

Tabla de Contenido

1	Introducción	1
1.1	Motivación	2
1.2	Hipótesis	3
1.3	Metodología	3
1.4	Objetivos	3
1.4.1	Objetivo general	4
1.4.2	Objetivos específicos	4
2	Antecedentes	5
2.1	Cáncer gástrico	5
2.1.1	Epidemiología del cáncer gástrico	6
2.1.2	Clasificación de cáncer gástrico	9
2.1.3	Proteína HER2	13
2.2	Imágenes digitales y patología digital	18
2.3	Aprendizaje de máquinas	19
2.3.1	Aprendizaje supervisado y no supervisado	21
2.3.2	Clasificación y regresión	21
2.3.3	Teoría de aprendizaje de máquinas	21
2.3.4	Conjuntos de entrenamiento, validación y evaluación	24
2.4	Redes Neuronales	25
2.5	Aprendizaje profundo	28
2.5.1	Transferencia de aprendizaje	32
2.6	Otros modelos de aprendizaje de máquinas	33
2.6.1	Máquinas de soporte vectorial	33
2.6.2	Árboles de decisión y bosques aleatorios	34
2.7	Métricas de evaluación	35
2.8	Estado del arte de clasificación de sobreexpresión de HER2 mediante ML	40
3	Bases de datos	43
3.1	Estudio PRECISO	43
3.2	Etiquetado previo	45
3.3	Etiquetado generado para esta tesis	49
4	Solución propuesta y resultados	52
4.1	Macroexperimento I	53
4.1.1	Procesamiento de datos	53
4.1.2	Metodología experimental	54
4.1.3	Resultados	59

4.2	Macroexperimento II	66
4.2.1	Metodología experimental	67
4.2.2	Resultados	70
5	Discusión	79
5.1	Comparación con otros modelos de ML	79
5.2	Macroexperimento I	80
5.3	Macroexperimento II	82
5.3.1	Estado del arte y limitaciones	85
6	Conclusiones	88
6.1	Trabajo futuro	89
	Bibliografía	91
	Apéndices	99
A	Matrices de confusión sin normalizar, macroexperimento I	99
B	Matrices de confusión sin normalizar, macroexperimento II	102
B.1	Evaluación de parches	102
B.2	Evaluación de biopsias	103

Índice de Tablas

2.1.1	Incidencia estandarizada por edad y mortalidad estandarizada por edad para Chile durante el año 2018, para ambos sexos, sólo hombres y sólo mujeres. Fuente: Elaboración propia con datos extraídos de Ferlay y col. [28].	9
2.1.2	Relación entre clasificaciones de OMS y Lauren. Fuente: Hu y col. [39]. . .	11
2.1.3	Etapas patológicas del cáncer gástrico. Fuente: González [33].	14
2.1.4	Etapas clínicas del cáncer gástrico. Fuente: González [33].	14
2.1.5	Pauta de clasificación para interpretación de inmunohistoquímica HER2 en carcinoma gástrico. *Un <i>cluster</i> de células tumorales se define como un grupo de 5 o más células neoplásicas. Traducido desde Bartley y col. [10].	17
2.7.1	Interpretación de distintos valores de κ de Cohen. Fuente: Landis y Koch [52].	40
3.1.1	Resumen de la evaluación realizada por patólogo 0, original del estudio PRECISO, desagregada por tipo de biopsia.	45
3.2.1	Resumen de las evaluaciones de biopsias realizadas por patólogos 1 y 2, desagregadas por tipo de biopsia.	45
3.2.2	Estadísticas de anotaciones realizadas por patólogos 1 y 2.	46
3.2.3	Concordancia entre patólogos 0, 1 y 2, desglosada por tipo de biopsia. α de Krippendorff fue calculado utilizando las clasificaciones realizadas por los tres patólogos.	47
3.2.4	Estadísticas de clasificaciones de biopsias y ROIs, tras aplicar proceso de filtrado basado en voto de mayoría y eliminar ROIs de patólogo 1.	47
3.3.1	Anotaciones realizadas por patólogo 3 y clasificación HER2 correspondiente, de acuerdo a método de Ruschoff/Hofmann para biopsias por resección. . .	49
3.3.2	Resumen de las evaluaciones realizadas por patólogo 3.	51
3.3.3	Estadísticas de anotaciones realizadas por patólogo 3.	51
3.3.4	Concordancia de patólogo 3 con patólogos 0, 1 y 2, utilizando esquemas de 4 clases (0, 1+, 2+ y 3+) y 3 clases (Negativo, Equívoco, Positivo). Todas las muestras evaluadas corresponden a biopsias por resección. * α de Krippendorff puede ser computado aún con datos faltantes; así, el α calculado para el grupo de todos los patólogos corresponde a las 34 biopsias evaluadas por el patólogo 3. ** N° de muestras es igual en ambos esquemas.	51
4.1.1	Estadísticas de dataset 1, formado por los parches extraídos de los ROIs anotados por los patólogos 1 y 2, desagregados por magnificación. Filtro aplicado corresponde a parches con proporción de tejido mayor a 20%. . . .	55
4.1.2	Resumen de cada configuración experimental del macroexperimento I. . . .	57
4.1.3	Hiperparámetros seleccionados mediante k -fold anidado, magnificación 10x.	60
4.1.4	Hiperparámetros seleccionados mediante k -fold anidado, magnificación 20x.	61

4.1.5	Hiperparámetros seleccionados mediante k -fold anidado, magnificación 40x.	61
4.1.6	Resumen de resultados conseguidos con magnificación 10x. En el caso de precisión, recuperación y F1-score, se calculó un promedio ponderado. . . .	62
4.1.7	Resumen de resultados conseguidos con magnificación 20x. En el caso de precisión, recuperación y F1-score, se calculó un promedio ponderado. . . .	63
4.1.8	Resumen de resultados conseguidos con magnificación 40x. En el caso de precisión, recuperación y F1-score, se calculó un promedio ponderado. . . .	64
4.2.1	Estadísticas de dataset 2, formado por los parches extraídos de los ROIs anotados por el patólogo 3. Todos los parches fueron extraídos a magnificación 10x. Filtro aplicado corresponde a parches con proporción de tejido mayor a 20%.	66
4.2.2	Resumen de cada configuración experimental del macroexperimento II. . . .	68
4.2.3	Resumen de resultados conseguidos en experimento <i>todo en uno</i> , evaluación de clasificación de parches de ROIs. En el caso de precisión, recuperación y F1-score, se calculó un promedio ponderado.	72
4.2.4	Resumen de resultados conseguidos en experimento <i>en cascada</i> , evaluación de clasificación de parches de ROIs. En el caso de precisión, recuperación y F1-score, se calculó un promedio ponderado.	72
4.2.5	Resumen de resultados para clasificación de biopsias, considerando clases 0, 1+, 2+ y 3+. En el caso de precisión, recuperación y F1-score, se calculó un promedio ponderado.	75
4.2.6	Resumen de resultados para clasificación de biopsias, considerando clases <i>Negativo</i> , <i>Equívoco</i> y <i>Positivo</i> . En el caso de precisión, recuperación y F1-score, se calculó un promedio ponderado.	76

Índice de Ilustraciones

2.1.1	Partes del estómago. Fuente: American Cancer Society [6].	6
2.1.2	Capas de la pared estomacal. Fuente: American Cancer Society [6].	6
2.1.3	Incidencia del cáncer gástrico en el mundo, para ambos sexos, estandarizada por edad; resaltada está la incidencia de Chile. Fuente: Elaboración propia con datos extraídos de Ferlay y col. [28].	7
2.1.4	Incidencia estandarizada por edad del cáncer gástrico en distintas partes del mundo para ambos sexos. Fuente: Bray y col. [13].	8
2.1.5	Incidencia y mortalidad estimada para distintos tipos de cáncer en Chile. Fuente: Ferlay y col. [28].	9
2.1.6	Clasificación de Lauren para carcinomas gástricos tempranos: A) tipo intestinal, B) tipo difuso y C) tipo mixto. Fuente: Chong y col. [18].	10
2.1.7	Análisis inmunohistoquímico en muestras representativas de expresión de HER2 en cáncer gástrico. A) 0, negativo. B) 1+, negativo. C) 2+, equívoco. D) 3+, positivo. Fuente: Bartley y col. [10].	15
2.1.8	Imágenes de cáncer gástrico y cáncer de mama teñidas inmunohistoquímicamente. Ambos tumores son HER2-positivos (IHC 3+), pero se aprecian diferencias importantes: mientras en el cáncer de mama se verifica que cada célula inmunopositiva presenta una tinción completa de su membrana, en el cáncer gástrico esto no ocurre, teniéndose que muchas membranas no están teñidas completamente y quedan “abiertas”. Fuente: Ross y Mulcahy [73].	16
2.1.9	ISH e IHC y sus respectivos objetivos de análisis; ejemplo en cáncer gástrico. Fuente: Dako [20].	17
2.2.1	Representación de una imagen digital como un conjunto de tres matrices. Fuente: Gonzalez y Woods [32].	18
2.2.2	Ejemplo de imagen piramidal con múltiples magnificaciones. En este caso, biopsia de cáncer gástrico con tinción inmunohistoquímica. A) 0.5x, B) 5x, C) 20x, D) 40x. Fuente: Elaboración propia con datos provenientes del estudio PRECISO [61].	19
2.3.1	Esquema básico de un problema típico de aprendizaje de máquinas, con ejemplos relativos a la pregunta de si aprobar o rechazar créditos bancarios a potenciales clientes. Fuente: Abu-Mostafa, Magdon-Ismail y Lin [2].	20
2.3.2	Izquierda: ejemplo de problema de clasificación binaria; el objetivo es encontrar una función que permita separar los datos correctamente. Derecha: ejemplo de problema de regresión; el objetivo es encontrar una función que se aproxime a la distribución de los datos. Fuente: Soni [84].	22
2.3.3	Esquema general de un problema de aprendizaje de máquinas supervisado, considerando distribuciones de probabilidad y ruido. Fuente: Abu-Mostafa, Magdon-Ismail y Lin [2].	23

2.3.4	Esquema de K -Fold, con $K = 10$. Fuente: Norena [66].	25
2.4.1	Arquitectura típica de una red neuronal artificial completamente conexas. Fuente: Dertat [25].	26
2.4.2	Esquemmatización de la heurística de descenso de gradiente. Es posible apreciar que dependiendo del punto de partida es posible alcanzar un mínimo local, mas no necesariamente el global. Fuente: Zhang [93].	27
2.5.1	Arquitectura típica de una red neuronal convolucional. Fuente: Saha [79].	30
2.5.2	Campos receptivos locales en una red neuronal convolucional.	30
2.5.3	Ejemplos de distintos filtros de tamaño 5x5 aprendidos por una red neuronal convolucional. Bloques más oscuros representan un mayor peso, lo cual implica que dicho filtro responde con mayor fuerza a los pixeles correspondientes. Fuente: Nielsen [64].	31
2.5.4	Ejemplos de aplicación de <i>max pooling</i> y <i>average pooling</i> , con una ventana de 2x2 neuronas. Fuente: Saha [79].	32
2.6.1	Izquierda: múltiples hiperplanos que dividen un conjunto de datos en base a sus clases. Derecha: hiperplano de margen máximo sobre el mismo conjunto de datos. Fuente: Gandhi [31].	34
2.6.2	Ejemplo de árbol de decisión en un espacio bidimensional, junto a las particiones generadas en dicho espacio. Fuente: James y col. [42].	35
2.7.1	Esquema de matriz de confusión, junto a fórmulas derivadas.	37
2.7.2	Ejemplo de curva ROC. Fuente: Brownlee [14].	38
3.1.1	Ejemplo de biopsias del estudio <i>PRECISO</i> . A) Tinción H&E, magnificación 0.21x B) Tinción IHC, magnificación 0.25x. Es posible apreciar que en la biopsia IHC, en el lado izquierdo se encuentra el tejido de control.	44
3.2.1	Ejemplo de anotaciones realizadas por patólogos 1 (azul) y 2 (verde) sobre biospia endoscópica. Magnificación 1.25x.	48
3.2.2	Ejemplo de ROIs resultantes tras aplicar proceso de filtrado, magnificación x10. A) 0, B) 1+, C) 2+, D) 3+.	48
3.3.1	Ejemplo de anotaciones realizadas por patólogo 3, todas provenientes de la misma biopsia por resección a magnificación 10x. A) No tumor, B) sin reactividad, C) reactividad positiva no lineal, D) reactividad lineal casi imperceptible, E) reactividad lineal débil, F) reactividad lineal fuerte.	50
4.1.1	Esquema de extracción de parches desde un ROI. Todos los parches extraídos de un ROI de tipo 3+ también son de tipo 3+.	53
4.1.2	Ejemplos de parches extraídos de ROIs, junto a la proporción de tejido presente en cada uno de ellos. Además, en la esquina de cada parche, se encuentra anotada la clase a la que pertenece.	54
4.1.3	Arquitectura de <i>Inception v3</i> . Fuente: Tsang [88].	55
4.1.4	Parche extraído de ROI (esquina superior izquierda) junto a ejemplos de transformaciones aleatorias aplicadas sobre dicho parche.	57
4.1.5	Esquema de generación de subconjuntos para entrenamiento con técnica de validación cruzada k -fold, con $k = 5$	58
4.1.6	Esquema de k -fold anidado. Adaptado desde Jin [43].	60
4.1.7	Matrices de confusión normalizadas de experimentos con mejores resultados para magnificación 10x.	62

4.1.8	Matrices de confusión normalizadas de experimentos con mejores resultados para magnificación 20x.	63
4.1.9	Matrices de confusión normalizadas de experimentos con mejores resultados para magnificación 40x.	64
4.1.10	Ejemplos de clasificación de parches extraídos de ROIs en macroexperimento I. Cada ROI está formado por varios parches, y sobre cada parche, se pintó con transparencia la clasificación predicha (verde: HER2 negativo, amarillo: HER2 equívoco, rojo: HER2 positivo). Visualizaciones de lado izquierdo fueron generadas usando las redes entrenadas en el experimento de reentrenamiento total a magnificación 10x, mientras que las del lado derecho corresponden a reentrenamiento total con magnificación 40x. A) y B) ROI de tipo 3+, C) y D), ROI de tipo 2+, E) y F) ROI de tipo 0. Todos los ROIs fueron anotados por el patólogo 2.	65
4.2.1	Esquema del experimento <i>todo en uno</i>	68
4.2.2	Esquema del experimento <i>en cascada</i>	69
4.2.3	Esquema de evaluación de parches extraídos directamente de biopsias en experimento <i>en cascada</i>	70
4.2.4	Análisis de umbral de decisión para clasificación binaria de <i>tumor / no tumor</i>	71
4.2.5	Matriz de confusión normalizada para experimento <i>todo en uno</i> , evaluación de clasificación de parches de ROIs.	73
4.2.6	Matrices de confusión normalizadas para experimento <i>en cascada</i> , evaluación de clasificación de parches de ROIs.	73
4.2.7	Matrices de confusión normalizadas para evaluación de clasificación de biopsias, usando clases 0, 1+, 2+ y 3+. En clasificación binaria de esquema <i>en cascada</i> se utilizó un umbral $T = 0,3$	75
4.2.8	Matrices de confusión normalizadas para evaluación de clasificación de biopsias, usando clases <i>Negativo</i> , <i>Equívoco</i> y <i>Positivo</i> . En clasificación binaria de esquema <i>en cascada</i> se utilizó un umbral $T = 0,3$	76
4.2.9	Biopsia con sobreexpresión HER2 negativa, correctamente clasificada. A) biopsia original, B) visualización generada utilizando la clasificación producida por el modelo <i>en cascada</i>	77
4.2.10	Biopsia con sobreexpresión HER2 equívoca, correctamente clasificada. A) biopsia original, B) visualización generada utilizando la clasificación producida por el modelo <i>en cascada</i>	78
4.2.11	Biopsia con sobreexpresión HER2 positiva, correctamente clasificada. A) biopsia original, B) visualización generada utilizando la clasificación producida por el modelo <i>en cascada</i>	78
5.2.1	Subexperimento de selección de parámetros simple utilizando parches a magnificación 40x. Entrenamiento con subconjuntos 2, 3 y 5, y validación con conjunto 1.	81
5.2.2	Subexperimento de selección de parámetros con <i>fine tuning</i> y <i>data augmentation</i> utilizando parches a magnificación 10x. Entrenamiento con subconjuntos 1, 4 y 5, y validación con conjunto 2.	81
5.2.3	Subexperimento de selección de parámetros con reentrenamiento total y <i>data augmentation</i> utilizando parches a magnificación 20x. Entrenamiento con subconjuntos 1, 3 y 4 y validación con conjunto 5.	82

5.3.1	Ejemplos de ROIs pertenecientes a clases que las redes suelen confundir. A) Reactividad no lineal, B) Reactividad lineal casi imperceptible, C) Reactividad lineal débil. Todos los ROIs provienen de la misma biopsia.	83
5.3.2	Biopsia con sobreexpresión HER2 negativa, incorrectamente clasificada como equívoca. Tejido de tipo 2+ o 3+ calculado por algoritmo corresponde a 16,5% del tejido canceroso. A) biopsia original, B) visualización generada utilizando la clasificación producida por el modelo <i>en cascada</i>	86
A.1	Matrices de confusión sin normalizar para macroexperimento I, magnificación 10x.	99
A.2	Matrices de confusión sin normalizar para macroexperimento I, magnificación 20x.	100
A.3	Matrices de confusión sin normalizar para macroexperimento I, magnificación 40x.	101
B.1	Matriz de confusión sin normalizar para experimento <i>todo en uno</i> , evaluación de clasificación de parches de ROIs.	102
B.2	Matrices de confusión sin normalizar para experimento <i>en cascada</i> , evaluación de clasificación de parches de ROIs.	103
B.3	Matrices de confusión sin normalizar para evaluación de clasificación de biopsias, usando clases 0, 1+, 2+ y 3+. En clasificación binaria de esquema <i>en cascada</i> se utilizó un umbral $T = 0,3$	103
B.4	Matrices de confusión sin normalizar para evaluación de clasificación de biopsias, usando clases <i>Negativo</i> , <i>Equívoco</i> y <i>Positivo</i> . En clasificación binaria de esquema <i>en cascada</i> se utilizó un umbral $T = 0,3$	104