



UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E
INFORMÁTICA



**MODELO DE RED NEURONAL ARTIFICIAL PARA PREDECIR
LA DESERCIÓN DE ESTUDIANTES DE LA MAESTRÍA EN
DERECHO DE LA UNIVERSIDAD NACIONAL DEL ALTIPLANO
PUNO – 2019**

TESIS

PRESENTADA POR:

Bach. ANA MILAGROS MAMANI QUISPE

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO ESTADÍSTICO E INFORMÁTICO

PUNO – PERÚ

2022



DEDICATORIA

Con mucho cariño a mi recordado padre German Paulo Mamani Pacco (†) por haber sido mi guardián, mi más fiel admirador, mi ejemplo a seguir, quien siempre será luz y guía en mi vida, por ser quien me impulso a iniciar mis estudios; y ahora, que estoy en la culminación de esta etapa sé, que este momento hubiera sido especial para él tanto como lo es para mí.

A mi querida madre Victoria Quispe Rojas por su constante sacrificio e incondicional apoyo, por ser el pilar fundamental en la culminación de mis estudios dándome siempre su amor, comprensión y consejos a realizarme como profesional y como persona.

A mis hermanos Hebeth, Jackeline y Deisy por ser compañía y consejeros en el camino de mi formación profesional, por brindarme apoyo moral y las ganas de luchar por algo mejor.

A mis sobrinos José y Teo por llenar mi vida de amor y alegría con cada una de sus ocurrencias.

Ana Milagros Mamani Quispe



AGRADECIMIENTOS

- *A la existencia misma que en su conexión e interacción de toda la diversidad de las cosas variables hacen posible este momento, el estar aquí en el presente, el vivir o la esencia de estar vivos.*
- *A la Universidad Nacional del Altiplano pionera de la educación superior en la región y ejemplo de calidad académica con licenciamiento y acreditación, por la oportunidad de convertirme en un profesional post graduado en el calor de sus aulas.*
- *A mis jurados por su experiencia, paciencia y comprensión en las vicisitudes para la evaluación y sustentación de la tesis.*
- *A mi asesor de tesis el Dr. Percy Huata, conspicuo guía en la consecución de este trabajo y por su aporte en cuestiones metodológicas.*
- *A mis docentes, distinguidos profesionales por su esfuerzo, contribución académica, y lo transformaron en calidad profesional*
- *A quienes, con sus investigaciones precedentes, hicieron posible tener un panorama más amplio sobre el tema materia de tesis*
- *A mis padres.*

Ana Milagros Mamani Quispe



ÍNDICE GENERAL

	Pág.
DEDICATORIA	
AGRADECIMIENTOS	
ÍNDICE GENERAL	
ÍNDICE DE TABLAS	
ÍNDICE DE FIGURAS	
ÍNDICE DE ACRÓNIMOS	
ÍNDICE DE ANEXOS	
RESUMEN	13
ABSTRACT.....	14
CAPÍTULO I	
INTRODUCCIÓN	
1.1.DESCRIPCIÓN DEL PROBLEMA	16
1.2.FORMULACIÓN DEL PROBLEMA	17
1.2.1.Problema general.....	17
1.2.2.Problemas específicos	17
1.3.JUSTIFICACIÓN DE LA INVESTIGACIÓN	18
1.4.OBJETIVOS DE LA INVESTIGACIÓN.....	19
1.4.1.Objetivo General	19
1.4.2.Objetivos Específicos.....	20
1.5.HIPOTESIS DE LA INVESTIGACIÓN.....	20
1.5.1 Hipótesis General.....	20
1.5.2 Hipótesis Específicas	20



CAPÍTULO II

REVISIÓN DE LITERATURA

2.1.ANTECEDENTES.....	21
2.1.1.Antecedentes locales	21
2.1.2.Antecedentes nacionales	24
2.1.3.Antecedentes internacionales	26
2.2.MARCO TEÓRICO	28
2.2.1.La deserción	28
2.2.2.Tipos de deserción y retención en estudiantes	29
2.2.3.Aprendizaje automático	32
2.2.4.Red neuronal artificial.....	36
2.3.TÉRMINOS UTILIZADOS.....	43

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1.TIPO DE INVESTIGACIÓN	44
3.2.DISEÑO DE INVESTIGACIÓN.....	44
3.3.POBLACIÓN Y MUESTRA	45
3.3.1.Población.....	45
3.3.2.Muestra.....	45
3.4.RECOLECCIÓN DE DATOS.....	46
3.4.1.Técnica.....	46
3.4.2.Instrumento	46
3.4.3.Método de recolección de datos.....	47
3.5.VARIABLES DE ESTUDIO.....	47
3.5.1.Variable Dependiente.....	48



3.5.2. Variables Independientes.	48
3.6. PROCESAMIENTO DE DATOS	49
3.6.1. Metodología de Redes Neuronales	49
CAPÍTULO IV	
RESULTADOS Y DISCUSIÓN	
4.1. VARIABLES QUE CONTRIBUYEN AL RIESGO DE DESERCIÓN.....	54
4.2. ARQUITECTURA DE RNA PARA PREDECIR LA DESERCIÓN	71
4.2.1. Modelo de Red Neuronal 1	74
4.2.2. Modelo de Red neuronal 2	77
4.2.3. Modelo de Red neuronal 3	80
4.2.4. Modelo de Red neuronal 4	83
4.3. ENTRENAMIENTO Y VALIDACIÓN DEL MODELO DE RNA PARA OBTENER UN MODELO DE PREDICCIÓN	85
4.3.1. Entrenamiento	86
4.3.2. Validación..	89
4.4. MODELO DE PREDICCIÓN PARA LA DESERCIÓN DE ESTUDIANTES MEDIANTE RNA.....	91
V. CONCLUSIONES	96
VI. RECOMENDACIONES	97
VII. REFERENCIAS.....	98
ANEXOS.....	106

Área : Inteligencia artificial
Tema : Redes Neuronales Artificiales

FECHA DE SUSTENTACIÓN: 15 de julio del 2022



ÍNDICE DE TABLAS

	Pág.
Tabla 01. Métricas de clasificación.	42
Tabla 02. Resumen descriptivo.....	54
Tabla 03. Deserción Estudiantil.....	55
Tabla 04. Sexo de los estudiantes de la Maestría en Derecho de la UNA.....	57
Tabla 05. Estado de cuenta de los estudiantes de la Maestría en Derecho de la UNA..	58
Tabla 06. Estado civil estudiantes de la Maestría en Derecho de la UNA	59
Tabla 07. Procedencia de los estudiantes de la Maestría en Derecho de la UNA	60
Tabla 08. Tipo de universidad que provienen los estudiantes de la Maestría en Derecho de la UNA.....	61
Tabla 09. Sexo y deserción estudiantil	62
Tabla 10. Estado civil y deserción estudiantil	65
Tabla 11. Procedencia y deserción estudiantil.....	67
Tabla 12. Estado de cuenta filtrado por la deserción.....	68
Tabla 13. Tipo de universidad y la deserción.	69
Tabla 14. Índice de las muestras de entrenamiento	70
Tabla 15. Índice de las muestras de validación.....	71
Tabla 16. Configuración de las 4 redes neuronales a comparar	72
Tabla 17. Número de iteraciones que minimiza la RNA 1	75
Tabla 18. Número de iteraciones que minimiza la RNA 2	78
Tabla 19. Número de iteraciones que minimiza la RNA 3	81
Tabla 20. Número de iteraciones que minimiza la RNA 4	84
Tabla 21. Matriz de confusión de la red neuronal 1	86
Tabla 22. Matriz de confusión de la red neuronal 2	86



Tabla 23. Matriz de confusión de la red neuronal 3	87
Tabla 24. Matriz de confusión de la red neuronal 4	87
Tabla 25. Métricas del error de entrenamiento en las 4 redes neuronales creadas	87
Tabla 26. Matriz de confusión del conjunto de validación para la red neuronal 1	89
Tabla 27. Matriz de confusión del conjunto de validación para la red neuronal 2	89
Tabla 28. Matriz de confusión del conjunto de validación para la red neuronal 3	90
Tabla 29. Matriz de confusión del conjunto de validación para la red neuronal 4	90
Tabla 30. Métricas de error de validación en las 4 redes neuronales creadas	90
Tabla 31. Parámetros de la RNA que mejor se ajustó para la predicción de la deserción estudiantes de la Maestría en Derecho de la UNA, mediante redes neuronales artificiales.....	92



ÍNDICE DE FIGURAS

	Pág.
Figura 01. Tipos de deserción de estudiantes.	30
Figura 02. Categorías generales de las causas de la deserción	31
Figura 03. Diagrama de regresión.	35
Figura 04. Diagrama de clasificación.	36
Figura 05. Sistema global de proceso de una red neuronal	37
Figura 06. Ejemplo de una red neuronal totalmente conectada.....	37
Figura 07. Funciones de activación.	39
Figura 08. Distribución de la deserción estudiantil.	55
Figura 09. Histograma del promedio de notas de los alumnos.....	56
Figura 10. Histograma de la edad que cuentan los alumnos.....	56
Figura 11. Sexo de los estudiantes de la Maestría en Derecho de la UNA.	57
Figura 12. Estado de cuenta de los estudiantes de la Maestría en Derecho - UNA.....	58
Figura 13. Estado civil estudiantes de la Maestría en Derecho de la UNA.	59
Figura 14. Procedencia de los estudiantes de la Maestría en Derecho de la UNA.	60
Figura 15. Tipo de universidad de la que provienen los estudiantes de la Maestría en Derecho de la UNA.....	61
Figura 16. Sexo filtrado por la deserción estudiantil.	62
Figura 17. Histograma de los promedios filtrados por la deserción	63
Figura 18. Edad de los estudiantes filtrados por la deserción.....	64
Figura 19. Estado civil filtrado por la deserción.....	66
Figura 20. Procedencia filtrada por la deserción	67
Figura 21. Estado de cuenta filtrado por la deserción	68
Figura 22. Tipo de universidad filtrado por la deserción.....	69



Figura 23. Modelo de Red neuronal 1	74
Figura 24. Número de iteraciones que minimiza la RNA 1 de la función de pérdida en los conjuntos de entrenamiento y validación.....	76
Figura 25. Modelo de Red neuronal 2	77
Figura 26. Número de iteraciones que minimiza la RNA 2 de la función de pérdida en los conjuntos de entrenamiento y validación.....	79
Figura 27. Modelo de Red neuronal 3	80
Figura 28. Número de iteraciones que minimiza la RNA 3 de la función de pérdida en los conjuntos de entrenamiento y validación.....	82
Figura 29. Red neuronal 4	83
Figura 30. Número de iteraciones que minimiza la RNA 4 de la función de pérdida en los conjuntos de entrenamiento y validación.....	85
Figura 31. F1_Score de las 4 redes neuronales en comparación en la muestra de entrenamiento.	88
Figura 32. F1_Score de las 4 redes neuronales en comparación en la muestra de validación.	91
Figura 33. La RNA que mejor se ajusto de predicción de la deserción.....	93



ÍNDICE DE ACRÓNIMOS

UNA:	Universidad Nacional del Altiplano.
ML:	Machine Learning (Aprendizaje automático)
RNA:	Red neuronal artificial.
Recall:	Precisión en la clase objetivo alumnos que desertan.
Accuracy:	Precisión en la clase no objetivo alumnos que no desertan.
F1_Score:	Promedio armónico entre Recall y Accuracy.



ÍNDICE DE ANEXOS

	Pág.
Anexo 1. Datos recopilados	108
Anexo 2. Código utilizado en R.....	122
Anexo 3. Ficha de Inscripción	130
Anexo 4. Codigo utilizado en Phyton.....	131



RESUMEN

El presente trabajo de investigación se realizó con la finalidad de determinar la Red Neuronal Artificial que proporcione un buen ajuste para predecir la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano de Puno. La población de estudio estuvo conformada por los registros de datos extraídos de las carpetas de ingreso a lo largo del tiempo de los alumnos que tuvo la Maestría en Derecho de la UNA Puno. La muestra estuvo compuesta de 334 registros de datos extraídos de las carpetas de ingreso de los alumnos durante los años 2015 – 2019. La investigación se desarrolló bajo el enfoque cuantitativo, con un nivel de investigación descriptivo, y por su naturaleza es de tipo documental, el diseño de investigación es no experimental y de corte transversal. Concluyendo que se logró determinar que la RNA que mejor se ajustó para la predicción de la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno, cuenta con una tasa de aprendizaje de 0.1, una sola neurona en la capa oculta, realizó 25 iteraciones, la función de activación es sigmoidea, el algoritmo para calcular los pesos sinápticos es de retro propagación, cuenta con un grado de acierto de: Accuracy del 92% en la clase de los no desertores, Recall del 86% en la clase de los alumnos desertores y en general cuenta con F1_Score de 89%.

PALABRAS CLAVE:

Deserción, Machine Learning, predicción, overfitting, redes neuronales artificiales.



ABSTRACT

The present research was carried out with the purpose of determining the Artificial Neural Network that provides a good fit to predict desertion in students of the Master of Laws of the Postgraduate School of the National University of the Altiplano of Puno. The study population was made up of the data records extracted from the admission folders over time of the students who had the Master's Degree in Law from UNA Puno. The sample consisted of 334 data records extracted from the student's admission folders during the years 2015 - 2019. The research was developed under the quantitative approach, with a descriptive level of research, and by its nature it is documentary, the research design is non-experimental and cross-sectional. Concluding that it was possible to determine that the ANN that was best adjusted for the prediction of the dropout of students of the Master of Laws of the Postgraduate School of the National University of the Altiplano - Puno, has a learning rate of 0.1, a single neuron in the hidden layer, performed 25 iterations, the activation function is sigmoid, the algorithm to calculate the synaptic weights is back propagation, it has a degree of success of: Accuracy of 92% in the class of non-dropouts, Recall of 86% in the class of dropout students and in general it has an F1_Score of 89%.

Key words: Attrition, Machine Learning, prediction, overfitting, artificial neural networks.



CAPÍTULO I

INTRODUCCIÓN

Las organizaciones de educación públicas o privadas, en los diferentes niveles escolares, se encuentran en la continua búsqueda de datos primarios que sirvan para la formulación de estrategias que ayuden a mejorar la calidad de su razón de ser (Hernandez et al., 2014). Según el personal de la unidad de Posgrado de Derecho los estudiantes desertan en su mayoría, en los primeros semestres, y esto se ve reflejado en la disminución de estudiantes matriculados en el siguiente semestre, muchos de estos por la normativa impuesta por la universidad ya que para un estudiante solo puede ser considerado aprobado con una mínima nota de 14, y si un estudiante tiene más de 3 asignaturas desaprobadas automáticamente se le elimina (EPG, 2020), otros estudiantes por motivos económicos dejan de estudiar, también por qué no se cumple con las expectativas que los alumnos tienen; sobre esta base la Maestría en Derecho de la Universidad Nacional del Altiplano (UNA) sufre una serie de pérdidas tanto económicas, como sociales ya que estos programas a pesar de ser autofinanciadas por los mismos estudiantes es parte de una entidad estatal, por ello se espera que tengan repercusiones sociales, tecnológicas y que aporten en el desarrollo local, regional y nacional; por ello intervenir con acciones preventivas y correctivas es una necesidad que debe ser atendida, para evitar que los estudiantes de posgrado abandonen sus estudios.

Hoy en día el auge de la inteligencia artificial, especialmente el aprendizaje automático y las redes neuronales son una alternativa para diversos problemas, en los cuales se requiere modelar cierta información para luego abordar el problema teniendo en cuenta el comportamiento de un determinado conjunto de individuos, en este caso se



utilizará las redes neuronales para predecir si un estudiante desertará o continuara con sus estudios de la Maestría en Derecho de la Universidad Nacional del Altiplano.

La presente tesis desarrollada cuenta con la siguiente estructura:

Capítulo I: Se presentó la descripción del problema, formulación del problema, justificación de esta, como también se mostró los objetivos e hipótesis de investigación.

Capítulo II: Se detalló los fundamentos teóricos y/o bases teóricas de la investigación, conceptos básicos y antecedentes referentes a la deserción estudiantil y las redes neuronales.

Capítulo III: Se presentó materiales y métodos, donde se detalló la población y muestra, diseño, alcance de la investigación, técnicas e instrumentos de recolección de datos, metodología estadística utilizada.

Capítulo IV: Se expuso los resultados y la discusión de estos, constituidos de cada objetivo planteado en la investigación, demostrando los resultados mediante cuadros y gráficos y redes neuronales en estudio, los mismos que han sido analizados, para poder demostrar las hipótesis planteadas, de este modo llegar a las conclusiones y recomendaciones.

Finalmente se presentó, la bibliografía y anexos en los que se detalla la base de datos y el código fuente desarrollado en R y Phyton.

1.1. DESCRIPCIÓN DEL PROBLEMA

Según Barrientos & Umaña (2010) la deserción estudiantil en estudiantes de posgrado afecta en el retraso de la graduación del 68% de los estudiantes, la deserción según reglamento académico es del 87%, la deserción por intención expresada por los estudiantes es de 19%, la deserción estudiantil se ha reflejado en su mayoría en los primeros años de estudio, de ahí se puede indicar, que es un proceso de abandono por completo de las actividades académicas lo que es un problema continuo, en el Perú, existe



dificultades para identificar y encontrar información que nos permita predecir la probabilidad de que un estudiante desierte. En las últimas décadas el avance y desarrollo de las redes neuronales artificiales surgen como una potente herramienta para enfrentar este tipo de problemas. Además, la maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno centra sus esfuerzos en la formación de sus alumnos y la mejora continua de su calidad de enseñanza ya que la educación superior pública es vista como un sistema planeado, organizado, estructurado y dirigido por el Estado, reconocido como un bien público (Hernandez et al., 2014). Por tal motivo es importante identificar el patrón de comportamiento de las variables que intervienen en la deserción de los estudiantes, para así proponer estrategias de retención que puedan ser implementados por la maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno.

1.2. FORMULACIÓN DEL PROBLEMA

1.2.1. Problema general

¿Cuál será la Red Neuronal Artificial que proporcione un buen ajuste para predecir la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno?

1.2.2. Problemas específicos

1. ¿Cuál será el comportamiento de los factores que contribuyen al riesgo de deserción del estudiante durante su periodo de estudios?
2. ¿Cuál será el diseño de la arquitectura de una red neuronal artificial para predecir la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano Puno?



3. ¿Es posible entrenar y validar el modelo de Red Neuronal Artificial para obtener un modelo de predicción para la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno?

1.3. JUSTIFICACIÓN DE LA INVESTIGACIÓN

Hoy en día las universidades peruanas están en la necesidad de mejorar su calidad educativa debido al proceso de Licenciamiento Institucional que es un procedimiento obligatorio para todas las universidades del país, a través del cual cada casa de estudios debe demostrar ante la SUNEDU que cumple con las Condiciones Básicas de Calidad para poder brindar el servicio educativo (SUNEDU, 2016); dicho esto tenemos entendido que la deserción se produce cuando el alumno abandona las aulas de clase antes de terminar su formación académica (Rodríguez & Londoño, 2010); y cuando hablamos de calidad en la educación se considera como uno de los principales indicadores a la eficiencia terminal, es decir que, el número de alumnos que ingresa a un programa académico debe ser el mismo que concluye. Es por ello que esta investigación posee relevancia científica al identificar el comportamiento de los factores que contribuyen al riesgo de deserción del estudiante, y aportará a establecer medidas para la retención de estudiantes.

La deserción de posgrado afecta tanto en los ámbitos personales como en los institucionales y económicos. En lo personal, implica una condición de fracaso que afecta emocionalmente por la discrepancia con las aspiraciones personales e incide en la trayectoria ocupacional de los individuos. En lo institucional, implica una disminución del rendimiento académico (Chavarri, 2018) de la Maestría. Y en lo económico, la Maestría en Derecho se financia con los pagos de las pensiones de enseñanza de los alumnos, es decir que se autofinancia, por lo tanto, para su correcto funcionamiento es



preferible que todos sus estudiantes culminen sus estudios cumpliendo sus deberes como estudiantes.

Además, cuando se habla de deserción estudiantil, son más los estudios que se hacen con estudiantes de colegio y de pregrado; y son muy poco los casos que tratan sobre deserción en estudiantes de posgrado, por ello la presente investigación aportará importante información al respecto; y que mejor utilizando la técnica de Redes Neuronales Artificiales y es que estas se han convertido en la familia de algoritmos de machine learning más populares y potentes de los últimos tiempos, con los que podemos modelar comportamientos inteligentes para poder determinar la Red Neuronal que mejor se ajuste para poder predecir la deserción en estudiantes de la Maestría en Derecho de la Universidad Nacional del Altiplano – Puno.

Los resultados del presente estudio servirán como base para la implementación de políticas de mejora en la Maestría en Derecho de la Una Puno, y por eso se hace necesario tener un mecanismo que permita determinar el riesgo de deserción de los alumnos, para aplicar medidas de mejora en forma temprana y oportuna. Por tal razón la investigación es viable y de mucha importancia por los resultados que traerá la misma.

Por ello el presente trabajo de investigación surge ante uno de los problemas que enfrenta la Maestría en Derecho de la Universidad Nacional del Altiplano que es la deserción de estudiantes.

1.4. OBJETIVOS DE LA INVESTIGACIÓN

1.4.1. Objetivo General

Determinar la Red Neuronal Artificial que proporcione un buen ajuste para predecir la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno.



1.4.2. Objetivos Específicos

1. Identificar el comportamiento de las variables que contribuyen al riesgo de deserción del estudiante durante su periodo de estudios.
2. Diseñar la arquitectura de la red neuronal artificial para predecir la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano Puno.
3. Entrenar y validar el modelo de Red Neuronal Artificial para obtener un modelo de predicción para la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno.

1.5. HIPOTESIS DE LA INVESTIGACIÓN

1.5.1. Hipótesis General

Con la aplicación de Redes Neuronales Artificiales es posible determinar un modelo óptimo de predicción para la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano - Puno.

1.5.2. Hipótesis Específicas

1. Todas las variables identificadas contribuyen al riesgo de deserción del estudiante durante su periodo de estudios.
2. El diseño de la arquitectura de una red neuronal artificial para predecir la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano Puno es de una capa oculta.
3. Se puede entrenar y validar el modelo de Red Neuronal Artificial para obtener un modelo de predicción con el 80% de los datos para la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno.



CAPÍTULO II

REVISIÓN DE LITERATURA

1.6. ANTECEDENTES DE LA INVESTIGACIÓN

1.6.1. Antecedentes locales

Paja (2017) consideró 23 variables predictoras; de ellas 17 variables fueron seleccionados por el método stepwise para la asignatura 009-05 con un $R^2= 0.72$, 0.11 de error para los modelos, 0.37 de error para la predicción y 16 variables para la asignatura 010-05 con de $R^2= 0.92$, 0.96 de error para los modelos, 0.0032 de error para la predicción. Mientras que las RNA's Perceptrón multicapa, algoritmo backpropagation, con arquitectura de 3 capas en función de las mismas variables que fueron seleccionadas en regresión múltiple; en donde de ambas asignaturas mencionadas anteriormente obtuvieron promedios de 0.0558 de error para los modelos y 0.00081 de error para las predicciones. Finalmente se midieron los promedios los errores obteniendo que las RNA's es la mejor técnica en predicción; puesto que tienen una diferencia de 0.534 error en modelo y 0.1307 error en predicción respecto a la regresión.

Chavarri (2018) Contó con una población de 1781 estudiantes de quienes se recogió sus datos de la base del sistema académico de la Universidad Peruana Unión Filial Juliaca. Para la implementación de este modelo de predicción se utilizó la regresión logística binaria y se concluyó que los factores determinantes para el riesgo de deserción de estudiantes en la Facultad de Ingeniería y Arquitectura de la Universidad Peruana Unión Filial Juliaca son: edad, cursos desaprobados, ciclo hasta donde avanzo la carrera, promedio ponderado, deuda contraída con la universidad y el riesgo de deserción se calcula con el modelo.



Calla (2019) utilizó la red neuronal mediante los procedimientos de (Kaastra & Boyd, 1996) y como arquitectura de tipo Perceptron Multicapa. La RNA dispuso de seis capas ($M = 6$), una, de entrada, con una neurona ($S1 = 1$), una, de salida con una neurona ($S6 = 10$) y cuatro, ocultas con diferente número de neuronas en cada una ($S2 = 2$, $S3 = 6$, $S4 = 7$, $S5 = 9$). Para determinar el número de neuronas de las capas ocultas se tuvo en cuenta que pocas neuronas no permiten a la RNA realizar un buen ajuste y que demasiadas neuronas pueden provocar overfitting. El error (MAPE) de entrenamiento mínimo fue de 2% y el error (MAPE) de validación mínimo fue de 4%, la RNA seleccionada fue (1, 2, 6, 7, 9, 10). Finalmente se obtuvieron el pronóstico de 12 meses en adelante; presentó un comportamiento estable en el siguiente año 2019 de enero a diciembre.

Sanchez (2017) tuvo como propósito encontrar un modelo estadístico que determine la deserción estudiantil en las escuelas profesionales de la Universidad Nacional del Altiplano – Puno, se trabajó con el muestreo no probabilístico considerando la existencia de las áreas académicas, para lo cual se entrevistó a 325 estudiantes. El diseño de la investigación es no experimental, es decir, es aquella que se realiza sin manipular deliberadamente variables. Solo se describe y diagnostica una situación problema a fin de determinar sus causas y ofrecer alternativas que minimicen la problemática. Para la determinación del modelo se utilizó la estadística multivariada, concretamente la regresión logística binaria, los resultados son: Se determinó y codifico las variables para el modelo estadístico de deserción estudiantil en las escuelas profesionales de la UNA-PUNO. Utilizando la regresión logística binaria. El porcentaje de clasificación correcta utilizando el modelo es del 91.2%. Cuando las variables interactúan en conjunto utilizando la prueba chi-cuadrada o prueba ómnibus. Se determinó los coeficientes significativos del modelo estadístico adecuado para la deserción estudiantil utilizando la regresión logística binaria.



Quispe (2018) en sus resultados los nodos que los modelos presentaron en la capa oculta se variaron entre 4 y 10 para oxígeno disuelto y clorofila respectivamente ambos casos. Las funciones de transferencia de la capa oculta fueron tansig, y en la capa de salida fue lineal o purelin, para contrastar fue realizado un análisis estadístico para la fase de entrenamiento y prueba que permitió asegurar al 95% de nivel de confianza, comprobándose el desempeño de los modelos obtenidos. Finalmente se realizaron las pruebas de validación con los modelos entrenados dando un performance de 0.97% para efectos de oxígeno disuelto y 0.98% para la clorofila alfa.

Levisi (2017) tuvo como objetivo determinar si los Factores socio económicos y personales influirán en la deserción de los estudiantes de la FTS, el método empleado es el hipotético deductivo; el tipo de investigación es explicativa, pues busca el porqué de los hechos, lo cual permitirá determinar los factores de mayor influencia en la deserción universitaria que se presentan la Facultad de Trabajo Social, para este diseño, se tomó una muestra considerando las variables dependientes e independientes, la población, son los estudiantes matriculados en el año 2014, el cual son 590 alumnos, de los cuales se eligió de una muestra de 118, alumnos, siendo 59 los que si desertaron y otros 59 que desertaron pero que retomaron sus estudios. Los resultados nos permiten concluir en lo siguiente: primer objetivo específico, consideramos la tabla N° 02 en el que observamos que el 28% de alumnas que dejaron de estudiar en la FTS, vivían solas al momento de dejar la carrera. Por otro lado, en relación al objetivo específico N° 02, se concluye que la principal razón por la que los estudiantes abandonan la carrera de trabajo social, es por el tema de orientación vocacional tal como lo demuestra la tabla N° 09, con un 39.8%. Otro motivo muy relevante es el que muestra la tabla N° 11, el, que el 34.7% de los estudiantes dejaron de estudiar a causa de un embarazo prematuro. A partir de estos resultados, concluimos que existe influencia significativa de los Factores socio



económicos y personales en la deserción de los estudiantes de la Facultad de Trabajo Social, de la Universidad Nacional del Altiplano de Puno.

1.6.2. Antecedentes nacionales

(Alania, 2018) concluye que el análisis de las técnicas de clasificación se puede aplicar en los datos educativos para predecir el resultado de abandono o no del estudiante a los estudios y conocer la cantidad tentativa de grupos o secciones que se pueden asignar para su planificación. Las pruebas realizadas en el caso de estudio con en la carrera profesional de Ingeniería de Sistemas de la Universidad Nacional Daniel Alcides Carrión, en cuanto a la relación de atributos Abandono y Categoría: demuestran que la mayor cantidad de estudiantes que abandonan la carrera es por las bajas notas. La relación de los atributos Abandono y cursos Aprobados demuestran que la mayor cantidad de estudiantes que abandonan son los han aprobado menos de 24 cursos Aprobados.

Carpio (2017) entre los factores dependientes de los estudiantes que podrían haber influido en la decisión de no continuar sus estudios, la edad, sexo, estado civil, ocupación, procedencia, o problemas familiares no influyeron en la deserción ($p > 0.05$), pero refirieron problemas laborales en 51% de los que abandonaron los estudios, lo que ocurrió en 29% de casos que no abandonaron ($p < 0.05$). En cuanto a los problemas asociados con el docente, hubo percepción de impuntualidad en 38% de los que abandonaron y en 17% en los que no abandonaron ($p < 0.05$). No hubo influencia de la exigencia o las tareas ni con la relación con el docente. Las características de infraestructura, horarios, o materiales didácticos no influyeron en la decisión de retiro, así como tampoco las dificultades con la realización de la tesis de grado. Se concluye que la frecuencia de deserción de los estudios de posgrado es alta, y motivada principalmente por dificultades laborales de los estudiantes.



Guevara (2018) los resultados muestran que la deserción forzosa en los primeros semestres está asociada a factores personales y académicos, y no tanto los factores socioeconómicos e institucionales. Predominan subfactores como problemas de salud, psicológicos y falta de organización por parte del estudiante. El análisis de los subfactores permite plantear que los hábitos académicos de los alumnos no están suficientemente desarrollados para asumir las tareas académicas de la universidad, las cuales requieren de autonomía en la organización del estudio y el desarrollo de habilidades de lectura y redacción. Se presenta una propuesta enfocada en la realización de talleres que fortalezcan sus hábitos de estudio en los que se aborden técnicas de concentración, elaboración de resúmenes y armado de horarios de estudio, y el desarrollo de la capacidad de comprensión lectora.

Ramírez (2017). El contraste Chi cuadrado de la Prueba Ómnibus revela valores significativos ($\chi^2= 28,179$; $gl=2$; $p=0.000$) entre las variables de estudio; lo cual es corroborado por la prueba de bondad de ajuste de Hosmer y Lemeshow ($\chi^2=14,1772$; $gl=8$; $p=0,064$). Estos resultados permiten concluir que los factores individuales y de contexto inciden significativamente en la deserción universitaria de los estudiantes del Programa Beca 18.

Zevallos (2017) Se creó una arquitectura de red neuronal, utilizando el perceptrón multicapa con algoritmo de aprendizaje de resilient backpropagation y función de activación la tangente hiperbólica. La arquitectura de la red neuronal tiene 3 capas ocultas, bias para mejorar la performance de la arquitectura y como parámetro de error de aceptación hasta 2%. Finalmente, la red neuronal artificial obtuvo resultados de un 2% de error absoluto en el entrenamiento y un grado de validez de un 84%.

Zarria et al. (2016) determinó que las variables académicas de ingreso de los estudiantes no resultan significativas para explicar la deserción de primer año. Con estos



resultados, la Universidad podrá generar mejoras en los procesos críticos y en las variables que pudiesen intervenir, haciendo más eficiente su gestión y mejorando el bienestar del estudiante, y, por ende, de la comunidad en la cual se encuentra inmersa.

Piscoya (2016) utilizó la Metodología CRISP DM, como guía para la construcción del modelo de minería de datos basado en series de tiempo logrando realizar las predicciones de deserción escolar en la región de Lambayeque donde solo se tomó como muestra la Ugel de Chiclayo en periodos anuales de manera automatizada dejando de lado el uso de herramientas ofimáticas que retrasan el proceso de los resultados; y usó de la metodología XP para el desarrollo del sistema como solución a la optimización de los procesos mostrando los resultados.

1.6.3. Antecedentes internacionales

Ahmad & Shahzadi (2018) en su investigación su objetivo es predecir la posibilidad de que los estudiantes estén en riesgo (AR) o no "No en riesgo" (NAR), la población de estudio consistió en todos los estudiantes al final del segundo semestre, Tiempo de estudio. Calificaciones previas, ambiente familiar, hábitos de estudio, habilidades de aprendizaje, trabajo duro e interacción académica. Al clasificar a los estudiantes en riesgo o no en riesgo, logró una tasa de clasificación correcta de más del 95% en la muestra de entrenamiento y más del 85% en la muestra reservada. Los modelos estimados se pueden usar para predecir si los estudiantes están en riesgo o no.

Miranda & Guzmán (2017) la deserción de los estudiantes de educación superior genera una serie de inconvenientes que afectan a los estudiantes y las universidades. Los resultados obtenidos a partir de los datos proporcionados por las carreras de Ingeniería de la Universidad Católica del Norte en Antofagasta y Coquimbo (Chile) determinan que las variables que mejor explican la deserción de un estudiante son, las razones socioeconómicas y el puntaje de ingreso a la universidad (PSU). Según el árbol de



decisión construido se concluye que la retención se sitúa en un 78,3%. La calidad de los clasificadores permite asegurar que sus predicciones son correctas, con niveles estadísticos de curva ROC de 76%, 75% y 83% de acierto para los clasificadores de red bayesiana, árbol de decisión y red neuronal respectivamente.

Terán et al. (2019) los resultados finales muestran una ventaja importante de los métodos causales sobre los métodos estadísticos para los casos de prueba. Consideramos que esta ventaja ocurre debido a que el modelo causal aprende los patrones de comportamiento de los datos de entrenamiento de forma independiente en vez de generalizar porcentajes de acreditación. Además de lo anterior, el método estadístico puede presentar problemas importantes al tratar de pronosticar porcentajes de acreditación para situaciones que no se encuentren en los datos de entrenamiento, mientras que el modelo causal utilizará la información aprendida para pronosticar dichas situaciones.

Salgado et al. (2018) en su investigación describe un modelo predicativo de rendimiento académico usando técnicas de redes neuronales sobre un conjunto de datos reales de 300 estudiantes de la carrera de Sistemas de la Universidad Central del Ecuador. Este registro fue suministrado por el entorno virtual de aprendizaje (<https://uvirtual.uce.edu.ec/>) desarrollado en Moodle y usado en dicha Universidad.

Miranda et al. (2020) indica que el estudio de bases de datos escolares es un área que ha sido poco estudiada y cuestionada desde el punto de vista de la minería de datos o la inteligencia artificial, actualmente, existen algunos trabajos que muestran su procesamiento mediante algoritmos de aprendizaje automático o “inteligentes”, sin embargo, no se detienen a analizar la pertinencia de procesar datos cualitativos como si fueran cuantitativos. En este trabajo, se estudia este problema con el uso de tres modelos de red neuronal. Los resultados evidencian la capacidad de estos modelos para clasificar



con un alto grado de acierto tendencias en los estudiantes, utilizando principalmente datos cualitativos.

Bedregal et al. (2020) concluye que al analizar el rendimiento académico de un estudiante no basta con las calificaciones obtenidas, se debe considerar su comportamiento académico, su rendimiento en relación a su cohorte y el ritmo de avance en la aprobación de las asignaturas. Es así que las técnicas empleadas permiten identificar oportunamente a los estudiantes en riesgo académico. Al aplicar técnicas de clasificación: redes neuronales y árboles de decisión se encuentra que las variables más influyentes son el índice de rendimiento exógeno y la relación de créditos aprobados en relación a los créditos que en teoría debieron aprobar.

1.7. MARCO TEÓRICO

1.7.1. La deserción

La **deserción** es la interrupción o desvinculación de los estudiantes de sus estudios, Bedregal (2020) define que es un estado en el que el estudiante se enfrenta a una situación en la que no logra concluir su proyecto educativo. Se refiere al abandono prematuro de un programa de estudios antes de alcanzar el título o grado, y se considera un tiempo suficientemente largo como para descartar la posibilidad que el estudiante se reincorpore (Himmel, 2002).

La educación es considerada como el factor crucial en las economías de los países. El sistema educativo, que cada país adopte, debe de estar diseñado para cubrir las necesidades de desarrollo y aspiraciones de la nación. En otras palabras, la competitividad se logra mediante la educación de la sociedad, principalmente la educación superior y el posgrado, como el grado de especialización de mayor impacto. La educación superior pública es vista como un sistema planeado, organizado, estructurado y dirigido por el Estado, reconocido como un bien público, a través del cual se forman a las personas para



una vida social y laboral, dicha tarea debe responder a indicadores de calidad que aseguren en la medida de lo posible, los macro objetivos que se plantean en el sector, que abonen al desarrollo de la nación y que sean reconocidos por la sociedad (Hernandez et al., 2014).

El estudio de la deserción de la educación superior es extremadamente complejo, pues implica no solo una variedad de perspectivas sino también una gama de diferentes tipos de abandono. Probablemente ninguna definición puede captar en su totalidad la complejidad de este fenómeno universitario (Tinto, 1982).

1.7.2. Tipos de deserción y retención en estudiantes

1.7.2.1. Tipos de Deserción

Según Rodríguez & Londoño (2010) el abandono se clasifica de acuerdo con cuatro variables: el tiempo (Tinto, 1982) el espacio y el carácter volitivo o no-volitivo (Rodríguez & Londoño, 2010).

De acuerdo con la variable tiempo, existe:

1. Deserción precoz: individuo que siendo admitido no se matricula.
2. Deserción temprana: estudiante que abandona sus estudios en los primeros semestres del programa.
3. Deserción tardía: individuo que abandona sus estudios en los últimos semestres.

Con respecto al espacio, propone la siguiente tipología:

1. Deserción institucional: el abandono de la institución.
2. Deserción interna: el estudiante no abandona la institución, pero cambia de programa.

Así mismo, Himmel (2002) indica que es necesario distinguir, además, entre la deserción voluntaria y la involuntaria. La deserción voluntaria puede adoptar la forma de renuncia a la carrera por parte del estudiante o del abandono no informado a la institución

de educación superior. La deserción involuntaria, en cambio, se produce como consecuencia de una decisión institucional, fundada en sus reglamentos vigentes, que obliga al alumno a retirarse de los estudios. En este último caso, la deserción puede estar fundamentada en un desempeño académico insuficiente o responder a razones disciplinarias de diversa índole.

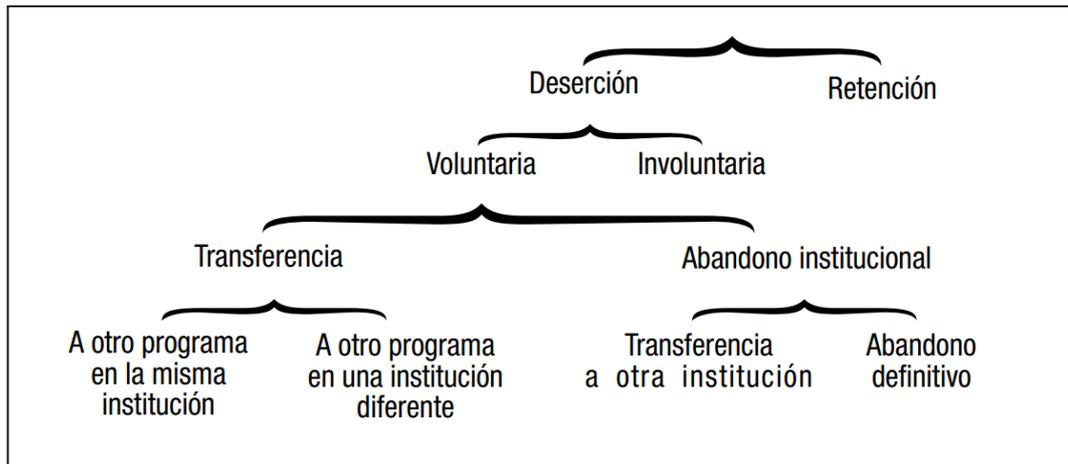


Figura 01. Tipos de deserción de estudiantes.

Fuente: (Himmel, 2002)

Y según (Hernandez et al., 2014) se puede considerar que los principales factores que influyen en el abandono escolar en nivel licenciatura y posiblemente en posgrado pueden ser clasificados de la siguiente forma:



Figura 02. Categorías generales de las causas de la deserción

Fuente: (Hernandez et al., 2014).

1.7.2.2. Estrategias para la retención del estudiante

Según Espinosa et al. (2020) se puede implementar las siguientes estrategias para disminuir la deserción estudiantil.

1. Factores Académicos: Acompañamiento individual al estudiante a través de tutorías para potenciar las condiciones académicas. Asesorías de carácter complementario para aclarar, afirmar o ampliar los conocimientos derivados del proceso de aprendizaje (monitorias). Actividades académicas encaminadas al refuerzo de los conocimientos, habilidades y competencias. Cursos remediales, especiales o de nivelación orientados hacia aquellos estudiantes que reprobaban materias o créditos académicos con el fin de nivelarlos para el siguiente semestre. Cursos de nivelación orientados hacia aquellos estudiantes que deseen adelantar materias o créditos académicos del siguiente semestre.



2. Factores Financieras: Becas y descuentos en el valor de la matrícula por méritos académicos, deportivos o artísticos. Descuentos en el valor de la matrícula por convenios interinstitucionales o por cooperación extranjera. Descuentos en el valor de la matrícula por acuerdos sindicales o con los empleados. Estímulos económicos por participación en actividades curriculares. Financiación directa del valor de la matrícula.
3. Factores Psicológicas: Programas de identificación y seguimiento a estudiantes en conductas de riesgo, como consumo de sustancias psicoactivas, embarazos no planeados y violencia, entre otras. Programas de detección y manejo de las principales características de la salud mental de los estudiantes. Programas para el fortalecimiento de las capacidades y recursos del estudiante en su proceso de formación humana.
4. Gestión Universitaria: Programas de movilidad estudiantil en pregrado y postgrado. Diversificación de ofertas curriculares educativas. Ampliación de las oportunidades de acceso. Diferenciación de las instituciones educativas. Consolidación de la formación por competencias. Promoción de la pertinencia y de la vinculación laboral.

1.7.3. Aprendizaje automático

Existen varias formas de adquirir el conocimiento necesario, una puede ser directamente a partir del humano, o a partir de problemas resueltos previamente. Los datos que se le proporcionan como entrada al sistema permiten que el algoritmo de aprendizaje pueda extraer de ellos la información necesaria para enfrentarse a nuevos datos y realizar la función para la cual fue diseñado. Según el resultado que se desea obtener a través del sistema a todo ello se le denomina aprendizaje automático (Valero, 2018).



Dey (2016) nos dice que el aprendizaje automático se utiliza para enseñar a las máquinas a manejar los datos de manera eficientemente. A veces, después de observar los datos, nos es imposible interpretar o extraer el patrón de los datos que nos genera información. Para ello el aprendizaje automático nos ayuda a procesar los datos para obtener información.

Erickson et al. (2017) sostiene que el aprendizaje automático ahora se está aplicando en muchas áreas además de la informática, y tiene un papel central en tareas como el reconocimiento de voz, la traducción entre idiomas, predicción de la deserción estudiantil, la navegación autónoma de vehículos y las recomendaciones de productos. Algunas de estas tareas no eran factibles anteriormente ahora los avances recientes en el aprendizaje automático los han hecho posibles.

Lo que hace Machine Learning es crear un conjunto de reglas a partir del conjunto de entradas evitando así que se tengan que definir las reglas a mano (Roche, 2014)

El Machine Learning, extrae conocimiento a partir de los datos para gradualmente mejorar el funcionamiento de los modelos predictivos y tomar decisiones dirigidas hacia un objetivo. Hoy en día, se está empleando el ML en el mundo en todas las ciencias dirigiéndonos rápidamente hacia una mejora del conocimiento y desarrollo (Marin, 2019).

1.7.3.1. Aprendizaje automático no supervisado

Dey (2016) Nos dice que los algoritmos de aprendizaje no supervisados aprenden pocas características de los datos. Cuando se introducen nuevos datos, utiliza las características previamente aprendidas para reconocer patrones desconocidos en los datos. Se utiliza principalmente para agrupamiento y reducción de características de dimensiones muy altas entre estas tenemos los siguientes algoritmos que no detallaremos porque no forman parte del estudio.

- Análisis de componentes principales



- K –medias
- Análisis clúster

1.7.3.2. Aprendizaje automático supervisado

Los algoritmos de aprendizaje automático supervisados son los algoritmos que necesitan ayuda externa. El conjunto de datos se divide un conjunto de datos de entrenamiento y prueba que tienen una variable de salida que debe predecirse un valor continuo o un valor de clase. Todos los algoritmos aprenden algún tipo de patrones del conjunto de datos de entrenamiento y aplicarlos al conjunto de datos de prueba (Dey, 2016).

Y Marin (2019) El objetivo principal del aprendizaje supervisado es el entrenar un modelo predictivo a partir de datos que han sido previamente etiquetados. Gracias al entrenamiento, el algoritmo obtendrá la información necesaria para hacer predicciones cuando se le suministre nuevas instancias. El término supervisado, hace referencia a que empleamos un grupo de muestras cuya señal de salida (output) previamente ya conocemos (marcadas o clasificadas). Dependiendo de la naturaleza de la variable que queremos predecir (variable dependiente) nos enfrentaremos a un problema de clasificación o de regresión. Cuando la variable dependiente o variable a predecir es categórica, se trata de un problema de clasificación, mientras que si la variable es continua estaremos ante un problema de regresión.

Entre los más destacados algoritmos de aprendizaje tenemos los siguientes:

- Naïve Bayes
- Árboles de decisión
- Redes neuronales
- Regresión lineal
- Métodos Boosting

Los métodos de clasificación y regresión son tipos propios del paradigma de aprendizaje supervisado.

1. **Regresión:** Al aplicar un algoritmo de regresión lo más frecuente es que tanto la variable objetivo sea de característica sea continua, en función a otras variables que pueden ser cuantitativas o cualitativas (Rodríguez & Mora, 2001).

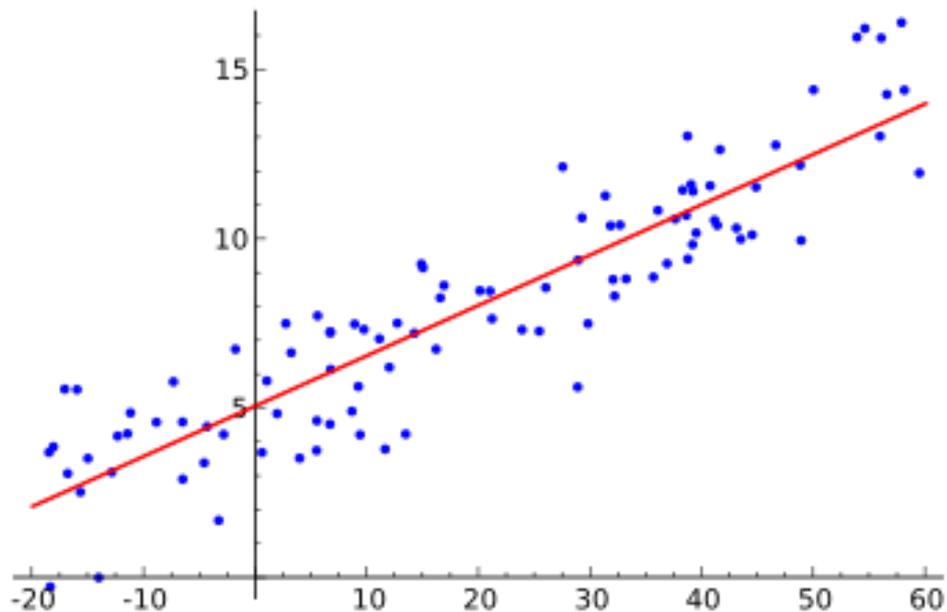


Figura 03. Diagrama de regresión.

Fuente: (Ahmad & Shahzadi, 2018)

2. **Clasificación:** Dado un conjunto de ejemplos de entrenamiento (de muestras) podemos etiquetar las clases y entrenar un algoritmo supervisado para construir un modelo que prediga la clase de una nueva muestra, la diferencia entre algoritmos de clasificación y regresión viene determinada por la naturaleza de la variable dependiente. Si ésta es cualitativa se denominan como de clasificación (Marín, 2019).

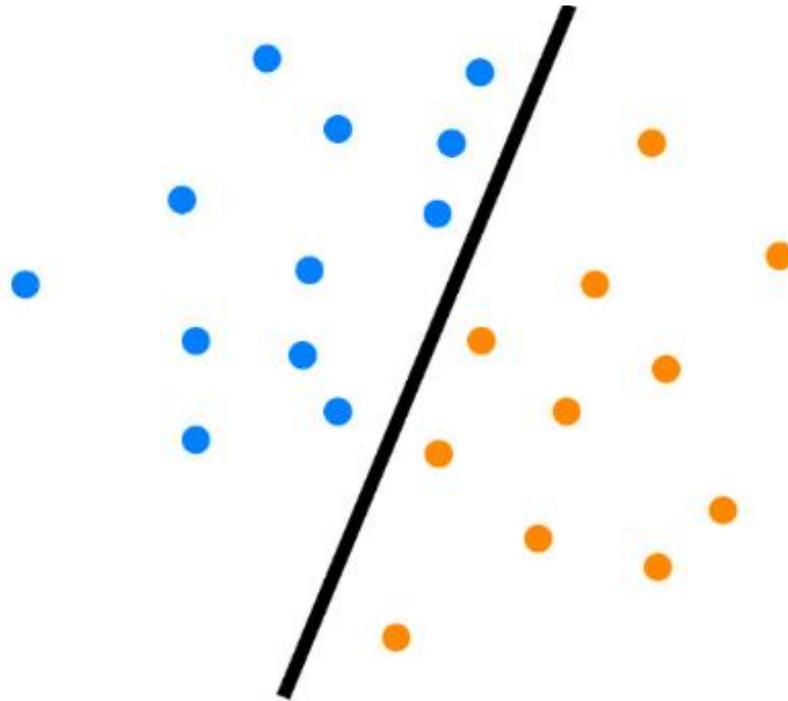


Figura 04. Diagrama de clasificación.

Fuente: (Ahmad & Shahzadi, 2018)

1.7.4. Red neuronal artificial

Las redes de neuronas artificiales se han convertido en la familia de algoritmos de machine learning (ML) más populares de los últimos tiempos, son un paradigma de aprendizaje automático inspirado en la forma en que funciona el sistema nervioso biológico. Según Matich (2021) es un sistema de computación compuesto por un gran número de elementos simples, elementos de procesos muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas. La RNA se trata de una familia de algoritmo muy potentes con el cual podemos modelar comportamientos inteligentes y se trata de un sistema de interconexión de neuronas que colaboran entre sí para producir un estímulo de salida.

Ravi et al. (2017) la red neuronal es generalmente utilizado para clasificación y regresión con éxito en muchas áreas de la ciencia, permite expresar hipótesis complejas (no lineales) el entrenamiento no es trivial porque una vez que los errores se propagan hacia atrás a las primeras capas, se vuelven minúsculos.

Así también son capaces de extraer patrones y detectar tramas que son muy difíciles de apreciar por el ser humano u otras técnicas computacionales. Una de las principales ventajas de utilizar las RNA es la capacidad de aprender a realizar tareas basadas en un entrenamiento o en una experiencia inicial (Matich, 2021).

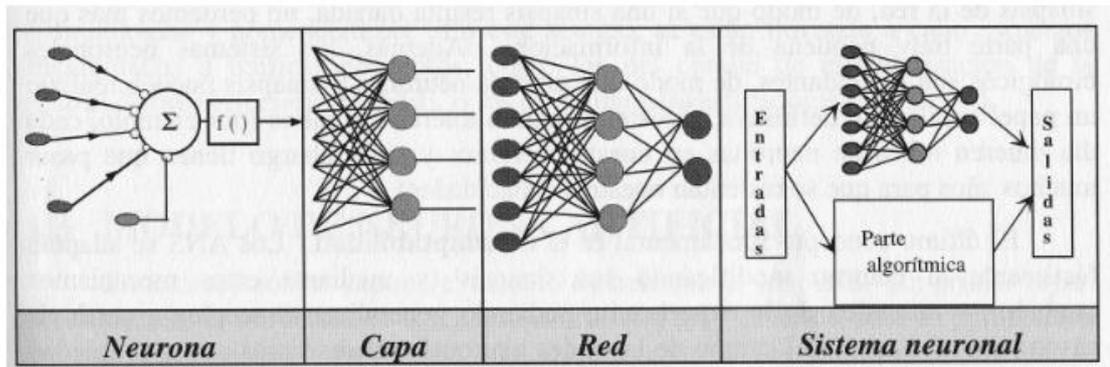


Figura 05. Sistema global de proceso de una red neuronal

Fuente: (Larrañaga et al., 2016)

1.7.4.1. Elementos que componen una Red Neuronal Artificial

En la figura 06 vemos un esquema de una red neuronal, la misma está constituida por neuronas interconectadas y arregladas en tres capas (esto último puede variar). Los datos ingresan por medio de la “capa de entrada”, pasan a través de la “capa oculta” y salen por la “capa de salida”. Cabe mencionar que la capa oculta puede estar constituida por varias capas (Matich, 2021).

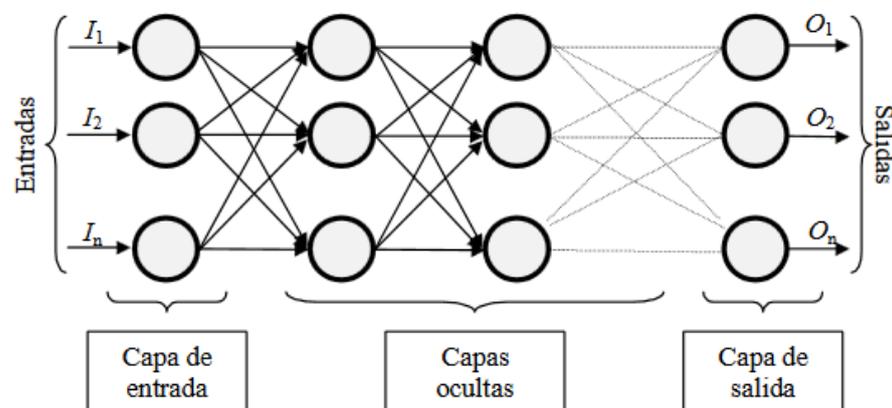


Figura 06. Ejemplo de una red neuronal totalmente conectada

Fuente: (Matich, 2021)



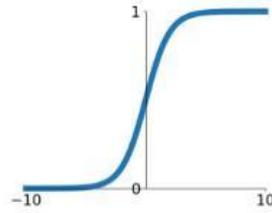
Función de entrada: Los valores de entrada se multiplican por los pesos anteriormente ingresados a la neurona. Por consiguiente, los pesos que generalmente no están restringidos cambian la medida de influencia que tienen los valores de entrada. Es decir, que permiten que un gran valor de entrada tenga solamente una pequeña influencia, si estos son lo suficientemente pequeños (Matich, 2021).

Función de activación: Matich (2021) dice que calcula el estado de actividad de una neurona; transformando la entrada global (menos el umbral) en un valor (estado) de activación, cuyo rango normalmente va de (0 a 1) o de (-1 a 1), porque una neurona puede estar totalmente inactiva (0 o -1) o activa (1). Las Redes neuronales dependiendo de la aplicación específica utiliza una función de activación, como la activación sigmoidea puede permitir a la red aprender más rápido, generalizar mejor, o ambas. Una de las alternativas para la unidad sigmoidea es la función de activación tanh, sin embargo, las más utilizadas actualmente son la función de activación Relu, Logística o sigmoidea (Repetur, 2019).

1. Función lineal
2. Función sigmoidea: Los valores de salida que proporciona esta función están comprendidos dentro de un rango que va de 0 a 1. Al modificar el valor de g se ve afectada la pendiente de la función de activación (Matich, 2021).
3. Función tangente hiperbólica: Los valores de salida están comprendidos dentro de -1 a 1.
4. Función ReLU y otros.

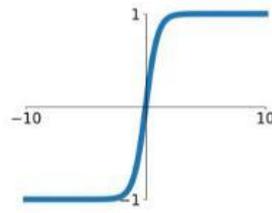
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



tanh

$$\tanh(x)$$



ReLU

$$\max(0, x)$$

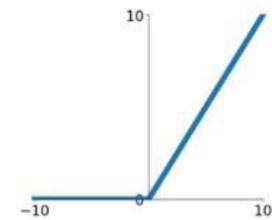


Figura 07. Funciones de activación.

Fuente: (Repetur, 2019).

Función de salida: El último componente que una neurona necesita es la función de salida. El valor resultante de esta función es la salida de la neurona; por ende, la función de salida determina que valor se transfiere a las neuronas vinculadas. Si la función de activación está por debajo de un umbral determinado, ninguna salida se pasa a la neurona subsiguiente. Normalmente, no cualquier valor es permitido como una entrada para una neurona, por lo tanto, los valores de salida están comprendidos en el rango $[0,1]$ o $[-1,1]$. También pueden ser binarios $\{0, 1\}$ o $\{-1,1\}$ (Matich, 2021)

1.7.4.2. Optimización de la red neuronal.

Repetur (2019) al trabajar con redes neuronales lo que realmente nos interesa optimizar durante el entrenamiento es alguna medida P que representa el rendimiento de la red. El inconveniente que surge es que P suele ser intratable y depende del conjunto de prueba, no del conjunto de entrenamiento. En consecuencia, lo que se hace es optimizar P de manera indirecta, recurriendo a otra función costo L , con la esperanza de que esto también mejore P .

El objetivo de un algoritmo de aprendizaje es reducir el error de generalización, que podemos escribir como:

$$R(\theta) = E_p[L(f(x, \theta), y)]$$

Donde p es la distribución de los pares (x, y) , L es la función que cuantifica la pérdida para cada par y $f(x, \theta)$ es la salida de la red. Esta cantidad no es más que la pérdida en la que se incurre para un determinado valor de los parámetros θ y es por esto que también se suele llamar riesgo. Generalmente no conocemos p , sino que tenemos un conjunto de entrenamiento, por lo que minimizamos el costo sobre este conjunto en vez de sobre toda la distribución, con lo cual ahora estamos minimizando el riesgo empírico.

$$J(\theta) = E_{\hat{p}}[L(f(x, \theta), y)] = \frac{1}{n} \sum_{k=1}^n L(f(x_k, \theta), y_k),$$

Y esperamos que esto a su vez nos permita disminuir $R(\theta)$ de manera significativa. Sin embargo, este enfoque es propenso al overfitting, ya que modelos con alta capacidad pueden simplemente memorizar el conjunto de entrenamiento. Además, muchas técnicas efectivas de optimización se basan en descenso por gradiente, mientras que funciones como la que usamos para cuantificar la exactitud de la red no tienen derivadas útiles. Debido a esto es raro que se recurra a la minimización directa del riesgo empírico en redes neuronales, sino que se utilizan funciones sustituto para la pérdida, como pueden ser el costo cuadrático, la entropía cruzada, o el negative log-likelihood.

1.7.4.3. Algoritmo de RetroPropagacion (Rprop)

El algoritmo Rprop emplea un esquema basado en el signo para actualizar los pesos con el fin de eliminar las influencias perjudiciales de la magnitud de las derivadas en las actualizaciones de los pesos, es decir, la dirección de la actualización a lo largo de cada dirección de peso sólo depende del signo de la derivada correspondiente. Este

enfoque se considera eminentemente adecuado para aplicaciones en las que el gradiente se estima numéricamente o el error es ruidoso (Igel & Husken, 2003).

El tamaño del paso de actualización a lo largo de una dirección de peso está determinado exclusivamente por un peso específico " valor de actualización "

$$\Delta w_{ij}^k = \begin{cases} -\Delta_{ij}^k \text{ if } \frac{\partial E(w^k)}{\partial w_{ij}} > 0, \\ +\Delta_{ij}^k \text{ if } \frac{\partial E(w^k)}{\partial w_{ij}} < 0, \\ 0 \text{ otherwise,} \end{cases}$$

Donde $\partial E(W^k) / \partial w_{ij}$ denota la derivada parcial del error del lote con respecto a w_{ij} en la iteración k . El segundo paso del aprendizaje Rprop es determinar los nuevos valores de actualización.

La suposición de Rprop de que un cambio de signo de la derivada parcial implica un salto

Sobre un mínimo local a lo largo de la dirección del peso w_{ij} no tiene en cuenta si la actualización del peso ha causado un aumento o una disminución del error del lote. Obsérvese que un cambio de signo de la derivada parcial también podría implicar un salto sobre un máximo local. También la estrategia de Rprop de aumentar ligeramente el valor de la actualización cuando las derivadas conservan su signo puede acelerar la convergencia en regiones poco profundas cuando las derivadas son *negativas*, pero puede ser ineficiente cuando las dos derivadas son positivas. En el caso de las derivadas positivas, las grandes actualizaciones del peso pueden llevar la trayectoria del peso lejos del mínimo o en regiones con valores más altos de la función de error de la función de error. Este tipo de problemas se trata empíricamente empleando dos parámetros heurísticos D_{max} y D_{min} , que actúan como condiciones de restricción del tamaño de los pasos de actualización (Anastasiadis et al., 2005).

1.7.4.4. Tasa de aprendizaje y número de iteraciones

La actualización de los parámetros θ depende de la tasa de aprendizaje (learning rate) valores que suelen ser 0.1,0.3,0.5; entre más pequeño sea menor será el gradiente por tanto se requerirá un mayor número de iteraciones (n_iters) por lo que debe configurarse en cada caso específico (Anastasiadis et al., 2005).

1.7.4.5. Métricas de clasificación

Consideramos las métricas que evalúan el rendimiento de los modelos en términos de su (estimado) tasas de verdaderos y falsos positivos. La Tabla 01 contiene fórmulas para las principales métricas consideradas (Flach, 2003).

Las fórmulas se pueden verificar sustituyendo:

$c = \text{pegativos} / \text{positivos}$,

$tpr = \text{verdaderos positivos} / \text{positivos}$

$fpr = \text{falsos positivos} / \text{negativos}$

Tabla 01. Métricas de clasificación.

Métrica	Formula
Accuracy	$\frac{tpr + c(1 - fpr)}{1 + c}$
Recall	$\frac{tpr}{tpr + c \cdot fpr}$
F1-Score	$\frac{2tp}{tpr + c \cdot fpr + 1}$

Fuente: (Flach, 2003).

Valores que varían entre 0-1 (Flach, 2003):



- Accuracy es la precisión de la clase 0.
- Recall es la precisión de la clase 1 (comúnmente la clase objetivo).
- F-Score es el promedio armónico entre Accuracy y Recall.

1.8. TÉRMINOS UTILIZADOS

Parámetro: Es aquella característica de un algoritmo que dependiendo de la variación de este obtendremos resultados distintos ya sea incrementando su precisión y costo computacional entre otros (Aguado, 2015).

Deserción estudiantil: Se considera deserción estudiantil cuando un estudiante por diversos motivos internos o externos ya no puede continuar sus estudios (Himmel, 2002).



CAPÍTULO III

MATERIALES Y MÉTODOS

1.9. TIPO DE INVESTIGACIÓN

El presente trabajo de investigación se desarrolló bajo el planteamiento metodológico del enfoque cuantitativo. El enfoque cuantitativo utiliza la recolección y el análisis de datos para contestar preguntas de investigación y probar hipótesis establecida previamente y confía en la mediación numérica, el conteo y frecuentemente en el uso de la estadística para establecer con exactitud patrones de comportamiento en una población (Hernández et al., 2003).

El tipo de investigación utilizada fue la Aplicada que consiste en la utilización de los conocimientos en la práctica; con un nivel de investigación descriptivo el cual describe las características de un fenómeno existente, recolecta datos que muestren un evento, hecho, contexto o situación que ocurre. Y la clasificación de la investigación por su naturaleza es de tipo documental por los medios para obtener los datos, puesto que es lo que mejor se adapta a las características y necesidades de la investigación.

El tipo de investigación documental se realizó apoyándose en fuentes documentales (primarias y secundarias) obtenidas a través de fuentes bibliográficas (libros), hemerográficas (artículos o ensayos en revistas y periódicos) o archivísticas (documentos de primera mano cómo cartas, oficios, circulares, expedientes, etcétera) los recursos donde se obtiene la información requerida (Prieto, 2017).

1.10. DISEÑO DE INVESTIGACIÓN

Dado que el objetivo del estudio es determinar la Red Neuronal Artificial que proporcione un buen ajuste para predecir la deserción en estudiantes de la Maestría en



Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno, el tipo de investigación es no experimental que se basa en la observación de los fenómenos en su entorno natural; de corte transversal que se utiliza para observar y analizar un momento exacto de la investigación. Y también descriptiva para conocer a detalle cual es el comportamiento de las variables que incitan a los estudiantes de la Maestría en Derecho de la UNA Puno a desertar sus estudios

Del enfoque cuantitativo se tomará la técnica de redes neuronales para establecer un modelo de predicción de deserción de estudiantes de la Maestría en Derecho.

1.11. POBLACIÓN Y MUESTRA

1.11.1. Población

La población es el conjunto de personas u objetos de los que se desea conocer algo en una investigación (Lopez, 2004).

La población de estudio es un conjunto de casos, definido, limitado y accesible, que formará el referente para la elección de la muestra, y que cumple con una serie de criterios predeterminados (Arias et al., 2016).

La población de estudio para el presente trabajo de investigación estuvo conformada por el registro de datos extraídos de las carpetas de ingreso a lo largo del tiempo de los alumnos de la Maestría en Derecho de la UNA Puno.

1.11.2. Muestra

Es un subconjunto de la población. Así mismo, Canales et al. (1994) señalan que es “un subconjunto de la población en que se llevara a cabo la investigación con el fin posterior de generalizar los hallazgos del todo”.

En el presente de investigación se utilizó el muestreo no probabilístico, en el cual, los elementos no depende de la probabilidad, sino de causas relacionadas con las



características de la investigación (Hernández et al., 2014) y según Canales et al. (1994) “se toman los casos o unidades que estén disponibles en un momento dado”. Aquí el procedimiento no es mecánico ni se basa en fórmulas de probabilidad, sino que depende del proceso de toma de decisiones de un investigador o de un grupo de investigadores y, desde luego, las muestras seleccionadas obedecen a otros criterios de investigación (Hernández et al., 2014). Y el método de muestreo no probabilístico utilizado fue muestreo intencional o de conveniencia que, según Arias et al., (2016) “consiste en la selección por métodos no aleatorios de una muestra cuyas características sean similares a las de la población objetivo. También puede ser que el investigador seleccione directa e intencionadamente los individuos de la población. El caso más frecuente de este procedimiento es utilizar como muestra los individuos a los que se tiene fácil acceso”.

La muestra estuvo conformada por un total de 334 registros de datos extraídos de las carpetas de ingreso de los alumnos durante los años 2015 – 2019, elegidas a criterio por la disponibilidad de los datos.

1.12. RECOLECCIÓN DE DATOS

1.12.1. Técnica

La técnica de recolección de datos es documental y esta se realiza apoyándose en fuentes documentales obtenidas a través de fuentes archivísticas los recursos donde se obtiene la información requerida,

1.12.2. Instrumento

El instrumento utilizado fue el denominado “documento de archivo”. Un documento de archivo es un instrumento de carácter contemporáneo que es creado por individuos y organizaciones en el desarrollo de sus actividades; con el transcurrir del tiempo estos documentos se convertirán en una ventana que nos permitirá tener acceso a



los eventos ocurridos en el pasado (International Council on Archives, 2016), que, en este caso sería la ficha de inscripción (Anexo 03).

1.12.3. Método de recolección de datos

Para esta investigación los datos a usar fueron proporcionados por la Maestría en Derecho de la UNA Puno en un archivo Excel a partir de la información de las fichas de inscripción extraídas de las carpetas de ingreso de los alumnos.

1.13. VARIABLES DE ESTUDIO

Las variables en un estudio de investigación constituyen todo aquello que se mide, la información que se colecta o los datos que se recaban con la finalidad de responder las preguntas de investigación, las cuales se especifican en los objetivos. (Villasís & Miranda, 2016)

Según Chen et al. (2020) es muy importante seleccionar la entrada adecuada de los datos, las variables se eligen en función al dominio del conocimiento por personas capacitadas en el tema en específico o de cualquier manera arbitraria pero que se garantice una asociación intuitiva de las variables X con respecto a Y.

Se toman principalmente tres puntos a considerar para seleccionar las variables de entrada.

1. El número de variables disponibles.
2. Correlaciones entre posibles variables cuantitativas de entrada, lo que crea redundancia.
3. Variables que tienen poco o ningún poder predictivo.

Si se considera estos tres puntos para analizar sus posibles variables de entrada, entonces puede decidir cuáles son las variables óptimas para el dominio de su problema.



Y tomando en cuenta lo expuesto por Bedregal et al., (2020) analizar el comportamiento de un estudiante no basta con las calificaciones obtenidas, se debe considerar su comportamiento académico, su rendimiento en relación a su cohorte y el ritmo de avance en la aprobación de las asignaturas. Es así que las técnicas empleadas permiten identificar oportunamente a los estudiantes en riesgo académico.

También se tomaron las variables en estudio según la opinión de los trabajadores de la Maestría en Derecho de la UNA Puno, así como la disponibilidad, y que tuvieran relación con la deserción estudiantil; así mismo tomando en cuenta los antecedentes citados.

1.13.1. Variable Dependiente

Y: Deserción estudiantil

- 0: Continúa estudiando
- 1: Desertó por diversos motivos

1.13.2. Variables Independientes.

X1: Promedio de notas

X2: Edad, valor numérico entero mayor que 0

X3: Sexo del estudiante

- 0: Mujer
- 1: Hombre

X4: Estado civil del estudiante



- 0: Soltero
- 1: Conviviente
- 2: Casado
- 3: Divorciado

X5: Procedencia del estudiante

- 0: Proviene de la provincia de puno
- 1: Proviene de otro lugar que no es la provincia de puno

X6: Estado de cuenta

- 0: debe
- 1: No debe.

X7: Tipo de universidad

- 0: Privado
- 1: Estatal

1.14. PROCESAMIENTO DE DATOS

1.14.1. Metodología de Redes Neuronales

1.14.1.1. Selección de Variables:

Detectar cuales son aquellas variables que influyen significativamente para que un estudiante pueda desertar sus estudios.



1.14.1.2. Pre procesamiento de Datos

Luego de la recopilación de datos se realiza exploración de datos.

1.14.1.3. Patrones de entrenamiento y prueba:

La finalidad de un modelo es predecir la variable respuesta en observaciones futuras que el modelo no haya visto antes. El error mostrado por defecto tras entrenar un modelo suele ser el error de entrenamiento, el error que comete el modelo al predecir las observaciones que ya ha visto. Si bien estos errores son útiles para entender cómo está aprendiendo el modelo, no es una estimación realista de cómo se comporta el modelo ante nuevas observaciones. Para conseguir una estimación más certera, se tiene que recurrir a un conjunto de test o emplear estrategias de validación basadas en resampling.

Los métodos de validación, también conocidos como resampling, son estrategias que permiten estimar la capacidad predictiva de los modelos cuando se aplican a nuevas observaciones, haciendo uso únicamente de los datos de entrenamiento. La idea en la que se basan todos ellos es la siguiente: el modelo se ajusta empleando un subconjunto de observaciones del conjunto de entrenamiento y se evalúa (accuracy) con las observaciones restantes. Este proceso se repite múltiples veces y los resultados se agregan y promedian. Gracias a las repeticiones, se compensan las posibles desviaciones que puedan surgir por el reparto aleatorio de las observaciones. La diferencia entre métodos suele ser la forma en la que se generan los subconjuntos de entrenamiento/validación.

Dado un conjunto de muestras (muestras), queremos:

- Aprender: el modelo w
- Evaluar el clasificador C_w o el detector c_{mplet}

Sin embargo:



Usar las muestras para aprender un modelo y evaluar su clasificador asociado solo nos sirve para obtener el error de clasificación, de hecho, lo normal es que, precisamente, ya sea ese error el usado (minimizado) para aprender el modelo.

Lo que nos interesa es conocer la capacidad de generalización de C_w y para eso los conjuntos de entrenamiento y evaluación han de ser diferentes.

1. Método Hold-out

- Entrenar: A% de las muestras
- Evaluar: (100-A%) de las muestras
- Usual A=80%

La lección de la muestra para entrenar/evaluar es aleatoria. Podemos hacerlo k veces para promediar las evaluaciones.

Es importante que el conjunto de datos de prueba tenga un volumen suficiente como para generar resultados estadísticamente significativos, y a la vez, que sea representativo del conjunto de datos global.

Normalmente el conjunto de datos se suele repartir en un 80% de datos de entrenamiento y un 20% de datos de test, pero se puede variar la proporción según el caso. Lo importante es ser siempre conscientes de que hay que evitar el sobreajuste u “overfitting”.

Sobreajuste u «overfitting»

El sobreajuste ocurre cuando un modelo está “sobre-entrenado”. Son modelos complejos que se ajusta tan milimétricamente al conjunto de datos a partir del cual se han creado, que pierden gran parte de su poder predictivo, y ya no son útiles para otros conjuntos de datos. Esto se debe a que los datos siempre tienen cierto grado de error o imprecisión, e intentar ajustarse demasiado a ellos, complica el modelo inútilmente al mismo tiempo que le resta utilidad. Lo entenderemos mejor con un ejemplo.



2. Método Cross-validation (Validación cruzada)

Patrones de Entrenamiento:

Para esta investigación estuvo conformado por el 80% de los datos de la serie. Se seleccionan en forma consecutiva y ordenada, este conjunto de datos es el que se utilizó para el entrenamiento de la red neuronal.

Patrones de Prueba:

Para esta investigación estuvo conformado por el 20% de los datos, que corresponden a los datos restantes, una vez que se han seleccionado los patrones de entrenamiento. Este conjunto de datos se utilizó para evaluar la capacidad de generalización o predicción de la red.

1.14.1.4. Arquitectura de la Red Neuronal Artificial

- N° de capas ocultas
- N° de nodos en la capa oculta
- N° de salidas
- Selección de los pesos iniciales: Escoger los pesos iniciales puede ser crucial y es recomendable probar con diferentes conjuntos de valores iniciales para tratar de obtener buenos resultados.
- Determinación del número de nodos de la capa oculta: Se asume que el valor inicial del número de nodos de la capa oculta sea igual al promedio entre el número de entradas y salidas (si el valor obtenido es decimal se redondea). Pueden realizarse pruebas por ensayo y error, agregando más nodos y comparando los errores de ajuste y predicción.
- Algoritmo de Entrenamiento: Retro propagación (Back Propagation)
- Entrenamiento de la RNA seleccionada: Para entrenar la red es necesario establecer los siguientes parámetros: El número máximo de ciclos y el error



permitido de convergencia se fijarán por ensayo y error. Tasa de aprendizaje, incremento de la tasa de aprendizaje y momento pueden ser fijados en 0.05, 1.05 y 0.95 respectivamente. Es conveniente realizar pruebas cambiando estos valores y evaluando el comportamiento de los errores de entrenamiento y generalización.

Una vez definida la RNA, con su ecuación se generó los valores ajustados o producida por la red, utilizando los patrones de entrenamiento. Luego se calcula el error de entrenamiento.

1.14.1.5. Criterio de Evaluación

Se refiere al rendimiento que se utiliza para indicar que tan bien aprendió la red neuronal artificial, estos resultados se calcula con el valor pronosticado y el valor real.

Comparación entre RNA con diferentes entradas: Mediante el error de entrenamiento y de generalización se comparan las RNA generadas y se selecciona aquella en la que ambos valores sean mínimos. No es conveniente que el error de entrenamiento sea muy pequeño en comparación con el error de generalización, pues esto indica un sobreajuste o memorización. La correlación entre los valores originales de la serie y los estimados por la RNA puede usarse como una medida de la exactitud de la predicción.

1.14.1.6. Implementación de la Red Neuronal para el Pronóstico:

La tabulación de los datos se obtendrá mediante el procedimiento Perceptrón Multicapa (MLP), que genera un modelo predictivo para una o más variables dependientes basada en los valores de las variables predictoras, para ello se utilizará el programa Microsoft Office Excel, el software estadístico R y phyton.

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

En este apartado se exponen los resultados obtenidos, seguidamente de su interpretación y análisis. Utilizando tablas, gráficos de histograma, barra, entre otros para mostrar los resultados y facilitar su comprensión detallado por objetivo.

1.15. VARIABLES QUE CONTRIBUYEN AL RIESGO DE DESERCIÓN.

Según al objetivo específico 1 que nos propusimos identificar el comportamiento de las variables que contribuyen al riesgo de deserción del estudiante durante su periodo de estudios.

Tabla 02. Resumen descriptivo

Variable	Muestras	Media	Desviación	Mediana	Mínimo	Máximo	Rango
Promedio de notas	334	14.09173653	2.785047373	14.7	3	19	16
Edad	334	31.04491018	4.207931288	31	24	50	26
Sexo	334	0.571856287	0.495552136	1	0	1	1
Estado civil	334	0.97005988	0.823273694	1	0	3	3
Procedencia	334	0.368263473	0.483057059	0	0	1	1
Estado de cuenta	334	0.613772455	0.487614302	1	0	1	1
Tipo de universidad	334	0.502994012	0.50074121	1	0	1	1
Deserción	334	0.24251497	0.42924715	0	0	1	1

Fuente: Elaboración Propia

En la tabla 02 se observa el resumen descriptivo de las variables en estudio de los estudiantes de la Maestría en Derecho de la UNA Puno como la media aritmética (teniendo en cuenta que muchas de las variables son categóricas), mediana, desviación estándar, mínimo valor obtenido y por últimos el rango.

Tabla 03. Deserción Estudiantil

Deserción estudiantil	Estudiantes	%
Continúa estudiando	253	75.74%
Desertó por diversos motivos	71	24.26%
Total	334	100.00%

Fuente: Elaboración Propia

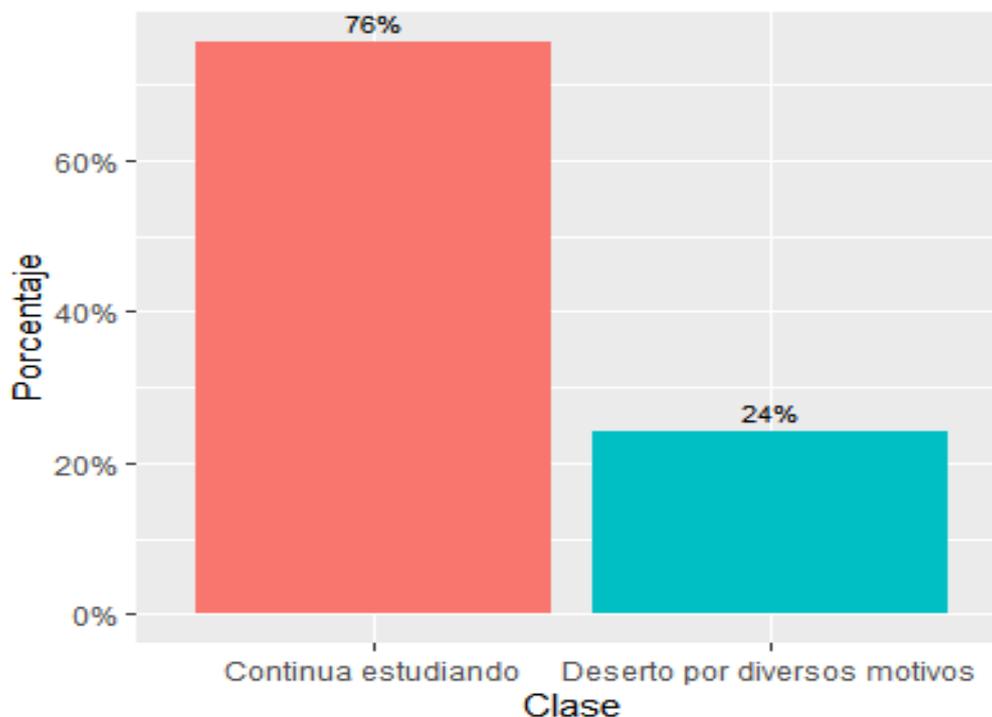


Figura 08. Distribución de la deserción estudiantil.

Fuente: Elaboración Propia

En la figura 08 observamos que el 76% de los estudiantes de la Maestría en Derecho de la UNA Puno continúan en sus estudios y que solo el 24% de ellos desertan, resultados de los cuales se puede inducir que aproximadamente 1 de cada cuatro estudiantes deserta por diversos motivos y el algoritmo a emplearse debería poder captar este patrón.

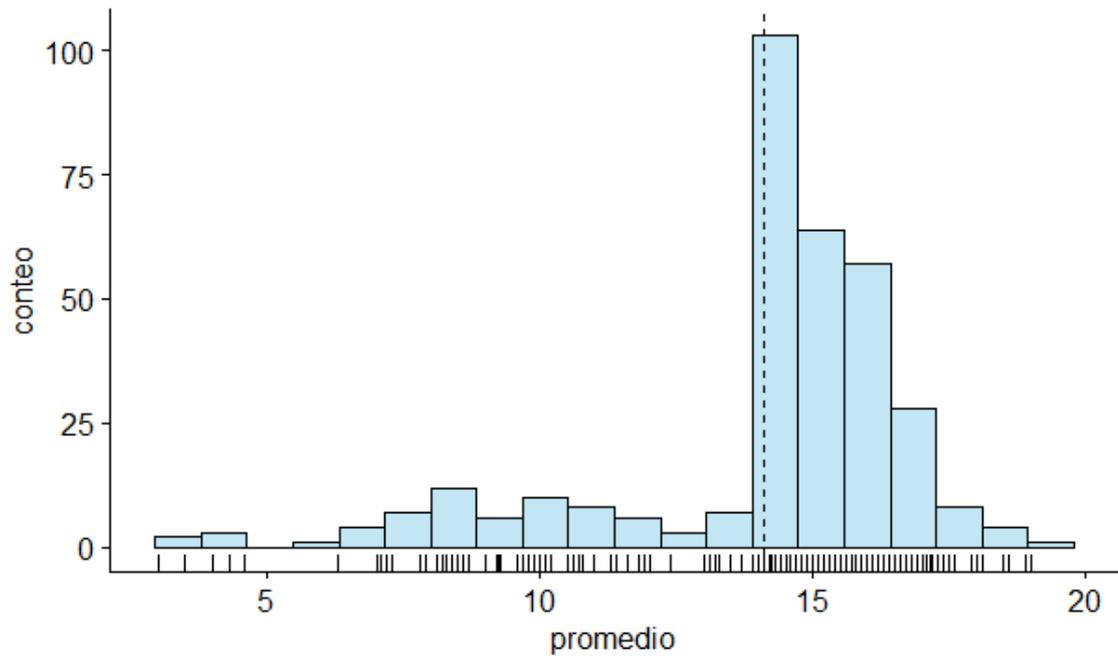


Figura 09. Histograma del promedio de notas de los alumnos.

Fuente: Elaboración Propia

De la figura 09 se detalla el histograma que la calificación promedio de los estudiantes de la Maestría en Derecho de la UNA tiende a una mayor concentración de información mayor a 14 muy cercano al promedio aritmético de 14.7, gráficamente sugiere una distribución normal.

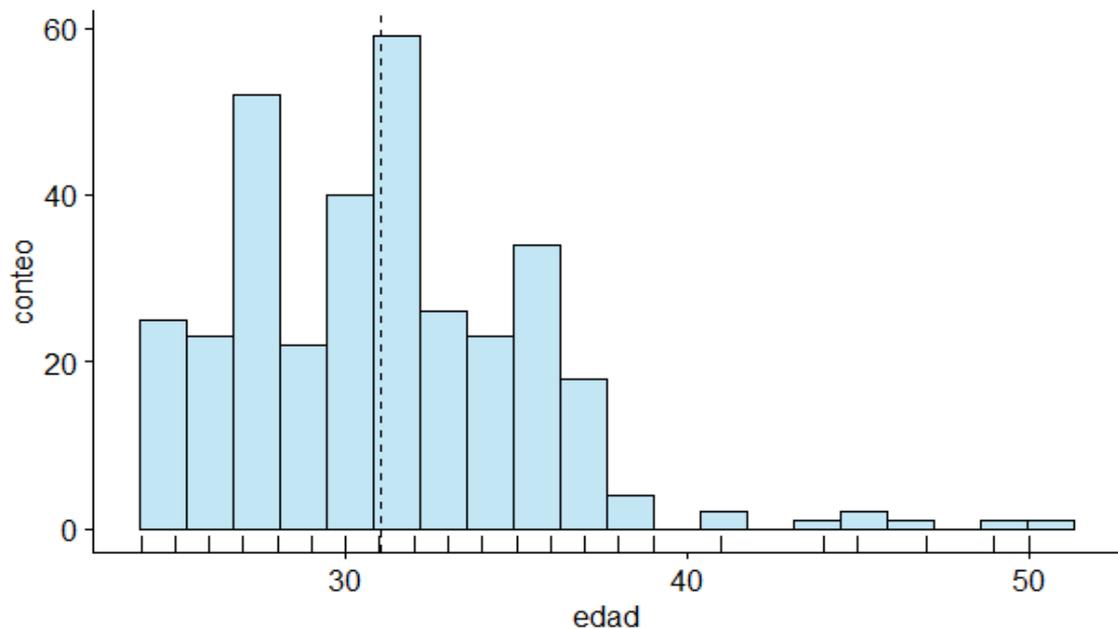


Figura 10. Histograma de la edad que cuentan los alumnos.

Fuente: Elaboración Propia

En la figura 10 observamos el histograma de edad de los estudiantes de la Maestría en Derecho de la UNA están próximos una gran mayoría al promedio de 31 años y algunos estudiantes mayores a los 40 años, variable que tiende a una distribución normal lo que beneficia la empleabilidad de las redes neuronales ya que no tendrá inconvenientes.

Tabla 04. Sexo de los estudiantes de la Maestría en Derecho de la UNA

Sexo	Estudiantes	%
Hombres	191	57.19%
Mujeres	143	42.81%
Total	334	100.00%

Fuente: Elaboración Propia

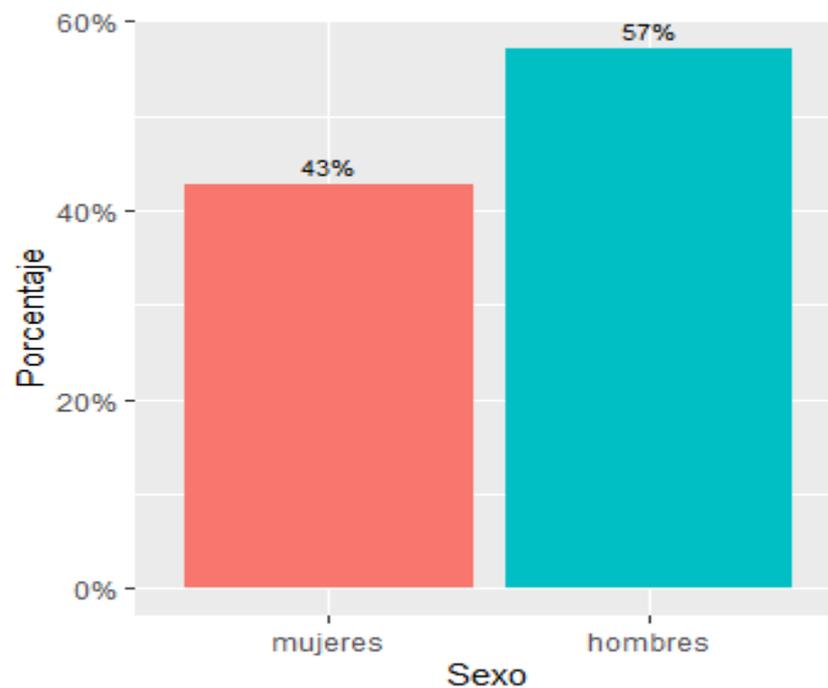


Figura 11. Sexo de los estudiantes de la Maestría en Derecho de la UNA.

Fuente: Elaboración Propia

En la figura 11 observamos un gráfico de barras del cual se observa que el 57% de los estudiantes de la Maestría en Derecho de la UNA fueron hombres y el 43% mujeres.

Tabla 05. Estado de cuenta de los estudiantes de la Maestría en Derecho de la UNA

Estado de cuenta	Estudiantes	%
Debe	129	38.62%
No debe	205	61.37%
Total	334	100.00%

Fuente:

Elaboración Propia

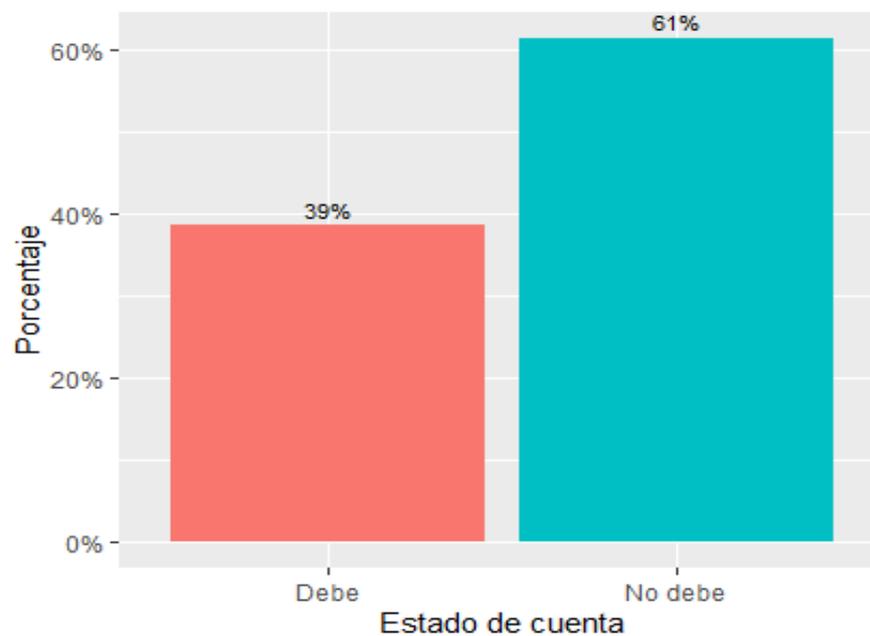


Figura 12. Estado de cuenta de los estudiantes de la Maestría en Derecho - UNA

Fuente: Elaboración Propia

En la figura 12 se aprecian el estado de cuenta de los estudiantes de la Maestría en Derecho de la UNA, el 61% no cuenta con deudas y el 39% tiene deudas pendientes, observando que la mayoría de los estudiantes tienen su estado de cuenta sin deudas.

Tabla 06. Estado civil estudiantes de la Maestría en Derecho de la UNA

Estado civil	Estudiantes	%
Solteros	111	33.23%
Convivientes	129	38.62%
Casados	87	26.05%
Divorciados	7	2.10%
Total	334	100.00%

Fuente: Elaboración Propia

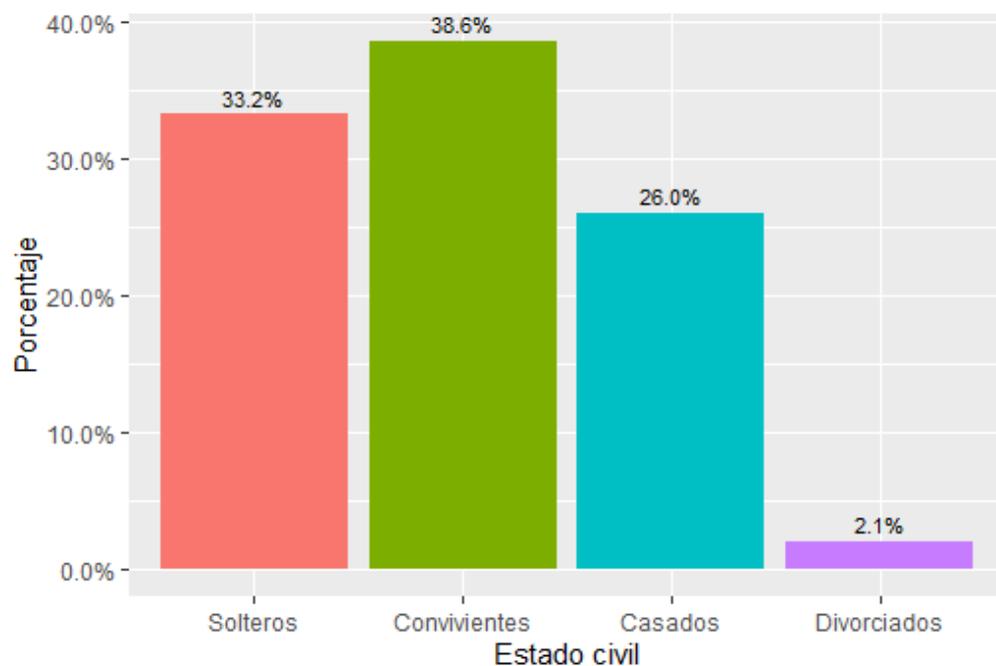


Figura 13. Estado civil estudiantes de la Maestría en Derecho de la UNA.

Fuente: Elaboración Propia

En la figura 13 se observa la información referida a el estado civil al que pertenecen los estudiantes de la Maestría en Derecho de la UNA siendo el 36.6% convivientes, el 32.2% solteros, el 26% casados y el 2.1% divorciados. Resultados que evidencian una gran concentración de estudiantes entre solteros y convivientes.

Resultado que es apoyado por Ramírez (2017) quien concluye que los factores individuales y de contexto inciden significativamente en la deserción universitaria de los estudiantes del Programa Beca 18.

Tabla 07. Procedencia de los estudiantes de la Maestría en Derecho de la UNA

Procedencia	Estudiantes	%
Provincia de Puno	211	63.17%
Otro lugar	123	36.83%
Total	334	100.00%

Fuente: Elaboración Propia

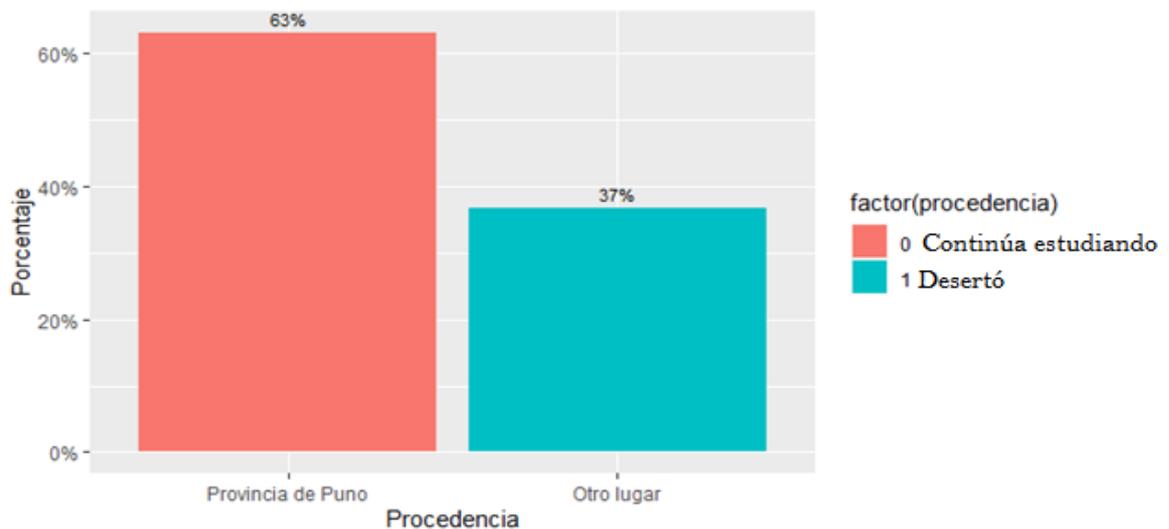


Figura 14. Procedencia de los estudiantes de la Maestría en Derecho de la UNA.

Fuente: Elaboración Propia

En la figura 14 observamos que la procedencia de los estudiantes de la Maestría en Derecho de la UNA mayormente es de la provincia de Puno siendo estos un 63% y provenientes de otros lugares ya sean de otras regiones alcanzando un 37%. De estos resultados se puede inducir que alrededor de 2/3 de los alumnos provienen de la provincia de Puno.

Tabla 08. Tipo de universidad que provienen los estudiantes de la Maestría en Derecho de la UNA

Tipo de universidad	Estudiantes	%
Privado	166	49.70%
Estatad	168	50.30%
Total	334	100.00%

Fuente: Elaboración Propia

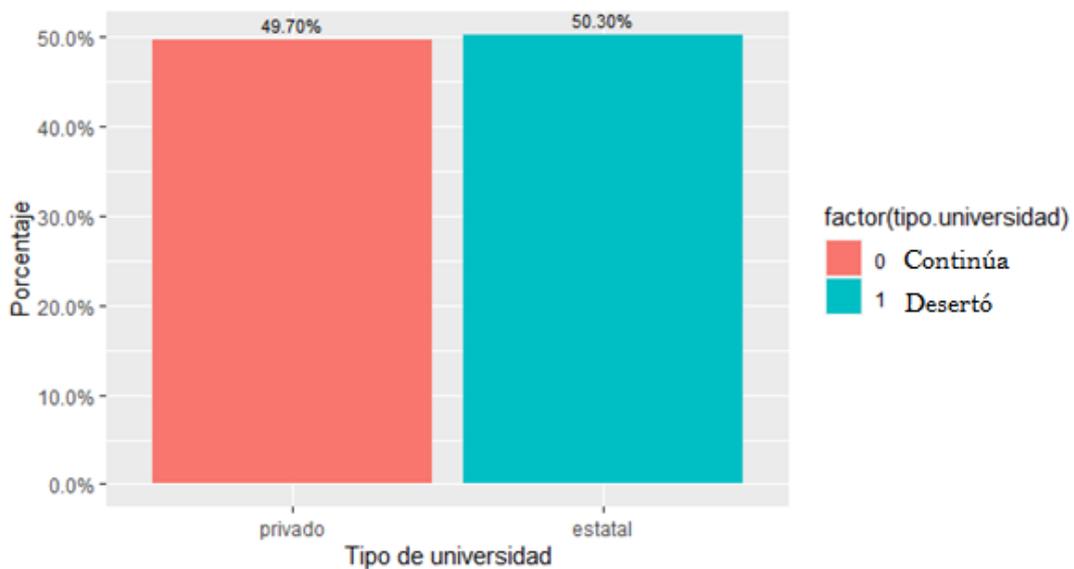


Figura 15. Tipo de universidad de la que provienen los estudiantes de la Maestría en Derecho de la UNA.

Fuente: Elaboración Propia

En la figura 15 se aprecia que el 50.30 % de los estudiantes de la Maestría en Derecho de la UNA provienen de universidades estatales, y el 49.70% de universidades privadas; siendo esta información que indica que casi se encuentran igual cantidad de estudiantes de universidades privadas y estudiantes de universidades estatales.

Tabla 09. Sexo y deserción estudiantil

Sexo/deserción	Mujeres		Hombres		Total
	Estudiantes	%	Estudiantes	%	
Deserto por diversos motivos	34	10.18%	47	14.07%	81
Continúa estudiando	109	32.63%	144	43.11%	253
Total	143	42.81%	191	57.19%	334

Fuente: Elaboración Propia

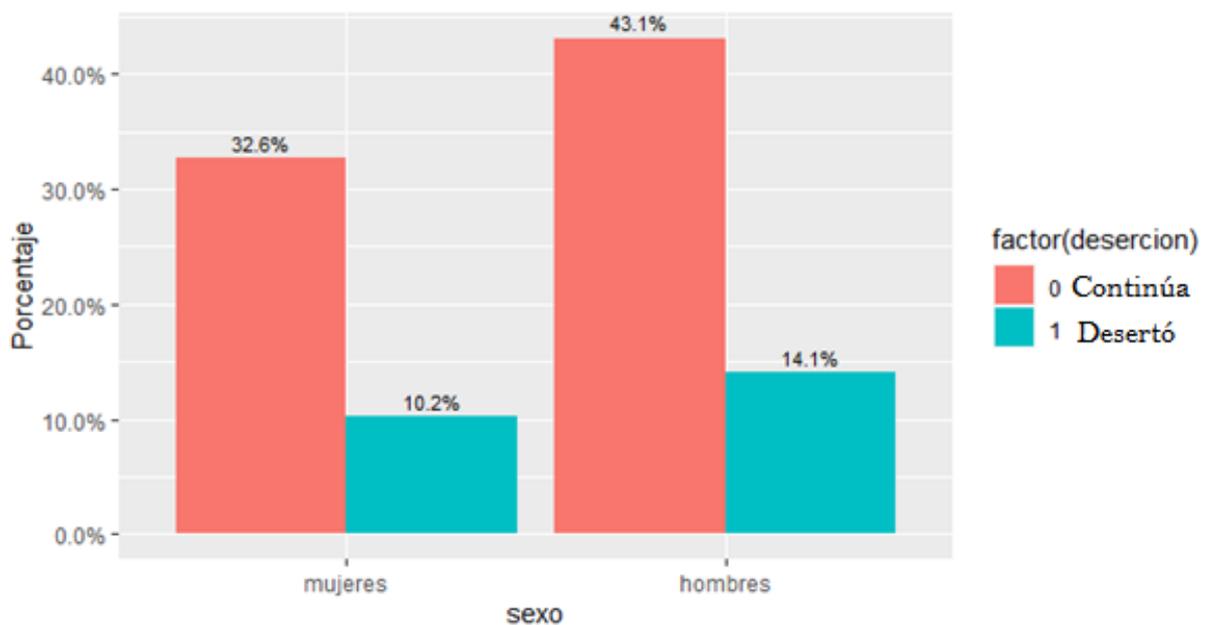


Figura 16. Sexo filtrado por la deserción estudiantil.

Fuente: Elaboración Propia

En la figura 16 observamos que el 43.1% de los estudiantes de la Maestría en Derecho de la UNA son hombres y no desertaron (0), el 32.6% son mujeres y no desertaron (0), el 14.1% son hombres que desertaron (1) y por último el 10.2% son

mujeres que desertaron (1). Se puede distinguir que tanto hombres como mujeres existen pocos que desertan patrón que la red neuronal debe ser capaz de aprender y generalizar.

Lo que es apoyado por Miranda et al. (2020) que tuvo como objetivo evaluar la efectividad de tres modelos de red neuronal artificial en la clasificación de alumnos de alto y bajo rendimiento, así como alumnos en riesgo de deserción, los resultados presentados muestran la efectividad de los modelos neuronales para clasificar bases a su escuela de procedencia, sexo, entre otros datos.

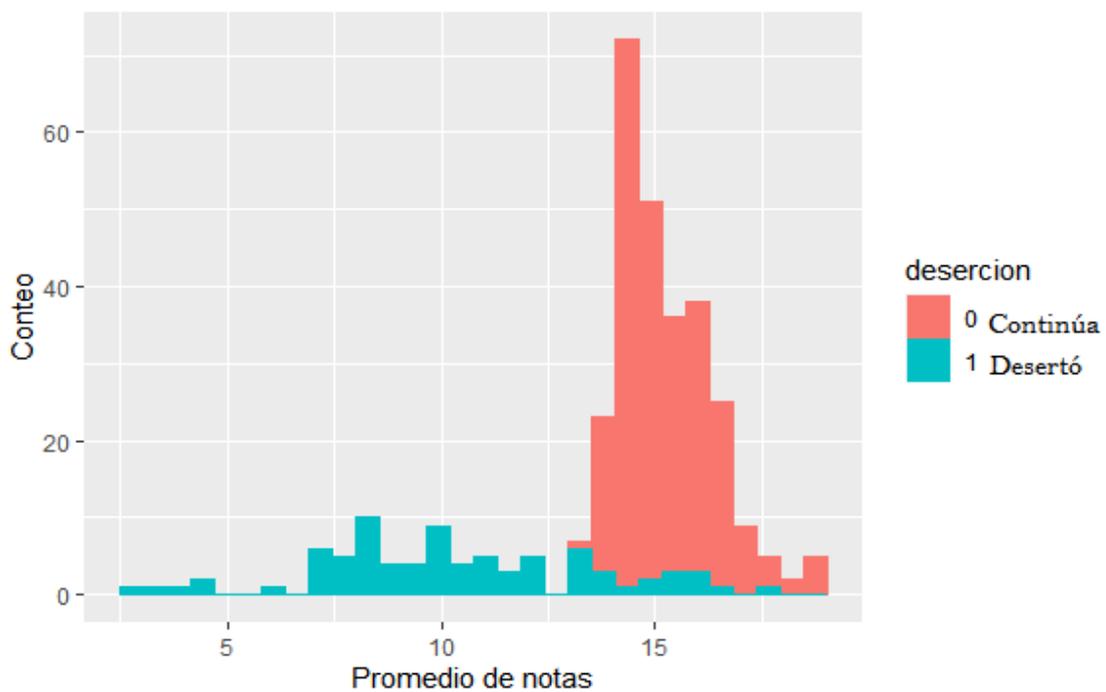


Figura 17. Histograma de los promedios filtrados por la deserción

Fuente: Elaboración Propia

En la figura 17 se observa que estudiantes con promedios de notas altos mayores a 14 aproximadamente tienden a no desertar (0) y estudiantes con bajas calificaciones son los que desertan (1). Observamos un claro patrón de comportamiento en la deserción según el promedio se esperará que un estudiante con excelentes calificaciones no deserte y un estudiante con bajas calificaciones deserte.

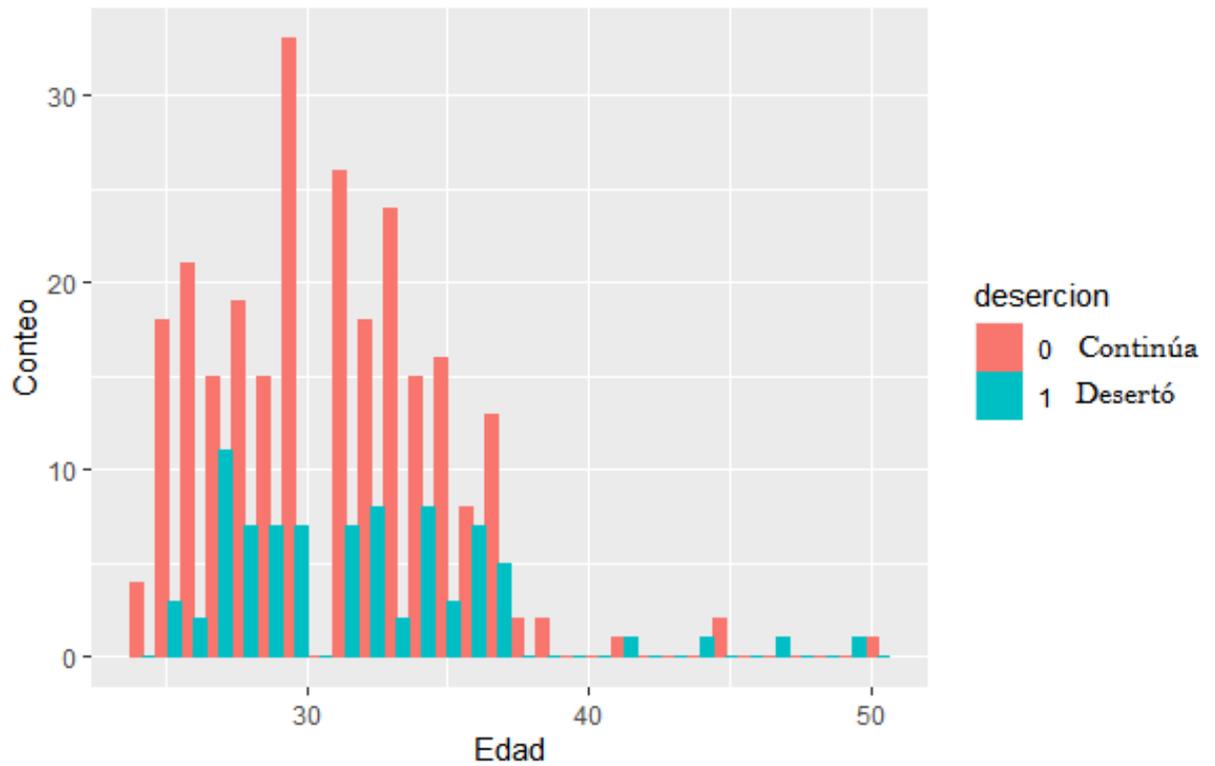


Figura 18. Edad de los estudiantes filtrados por la deserción

Fuente: Elaboración Propia

En la figura 18 apreciamos que la edad de los % de los estudiantes de la Maestría en Derecho de la UNA, filtrados por la deserción fue indistinta solo se aprecia una mayor concentración de estudiantes que no desertan (0).



Tabla 10. Estado civil y deserción estudiantil

Estado civil/deserción	Soltero		Conviviente		Casado		Divorciado	
	Estudiantes	%	Estudiantes	%	Estudiantes	%	Estudiantes	%
Desertero por diversos motivos	31	9.28%	31	9.28%	18	5.39%	1	0.30%
Continúa estudiando	80	23.95%	98	29.34%	69	20.66%	6	1.80%
Total	111	33.23%	129	38.62%	87	26.05%		2.10%

Fuente: Elaboración Propia

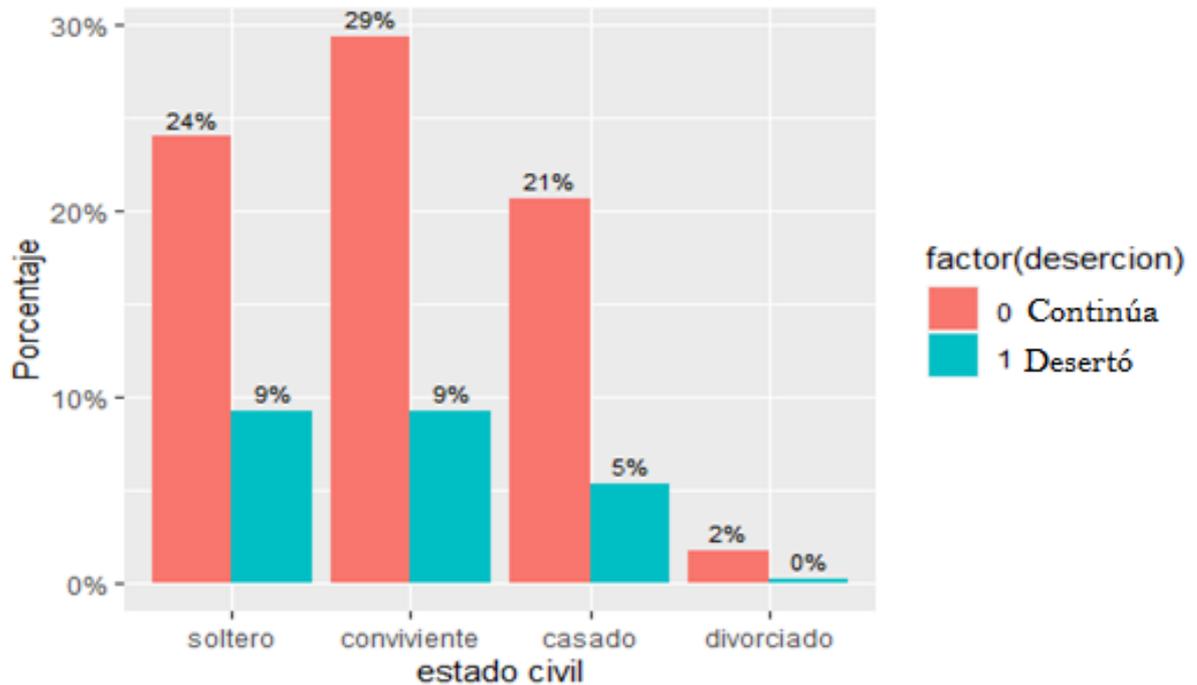


Figura 19. Estado civil filtrado por la deserción

Fuente: Elaboración Propia

En la figura 19 observamos que el 29% de los estudiantes de la Maestría en Derecho de la UNA fueron de estado civil conviviente que no desartaron (0), el 23% solteros que no desartaron (0), 21% casados que no desartaron, 9% solteros que desartaron (1), 9% convivientes que desartaron (1), 5% casados que desartaron, 2% divorciados que no desartaron y por último el 0% de los divorciados desartaron (1).

Hay un claro patrón de comportamiento en el estado civil que indistintamente del que tengan la gran mayoría no deserta solo en los divorciados ninguno de ellos deserto.

Tabla 11. Procedencia y deserción estudiantil

Procedencia/ deserción	Provincia de Puno		Otro lugar		Total
	Estudiantes	%	Estudiantes	%	
Deserto por diversos motivos	52	15.57%	29	8.68%	81
Continúa estudiando	159	47.60%	94	28.14%	253
Total	211	63.17%	123	36.83%	334

Fuente: Elaboración Propia

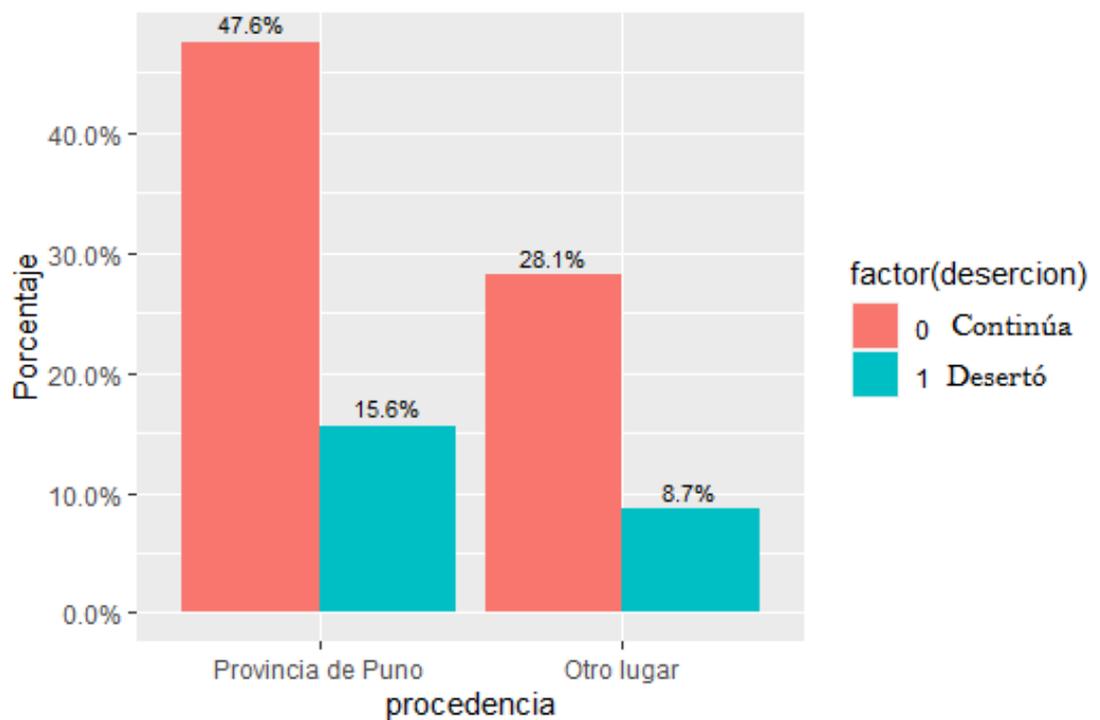


Figura 20. Procedencia filtrada por la deserción

Fuente: Elaboración Propia

En la figura 20 observamos que el 47.6% de los estudiantes de la Maestría en Derecho de la UNA provinieron de la provincia de Puno que no desertaron (0) y continua

con sus estudios, el 28.1% proviene de otro lugar diferente no desertaron (0), el 15.6% provienen de la provincia de Puno y desertaron (1), el 8.7% provienen de otro lugar y desertaron (1). Distinguiéndose claramente que según la procedencia desiertan más si provienen de la provincia de Puno, y en menos medida si provienen de otro lugar.

Tabla 12. Estado de cuenta filtrado por la deserción

Estado de cuenta/deserción	Debe		No debe		Total
	Estudiantes	%	Estudiantes	%	
Deserto por diversos motivos	39	11.68%	42	12.57%	81
Continúa estudiando	90	26.95%	163	48.80%	253
Total	129	38.62%	205	61.38%	334

Fuente: Elaboración Propia

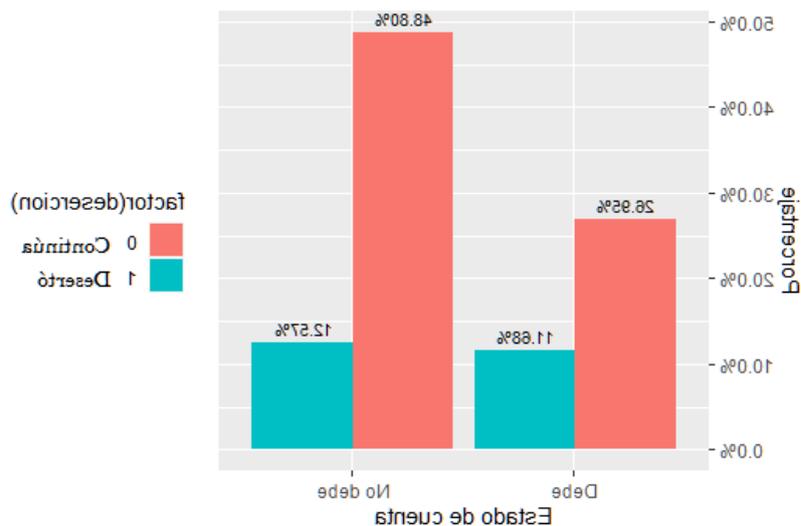


Figura 21. Estado de cuenta filtrado por la deserción

Fuente: Elaboración Propia

En la figura 21 observamos que el 48.8% de los estudiantes de la Maestría en Derecho de la UNA contó con un estado de cuenta sin deudas que no desertaron (0), el 26.95% cuenta con un estado de cuenta pendiente de deudas y que desertaron (1), el 12.57% en estado de cuenta sin deudas pero que desertaron (1) y el 11.68% con estado de cuentas con deudas pendientes pero que no desertaron (0). Resultados que dan a entender que una gran cantidad de estudiantes que no deben tienden a no desertar.

Tabla 13. Tipo de universidad y la deserción.

Tipo de universidad/deserción	Estatal		Privado		Total
	Estudiantes	%	Estudiantes	%	
Deserto por diversos motivos	35	10.48%	46	13.77%	81
Continúa estudiando	131	39.22%	122	36.53%	253
Total	166	49.70%	168	50.30%	334

Fuente: Elaboración Propia

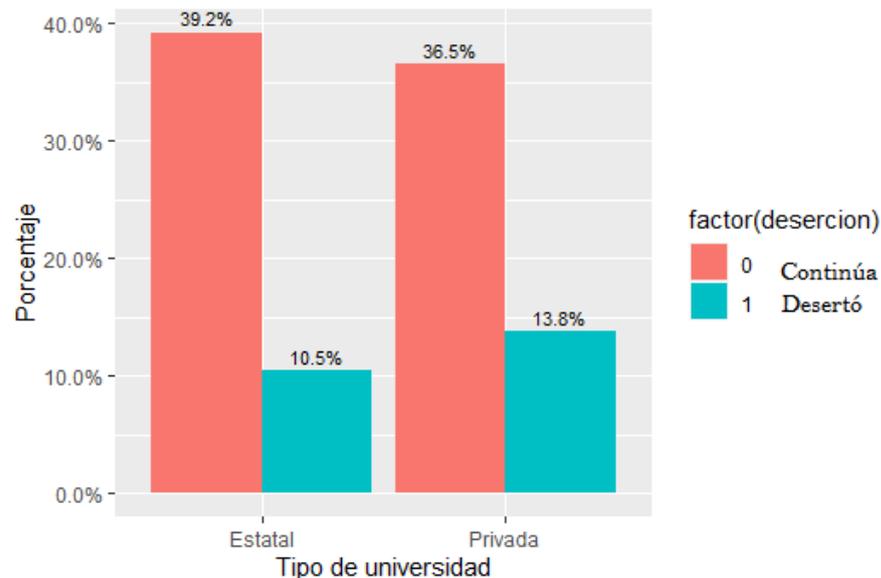


Figura 22. Tipo de universidad filtrado por la deserción

Fuente: Elaboración Propia

En la figura 22 observamos que el 39.02% de los estudiantes de la Maestría en Derecho de la UNA son provenientes de universidades estatales que no desertaron, el 36.5% de universidades privadas que no desertaron, el 10.5% de universidades estatales que desertaron y por último el 13.8% de universidades privadas que desertaron, observándose que existe una pequeña mayor deserción de universidades privadas.

Dividiendo el conjunto de datos en entrenamiento y validación en 80% (267 muestras de entrenamiento) y 20% (67 muestras de validación) teniendo en cuenta la disponibilidad de la información. Así tener suficientes datos de entrenamiento para captar los patrones en la RNN y tener suficientes datos de validación para garantizar los resultados de estas.

Tabla 14. Índice de las muestras de entrenamiento

Datos de entrenamiento								
224	209	79	112	44	245	32	43	169
78	93	69	202	307	0	249	217	58
295	30	208	194	65	299	162	190	48
17	22	96	117	292	286	256	259	235
24	284	97	319	85	303	285	105	252
250	294	198	322	186	100	242	53	21
326	56	68	302	70	192	138	1	313
238	175	23	164	159	287	62	49	160
177	10	37	223	12	218	215	80	276
82	225	246	226	35	213	135	205	191
111	227	67	154	28	103	128	34	293
143	81	266	231	170	171	8	263	257
240	153	144	2	142	98	64	91	308
116	39	125	115	244	36	14	52	149
140	310	221	173	131	255	156	264	130
176	104	86	272	268	61	40	241	151
90	155	282	120	51	150	306	13	99
314	193	152	318	95	47	271	88	87
258	158	283	289	254	247	216	273	214
31	124	148	127	178	315	290	166	121
300	228	291	74	204	11	236	328	333
185	66	239	29	123	330	207	134	20
167	210	222	83	41	200	212	321	188
234	18	157	277	89	298	304	243	71
331	196	219	107	260	27	251	54	106
145	267	129	183	206	262	325	50	270
59	139	38	279	136	230	305	174	102
119	19	237	133	26	233	269	189	x
15	329	203	184	327	4	201	332	x
6	92	165	232	141	122	161	187	x

Fuente: Elaboración Propia



Fuente: Elaboración Propia

En la tabla 14 se aprecia los índices de los registros que pertenecen a la muestra de entrenamiento así el valor 1 se interpreta como el primer registro que contengan las variables (promedio de notas, edad, sexo estado civil, procedencia, estado de cuenta y tipo de universidad).

Tabla 15. Índice de las muestras de validación

Datos de validación						
25	75	109	168	33	94	297
309	9	137	301	42	229	317
73	320	316	63	172	220	280
195	126	199	84	181	16	3
57	265	5	278	113	108	77
296	132	110	288	72	248	311
281	114	261	182	274	7	60
180	146	324	253	45	55	x
211	76	163	46	101	147	x
197	275	323	118	312	179	x

Fuente: Elaboración Propia

En la tabla 15 se observó los índices en la muestra de validación son todos aquellos registros que no son contemplados en la muestra de entrenamiento siendo estos un total de 67 registros que contengan las variables (promedio de notas, edad, sexo estado civil, procedencia, estado de cuenta y tipo de universidad).

1.16. ARQUITECTURA DE RED NEURONAL PARA PREDECIR LA DESERCIÓN

Según el objetivo específico 2. En el cual nos propusimos diseñar la arquitectura de una red neuronal artificial para predecir la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano Puno.

Se presentan los siguientes resultados.

Tabla 16. Configuración de las 4 redes neuronales a comparar

Modelo	tasa de aprendizaje	N° de neuronas en la capa oculta	numero de iteraciones realizado	Función de activación	Algoritmo de entrenamiento
1	0.01	1	500	Sigmoidea	rprop
2	0.03	1, 3	70	Sigmoidea	rprop
3	0.3	1, 2	70	Sigmoidea	rprop
4	0.1	1	25	Sigmoidea	rprop

Fuente: Elaboración Propia

En la tabla 16 se muestra la configuración de los 4 modelos en el cual el modelo 1 tiene una tasa de aprendizaje de 0.01, 1 neurona en la capa oculta , realizó 500 iteraciones, con función de activación Sigmoidea y utilizo el algoritmo rprop para calcular los pesos en las neuronas; el modelo 2 con tasa de aprendizaje de 0.03, 3 neuronas en la capa oculta realizo 70 iteraciones con función de activación Sigmoidea y algoritmo rprop; el modelo 3 con 0.3 de tasa de aprendizaje, 1,2 neuronas en la capa oculta 70 iteraciones, función de activación logística y algoritmo rprop; el modelo 4 con una tasa de aprendizaje de 0.1, con 1 neurona en la capa oculta, 25 iteraciones , función de activación Sigmoidea y algoritmo rprop.

Se tomó en cuenta los valores de aprendizaje de modo que la actualización de los parámetros θ depende de la tasa de aprendizaje (learning rate) valores que suelen ser 0.1 ,0.3 ,0.5; entre más pequeño sea menor será el gradiente por tanto se requerirá un mayor número de iteraciones (n_iters) por lo que debe configurarse en cada caso específico (Anastasiadis et al., 2005).

Se utilizó en toda función sigmoidea según Brownlee (2017) es una función de activación ideal para problemas de clasificación binaria donde la salida se interpreta como una distribución de probabilidad.



Igel & Husken (2003) recomiendan la estrategia de Rprop porque es capaz de aumentar ligeramente el valor de la actualización cuando las derivadas conservan su signo, esto ayuda a acelerar la convergencia en regiones poco profundas cuando las derivadas son negativas, pero puede ser ineficiente cuando las dos derivadas son positivas. En el caso de las derivadas positivas

Las grandes actualizaciones de la tasa de aprendizaje pueden llevar la trayectoria del peso lejos del mínimo o en regiones con valores más altos de la función de error de la función de error (Igel & Husken, 2003)

Arquitectura similar a la utilizada por Zevallos (2017) quien utilizó una arquitectura de red neuronal, utilizando el perceptrón multicapa con algoritmo de aprendizaje de rprop y función de activación la tangente hiperbólica. La arquitectura de la red neuronal tiene 3 capas ocultas.

1.16.1. Modelo de Red Neuronal 1

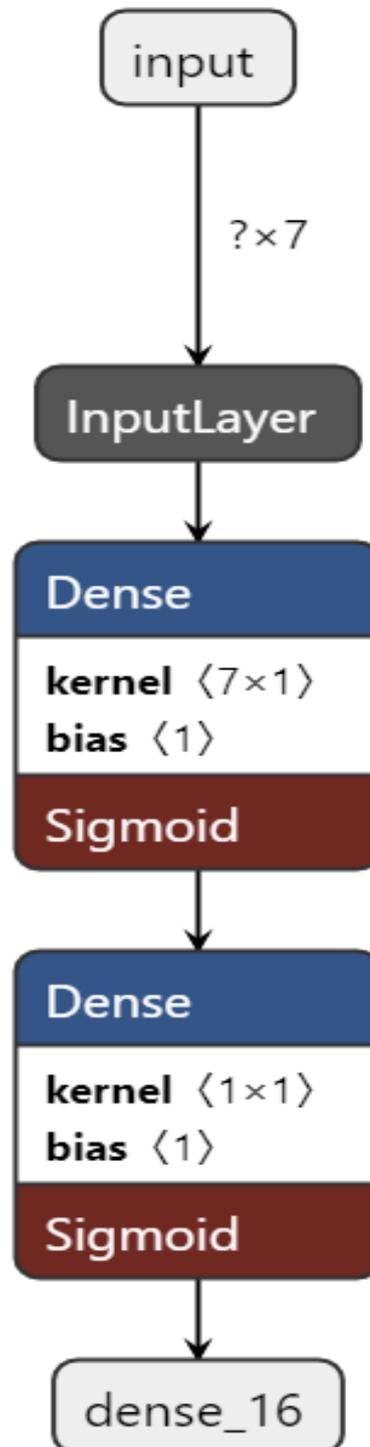


Figura 23. Modelo de Red neuronal 1

Fuente: Elaboración Propia

En la Figura 23 se observa la red neuronal 1 que tiene como entradas las 7 variables en estudio que son: Promedio de notas, edad, sexo, estado civil, procedencia,

estado de cuenta y tipo de universidad cuenta con una sola neurona en la capa oculta de función de activación sigmoidea, una neurona de salida con función sigmoidea.

Considerado como una arquitectura de perceptrón simple.

Número de iteraciones que minimiza la RNA 1 de la función de pérdida en los conjuntos de entrenamiento y validación

Tabla 17. Número de iteraciones que minimiza la RNA 1

Iteración	Entrenamiento	Validación
	Función de pérdida	
1	0.97556	0.95652
50	0.32479	0.31772
100	0.20501	0.20682
150	0.17457	0.18276
200	0.16368	0.17616
250	0.15856	0.17324
300	0.15554	0.17039
400	0.15172	0.16318
450	0.15031	0.15981
490	0.14932	0.15752
491	0.14929	0.15747
492	0.14927	0.15742
493	0.14925	0.15736
494	0.14922	0.15731
495	0.14920	0.15726
496	0.14918	0.15721
497	0.14915	0.15716
498	0.14913	0.15711
499	0.14911	0.15706
500	0.14908	0.15701

Fuente: Elaboración propia

De la tabla presentada se resume en la figura.

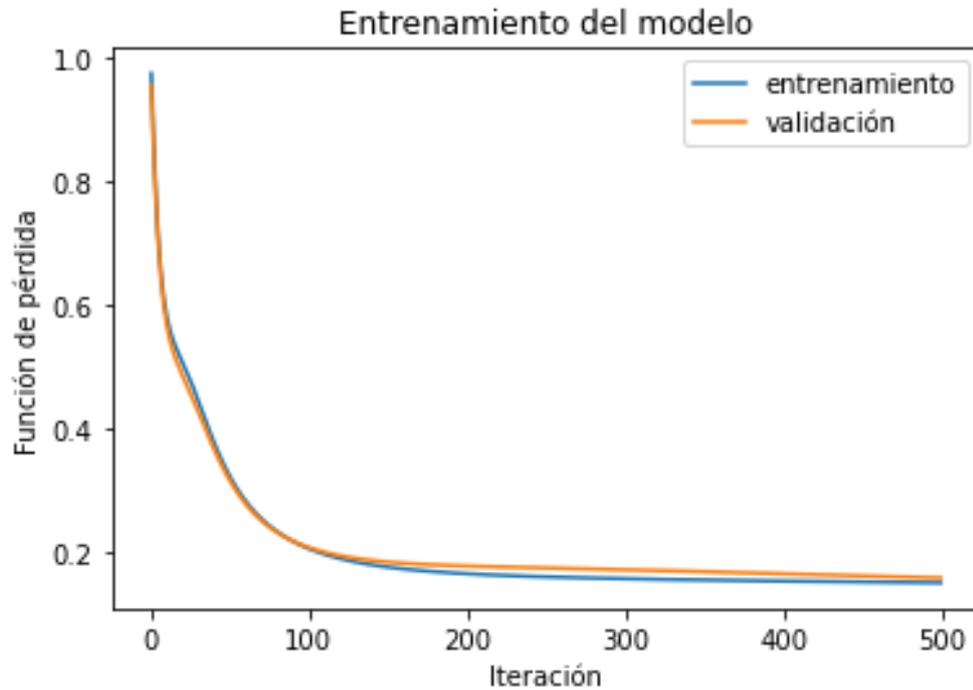


Figura 24. Número de iteraciones que minimiza la RNA 1 de la función de pérdida en los conjuntos de entrenamiento y validación

Fuente: Elaboración propia

En la figura 24 se presenta la reducción de la función de pérdida para la RNA 1 que minimiza lentamente a una tasa de aprendizaje de 0.01 con una sola neurona en la capa oculta de función sigmoidea y la capa de salida con una neurona de función sigmoidea que facilita la predicción de la probabilidad de deserción entre 0 y 1.

1.16.2. Modelo de Red neuronal 2

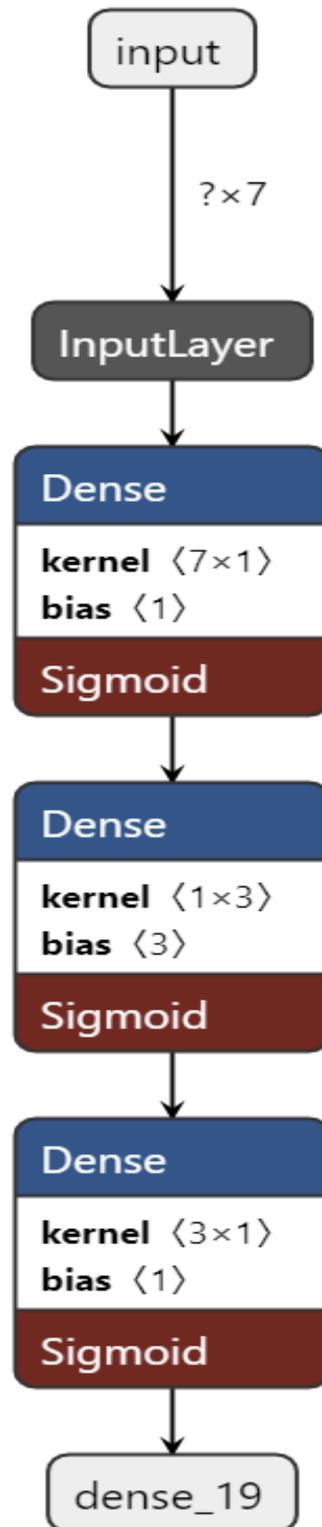


Figura 25.Modelo de Red neuronal 2

Fuente: Elaboración Propia

En la Figura 25 se observa la red neuronal 2 que tiene como entradas las 7 variables promedio de notas, edad, sexo, estado civil, procedencia, estado de cuenta y tipo de universidad, esta RNA presenta dos neuronas ocultas que tienen 1 y 3 neuronas respectivamente con función sigmoidea. Una única neurona de salida con función sigmoidea para garantizar que la RNA pueda predecir la probabilidad de deserción valores que oscilan entre 0 y 1.

Número de iteraciones que minimiza la RNA 2 de la función de pérdida en los conjuntos de entrenamiento y validación

Tabla 18. Número de iteraciones que minimiza la RNA 2

Iteración	Función de pérdida	
	Entrenamiento	Validación
1	0.68001	0.56964
11	0.41862	0.37343
20	0.18754	0.19051
30	0.16114	0.17151
40	0.15442	0.16262
50	0.15097	0.15753
60	0.14874	0.15444
70	0.14714	0.15242
71	0.14700	0.15225
72	0.14687	0.15210
73	0.14674	0.15194
74	0.14661	0.15180
75	0.14649	0.15165

Fuente: Elaboración propia

A partir de la tabla se resume en la figura siguiente.

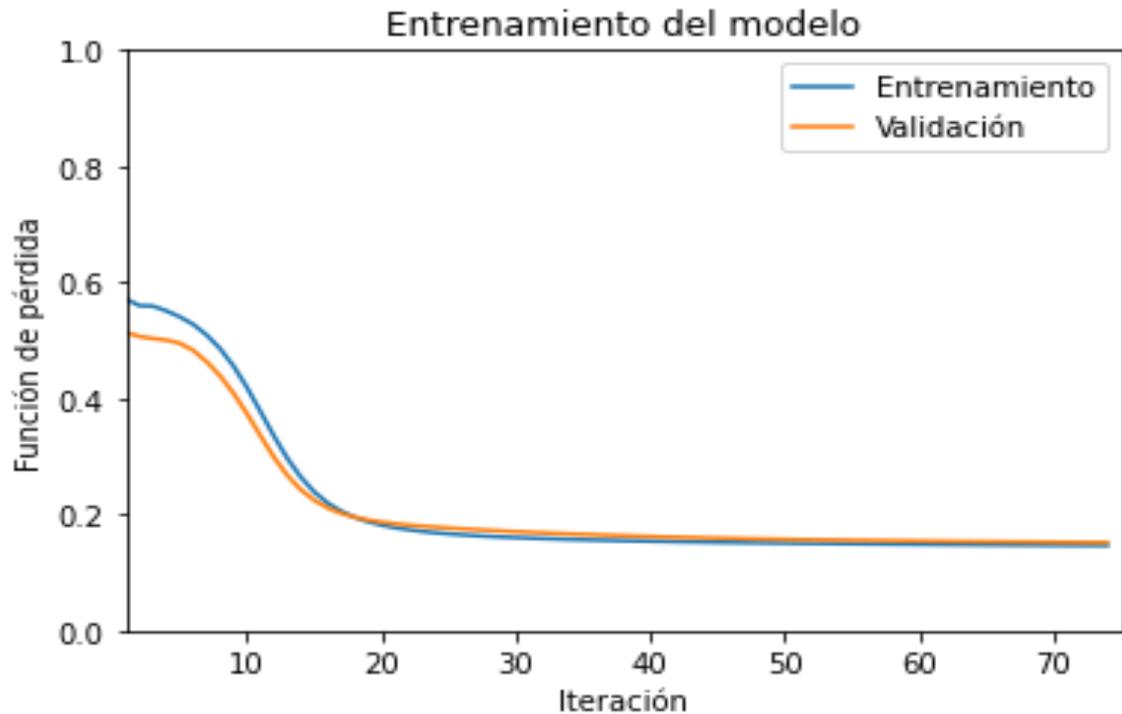


Figura 26. Número de iteraciones que minimiza la RNA 2 de la función de pérdida en los conjuntos de entrenamiento y validación

Fuente: Elaboración propia

De acuerdo a la figura 26 la RNA 2 minimiza rápidamente la función de pérdida con una tasa de aprendizaje de 0.03, con dos capas ocultas la primera de una sola neurona, la segunda con tres neuronas siendo un perceptrón multicapa; se observa que la función de pérdida no disminuye a partir de 50 de iteraciones.

1.16.3. Modelo de Red neuronal 3

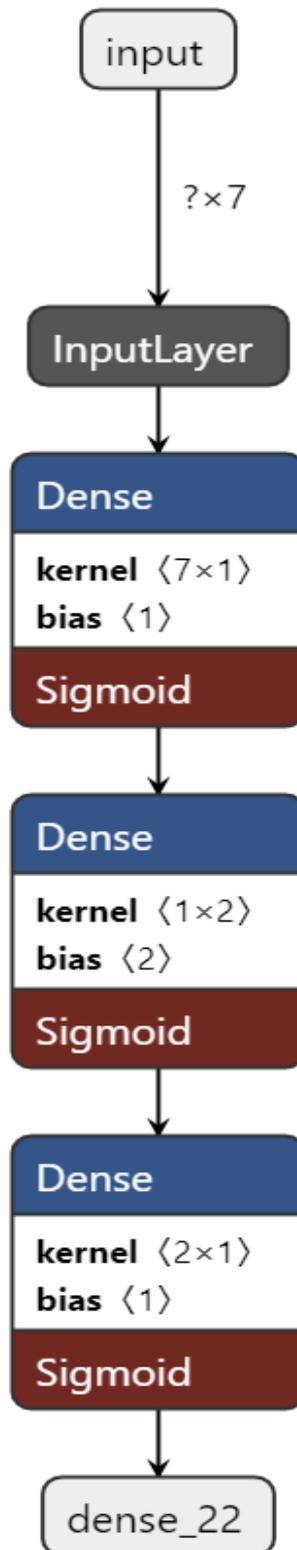


Figura 27. Modelo de Red neuronal 3

Fuente: Elaboración Propia

En la Figura 27 se observa la red neuronal 3 que tiene como entradas las 7 variables promedio de notas, edad, sexo, estado civil, procedencia, estado de cuenta y tipo de universidad cuenta con dos capas ocultas, la primera con una sola neurona que tiene la función de activación sigmoidea, la segunda con dos neuronas con función de activación sigmoidea y por último una capa de salida con una sola neurona con función de activación sigmoidea que predice la probabilidad de deserción. Contando con un número de parámetros = 15

Número de iteraciones que minimiza la RNA 3 de la función de pérdida en los conjuntos de entrenamiento y validación.

Tabla 19. Número de iteraciones que minimiza la RNA 3

Iteración	Entrenamiento	Validación
	Función de pérdida	
1	0.68001	0.56964
10	0.15133	0.13479
20	0.14302	0.15617
30	0.14257	0.14791
40	0.14213	0.14510
50	0.14188	0.14323
60	0.14173	0.14187
61	0.14171	0.14175
62	0.14170	0.14163
63	0.14169	0.14152
64	0.14168	0.14141
65	0.14167	0.14131
66	0.14166	0.14121
67	0.14165	0.14110
68	0.14164	0.14101
69	0.14163	0.14091
70	0.14162	0.14082
71	0.14161	0.14072
72	0.14160	0.14063
73	0.14159	0.14055
74	0.14158	0.14046
75	0.14158	0.14038

Fuente: Elaboración propia

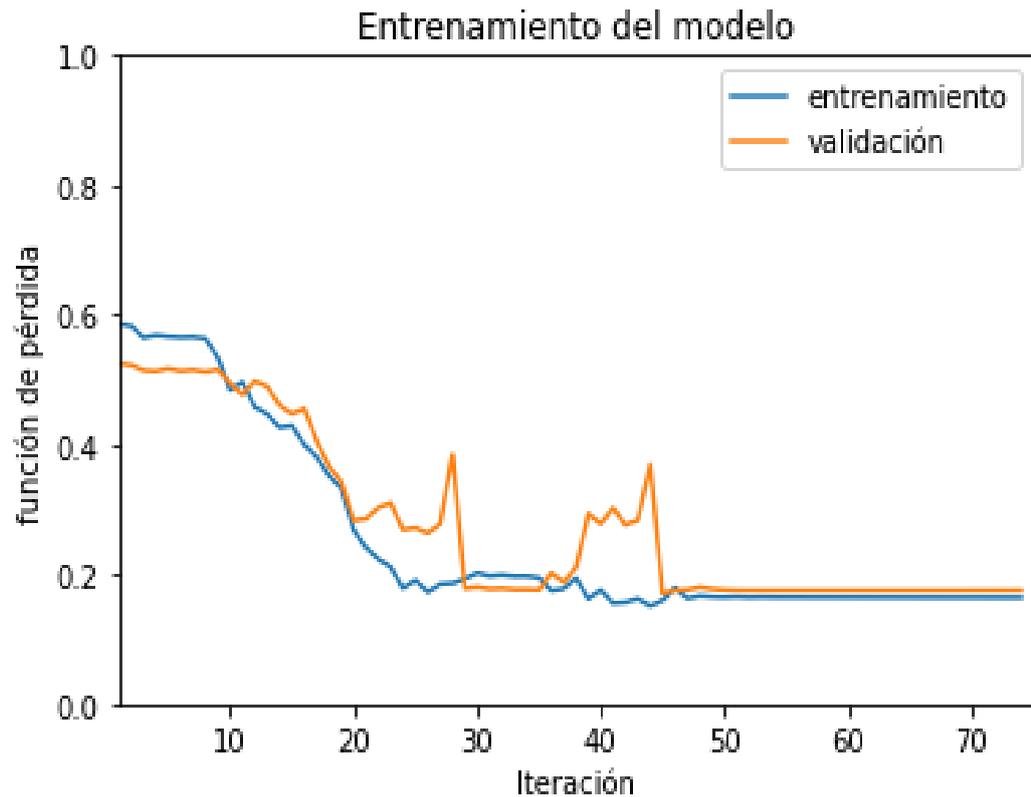


Figura 28. Número de iteraciones que minimiza la RNA 3 de la función de pérdida en los conjuntos de entrenamiento y validación.

Fuente: Elaboración Propia

Presentando las iteraciones la RNA 3 minimiza rápidamente la función de pérdida ya que su tasa de aprendizaje es 0.3; con dos capas ocultas la primera de una sola neurona, la segunda con dos neuronas siendo un perceptrón multicapa; con estos hiperparámetros propician que la minimización ya no disminuya a partir de 50 iteraciones manteniéndose robusto al no sobre ajustarse en validación.

1.16.4. Modelo de Red neuronal 4

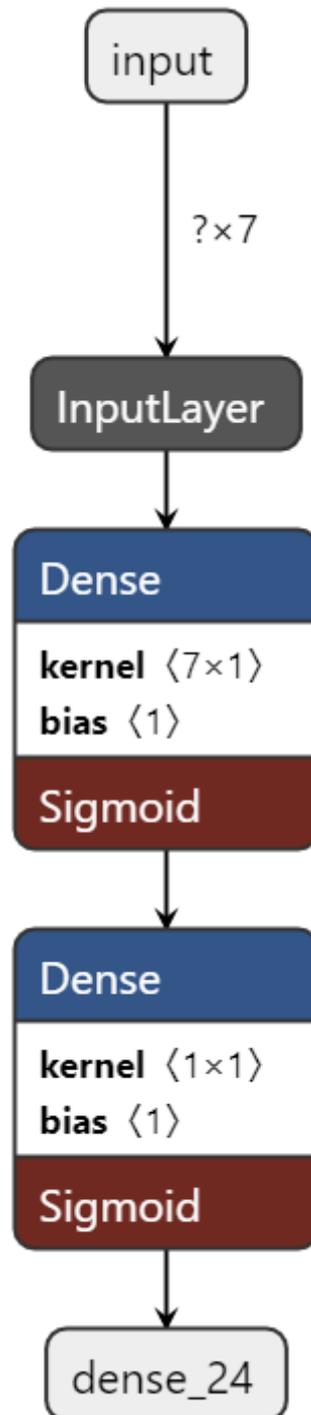


Figura 29. Red neuronal 4

Fuente: Elaboración Propia

En la Figura 29 se observa la arquitectura de la red neuronal 4 que tiene como entradas las 7 variables promedio de notas, edad, sexo, estado civil, procedencia, estado de cuenta y tipo de universidad cuenta con una sola neurona en la capa oculta con una

función de activación sigmoidea y una sola neurona en la capa de salida con función de activación sigmoidea cuya salida da la probabilidad de deserción valores ente 0 y 1.

Número de iteraciones que minimiza la RNA 4 de la función de pérdida en los conjuntos de entrenamiento y validación

Tabla 20. Número de iteraciones que minimiza la RNA 4

Iteración	Entrenamiento	Validación
	Función de pérdida	
1	0.76544	0.54570
2	0.53656	0.48511
3	0.51811	0.45747
4	0.46871	0.41096
5	0.40456	0.36102
6	0.34758	0.31285
7	0.30029	0.27436
8	0.26510	0.24579
9	0.24013	0.22702
10	0.22224	0.21406
11	0.20915	0.20461
12	0.19936	0.19816
13	0.19192	0.19335
14	0.18622	0.18956
15	0.18172	0.18698
16	0.17811	0.18518
17	0.17518	0.18372
18	0.17275	0.18259
19	0.17070	0.18174
20	0.16896	0.18106
21	0.16747	0.18048
22	0.16618	0.17999
23	0.16504	0.17955
24	0.16403	0.17915
25	0.16314	0.17876

Fuente: Elaboración propia.

De la figura tabla se resume en la siguiente figura.

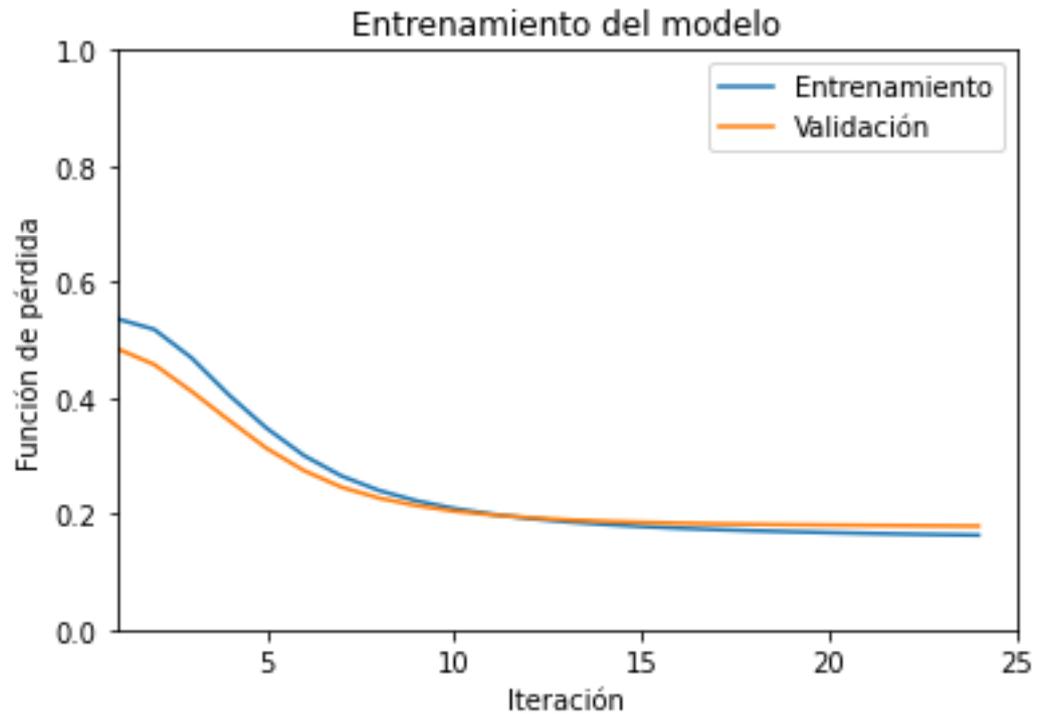


Figura 30. Número de iteraciones que minimiza la RNA 4 de la función de pérdida en los conjuntos de entrenamiento y validación

Fuente: Elaboración Propia

Se observa que la RNA 4 minimiza rápidamente la función de pérdida ya que su tasa de aprendizaje es de 0.1 y el número de neuronas ocultas son solo 1 basado en el perceptrón simple; estos hiperparámetros propician que la minimización ya no sea significativa a partir de 20 iteraciones sin embargo los resultados indican que pudo minimizar la función de pérdida por debajo de 0.20 tanto en entrenamiento y validación por lo que esta RNA es bastante robusta.

1.17. ENTRENAMIENTO Y VALIDACIÓN DEL MODELO DE RNA PARA OBTENER UN MODELO DE PREDICCIÓN

Según el objetivo específico 3, nos propusimos entrenar y validar el modelo de Red Neuronal Artificial para obtener un modelo de predicción para la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno.

1.17.1. Entrenamiento

Utilizando el límite de 0.5 para predecir si una estudiante desertara en los modelos se tiene los siguientes resultados en la muestra de entrenamiento:

Red neuronal 1

Tabla 21. Matriz de confusión de la red neuronal 1

Matriz de confusión		
Real/red neuronal 1	Continúan	Deserción
Continúan	200	0
Deserción	9	58

Fuente: Elaboración Propia

En la tabla 21 observamos que la red neuronal 1 acertó en 200 alumnos que no desertaron y ninguno de ellos fue clasificado erróneamente como desertor, también esta red acertó en 58 alumnos que desertaron (1) pero que a 9 desertores los clasifíco como alumnos que continúan estudiando.

Red neuronal 2

Tabla 22. Matriz de confusión de la red neuronal 2

Matriz de confusión		
Real/red neuronal 2	Continúan	Deserción
Continúan	200	0
Deserción	9	58

Fuente: Elaboración propia

Red neuronal 3

Tabla 23. Matriz de confusión de la red neuronal 3

Matriz de confusión		
Real/red neuronal 3	Continúan	Deserción
Continúan	200	0
Deserción	9	58

Fuente: Elaboración propia

Red neuronal 4

Tabla 24. Matriz de confusión de la red neuronal 4

Matriz de confusión		
Real/red neuronal 4	Continúan	Deserción
Continúan	200	0
Deserción	9	58

Fuente: Elaboración propia

Tabla 25. Métricas del error de entrenamiento en las 4 redes neuronales creadas

Red	Accuracy	Recall	F1_Score
1	1	0.86567	0.92800
2	1	0.86567	0.92800
3	1	0.86567	0.92800
4	1	0.86567	0.92800

Fuente: Elaboración propia

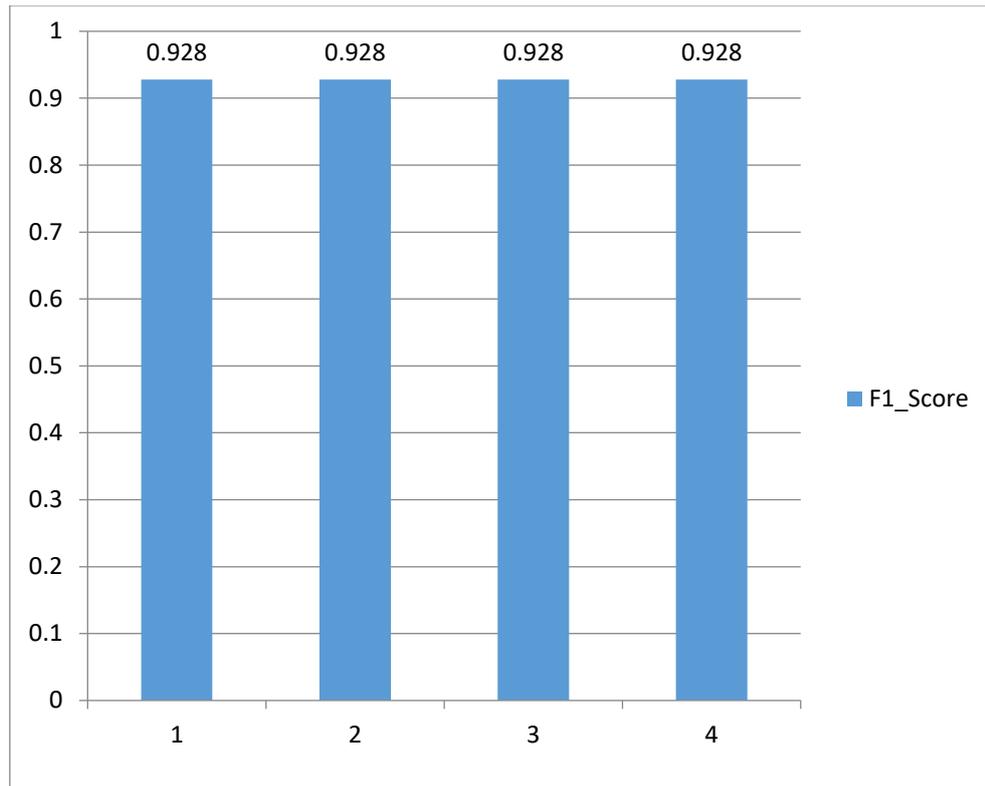


Figura 31. F1_Score de las 4 redes neuronales en comparación en la muestra de entrenamiento.

Fuente: Elaboración Propia

Las 4 redes neuronales en la muestra de entrenamiento obtienen un Accuracy de a 1 interpretándose que todas las RNAs aciertan a los que no son desertores en un 100% de las veces, observando los valores de Recall las redes neuronales no aciertan en todas las predicciones de los alumnos desertores observando aproximadamente un 87% de asertividad en los desertores , por ello se aplica la métrica F1-Score se ve afectada tanto por el % de aciertos en los desertores y el % de acierto por los no desertores las 4 redes neuronales obtienen un valor de F1-Score de 0.92800 que se interpreta que aproximadamente los 4 modelos aciertan en un 93% considerando ambas clases desertores y no desertores.

1.17.2. Validación

Utilizando 0.5 como límite de decisión para detectar la deserción

Tabla 26. *Matriz de confusión del conjunto de validación para la red neuronal 1*

Matriz de confusión		
Real/red neuronal 1	Continúan	Deserción
Continúan	52	1
Deserción	2	12

Fuente: Elaboración propia

En tabla 26 observamos la matriz de confusión para la red neuronal 1 en el conjunto de validación que logro clasificar 52 alumnos correctamente que no fueron desertores, 1 que continúa estudiando lo clasifiko erróneamente como desertor, 12 desertores fueron clasificados correctamente y solo 2 desertores fueron erróneamente clasificados como si continuaran estudiando.

Tabla 27. *Matriz de confusión del conjunto de validación para la red neuronal 2*

Matriz de confusión		
Real/red neuronal 2	Continúan	Deserción
Continúan	52	1
Deserción	2	12

Fuente: Elaboración propia

En la tabla 27 se observa la matriz de confusión para la red neuronal 2 en el conjunto de validación, se clasifica a 52 alumnos correctamente que continúan su estudios considerados como verdaderos positivos, 1 estudiante que continua estudiando lo clasifiko erróneamente como desertor considerado denominado falso positivo, 12 desertores fueron clasificados correctamente llamados verdaderos negativos y se clasifiko 2 desertores erróneamente como si continuaran estudiando llamados falsos negativos.

Tabla 28. Matriz de confusión del conjunto de validación para la red neuronal 3

Matriz de confusión		
Real/red neuronal 3	Continúan	Deserción
Continúan	51	2
Deserción	3	11

Fuente: Elaboración propia

En la tabla 28 se presenta en la matriz de confusión para la red neuronal 3 en el conjunto de validación que logro clasificar 51 alumnos como verdaderos positivos que continúan sus estudios, 2 que continúa estudiando lo clasifíco erróneamente como desertores llamados falsos negativos, simultáneamente 11 desertores fueron clasificados correctamente y 3 desertores fueron erróneamente clasificados.

Tabla 29. Matriz de confusión del conjunto de validación para la red neuronal 4

Matriz de confusión		
Real/red neuronal 4	Continúan	Deserción
Continúan	52	1
Deserción	2	12

Fuente: Elaboración propia

Tabla 30. Métricas de error de validación en las 4 redes neuronales creadas

Red	Accuracy	Recall	F1_Score
1	0.92307	0.85714	0.88888
2	0.92307	0.85714	0.88888
3	0.84615	0.78571	0.81481
4	0.92307	0.85714	0.88888

Fuente: Elaboración Propia

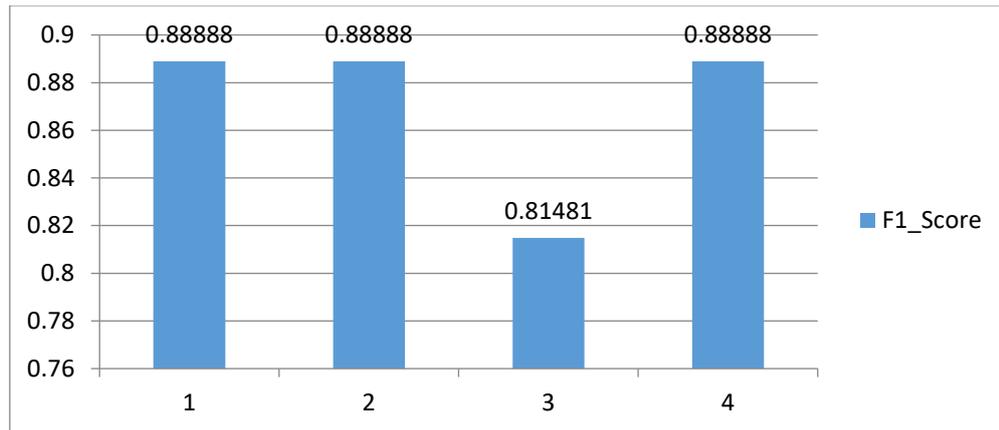


Figura 32. F1_Score de las 4 redes neuronales en comparación en la muestra de validación.

Fuente: Elaboración Propia

En la tabla 30 y figura 32 se observó que las RNAs 1,2 y 4 presentan las mismas métricas, siendo el porcentaje de aciertos para los no desertores del 92% (Accuracy), el porcentaje de aciertos para los estudiantes que desertan son aproximadamente del 86% (Recall) y en general considerando tanto el % de acierto de los desertores y no desertores es de 89% (F1_Score).

La RNA 3 presenta un % de aciertos a los no desertores del 85% (Accuracy), el % de aciertos en los desertores es de 79% (Recall), y en promedio tanto como desertores y no desertores acierta en el 81%.

Los resultados presentados son similares porque las RNAs están minimizando correctamente la función de pérdida

1.18. MODELO DE PREDICCIÓN PARA LA DESERCIÓN DE ESTUDIANTES MEDIANTE RNA

Según el objetivo general que nos planteamos fue determinar la RNA que mejor se ajuste para la predicción de deserción de los estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno.

Repetur (2019) manifiesta que el objetivo de un algoritmo de aprendizaje es reducir el error de generalización, con el menor costo computacional posible.

Para ello considerando las métricas de F1_Score las RNAs 1,2 y 4 presentando el mismo valor de aciertos del 89% siendo superiores la RNA 3 con 81% de aciertos. Se elige la RNA 4 por tener el menor número de iteraciones, al ser la más simple facilita su comprensión siendo un perceptrón simple y con un menor coste computacional.

Tabla 31. *Parámetros de la RNA que mejor se ajustó para la predicción de la deserción en estudiantes de la Maestría en Derecho de la UNA, mediante redes neuronales artificiales*

Modelo	tasa de aprendizaje	Número de neuronas en la capa oculta	Número de iteraciones realizado	Función de activación	algoritmo	Accurac y	Recal l	F1_scor e
4	0.1	1	25	sigmoidea	Rprop	92%	86%	89%

Fuente: Elaboración Propia

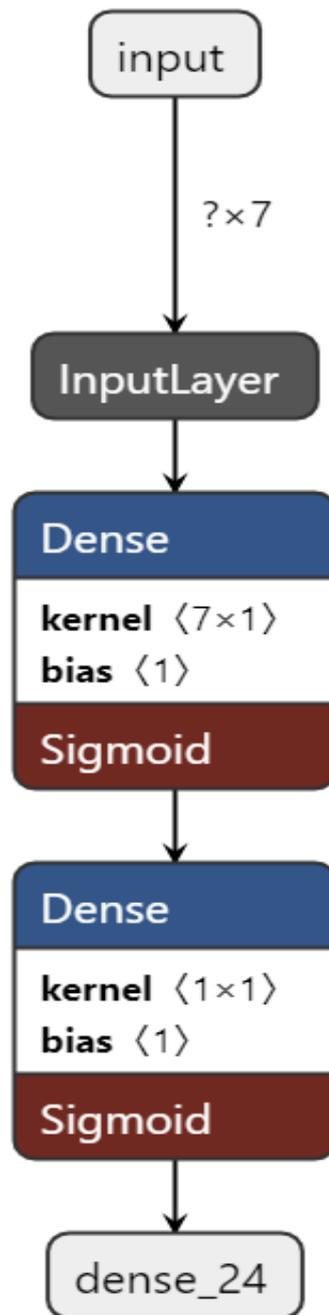


Figura 33. La RNA que mejor se ajustó de predicción de la deserción.

Fuente: Elaboración Propia

Según la tabla 31 y figura 33. La RNA que mejor se ajustó para la predicción de la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno, cuenta con una tasa de aprendizaje de 0.1, una sola neurona en la capa oculta, realizó 25 iteraciones, la función de activación es sigmoidea, el algoritmo para calcular los pesos sinápticos es de rprop, cuenta con un grado



de acierto de: Accuracy del 92% en la clase de los no desertores, Recall del 86% en la clase de los alumnos desertores y en general cuenta con F1_Score de 89%.

Resultados superiores a los de Ahmad & Shahzadi (2018) en el conjunto de validación al clasificar a los estudiantes en riesgo / no en riesgo, logrando una tasa de clasificación correcta del 85% en la muestra reservada.

Porcentaje de aciertos similares Zarria et al. (2016) sin embargo el uso del algoritmo de Random Forest es el algoritmo de mejor desempeño general, con un 88.9% de exactitud.

Por su parte Miranda & Guzmán (2017) la calidad de los clasificadores permite asegurar que sus predicciones son correctas, con niveles estadísticos de curva ROC de 76%, 75% y 83% de acierto para los clasificadores de red bayesiana, árbol de decisión y red neuronal respectivamente.

Asimismo Miranda et al. (2020) se encuentran a favor de la validez de estos modelos matemáticos de RNAs para tratar datos cualitativos, más allá de su efectividad o precisión en la clasificación. La principal conclusión a la que se llegó en este autor es de la gran utilidad e importancia de las RNAs para tratar variables cualitativas.



Cuya función matemática es:

$$\text{Sigmoide}(z) = \frac{1}{1 + e^{-z}}$$

$$\hat{y} = \text{Sigmoide}(((\text{Sigmoide}(-6.948304 * \text{promedio} + -1.1010511 * \text{edad} + \\ -0.63760865 * \text{sexo} + 0.37375426 * \text{estado civil} + 0.3536626 * \text{procedencia} + 0.7 \\ 231702 * \text{debe} + 0.46687895 * \text{tipo universidad} - 2.8151066)) * 7.1241055) + -3 \\ .0387094)$$



V. CONCLUSIONES

Se logró determinar que la RNA que mejor se ajustó para la predicción de la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno, cuenta con una tasa de aprendizaje de 0.1, una sola neurona en la capa oculta, realizó 25 iteraciones, la función de activación es sigmoidea, el algoritmo para calcular los pesos sinápticos es de rprop, cuenta con un grado de acierto de: Accuracy del 92% en la clase de los no desertores, Recall del 86% en la clase de los alumnos desertores y en general cuenta con F1_Score de 89%.

Se pudo identificar las variables que contribuyan al riesgo de deserción del estudiante durante su periodo de estudios ya se puede distinguir que tanto hombres 14.1% como mujeres 10.2% existen pocos que desertan, también se detectó un claro patrón de comportamiento en la deserción según el promedio de notas se espera que un estudiante con excelentes calificaciones mayores a 14 no deserte y un estudiante con bajas calificaciones deserte, también una gran cantidad de estudiantes que no tienen deudas pendientes con la universidad tienden a no desertar siendo estos un el 48.8% .

El Diseño de la arquitectura de una red neuronal artificial para predecir la deserción de estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano Puno es de 7 entradas (promedio de notas, edad, sexo estado civil, procedencia, estado de cuenta y tipo de universidad) con una neurona en la capa oculta de función sigmoidea y una neurona de salida con función sigmoidea

El entrenamiento de la red neuronal fue con el 80% de los datos y 20% de validación del modelo de Red Neuronal Artificial para obtener un modelo de predicción para la deserción en estudiantes de la Maestría en Derecho de la Escuela de Posgrado de la Universidad Nacional del Altiplano – Puno.



VI. RECOMENDACIONES

Al coordinador de estudios de la maestría en derecho de la universidad nacional del altiplano, utilizar la red neuronal desarrollada en el presente trabajo de investigación ya que se logró capturar el patrón de comportamiento con excelentes resultados (Accuracy del 92% en la clase de los no desertores, Recall del 86% en la clase de los alumnos desertores y en general cuenta con F1_Score de 89%). para así retener y de evitar la deserción de estudiantes, ya que esto les afecta económicamente y socialmente, para que las acciones tomadas sean efectivas al detectar potenciales alumnos desertores.

A los estudiantes y docentes de la maestría en derecho de la universidad nacional del altiplano tener en cuenta que al existir un claro patrón de comportamiento en la deserción según el promedio de notas se espera que un estudiante con excelentes calificaciones mayores a 14 no deserte y un estudiante con bajas calificaciones deserte; se recomienda mitigar, apoyar con el propósito de disminuir estos estudiantes con bajas calificaciones que son propensos a desertar.

A futuros investigadores tener en cuenta el presente trabajo como antecedente en el Diseño de la arquitectura de una red neuronal artificial para predecir en diversos campos de la ciencia y tecnología, ya que su gran versatilidad nos permite modelar datos de diferente comportamiento.

Se recomienda a persona interesadas en el tema de las redes neuronales seleccionar bien el conjunto de entrenamiento y validación porque de ellos depende saber que si una red neuronal está captando correctamente el patrón de los datos en cuestión



VII. REFERENCIAS BIBLIOGRÁFICAS

- Aguado, G. (2015). *Aplicación de técnicas de aprendizaje automático sobre juegos* [Universidad Politecnica de Valencia].
<https://riunet.upv.es/bitstream/handle/10251/55343/AGUADO - Aplicación de técnicas de aprendizaje automático sobre juegos.pdf?sequence=1>
- Ahmad, Z., & Shahzadi, E. (2018). *Prediction of Students' Academic Performance using Artificial Neural Network*. 40(3), 157–164. <https://eric.ed.gov/?id=EJ1209686>
- Alania, P. (2018). Aplicación de técnicas de minería de datos para predecir la deserción estudiantil de la facultad de ingeniería de la Universidad Nacional Daniel Alcides Carrión [Universidad Nacional Daniel Alcides Carrion]. In *Universidad Nacional "Daniel Alcides Carrión."*
http://repositorio.undac.edu.pe/bitstream/undac/829/1/T026_40573846_M.pdf
- Anastasiadis, A. D., Magoulas, G. D., & Vrahatis, M. N. (2005). New globally convergent training scheme based on the resilient propagation algorithm. *ELSEVIER*.
<https://www.sciencedirect.com/science/article/abs/pii/S0925231204005168>
- Arias-Gómez, J., Villasís-Keever, M. Á., & Miranda-Novales, M. G. (2016). El protocolo de investigación III. La población de Estudio. *Revista Alergia Mexico*, 63(2), 201–206. <https://doi.org/10.29262/ram.v63i2.181>
- Barrientos, Z., & Umaña, R. (2010). Deserción estudiantil en posgrados semipresenciales de la Universidad Estatal a Distancia (UNED), Costa Rica: ¿Deserción o retraso? *UNED Research Journal*, 1(2), 141–149. <https://doi.org/10.22458/urj.v1i2.227>
- Bedregal-Alpaca, N., Tupacyupanqui-Jaén, D., & Cornejo-Aparicio, V. (2020). Análisis



del rendimiento académico de los estudiantes de Ingeniería de Sistemas, posibilidades de deserción y propuestas para su retención. *Ingeniare. Revista Chilena de Ingeniería*, 28(4), 668–683. <https://doi.org/10.4067/s0718-33052020000400668>

Brownlee, J. (2017). *How to Develop Word Embeddings in Python with Gensim*. <https://machinelearningmastery.com/develop-word-embeddings-python-gensim/>

Calla Charrez, V. (2019). *Modelo de pronóstico para número de egresos hospitalarios usando redes neuronales - Hospital Regional Manuel Nuñez Butron - Puno, 2009 - 2018* [Universidad Nacional del Altiplano Puno]. <http://repositorio.unap.edu.pe/handle/UNAP/11474>

Canales, F., Alvarado, E. L., & Pineda, E. B. (1994). *Metodología de la investigación. Manual para el desarrollo de personal de salud* (2da ed.). Organización panamericana de la salud. <https://www.passeidireto.com/arquivo/91913304/pineda-metodologia-de-la-investigacion-manual-para-el-desarrollo-de-personal-de->

Carpio Contreras, A. L. (2017). *Frecuencia y factores asociados a la deserción universitaria en estudiantes de la escuela de posgrado modalidad a distancia de la universidad Católica de Santa María. Arequipa 2013 - 2015* [Universidad Católica de Santa María]. <http://tesis.ucsm.edu.pe/repositorio/bitstream/handle/UCSM/6696/60.1332.EN.pdf?sequence=1&isAllowed=y>

Chavarri Becerra, J. (2018). *Modelo de predicción para determinar el riesgo de deserción en estudiantes de la facultad de Ingeniería y Arquitectura de la Universidad*



Peruana Unión Filial Juliaca 2017.

<http://repositorio.unap.edu.pe/handle/UNAP/7748>

Chen, Y., Song, L., Liu, Y., Yang, L., & Li, D. (2020). A review of the artificial neural network models for water quality prediction. *Applied Sciences (Switzerland)*, *10*(17). <https://doi.org/10.3390/app10175776>

Dey, A. (2016). Machine Learning Algorithms: A Review. *International Journal of Computer Science and Information Technologies*, *7*(3), 1174–1179. <https://ijcsit.com/docs/Volume 7/vol7issue3/ijcsit2016070332.pdf>

EPG. (2020). REGLAMENTO GENERAL DE LA ESCUELA DE POSGRADO. *2*, *1*(051), 144. <https://transparencia.unap.edu.pe/documentos/academicos/REGLAMENTO GENERAL EPG.pdf>

Erickson, B. J., Korfiatis, P., Akkus, Z., & Kline, T. L. (2017). Machine learning for medical imaging. *Radiographics*, *37*(2), 505–515. <https://doi.org/10.1148/rg.2017160130>

Espinosa-Castro, J.-F., Hernández-Lalinde, J., & Mariño Castro, L. (2020). Estrategias de permanencia universitaria. *AVFT – Archivos Venezolanos de Farmacología y Terapéutica*, *39*(1), 88–97. https://www.revistaavft.com/images/revistas/2020/avft_1_2020/16_estrategias.pdf

Flach, P. A. (2003). The Geometry of ROC Space: Understanding Machine Learning Metrics through ROC Isometrics. *Proceedings, Twentieth International Conference on Machine Learning*, *1*, 194–201.



<https://www.aaai.org/Papers/ICML/2003/ICML03-028.pdf>

Guevara Paredes, C. J. (2018). *Factores de la deserción forzosa institucional de estudiantes de Estudios Generales Letras de la PUCP* [Pontificia Universidad Católica del Perú]. <http://hdl.handle.net/20.500.12404/13561>

Hernández-Samipieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2003). El proceso de investigación y los enfoques cuantitativo y cualitativo. Hacia un modelo integral. In *Metodología de la investigación* (1 edición, Vol. 7, Issue 1, pp. 37–72). <http://metodos-comunicacion.sociales.uba.ar/wp-content/uploads/sites/219/2014/04/Hernandez-Sampieri-Cap-1.pdf>

Hernández-Samipieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2014). Selección de la muestra. In *Metodología de la investigación: Vol. 6ta edición* (6ta. ed, pp. 170–196). Espacio de formación multimodal. http://metabase.uaem.mx/xmlui/bitstream/handle/123456789/2776/506_6.pdf

Hernandez, O. K., Perez, R., & Gonzáles, G. (2014). La deserción en los posgrados, un problema no menor. *Red de Revistas Científicas de América Latina, El Caribe, España y Portugal*, 19. <https://www.redalyc.org/pdf/5534/553457062009.pdf>

Himmel, E. (2002). Modelo de análisis de la deserción estudiantil en la educación superior. *Calidad En La Educación*, 17, 91. <https://doi.org/10.31619/caledu.n17.409>

Igel, C., & Husken, M. (2003). Empirical evaluation of the improved Rprop learning algorithms. *ELSEVIER*. [https://doi.org/https://doi.org/10.1016/S0925-2312\(01\)00700-7](https://doi.org/https://doi.org/10.1016/S0925-2312(01)00700-7)



- International Council on Archives. (2016). *Los archivos*. International Council on Archives. <https://www.ica.org/es/que-es-un-documento-de-archivo>
- Larrañaga, P., Inza, I., & Moujahid, A. (2016). Tma 8. Redes Neuronales. In *Redes Neuronales* (pp. 1–19). https://www.researchgate.net/profile/Pedro-Larranaga/publication/268291232_Tema_8_Redес_Neuronales/links/55b7b5c408ae9289a08c0c68/Tema-8-Redes-Neuronales.pdf
- Levisi Quispe, C. J. (2017). Factores que influyen en la deserción de los estudiantes de la Facultad de Trabajo Social de la Universidad Nacional del Altiplano de Puno 2014 - 2015. [Universidad Nacional del Altiplano]. In *Universidad Nacional del Altiplano Facultad de Trabajo Social Curso de Actualización*. http://repositorio.unap.edu.pe/bitstream/handle/UNAP/8037/Levisi_Quispe_Carmen_Julia.pdf?sequence=1&isAllowed=y
- Lopez, P. L. (2004). Población, muestra y muestreo. *Punto Cero*, 129–139. <http://www.scielo.org.bo/pdf/rpc/v09n08/v09n08a12.pdf>
- Marin Mendez, J. J. (2019). *Desarrollo de un algoritmo con machine learning para la clasificación de pacientes con Parkinson* [Universidad Autónoma de Madrid]. https://repositorio.uam.es/bitstream/handle/10486/688837/marin_mendez_juan_jesús_tfm.pdf?sequence=1&isAllowed=y
- Matich, D. (2021). Redes Neuronales: Conceptos Básicos y Aplicaciones. *Facultad Tecnológica Nacional*, 55. <ftp://decsai.ugr.es/pub/usuarios/castro/Material-Redes-Neuronales/Libros/matich-redesneuronales.pdf>
- Miranda, B., Gonzales, V., Reyes, A., Alejo, R., & Rendon, E. (2020). Procesamiento de bases de datos escolares por medio de redes neuronales artificiales. *Ciencia Ergo-*



Sum, 27(3). <https://cienciaergosum.uaemex.mx/article/view/13136>

Miranda, M. A., & Guzmán, J. (2017). Análisis de la deserción de estudiantes universitarios usando técnicas de minería de datos. *Formacion Universitaria*, 10(3), 61–68. <https://doi.org/10.4067/S0718-50062017000300007>

Paja Dominguez, H. E. (2017). *Prediccion de rendimiento academico mediante regresion y redes neuronales en los estudiantes de la Escuela profesional de Ingenieria Estadistica e Informatica de la Universidad Nacional del Altiplano - Puno, 2015*. [Universidad Nacional del Altiplano Puno]. <http://repositorio.unap.edu.pe/handle/UNAP/6337>

Piscoya, L. (2016). *APLICACIÓN DE TÉCNICAS DE MINERÍA DE ESTUDIANTIL EN LA EDUCACION BASICA REGULAR EN LA REGION DE LAMBAYEQUE* [Universidad Señor de Sipán]. <https://repositorio.uss.edu.pe/bitstream/handle/20.500.12802/4066/Tesisterminada.pdf?sequence=1&isAllowed=y>

Prieto Hernandez, A. M. (2017). *Tipos de Investigación*. Universidad Abierta y a Distancia de Mexico. <https://jose-mario-balderas-reyes-unadm.blogspot.com/2017/08/tipos-de-investigacion.html>

Quispe, A. (2018). *Formulacion de un modelos de red neuronal artificial para el pronostico de concentracion de oxigeno disuelto y clorofina en la bahia interior de Puno* [Universidad Nacional del Altiplano Puno]. <http://repositorio.unap.edu.pe/handle/UNAP/11018>

Ramírez Yparraguirre, M. Y. (2017). Factores individuales y de contexto que inciden en la deserción universitaria de los estudiantes del programa Beca 18. *Universidad*



César Vallejo. <http://repositorio.ucv.edu.pe/handle/UCV/5333>

- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2017). Deep Learning for Health Informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1), 4–21. <https://doi.org/10.1109/JBHI.2016.2636665>
- Repetur, A. (2019). Redes Neuronales Artificiales [Universidad Nacional del Centro]. In *RIDAA UNICEN*. [https://doi.org/10.1016/S0210-5691\(05\)74198-X](https://doi.org/10.1016/S0210-5691(05)74198-X)
- Roche, A. (2014). *Introducción a machine learning*. GPMESS. <https://gpmess.com/blog/2014/07/03/introduccion-machine-learning/>
- Rodríguez, M., & Mora, R. (2001). Análisis de Regresión Múltiple. In *Técnicas de investigación social II*. (pp. 3–17). <https://doi.org/84-7908-638-6>
- Rodríguez Núñez, L. H., & Londoño Londoño, F. J. (2010). Estudio sobre deserción estudiantil en los programas de Educación de la Católica del Norte Fundación Universitaria A Study on Student's Dropout in Education Programs of the Catholic of the North University Foundation Étude sur l'abandon scolaire aux progr. *Católica Del Norte Fundacio Universitaria*, 33. <https://www.redalyc.org/pdf/1942/194218961018.pdf>
- Salgado Reyes, N., Beltrán Morales, J., Guaña Moya, J., Escobar Teran, C., Nicolalde Rodríguez, D., & Chafía Altamirano, G. (2018). Modelo para predecir el rendimiento académico basado en redes neuronales y analítica de aprendizaje. *Revista Ibérica de Sistemas e Tecnologías de Información*, 258–266. <https://search.proquest.com/openview/5720c78f2e17a27355a8766fe81feb1a/1?q-origsite=gscholar&cbl=1006393>



- Sanchez Nina, A. (2017). *Modelo estadístico para determinar la deserción estudiantil de las escuelas profesionales de la UNA - PUNO, 2017* [Universidad Nacional del Altiplano Puno]. <http://repositorio.unap.edu.pe/handle/UNAP/6297>
- SUNEDU. (2016). *Licenciamiento Institucional*.
<https://www.sunedu.gob.pe/licenciamiento-institucional/>
- Terán-villanueva, J. D., Ibarra-martínez, S., Laria-menchaca, J., Castán-rocha, J. A., García-ruiz, A. H., & Martínez-infante, J. E. (2019). Estudio de redes neuronales para el pronóstico de la demanda de asignaturas. *Tunja-Boyaca*, 28(50).
<http://www.scielo.org.co/pdf/rfing/v28n50/0121-1129-rfing-28-50-00034.pdf>
- Tinto, V. (1982). Definir la deserción: una cuestión de perspectiva. *Revista de La Educación Superior*, 1–9.
http://publicaciones.anui.es.mx/pdfs/revista/Revista71_S1A3ES.pdf
- Valero, C. S. (2018). Aplicación de métodos de aprendizaje automático en el análisis y la predicción de resultados deportivos. *Retos*, 2041(34), 377–382.
<https://doi.org/10.47197/retos.v0i34.58506>
- Villasís-Keever, M. Á., & Miranda-Navales, M. G. (2016). El protocolo de investigación IV: las variables de estudio. *Revista Alergia México*, 63(3), 303–310.
<https://doi.org/10.29262/ram.v63i3.199>
- Zarria Torres, C., Arce Ramos, C., & Lam Moraga, J. (2016). *Estudio de variables que influyen en la deserción de estudiantes universitarios de primer año, mediante minería de datos*. 6(1), 73–84.
<https://dialnet.unirioja.es/servlet/articulo?codigo=5608574>



Zevallos Salazar, R. J. (2017). *Predicción del rendimiento académico mediante redes neuronales* [Universidad Nacional del Altiplano Puno].
[http://repositorio.unac.edu.pe/bitstream/handle/UNAC/2728/Zevallos Salazar_TESIS_2017.pdf?sequence=1&isAllowed=y](http://repositorio.unac.edu.pe/bitstream/handle/UNAC/2728/Zevallos_Salazar_TESIS_2017.pdf?sequence=1&isAllowed=y)



ANEXOS



Anexo 1. Datos recopilados

muestra	promedio	edad	sexo	estado civil	procedencia	debe	tipo universidad	deserción
1	9	34	1	0	0	0	1	1
2	18	33	1	1	0	0	1	0
3	4.6	35	1	0	0	0	0	1
4	9.29	36	1	0	0	0	1	1
5	14.25	30	0	0	0	0	0	0
6	14.5	32	1	2	0	0	0	0
7	9.25	47	1	0	0	0	0	1
8	18.5	26	0	0	0	0	0	0
9	17.15	50	1	0	0	1	1	0
10	15.7	28	1	0	0	1	1	0
11	16.7	32	1	0	0	0	1	0
12	15.3	28	1	0	0	0	1	0
13	7	34	0	1	0	0	1	1
14	7	34	0	0	1	1	1	1
15	8.2	35	0	0	0	0	0	1
16	14.3	26	1	0	1	1	1	0
17	7	37	1	0	0	0	0	1
18	7.9	44	1	0	0	1	1	1
19	7.1	37	0	1	0	0	1	1
20	7.9	37	1	1	1	1	1	1
21	8.4	49	0	1	0	0	1	1
22	8.2	36	1	1	0	0	1	1



23	7.9	36	0	0	1	0	0	1
24	8.7	35	1	1	0	0	1	1
25	8.3	36	1	0	0	0	1	1
26	7.9	41	1	1	0	1	0	1
27	8.6	34	1	2	1	1	1	1
28	18.6	39	1	2	0	1	1	0
29	17.9	38	1	2	0	1	0	0
30	16.5	29	0	2	0	1	1	0
31	7.3	34	1	2	0	0	0	1
32	9.6	27	0	2	0	1	0	1
33	8.3	29	0	2	1	1	0	1
34	3.5	37	1	2	1	0	0	1
35	14.6	37	1	2	0	0	0	0
36	15.7	37	0	2	0	1	0	0
37	15.6	37	1	2	0	0	1	0
38	17.1	34	0	2	1	1	1	0
39	19	30	1	1	1	0	1	0
40	15.2	30	0	0	0	1	1	0
41	18.9	33	1	0	0	1	1	0
42	16.1	28	1	0	0	1	1	0
43	18.9	25	0	0	0	1	1	0
44	14.2	31	0	1	0	1	1	0
45	15	24	1	0	0	1	0	0
46	16.7	31	0	1	0	1	1	0



47	18.1	26	0	0	0	1	1	0
48	16.9	29	1	1	0	1	0	0
49	14.6	25	0	0	0	1	0	0
50	14.6	32	1	1	1	1	0	0
51	14.9	31	0	0	0	1	0	0
52	15	28	0	0	0	1	0	0
53	15	25	0	0	0	1	0	0
54	15.8	26	0	0	1	1	1	0
55	14.6	31	1	1	0	1	0	0
56	15.8	31	1	1	0	1	0	0
57	16.2	33	0	0	0	1	1	0
58	15.4	35	1	2	0	1	1	0
59	15.7	24	1	0	1	1	0	0
60	14.9	26	1	0	0	1	0	0
61	17.2	26	1	0	0	1	0	0
62	4	30	1	1	0	0	0	1
63	15.2	35	0	1	1	1	0	0
64	14.5	26	1	0	0	0	0	0
65	15.4	25	0	0	0	1	1	0
66	8.5	32	1	1	0	1	0	1
67	17.4	33	0	2	0	1	1	0
68	14.3	26	0	0	1	1	0	0
69	14.2	31	1	1	0	1	0	0
70	16.5	35	1	2	0	0	0	0



71	14.6	24	1	0	0	1	1	0
72	14.8	25	0	0	1	1	0	0
73	14.7	39	0	2	0	1	0	0
74	14.5	34	0	0	0	1	1	0
75	16.4	25	1	0	0	1	0	0
76	15.1	26	1	0	0	1	1	0
77	17.2	28	0	0	0	1	0	0
78	3	32	1	2	0	1	0	1
79	14.9	31	1	1	0	1	1	0
80	15.7	24	0	0	0	1	0	0
81	9	29	1	1	0	0	1	1
82	16.3	33	0	2	0	1	0	0
83	16	26	1	2	1	1	0	0
84	16.5	33	1	2	0	1	0	0
85	15.8	36	1	2	0	1	0	0
86	16.8	38	0	2	0	1	0	0
87	16.5	30	1	1	1	1	1	0
88	17.1	30	1	1	0	1	1	0
89	16.2	33	1	1	0	1	1	0
90	16.5	26	1	0	0	1	1	0
91	8.3	25	1	1	1	0	1	1
92	16.5	30	1	1	0	0	0	0
93	17.4	33	0	1	0	1	0	0
94	15	25	1	0	0	1	0	0



95	14.3	34	0	1	1	0	1	0
96	16	26	0	0	0	1	0	0
97	17.1	29	0	1	0	0	1	0
98	16	27	0	1	0	1	0	0
99	16.8	30	0	1	1	1	1	0
100	16.6	27	1	1	0	1	1	0
101	16.8	32	1	1	0	1	0	0
102	14.7	30	1	1	0	1	1	0
103	14.9	25	1	0	0	1	0	0
104	14.6	35	1	1	1	0	1	0
105	14.6	26	0	0	0	1	0	0
106	15.1	27	0	3	0	1	1	0
107	15.2	29	1	3	0	1	0	0
108	16	25	1	2	1	1	1	0
109	10.2	32	0	0	0	1	0	1
110	15.6	30	0	1	1	0	0	0
111	14.9	28	1	3	0	1	0	0
112	14.6	25	1	0	0	1	0	0
113	15.6	31	0	1	0	0	0	0
114	15.4	31	0	0	0	1	1	0
115	14.4	27	0	1	1	1	1	0
116	15	30	0	1	1	1	0	0
117	14.6	31	1	0	1	0	0	1
118	14.2	27	1	2	1	0	0	0



119	14.2	28	0	0	0	0	1	0
120	14.4	31	0	0	0	0	0	0
121	14.7	32	1	0	1	1	1	0
122	16.6	32	1	0	1	0	1	1
123	16.6	29	0	2	1	1	0	0
124	14	26	0	1	1	1	0	1
125	8.5	31	0	2	1	1	1	1
126	6.3	28	1	0	0	1	1	1
127	14.4	30	1	0	0	0	1	0
128	14.4	29	1	2	1	1	0	0
129	14.6	26	0	1	1	1	0	0
130	14.3	28	0	2	0	1	1	0
131	10.8	30	0	2	0	1	0	1
132	14.9	31	1	2	1	1	0	0
133	10.2	28	0	0	1	0	1	1
134	14.3	30	0	0	1	1	0	0
135	10.6	28	1	0	1	1	0	1
136	14.9	25	0	1	1	1	1	1
137	15.3	27	0	2	0	0	1	1
138	15.2	25	0	1	0	1	1	1
139	15.1	27	0	2	0	1	1	1
140	15.3	31	0	2	1	1	0	1
141	14.7	25	0	2	0	1	0	0
142	16.2	32	1	1	1	1	1	0



143	10.7	31	1	1	0	0	0	1
144	15.1	27	1	1	0	0	0	0
145	16.4	26	1	0	1	1	0	0
146	14.6	41	1	1	1	1	1	0
147	14.6	30	1	2	1	1	0	0
148	15.9	25	0	0	1	0	0	0
149	13.2	32	1	2	1	1	1	1
150	17.3	29	0	1	1	1	0	0
151	14.3	29	0	2	1	1	0	0
152	15.8	32	0	0	1	1	1	0
153	14.3	25	1	0	0	1	1	0
154	13.1	32	1	2	0	1	0	1
155	14.1	28	0	0	1	0	1	0
156	16.2	27	0	1	1	1	0	0
157	14.7	25	1	1	0	1	1	0
158	14	27	1	1	1	1	0	0
159	14.6	28	0	2	0	1	1	0
160	16.3	26	1	0	0	1	1	0
161	14.8	28	0	2	0	0	0	0
162	14.2	31	0	2	1	1	0	0
163	16	32	0	1	0	1	0	0
164	14.4	29	1	1	0	1	0	0
165	14.6	28	0	0	1	1	1	0
166	14.6	32	1	0	1	0	0	0



167	16	27	1	2	1	0	0	0
168	14.9	30	1	2	0	0	1	0
169	14.6	26	0	1	1	1	1	0
170	16	28	0	1	1	0	1	0
171	16.6	32	0	0	0	1	0	0
172	14.9	27	0	2	1	1	0	0
173	14.6	25	0	0	1	1	0	0
174	14.8	25	0	1	1	0	0	0
175	14.3	30	1	1	0	1	1	0
176	11.3	27	0	2	1	1	0	1
177	14	27	1	2	1	1	1	0
178	13.2	29	0	1	1	1	0	1
179	14.3	26	1	0	0	0	1	0
180	14.6	25	0	0	1	1	0	0
181	14	28	1	2	0	0	0	0
182	15	26	0	0	0	1	1	0
183	16	25	0	2	0	0	0	0
184	12	32	1	2	0	0	1	1
185	10	36	1	2	0	0	1	1
186	7.2	31	1	1	0	0	0	1
187	9.2	37	1	3	1	0	1	1
188	10.6	30	1	2	0	0	0	1
189	16	30	1	2	1	0	1	0
190	7.8	31	0	1	0	0	0	1



191	4.3	36	0	1	0	0	1	1
192	8.2	36	0	2	0	0	1	1
193	8.1	31	1	1	0	0	1	1
194	15	32	1	1	1	0	1	0
195	14.6	30	1	1	1	0	0	0
196	15.7	37	1	3	0	0	1	0
197	14.9	30	0	1	0	0	1	0
198	15.9	30	1	3	0	0	1	0
199	15.8	37	1	1	0	0	0	0
200	15	31	1	2	0	0	1	0
201	14.2	30	1	2	1	0	1	0
202	14.2	36	1	1	1	1	1	0
203	15.2	35	1	1	0	0	0	0
204	13.7	30	1	2	0	1	0	0
205	14.3	32	0	2	0	1	1	0
206	15.3	35	1	2	0	1	1	0
207	16.1	30	1	1	0	0	1	0
208	15.7	31	1	2	1	1	0	0
209	14.6	34	0	2	1	1	1	0
210	15.5	32	1	2	1	0	0	0
211	14.2	37	0	2	0	1	0	0
212	14.8	36	1	1	1	1	0	0
213	14	30	1	1	1	1	0	0
214	14.2	34	1	1	0	1	0	0



215	14.3	35	0	2	0	1	1	0
216	15.2	30	1	1	1	0	0	0
217	14.3	31	1	1	0	0	1	0
218	16.6	30	0	0	0	0	1	0
219	16.5	33	0	0	1	0	0	0
220	15.6	45	1	3	0	1	0	0
221	15.2	33	1	2	0	1	0	0
222	14.1	34	1	1	0	1	0	0
223	15.3	33	1	1	1	0	0	0
224	14.9	35	0	1	0	1	1	0
225	15.6	33	1	0	0	0	0	0
226	15.4	32	1	0	0	0	0	0
227	14.7	31	1	2	0	0	1	0
228	14.6	33	0	1	0	1	1	0
229	14.2	34	0	2	0	1	0	0
230	14.2	33	1	0	1	0	1	0
231	14.7	31	1	1	1	1	0	0
232	15.6	32	1	1	1	1	1	0
233	15.2	35	1	1	0	1	1	0
234	15.2	37	1	2	1	1	0	0
235	15.6	30	1	1	1	0	1	0
236	15.2	33	1	2	1	1	1	0
237	14.3	37	1	2	0	1	1	0
238	14.8	45	1	2	1	0	1	0



239	16	34	0	2	1	0	1	0
240	14	36	1	0	1	0	0	0
241	14.1	30	1	0	1	1	0	0
242	16	32	1	1	0	1	1	0
243	16.5	34	0	1	0	1	1	0
244	14.9	32	1	2	0	1	0	0
245	16.3	37	0	1	0	1	0	0
246	14.1	35	1	1	0	0	0	0
247	15.2	31	0	0	0	1	0	0
248	14.2	37	1	2	0	0	1	0
249	15.1	31	0	2	0	0	1	0
250	14	31	1	1	0	1	1	0
251	16	30	1	0	1	0	0	0
252	14.6	34	1	2	0	1	1	0
253	14.4	37	0	2	1	0	1	0
254	16	30	0	1	0	1	1	0
255	14.3	37	0	2	0	1	1	0
256	14	34	1	2	0	1	1	0
257	16.8	33	1	2	0	1	1	0
258	14.7	35	1	2	0	0	1	0
259	14.6	36	0	2	0	0	0	0
260	13.3	36	1	2	0	1	0	0
261	15.5	31	1	2	1	0	1	0
262	14	37	0	1	0	0	1	0



263	15	35	1	2	0	0	0	0
264	16	36	1	1	0	1	0	0
265	15	34	0	1	1	1	1	0
266	14.9	35	1	1	1	0	0	0
267	15	36	1	2	0	1	1	0
268	14	31	1	0	0	1	1	0
269	14.4	30	0	2	0	0	1	0
270	14	34	1	0	0	1	0	0
271	14	28	0	1	0	1	0	0
272	14.1	33	0	1	0	0	0	0
273	9.9	27	0	0	1	0	0	1
274	14.6	35	1	1	1	0	1	0
275	14.6	33	1	1	0	0	1	0
276	16	29	0	0	0	0	1	0
277	17	28	0	1	1	0	0	0
278	14.1	34	0	0	0	1	1	0
279	15	32	0	0	0	0	1	0
280	16.2	33	0	1	0	0	0	0
281	16.2	31	1	1	0	1	1	0
282	16.2	29	1	0	0	0	0	0
283	16.2	29	1	1	0	0	0	1
284	17.5	29	1	1	0	1	1	1
285	16	28	1	1	0	1	1	1
286	15	34	1	1	1	1	1	0



287	13	26	0	0	0	1	0	1
288	15.8	35	0	0	0	1	0	0
289	14.8	35	1	1	0	1	1	0
290	10.1	28	1	1	1	0	1	1
291	11	30	1	0	0	0	1	1
292	14	26	1	1	1	0	1	0
293	14	27	0	1	1	1	1	1
294	11.3	32	1	0	0	1	0	1
295	16	28	0	1	1	1	0	0
296	17.6	33	0	1	1	0	1	0
297	14	33	1	1	0	1	1	0
298	14.5	29	0	1	1	0	0	0
299	10.5	30	1	0	0	0	1	1
300	14	33	1	1	1	1	0	0
301	15.9	33	1	1	1	1	1	0
302	14.9	28	0	1	0	1	0	0
303	14	27	1	1	1	0	1	0
304	14.6	30	0	0	1	1	0	0
305	14	31	0	1	0	0	0	0
306	14	29	0	1	1	0	1	0
307	13.9	27	0	1	0	1	1	1
308	14.7	29	1	1	1	0	0	0
309	16.5	29	0	0	1	0	1	0
310	15	26	0	1	1	0	0	0



311	11.6	30	0	1	0	1	1	1
312	12	28	1	1	1	1	1	1
313	12.4	34	0	0	1	1	1	1
314	14.3	31	0	0	1	1	1	0
315	14	30	0	1	0	1	0	0
316	10.8	30	0	1	0	0	1	1
317	13.2	27	1	0	1	1	0	1
318	13.5	33	0	0	0	1	1	1
319	14.9	33	1	1	1	1	0	0
320	16	27	1	1	0	1	0	0
321	16	34	1	0	0	1	1	1
322	9.9	34	0	0	0	1	0	1
323	11.4	27	1	1	1	1	0	1
324	10.2	27	1	0	0	1	0	1
325	14.6	31	1	0	0	1	0	0
326	15	28	1	1	1	0	0	0
327	14	27	1	0	1	0	1	0
328	9.7	29	0	0	0	1	0	1
329	9.8	33	1	0	1	1	1	1
330	11.8	28	1	1	0	1	1	1
331	12.4	29	1	0	1	0	1	1
332	15	27	1	0	0	0	0	0
333	15.6	30	0	0	0	1	0	0
334	11.9	27	0	0	1	1	1	1



Anexo 2. Código utilizado en R

```
#análisis descriptivo
```

```
library(ggplot2)
```

```
library(caret)
```

```
library(neuralnet)
```

```
library(dplyr)
```

```
library(MLmetrics)
```

```
library(ggpubr)
```

```
datos <- read.csv(" datos.csv", sep=";")
```

```
datos$desercion<-as.factor(datos$desercion)
```

```
common_theme <- theme(plot.title = element_text(hjust = 0.5, face = "bold"))
```

```
gghistogram(datos, x = "promedio", fill = "lightgray",
```

```
add = "mean", rug = TRUE)+
```

```
labs(x = 'Promedio de notas', y = 'Conteo')
```

```
gghistogram(datos, x = "edad", fill = "lightgray",
```

```
add = "mean", rug = TRUE)+
```

```
labs(x = 'Edad del estudiante', y = 'Conteo')
```

```
ggplot(data = datos, aes(x = factor(sexo),
```

```
  y = prop.table(stat(count)), fill = factor(sexo),
```

```
  label = scales::percent(prop.table(stat(count)))))) +
```

```
geom_bar(position = "dodge") +
```

```
geom_text(stat = 'count',
```

```
position = position_dodge(.9),
```



```
vjust = -0.5,  
  
size = 3) +  
  
scale_x_discrete(labels = c("mujeres", "hombres"))+  
  
scale_y_continuous(labels = scales::percent)+  
  
labs(x = 'Sexo', y = 'Porcentaje') +  
  
ggtitle("") +  
  
common_theme  
  
ggplot(data = datos, aes(x = factor(debe),  
  
  y = prop.table(stat(count)), fill = factor(debe),  
  
  label = scales::percent(prop.table(stat(count)))))) +  
  
geom_bar(position = "dodge") +  
  
geom_text(stat = 'count',  
  
position = position_dodge(.9),  
  
vjust = -0.5,  
  
size = 3) +  
  
scale_x_discrete(labels = c("Debe", "No debe"))+  
  
scale_y_continuous(labels = scales::percent)+  
  
labs(x = 'Estado de cuenta', y = 'Porcentaje') +  
  
ggtitle("") +  
  
common_theme  
  
ggplot(data = datos, aes(x = factor(estado.civil),  
  
  y = prop.table(stat(count)), fill = factor(estado.civil),  
  
  label = scales::percent(prop.table(stat(count)))))) +  
  
geom_bar(position = "dodge") +  
  
geom_text(stat = 'count',
```



```
position = position_dodge(.9),  
  
vjust = -0.5,  
  
size = 3) +  
  
scale_x_discrete(labels = c("Solteros", "Convivientes", 'Casados', 'Divorciados'))+  
  
scale_y_continuous(labels = scales::percent)+  
  
labs(x = 'Estado civil', y = 'Porcentaje') +  
  
ggtitle("") +  
  
common_theme  
  
ggplot(data = datos, aes(x = factor(procedencia),  
  
  y = prop.table(stat(count)), fill = factor(procedencia),  
  
  label = scales::percent(prop.table(stat(count)))))) +  
  
geom_bar(position = "dodge") +  
  
geom_text(stat = 'count',  
  
position = position_dodge(.9),  
  
vjust = -0.5,  
  
size = 3) +  
  
scale_x_discrete(labels = c("Provincia de Puno", "Otro lugar"))+  
  
scale_y_continuous(labels = scales::percent)+  
  
labs(x = 'Procedencia', y = 'Porcentaje') +  
  
ggtitle("") +  
  
common_theme  
  
ggplot(data = datos, aes(x = factor(tipo.universidad),  
  
  y = prop.table(stat(count)), fill = factor(tipo.universidad),  
  
  label = scales::percent(prop.table(stat(count)))))) +  
  
geom_bar(position = "dodge") +
```



```
geom_text(stat = 'count',  
  
position = position_dodge(.9),  
  
vjust = -0.5,  
  
size = 3) +  
  
scale_x_discrete(labels = c("privado", "estatal"))+  
  
scale_y_continuous(labels = scales::percent)+  
  
labs(x = 'Tipo de universidad', y = 'Porcentaje') +  
  
ggtitle("") +  
  
common_theme  
  
ggplot(data = datos, aes(x = factor(tipo.universidad),  
  
y = prop.table(stat(count)), fill = factor(tipo.universidad),  
  
label = scales::percent(prop.table(stat(count)))))) +  
  
geom_bar(position = "dodge") +  
  
geom_text(stat = 'count',  
  
position = position_dodge(.9),  
  
vjust = -0.5,  
  
size = 3) +  
  
scale_x_discrete(labels = c("privado", "estatal"))+  
  
scale_y_continuous(labels = scales::percent)+  
  
labs(x = 'Tipo de universidad', y = 'Porcentaje') +  
  
ggtitle("") +  
  
common_theme  
  
ggplot(data = datos, aes(x = factor(desercion),  
  
y = prop.table(stat(count)), fill = factor(desercion),  
  
label = scales::percent(prop.table(stat(count)))))) +
```



```
geom_bar(position = "dodge") +  
geom_text(stat = 'count',  
position = position_dodge(.9),  
vjust = -0.5,  
size = 3) +  
scale_x_discrete(labels = c("Continua estudiando", "Deserto por diversos motivos"))+  
scale_y_continuous(labels = scales::percent)+  
labs(x = 'Clase', y = 'Porcentaje') +  
ggtitle("") +  
common_theme  
  
ggplot(data=datos, aes(x=promedio,color=desercion,fill=desercion))+  
geom_histogram()+common_theme  
  
ggplot(data=datos, aes(x=edad,color=desercion,fill=desercion))+  
geom_histogram(position="dodge")+common_theme+  
labs(x = 'Edad', y = 'Conteo')  
  
ggplot(data=datos, aes(x=promedio,color=desercion,fill=desercion))+  
geom_histogram()+common_theme+ labs(x = 'Promedio de notas', y = 'Conteo')  
  
ggplot(data=datos, aes(x=edad,color=desercion,fill=desercion))+  
geom_histogram(position="dodge")+common_theme  
  
ggplot(data = datos, aes(x = factor(sexo),  
y = prop.table(stat(count)), fill = factor(desercion),  
label = scales::percent(prop.table(stat(count)))))) +
```



```
geom_bar(position = "dodge") +  
geom_text(stat = 'count',  
position = position_dodge(.9),  
vjust = -0.5,  
size = 3) +  
scale_x_discrete(labels = c("mujeres", "Hombres"))+  
scale_y_continuous(labels = scales::percent)+  
labs(x = 'Sexo', y = 'Porcentaje') +  
ggtitle("") +  
common_theme  
  
ggplot(data = datos, aes(x = factor(estado.civil),  
y = prop.table(stat(count)), fill = factor(desercion),  
label = scales::percent(prop.table(stat(count)))))) +  
geom_bar(position = "dodge") +  
geom_text(stat = 'count',  
position = position_dodge(.9),  
vjust = -0.5,  
size = 3) +  
scale_x_discrete(labels = c("soltero", "conviviente", 'casado', 'divorciado'))+  
scale_y_continuous(labels = scales::percent)+  
labs(x = 'estado civil', y = 'Porcentaje') +  
ggtitle("") +  
common_theme  
  
ggplot(data = datos, aes(x = factor(procedencia),
```



```
y = prop.table(stat(count)), fill = factor(desercion),  
label = scales::percent(prop.table(stat(count)))) +  
geom_bar(position = "dodge") +  
geom_text(stat = 'count',  
position = position_dodge(.9),  
vjust = -0.5,  
size = 3) +  
scale_x_discrete(labels = c('Provincia de Puno','Otro lugar'))+  
scale_y_continuous(labels = scales::percent)+  
labs(x = 'procedencia', y = 'Porcentaje') +  
ggtitle("") +  
common_theme
```

```
ggplot(data = datos, aes(x = factor(debe),  
y = prop.table(stat(count)), fill = factor(desercion),  
label = scales::percent(prop.table(stat(count)))) +  
geom_bar(position = "dodge") +  
geom_text(stat = 'count',  
position = position_dodge(.9),  
vjust = -0.5,  
size = 3) +  
scale_x_discrete(labels = c('Debe','No debe'))+  
scale_y_continuous(labels = scales::percent)+  
labs(x = 'Estado de cuenta', y = 'Porcentaje') +  
ggtitle("") +
```



common_theme

```
ggplot(data = datos, aes(x = factor(tipo.universidad),  
  y = prop.table(stat(count)), fill = factor(desercion),  
  label = scales::percent(prop.table(stat(count)))))) +  
geom_bar(position = "dodge") +  
geom_text(stat = 'count',  
position = position_dodge(.9),  
vjust = -0.5,  
size = 3) +  
scale_x_discrete(labels = c('Estatad', 'Privada'))+  
scale_y_continuous(labels = scales::percent)+  
labs(x = 'Tipo de universidad', y = 'Porcentaje') +  
ggtitle("") +  
common_theme
```

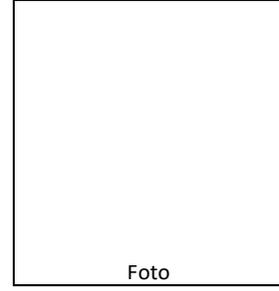


Anexo 3. Ficha de Inscripción
UNIVERSIDAD NACIONAL DEL ALTIPLANO
ESCUELA DE POSGRADO



FICHA DE INSCRIPCION

CODIGO Y PASSWORD				
CODIGO PAGO BANCO SCOTIABANK		CONTRASEÑA		
DATOS PERSONALES				
APELLIDO PATERNO		APELLIDO MATERNO		
NOMBRES				
DNI	FECHA DE NACIMIENTO	SEXO	ESTADO CIVIL	
EMAIL				
CELULAR	DIRECCION			
LUGAR DE NACIMIENTO				
DEPARTAMENTO		PROVINCIA	DISTRITO	
DATOS ACADEMICOS				
TITULO O NIVEL ACADEMICO				
CENTRO DE ESTUDIOS				
PROGRAMA Y MENCION A LOS QUE POSTULA				
PROGRAMA		MENCION		
ESTUDIOS DE POSGRADO				
ESTUDIOS DE POSGRADO				
DATOS DE LA INSTITUCION DE PROCEDENCIA (CENTRO DE TRABAJO)				
INSTITUCION				
DIRECCION				
UBIGEO INSTITUCION PROCEDENCIA				
DEPARTAMENTO	PROVINCIA		DISTRITO	
REGISTRO				
FECHA DE REGISTRO			HORA DE REGISTRO	



Foto

Declaro bajo juramento que la informacion presentada es completa y correcta

FIRMA



Anexo 4. Código utilizado en Python

#Modelado

```
import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split
from tensorflow.keras import Sequential
from tensorflow.keras.layers import Dense
import matplotlib.pyplot as plt
from sklearn.metrics import precision_score, recall_score, f1_score, confusion_matrix
import tensorflow.keras.optimizers
import tensorflow.keras.callbacks
from tensorflow import keras
import tensorflow as tf
import seaborn as sns

df = pd.read_csv("datos2.csv", sep=";")

X= df[['promedio','edad','sexo','estado civil','procedencia','debe','tipo universidad']]
y = df['desercion']

X_train, X_test, y_train, y_test = train_test_split( X, y, test_size=.2, random_state=42)

test = pd.Series(y_test)
train = pd.Series(y_train)
test.to_csv("test.csv")
train.to_csv("train.csv")

from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.fit_transform(X_test)
print("conjunto de entrenamiento = a { } registros por { }
variables".format(X_train.shape[0],X_train.shape[1]))
print("conjunto de validación = a { } registros por { }
variables".format(X_test.shape[0],X_test.shape[1]))

# se define la tasa de aprendizaje 0.01
optimizer = tensorflow.keras.optimizers.Adam(lr=0.01)

#semilla aleatoria para la inicializacion de los pesos
tf.random.set_seed(333330)
# iniciacion de la RNN
model1 = Sequential()
#capa oculta
model1.add(Dense(1, input_shape=(X_train.shape[1],), activation='sigmoid'))
#capa de salida
model1.add(Dense(1,activation="sigmoid"))
#se optimiza por retroprop automáticamente
```



```
model1.compile(optimizer=optimizer, loss='binary_crossentropy',metrics=['accuracy'])
#se realiza el entrenamiento con 500 iteraciones
history1 = model1.fit(X_train, y_train, epochs=500, validation_data=(X_test, y_test),

verbose=1, shuffle=False)

#gráfico del entrenamiento en cada iteración
plt.plot(history1.history['loss'])
plt.plot(history1.history['val_loss'])
plt.title('Entrenamiento del modelo')
plt.ylabel('Función de pérdida')
plt.xlabel('Iteración')
plt.legend(['Entrenamiento', 'Validación'], loc='upper right')
plt.show()

losstrain1 = pd.Series(history1.history['loss'])
losstest1 = pd.Series(history1.history['val_loss'])
losstest1.to_csv("losstest1.csv")
losstrain1.to_csv("losstrain1.csv")

##matriz de confucion train
p_train= model1.predict(X_train)
p_train[p_train<=0.5]=0
p_train[p_train>0.5]=1
matriz_train = confusion_matrix(y_train,p_train)
##matriz de confusion test
p_test = model1.predict(X_test)
p_test [p_test <= 0.5] = 0
p_test [p_test > 0.5]= 1
matriz_test = confusion_matrix(y_test,p_test)

## precision train y test
precision_train = precision_score(y_train,p_train)
precision_test = precision_score(y_test,p_test)
## recall train y test
recall_train = recall_score(y_train,p_train)
recall_test = recall_score(y_test,p_test)

## f1_score train y test

f1_train = f1_score(y_train,p_train)
f1_test = f1_score(y_test,p_test)
print("_____ Modelo 1
_____")
print("RESULTADOS DE ENTRENAMIENTO")
print("Matriz de confucion para entrenamiento")
print(matriz_train)
print("precision entrenamiento es: {}".format(precision_train))
print("recall para entrenamiento es: {}".format(recall_train))
print("F1_score para entrenamiento es: {}".format(f1_train))
```



```
print("_____")
print("Matriz de confucion para validación")
print(matriz_test)
print("precision validación es: {}".format(precision_test))
print("recall para validación es: {}".format(recall_test))
print("F1_score para validación es: {}".format(f1_test))

model1.summary()

model1.save('model1.h5')
model1.get_weights()

optimizer2 = tensorflow.keras.optimizers.Adam(lr=0.03)

#semilla aleatoria para la inicializacion de los pesos
tf.random.set_seed(2345)
# iniciacion de la RNN
model2 = Sequential()
#capa oculta
model2.add(Dense(1, input_shape=(X_train.shape[1],), activation='sigmoid'))
model2.add(Dense(3, activation='sigmoid'))
#capa de salida
model2.add(Dense(1,activation="sigmoid"))
#se optimiza por retroprop automáticamente
model2.compile(optimizer=optimizer2, loss='binary_crossentropy',metrics=['accuracy'])
#se realiza el entrenamiento con 35 iteraciones
history2 = model2.fit(X_train, y_train, epochs=75, validation_data=(X_test, y_test),

verbose=2, shuffle=False)

#gráfico del entrenamiento en cada iteración
plt.plot(history2.history['loss'])
plt.plot(history2.history['val_loss'])
plt.title('Entrenamiento del modelo')
plt.ylabel('Función de pérdida')
plt.xlabel('Iteración')
plt.legend(['Entrenamiento', 'Validación'], loc='upper right')
plt.axis([1, 75, 0, 1])
plt.show()

losstrain2 = pd.Series(history2.history['loss'])
losstest2 = pd.Series(history2.history['val_loss'])
losstest2.to_csv("losstest2.csv")
losstrain2.to_csv("losstrain2.csv")

losstrain2 = pd.Series(history2.history['loss'])
losstest2 = pd.Series(history2.history['val_loss'])
losstest2.to_csv("losstest2.csv")
losstrain2.to_csv("losstrain2.csv")
```



```
model2.summary()

model2.save('model2.h5')
model2.get_weights()

optimizer3 = tensorflow.keras.optimizers.Adam(lr=0.3)

# semilla aleatoria para la inicializacion de los pesos
tf.random.set_seed(222)
# iniciacion de la RNN
model3 = Sequential()
#capa oculta
model3.add(Dense(1, input_shape=(X_train.shape[1],), activation='sigmoid'))
model3.add(Dense(2, activation='sigmoid'))
#capa de salida
model3.add(Dense(1,activation="sigmoid"))
#se optimiza por retroprop automáticamente
model3.compile(optimizer=optimizer3, loss='binary_crossentropy',metrics=['accuracy'])
#se realiza el entrenamiento con 20 iteraciones
history3 = model3.fit(X_train, y_train, epochs=75, validation_data=(X_test, y_test),

verbose=1, shuffle=False)

#gráfico del entrenamiento en cada iteración
plt.plot(history3.history['loss'])
plt.plot(history3.history['val_loss'])
plt.title('Entrenamiento del modelo')
plt.ylabel('Función de pérdida')
plt.xlabel('Iteración')
plt.legend(['Entrenamiento', 'Validación'], loc='upper right')
plt.axis([1, 75, 0, 1])
plt.show()

losstrain3 = pd.Series(history3.history['loss'])
losstest3 = pd.Series(history3.history['val_loss'])
losstest3.to_csv("losstest3.csv")
losstrain3.to_csv("losstrain3.csv")

p_train3= model3.predict(X_train)
p_train3[p_train3<=0.5] = 0
p_train3[p_train3>0.5]= 1
matriz_train3 = confusion_matrix(y_train,p_train3)
matriz_train3

##matriz de confucion train
p_train3= model3.predict(X_train)
p_train3[p_train3<=0.5] = 0
p_train3[p_train3>0.5]= 1
matriz_train3 = confusion_matrix(y_train,p_train3)
##matriz de confucion test
```



```
p_test3 = model3.predict(X_test)
p_test3 [p_test3 <= 0.5] = 0
p_test3 [p_test3 > 0.5]= 1
matriz_test3 = confusion_matrix(y_test,p_test3)

## precision train y test
precision_train3 = precision_score(y_train,p_train3)
precision_test3 = precision_score(y_test,p_test3)
## recall train y test
recall_train3 = recall_score(y_train,p_train3)
recall_test3 = recall_score(y_test,p_test3)

## f1_score train y test

f1_train3 = f1_score(y_train,p_train3)
f1_test3= f1_score(y_test,p_test3)
print("_____")
print("RESULTADOS DE ENTRENAMIENTO modelo 3")
print("Matriz de confucion para entrenamiento")
print(matriz_train3)
print("precision entrenamiento es: {}".format(precision_train3))
print("recall para entrenamiento es: {}".format(recall_train3))
print("F1_score para entrenamiento es: {}".format(f1_train3))
print("_____")
print("Matriz de confucion para validación")
print(matriz_test3)
print("precision validación es: {}".format(precision_test3))
print("recall para validación es: {}".format(recall_test3))
print("F1_score para validación es: {}".format(f1_test3))

model3.summary()

model3.save('model3.h5')
model3.get_weights()

optimizer4 = tensorflow.keras.optimizers.Adam(lr=0.1)

#semilla aleatoria para la inicializacion de los pesos
tf.random.set_seed(333330)
# iniciacion de la RNN
model4 = Sequential()
#capa oculta
model4.add(Dense(1, input_shape=(X_train.shape[1],), activation='sigmoid'))
#capa de salida
model4.add(Dense(1,activation="sigmoid"))
#se optimiza por retroprop automáticamente
model4.compile(optimizer=optimizer4, loss='binary_crossentropy',metrics=['accuracy'])
#se realiza el entrenamiento con 500 iteraciones
```



```
history4 = model4.fit(X_train, y_train, epochs=25, validation_data=(X_test, y_test),  
  
shuffle=False)  
  
#semilla aleatoria para la inicializacion de los pesos  
tf.random.set_seed(333330)  
# iniciacion de la RNN  
model4 = Sequential()  
#capa oculta  
model4.add(Dense(1, input_shape=(X_train.shape[1],), activation='sigmoid'))  
#capa de salida  
model4.add(Dense(1,activation="sigmoid"))  
#se optimiza por retroprop automáticamente  
model4.compile(optimizer=optimizer4, loss='binary_crossentropy',metrics=['accuracy'])  
#se realiza el entrenamiento con 500 iteraciones  
history4 = model4.fit(X_train, y_train, epochs=25, validation_data=(X_test, y_test),  
  
shuffle=False)  
  
#gráfico del entrenamiento en cada iteración  
plt.plot(history4.history['loss'])  
plt.plot(history4.history['val_loss'])  
plt.title('Entrenamiento del modelo')  
plt.ylabel('Función de pérdida')  
plt.xlabel('Iteración')  
plt.legend(['Entrenamiento', 'Validación'], loc='upper right')  
plt.axis([1, 25, 0,1])  
plt.show()  
  
losstrain4 = pd.Series(history4.history['loss'])  
losstest4 = pd.Series(history4.history['val_loss'])  
losstest4.to_csv("losstest4.csv")  
losstrain4.to_csv("losstrain4.csv")  
  
##matriz de confucion train  
p_train4= model4.predict(X_train)  
p_train4[p_train4<=0.5]=0  
p_train4[p_train4>0.5]=1  
matriz_train4 = confusion_matrix(y_train,p_train4)  
##matriz de confucion test  
p_test4 = model4.predict(X_test)  
p_test4 [p_test4 <= 0.5] = 0  
p_test4 [p_test4 > 0.5]= 1  
matriz_test4 = confusion_matrix(y_test,p_test4)  
  
## precision train y test  
precision_train4 = precision_score(y_train,p_train4)  
precision_test4 = precision_score(y_test,p_test4)  
## recall train y test  
recall_train4 = recall_score(y_train,p_train4)
```



```
recall_test4 = recall_score(y_test,p_test4)

## f1_score train y test

f1_train4 = f1_score(y_train,p_train4)
f1_test4 = f1_score(y_test,p_test4)
print("_____ modelo 4
_____")
print("RESULTADOS DE ENTRENAMIENTO")
print("Matriz de confucion para entrenamiento")
print(matriz_train4)
print("precision entrenamiento es: {}".format(precision_train4))
print("recall para entrenamiento es: {}".format(recall_train4))
print("F1_score para entrenamiento es: {}".format(f1_train4))
print("_____
_____")
print("Matriz de confucion para validación")
print(matriz_test4)
print("precision validación es: {}".format(precision_test4))
print("recall para validación es: {}".format(recall_test4))
print("F1_score para validación es: {}".format(f1_test4))

model4.summary()

model4.save('model4.h5')
model4.get_weights()
```