melanoma bloque 1.....	18
Tabla 3.3. Configuración redes melanoma bloque 2.....	20
Tabla 3.4. Resultados redes melanoma bloque 2.....	20
Tabla 3.5. Configuración redes melanoma bloque 3.....	21
Tabla 3.6. Resultados redes melanoma bloque 3.....	21
Tabla 3.7. Configuración redes melanoma bloque 4.....	22
Tabla 3.8. Resultados redes melanoma bloque 3.....	22
Tabla 3.9. Comparación final modelos melanoma.....	24
Tabla 3.10. Configuración redes pulmón bloque 1.....	24
Tabla 3.11. Resultados redes pulmón bloque 1.....	25
Tabla 3.12. Configuración redes pulmón bloque 2.....	26
Tabla 3.13. Resultados redes pulmón bloque 2.....	26
Tabla 3.14. Configuración redes pulmón bloque 3.....	27
Tabla 3.15. Resultados redes pulmón bloque 3.....	27
Tabla 3.16. Configuración redes pulmón bloque 4.....	28
Tabla 3.17. Resultados redes pulmón bloque 4.....	28
Tabla 3.18. Comparación final modelos pulmón.....	30
Tabla 3.19. Configuración redes mama bloque 1.....	31
Tabla 3.20. Resultados redes mama bloque 1.....	31
Tabla 3.21. Configuración redes mama bloque 2.....	32
Tabla 3.22. Resultados redes mama bloque 2.....	33
Tabla 3.23. Configuración redes mama bloque 3.....	34
Tabla 3.24. Resultados redes mama bloque 3.....	34
Tabla 3.25. Configuración redes mama bloque 4.....	35
Tabla 3.26. Resultados redes mama bloque 4.....	35
Tabla 3.27. Comparación final modelos mama.....	37
Tabla 3.28. Comparación Melanoma con HEAL.....	42
Tabla 3.29. Comparación mama con resultados de HEAL....	43
Tabla 3.30. Comparación pulmón con resultados de HEAL..	42
Tabla 6.1. Presupuesto.....	46

Capítulo 1.

Introducción

1.1 Antecedentes

Las redes neuronales artificiales han experimentado un importante auge en los últimos años, grandes empresas como Microsoft, Amazon, IBM o Google han apostado por desarrollar plataformas relacionadas con este campo, sistemas, por lo general, especialmente diseñados para compartir la información generada y utilizarla como piedra angular en su propio desarrollo.

En el campo de la medicina el estudio de redes neuronales artificiales no es nada nuevo, existen multitud de desarrollos que tratan de facilitar nuevos sistemas de detección precoz de enfermedades para optimizar y facilitar la labor de los profesionales del sector.

1.2 Objetivos

El objetivo de este trabajo de fin de grado es realizar un estudio sobre el uso de sistemas inteligentes en la predicción de cáncer con determinada herramienta. Se analizarán principalmente diferentes tipos de redes neuronales artificiales y se compararán con otros modelos utilizados en la inteligencia artificial como son los árboles de decisión en búsqueda del mejor modelo de detección precoz de cáncer con el conjunto de datos que se tiene.

1.3 Programa de Apoyo a Trabajos Finales Libres

Este trabajo de fin de grado está adherido al programa de apoyo a trabajos finales libres de la Oficina de Software Libre de La Universidad. [1]

1.4 Alcance

Se establece una serie de objetivos o metas que servirán para identificar el alcance de este trabajo de fin de grado:

- a) Preparación.
 - a. Búsqueda de información sobre redes neuronales.
 - b. Estudio sobre software relacionado con redes neuronales.
 - c. Estudio del trabajo del laboratorio HEAL.
- b) Preprocesado de datos.
 - a. Análisis de los datos.
 - b. Elección de herramienta de preprocesado.
 - c. Realizar preprocesado.
 - d. Seleccionar particiones del conjunto de datos.
- c) Diseño y desarrollo de la red neuronal.
 - a. Elegir modelos de red neuronal artificial.
 - b. Diseño de la red.
 - c. Entrenamiento y validación.
 - d. Optimización del número de neuronas de la red.
- d) Evaluación de la red neuronal.
 - a. Obtener porcentaje de acierto real.
 - b. Crear gráficas de errores y precisión.
 - c. Comparar resultado con resultado del equipo austriaco.
- e) Finalización.
 - a. Preparación de la memoria del TFG.
 - b. Preparación de la defensa de la memoria.

Capítulo 2.

Redes Neuronales Artificiales, Datos y Herramientas

2.1 Redes Neuronales Artificiales

2.1.1 Historia

Una red neuronal artificial, también llamada RNA por sus iniciales, es un paradigma informático de procesamiento de información inspirado en los sistemas neuronales biológicos.

2.1.2 Funcionamiento

Las RNA tratan la información de forma similar a como lo hace un cerebro humano, aprenden con ejemplos y no son programadas para que hagan una tarea específica.

El aprendizaje en un sistema biológico se realiza mediante el ajuste de las conexiones sinápticas entre las neuronas, lo que también se hace en las redes neuronales artificiales.

2.1.3 Estado actual

FALTA

2.2 Base de Datos

2.2.1 Datos en RNA

La propia naturaleza de funcionamiento de las redes neuronales artificiales requiere una gran cantidad de datos de entrenamiento para ser ajustada y que consiga funcionar correctamente. Los datos necesarios para este estudio son de carácter sensible y la disponibilidad certera de los mismos no es alta ya que se tratan de datos clínicos personales.

Esta dificultad se ha visto salvada gracias al contacto establecido con el "Heuristic and Evolutionary Algorithms Laboratory", también llamado HEAL.

2.2.2 Heuristic and Evolutionary Algorithms Laboratory (HEAL)

El Heuristic and Evolutionary Algorithms Laboratory es un grupo de investigadores austriaco que ha realizado un estudio similar en cuanto a objetivo al pretendido en este trabajo, han desarrollado su propio sistema de detección de cáncer usando árboles de decisión. Los árboles de decisión no son más que otro paradigma de predicción usado en el ámbito de la inteligencia artificial. Este grupo de investigación nos ha cedido los datos usados en este estudio para poder entrenar a nuestros sistemas y a la vez tenemos los resultados en porcentaje de acierto obtenidos por ellos, algo de gran utilidad si queremos comparar nuestros resultados finales con los obtenidos en pruebas de naturaleza similar.

2.2.3 Resultados de obtenidos por HEAL

Además de los datos de análisis sanguíneos, HEAL nos ha cedido sus resultados, frutos de una amplia investigación en el tema. Estos resultados resultan de gran utilidad para este estudio ya que nos sirve para comparar los resultados obtenidos con un estudio similar hecho con los mismos datos.

	Training			Test		
Criteria	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Best Accuracy	81.87%	89.79%	72.53%	74.36%	84.29%	62.65%
Best Sensitivity	76.63%	95.29%	54.63%	74.08%	92.67%	52.16%
Best of Both	77.05%	95.55%	55.25%	74.36%	91.88%	53.70%

Figura 1.2. Resultados HEAL cáncer de mama

	Training			Test		
Criteria	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Best Accuracy	83.20%	79.52%	86.39%	73.92%	66.19%	80.62%
Best Sensitivity	83.65%	81.43%	85.57%	73.04%	67.14%	78.14%
Best of Both	83.20%	80.24%	85.77%	73.15%	66.67%	78.76%

Figura 3.4. Resultados HEAL melanoma

	Training			Test		
Criteria	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Best Accuracy	92.47%	87.25%	96.27%	87.26%	76.71%	94.95%
Best Sensitivity	92.85%	89.06%	95.61%	87.64%	78.31%	94.44%
Best of Both	92.93%	88.96%	95.83%	87.30%	77.61%	94.37%

Figura 5.6. Resultados HEAL cáncer de pulmón

En las tres figuras anteriores se muestran los resultados finales obtenidos por el HEAL y seguidamente se encuentran la relación de índices que han usado para mostrar sus resultados:

$$\text{Sensitivity} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{Specifity} = \text{TN} / (\text{TN} + \text{FP})$$

$$\text{Accuracy} = \text{TP} / \text{All Predictions}$$

TP = True Positive

TN = True Negative

FN = False Negative

FP = False Positive

2.3 Herramientas

2.3.1 Herramientas de gestión de RNA

Existen multitud de herramientas dedicadas a la implementación de RNAs, para la realización de este trabajo se ha hecho un amplio estudio del campo y se han seleccionado un total de 7 herramientas que se han dividido en dos grupos. Por un lado librerías, generalmente de código abierto, que ofrecen un amplio abanico de algoritmos y funcionalidades en distintos lenguajes de programación entre los que destacan Python, C/C++ y R. Por otro lado se encuentran una serie de entornos de desarrollo ya sean en versión software o web que ofrecen una serie de algoritmos de aprendizaje genéticos junto con una GUI diseñada para facilitar cada paso.

1. Librerías

- NeuroLab: Simple y potente librería para RNAs bajo Python. Incluye modelos supervisados, no supervisados y recurrentes.
- PyBrain: Al igual que la anterior se trata de una librería Python que incluye modelos supervisados, no supervisados y recurrentes. Tiene la ventaja de que separa la creación de la red neuronal del hecho de entrenarla, cosa que es imposible en NeuroLab. Además PyBrain posee su propio tipo de datos.
- NeuralNet: Se trata de una librería en lenguaje R que incluye diferentes algoritmos de aprendizaje supervisado.
- Fast Artificial Neural Network: También conocida como FANN es una librería de código abierto escrita en C reconocida por su alta velocidad de procesamiento.

2. Entornos de desarrollo

- Weka: Software de escritorio para el aprendizaje automático y minería de datos diseñado en Java. Herramienta versátil que soporta multitud de tareas de procesamiento de datos y un amplio conjunto de algoritmos de aprendizaje de máquina, incluidos varios de RNAs. Incorpora la posibilidad de ser utilizado mediante llamadas a su librería.
- IBM SPSS Modeler: Se trata de un paquete software dedicado al modelado de datos que también proporciona diferentes algoritmos de aprendizaje como RNAs.
- Microsoft Azure Machine Learning: Microsoft plantea esta herramienta como un servicio de cloud computing. Accedemos a nuestro espacio de trabajo a través de un portal web donde tenemos alojadas las diferentes redes neuronales artificiales así como los datos que usamos. Esta característica nos ofrece la posibilidad de ejecutar los costos algoritmos de aprendizaje en servidores de la compañía especialmente diseñados para esto. Además, el abanico de modelos de redes neuronales que ofrece es muy amplio y la versión gratuita de uso no es nada limitada. Asimismo destacar también las múltiples funcionalidades extra que incorpora como la posibilidad de usar nuestros propios scripts en R o Python y un potente conjunto de herramientas de modelado de datos. Por último añadir también que la interfaz que ofrece es muy cómoda e intuitiva.

2.3.2 Microsoft Azure Machine Learning

Microsoft Azure Machine Learning es una herramienta que sigue en desarrollo, no obstante ya cuenta con un abanico amplio de algoritmos de aprendizaje que han clasificado en los 5 tipos que pueden ver en la siguiente imagen.

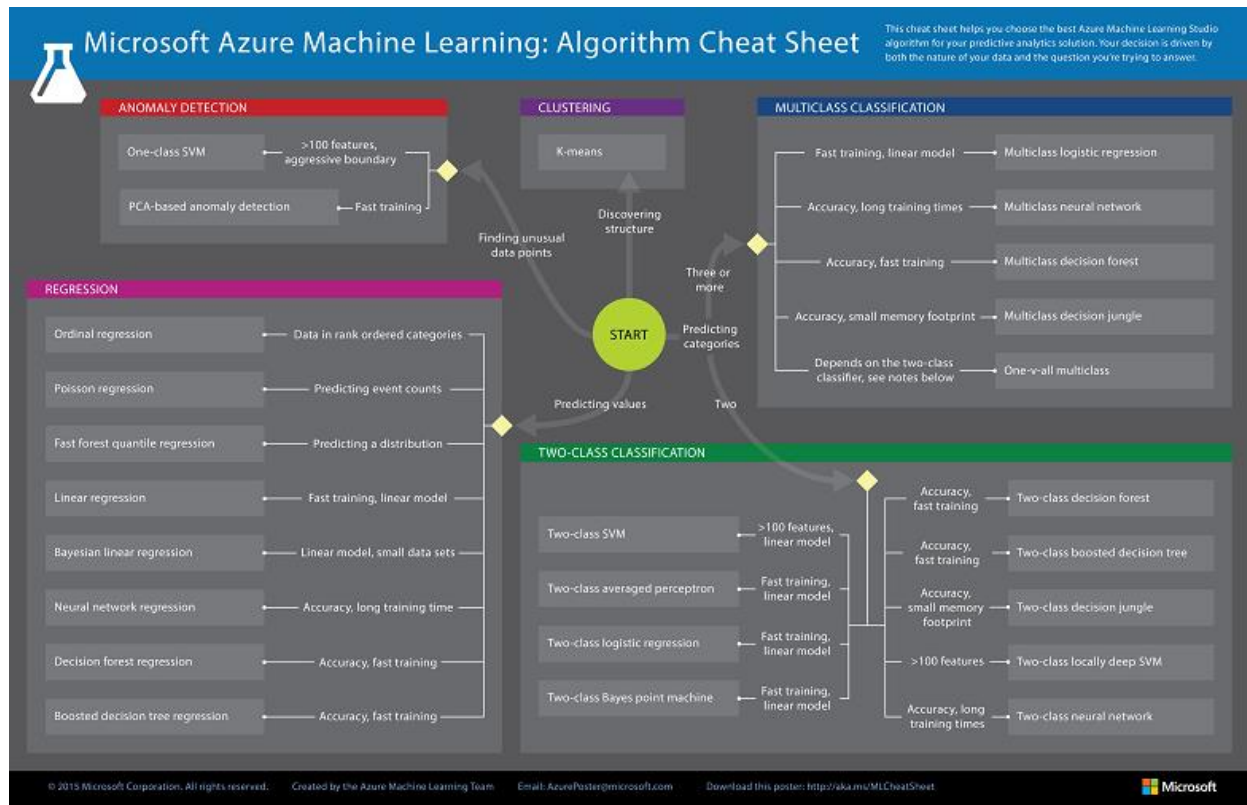


Figura 7.8. Algoritmos Azure.

Sin embargo, se analizarán tan solo los algoritmos supervisados de tipo "two-class classification" ya que es la categoría que se adapta a las necesidades del estudio, es decir, poder predecir si un sujeto tiene cáncer o no.

Modelos de redes neuronales de Azure ML

- Two-Class Averaged Perceptron: Se trata del algoritmo de clasificación más simple, es una versión temprana de lo que serían las redes neuronales. Las entradas se clasifican en varias salidas posibles en base a una función lineal y después de ajustar una serie de pesos se obtiene una salida.

- Two-Class SVM: Dado un conjunto de puntos en el espacio, los cuales cada uno de ellos pertenece a una de dos posibles categorías, un algoritmo basado en SVM construye un modelo capaz de predecir si un punto nuevo (cuya categoría desconocemos) pertenece a una categoría o a la otra.
- Two-Class Neural Network: Modelo de red neuronal clásica. Este algoritmo permite definir el número de capas de la red neuronal, número de neuronas por capa y como están definidas sus relaciones.
- Two-Class Locally Deep Support Vector Machine: Los algoritmos de aprendizaje profundo contrastan con los algoritmos de aprendizaje poco profundo por el número de transformaciones aplicadas a la señal mientras se propaga desde la capa de entrada a la capa de salida. Cada una de estas transformaciones incluyen parámetros que se pueden entrenar como pesos y umbrales. Esta es la fórmula de optimización que usa este algoritmo:

$$\min_{W, \theta, \theta'} P(W, \theta, \theta') = \frac{\lambda_W}{2} \text{Tr}(W^t W) + \frac{\lambda_\theta}{2} \text{Tr}(\theta^t \theta) + \frac{\lambda_{\theta'}}{2} \text{Tr}(\theta'^t \theta') + \sum_{i=1}^N L(y_i, \phi_L^t(x_i) W^t x_i)$$

Algoritmos basados en árboles de decisión

Los árboles de decisión son construcciones lógicas que sirven para categorizar y representar condiciones para la resolución de un problema.

Para completar este estudio resulta de utilidad hacer una comparación de los resultados obtenidos de las redes neuronales artificiales con resultados obtenidos de otro modelo inteligente, como son los árboles de decisión.

- Two-Class Boosted Decision Tree: Se trata de un algoritmo supervisado que es capaz de crear un número determinado de árboles que conectados secuencialmente son capaces de corregir los errores del inmediatamente anterior. De esta manera se toma una decisión de clasificación lo más acertada posible.

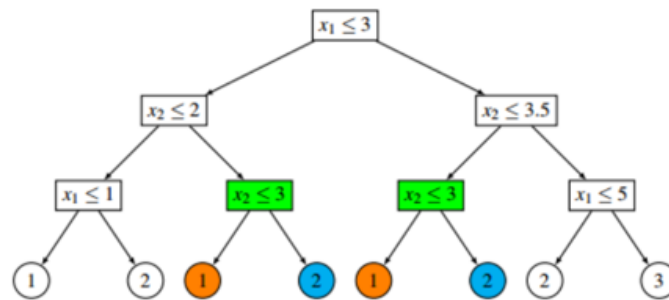


Figura 9.10. Árbol de decisión

- Two-Class Decision Forest: Este algoritmo crea múltiples árboles de decisión y después vota por la salida más “popular” entre las salidas de los diferentes árboles.
- Two-Class Decision Jungle: Se basa en la idea del algoritmo anterior pero en vez de usar árboles de decisión como unidades de voto usa un grafo acíclico dirigido (DAG por sus iniciales en inglés). El DAG es un sistema que trata de minimizar la gran cantidad de memoria generado por los árboles de decisión.

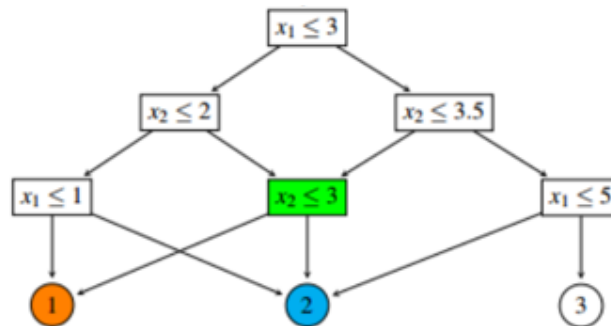


Figura 11.12. Grafo acíclico dirigido

Funcionamiento de la herramienta

Microsoft ofrece una intuitiva interfaz de módulos que a través de un sistema de “drag and drop” se podrá usar, ajustar e interconectar. Este es un diagrama del funcionamiento estándar de la herramienta:

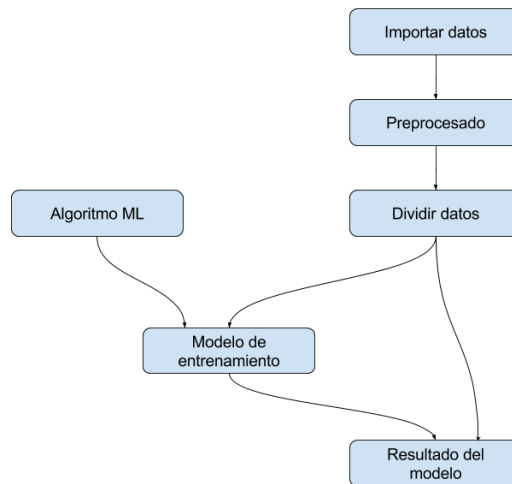


Figura 13.14. Grafo acíclico dirigido

Como se puede observar la herramienta comienza importando los datos, preprocesándolos y dividiéndolos en dos subconjuntos para el posterior entrenamiento del modelo. Por otro lado se ha de elegir cuál es el algoritmo que utilizaremos para nuestro proyecto y finalmente obtendremos un resultado.

A estas funcionalidades básicas se le pueden sumar un amplio abanico de posibilidades como pueden ser la representación de los resultados en gráficas o la utilización de lenguaje Python o R en cualquier parte del proceso de modelado de datos.

Análisis de los gráficos y datos de Azure

En la siguiente figura se muestra un ejemplo de cómo Azure ML muestra las estadísticas obtenidas de un modelo por defecto.

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
46	51	0.650	0.622	0.5	0.683
False Positive	True Negative	Recall	F1 Score		
28	101	0.474	0.538		
Positive Label	Negative Label				
1	0				

Figura 15.16. Ejemplo resultados Azure

Los siguientes son los índices usados por Azure ML para mostrar los resultados obtenidos en las pruebas:

- **Accuracy:**
Es simplemente el ratio de predicciones correctas.
$$\text{True Predictions} / \text{All Predictions}$$
- **Precision o precisión:**
Porcentaje de predicciones positivas correctas:
$$\text{True Positives} / (\text{True Positives} + \text{False Positives})$$
- **Recall o sensibilidad:**
Es la relación entre las predicciones positivas correctas y los falsos negativos.
$$\text{True Positives} / (\text{True Positive} + \text{False Negatives})$$
- **F1 Score:**
Es la media ponderada de la precisión y la sensibilidad.
$$2 * (\text{Recall} * \text{Precision}) / \text{Recall} + \text{Precision}$$
- **Threshold o umbral:**
El modelo de entrenamiento genera un número real entre 0 y 1 que después es usado para determinar si dicho número se traduce en un "positive" o en un "negative". Por defecto este valor se marca como 0,5 lo que significa que todos los índices recogidos mayor que este valor serán interpretados como positivos y los menores como negativos. No obstante Azure nos permite la posibilidad de modificarlo dinámicamente.
- **AUC:**
AUC es un valor entre 0 y 1 que representa el área debajo de la curva ROC. Cuanto mayor sea significa, por lo general, que el modelo tiene un porcentaje de acierto mayor.

Curva ROC:

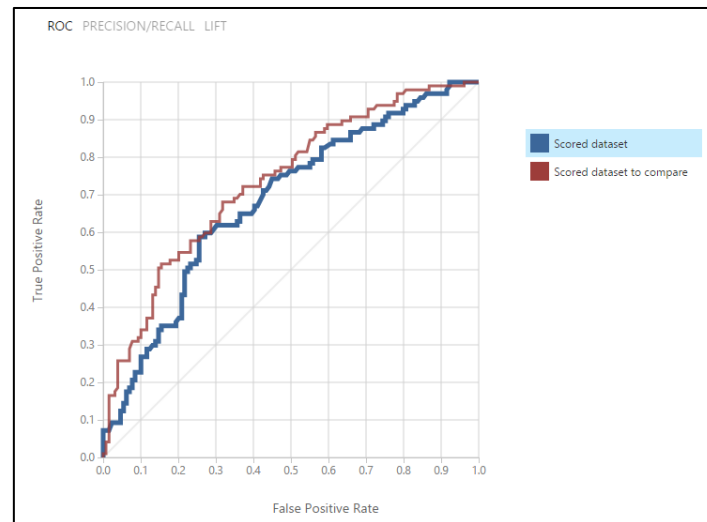


Figura 17.18. Ejemplo grafica ROC Azure

En esta gráfica dos modelos de entrenamientos son comparados, la curva azul representa el modelos de la izquierda y la roja el de la derecha. Cada posición en la línea representa el resultado del modelo para cada threshold o umbral seleccionado. La esquina inferior izquierda (0,0) representa un umbral de 1 mientras que la esquina superior derecha (1,1) representa un umbral de 0. Las dimensiones donde se mueve son un eje vertical que representa los "true positive" y un eje horizontal que representa los "false positive".

Al cambiar el umbral, disminuye la frecuencia de un tipo de error a expensas de aumentar el otro tipo. A la hora de comparar dos modelos nos interesa saber cuál curva se acerca más a la esquina superior izquierda ya que este modelo proporciona una mejor predicción.

Precision/recall:

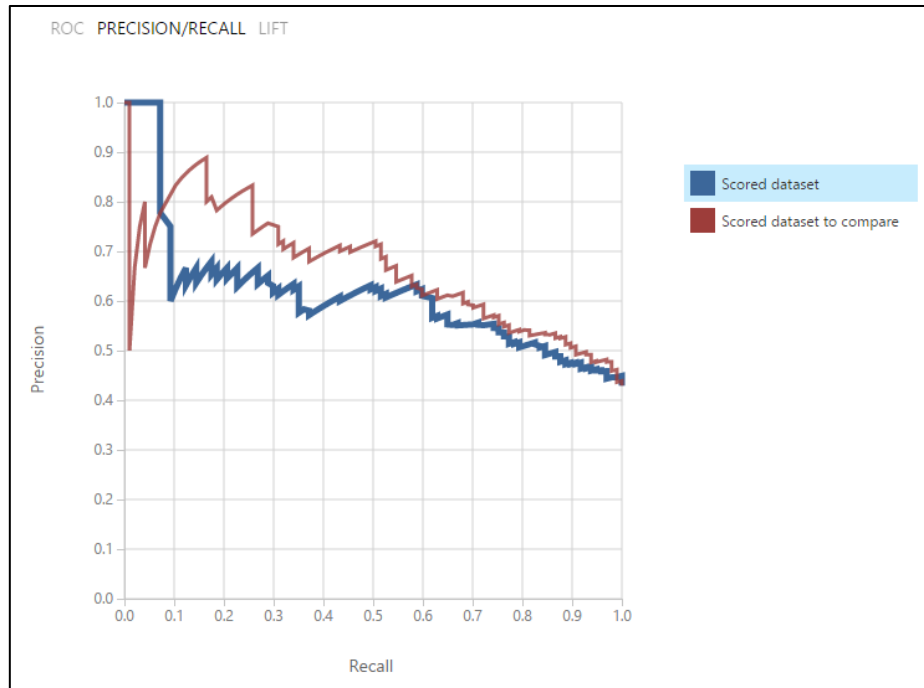


Figura 19.20. Ejemplo grafica P/R Azure

En esta gráfica ambas curvas se mueven bajo los ejes vertical "precisión" y horizontal "recall". Cada punto representa el porcentaje de acierto para datos positivos y negativos cuando variamos el umbral. La curva que más se acerque a la esquina superior derecha será la que ostente mejores resultados ya que se traduce en el modelo con más aciertos.

Capítulo 3.

Estudio

3.1 Principios de los estudios

Los estudios se han realizado independientemente para cada tipo de cáncer, no obstante todos siguen un patrón, el cual será explicado en esta sección.

En la siguiente imagen se muestra de forma gráfica cómo se han desarrollado las pruebas con la interfaz de Azure ML. Este sistema es análogo para cada tipo de cáncer y tan solo cambian los datos introducidos y los resultados obtenidos.

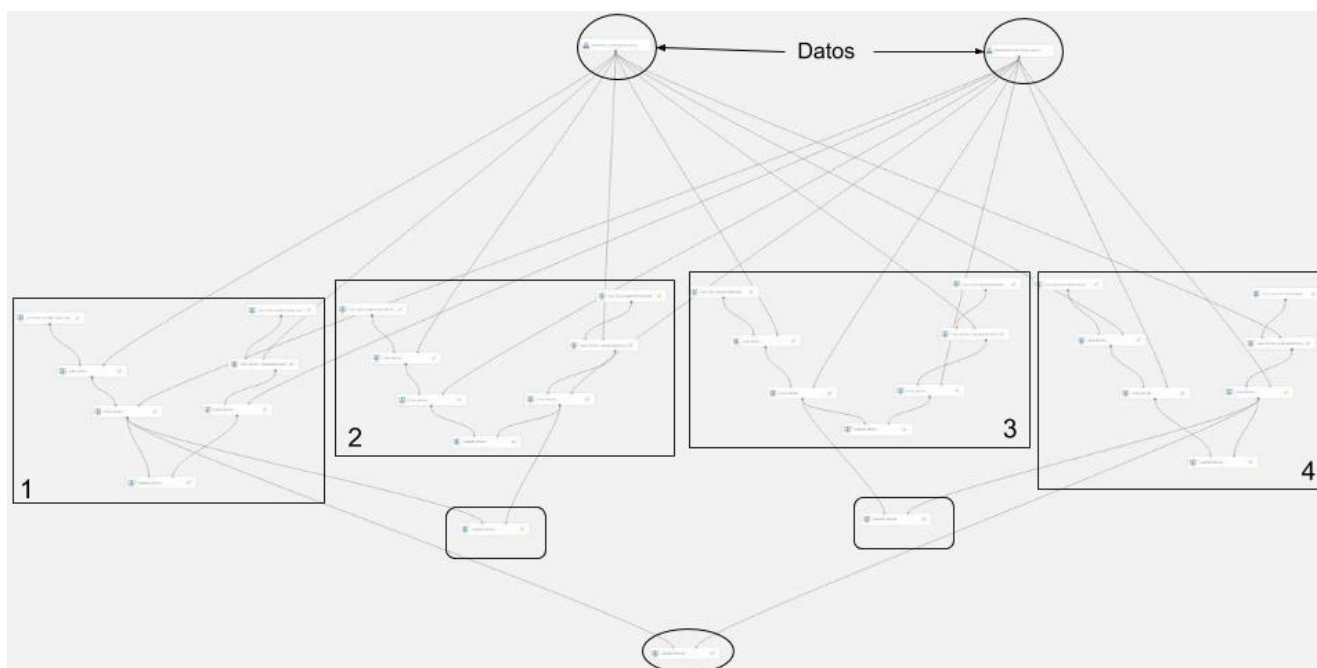


Figura 1.2. Esquema de las pruebas

Los dos módulos superiores son las entradas de datos, representan al subconjunto de datos de entrenamiento (izquierda) y de validación (derecha).

En los siguientes cuatro grandes bloques de módulos se encuentran los modelos inteligentes definidos. Cada bloque, a su vez, tiene una estructura similar al resto:

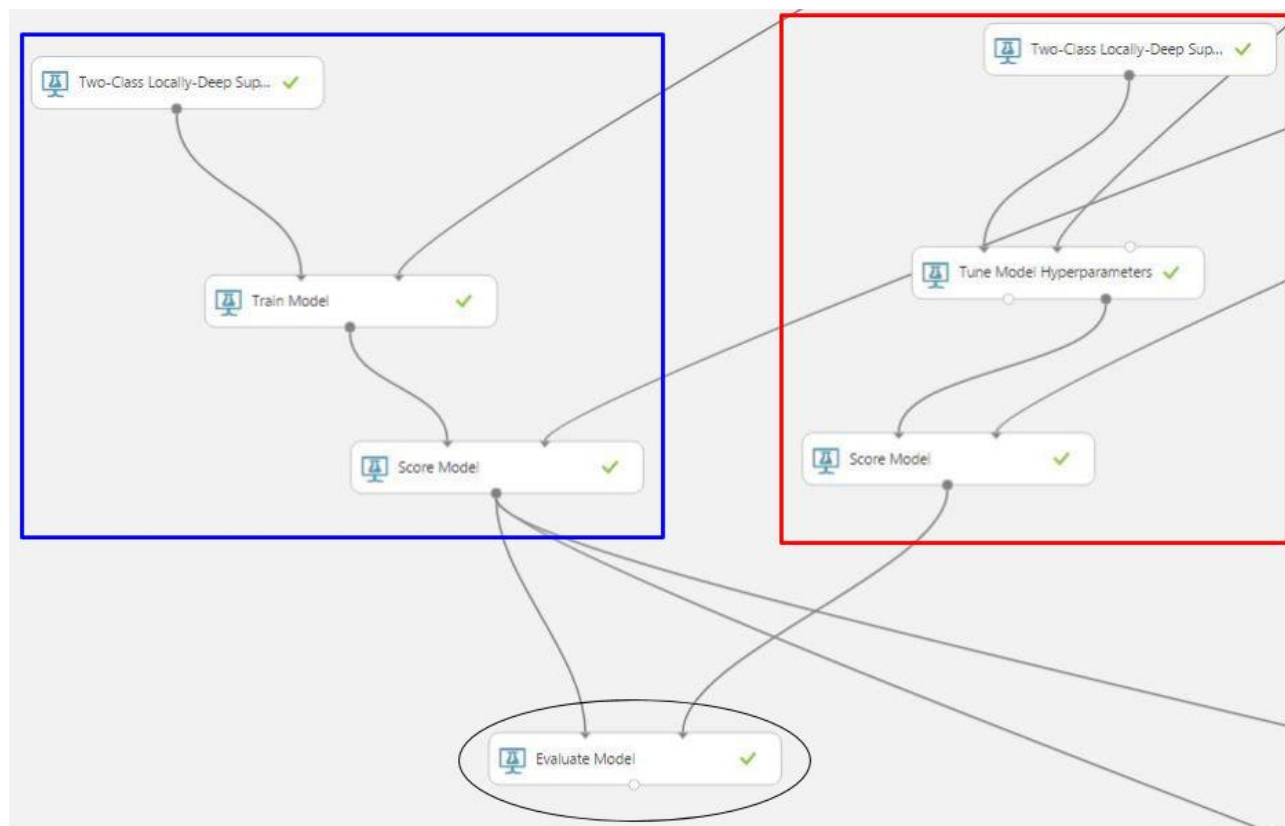


Figura 3.4. Grafico bloque prueba.

En cada bloque se sigue una estructura similar a la mostrada en la Figura 3.2, con la diferencia de que en cada bloque se cambia el modelo de red neuronal definida, en este ejemplo es un modelo "Two-Class Locally-Deep Support Vector Machine".

Como pueden observar se han dividido en dos cada bloque, esto se debe a que dentro de cada bloque se han creado dos redes neuronales distintas que después son comparadas en el módulo que se observa debajo, marcado con una elipse. Los colores definidos ayudarán después a la hora de comparar los resultados de tal forma que siempre la red azul será la de la izquierda y la roja la de la derecha.

La red azul en cada caso es el modelo definido con los parámetros por defecto de Azure ML y la red roja es una red obtenida de la comparación de veinte modelos definidos con parámetros aleatorios en búsqueda del mayor porcentaje de aciertos. Una vez entrenadas estas dos redes se comparan en el módulo "Evaluate Model".

Esta primera comparación entre modelos se realiza dentro de cada bloque para obtener el mejor modelo de red. Como se ha mencionado anteriormente en cada bloque se define un modelo:

- Bloque 1: Two-Class Locally-Deep Support Vector Machine.
- Bloque 2: Two-Class Support Vector Machine.
- Bloque 3: Two-Class Neural Network.
- Bloque 4: Two-Class Decision Jungle.

Una vez obtenida la configuración de modelo más óptimo para cada tipo de modelo los comparamos entre sí para obtener qué modelo es el que obtiene los mejores resultados. Se compara el bloque 1 con el 2 y el 3 con el 4. Los mejores resultados de estas dos comparaciones son evaluados en una nueva y última comparación de la que resultará ganador un solo modelo, el que será elegido como mejor.

Para decidir qué RNA obtiene mejores resultados se analiza el global de índices obtenidos en las pruebas, no obstante, nos centraremos en los valores de Accuracy y Recall.

3.2 Resultados para cada tipo de cáncer analizado

3.2.1 Melanoma

Bloque 1: Búsqueda de la configuración del modelo "Two-Class Locally-Deep Support Vector Machine" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Tree Depth	Lambda W	Lambda Theta	Lambda Theta Prime	Sigma	Number of Iterations
Azul:	3	0.1	0.01	0.01	1	15000
Roja:	6	0.057241	0.064953	0.05212	0.970426	10388

Tabla 5.1. Configuración redes melanoma bloque 1.

Resultados obtenidos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.739	0.670	0.707	0.688	0.804
Roja:	0.721	0.639	0.689	0.663	0.780

Tabla 6.2. Resultados redes melanoma bloque 1.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:

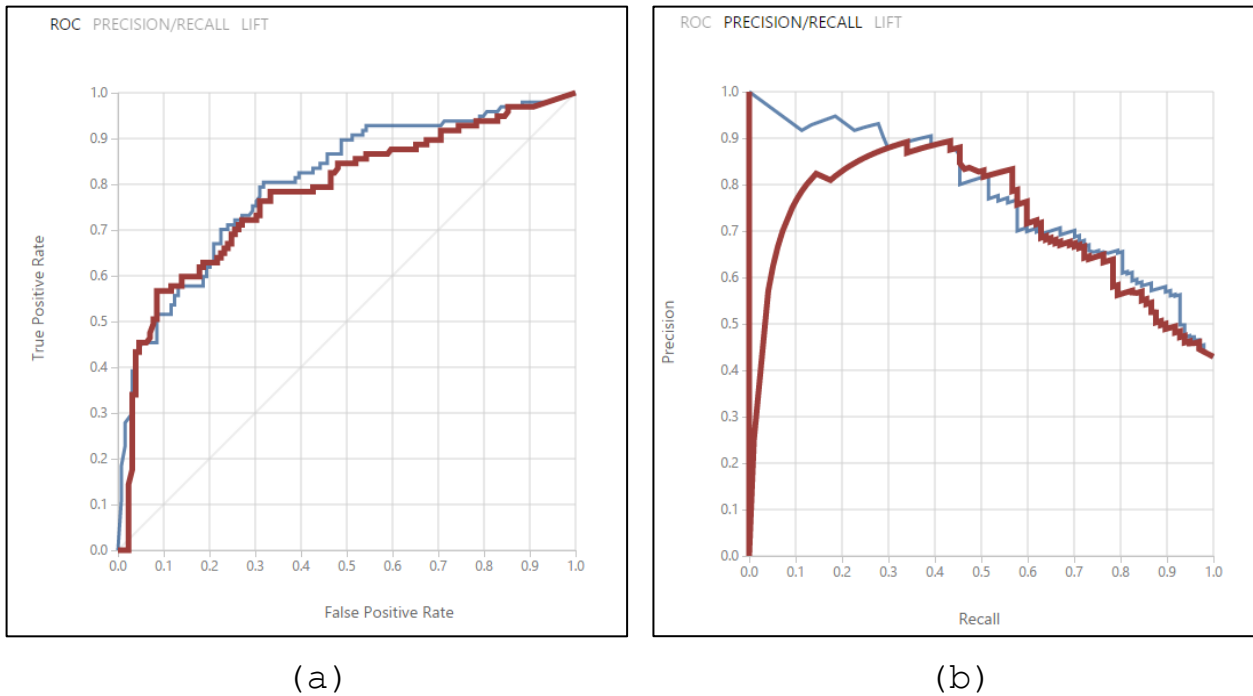


Figura 7.8. Melanoma bloque 1. (a)ROC. (b)P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Azul por ofrecer mejores valores tanto de accuracy como de recall.

Bloque 2: Búsqueda de la configuración del modelo "Two-Class Support Vector Machine" que ofrezca un mejor valor de Accuracy.

Red Azul configurada con los siguientes parámetros:

Configuración de las redes:

Red	Number of iterations	Lambda
Azul:	1	0.001
Roja:	17	0.000895

Tabla 9.10. Configuración redes melanoma bloque 2.

Resultados obtenidos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.664	0.557	0.621	0.587	0.751
Roja:	0.690	0.629	0.642	0.635	0.751

Tabla 11.12. Resultados redes melanoma bloque 2.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:

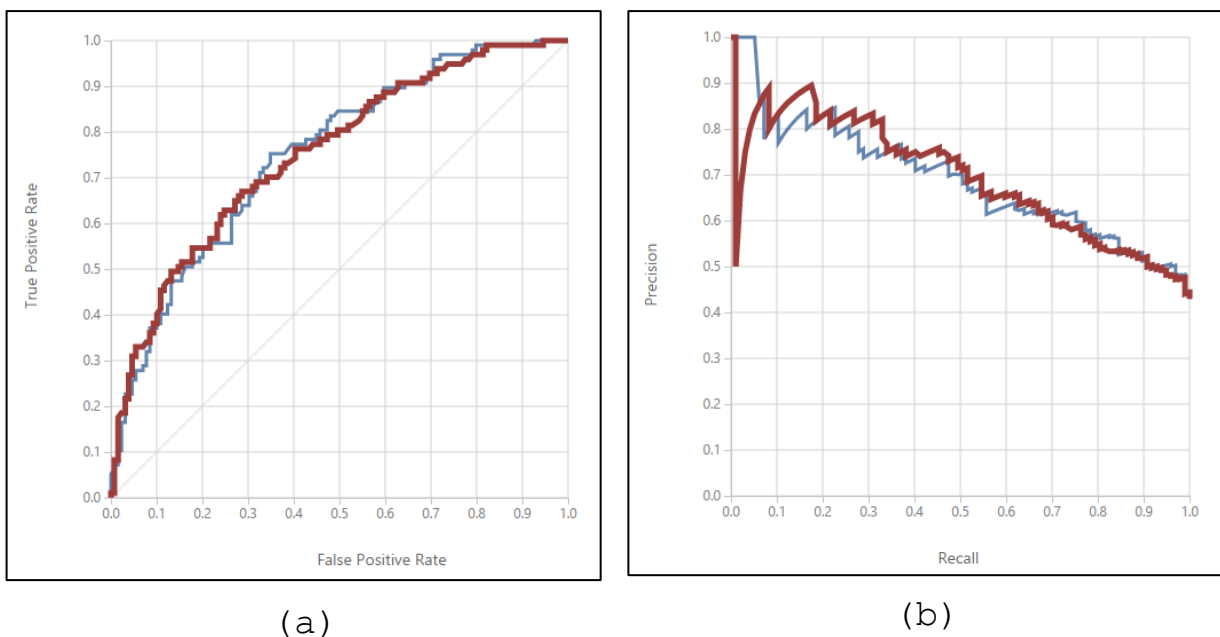


Figura 13.14. Melanoma bloque 2. (a)ROC. (b) P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Roja por ofrecer mejores valores tanto de accuracy como de recall

Bloque 3: Búsqueda de la configuración del modelo "Two-Class Neural Network" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Learning rate	Number of iterations
Azul:	0.1	100
Roja:	0.016576	54

Tabla 15.16. Configuración redes melanoma bloque 3.

Resultados obtenidos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.726	0.619	0.706	0.659	0.791
Roja:	0.681	0.588	0.640	0.613	0.739

Tabla 17.18. Resultados redes melanoma bloque 3.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:

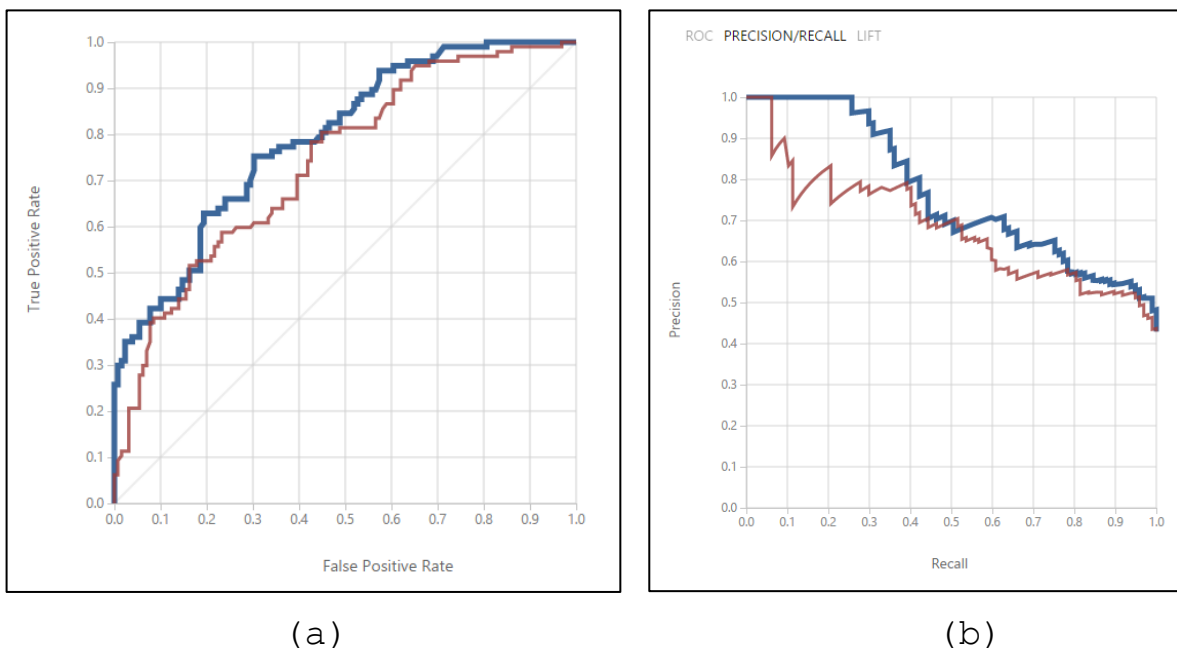


Figura 19.20. Melanoma bloque 3. (a) ROC. (b) P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Azul por ofrecer mejores valores tanto de accuracy como de recall.

Bloque 4: Búsqueda de la configuración del modelo "Two-Class Decision Jungle" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Optimization steps	Maximum width of the decision DAGs	Maximum depth of the decision DAGs	Number of decision DAGs
Azul:	2048	128	32	8
Roja:	2235	150	104	30

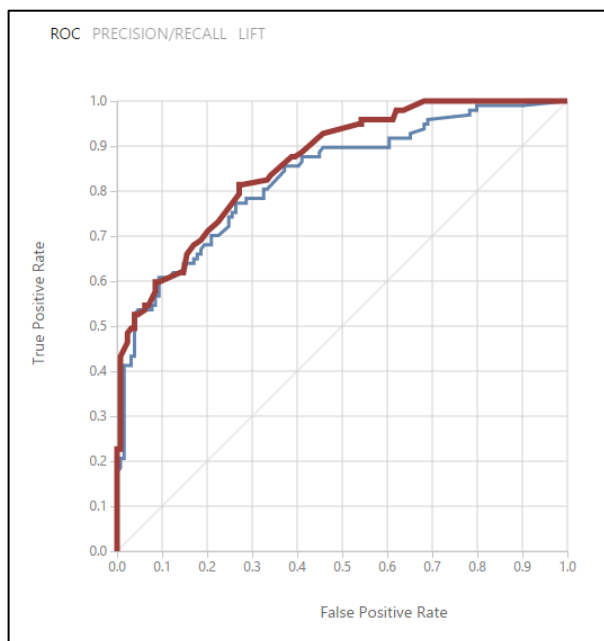
Tabla 21.22. Configuración redes melanoma bloque 4.

Resultados obtenidos:

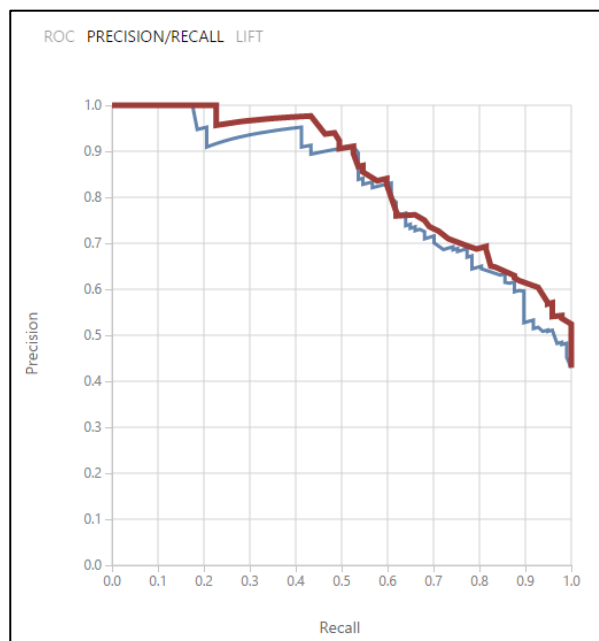
Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.739	0.722	0.686	0.704	0.832
Roja:	0.761	0.691	0.736	0.713	0.862

Tabla 23.24. Resultados redes melanoma bloque 3.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:



(a)



(b)

Figura 25.26. Melanoma bloque 4. (a) ROC. (b) P/R.

Resultado de la comparación:

En esta comparación la red Azul tiene un Recall mayor a la red Roja en 0,032 puntos, no obstante en el parámetro de Accuracy gana ventaja la red Roja por 0,022 puntos. Adicionalmente el parámetro AUC de la curva roja es mejor por lo tanto y teniendo en cuenta que la diferencia de Recall no es significativa escogemos a la red roja como vencedora.

Una vez realizada la búsqueda de las 4 mejores configuraciones para cada modelo seleccionado para este análisis procederemos a comparar dichas redes unas con otras.

Resultados de las mejores configuraciones de modelos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Bloque 1:	0.739	0.670	0.707	0.688	0.804
Bloque 2:	0.690	0.629	0.642	0.629	0.751
Bloque 3:	0.726	0.619	0.706	0.659	0.791
Bloque 4:	0.761	0.691	0.736	0.713	0.862

Tabla 27.28. Comparación final modelos melanoma.

Estos resultados se traducen en que el sistema que ha obtenido mejores resultados de validación es el procedente del bloque 4, un modelo "Two-Class Decision Jungle" con estos resultados:

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
67	30	0.761	0.736	0.5	0.862
False Positive	True Negative	Recall	F1 Score		
24	105	0.691	0.713		
Positive Label	Negative Label				
1	0				

Figura 29.30. Resultados finales melanoma

3.2.2 Cáncer de pulmón

Bloque 1: Búsqueda de la configuración del modelo "Two-Class Locally-Deep Support Vector Machine" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Tree Depth	Lambda W	Lambda Theta	Lambda Theta Prime	Sigma	Number of Iterations
Azul:	3	0.1	0.01	0.01	1	15000
Roja:	6	0.057241	0.064953	0.05212	0.970426	10388

Tabla 31.32. Configuración redes pulmón bloque 1.

Resultados obtenidos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.883	0.813	0.891	0.850	0.916
Roja:	0.851	0.801	0.828	0.814	0.946

Tabla 33.34. Resultados redes pulmón bloque 1.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:

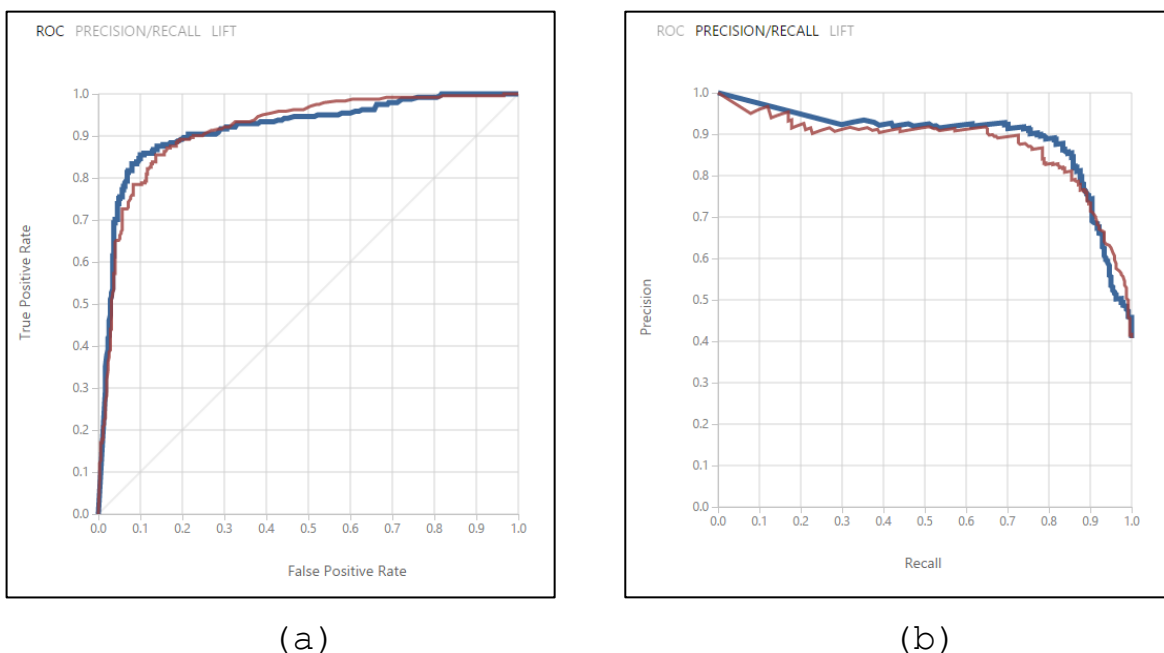


Figura 35.36. Pulmón bloque 1. (a)ROC. (b) P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Azul por ofrecer mejores valores tanto de accuracy como de recall.

Bloque 2: Búsqueda de la configuración del modelo "Two-Class Support Vector Machine" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Number of iterations	Lambda
Azul:	1	0.001
Roja:	17	0.000895

Tabla 37.38. Configuración redes pulmón bloque 2.

Resultados obtenidos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.770	0.689	0.731	0.7709	0.824
Roja:	0.787	0.710	0.753	0.731	0.836

Tabla 39.40. Resultados redes pulmón bloque 2.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:

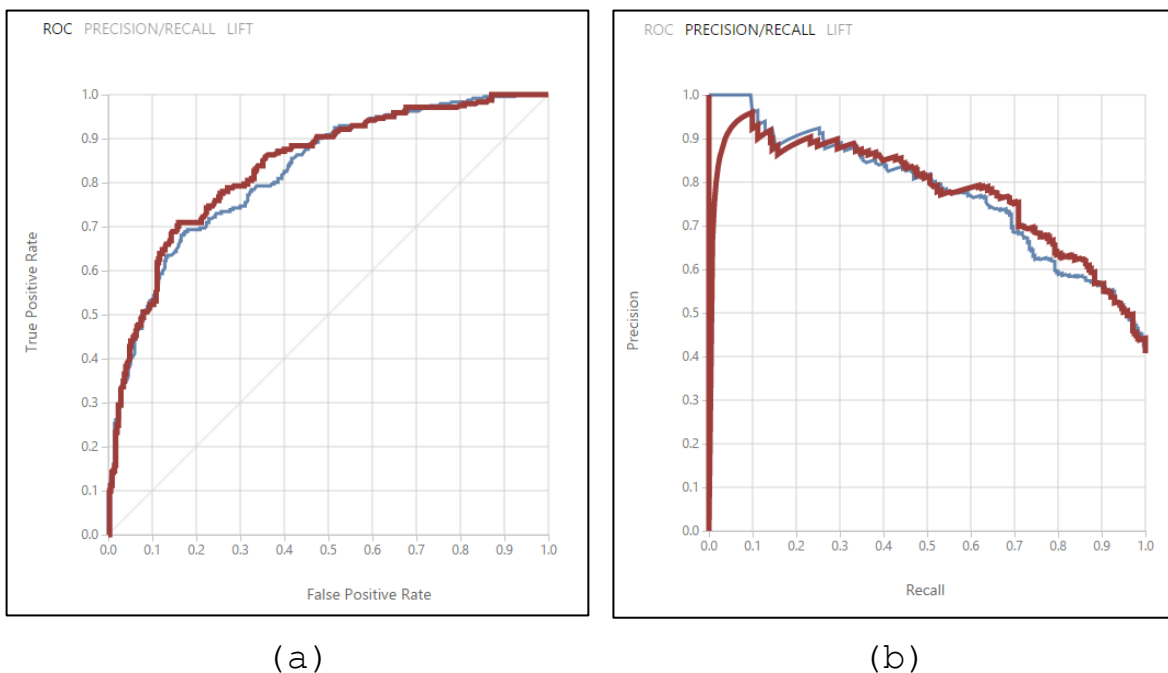


Figura 41.42. Pulmón bloque 2. (a)ROC. (b)P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Roja por ofrecer mejores valores tanto de accuracy como de recall.

Bloque 3: Búsqueda de la configuración del modelo "Two-Class Neural Network" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Learning rate	Number of iterations
Azul:	0.1	100
Roja:	0.037861	129

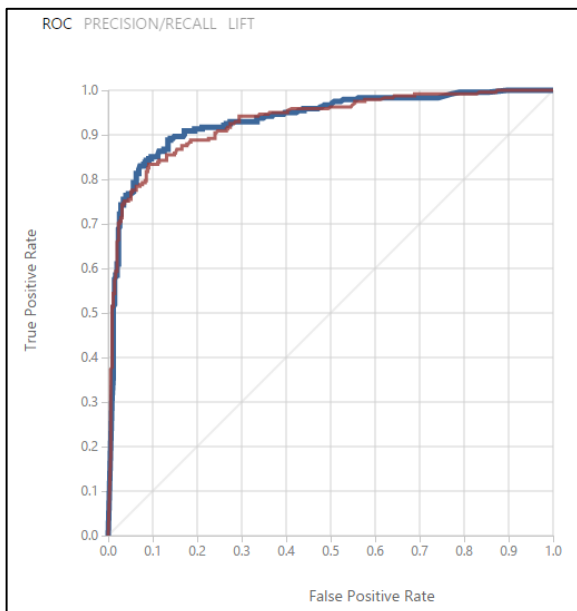
Tabla 43.44. Configuración redes pulmón bloque 3.

Resultados obtenidos:

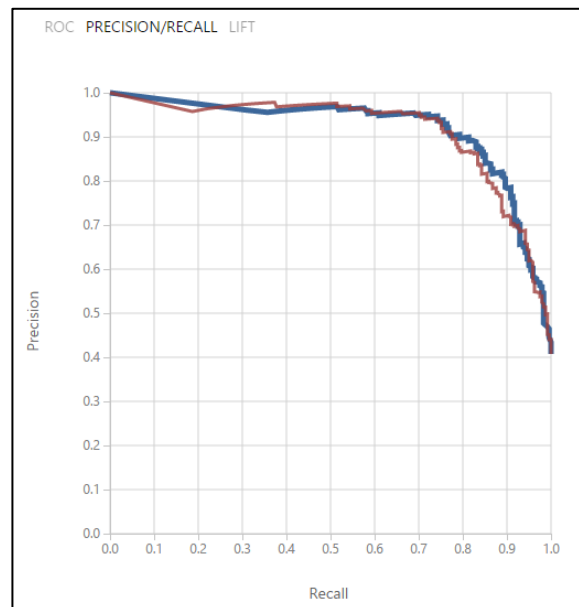
Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.883	0.842	0.868	0.855	0.934
Roja:	0.873	0.817	0.864	0.840	0.931

Tabla 45.46. Resultados redes pulmón bloque 3.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:



(a)



(b)

Figura 47.48. Pulmón bloque 3. (a) ROC. (b) P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Azul por ofrecer mejores valores tanto de accuracy como de recall.

Bloque 4: Búsqueda de la configuración del modelo "Two-Class Decision Jungle" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Optimization steps	Maximum width of the decision DAGs	Maximum depth of the decision DAGs	Number of decision DAGs
Azul:	2048	128	32	8
Roja:	2235	150	104	30

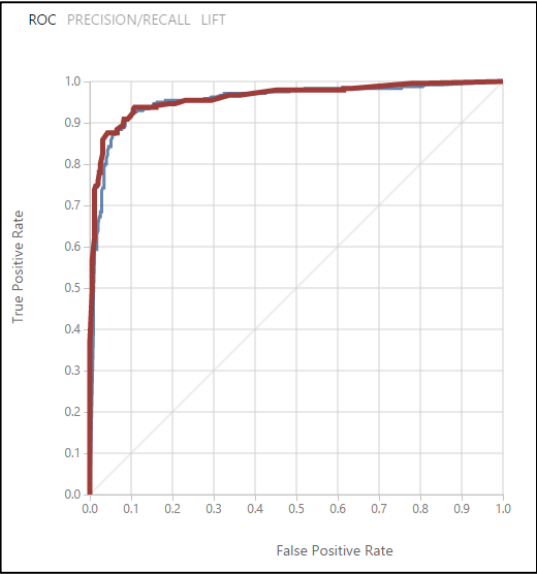
Tabla 49.50. Configuración redes pulmón bloque 4.

Resultados obtenidos:

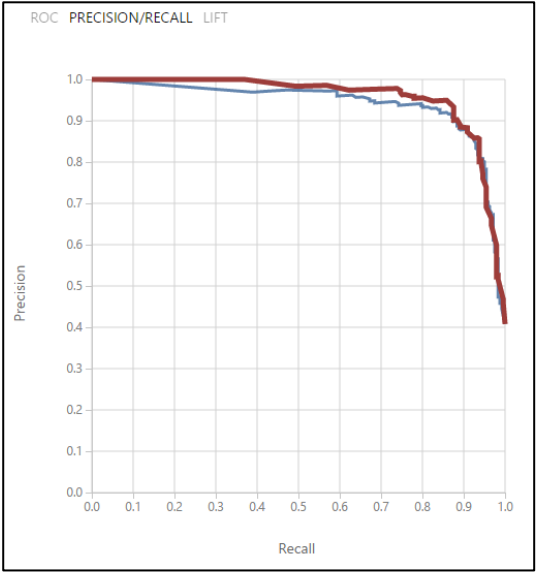
Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.907	0.834	0.931	0.880	0.952
Roja:	0.924	0.876	0.934	0.904	0.961

Tabla 51.52. Resultados redes pulmón bloque 4.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:



(a)



(b)

Figura 53.54.Pulmón bloque 4. (a)ROC. (b) P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Roja por ofrecer mejores valores tanto de accuracy como de recall.

Una vez realizada la búsqueda de las 4 mejores configuraciones para cada modelo seleccionado para este análisis procederemos a comparar dichas redes unas con otras.

Resultados de las mejores configuraciones de modelos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Bloque 1:	0.883	0.813	0.891	0.850	0.916
Bloque 2:	0.787	0.710	0.753	0.753	0.836
Bloque 3:	0.883	0.842	0.868	0.855	0.934
Bloque 4:	0.924	0.876	0.934	0.904	0.961

Tabla 55.56. Comparación final modelos pulmón.

Estos resultados se traducen en que el sistema que ha obtenido mejores resultados de validación es el procedente del bloque 4, un modelo "Two-Class Decision Jungle" con estos resultados:

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
211	30	0.924	0.934	0.5	0.961
False Positive	True Negative	Recall	F1 Score		
15	335	0.876	0.904		
Positive Label	Negative Label				
1	0				

Figura 57.58. Resultado final pulmón

3.2.3 Cáncer de mama

Bloque 1: Búsqueda de la configuración del modelo "Two-Class Locally-Deep Support Vector Machine" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Tree Depth	Lambda W	Lambda Theta	Lambda Theta Prime	Sigma	Number of Iterations
Azul:	3	0.1	0.01	0.01	1	15000
Roja:	6	0.057241	0.024971	0.05212	0.970426	10388

Tabla 59.60. Configuración redes mama bloque 1.

Resultados obtenidos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.767	0.777	0.785	0.781	0.849
Roja:	0.773	0.787	0.787	0.787	0.847

Tabla 61.62. Resultados redes mama bloque 1.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:

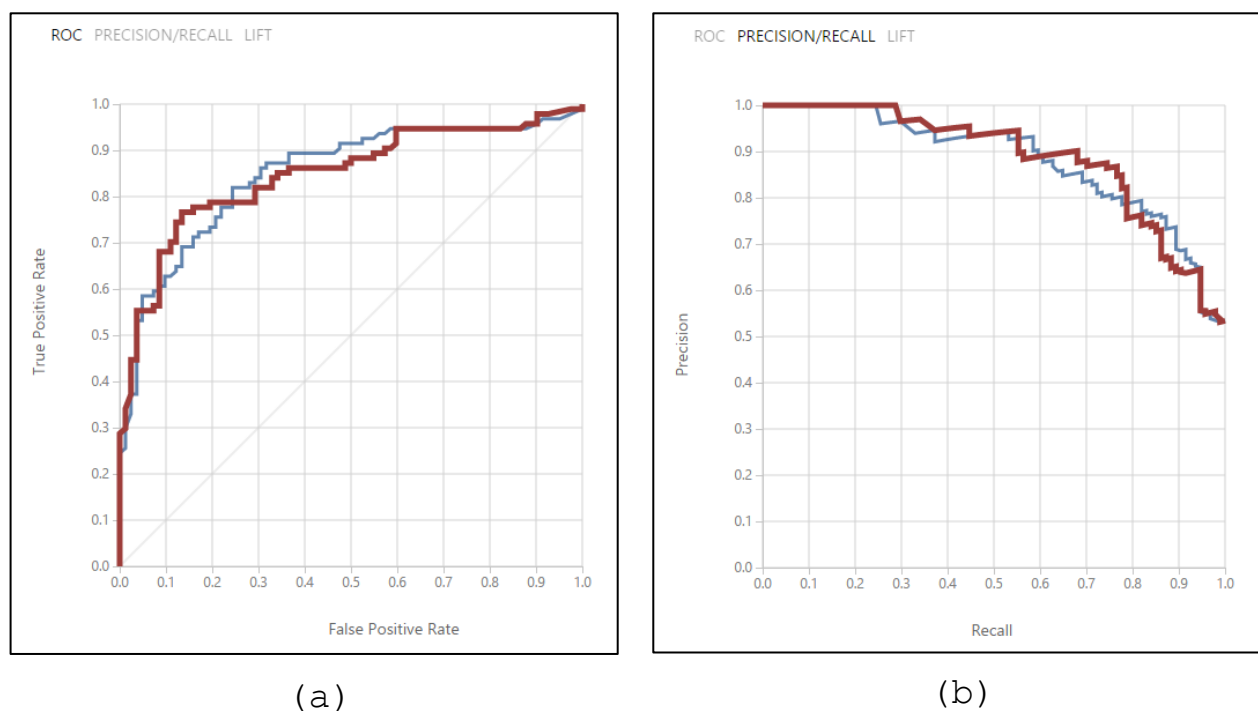


Figura 63.64. Mama bloque 1. (a)ROC. (b)P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Roja por ofrecer mejores valores tanto de accuracy como de recall.

Bloque 2: Búsqueda de la configuración del modelo "Two-Class Support Vector Machine" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Number of iterations	Lambda
Azul:	1	0.001
Roja:	6	0.007497

Tabla 65.66. Configuración redes mama bloque 2.

Resultados obtenidos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.756	0.787	0.763	0.775	0.838
Roja:	0.733	0.809	0.724	0.764	0.830

Tabla 67.68. Resultados redes mama bloque 2.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:

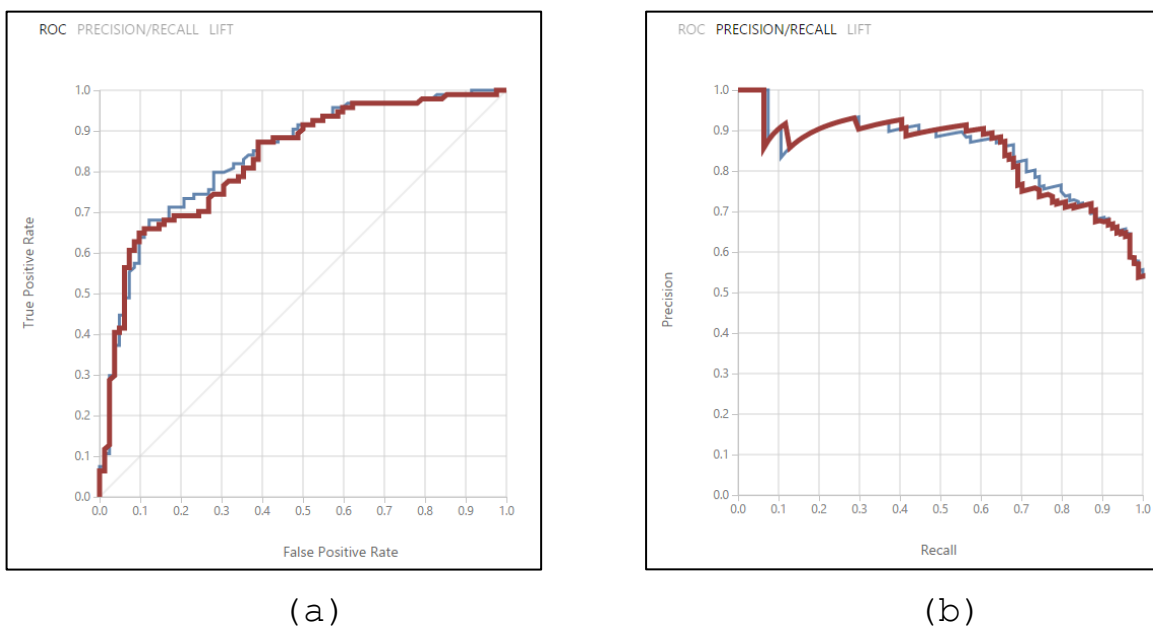


Figura 69.70. Mama bloque 2. (a)ROC. (b)P/R.

Resultado de la comparación:

En esta comparación la red Roja tiene un Recall mayor a la red Azul en 0,022 puntos, no obstante en el parámetro de Accuracy gana ventaja la red Azul por 0,023 puntos. Adicionalmente el parámetro AUC de la curva azul es mejor por lo tanto y teniendo en cuenta que la diferencia de Recall no es significativa escogemos a la red azul como vencedora.

Bloque 3: Búsqueda de la configuración del modelo "Two-Class Neural Network" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Learning rate	Number of iterations
Azul:	0.1	100
Roja:	0.016576	54

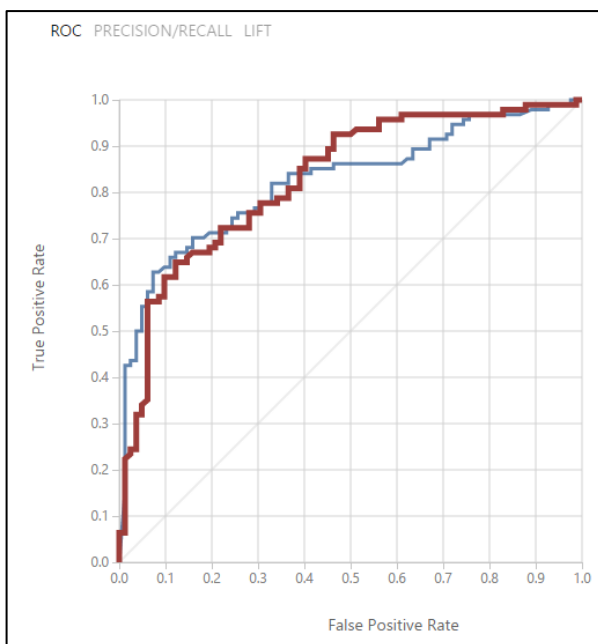
Tabla 71.72. Configuración redes mama bloque 3.

Resultados obtenidos:

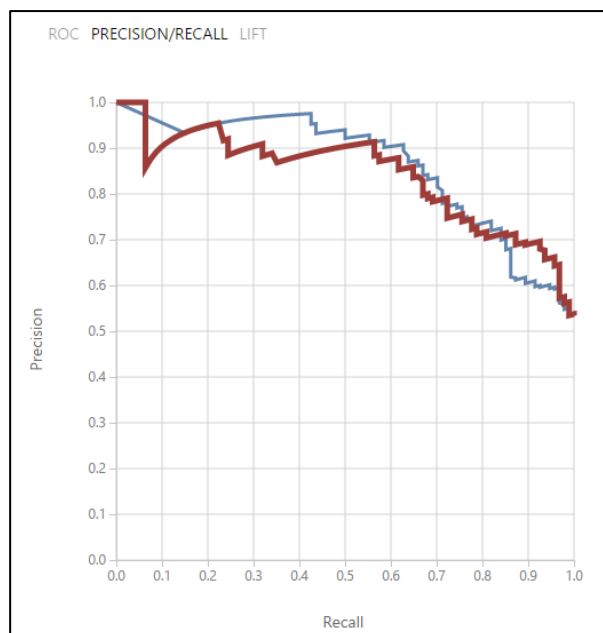
Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.744	0.809	0.738	0.772	0.822
Roja:	0.727	0.830	0.709	0.765	0.828

Tabla 73.74. Resultados redes mama bloque 3.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:



(a)



(b)

Figura 75.76. Mama bloque 3. (a) ROC. (b) P/R.

Resultado de la comparación:

En esta comparación la red Roja tiene un Recall mayor a la red Azul en 0,021 puntos, no obstante en el parámetro de Accuracy gana ventaja la red Azul por 0,017 puntos. Adicionalmente el parámetro AUC de la curva roja es mejor por lo tanto y teniendo en cuenta que la diferencia de Accuracy no es significativa escogemos a la red roja como vencedora.

Bloque 4: Búsqueda de la configuración del modelo "Two-Class Decision Jungle" que ofrezca un mejor valor de Accuracy.

Configuración de las redes:

Red	Optimization steps	Maximum width of the decision DAGs	Maximum depth of the decision DAGs	Number of decision DAGs
Azul:	2048	128	32	8
Roja:	16384	8	128	32

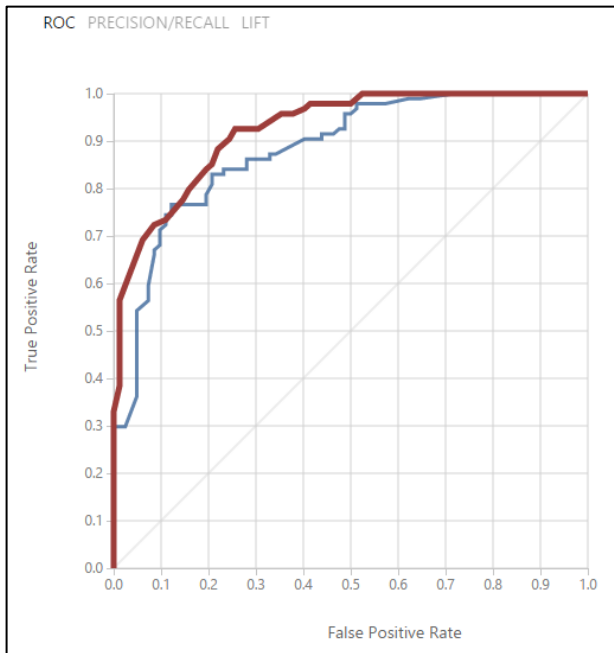
Tabla 77.78. Configuración redes mama bloque 4.

Resultados obtenidos:

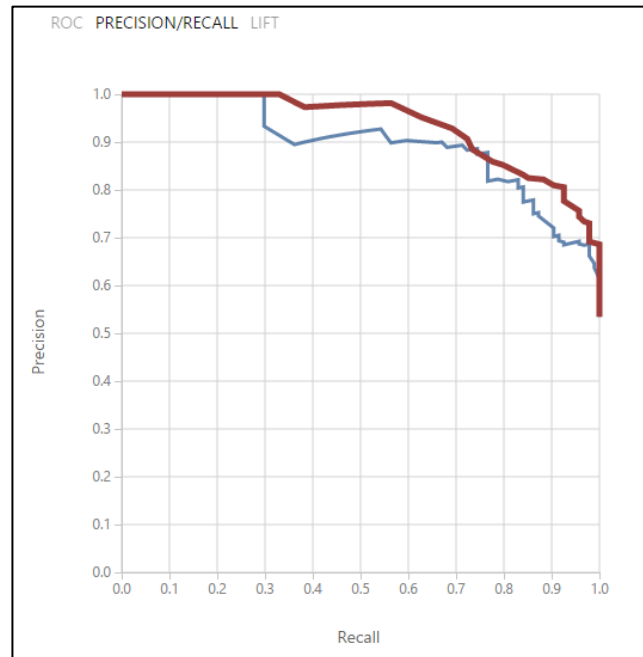
Red	Accuracy	Recall	Precision	F1 Score	AUC
Azul:	0.790	0.840	0.782	0.810	0.882
Roja:	0.835	0.904	0.810	0.854	0.924

Tabla 79.80. Resultados redes mama bloque 4.

Esta es la comparación grafica de los resultados obtenidos por cada red neuronal:



(a)



(b)

Figura 81.82. Mama bloque 4. (a) ROC. (b) P/R.

Resultado de la comparación:

En esta comparación sale vencedora la red Roja por ofrecer mejores valores tanto de accuracy como de recall.

Una vez realizada la búsqueda de las 4 mejores configuraciones para cada modelo seleccionado para este análisis procederemos a comparar dichas redes unas con otras.

Resultados de las mejores configuraciones de modelos:

Red	Accuracy	Recall	Precision	F1 Score	AUC
Bloque 1:	0.773	0.787	0.787	0.787	0.847
Bloque 2:	0.756	0.787	0.763	0.775	0.838
Bloque 3:	0.727	0.830	0.709	0.765	0.828
Bloque 4:	0.790	0.840	0.782	0.810	0.882

Tabla 83.84. Comparación final modelos mama.

Estos resultados se traducen en que el sistema que ha obtenido mejores resultados de validación es el procedente del bloque 4, un modelo "Two-Class Decision Jungle" con estos resultados:

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
79	15	0.790	0.782	0.5	0.882
False Positive	True Negative	Recall	F1 Score		
22	60	0.840	0.810		
Positive Label	Negative Label				
1	0				

Figura 85.86. Resultado final mama

3.3 Estudio de los sistemas bajo el subconjunto de test

Los modelos obtenidos en los apartados anteriores serán puestos a prueba en esta sección bajo condiciones reales usando para ello el subconjunto de datos de test. Este subconjunto no ha sido usado antes para entrenar el modelo por lo que son datos que no conoce y por lo tanto nos servirán para obtener unos porcentajes de aciertos más próximos a la realidad.

Para este estudio utilizaremos el siguiente esquema de módulos de Azure:

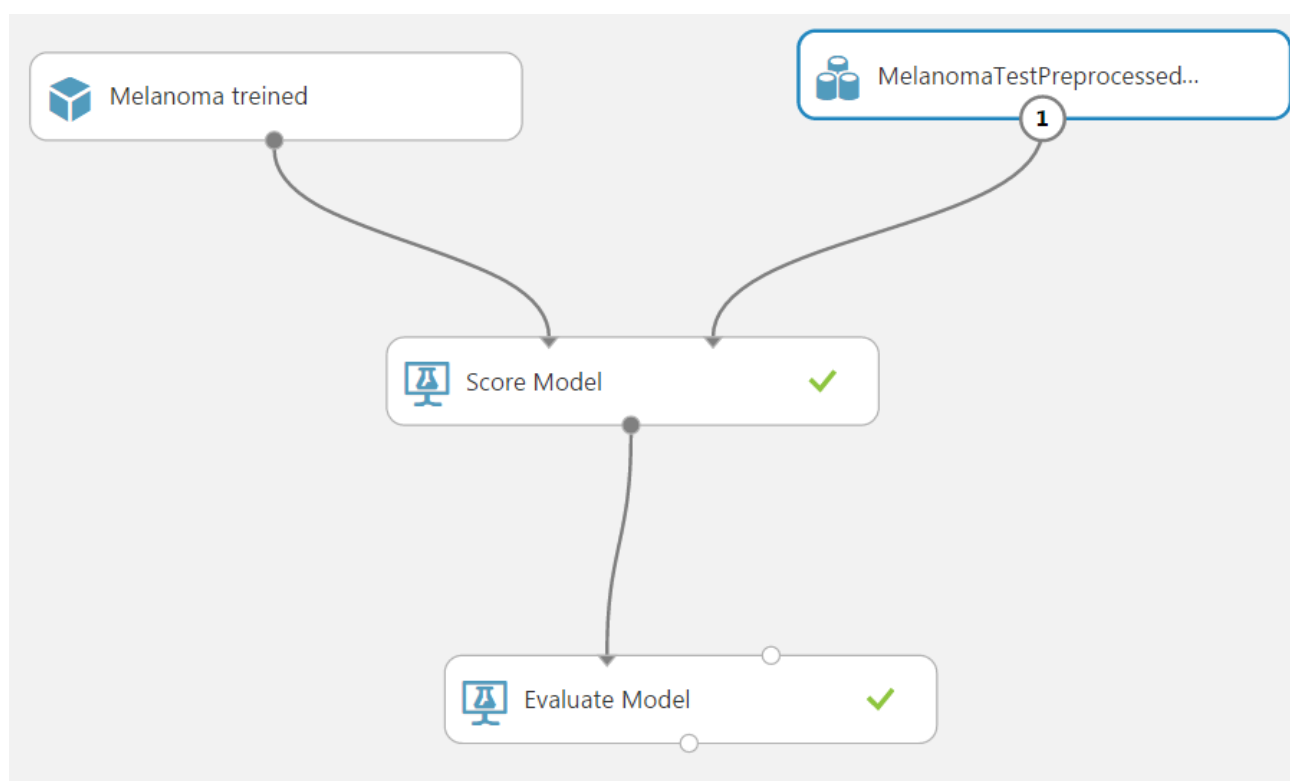
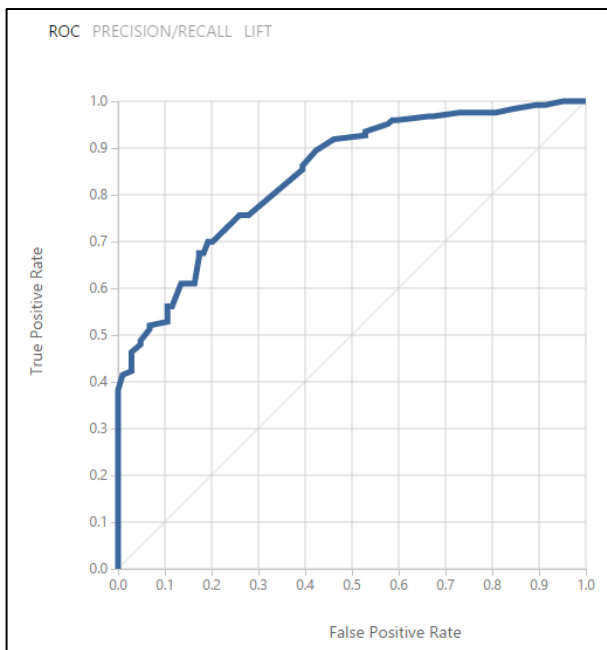
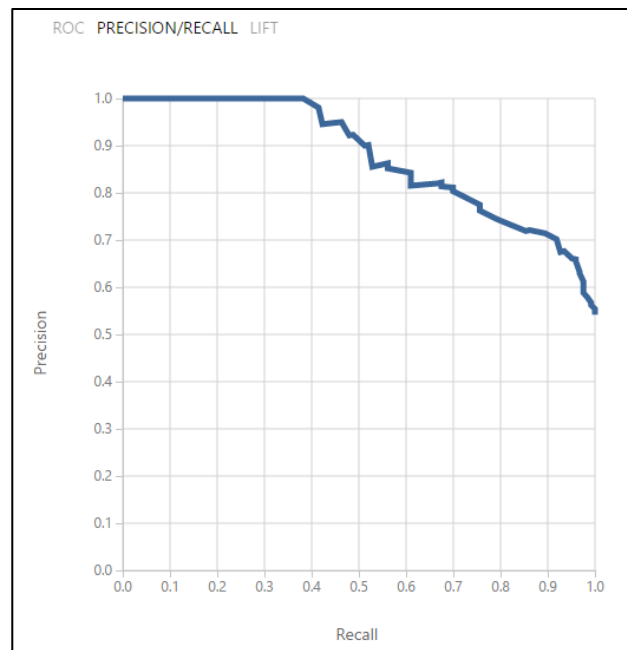


Figura 87.88. Esquema de módulos subconjunto de test.

3.3.1 Melanoma



(a)



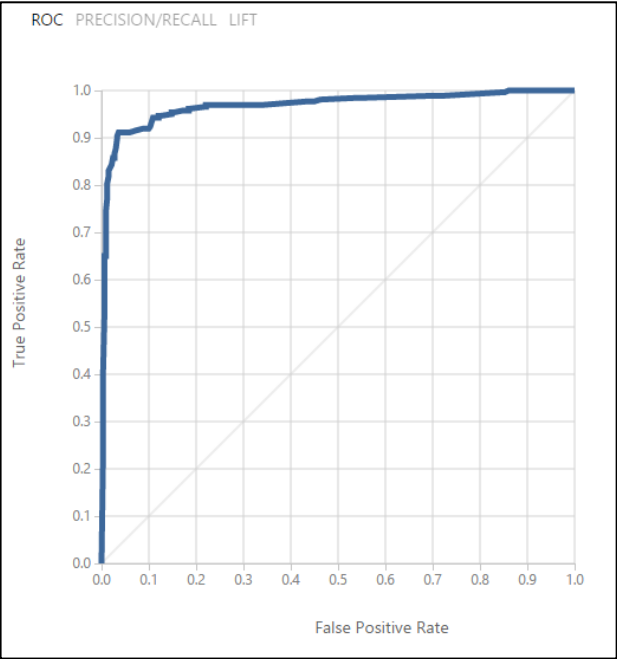
(b)

Figura 89.90. Melanoma test. (a) ROC. (b) P/R.

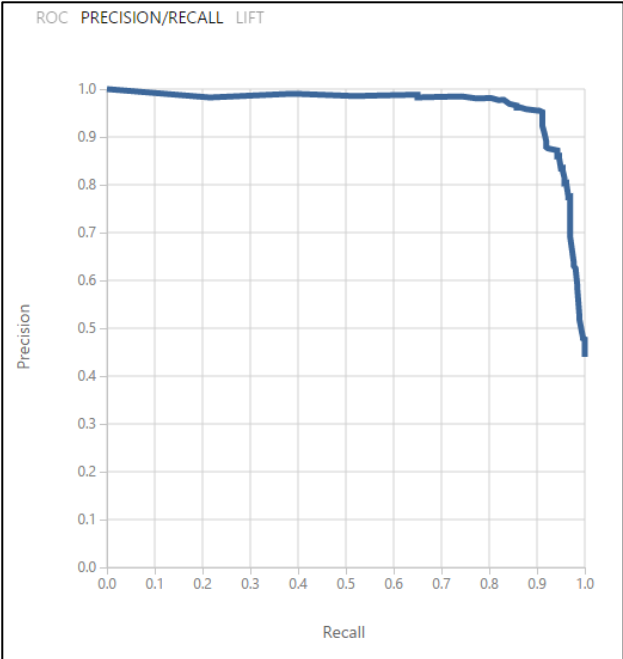
True Positive	False Negative	Accuracy	Precision	Threshold	AUC
86	37	0.749	0.811	0.5	0.842
False Positive	True Negative	Recall	F1 Score		
20	84	0.699	0.751		
Positive Label	Negative Label				
1	0				

Figura 91.92. Melanoma test, índices.

3.3.2Cáncer de pulmón



(a)



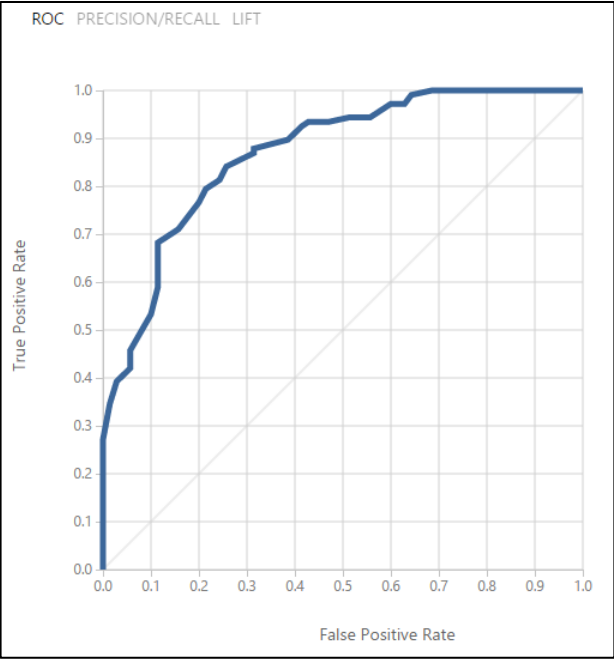
(b)

Figura 93.94. Mama test. (a) ROC. (b) P/R.

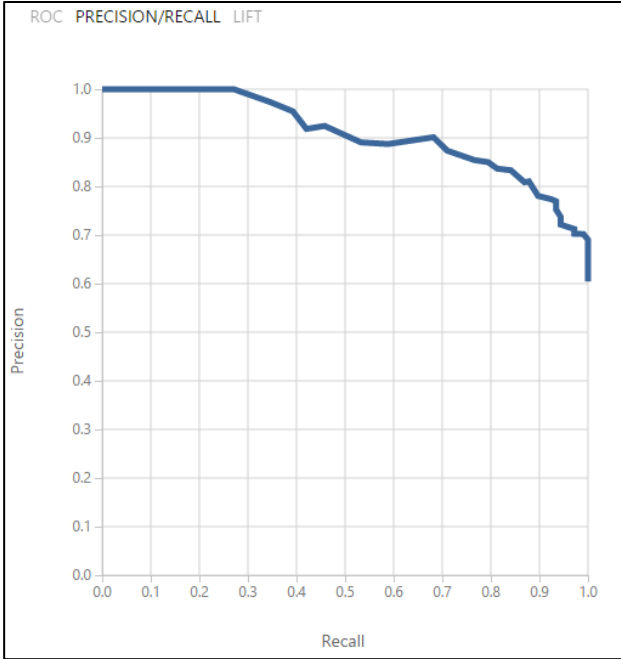
True Positive	False Negative	Accuracy	Precision	Threshold	AUC
237	23	0.941	0.952	0.5	0.967
False Positive	True Negative	Recall	F1 Score		
12	319	0.912	0.931		
Positive Label	Negative Label				
1	0				

Figura 95.96. Mama test índices.

3.3.3 Cáncer de mama



(a)



(b)

Figura 97.98. Mama test. (a) ROC. (b) P/R.

True Positive	False Negative	Accuracy	Precision	Threshold	AUC
93	14	0.797	0.809	0.5	0.870
False Positive	True Negative	Recall	F1 Score		
22	48	0.869	0.838		
Positive Label	Negative Label				
1	0				

Figura 99.100. Mama test. (a) ROC. (b) P/R.

3.4 Comparación de resultados con HEAL

Hecho el estudio y obtenidos los resultados finales se procede a la comparación con los resultados obtenidos por el HEAL sobre el mismo conjunto de datos.

Importante recordar que el parámetro "sensitive" que estudia HEAL es el mismo que hemos denominado hasta ahora como "recall" a lo largo del estudio. Se usará este parámetro junto con el "accuracy" para medir y comparar ambos resultados.

3.4.1 Melanoma

Resultados	Training		Test	
	Accuracy	Recall/Sensitive	Accuracy	Recall/Sensitive
TFG	0.761	0.691	0.749	0.699
HEAL	0.8320	0.8024	0.7315	0.6667

Tabla 101.102. Comparación Melanoma con resultados de HEAL.

En el caso del análisis realizado para la detección de melanoma se concluye que el estudio realizado por el HEAL ha obtenido mejores resultados con el subconjunto de datos de entrenamiento que el estudio realizado para este trabajo de fin de grado. No obstante sobre el subconjunto de datos de testeo se han obtenido mejores resultados con el estudio realizado en este trabajo.

3.4.2 Cáncer de pulmón

Resultados	Training		Test	
	Accuracy	Recall/Sensitive	Accuracy	Recall/Sensitive
TFG	0.924	0.876	0.941	0.912
HEAL	0.9293	0.8896	0.8730	0.7761

Tabla 103.104. Comparación pulmón con HEAL.

En cuanto al cáncer de pulmón los resultados globales se asemejan a los obtenidos para el melanoma. En el subconjunto de datos de entrenamiento el HEAL obtiene mejores resultados pero en el subconjunto de testeo es el estudio desarrollado para este trabajo de fin de grado el que obtiene un mejor rendimiento.

3.4.3Cáncer de mama

Resultados	Training		Test	
	Accuracy	Recall/Sensitive	Accuracy	Recall/Sensitive
TFG	0.790	0.840	0.797	0.869
HEAL	0.7705	0.9555	0.7436	0.9188

Tabla 105.106. Comparación mama con resultados de HEAL.

Para el cáncer de mama se concluye que tanto el subconjunto de entrenamiento como el de testeo se han obtenido mejores índices de sensibilidad en los estudios realizados por el HEAL. Por otro lado se observa un accuracy mayor en el estudio realizado para este trabajo de fin de grado

Capítulo 4. Conclusiones y líneas futuras

La inteligencia artificial ha llegado al día a día de nuestras vidas para quedarse. Son muchos ya los ámbitos donde se encuentra y muchas más el número de puertas que nos abre de cara al futuro de la vida de la sociedad humana tal y como la entendemos.

Capítulo 5.

Summary and Conclusions

FALTA

Capítulo 6.

Presupuesto

6.1 Conceptos

Las horas dedicadas por un ingeniero informático se ha estimado en un total de 300, valoradas en 10 euros/hora resulta un total de 3000 euros. Esta es la división de horas:

- a) Seminarios: 15
- b) Estudio autónomo: 65
- c) Tareas y trabajos: 200
- d) Reuniones y tutorías: 20

Por otro lado se han estimado un total de 40 euros destinados a transporte público para realizar los seminarios y reuniones necesarias.

Concepto	Coste
Horas Ingeniero Informático	3000 €
Transporte	40 €
Total	3040 €

Tabla 1.2. Presupuesto.

Bibliografía

- [1] **Enlace** **GitHub.**
<https://github.com/etsiiull/azuremedicalnn>
- [2] **Redes Neuronales Artificiales, un enfoque práctico.**
Pedro Isasi Viñuela e Inés M. Galván León.
- [3] **Redes Neuronales y Sistemas Borrosos.** Bonifacio Marín
del Brío y Alfredo Sanz Molina.
- [4] **Documentación Azure Machine Learning.**
<https://msdn.microsoft.com/en-us/library/azure/dn578280.aspx>
- [5]