



FACULTAD DE INGENIERÍA Y ARQUITECTURA
ESCUELA PROFESIONAL DE INGENIERÍA DE SISTEMAS E INFORMÁTICA

TESIS

**REDES NEURONALES ARTIFICIALES Y LA PREDICCIÓN DE LA
VIABILIDAD LEGISLATIVA EN LAS PROPOSICIONES
PARLAMENTARIAS, PERÚ, AÑO 2020**

**PARA OPTAR:
EL TÍTULO PROFESIONAL DE
INGENIERO DE SISTEMAS E INFORMÁTICA**

**PRESENTADO POR:
BACH. DIOGENES LUIS NAZARIO HUANACO**

**ASESOR:
MG. WILINTON CHILIMAN ORTIZ COTRINA**

**LIMA – PERÚ
2021**

DEDICATORIA

A mis padres, a ellos dedico mis logros. Agradezco la labor y entrega de mis padres, mi formación profesional y por haber siempre confiado en mí.

AGRADECIMIENTO

A la Mg. Martha Medalit Campos Villegas, al Ing. Enrique Carlos Garrido-Lecca Risco, al Ing. Henry Hermes Veneros Manrique y al Mg. Jaime Moshé Farfán Madariaga por sus valiosos conocimientos compartidos.

RECONOCIMIENTO

Al Ph. D. Luis Alex Alzamora de los Godos Urcia por compartir sus conocimientos y aportes en el desarrollo de la investigación científica.

ÍNDICE

CARÁTULA	i
DEDICATORIA	ii
AGRADECIMIENTO	iii
RECONOCIMIENTO	iv
ÍNDICE	v
ÍNDICE DE TABLAS	vii
ÍNDICE DE FIGURAS	ix
RESUMEN	xii
ABSTRACT	xiii
INTRODUCCIÓN	1
Descripción de la Situación Problemática.....	3
Formulación del Problema.....	15
<i>Problema Principal</i>	15
<i>Problemas Derivados</i>	15
Objetivos de la Investigación	16
<i>Objetivo General</i>	16
<i>Objetivos Específicos</i>	16
Justificación de la Investigación.....	17
<i>Importancia de la Investigación</i>	17
<i>Viabilidad de la Investigación</i>	19
Limitaciones del Estudio	21
CAPÍTULO I: MARCO TEÓRICO CONCEPTUAL	23
1.1. Antecedentes de la Investigación.....	23
1.2. Bases Teóricas	28
1.2.1. <i>Redes Neuronales Artificiales</i>	28
1.2.2. <i>Viabilidad legislativa</i>	47

1.3.	Definición de Términos Básicos	60
CAPÍTULO II: PROPUESTA DE LA ELABORACIÓN DEL MODELO PARA DEMOSTRAR SU CAPACIDAD PREDICTIVA EN LA VIABILIDAD LEGISLATIVA.....		66
2.1.	Fase 2 - Comprensión de los Datos.....	66
2.2.	Fase 3 - Preparación de Datos	70
2.3.	Fase 4 - Modelamiento	71
2.4.	Fase 5 - Evaluación	77
CAPÍTULO III: METODOLOGÍA.....		79
3.1.	Diseño Metodológico.....	79
3.1.1.	<i>Tipo y Nivel de la Investigación.....</i>	<i>79</i>
3.1.2.	<i>Métodos y Diseño de la Investigación.....</i>	<i>81</i>
3.2.	Técnicas de Recolección de Datos	84
3.3.	Técnicas Estadísticas para el Procesamiento de la Información	85
3.4.	Diseño Muestral	85
3.5.	Aspectos Éticos.....	86
CAPÍTULO IV: RESULTADOS.....		88
4.1.	Presentación de los Resultados Descriptivos	88
4.2.	Presentación de los Resultados de la Predicción	103
CAPÍTULO V: DISCUSIÓN.....		113
CONCLUSIONES		117
RECOMENDACIONES		119
FUENTES DE INFORMACIÓN		121
ANEXOS		128

ÍNDICE DE TABLAS

Tabla 1 Cálculo estimado de costos (meses) que incurre una proposición en tiempos promedios (meses) en el flujo legislativo	8
Tabla 2 Presupuesto de la investigación	19
Tabla 3 Técnicas de aprendizaje.....	37
Tabla 4 Variables de la Calidad del ACB Convencional (IC-ACB-C).....	48
Tabla 5 Variables de la calidad del ACB Legislativo (IC-ACB-L).....	49
Tabla 6 Variables de la Viabilidad Legislativa sin ACB.	50
Tabla 7 Variables Relacionadas a los Proponentes y las Iniciativas.	50
Tabla 8 Comparativa de precisión entre las técnicas más usadas en los estudios de los antecedentes nacionales e internacionales	71
Tabla 9 Precisión de la red neuronal de función de base radial mediante SPSS Statistic.....	76
Tabla 10 Precisión de la red neuronal de perceptrón multicapa mediante SPSS Statistic.....	76
Tabla 11 Comparación de precisión del modelo ANN de perceptrón multicapa con otras técnicas algorítmicas.....	77
Tabla 12 Niveles, objetivos y tipos en la investigación holística.....	80
Tabla 13 Tipo de proposición legislativa	89
Tabla 14 Número de comisiones que fue enviada la proposición	91
Tabla 15 Comisión de carácter económico de la proposición	92
Tabla 16 La proposición tiene incidencia económica directa	93
Tabla 17 Número de congresistas que firman la propuesta	95

Tabla 18 El proponente forma parte de la mayoría parlamentaria	97
Tabla 19 El proponente tiene experiencia parlamentaria	99
Tabla 20 Sexo del proponente	100
Tabla 21 Número de proposiciones presentadas por el congresista.....	102
Tabla 22 Resumen de procesamiento de casos	103
Tabla 23 Información de red.....	104
Tabla 24 Estimaciones de parámetro.....	108
Tabla 25 Importancia de las variables independientes.....	109
Tabla 26 Área bajo la curva.....	110
Tabla 27 Resumen del modelo.....	111
Tabla 28 Clasificación	112
Tabla 29 Matriz de consistencia.....	129

ÍNDICE DE FIGURAS

Figura 1 Página web de seguimiento de proyectos ley con pronóstico de viabilidad en el Congreso de los Estados Unidos	5
Figura 2 Proyectos de ley presentados en el Parlamento Latinoamericano entre el año 2008 y 2017.	6
Figura 3 Escala de tiempo del proyecto de ley 00051 en el flujo legislativo	7
Figura 4 Propositiones viables y no viables entre el periodo parlamentario 1995 y 2016 presentadas por el poder legislativo.....	9
Figura 5 Propositiones viables y no viables entre el periodo parlamentario 1995 y 2016 presentadas por el poder ejecutivo	10
Figura 6 Meses promedio que demora una proposición en el proceso legislativo por periodo parlamentario	10
Figura 7 Cálculo estimado de costos(sueldos) que incurre una proposición en tiempos promedios(meses) en el proceso legislativo	11
Figura 8 Diagrama del árbol de problemas	12
Figura 9 Izquierda: neurona en el cerebro. Derecha: red neuronal artificial con tres neuronas en la capa de entrada y una neurona en la capa de salida.	29
Figura 10 Infografía de la historia de las redes neuronales artificiales.....	32
Figura 11 Red feedforward con una sola capa de neuronas.....	33
Figura 12 Red feedforward completamente conectada con una capa oculta y una capa de salida.	34
Figura 13 Red recurrente con neuronas ocultas.	35
Figura 14 Diagrama de bloques de aprendizaje supervisado.....	36
Figura 15 Diagrama de bloques de aprendizaje no supervisado.....	36

Figura 16 Diagrama de bloques de aprendizaje por refuerzo; el sistema de aprendizaje y el entorno están ambos dentro del circuito de retroalimentación.	37
Figura 17 Machine learning un nuevo paradigma de programación.....	41
Figura 18 Fases de la metodología CRISP-DM	42
Figura 19 Fases y actividades de la metodología CRISP-DM.....	45
Figura 20 Clasificación del ACB.....	55
Figura 21 El flujo legislativo.....	58
Figura 22 Las proposiciones parlamentarias.....	59
Figura 23 Uso de la herramienta Octoparse versión 8	67
Figura 24 Exportación de datos al archivo Excel.....	67
Figura 25 Usando la herramienta SPSS Modeler versión 18 para exportar el archivo Excel a formato *.sav.....	68
Figura 26 Verificación de tipos de variables y asignación de etiquetas a sus valores en SPSS Modeler	68
Figura 27 Exploración de datos mediante SPSS Modeler.....	69
Figura 28 Generación de resultados descriptivos con SPSS Modeler	70
Figura 30 Proceso de ejecución de la técnica predictiva de redes neuronales artificiales (ANNs) mediante SPSS Modeler	74
Figura 31 Generación de la red neuronal mediante SPSS Modeler.....	74
Figura 29 Comparación de precisión de otras técnicas algorítmicas mediante SPSS Modeler.....	75
Figura 32 Esquema para una investigación predictiva	81
Figura 33 La espiral holística y los métodos en investigación	82

Figura 34 Esquema para seleccionar los diseños de investigación	83
Figura 35 Tipo de proposición legislativa	90
Figura 36 Número de comisiones que fue enviada la proposición	91
Figura 37 Comisión de carácter económico de la proposición	93
Figura 38 La proposición tiene incidencia económica directa	94
Figura 39 Número de congresistas que firman la propuesta.....	96
Figura 40 El proponente forma parte de la mayoría parlamentaria	97
Figura 41 El proponente tiene experiencia parlamentaria	99
Figura 42 Sexo del proponente	101
Figura 43 Número de proposiciones presentadas por el congresista.....	102
Figura 44 Diagrama de la red neuronal artificial.....	106
Figura 45 Curva ROC (curva rendimiento diagnóstico).....	110

RESUMEN

La presente investigación trata acerca de las redes neuronales artificiales y la viabilidad legislativa. El objetivo del estudio es predecir la viabilidad legislativa mediante el uso de las redes neuronales artificiales en las proposiciones parlamentarias (proyectos de ley y resoluciones legislativas) del Congreso de la República. El método de la investigación es inductivo y el diseño es transeccional retrospectivo documental de rasgo. La población de estudio fue de 377 proposiciones registradas en el periodo anual de sesiones 2019, correspondiente al periodo parlamentario 2016. En base a los resultados, el modelo obtuvo el 96.1% de precisión de aciertos. Finalmente, en la investigación se concluyó que, aplicando las técnicas algorítmicas de las redes neuronales artificiales, se obtuvo un modelo con un alto porcentaje de predicción.

Palabras claves: Congreso, legislativo, predicción, proposición, proyecto de ley, redes neuronales, viabilidad.

ABSTRACT

This research is about artificial neural networks and legislative viability. The objective of the study is to predict legislative viability through the use of artificial neural networks in the parliamentary proposal (bills and parliamentary resolutions) of the Congress of the Republic. The research method is inductive, and the design is a feature documentary retrospective transectional. The study population was 377 proposals registered in the 2019 annual period of sessions, corresponding to the 2016 parliamentary period. Based on the results, the model obtained 96.1% accuracy of hits. Finally, the research concluded that, applying algorithmic techniques of artificial neural networks, a model with a high percentage of prediction was obtained.

Keywords: Congress, legislative, prediction, proposal, bill, neural networks, viability.

INTRODUCCIÓN

La inteligencia artificial puede cambiar el modo de actuar de los gobiernos, nos encontramos en la era de la industria 4.0, frente a esto, es una obligación tomar acciones para dejar lo tradicional y buscar mejores alternativas en la prestación de servicios más eficientes y de mayor calidad a los ciudadanos. El presente estudio, en su afán de aportar conocimientos, se centra en demostrar la capacidad del poder predictivo de la inteligencia artificial mediante el uso de las redes neuronales artificiales para predecir la viabilidad legislativa en las proposiciones parlamentarias. La aplicación de la inteligencia artificial puede transformarse en un soporte clave en la elaboración de las leyes, la evaluación de su viabilidad y la gestión del uso óptimo de los recursos en el quehacer legislativo. Para la elaboración de la presente investigación se ha consultado diversas fuentes que enriquecen ciertos conceptos y técnicas de la manera más dicáctica posible. A continuación, se detallan los capítulos que la componen.

En la parte introductoria, se enuncia el planteamiento del problema situacional a nivel mundial, regional y nacional, describiendo cómo las técnicas predictivas se han introducido inicialmente en los rubros del derecho y la legislación, se mencionan también las causas y consecuencias de no contar con proposiciones viables, asimismo, el beneficio de aplicar las redes neuronales en la predicción de la viabilidad legislativa.

En el CAPÍTULO I: MARCO TEÓRICO CONCEPTUAL, se desarrolla el marco teórico y los antecedentes de la investigación, presentando las últimas investigaciones realizadas con respecto a la viabilidad legislativa y similares casos en la que se aplicaron modelos algorítmicos de predicción. Se ha hecho una

recopilación de trabajos muy interesantes en la que se obtuvieron una alta tasa de aciertos predictivos.

En el CAPÍTULO II: PROPUESTA DE LA ELABORACIÓN DEL MODELO PARA DEMOSTRAR SU CAPACIDAD PREDICTIVA EN LA VIABILIDAD LEGISLATIVA, se desarrolla la propuesta de la investigación, describiendo el desarrollo de las fases de la metodología seleccionada.

En el CAPÍTULO III: METODOLOGÍA, se describe la metodología de la investigación, tipo, nivel y su diseño para la recolección de datos. Se detalla la población y la muestra que se usó como dato para el entrenamiento y la evaluación del modelo.

En el CAPÍTULO IV: RESULTADOS, se describen los resultados obtenidos, el análisis descriptivo de los atributos de la variable endógena y los resultados del estudio predictivo mediante tablas y figuras.

En el CAPÍTULO V: DISCUSIÓN, referente a la discusión, se realizó el cruce de información entre las últimas investigaciones realizadas (presentadas en el CAPÍTULO I: MARCO TEÓRICO CONCEPTUAL) y la presente investigación, resaltando las ventajas de cada una mediante la comparación de sus resultados.

Finalmente, se presenta las conclusiones de la investigación, confirmando lo más relevante de los resultados obtenidos, y las recomendaciones, en la que se detalla de manera reflexiva, los puntos destacables de la investigación brindando aportes y sugerencias con la finalidad de mejoras futuras.

Descripción de la Situación Problemática

A nivel mundial, el uso de la inteligencia artificial (IA) se está convirtiendo en un soporte de alta prioridad al momento de tomar decisiones, las grandes organizaciones lo están incluyendo dentro de sus actividades principales para proyectarse, optimizar y alcanzar el máximo beneficio posible. Gracias a la evolución de los dispositivos, la infraestructura de las redes móviles, las técnicas algorítmicas y el uso de enormes cantidades de datos, anticiparse a la conducta del cliente o al comportamiento del mercado, ahora es posible, por ejemplo, las empresas como Amazon, AliExpress, YouTube o Netflix hacen uso de la IA para predecir las preferencias de sus clientes, ofreciéndoles un servicio personalizado y a la medida de sus gustos, otro ejemplo también puede ser Cinelytic, que es una plataforma de la industria cinematográfica que provee datos, análisis e inteligencia predictiva, integrada a un sistema de gestión de proyectos, en la que es capaz de predecir el éxito de una película. Kaur & Bozié (2019) estiman que, en los últimos años, el modelado predictivo ha traído cambios del tipo preventivo en el funcionamiento de varias industrias, los avances en el estudio del aprendizaje profundo (DL) y procesamiento del lenguaje natural (NLP) se han aplicado progresivamente dando resultados altamente prometedores, el ámbito legal no es ajeno a estos avances. Una de las herramientas que están dando sus frutos en las ciencias del derecho es la jurimetría, (Loevinger, 1949) acuña el término jurimetría, basado en los conceptos de jurisprudencia y medición, la multinacional (Wolters Kluwer, 2020) lo define como una herramienta de analítica jurisprudencial que permite definir estrategias procesales idóneas para el éxito de un caso.

En los países europeos y asiáticos, de acuerdo con la publicación de (Iglesias Rodríguez, García Zaballos, Puig Gabarró, & Benzaqué, 2020), han tomado un gran interés en introducir planes estratégicos a nivel nacional desde el año 2014, con el objetivo de fortalecer la investigación e incrementar los estándares de IA en el sector público y privado, así como, de la creación de marcos jurídicos que regulen los efectos de la IA en la sociedad, también mencionan en su informe que, en la actualidad, los países asiáticos que lideran la producción de IA son China e Israel, bajo un ecosistema basado en políticas gubernamentales, talento técnico y emprendedores en serie. En el caso de Europa, mencionan que, Francia ha financiado una plataforma IA para conglomerar conocimientos y recursos disponibles para expandir su desarrollo, así como, de un proyecto multinacional europeo de investigación para la retención de talentos. Francia actualmente posee el laboratorio de startups más grande del mundo que fue inaugurado en junio de 2017.

Sonderegger (2018) postuló en la página web de About The Law, que, en los Estados Unidos, empresas dedicadas al desarrollo de programas de IA, han desarrollado modelos para predecir la viabilidad de los proyectos de ley en el Congreso norteamericano, el uso de la IA ha permitido estimar una tasa de precisión del 99% para los proyectos de ley que pasaron a la Cámara de Diputados o al Senado y un 98% de una tasa de precisión para la promulgación de una ley. En la Figura 1, se puede observar la página web actual del Congreso de los Estados Unidos que muestra el pronóstico de viabilidad por cada proyecto de ley.

Figura 1

Página web de seguimiento de proyectos ley con pronóstico de viabilidad en el Congreso de los Estados Unidos

The screenshot shows the GovTrack website interface. At the top, there is a navigation bar with links for Home, Bills & Votes, Legislators, Using GovTrack, and About Us. Below this is a search bar with a magnifying glass icon and a language selection dropdown labeled 'Seleccionar idioma'. The main heading is 'Advanced Search for Legislation' with a sub-heading 'Congress / Bills'. A descriptive paragraph explains the search capabilities. Below the text is a search filter panel on the left with fields for 'search title & full text', 'congress' (set to '117th Con'), 'sponsor' (set to 'All'), and 'current status' (with checkboxes for 'All', 'Introduced (5,887 bills)', 'Ordered Reported (241 bills)', 'Passed House (Senate next) (66 bills)', and 'Passed Senate (House next) (25 bills)'). To the right of the filters, it says '93 bills matched your search.' and 'sort by Secret Sauce'. Two search results are displayed as cards. The first card is for 'H.R. 1899: Ensuring Compliance Against Drug Diversion Act of 2021' by Rep. Morgan Griffith [R-VA9], showing it was introduced on Mar 16, 2021, passed the House, and has a 32% prognosis. The second card is for 'H.R. 1629: Fairness in Orphan Drug Exclusivity Act' by Rep. Madeleine Dean [D-PA4], showing it was introduced on Mar 8, 2021, passed the House, and has a 4% prognosis.

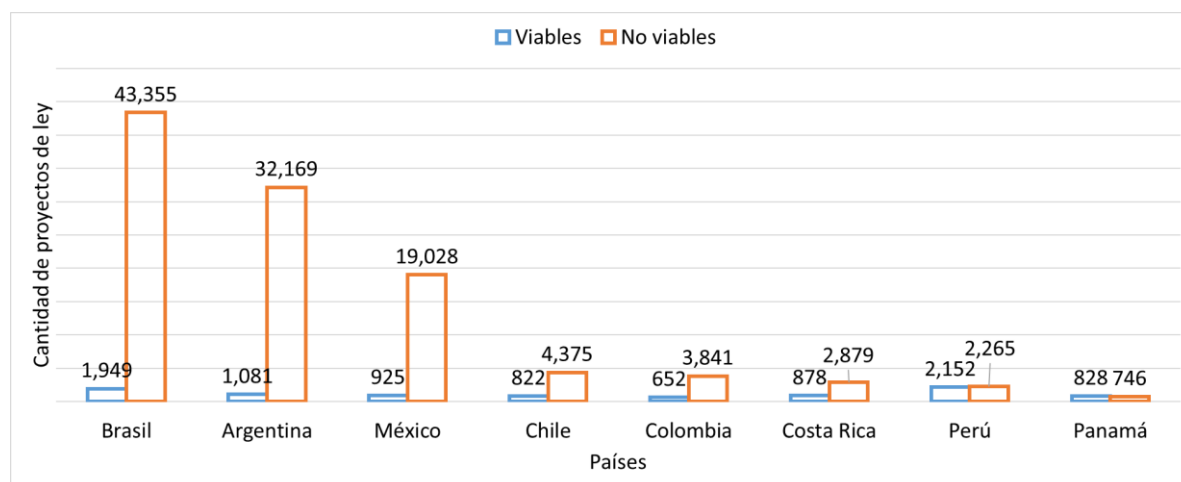
Fuente: (United States Congress, 2021)

En Latinoamérica, la (Fundación Directorio Legislativo, 2018) presentó la situación de los proyectos de ley viables y no viables de los países latinoamericanos entre los años 2008 y 2017, en el caso de Panamá, de los 1,574 proyectos de ley presentados y se promulgaron 828, representando el 52.60%. En Costa Rica se presentaron 3,757 y se promulgaron 878 proyectos de ley, representando el 23.37%. En Chile se presentaron 5,197 y promulgaron 822 proyectos de ley, representando el 15.82%. En Colombia se presentaron 4,493 y se promulgaron 652 proyectos de ley, representando el 14.51%. En México se

presentaron 19,953 y se promulgaron 925 proyectos de ley, representando el 4.64%. En Brasil se presentaron 45,304 y se promulgaron 1,949 proyectos de ley, representando el 4.30%. En Argentina se presentaron 33,250 y se promulgaron 1,081 proyectos de ley, representando el 3.25%. Como resultado comparativo, el país con la mayor tasa de éxito legislativo fue Panamá, con más del 50% de las iniciativas que se convirtieron en leyes. De acuerdo con los datos estadísticos presentados, los parlamentos latinoamericanos tienen un alto porcentaje de proyectos de ley no viables, entre ellos Brasil, Argentina y México se encuentran en las primeras posiciones. Con respecto a la manera de reducir este problema, (López, 2018) cita en la revista Tiempo Judicial que, Brasil concentra un mayor cúmulo de startups que brindan sus servicios de desarrollo de IA a los estados o a las empresas del rubro legal, en Argentina, a partir del año 2018, la Corte Suprema de Justicia procedió con el inicio de la elaboración de una serie de propuestas para modernizar y agilizar la justicia. En la Figura 2, se puede apreciar la cantidad de proyectos de ley viables y no viables en Latinoamérica.

Figura 2

Proyectos de ley presentados en el Parlamento Latinoamericano entre el año 2008 y 2017.

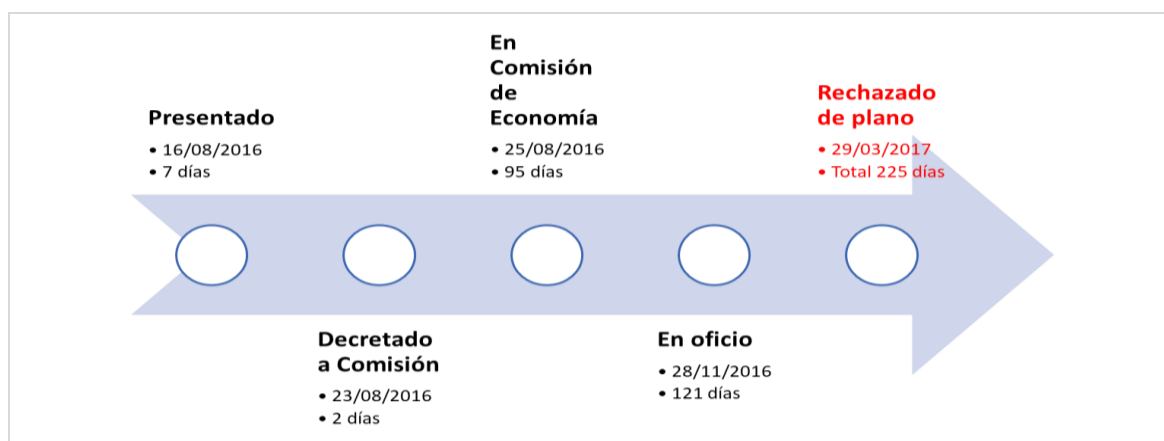


Fuente: elaboración propia con datos de (Fundación Directorio Legislativo, 2018)

En el Perú, similar a los parlamentos latinoamericanos y del mundo, cuando las proposiciones no viables se incrementan, también se incrementan los costos en recursos humanos, en tiempo dedicado a su elaboración y su seguimiento en el flujo legislativo y los costos en que afectan al presupuesto. Por ejemplo, para una demostración didáctica se tomó el proyecto de ley 00051, presentado por (Glave Remy, y otros, 2016). La presentación del proyecto de ley fue el 16 de agosto de 2016, quedando siete días en ese estado; después el día 23 del mismo mes permaneció por dos días en el estado de decretado a comisión; el 25 de agosto ingresó a la Comisión de Economía quedando en estudio por 95 días; el 28 de noviembre del mismo año pasa al estado de oficio y permanece por 121 días; finalmente, el 29 de marzo del 2017 es rechazado; el tiempo total que duró en el flujo legislativo fue de 225 días, aproximadamente 7.5 meses (ver Figura 3), de acuerdo al tiempo promedio de elaboración de las proposiciones no viables (ver Tabla 1), en la que se calculó solamente los costos de pagos del personal involucrado (simulando montos iguales en los diferentes periodos parlamentarios), se obtuvo como resultado un costo aproximado de S/ 336,030.00.

Figura 3

Escala de tiempo del proyecto de ley 00051 en el flujo legislativo



Fuente: elaboración propia con datos del (Congreso de la República, 2020)

Tabla 1

Cálculo estimado de costos (meses) que incurre una proposición en tiempos promedios (meses) en el flujo legislativo

		Periodo Parlamentario					
		1995	2000	2001	2006	2011	2016
Costos en pago al personal por mes	Congresista	S/ 15,326.00	S/ 15,326.00	S/ 15,326.00	S/ 15,326.00	S/ 15,326.00	S/ 15,326.00
	Asesor I	S/ 8,899.00	S/ 8,899.00	S/ 8,899.00	S/ 8,899.00	S/ 8,899.00	S/ 8,899.00
	Asesor II	S/ 7,189.00	S/ 7,189.00	S/ 7,189.00	S/ 7,189.00	S/ 7,189.00	S/ 7,189.00
	Técnico	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00
	Técnico	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00	S/ 4,697.00
	Asistente	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00
	Auxiliar	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00	S/ 1,998.00
	Total Costos	S/ 44,804.00	S/ 44,804.00	S/ 44,804.00	S/ 44,804.00	S/ 44,804.00	S/ 44,804.00
Tiempo promedio de duración de un PL en el flujo legislativo (meses)	Viable	7.45	2.97	12.55	13.86	13.58	11.66
	No Viable	10.17	4.58	20.08	28.87	22.38	13.46
Total de costos por duración	Viable	S/ 334,009.26	S/ 133,248.11	S/ 562,356.78	S/ 620,924.91	S/ 608,476.02	S/ 522,634.38
	No Viable	S/ 455,719.92	S/ 205,030.22	S/ 899,512.68	S/ 1,293,441.70	S/ 1,002,526.61	S/ 603,141.96

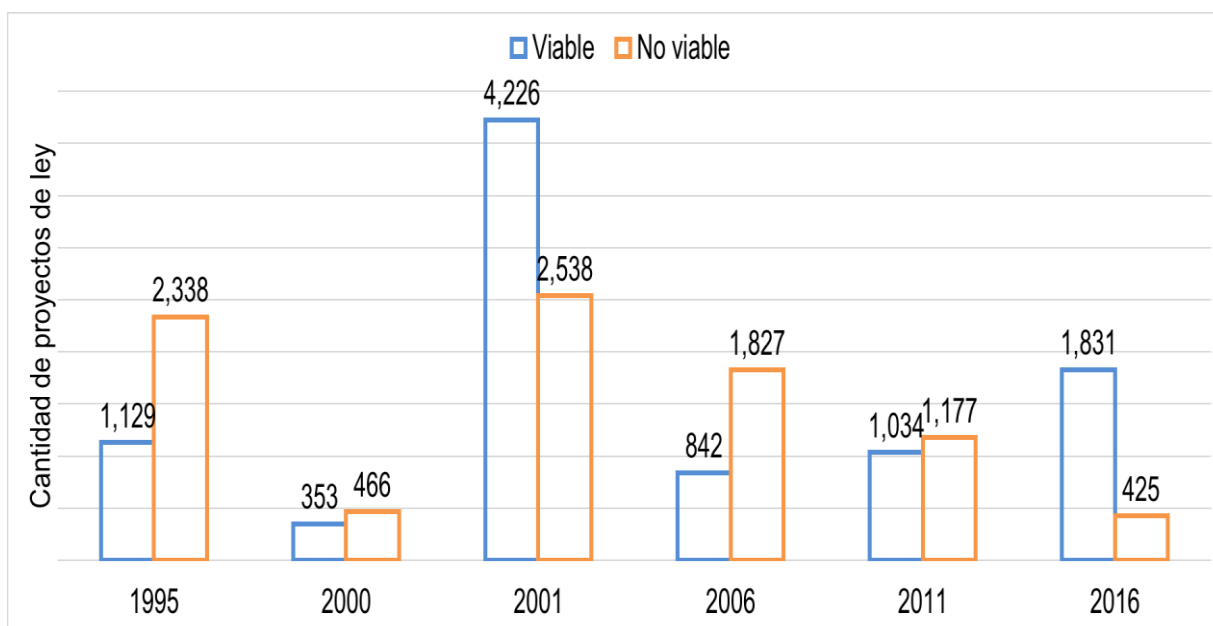
Fuente: elaboración propia con datos del (Congreso de la República, 2012) y (Congreso de la República, 2021)

En Figura 4, se muestra la cantidad de proposiciones elaboradas durante entre el periodo parlamentario 1995 y 2016, se puede observar que el periodo 2001 se incrementó las proposiciones viables, esta característica particular pudo estar influenciada por la situación coyuntural de ese entonces. En la Figura 5, se muestra las proposiciones presentadas por el poder ejecutivo, haciendo una comparación con la Figura 4, se puede observar la alta eficacia del poder ejecutivo en la producción de leyes. En la Figura 6, muestra la comparación de los meses promedio por periodo parlamentario que se demoró una proposición para

ser viable o no. En la Figura 7, se muestra los promedios en costos por periodo parlamentario.

Figura 4

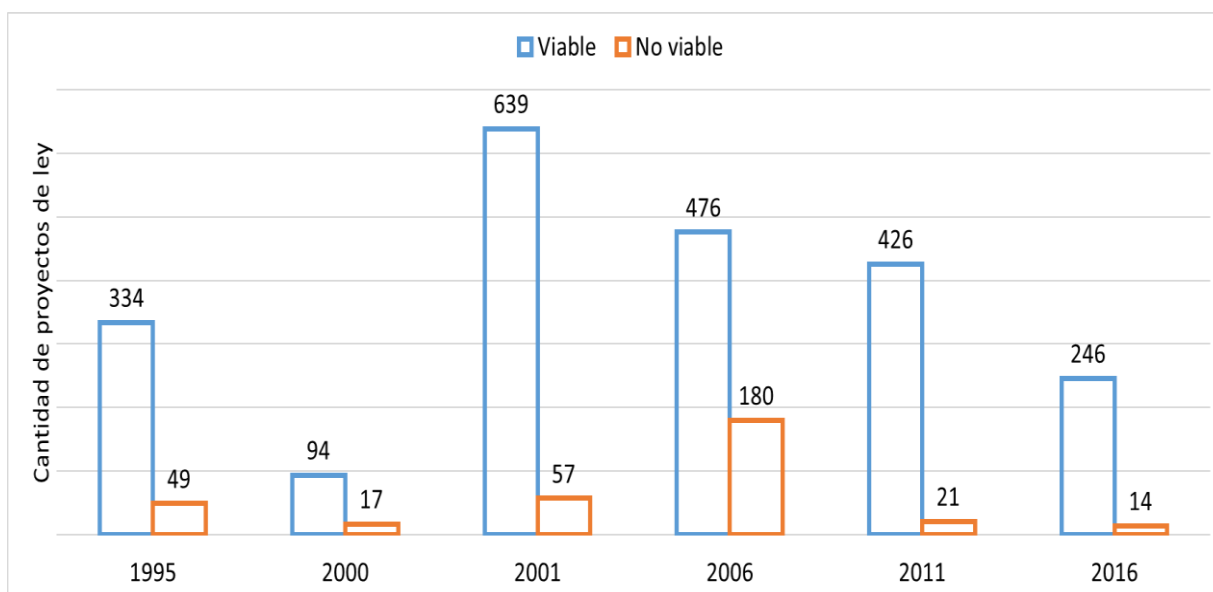
Proposiciones viables y no viables entre el periodo parlamentario 1995 y 2016 presentadas por el poder legislativo.



Fuente: elaboración propia con datos del (Congreso de la República, 2020)

Figura 5

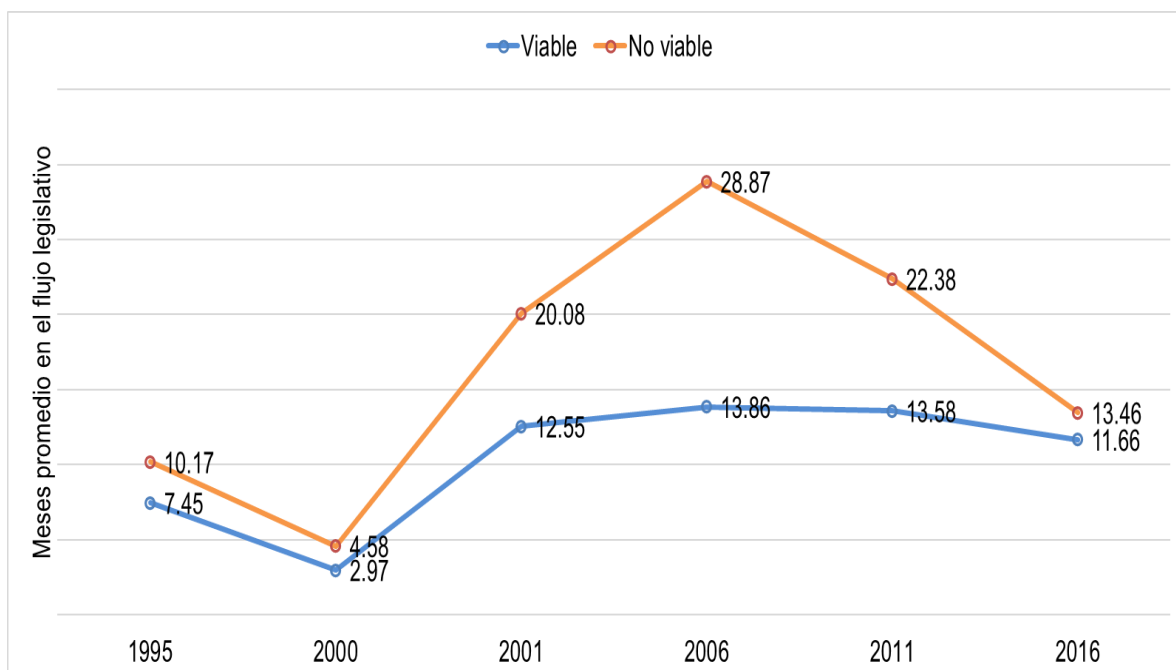
Proposiciones viables y no viables entre el periodo parlamentario 1995 y 2016 presentadas por el poder ejecutivo



Fuente: elaboración propia con datos del (Congreso de la República, 2020)

Figura 6

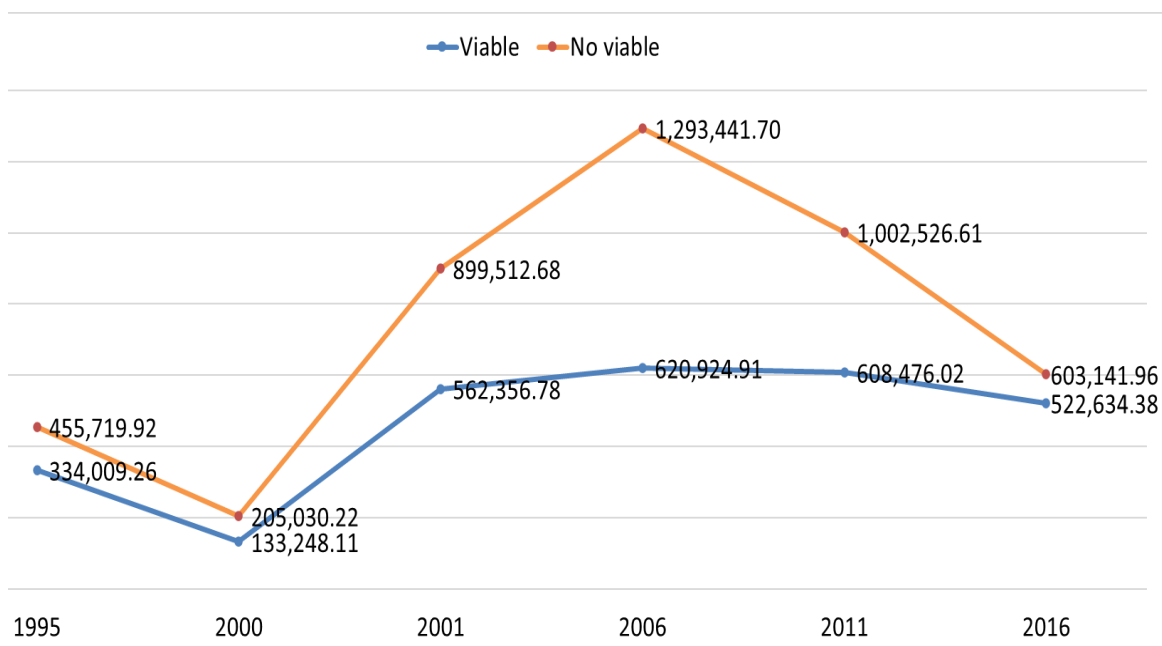
Meses promedio que demora una proposición en el proceso legislativo por periodo parlamentario



Fuente: elaboración propia con datos del (Congreso de la República, 2020)

Figura 7

Cálculo estimado de costos(sueldos) que incurre una proposición en tiempos promedios(meses) en el proceso legislativo



Fuente: elaboración propia con datos del (Congreso de la República, 2012) y (Congreso de la República, 2021)

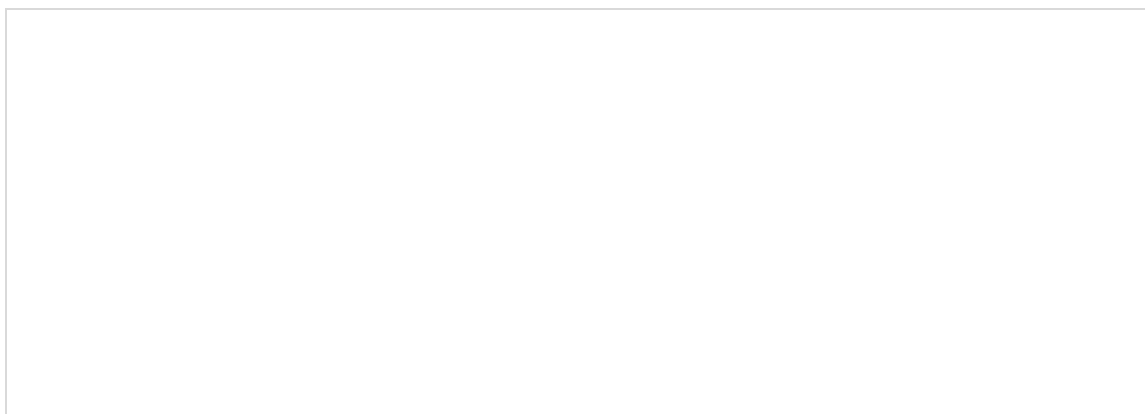
En el parlamento peruano se consideran ciertos factores para que una proposición parlamentaria sea considerada viable y lo señala el artículo 75 del Reglamento del Congreso de la República (Congreso de la República, 2021c) y en la publicación de Proyectos de ley archivados (Congreso de la República, 2020b) relacionados al motivo de su fundamento, su efecto, la inclusión del Análisis Costo Beneficio (ACB) y de un comentario acerca de su incidencia ambiental. Cabe considerar también otros factores como los establecidos en el Capítulo IX del Manual de la Técnica Legislativa acerca de la necesidad y viabilidad de la ley (Congreso de la República, 2021b) que incluye un bloque de once pasos a seguir. Asimismo, también es importante mencionar los factores referentes a las 35 variables definidas en el estudio de (Arias-Schreiber Barba, Peña Jumba, & Valdivieso Serrano, 2019) y, como factores externos, se pueden

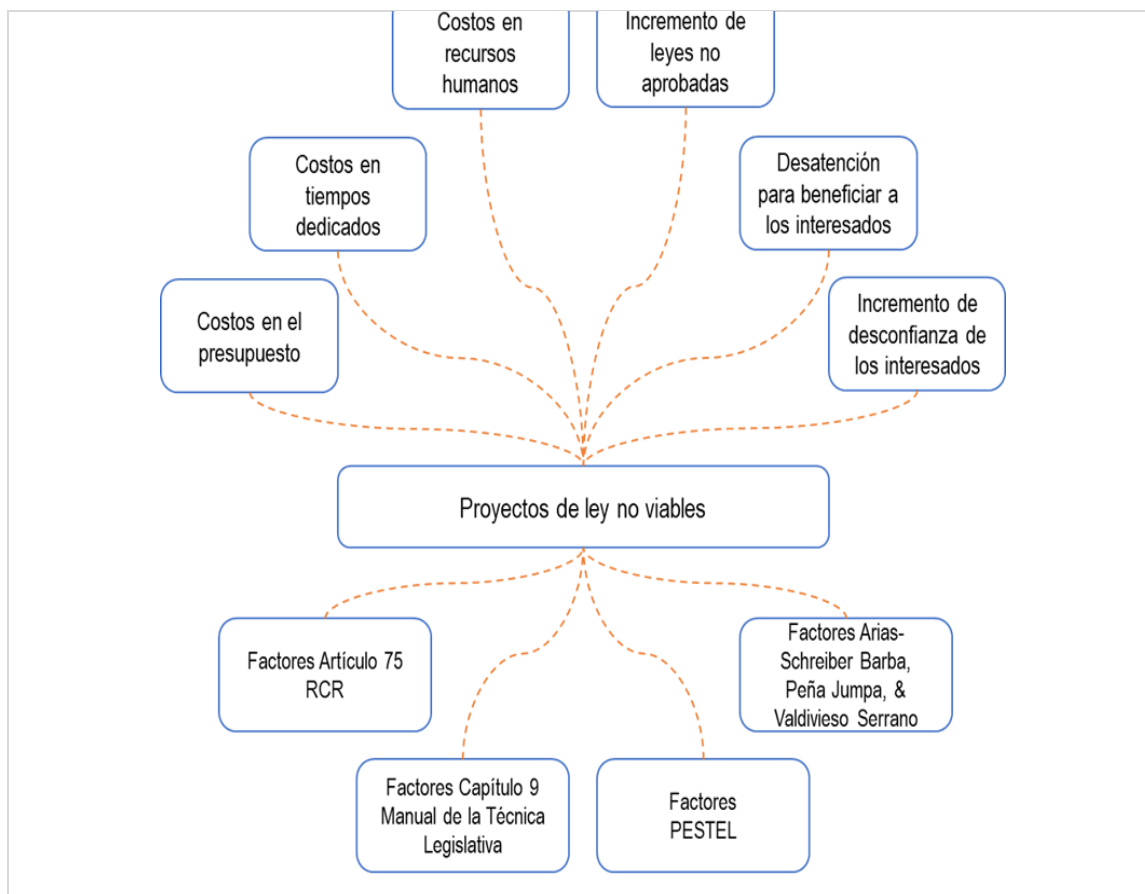
considerar los factores PESTEL (políticos, económicos, sociales, tecnológicos, ecológicos y legales). De acuerdo con los factores descritos anteriormente, se ha elaborado un diagrama de árbol de problemas que se muestra en la Figura 8.

En el Perú, en el año 2019, se llevó a cabo el primer evento de IA denominado IAcción Red Gealc, organizado por la Secretaría de Gobierno Digital para generar proyectos que mejoren los servicios del Estado soportados por la IA, la propuesta ganadora combinó la IA e imágenes médicas para el prediagnóstico de enfermedades respiratorias. En la conferencia online, “La Inteligencia Artificial como Herramienta de Desarrollo en Perú”, organizada por UNESCO Perú, realizada en septiembre del 2020, (Beltrán Castañón, 2020) expuso las limitaciones para el desarrollo de la IA, entre ellas mencionó la falta de talento especializado en IA, falta de datos para entrenar IA, desconocimiento de la IA por parte de la sociedad y la dificultad para encontrar fondos para proyectos de IA. En resumen, se puede decir que el Perú aún no inicia un estudio científico de la IA, en la que el estado participe activamente con las comunidades académicas y la empresa privada. Esto representa una brecha que hace lento el desarrollo de la IA en el Perú.

Figura 8

Diagrama del árbol de problemas





Fuente: elaboración propia

La aplicación de la IA a través del uso de las redes neuronales artificiales y otras técnicas algorítmicas del aprendizaje automático y del aprendizaje profundo, han demostrado resultados aceptables como herramientas de predicción en las diversas actividades a nivel mundial en el entorno privado, y últimamente al entorno público. El uso de la IA en los parlamentos del mundo, según los estudios realizados en otros países, está marcando un cambio de paradigma, por ejemplo, es posible predecir los votos de aprobación de las leyes (mediante atributos de los proyectos de ley y de los congresistas), predecir la probabilidad de corrupción de un congresista basado en los proyectos de ley aprobados, predecir el tiempo de evaluación del proceso legislativo, etc. Mediante el uso de las redes neuronales, predecir un futuro altamente probable representa una ventaja competitiva para anticiparse

a situaciones imprevistas y acertar en la toma de decisiones, no sólo para los legisladores sino también para los gobiernos regionales y locales. Para un ciudadano, conocer la probabilidad si un proyecto de ley de su interés será exitoso, incrementa la confianza y transparencia con respecto a la gestión parlamentaria y también se transforma en un insumo para la toma de decisiones de los legisladores al momento de elaborar los proyectos de ley en la que se incluyen costos presupuestales, costos en tiempos dedicados y costos en recursos humanos. Esto nos lleva a reflexionar acerca de la necesidad de predecir la viabilidad legislativa para la presentación de propuestas parlamentarias que tengan una alta probabilidad de éxito, para reducir costos e incrementar los beneficios a la sociedad realizando una adecuada gestión de las técnicas de IA y de los datos masivos.

Formulación del Problema

Problema Principal

¿En qué medida las redes neuronales artificiales predicen la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020?

Problemas Derivados

- ¿Se ha determinado la estructura de la información para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020?
- ¿Se ha determinado la calidad de la información para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020?
- ¿Se ha determinado la técnica algorítmica para su aplicación para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020?
- ¿Se ha determinado los resultados para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020?

Objetivos de la Investigación

Objetivo General

Predecir la viabilidad legislativa en las proposiciones parlamentarias aplicando las redes neuronales artificiales, Perú, año 2020.

Objetivos Específicos

- Determinar la estructura de la información mediante la fase CRISP-DM de comprensión de los datos para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020.
- Determinar la calidad de la información mediante la fase CRISP-DM de preparación de los datos para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020.
- Determinar la técnica algorítmica para su aplicación mediante la fase CRISP-DM de modelamiento para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020.
- Determinar los resultados mediante la fase CRISP-DM de evaluación para predecir la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020.

Justificación de la Investigación

Importancia de la Investigación

El hombre como ser sociable requiere vivir en sociedad, y vivir en sociedad significa contar con leyes que indiquen qué está permitido y qué no lo está. Las leyes son prioritarias para que exista un orden, de tal manera, exista un respeto mutuo de los derechos y las libertades de todos los integrantes de un grupo humano. Los parlamentos se encargan de crear leyes que impactan la vida de millones de ciudadanos, esta labor trasciende los cambios de una sociedad. Las proposiciones parlamentarias son la base para la creación de las leyes, éstos contienen un conjunto de disposiciones que deben ser cumplidos, pero antes que todo pasan por un proceso de estudio y aceptación en el Poder Legislativo.

El estudio de la actividad legislativa según (Huertas Hernández, 2019) es un tema que data a inicios de la década del 50 del siglo pasado, en la que nace el interés por los conceptos y el afán de medir el éxito, la efectividad y la productividad legislativa. Por otra parte, García Montero (2007) hace mención que, en mayoría, los trabajos realizados acerca de la actividad legislativa toman como caso de análisis al Congreso estadounidense y, en forma escasa, a los países latinoamericanos. Los poderes legislativos, según el estudio realizado por Perez Paredes (2020), aprueban leyes apresuradamente y no consideran la prognosis de impactos ni verifican si estas impactan en forma positiva o negativa, también refiere que nadie en su sano juicio puede negar que un buen proyecto de ley, sustentado por una adecuada prognosis de impactos, generalmente llegará a ser una buena ley. De acuerdo con lo mencionado anteriormente por los autores, se hace necesario ampliar la investigación de la actividad legislativa en este caso,

de la viabilidad legislativa, que podría considerarse como una actividad que guarda un problema “invisible”, no tomado en cuenta al momento de la elaboración y seguimiento de los proyectos de ley.

Conocer y aplicar metodologías y técnicas algorítmicas nos abre un abanico de oportunidades y, el sector público no puede ser ajeno a estas tecnologías. Referirse a las redes neuronales artificiales, no es más que hablar de técnicas de algoritmos que hoy en día se está aplicando en la vida cotidiana, dentro de este contexto, estas técnicas requieren el manejo de grandes cantidades de datos que implican una adecuada gestión de la información. Es imprescindible mencionar que, una predicción altamente aceptable permite evaluar anticipadamente, el uso óptimo y oportuno de los recursos que son necesarios en la elaboración de las leyes.

Mediante esta investigación, se espera incentivar el uso de las herramientas de IA, que en la actualidad, empiezan a tener un impacto a nivel mundial. Predecir se convertirá en una actividad principal en el quehacer legislativo para mejorar la optimización de recursos (tiempo en las actividades, tiempo en la elaboración de un proyecto de ley, costos en recursos humanos y contrato de especialistas, costos presupuestales), incrementar la transparencia y ser en gran medida beneficioso para la sociedad. Por ende, es importante el estudio de las proposiciones parlamentarias en todos los niveles investigativos. Es por ello que, en el presente estudio, el investigador incursiona en la problemática de la viabilidad legislativa y busca demostrar que mediante el uso de las redes neuronales artificiales se puede alcanzar un porcentaje de predicción aceptable.

Viabilidad de la Investigación

Concerniente a la viabilidad del estudio, el investigador cuenta con conocimientos y con una amplia experiencia en el campo de la informática y los procesos legislativos. Con respecto a la recolección de la información, no existió restricción para uso investigativo, ya que los registros de las proposiciones parlamentarias son de acceso público y se encuentran publicados en el portal del Congreso de la República. Referente a la obtención de la información del marco teórico, se contó con una variada bibliografía, como los estudios publicados recientemente por otros investigadores mediante canales físicos y digitales, no existiendo limitaciones económicas para su adquisición. En la Tabla 2 se muestra el presupuesto ejecutado.

Para complementar los conocimientos acerca del desarrollo metodológico de la investigación, se tomó en cuenta llevar cursos y especializaciones para su profundización. Haciendo uso de la libertad para elegir a los autores metodológicos, el investigador ha considerado en esta investigación, no tomar en cuenta la metodología clásica de los autores como Roberto Hernández Sampieri y Gordon Dankhe, dado que, no comparte ciertos conceptos y criterios metodológicos.

Tabla 2

Presupuesto de la investigación

	Cantidad	Precio	
		Unitario	Subtotal
1. Materiales de oficina			
Paquete 500 hojas A4	4	S/ 11.00	S/ 44.00
Grapadora	2	S/ 15.00	S/ 30.00
Perforadora	2	S/ 15.00	S/ 30.00
Cuaderno de apuntes	2	S/ 22.00	S/ 44.00
Resaltadores	2	S/ 4.00	S/ 8.00

	Cantidad	Precio	
		Unitario	Subtotal
Lapiceros Frixion Ball	12	S/ 7.00	S/ 84.00
1/2 litro tinta impresora	4	S/ 45.00	S/ 180.00
Laptop HP Pavilion AMD Rayzen 3	1	S/ 2,300.00	S/ 2,300.00
			S/ 2,720.00
2. Servicios			
Internet Movistar Total de 150 mbps 12 meses	12	S/ 217.00	S/ 2,604.00
Espiralados	10	S/ 5.00	S/ 50.00
Empastados	3	S/ 50.00	S/ 150.00
			S/ 2,804.00
3. Materiales de investigación			
Libro "La evaluación de las leyes en el Perú" PUCP	1	S/ 170.00	S/ 170.00
Libro "Metodología de la Investigación" Dr. Supo	1	S/ 180.00	S/ 180.00
Libro "El secreto detrás de una tesis" Dra. Rosario Martínez	1	S/ 40.00	S/ 40.00
Libro "Epidemiología y estadística aplicada a la salud" Luis Alex Alzamora de los Godos Urcia	1	S/ 230.00	S/ 230.00
Libro formato PDF "Metodología de la investigación científica en posgrado" Luis Alex Alzamora de los Godos Urcia	1	S/ 45.00	S/ 45.00
Libro formato EPUB "La evaluación de las leyes en el Perú" PUCP	1	S/ 50.00	S/ 50.00
Libro formato PDF "El aprendizaje automático" PUCP	1	S/ 36.00	S/ 36.00
Libro formato Kindle "El proyecto de investigación" Jacqueline Hurtado de Barrera	1	S/ 38.00	S/ 38.00
Libro formato Kindle "Cómo formular objetivos de investigación" Jacqueline Hurtado de Barrera	1	S/ 38.00	S/ 38.00
Libro formato Kindle "Filosofía de la ciencia" Marcos Fidel Barrera Morales	1	S/ 38.00	S/ 38.00
Libro formato Kindle "Inteligencia artificial" Lasse Rouhiainen	1	S/ 20.00	S/ 20.00
Libro formato Kindle "Máquinas predictivas" Ajay Agrawal, Joshua Gans, Avi Goldfars	1	S/ 38.00	S/ 38.00
Libro formato Kindle "Machine Learning" Rudolph Russell	1	S/ 23.00	S/ 23.00
Libro formato Kindle "Deep Learning" Rudolph Russell	1	S/ 23.00	S/ 23.00
Libro formato Kindle "Inteligencia artificial" Neil Wikings	1	S/ 38.00	S/ 38.00
			S/ 1,007.00
4. Capacitaciones			
Especialidad en Docencia e Investigación Científica – SINCIE	1	S/ 4,330.00	S/ 4,330.00
Diplomado en Estadística e Investigación Científica – SINCIE	1	S/ 4,330.00	S/ 4,330.00
Curso de Estadística – Univ. Politécnica de Valencia	1	S/ 350.00	S/ 350.00
Curso de Estadística – UPCH	1	S/ 450.00	S/ 450.00
Curso de Investigación Social – UPCH	1	S/ 450.00	S/ 450.00
Curso "Mi tesis de éxito" – Dra. Rocío Lima Orbegoso	1	S/ 460.00	S/ 460.00

	Cantidad	Precio	
		Unitario	Subtotal
Curso "Tesis online" – Dra. Rosario Martínez González	1	S/ 500.00	S/ 500.00
			S/ 10,870.00
5. Software			
SPSS Modeler versión 18	2 meses	S/ 1,896.00	S/ 3,792.40
SPSS Estadísticos versión 25	2 meses	S/ 376.20	S/ 752.40
Octoparse versión 8	2 meses	S/ 338.20	S/ 676.40
Microsoft Office 365 Excel	1	S/ 0.00	S/ 0.00
Microsoft Office 365 Word	1	S/ 0.00	S/ 0.00
			S/ 5,221.20
		Total	S/ 22,622.20

Fuente: elaboración propia

Limitaciones del Estudio

Entre las limitaciones encontradas en la presente investigación se consideraron: el formato de los registros, la obtención de la información, el software para sustracción de datos y el acceso al portal del Congreso de la República. A continuación, se detallan las limitaciones:

Formato de los datos

Las proposiciones parlamentarias publicadas en el portal del Congreso de la República están en formato HTML y no en un formato de CSV, una base de datos o un archivo Excel que facilitaría la obtención de los datos y brindaría una fidelidad del contenido de las proposiciones.

Obtención de datos adicionales

Las proposiciones parlamentarias publicadas en el portal del Congreso de la República son de una cantidad considerable, al realizar el levantamiento de información se verificó que la página web no cuenta con todos los datos

necesarios para realizar el estudio de acuerdo con las variables definidas por la investigación realizada por (Arias-Schreiber Barba, Peña Jumba, & Valdivieso Serrano, 2019), se requiere buscar en otras páginas web para poder complementarlo, caso contrario, se propuso obviar las variables con datos no obtenidos.

Software para la extracción de datos

Actualmente se cuenta con herramientas tecnológicas que ayudan a extraer la información de una página web, pero no se garantiza una extracción deseada, normalmente se realiza manualmente la verificación de los datos extraídos.

Acceso al Portal del Congreso de la República

En ocasiones se presentaron caídas en sus servidores y ataques externos que dificultaron la extracción de datos de las proposiciones parlamentarias, por ejemplo, en las primeras semanas de noviembre de 2020 durante la transición de gobierno ocurrieron una serie de ataques cibernéticos contra el portal institucional manteniéndola inactiva durante varios días.

CAPÍTULO I

MARCO TEÓRICO CONCEPTUAL

1.1. Antecedentes de la Investigación

Antecedentes Nacionales

El trabajo de (Zevallos Salazar, 2017) trata acerca del rendimiento académico y la aplicación de las redes neuronales. El objetivo de la investigación fue predecir el rendimiento académico de los estudiantes del nivel primario pertenecientes al distrito de Ventanilla. El autor propuso la utilización de la técnica del perceptrón multicapa cuyos valores de entrada fueron 33 factores del rendimiento académico. La población de estudio lo conformó 1,073 estudiantes de tres centros educativos. Como resultado se obtuvo un 84% de precisión predictiva. El autor concluye que es posible predecir el rendimiento académico mediante el modelo generado, dado su alto grado de validez.

El estudio de (Aguilar Vilca & Camargo Ramos, 2021) fue referente a una aplicación basada en el uso de tres técnicas algorítmica avanzadas de predicción y la deserción de clientes. El objetivo del estudio fue predecir la deserción de clientes de microcréditos de bancos. Los autores propusieron un modelo híbrido para mejorar la capacidad de predicción aplicando las técnicas de máquinas de soporte vectorial, redes neuronales y random forest en una población de 24,420 clientes y 37 atributos, como metodología para el trabajo de investigación aplicaron Knowledge Discovery in Databases (KDD). El resultado obtenido en las pruebas de la predicción fue del 97.38%.

Los autores concluyen que la combinación de las tres técnicas mejora los resultados de la predicción dando un porcentaje altamente aceptable.

Antecedentes Internacionales

Kornilova, Argyle, & Eidelman (2018) realizaron el estudio acerca de la predicción de votos de los congresistas. Este estudio tuvo como objetivo predecir cómo los legisladores del Congreso votarán para la introducción de un proyecto de ley. La metodología consistió en la evaluación de ocho modelos predictivos mediante mecanismo de las redes neuronales. La población de estudio estaba conformada por las votaciones presentadas en la 106^o a 111^o períodos de sesiones del Congreso de Estados Unidos (cerca de 20,000 proyectos de ley en cada sesión), cada proyecto de ley contó con un promedio de 2,683 palabras. En los resultados se obtuvo un 86.21% de precisión entre los ocho modelos evaluados, siendo el de más alta precisión las CNN+Meta . Los autores concluyeron que, los legisladores votan en base al tema y a la ideología introducida en el proyecto de ley.

Khashman & Khashman (2016) trataron acerca del estudio de la predicción de votos del Congreso de Estados Unidos. El objetivo de este estudio consistió en la predicción de votos de partidos políticos utilizando la inteligencia artificial. La metodología propuesta asoció las opiniones de los congresistas estadounidenses sobre temas nacionales y se relacionó a sus partidos políticos como el republicano o el demócrata, es decir, se basó en el conocimiento previo de cómo los miembros del congreso percibían los problemas nacionales, la población de estudio constó de 435 instancias (267 demócratas, 168 republicanos) y 16 atributos. Como resultado, el valor

global obtenido fue del 96.33% utilizando el modelo de redes neuronales, el modelo de máquina de vectores de soporte obtuvo un valor global del 95.29%. En conclusión, en base a los resultados experimentales obtenidos se demostró que el método aplicado se puede usar en más encuestas de votación similares con el fin de anticipar la inclinación de voto de los miembros del partido.

Zolghadr, Niaki, & Akhavan Niaki (2016) realizaron su investigación acerca del pronóstico de las elecciones presidenciales de los Estados Unidos. El objetivo principal de esta investigación fue obtener un modelo de pronóstico preciso para las elecciones presidenciales mediante algoritmos de aprendizaje. En la metodología, para inicializar el desarrollo del modelo, los métodos de preprocesamiento utilizados en esta investigación fueron la transformación de datos, la reducción de datos, y el agrupamiento, además, los algoritmos de aprendizaje empleados para el pronóstico fueron regresión lineal, regresión de vectores de soporte y redes neuronales. El modelo de pronóstico usó datos de las elecciones presidenciales de 2004, 2008 y 2012, con una población de 791 votos, segmentados en 16 conjuntos de datos y ocho atributos. Los resultados obtenidos fueron de un error porcentual absoluto medio (MAPE) de 1.86 para SVR y 2.59 para las ANN, tomando en cuenta que la precisión se incrementó al 50% se ha considerado un 95% para ambos resultados, restando el porcentaje de error sería para SVR una precisión de 93.14% y para la ANN un 92%. Los autores concluyen que la comparación demostró que los algoritmos de aprendizaje son mucho mejores para reducir el error en comparación con la regresión lineal.

Además, el modelo SVR fue identificado como el modelo de mejor rendimiento y resultó exitoso en pronosticar las elecciones presidenciales.

Colliri & Zhao (2019) realizaron un estudio acerca de los proyectos de ley y la predicción de condenas por corrupción. Su objetivo del estudio fue de analizar los proyectos de ley y predecir las condenas por corrupción de los congresistas brasileños aplicando redes temporales. La metodología propone una técnica basada en la red para analizar los datos de votación de proyectos de ley que comprendían votos de los congresistas brasileños, recopilados del sitio web oficial de la Cámara de Representantes de Brasil, estos conjuntos de datos comprenden 3,407 sesiones de votación de proyectos de ley en la Cámara de Representantes, desde el 22 de mayo de 1991 hasta el 14 de febrero de 2019 y de 10 atributos (6 votantes y 4 votos). En los resultados de los seis modelos predictivos aplicados, se obtuvo un 90%, entre los tres primeros modelos seguidos de un 70%, 50% y 0% respectivamente. El autor concluye que, se puede obtener una alta precisión de predicción de convicciones utilizando datos de votación de proyectos de ley, lo que implica que es posible revelar el comportamiento ilegal de los políticos a través de sus actividades públicas legales.

Kaur & Bozié (2019) presentaron un trabajo acerca de la predicción de las sentencias. Este estudio tuvo como objetivo, predecir las sentencias del Tribunal Europeo de Derechos Humanos basadas en redes neuronales convolucionales. Como metodología para el trabajo de investigación implementaron el modelo CRISP-D, los pasos principales incorporados fueron la comprensión de datos, preparación de datos, modelado y

evaluación. El total de casos judiciales para el entrenamiento y el conjunto de datos de prueba fueron 3,132 y 8,400 respectivamente y nueve atributos. Con base en los resultados obtenidos, se estableció que en general, los modelos de las redes neuronales convolucionales superaron a los modelos de soporte de vectores ya que el primero alcanzó una precisión media de 82% mientras que este último alcanzó el 75%. Los autores concluyeron que, específicamente, los modelos de redes convolucionales lograron una precisión estadísticamente mayor que modelos de máquina de vectores de soporte.

1.2. Bases Teóricas

1.2.1. *Redes Neuronales Artificiales*

El presente estudio está basado en el uso del algoritmo de las redes neuronales artificiales (RNA), denominadas en inglés artificial neural networks (ANN). La base teórica de esta variable se sustenta en (Haykin, 2016), que describe una red neuronal como un procesador distribuido masivamente en paralelo compuesto por unidades de procesamiento simples con una tendencia natural para almacenar experiencias y disponerlo para su propio uso. El autor lo describe lo parecido al cerebro humano en dos aspectos:

- El conocimiento se obtiene por la red de su entorno a través de un proceso de aprendizaje.
- La conexión entre neuronas, conocidas como pesos sinápticos, se utilizan para almacenar el conocimiento adquirido.

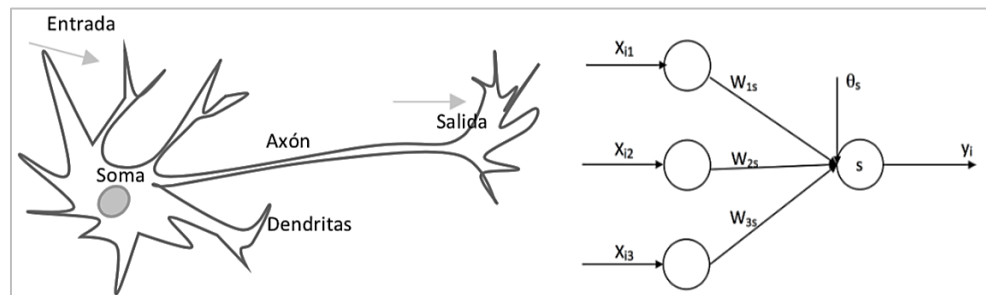
Una red neuronal es una máquina diseñada para ayudar a modelar una determinada tarea o función de interés; la red generalmente es implementada mediante la simulación usando un software. Para lograr un rendimiento aceptable, las ANN emplean una interconexión de células informáticas llamadas neuronas.

El procedimiento utilizado para realizar el proceso de aprendizaje es denominado algoritmo de aprendizaje, cuyo fin es modificar los pesos sinápticos de la red de manera ordenada para

lograr un modelo deseado. Una red neuronal es capaz de modificar su propia topología, similar a las neuronas del cerebro humano, pueden morir y pueden nacer conexiones sinápticas.

Figura 9

Izquierda: neurona en el cerebro. Derecha: red neuronal artificial con tres neuronas en la capa de entrada y una neurona en la capa de salida.



Fuente: (Véliz Capuñay, 2020)

Historia de las Redes Neuronales. A continuación (Stanford University's Computer Science Department, 2020) describe la historia de las ANN.

En 1943, el neurofisiólogo Warren McCulloch y el matemático Walter Pitts publicaron un artículo sobre el funcionamiento de las neuronas y modelaron una red neuronal simple utilizando circuitos eléctricos.

En 1949, Donald Hebb redacta *The Organization of Behavior*, un trabajo que mencionaba que las vías neuronales se fortalecen cada vez que se utilizan al mismo tiempo, argumentó que, la conexión entre ellos mejoraba más.

En la década del 1950, a medida que las computadoras evolucionaban, se pudo simular una red neuronal hipotética. Nathaniel Rochester lo desarrolló en los laboratorios de investigación de IBM.

En 1959, Bernard Widrow y Marcian Hoff de Stanford crean los modelos llamados "ADALINE" y "MADALINE". ADALINE se desarrolló para reconocer patrones binarios. MADALINE fue la primera red neuronal aplicada a un problema del mundo real.

En 1962, Widrow y Hoff desarrollaron un procedimiento de aprendizaje para el perceptrón que examina el valor antes de que el peso lo ajuste, se basa en la idea de que, se pueden ajustar los valores de peso para distribuirlo a través de la red. John von Neumann propuso la imitación de funciones neuronales a través del uso de relés telegráficos o tubos de vacío. La escasa información y las cuestiones filosóficas dieron como resultado, que los estudios y las financiaciones de las investigaciones se redujeran drásticamente.

En 1972, Kohonen y Anderson crearon una red de forma independiente entre sí. Aplicaron matemáticas matriciales para crear una serie de circuitos ADALINE analógicos.

En 1975, se desarrolló la primera red multicapa que era una red no supervisada.

En 1982, John Hopfield de Caltech publicó un artículo a la Academia Nacional de Ciencias acerca de crear máquinas más

útiles mediante el uso de líneas bidireccionales. Anteriormente, las conexiones entre neuronas eran solo de una forma. Reilly y Cooper utilizan una "red híbrida" con múltiples capas. Se realiza una conferencia entre Estados Unidos y Japón acerca de las ANN, Japón anunció su quinta generación sobre ANN, y generó preocupación de que Estados Unidos pudiera quedarse atrás en el campo.

En 1986, David Rumelhart, un ex miembro del departamento de psicología de Stanford, propuso ideas que ahora se denominan redes de retropropagación porque distribuye errores de reconocimiento de patrones.

Analytics Vidhya (2019) menciona también que en 1987, se realiza la primera Conferencia Internacional sobre Redes Neuronales organizada por el Instituto de Ingenieros Eléctricos y Electrónicos (IEEE). Esta reunión atrajo a más de 1,800 participantes.

En 1997, Schmidhuber y Hochreiter introducen un marco de red neuronal recurrente denominada Long Short-Term Memory (LSTM).

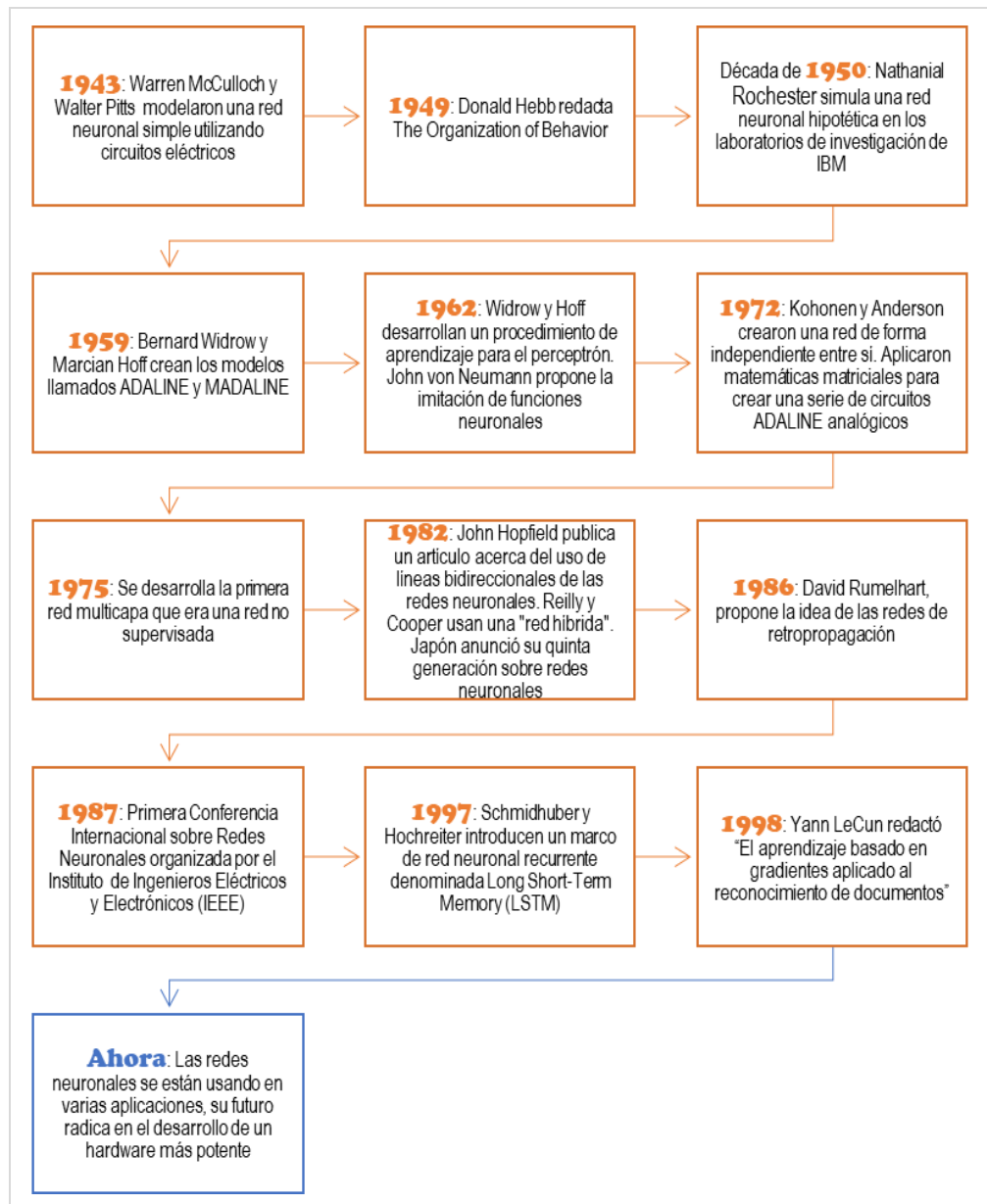
En 1998, Yann LeCun redactó "El aprendizaje basado en gradientes aplicado al reconocimiento de documentos".

Finalmente, Stanford University's Computer Science Department (2020) refiere que en la actualidad, las ANN se están usando en varias aplicaciones, su futuro radica en el desarrollo del hardware, debido a las limitaciones de los procesadores. Los

diversos chips que se están desarrollando son los digitales, analógicos y ópticos, a pesar de eso, va transcurrir un tiempo antes que los chips ópticos puedan ser usados a nivel comercial.

Figura 10

Infografía de la historia de las redes neuronales artificiales



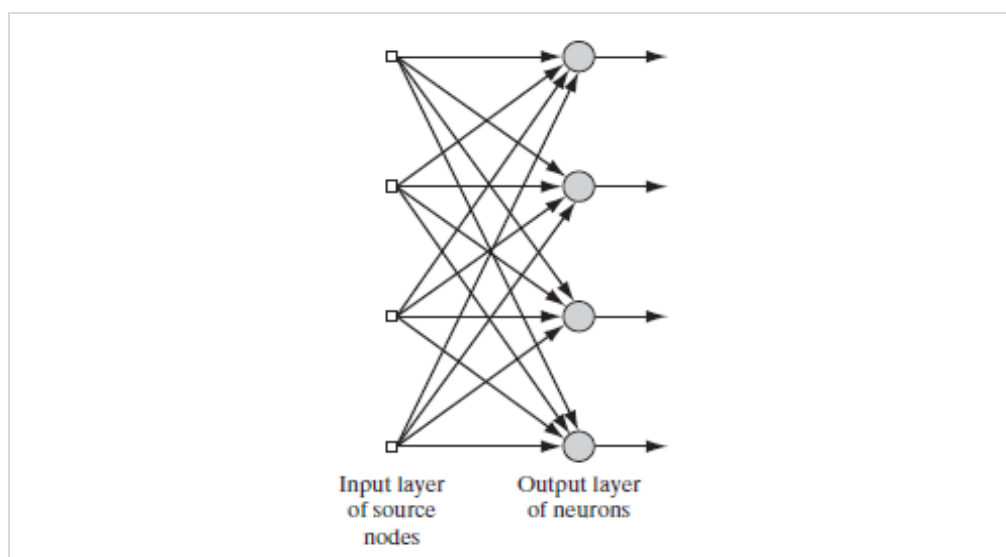
Fuente: elaboración propia con datos de (Stanford University's Computer Science Department, 2020) y (Analytics Vidhya, 2019)

Arquitecturas de las Redes Neuronales Artificiales. Según la definición de (Haykin, 2016) la arquitectura de las neuronas de una red neuronal está relacionada al algoritmo de aprendizaje utilizado para entrenar la red, por lo que podemos hablar de algoritmos de aprendizaje (reglas) utilizados en el diseño de ANN como estructurados. Haykin (2016) clasifica tres diferentes arquitecturas de red:

Redes Feedforward de Una Sola Capa. Es la más simple de una red en capas, una capa de entrada de nodos de origen se proyecta directamente sobre una capa de salida de neuronas (nodos de cálculo). La denominación de "una sola capa" hace referencia a la capa de salida de los nodos de cálculo (neuronas). La capa de entrada de los nodos de origen no se considera porque allí no se realiza ningún cálculo.

Figura 11

Red feedforward con una sola capa de neuronas.

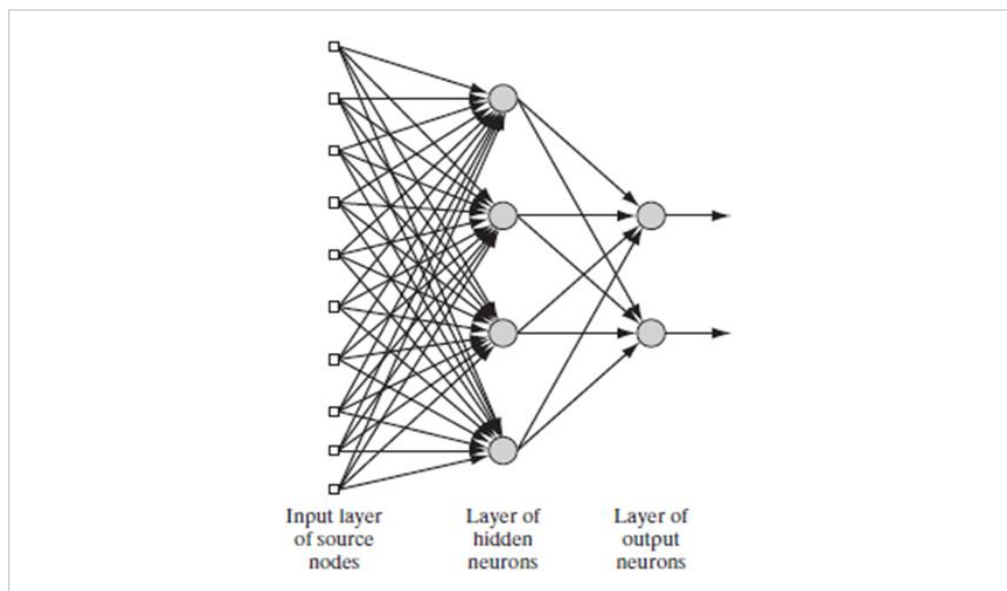


Fuente: (Haykin, 2016)

Redes Feedforward Multicapa. Se caracteriza por contener una o más capas ocultas, sus nodos de cálculo se denominan neuronas ocultas; el término "oculto" significa que esta parte de la red neuronal no se ve directamente sobre una capa de entrada ni salida de una red. Al agregar a la red más capas ocultas, la red está habilitada para la extracción de datos estadísticos de orden superior de su entrada. Cuando cada nodo de cada capa de la red está conectado (conexiones sinápticas) a todos los demás nodos de la capa adyacente se dice que está totalmente conectada. De otra manera, si existen algunos nodos no conectados se dice que la red está parcialmente conectada.

Figura 12

Red feedforward completamente conectada con una capa oculta y una capa de salida.

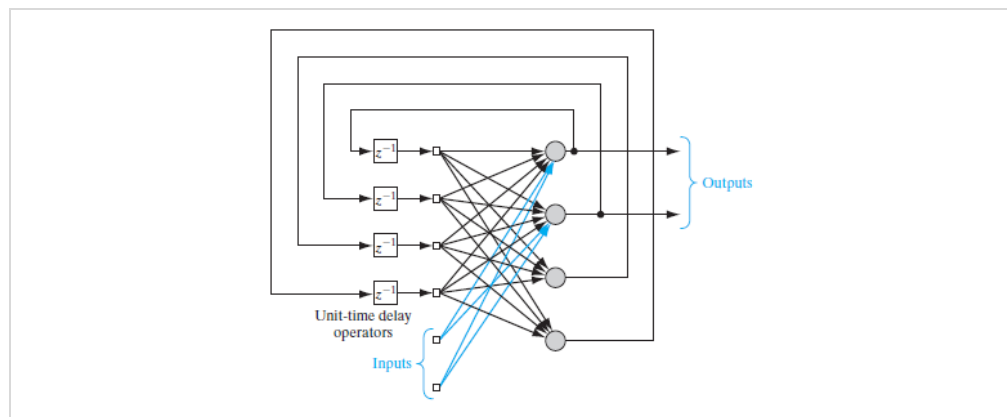


Fuente: (Haykin, 2016)

Redes Recurrentes. Una red neuronal recurrente se diferencia de una red neuronal feedforward cuando posee al menos un feedback loop. Por ejemplo, una red recurrente de una sola capa puede tener neuronas que conectan su señal de salida a las entradas de todas las demás neuronas, formado así, un bucle de conexiones sinápticas.

Figura 13

Red recurrente con neuronas ocultas.



Fuente: (Haykin, 2016)

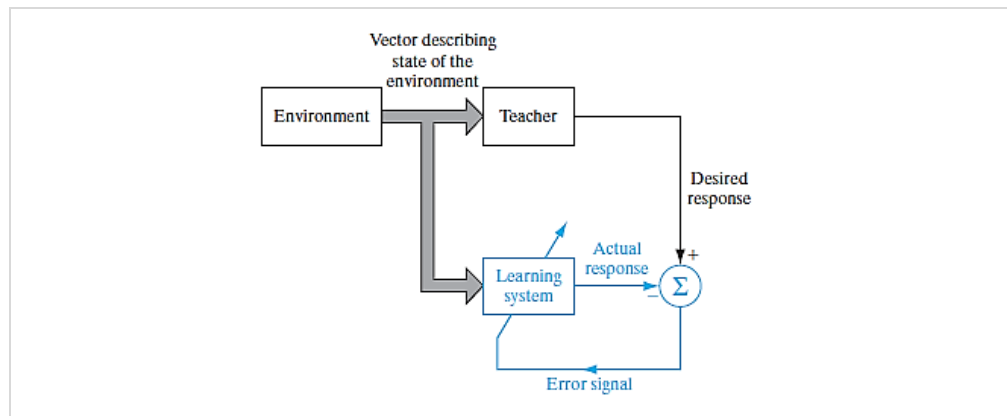
Los Procesos de Aprendizaje. Haykin (2016) menciona que hay dos procesos de aprendizaje: el proceso de aprendizaje con instructor y el proceso de aprendizaje sin instructor. Dentro del proceso de aprendizaje con instructor encontramos el aprendizaje supervisado. En el proceso de aprendizaje sin instructor encontramos dos tipos de procesos: el aprendizaje no supervisado y el aprendizaje por refuerzo.

El Aprendizaje Supervisado. Se entrena con datos etiquetados. Las etiquetas clasifican cada dato de diferentes grupos.

El programa aprende cómo se construyen estos datos, a estos datos se les ha denominado datos de entrenamiento, y los utiliza para predecir las categorías de los datos nuevos, conocidos como datos de prueba.

Figura 14

Diagrama de bloques de aprendizaje supervisado.



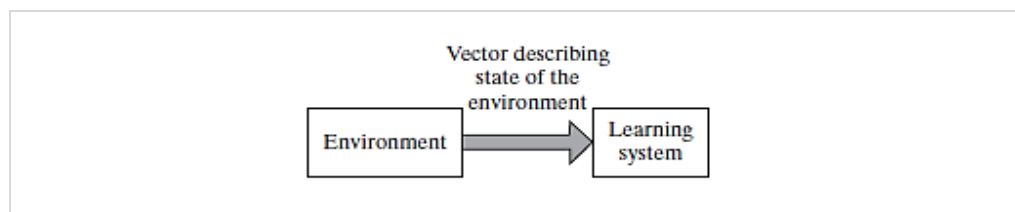
Fuente: (Haykin, 2016)

El Aprendizaje No Supervisado. Es aprender sin etiquetas.

Su objetivo es detectar los atributos de los datos que sean más o menos similares entre sí, por ejemplo, clasificando grupos y asignando datos a estos grupos.

Figura 15

Diagrama de bloques de aprendizaje no supervisado.

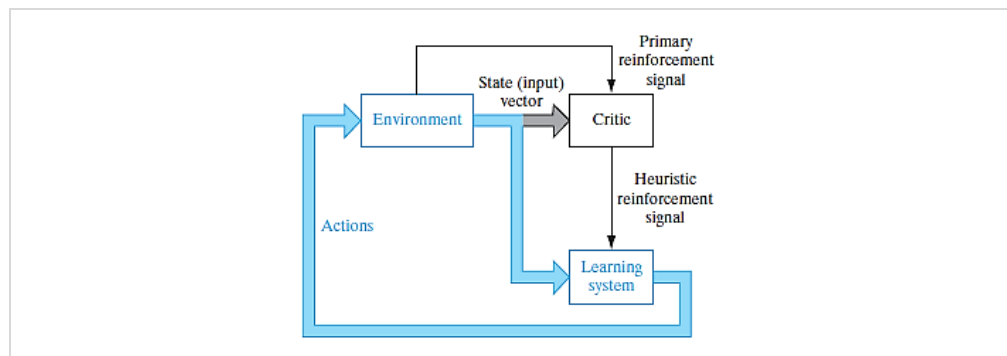


Fuente: (Haykin, 2016)

El Aprendizaje por Refuerzo. Se orienta en aprender de la experiencia y usa técnicas del aprendizaje supervisado y el no supervisado. En el aprendizaje por refuerzo, un agente interactúa con su entorno y recibe una función de recompensa que va ir optimizándolo, por ejemplo, un programa podría ser recompensado por ganar un juego. El objetivo del agente es aprender qué acciones fueron más importantes para ganar un juego, de acuerdo a este aprendizaje buscar estrategias que maximicen sus recompensas.

Figura 16

Diagrama de bloques de aprendizaje por refuerzo; el sistema de aprendizaje y el entorno están ambos dentro del circuito de retroalimentación.



Fuente: (Haykin, 2016)

Las Técnicas de Aprendizaje. Berzal (2018) clasifica en cinco técnicas de aprendizaje que son: las simbólicas, analógicas, bayesianas, evolutivas y las conexionistas. A continuación, se detallan los algoritmos que se derivan de la clasificación.

Tabla 3

Técnicas de aprendizaje.

Técnicas	Algoritmos
1. Las técnicas simbólicas (IA simbólica)	Los algoritmos de árboles de decisión, clasificación o identificación. Los algoritmos de inducción de reglas
2. Técnicas analógicas (Reconocimiento de patrones)	Vecinos más cercanos, k-NN. Máquinas de vectores de soporte (SVM), SVR
3. Técnicas bayesianas (Modelos probabilísticos)	Naïve Bayes Redes bayesianas Redes de Markov
4. Técnicas evolutivas (Computación evolutiva)	Algoritmos genéticos Programación evolutiva Estrategias de evolución Programación genética
5. Técnicas conexionistas (IA conexionista)	Redes neuronales artificiales (ANN, CNN, RNN)
6. Combinación de múltiples modelos (Ensembles)	Bagging Boosting Stacking

Fuente: elaboración propia con datos de (Berzal, 2018).

Las Técnicas Simbólicas (IA Simbólica). Interpretan los problemas desde un punto de vista abstracto y, al resolver un problema, incorporan el conocimiento ya existente. Existen innumerables algoritmos de aprendizaje simbólico, entre ellos se encuentran los dos más conocidos:

- Los algoritmos de árboles de decisión, clasificación o identificación.
- Los algoritmos de inducción de reglas.

Técnicas Analógicas (Reconocimiento de Patrones). Basan el aprendizaje en la extrapolación a partir de casos conocidos, mediante la realización de juicios de similitud. Su

argumentación por analogía, en ocasiones, se acaba resolviendo un problema matemático de optimización. Entre los algoritmos tenemos:

- Vecinos más cercanos, k-NN.
- Máquinas de vectores de soporte (SVM), SVR.

Técnicas Bayesianas (Modelos Probabilísticos). Aplican técnicas estadísticas de inferencia probabilística, apoyándose en el teorema de Bayes. Entre los algoritmos tenemos:

- Naïve Bayes
- Redes bayesianas
- Redes de Markov

Técnicas Evolutivas (Computación Evolutiva). Se fundamentan en la teoría de la evolución de Darwin y de las bases de la genética. Entre los algoritmos tenemos:

- Algoritmos genéticos
- Programación evolutiva
- Estrategias de evolución
- Programación genética

Técnicas Conexionistas (IA Conexionista). Se inspiran en el cerebro humano, tratan de hacer ingeniería inversa de su funcionamiento con el soporte de físicos y neurocientíficos. La

técnica utilizada es las redes neuronales artificiales (ANN, CNN, RNN).

Combinación de Múltiples Modelos (Ensembles). Se refiere al uso de una colección de modelos con el propósito de realizar una predicción y de mejorar el rendimiento de cada uno de esos modelos por separado.

- Bagging, varios clasificadores diferentes votan para decidir la clase de un caso particular, la votación es por mayoría simple.
- Boosting, los clasificadores tienen distintos pesos en la votación dependiendo de su precisión, usan una votación ponderada.
- Stacking, se entrenan múltiples modelos y se entrena un algoritmo de combinación para realizar una predicción final a partir de las predicciones de los modelos entrenados.

Diferencias entre Inteligencia Artificial, Machine Learning y Deep Learning. Management Solutions (2018) menciona que la Inteligencia Artificial (IA), Machine Learning (ML) y Deep Learning (DL) son conceptos que están relacionados. Con el objetivo de ilustrar dichas diferencias se proporcionan a continuación las siguientes definiciones:

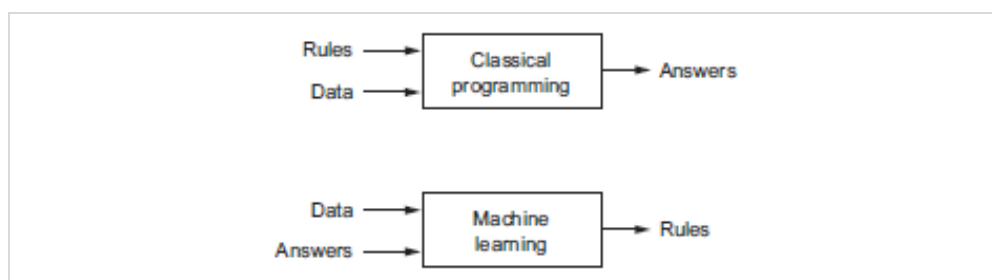
Inteligencia Artificial (IA). La inteligencia artificial tiene una connotación amplia, su objetivo es que las máquinas realicen tareas

de la misma forma que un ser humano. Mayormente, esto se desarrolla a través de la ejecución de reglas previamente programadas. En la Conferencia de IA de Dartmouth (1956) se le definió como “todas las formas de aprendizaje u otra característica de la inteligencia que una máquina sea capaz de realizarla”. A inicios del siglo XX hay algunos ejemplos como Alan Turing, quien descifró la máquina Enigma, fue la primera muestra de lo que hoy llamaríamos las redes neuronales.

Machine Learning (ML). Por otra parte, el machine learning o aprendizaje automático es una rama de la IA, y se define como un conjunto de métodos que realiza la detección de patrones automáticamente en un conjunto de datos para predecir datos futuros, o para ejecutar otras decisiones en entornos de incertidumbre.

Figura 17

Machine learning un nuevo paradigma de programación.



Fuente: (Chollet, 2018)

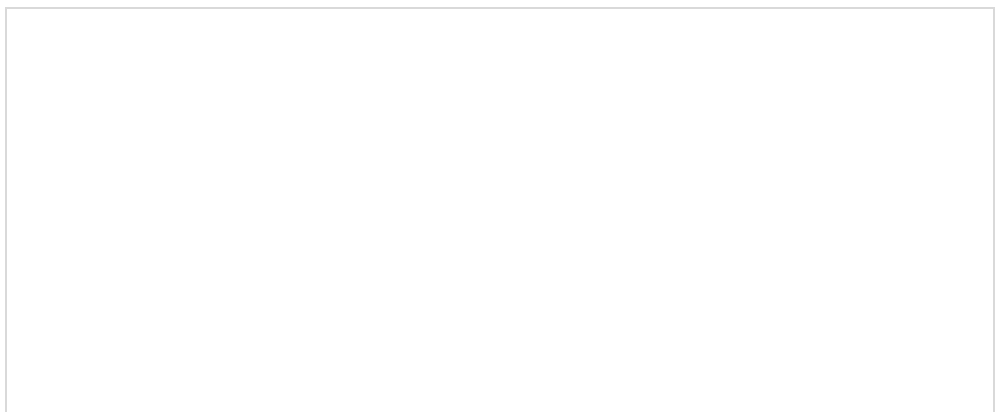
Deep Learning (DL). Por último, el Deep Learning o aprendizaje profundo es una rama del Machine Learning que se puede definir como un sistema que permite a los modelos

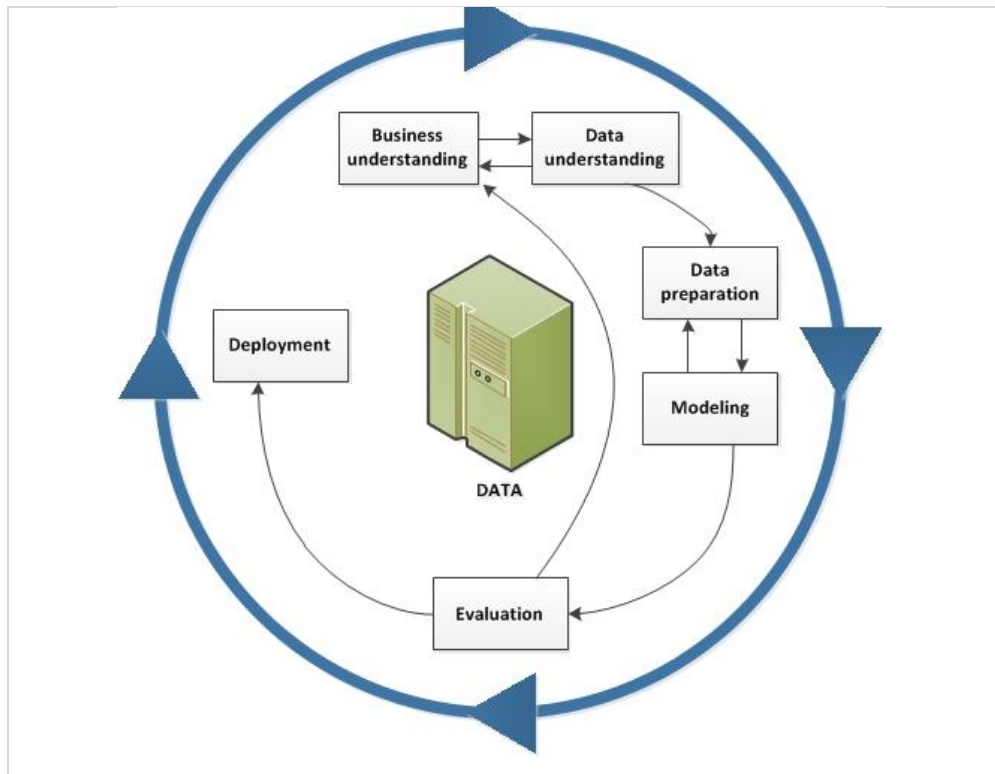
computacionales que están compuestos de múltiples capas de procesamiento aprender sobre datos con múltiples niveles de abstracción.

La Metodología CRISP-DM. IBM (2020) refiere que CRISP-DM son las siglas de Cross-Industry Standard Process for Data Mining, es un método utilizado para orientar trabajos de minería de datos. La (Universitat Oberta de Catalunya, 2020) señala que esta metodología nació de dos empresas, DaimlerChrysler y SPSS, precursoras en la aplicación de técnicas de minería de datos en los procesos de negocio. El ciclo del método lo conforman seis fases con dependencias secuenciales de mayor prioridad y frecuencia entre fases, la secuencia entre las fases no suele ser estricta y la mayoría de los proyectos avanzan y retroceden entre fases según sus necesidades.

Figura 18

Fases de la metodología CRISP-DM





Fuente: (IBM, 2020)

A continuación, de acuerdo con las definiciones de la (Universitat Oberta de Catalunya, 2020) se detalla las seis fases:

Fase 1 - Comprensión del Negocio o Comprensión Comercial. Esta fase trata de alcanzar desde una clara perspectiva de negocio cuáles son los objetivos, tratando de poder evitar el gran error de dedicar el esfuerzo de todo el proyecto a ofrecer respuestas correctas a preguntas equivocadas. Con los objetivos de negocio definidos, se elabora un estudio de la situación actual del negocio acerca de los objetivos planteados, en este punto, se trata de clarificar recursos, requerimientos y limitaciones, para así concretar objetivos de la minería de datos que apoyen claramente a la consecución de los objetivos primarios. Finalmente, se elabora un

plan de proyecto en el que se detalla las fases, tareas y actividades que se deberán llevar para concluir los objetivos planteados.

Fase 2 - Comprensión de los Datos. Se trata de trabajar los datos con el objetivo de conocerlos al máximo, saber su origen, en qué estados llegan, cuál es su estructura, qué propiedades contienen, qué inconvenientes presentan y cómo se puede mitigar o eliminar. Es una fase crítica puesto que se trabaja enteramente con la calidad de los datos, que por otro lado se debe ver como la materia prima para la minería de datos. Poseer una buena calidad de los datos es una condición necesaria, aunque no suficiente para alcanzar el éxito del proyecto.

Fase 3 - Preparación de Datos. El objetivo de esta fase es el de disponer completamente del juego de datos sobre el que se ejecutarán los modelos. También se desarrolla la documentación descriptiva que es necesaria para el conjunto de datos.

Fase 4 - Modelamiento. El objetivo de esta fase es el de disponer de un modelo que nos ayude a concretar los objetivos de la minería de datos y los objetivos de negocio detallados en el proyecto. Se entiende por modelo a la habilidad de aplicar una técnica a un conjunto de datos con la finalidad de predecir la variable objetivo o encontrar un patrón desconocido. La iteración de esta fase, tanto con su antecesora, acerca de la preparación de los datos, así como con su sucesora, acerca de la evaluación del modelo,

demuestra la importancia de la fase en términos de la calidad del proyecto.

Fase 5 - Evaluación. En las fases se asegura la fiabilidad y plausibilidad del modelo, en cambio, en esta fase se centra en evaluar el grado de aproximación de los objetivos de negocio y en la búsqueda, si se plantearon, por las cuales el modelo es ineficiente. Una forma esquemática y gráfica de visualizar el fin de un proyecto de minería de datos es pensar en la ecuación:

$$\text{Resultados} = \text{Modelos} + \text{Descubrimientos}$$

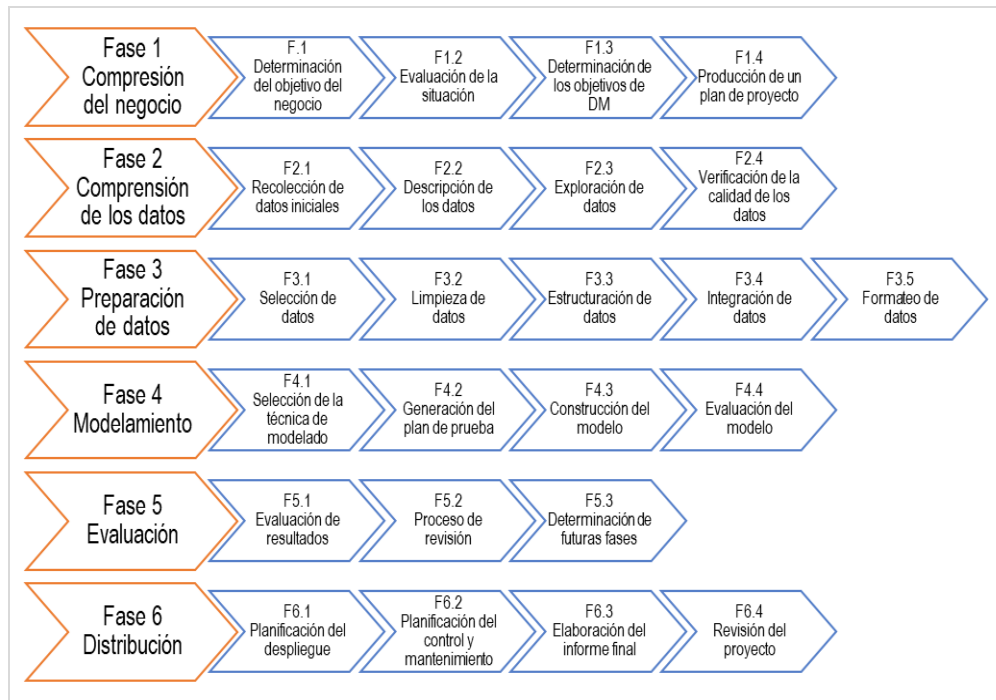
Esto representa que el propósito de un proyecto de minería de datos no sólo se refiere a los modelos, sino también los descubrimientos, que se podría definir como algo aparte del modelo que contribuye a obtener los objetivos de negocio o que ayude a plantear nuevas preguntas y que a la vez son decisivas para conseguir los objetivos de negocio.

Fase 6 - Distribución. En esta fase se organizan y ejecutan las tareas del despliegue de los resultados y del mantenimiento de las nuevas funcionalidades, cuando el despliegue haya finalizado.

En la Figura 19, se muestran las fases y las actividades o sub-fases que integran a la metodología CRISP-DM.

Figura 19

Fases y actividades de la metodología CRISP-DM



Fuente: elaboración propia con datos de (IBM, 2020)

1.2.2. Viabilidad legislativa

Arias-Schreiber Barba, Peña Jumpa, & Valdivieso Serrano (2019) consideran que una propuesta parlamentaria es viable legislativamente si cumple las siguientes condiciones:

- Fue aprobada como ley, a pesar de haber podido presentar un Análisis Costo Beneficio (ACB) inadecuado.
- Se encuentra en el proceso de orden del día.
- Alcanzó un dictamen favorable.
- Se encuentra en evaluación en las comisiones, sin importar su antigüedad, considerando que, no haya sido declarada inadmisibile, rechazada de plano, mandada al archivo, retirada, devuelta al autor o dictaminada negativamente.

Arias-Schreiber Barba, Peña Jumpa, & Valdivieso Serrano (2019) en su investigación identificaron 35 variables como factores de la viabilidad legislativa que fueron clasificadas en 4 grupos y son los siguientes:

Calidad del ACB Convencional (IC-ACB-C). Incluye 18 variables (17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33 y 34).

Calidad del ACB Legislativo (IC-ACB-L). Incluye 5 variables (17, 18, 19, 22 y 24).

Viabilidad legislativa sin ACB. Incluye una variable (35).

Relacionadas a los Proponentes y las Iniciativas. Incluye

15 variables (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 y 15).

A continuación, en la *Tabla 4*, *Tabla 5*, *Tabla 6* y la *Tabla 7* se muestran de manera más detallada, las variables y sus valores segmentados de acuerdo con la clasificación realizada por los autores.

Tabla 4

Variables de la Calidad del ACB Convencional (IC-ACB-C).

Variable	Valores
17 ¿Contiene la exposición de motivos de la proposición legislativa una sección específica en la que se presenta el ACB?	0= No contiene 1 = Sí contiene
18 ¿Describe la exposición de motivos de la proposición legislativa el problema que su autor pretende enfrentar?	0 = No hay descripción 1 = La descripción es deficiente 2 = Sí hay descripción y es adecuada
19 ¿Se define en la exposición de motivos con claridad el objetivo de la proposición legislativa?	0 = No hay una definición 1 = Sí hay definición, pero es inadecuada 2 = Sí hay una definición adecuada
20 ¿Se muestra en la exposición de motivos de la proposición legislativa que se ha analizado la necesidad de dar una ley para alcanzar el objetivo propuesto?	0 = No lo muestra 1 = Sí lo muestra, pero es insuficiente 2 = Sí lo muestra adecuadamente
21 ¿Muestra la exposición de motivos de la proposición legislativa que su autor ha estudiado y descartado otras alternativas regulatorias?	0 = No lo muestra 1 = Sí lo muestra, pero en forma inadecuada 2 = Sí lo muestra adecuadamente
22 ¿Se identifica en la exposición de motivos de la proposición legislativa a los beneficiarios?	0 = No se identifican o solo muy genéricamente 1 = Sí se identifican, pero inadecuadamente 2 = Sí se identifican adecuadamente
23 ¿Se cuantifica en la exposición de motivos a los beneficiarios de la proposición legislativa?	0 = No se cuantifican 1 = Sí se cuantifican, pero inadecuadamente 2 = Sí se cuantifican adecuadamente
24 ¿Se identifica en la exposición de motivos a los perjudicados de la proposición legislativa?	0 = No se identifican 1 = Sí se identifican, pero inadecuadamente 2 = Sí se identifican adecuadamente
25 ¿Se cuantifica en la exposición de motivos a los perjudicados de la proposición legislativa?	0 = No se cuantifican 1 = Sí se cuantifican, pero inadecuadamente 2 = Sí se cuantifican adecuadamente
26 ¿Muestra la exposición de motivos que se ha analizado la relevancia de identificar a beneficiarios y/o perjudicados indirectos de la proposición legislativa?	0 = No lo muestra 1 = Sí lo muestra, pero inadecuadamente 2 = Sí lo muestra adecuadamente

Variable	Valores
27 ¿Señala la exposición de motivos, cuando corresponde, los beneficios indirectos que genera la proposición legislativa?	0 = No los señala 1 = Sí los señala
28 ¿Señala la exposición de motivos, cuando corresponde, los perjuicios indirectos que genera la proposición legislativa?	0 = No los señala 1 = Sí los señala
29 ¿Se monetizan en la proposición legislativa los beneficios directos?	0 = No se monetizan 1 = Sí se monetizan, pero inadecuadamente 2 = Sí se monetizan adecuadamente
30 ¿Se monetizan en la proposición legislativa los perjuicios directos?	0 = No se monetizan 1 = Sí se monetizan, pero inadecuadamente 2 = Sí se monetizan adecuadamente
31 ¿Se monetizan en la proposición legislativa los efectos indirectos?	0 = No se monetizan 1 = Sí se monetizan, pero inadecuadamente 2 = Sí se monetizan adecuadamente
32 ¿Se indican en la exposición de motivos, cuando corresponde, los costos que la proposición genera para el Estado?	0 = No Los valoriza 1 = Sí los valoriza, pero inadecuadamente 2 = Sí los valoriza adecuadamente
33 ¿Muestra la exposición de motivos que se ha hecho uso de información bibliográfica, estadística, proveniente de consultores y/o expertos y/o se ha coordinado con el sector especializado en la materia, incluyendo la consulta sobre la conformidad de la proposición legislativa?	0 = No se muestra 1 = Sí se muestra, pero en forma inadecuada 2 = Sí se muestra adecuadamente
34 ¿Presenta la proposición legislativa el análisis de la relación entre beneficios y costos, identificando los efectos redistributivos como resultado de los pasos anteriores?	0 = No hay un análisis de efectos 1 = Sí hay un análisis de efectos, pero en forma inadecuada 2 = Sí hay un análisis adecuadamente

Fuente: elaboración propia con datos de (Arias-Schreiber Barba, Peña Jumpa, & Valdivieso Serrano, 2019).

Tabla 5

Variables de la calidad del ACB Legislativo (IC-ACB-L).

Variable	Valores
17 ¿Contiene la exposición de motivos de la proposición legislativa una sección específica en la que se presenta el ACB?	0= No contiene 1 = Sí contiene
18 ¿Describe la exposición de motivos de la proposición legislativa el problema que su autor pretende enfrentar?	0 = No hay descripción 1 = La descripción es deficiente 2 = Sí hay descripción y es adecuada
19 ¿Se define en la exposición de motivos con claridad el objetivo de la proposición legislativa?	0 = No hay una definición 1 = Sí hay definición, pero es inadecuada 2 = Sí hay una definición adecuada
22 ¿Se identifica en la exposición de motivos de la proposición legislativa a los beneficiarios?	0 = No se identifican o solo muy genéricamente 1 = Sí se identifican, pero inadecuadamente 2 = Sí se identifican adecuadamente

Variable	Valores
24 ¿Se identifica en la exposición de motivos a los perjudicados de la proposición legislativa?	0 = No se identifican 1 = Sí se identifican, pero inadecuadamente 2 = Sí se identifican adecuadamente

Fuente: elaboración propia con datos de (Arias-Schreiber Barba, Peña Jumba, & Valdivieso Serrano, 2019).

Tabla 6

Variables de la Viabilidad Legislativa sin ACB.

Variable	Valores
35 Aun cuando en la exposición de motivos de la proposición legislativa no se haya llegado a establecer el resultado del ACB como consecuencia de los pasos metodológicos correspondientes (ver la pregunta previa; es decir, la 34), ¿presenta, sin embargo, un análisis equilibrado entre beneficiarios y perjudicados?	0 = No presenta un análisis balanceado 1 = Sí presenta un análisis balanceado, pero en forma inadecuada 2 = Sí presenta un análisis balanceado

Fuente: elaboración propia con datos de (Arias-Schreiber Barba, Peña Jumba, & Valdivieso Serrano, 2019).

Tabla 7

Variables Relacionadas a los Proponentes y las Iniciativas.

Variable	Valores
1 ¿Quién es el proponente?	0 = Congresista de la República 1 = Poder Ejecutivo 2 = Otra entidad o grupo que tiene iniciativa legislativa
2 ¿Qué tipo de proposición legislativa es?	1 = Proyecto de ley 2 = Proyecto de resolución legislativa 3 = Proyecto de ley de reforma constitucional
3 ¿A cuántas comisiones ha sido enviada la proposición legislativa?	0 = A ninguna comisión 1 = A una sola comisión 2 = A dos comisiones
4 ¿A qué comisión fue enviada la proposición legislativa?	Número de la(s) comisión(es) a la(s) que fue(ron) enviada(s)
5 ¿Es la comisión de carácter económico o no?	0 = No es de carácter económico 1 = Sí es de carácter económico
6 ¿Tiene la proposición legislativa incidencia económica directa?	0 = No tiene incidencia económica directa 1 = Sí tiene incidencia económica directa
7 ¿Cuál es el nombre del congresista proponente?	Nombre del congresista proponente

Variable	Valores
8 ¿Cuántos congresistas firman la propuesta?	0 = Seis congresistas o menos, cuando el grupo parlamentario está conformado por menos de seis congresistas 1 = Más de seis congresistas 2 = Más de diez congresistas
9 ¿A qué grupo parlamentario pertenece el proponente principal?	El grupo parlamentario del proponente principal
10 ¿Es el autor de la proposición legislativa un congresista que forma parte de la mayoría parlamentaria y apoya normalmente las iniciativas del gobierno?	0 = Es de la mayoría parlamentaria 1 = Es de la oposición 2 = Ni lo uno ni lo otro
11 ¿Cuál es la profesión del proponente principal? ¿Ha recibido formación para entender la metodología del ACB?	0 = Es de la mayoría parlamentaria 1 = Ni lo uno ni lo otro
12 ¿Es el proponente un congresista reelecto o con experiencia parlamentaria previa?	0 = No tiene experiencia parlamentaria previa 1 = Sí tiene experiencia parlamentaria previa
13 ¿Tiene el congresista proponente experiencia administrativa pública o privada?	0 = No tiene experiencia administrativa 1 = Sí tiene experiencia administrativa
14 ¿El congresista autor de la proposición es hombre o mujer?	0 = Femenino 1 = Masculino
15 ¿Cuántas proposiciones legislativas ha presentado el congresista en el periodo anual bajo análisis?	Número de proyectos de ley presentado por el congresista

Fuente: elaboración propia con datos de (Arias-Schreiber Barba, Peña Jumpa, & Valdivieso Serrano, 2019).

En el artículo 75 del Reglamento del Congreso de la República (Congreso de la República, 2021c) y en la publicación de Proyectos de ley archivados (Congreso de la República, 2020b) señalan que, una proposición parlamentaria para ser viable debe cumplir con los requisitos:

- Relacionados al motivo de su fundamento (debe ser fundamentado por medio de especialistas en la materia tomada),
- Su efecto (acerca de las mejoras de su aplicación en la sociedad),

- La inclusión del Análisis Costo Beneficio (ACB), si fuera necesario, y
- La inclusión de un comentario acerca de su incidencia ambiental (si lo amerita).

En el Capítulo IX del Manual de la Técnica Legislativa referente a la necesidad y viabilidad de la ley (Congreso de la República, 2021b) también lista ciertos requisitos que, a continuación, se mencionan:

- El análisis de la información especializada.
- El estudio del marco normativo referente a la materia.
- La relación con las políticas públicas.
- El estudio de los antecedentes parlamentarios relacionados.
- El análisis de Impactos acerca de la propuesta parlamentaria.
- La evaluación institucional sobre la propuesta parlamentaria.
- La determinación de los destinatarios de la propuesta parlamentaria y de los actores que la impulsan.
- La obtención de la opinión de los destinatarios de la propuesta parlamentaria.

- El estudio acerca de la constitucionalidad de la ley.
- La determinación de materia que se pretende regular se relaciona con la necesidad de la aprobación de una ley.
- La forma de la propuesta parlamentaria.
- La consulta con el decisor político.

En lo referente, a las causas externas, se pueden considerar los factores PESTEL (políticos, económicos, sociales, tecnológicos, ecológicos y legales).

Como atributos en el presente estudio, se tomaron en cuenta las 15 variables definidas por (Arias-Schreiber Barba, Peña Jumpa, & Valdivieso Serrano, 2019), específicamente del grupo de variables relacionadas a los proponentes y las iniciativas, y como etiqueta, se estimó la condición de una propuesta aprobada como ley.

El Análisis Costo Beneficio (ACB). Haro Carranza (2012) menciona que el ACB es un procedimiento que permite plantear mejores decisiones, ya sea de origen personal, empresarial, social, cultural o jurídica, facilita conocer lo que se puede ganar y lo que se puede perder frente a una decisión. Algunos autores consideran que permite identificar el “costo de oportunidad”; es decir, lo que dejamos de ganar por elegir otra opción. El ACB se encuentra regulado en el artículo 75 del Reglamento del Congreso de la República y en el artículo 3 del Reglamento de la Ley Marco para la Producción y Sistematización Legislativa.

Congreso de la República (2017) señala que la metodología del ACB propone criterios para una elección en condiciones de escasez de recursos: ejecutar una acción si es que el beneficio adicional supera al costo adicional, es decir, se ejecuta si las ganancias individuales o sociales utilizan menores recursos, en términos económicos, tomar una decisión racional y eficiente. A veces, no siempre es deseable una decisión eficiente con un menor uso de recursos económicos, también se busca que los beneficios que genera sean distribuidos equitativamente, en estos casos, el principio del costo beneficio permite elegir entre la opción que optimiza la cantidad de recursos, de manera que se asegure el criterio de equidad en la sociedad y el criterio de eficiencia económica.

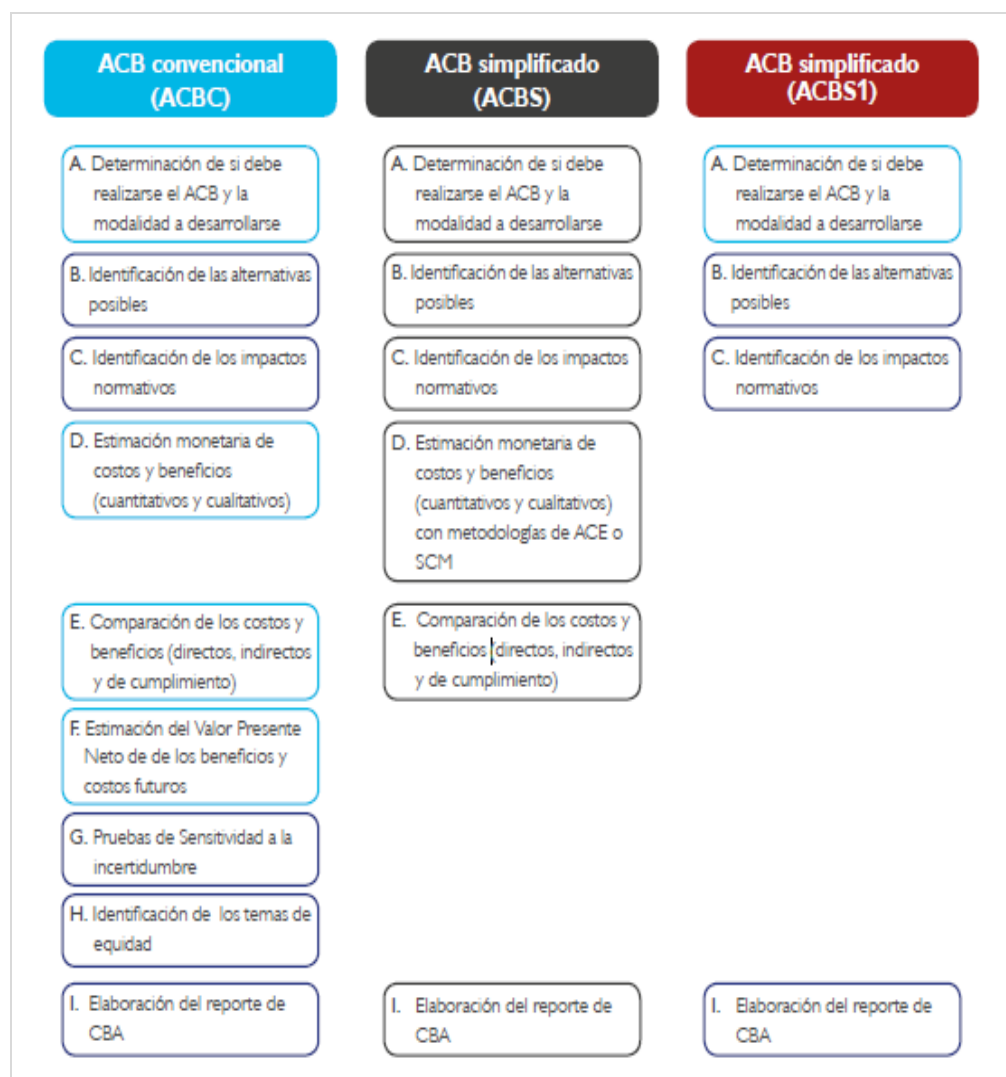
El concepto de racionalidad económica del ACB está basado en la utilización como guía para las tomas de decisiones gubernamentales, exigiendo hacer el máximo esfuerzo para un uso eficiente de los recursos públicos. El ACB se clasifica según la necesidad de los objetivos, entre ellos (Congreso de la República, 2017) menciona:

El ACB Convencional (ACB Completo - ACBC). Usado para la evaluación de un número reducido de propuestas, con un alto impacto en sectores económicos o en objetivo de lineamientos políticos económicos, sociales o ambientales.

El ACB Simplificado (ACBS). Similar al ACBC, pero prioriza la aplicación del Análisis Costo Efectividad (ACE) y el Modelo de Costos Estándar (Standard Cost Model - SCM).

El ACBS1. Considera la experiencia de la aplicación del ACB al trabajo legislativo peruano. Es conocida también como ACB Legislativo (ACBL), que fue desarrollado por el equipo de técnicos del Servicio Parlamentario del Congreso peruano.

Figura 20
Clasificación del ACB



Fuente: (Congreso de la República, 2017)

El Flujo Legislativo. Según la publicación del (Congreso de la República, 2016) el poder legislativo tiene como tarea principal la creación de las leyes. La Constitución Política del Perú menciona las vías para llegar a la formulación y promulgación de las leyes. A continuación se hace un breve resumen del flujo legislativo:

Oficina de Trámite Documentario. Esta oficina se encarga de recepcionar, registrar y numerar el proyecto.

Oficialía Mayor. El Oficial Mayor se encarga de dar cuenta al Consejo Directivo, y de remitir el decreto de envío.

Comisiones. La integran congresistas de diferentes bancadas, se encarga de realizar un exhaustivo estudio de los proyectos de ley propuesto, su función es establecer puntos de vista y encontrar el consenso entre los diferentes grupos parlamentarios. Las comisiones emiten sus dictámenes en 30 días útiles, desde la fecha de presentación. El consenso de todos sus integrantes termina un veredicto, ya sea por unanimidad o mayoría y minoría. Los dictámenes pueden ser: dictamen favorable, dictamen desfavorable, rechazado de plano.

Consejo Directivo. Este órgano con apoyo del Oficial Mayor, del Director General Parlamentario y del relator, su tarea es ordenar los proyectos de ley y programar la agenda parlamentaria para su debate en el pleno:

Pleno. Su función es debatir y puede aprobar la ley o rechazarla, enviándola al archivo. El debate es registrado minuciosamente, detallando todas sus incidencias y acuerdos.

Oficina de Relatoría y Agenda. Se encarga de elaborar la autógrafa de la ley y la prepara para su remisión al Poder Ejecutivo.

Oficialía Mayor. Este órgano revisa y certifica la autógrafa de la ley, y ordena su envío al Presidente de la República para su promulgación en el término de 15 días útiles.

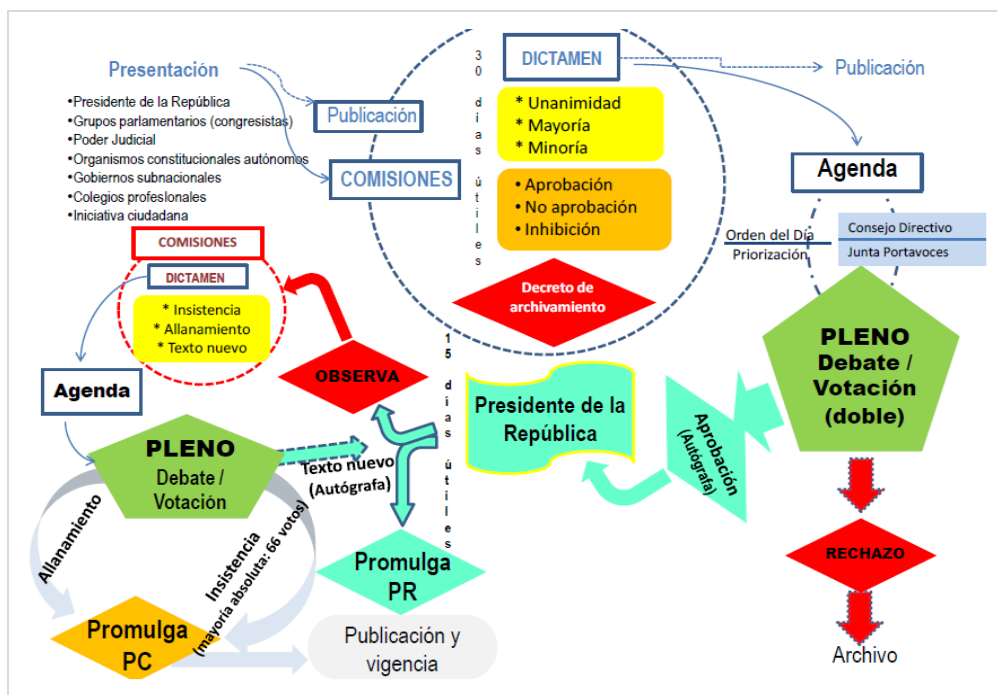
Poder Ejecutivo. Si no existen observaciones, el Presidente de la República promulga la ley, y dispone su publicación en el diario El Peruano. Si el Presidente de la República encuentra observaciones en la ley o una parte de la proposición aprobada, procede a presentar al Congreso en el mencionado término (15 días útiles). Si se vence el plazo, el Presidente de la República no llega a promulgar la proposición de ley, entonces, el Presidente del Congreso o el presidente de la Comisión Permanente, ejecuta la promulgación del proyecto de ley.

Promulgación y Publicación. Las leyes promulgadas son publicadas en la sección "Normas Legales" del diario oficial "El Peruano".

Vigencia de la Ley. La ley es obligatoria desde el día siguiente de su publicación en el diario oficial, salvo disposición

excepcional de la misma ley que indique un mayor período en todo o una parte de la ley.

Figura 21
El flujo legislativo



Fuente: (Delgado Guembes, 2020)

Las Proposiciones Parlamentarias. Según los artículos 65 y 66 del Reglamento del Congreso de la República (Congreso de la República, 2021c), las proposiciones son instrumentos procesales parlamentarios junto con los dictámenes e informes de las Comisiones. Se consideran proposiciones parlamentarias a las proposiciones de ley (también llamados proyectos de ley), las proposiciones de resolución legislativa, mociones de orden del día y los pedidos de información.

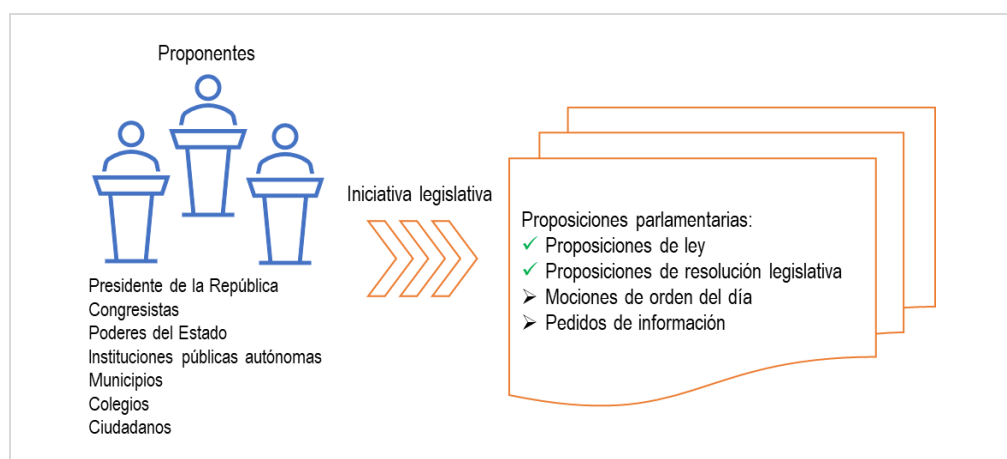
En la publicación de (Nicolini Del Castillo, 2013), se cita el concepto de iniciativa legislativa como el derecho a presentar

proposiciones de ley, no sólo están facultados el Presidente de la República y los congresistas, sino también, los proponentes pueden ser los otros poderes del Estado, las instituciones públicas autónomas, los municipios, los colegios e incluye a los ciudadanos. Además, Nicolini del Castillo (2013) menciona que cuando se ejerce el derecho de la iniciativa legislativa, se inicia el procedimiento legislativo para alcanzar la aprobación de una ley o resolución legislativa.

Para el presente estudio se consideró las proposiciones parlamentarias que recorren el flujo legislativo como son los proyectos de ley y las resoluciones legislativas.

Figura 22

Las proposiciones parlamentarias



Fuente: elaboración propia con datos de (Congreso de la República, 2021c)

1.3. Definición de Términos Básicos

Big Data

The Royal Society (2017) lo describe como datos grandes y heterogéneos recopilados sin un diseño experimental estricto, su uso es más común debido a la proliferación del almacenamiento digital, por ejemplo, a través de teléfonos móviles, y el mayor grado de interconexión entre dispositivos.

Comisiones

Nicolini del Castillo (2013) menciona que las comisiones a los grupos de trabajo especializados que tienen como función la supervisión del seguimiento de la estructura del Estado, así como, la realización del estudio y dictamen de los proyectos de ley y de la absolución de consultas de acuerdo con su especialidad o materia.

Curva ROC (Receiver Operating Characteristic)

Según (Management Solutions, 2018) es una curva empleada para analizar la capacidad predictiva de un modelo de salida binaria. Representa la relación entre el error de tipo 1 (clasificar incorrectamente sucesos adversos) y el error tipo 2 (clasificar incorrectamente sucesos favorables).

Datos de Entrenamiento

The Royal Society (2017) define que son datos que se pueden usar para entrenar sistemas de aprendizaje automático, que ya han sido etiquetados o categorizados en uno o más grupos.

Datos de Prueba

The Royal Society (2017) denomina así, a los datos que se utilizan para probar el funcionamiento de un sistema de aprendizaje automático o verificar sus resultados.

Entropía Cruzada

Google Developers (2019) lo define como una generalización de pérdida logística en eventos referentes a clasificación de clases múltiples. La entropía cruzada realiza la cuantificación de la diferencia entre dos distribuciones de probabilidad.

Función de Activación

Analytics Steps (2019) señala que la función de activación define el nodo de la salida del nodo de las entradas. Desactivan las neuronas o las activan para alcanzar una salida deseada.

Jurimetría

Bianchi (2018) menciona que el término jurimetría apareció en 1949, cuando el abogado Lee Loevinger lo presentó como la utilización de computadoras en el rubro del derecho y a su aplicación en el estudio de la legislación de libre competencia de Estados Unidos.

Máquina de Vectores de Soporte (SVM)

Goodfellow, Bengio, & Courville (2016) lo definen como una técnica de clasificación y regresión, a diferencia de la regresión logística, la máquina

de vectores de soporte no proporciona probabilidades. SVM es ideal para analizar datos con un gran número de campos de predictores.

Minería de Datos (Data mining)

Según (Shmueli, Bruce, Stephens, & Patel, 2017) lo describen como métodos analíticos de negocios que van más allá de los recuentos, técnicas descriptivas, reportes o métodos basados en reglas de negocio, mencionan también que en la era del big data se ha acelerado su uso y que los métodos de minería de datos tienen la capacidad para hacer frente a la extracción de enormes cantidades de datos.

Modelo

Google Developers (2019) menciona que es la representación de lo que un algoritmo de inteligencia artificial aprendió en base a datos de entrenamiento.

Perceptrón

Véliz Capuñay (2020) cita que a una simple de red neuronal artificial lo conforma el perceptrón, red que tiene una capa de entrada y una capa de salida con un único nodo o neurona.

Período Anual de Sesiones

Según (Congreso de la República, 2021c) este período está comprendido desde el 27 de julio de un determinado año hasta el 26 de julio del siguiente año.

Período Ordinario de Sesiones

Según (Congreso de la República, 2021c) existen dos períodos ordinarios de sesiones o legislaturas dentro de un periodo anual de sesiones: el primero es entre el 27 de julio al 15 de diciembre y el segundo es entre el 01 de marzo al 15 de junio.

Período Parlamentario

En el reglamento del (Congreso de la República, 2021c) el período parlamentario comienza desde la instalación de un nuevo Congreso, hasta la instalación del Congreso elegido en el siguiente proceso electoral. El período parlamentario comprende una duración de cinco años.

Procesamiento del Lenguaje Natural (NLP)

Según (Goodfellow, Bengio, & Courville, 2016) lo definen como el uso de lenguajes humanos, como el inglés o el español, por una computadora. El procesamiento del lenguaje natural contiene aplicaciones como la traducción automática y simultánea. Mayormente las aplicaciones de PNL se basan en modelos de lenguaje que definen una distribución de probabilidad en secuencias de palabras, caracteres o bytes en un lenguaje natural.

Random Forest (bosque aleatorio)

Según (Analytics Steps, 2019) lo define como métodos de conjunto que combinan diferentes algoritmos o similares que se usan en las tareas de clasificación. Esta técnica se aplica bajo un algoritmo supervisado que se

puede usar tanto para clasificaciones como para regresiones, siendo uno de los más utilizados en tareas de clasificación.

Redes Convolucionales (CNN)

Según el enunciado de (Goodfellow, Bengio, & Courville, 2016) son un tipo especializado de red neuronal para procesar datos que tiene una topología conocida parecido a una cuadrícula. El nombre "red neuronal convolucional" señala que la red usa una operación matemática llamada convolución (operación lineal). Las redes convolucionales son redes neuronales que utilizan la convolución en lugar de la multiplicación de matrices al menos en una de sus capas. Las redes convolucionales pueden escalar fácilmente a imágenes con gran ancho y alto.

Redes Neuronales Recurrentes (RNN)

Goodfellow, Bengio, & Courville (2016) refieren que son una familia de redes neuronales que procesan datos secuenciales. Así como una red convolucional es una red neuronal especializada para procesar una cuadrícula de valores x , una red neuronal recurrente es una red neuronal especializada para procesar una secuencia de valores $x(1), \dots, x(i)$. La mayoría de las redes recurrentes también pueden procesar secuencias de longitud variable. Las redes neuronales recurrentes pueden escalar imágenes a secuencias mucho más largas.

Regresión de Vectores de Soporte (SVR)

Awad & Khanna (2015) lo definen como una herramienta eficaz en la estimación de funciones de valor real. SVR se entrena aplicando una función

de pérdida simétrica, que penaliza las estimaciones erróneas altas y bajas. Usando el enfoque de Vapnik, un área similar a un tubo flexible de radio mínimo se forma simétricamente alrededor de la función estimada, de esta manera los valores absolutos de errores menores en un cierto umbral se ignoran tanto por encima como por debajo de la estimación.

Sesgo (Bias)

Berzal (2018) explica que un sesgo se encarga de medir la desviación de las predicciones del modelo y los valores correctos que se desea predecir.

Softmax

Analytics Steps (2019) refiere que es una función de activación que se usa en la última capa, similar a una función sigmoidea (enfoque probabilístico para la toma de decisiones con valores que oscilan entre 0 a 1), el softmax da valor a la variable de entrada de acuerdo con su peso, la suma de estos pesos es igual a uno.

Tangente Hiperbólica

Analytics Steps (2019) hace mención que es una función de activación regularmente mejor que la función sigmoidea, también se aplica para predecir o diferenciar dos clases, pero con la diferencia de asignar a las entradas negativas solamente cantidades negativas y oscila entre -1 y 1.

CAPÍTULO II

PROPUESTA DE LA ELABORACIÓN DEL MODELO PARA DEMOSTRAR SU CAPACIDAD PREDICTIVA EN LA VIABILIDAD LEGISLATIVA

La propuesta describe el desarrollo de un modelo predictivo mediante el uso de las ANN multicapa, su finalidad es demostrar que las ANN son técnicas capaces de obtener un alto grado de predicción, no se ha realizado su implementación ya que el estudio se orienta a un nivel predictivo y no aplicativo para efectuar mejoras al problema situacional. El desarrollo del modelo esta soportado por la metodología CRISP-DM. IBM (2020) cita el CRISP-DM (Cross-Industry Standard Process for Data Mining) como un método usado para orientar trabajos de minería de datos, siendo personalizable e integrado por fases principales. De las seis fases integradas en la metodología, sólo se consideraron cuatro fases. A continuación, se describen las fases desarrolladas.

2.1. Fase 2 - Comprensión de los Datos

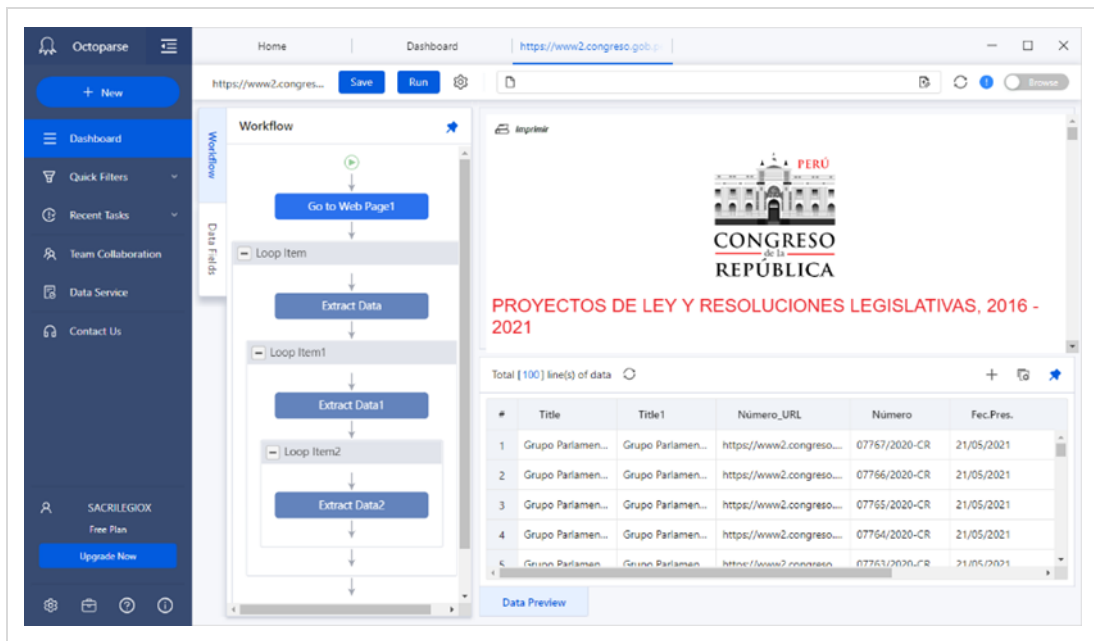
F2.1 - Recolección de Datos Iniciales

Los datos son de acceso público, por lo tanto, no existió restricción para su obtención. Los datos se localizan en el Portal del Congreso en la sección de proyectos de ley. Al encontrarse en formato HTML, los datos no se encuentran disponibles para una fácil extracción, por ende, fue necesario usar herramientas especiales para su obtención, se requirió el uso de un programa scraping como el Octoparse versión 8 de prueba gratuita. Se generó un archivo Excel para almacenar los datos, los atributos fueron facilitados por la base teórica del estudio realizado por (Arias-Schreiber

Barba, Peña Jumpa, & Valdivieso Serrano, 2019). La información almacenada comprendió los periodos parlamentarios a partir de 1995.

Figura 23

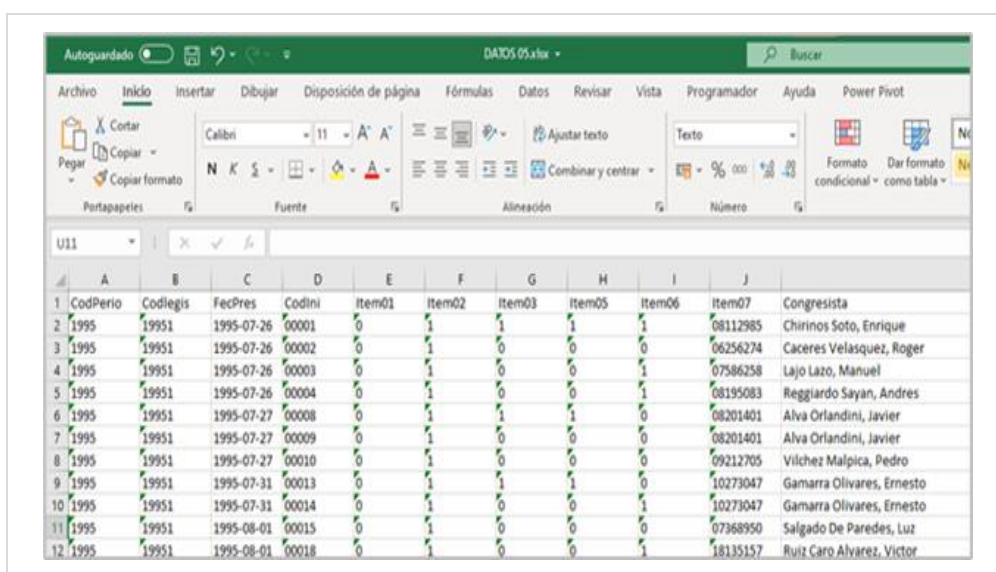
Uso de la herramienta Octoparse versión 8



Fuente: elaboración propia.

Figura 24

Exportación de datos al archivo Excel



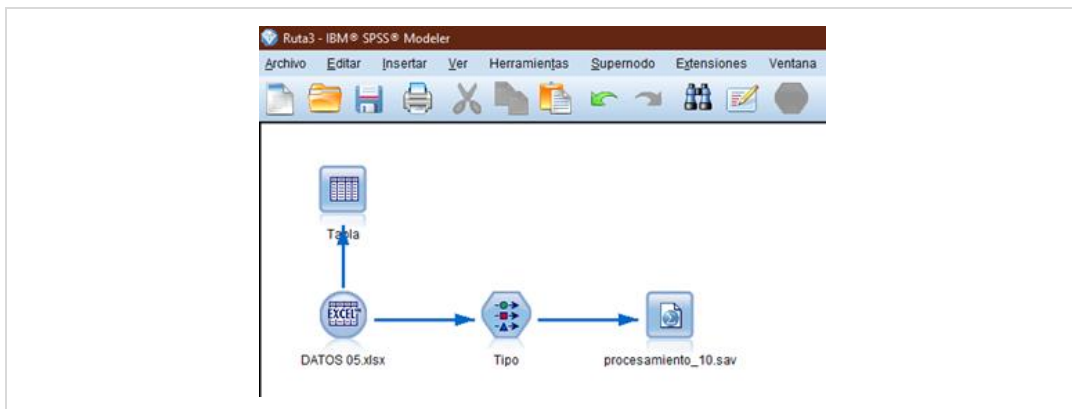
Fuente: elaboración propia.

F2.2 Descripción de los Datos

Haciendo uso de la herramienta SPSS Modeler versión 18 se importó los datos del archivo Excel al archivo de extensión *.sav. A las variables cualitativas se les asignó valores ordinales del tipo numérico y se agregó etiquetas a cada valor. A las variables cuantitativas se les recodificó y se les asignó valores ordinales del tipo numérico y se agregó etiquetas a cada valor.

Figura 25

Usando la herramienta SPSS Modeler versión 18 para exportar el archivo Excel a formato *.sav

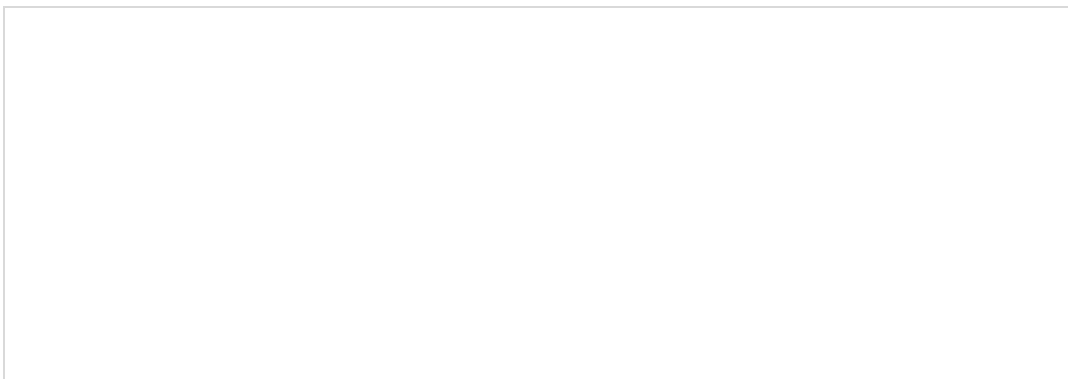


Fuente: elaboración propia.

Figura 26

Verificación de tipos de variables y asignación de etiquetas a sus valores en SPSS Modeler

Campo	Medida	Valores	No se enc...	Comprobar	Rol
CodPerio	Nominal	"1995","20...		Ninguno	Entrada
Codlegis	Nominal	"19951","1...		Ninguno	Entrada
FecPres	Sin tipo			Ninguno	Ninguno
Codini	Sin tipo			Ninguno	Ninguno
Item01	Marca	"070"		Ninguno	Entrada
Item02	Nominal	"1","2","3"		Ninguno	Entrada
Item03	Nominal	"0","1","2"		Ninguno	Entrada
Item05	Marca	"170"		Ninguno	Entrada
Item06	Marca	"170"		Ninguno	Entrada



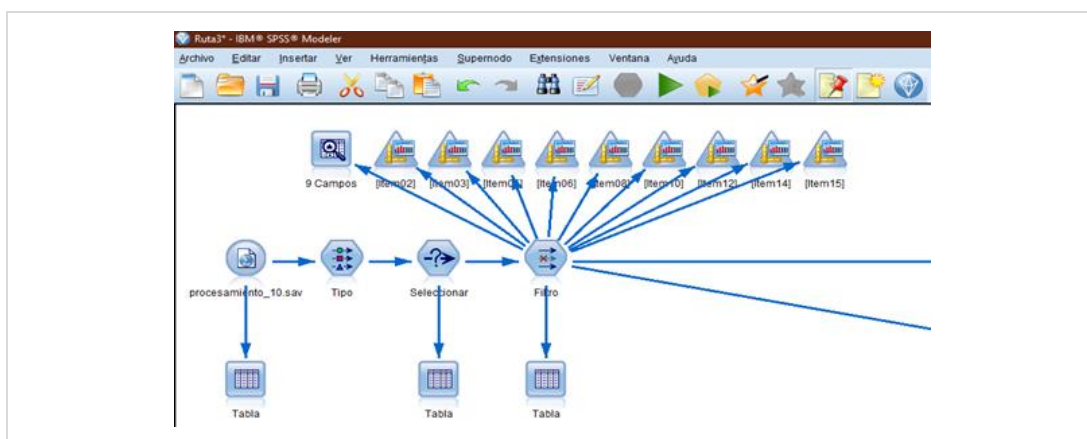
Fuente: elaboración propia.

F2.3 - Exploración de Datos

Mediante la herramienta SPSS Modeler versión 18 de prueba gratuita se importaron los registros de los archivos Excel para proceder con la generación de tablas y gráficos descriptivos. Los resultados se muestran en el CAPÍTULO IV: RESULTADOS, en el subtítulo de Presentación de los Resultados Descriptivos.

Figura 27

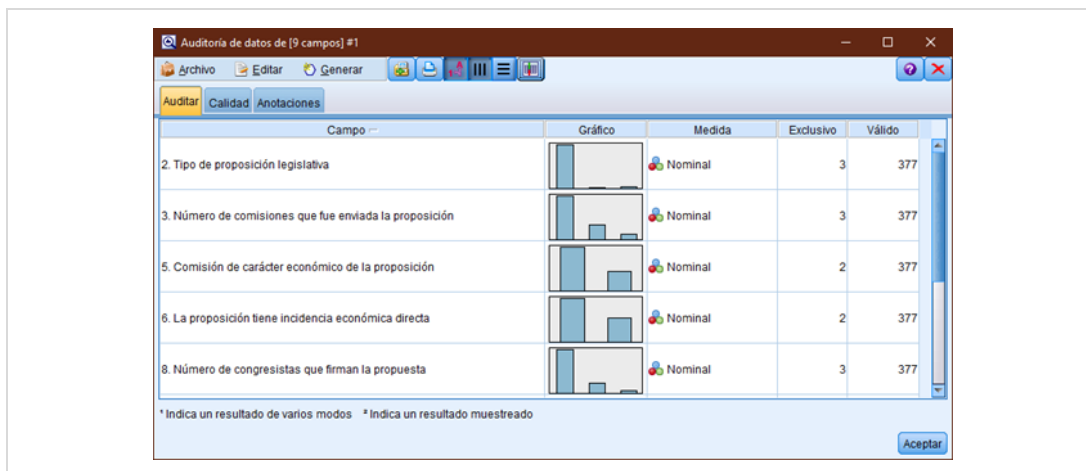
Exploración de datos mediante SPSS Modeler



Fuente: elaboración propia.

Figura 28

Generación de resultados descriptivos con SPSS Modeler



Fuente: elaboración propia.

F2.4 - Verificación de la Calidad de los Datos

De los quince atributos relacionados a los proponentes y a los proyectos de ley, no se consideraron seis atributos, ya que sus valores eran nulos y de poca significancia para el estudio.

2.2. Fase 3 - Preparación de Datos

F3.1 - Selección de Datos

Se seleccionaron los proyectos de ley cuyas fechas de presentación fueron correspondientes al periodo anual de sesiones 2019 perteneciente al periodo parlamentario 2016.

F3.2 - Limpieza de Datos

Al haberse realizado el proceso de calidad en la fase anterior no se encontró registros para su limpieza de la base de datos importada al programa SPSS.

F3.3 - Estructuración de Datos

Se agregó el atributo del tipo cualitativo “viabilidad” que representará el estado final del proyecto de ley, se les asignó valores ordinales del tipo numérico (0 y 1) y se agregó etiquetas a cada valor (viable y no viable).

F3.4 - Integración de Datos

No fue necesario realizar la integración de datos debido a que la información se importó de un mismo origen.

F3.5 - Formateo de Datos

No fue necesario formatear los valores debido a que el programa Octoparse versión 8 extrajo los datos limpios.

2.3. Fase 4 - Modelamiento

F4.1 - Selección de la Técnica de Modelado

Para proceder con la selección de la técnica del modelo se realizó una comparativa de precisión entre las técnicas más usadas en los estudios presentados en los antecedentes nacionales e internacionales (Ver Tabla 8).

Tabla 8

Comparativa de precisión entre las técnicas más usadas en los estudios de los antecedentes nacionales e internacionales

Referencia	Descripción	Registros	Técnica	Precisión
(Aguilar Vilca & Camargo)	Predicción de la deserción de clientes de microcréditos de	24,420 clientes	ANN RamdonForest	97.38%

Referencia	Descripción	Registros	Técnica	Precisión
Ramos, 2021)	bancos		SVM	
(Khashman & Khashman, 2016)	Predicción de los votos en el Congreso Estadounidense	435 legisladores	ANN SVM	96.33% 95.29%
(Zolghadr, Niaki, & Akhavan Niaki, 2016))	Predicción de las elecciones presidenciales de los Estados Unidos	791 votos	SVR ANN	93.14% 92%
(Colliri & Zhao, 2019))	Predicción de las condenas por corrupción de los congresistas brasileños	3,407 sesiones de votación de proyectos de ley	ANN	90%
(Kornilova, Argyle, & Eidelman, 2018)	Predicción de los votos de los congresistas	20,000 proyectos de ley	CNN	86.21%
(Zevallos Salazar, 2017)	Predicción del rendimiento académico de los estudiantes de tres colegios del nivel primario del distrito de Ventanilla	1,073 estudiantes	ANN	84%
(Kaur & Bozié, 2019)	Predicción de las sentencias del Tribunal Europeo de Derechos Humanos	11,532 sentencias	CNN SVM	82% 75%

Fuente: Elaboración propia

En la Tabla 8, se puede observar que, las técnicas más usadas son las redes neuronales artificiales (ANN), seguidas de las máquinas de soporte vectorial (SVM), las redes neuronales convolucionales (CNN), los bosques aleatorios (Random Forest), y por último, la regresión de vectores de soporte (SVR). Las técnicas que muestran mejores resultados son las ANN con un 97.38%, Random Forest con un 97.38%, SVM presenta un 97.38%, SVR un 93.14% y las CNN obtuvieron una precisión del 86.21%. El porcentaje de acierto más alto obtenido es el propuesto por (Aguilar Vilca & Camargo Ramos, 2021) quienes utilizaron una técnica de hibridación con las ANN,

SVM y Random Forest. Para culminar con la selección de la técnica del modelado, se puede concluir que, las ANN han demostrado ser las más usadas y de mayor precisión en las investigaciones presentadas en los antecedentes nacionales e internacionales. Por lo tanto, la técnica seleccionada para el presente estudio se le atribuye a las ANN.

F4.2 - Generación del Plan de Prueba

No se requirió generar un plan de pruebas, sólo se optó por realizar una revisión de las opciones de la herramienta, específicamente en la opción de redes neuronales.

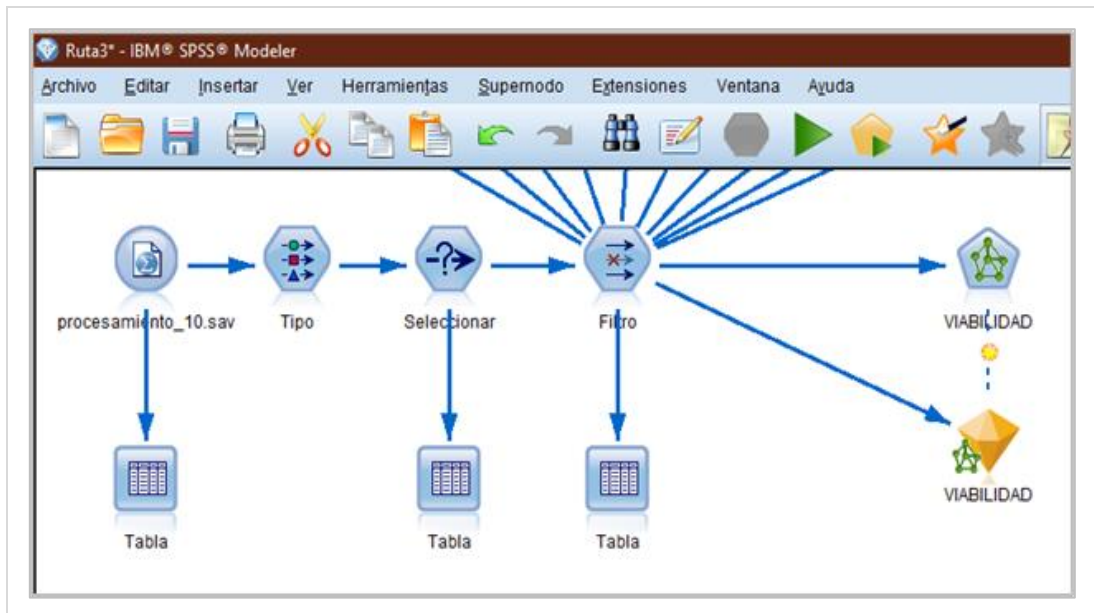
F4.3 - Construcción del Modelo

Para la construcción del modelo se consideró nueve atributos y una etiqueta mediante el uso de la herramienta SPSS Modeler y para obtener información más detallada del modelo se hizo uso de la herramienta estadística SPSS versión 25. La generación de unidades de las capas ocultas se dejó con la opción de “Selección automática de arquitectura” (de 1 a 50 capas ocultas) y referente a la asignación aleatoria de los casos para el entrenamiento y prueba se optó por los valores por defecto, 70% para casos de entrenamiento y 30% para casos de pruebas.

En la Figura 29, se muestra la construcción de red neuronal, se eligió la opción de modelado de red neuronal, procediendo con ingresar los nueve atributos y la etiqueta. En la Figura 30, se muestra la red neuronal generada.

Figura 29

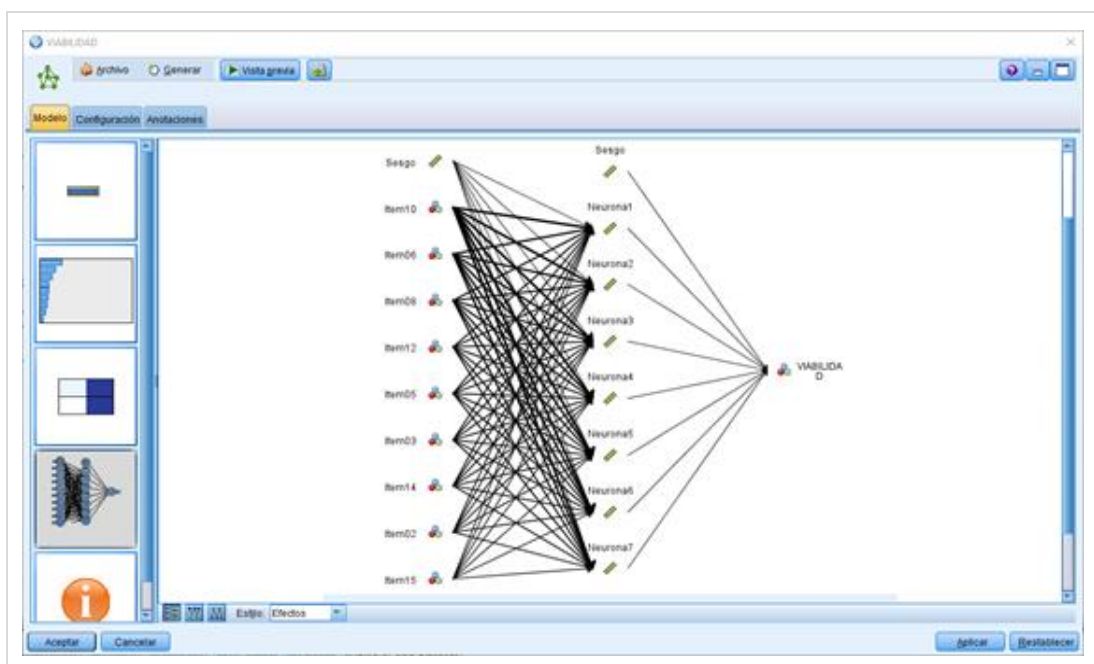
Proceso de ejecución de la técnica predictiva de redes neuronales artificiales (ANNs) mediante SPSS Modeler



Fuente: elaboración propia.

Figura 30

Generación de la red neuronal mediante SPSS Modeler



Fuente: elaboración propia.

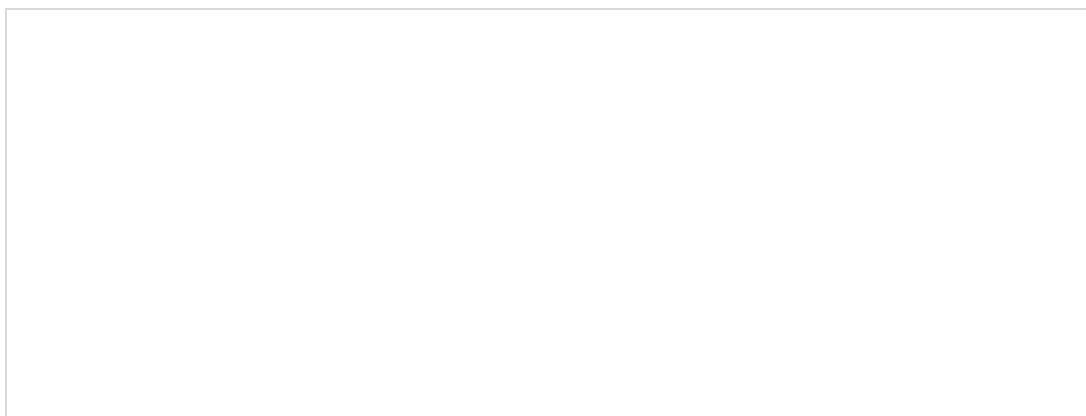
F4.4 - Evaluación del Modelo

Mediante el uso del SPSS Modeler, se ejecutaron tres técnicas algorítmicas: las redes bayesianas, el algoritmo KNN y las SVM (ver Figura 31); y mediante el SPSS Statistic, se ejecutaron la técnica de la red neuronal de función de base radial y la red neuronal de perceptrón multicapa para comparar su capacidad predictiva con otras técnicas.

A continuación, los resultados fueron los siguientes, las ANN de perceptrón multicapa alcanzaron un porcentaje de 96.1% (ver Tabla 10), las redes bayesianas alcanzaron un 96.02% de precisión, las redes neuronales de función de base radial un 93.7% (ver Tabla 9), el algoritmo KNN presentó un 90.45% y la SVM obtuvo una precisión del 89.39%. Según los resultados presentados, las ANN de perceptrón multicapa, las redes bayesianas y las ANN de función de base radial han alcanzado un alto porcentaje de precisión como se muestra en la Tabla 11.

Figura 31

Comparación de precisión de otras técnicas algorítmicas mediante SPSS Modeler





Fuente: elaboración propia.

Tabla 9

Precisión de la red neuronal de función de base radial mediante SPSS Statistic

Muestra	Observado	Pronosticado		
		0	1	Porcentaje correcto
Entrenamiento	0	0	8	0.0%
	1	0	258	100.0%
	Porcentaje global	0.0%	100.0%	97.0%
Pruebas	0	0	7	0.0%
	1	0	104	100.0%
	Porcentaje global	0.0%	100.0%	93.7%

Variable dependiente: Viabilidad

Fuente: elaboración propia.

Tabla 10

Precisión de la red neuronal de perceptrón multicapa mediante SPSS Statistic

Muestra	Observado	Pronosticado	
		0	1

		0	1	Porcentaje correcto
Entrenamiento	0	0	11	0,0%
	1	0	263	100,0%
	Porcentaje global	0,0%	100,0%	96,0%
Pruebas	0	0	4	0,0%
	1	0	99	100,0%
	Porcentaje global	0,0%	100,0%	96,1%

Variable dependiente: Viabilidad

Fuente: elaboración propia.

Tabla 11

Comparación de precisión del modelo ANN de perceptrón multicapa con otras técnicas algorítmicas

Modelo	Tiempo de ejecución	Atributos	Precisión
ANN de perceptrón multicapa	00:00:00.57	9	96.10%
Red bayesiana	00:00:19.00	9	96.02%
ANN de función de base radial	00:00:00.52	9	93.70%
Algoritmo KNN	00:00:20.00	9	90.45%
SVM	00:00:21.00	9	89.39%

Fuente: elaboración propia.

Por lo tanto, para el presente caso, las ANN de perceptrón multicapa han demostrado tener un alto porcentaje de precisión y un menor tiempo de ejecución (a excepción de la ANN de función de base radial) frente a la aplicación de otras técnicas algorítmicas.

2.4. Fase 5 - Evaluación

F5.1 - Evaluación de Resultados

El modelo ANN de perceptrón multicapa presentó un 96.1% de acierto de probabilidad de predicción de la viabilidad legislativa y un 3.9% de falencia en la que el modelo falla. A comparación de otras técnicas algorítmicas y de las investigaciones citadas en los Antecedentes de la Investigación en el CAPÍTULO I: MARCO TEÓRICO CONCEPTUAL, se ha demostrado que la aplicación de la ANN ha alcanzado una precisión alta, mayor a lo esperado. El detalle de los resultados se muestran en el CAPÍTULO IV: RESULTADOS, en el subtítulo de Presentación de los Resultados de la Predicción.

F5.2 - Proceso de Revisión

El proceso de la metodología CRISP-DM se adecúa a la necesidad de la investigación, no se requiere una revisión.

F5.3 - Determinación de Futuras Fases

Para el presente estudio no se consideró la fase comprensión del negocio ni la fase de implementación, esta última fase no se ejecutó dado que, el propósito de la investigación no se orienta a una investigación de nivel aplicativo.

CAPÍTULO III

METODOLOGÍA

3.1. Diseño Metodológico

3.1.1. *Tipo y Nivel de la Investigación*

Tipo de Investigación. La investigación se caracteriza por ser del tipo básico. Según (Bunge, 2001) menciona que cuando una investigación está orientada a alcanzar un nuevo conocimiento de una determinada realidad se denomina básica, mientras que la finalidad de una investigación aplicada es de “mejorar”. Según (Supo Condori & Zacarías Ventura, 2020) corresponde a una investigación básica o pura ya que se trata de un estudio transversal y no requiere de medidas repetitivas de monitoreo para evaluar la mejora que se desea obtener. Alzamora de los Godos Urcia & Calderón Saldaña (2010) manifiestan que la investigación básica está destinada a aportar un contenido organizado de conocimientos científicos y no produce necesariamente resultados de utilidad práctica inmediata. Kumar (2019), indaga que, una investigación del tipo pura o básica trata del desarrollo, examinación, verificación y perfección de los métodos, procedimientos o herramientas, mientras que la del tipo aplicada, usan estos métodos, técnicas o herramientas para ponerla en práctica y solucionar problemas.

Nivel de Investigación. La investigación se ubica en el nivel predictivo. Según (Supo Condori & Zacarías Ventura, 2020) explican que, en el nivel predictivo se crean modelos matemáticos para predecir la probabilidad de la ocurrencia de un fenómeno, hecho o acontecimiento. Para construir un modelo predictivo se deben conocer las causas y se representan como variables exógenas o predictivas y el valor a predecir se representa con la variable endógenas o variable a predecir. Hurtado de Barrera (2014) lo clasifica como nivel comprensivo, debido a que se trata de una investigación predictiva, como se muestra en la Tabla 12 y la Figura 32. En este tipo de estudio, el investigador observa un evento (variable) durante cierto tiempo, describiendo, analizando, buscando explicaciones y factores relacionados entre sí, logrando anticipar cuál será la tendencia futura de ese evento.

Tabla 12

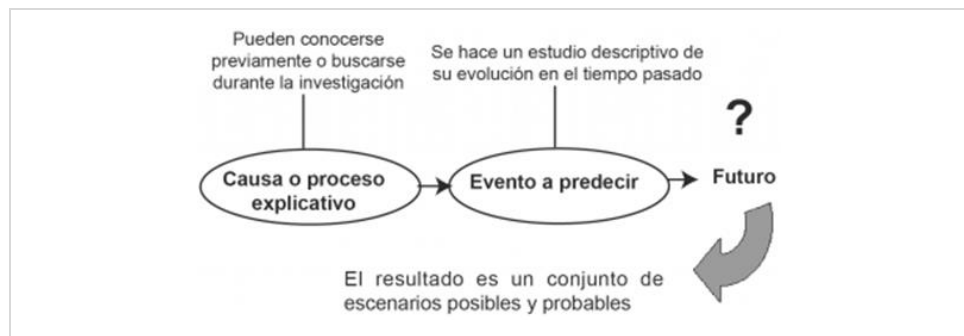
Niveles, objetivos y tipos en la investigación holística

Nivel	Objetivo	Tipo de investigación
Perceptual	Explorar	Investigación exploratoria
	Describir	Investigación descriptiva
Aprehensivo	Analizar	Investigación analítica
	Comparar	Investigación comparativa
	Explicar	Investigación explicativa
Comprensivo	Predecir	Investigación predictiva
	Proponer	Investigación proyectiva
	Modificar	Investigación interactiva
Integrativo	Confirmar	Investigación confirmatoria
	Evaluar	Investigación evaluativa

Fuente: elaboración propia con datos de (Hurtado de Barrera, 2014)

Figura 32

Esquema para una investigación predictiva



Fuente: (Hurtado de Barrera, 2014)

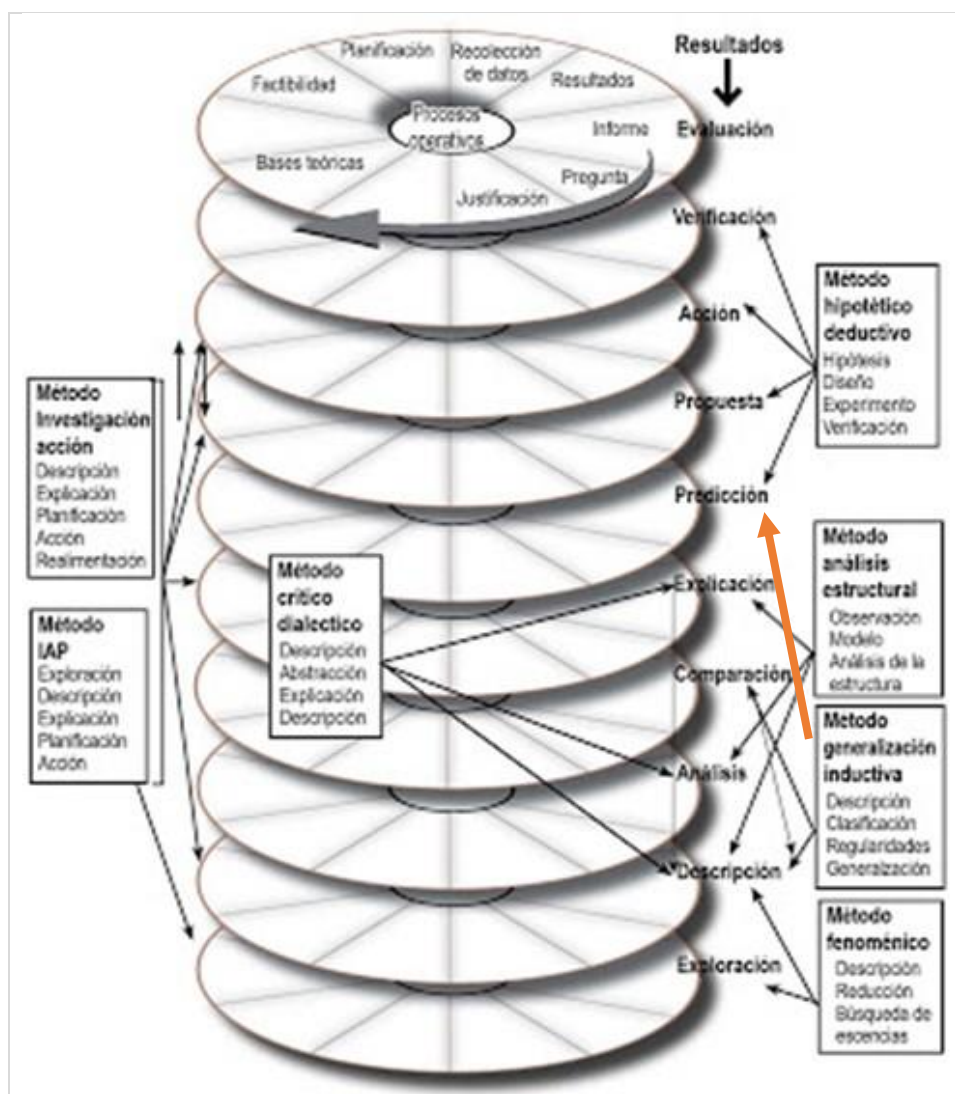
3.1.2. Métodos y Diseño de la Investigación

Métodos de Investigación. El método de investigación del presente estudio aplica el método inductivo, perteneciente al modelo epistémico del empirismo. Barrientos Carles (2020) manifiesta que el procesamiento de grandes cantidades de datos mediante algoritmos avanzados que las computadoras están “dominando”, los seres humanos no somos capaces de procesar. De esta manera, el autor concluye que existe una naturaleza puramente inductiva, también agrega que, los procesos computacionales elementales son de naturaleza deductiva ya que se basan en reglas de inferencia. En el informe de (Gershenson, 2020) considera que con el desarrollo de las computadoras podemos desarrollar nuevos métodos científicos, por ejemplo, el desarrollo de modelos por computadora o el análisis estadístico de datos masivos que permite estudiar modelos con millones de variables, contrastar de manera exhaustiva diferentes condiciones, y de manera sistemática, monitorear qué sucede

cuando varían ciertos parámetros. Hurtado de Barrera (2010) menciona que para los estudios proyectivos, predictivos e interactivos no es necesario formular hipótesis ya que no se van a someter a verificación, solamente se realiza la formulación de la hipótesis en los estudios confirmatorios y en los evaluativos.

Figura 33

La espiral holística y los métodos en investigación



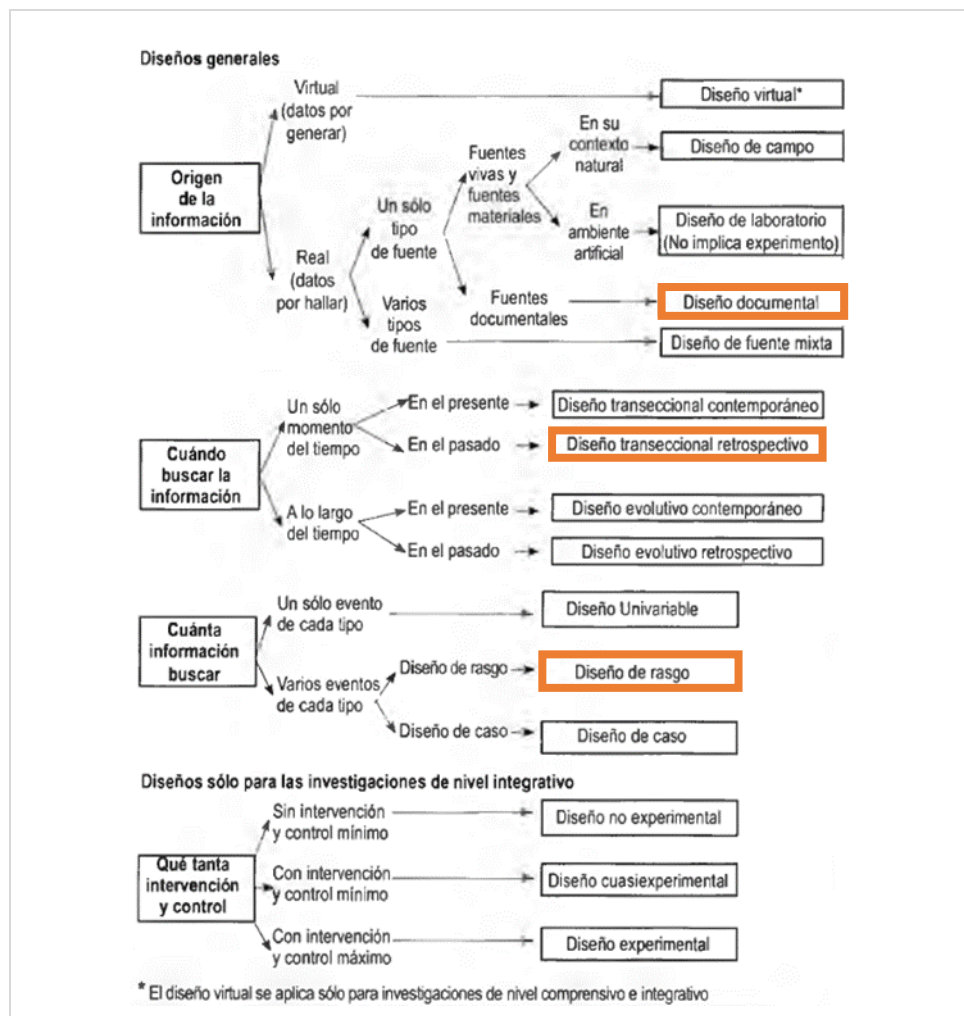
Fuente: (Hurtado de Barrera, 2010)

Diseño de la Investigación. Hurtado de Barrera (2014)

define que, el diseño de investigación trata acerca de las decisiones que se toman en cuanto al proceso de recolección de datos, la cual, hace referencia al dónde y cuándo se realiza la recopilación de la información, así como la cantidad a recopilar. De acuerdo con lo definido, la presente investigación su diseño se caracteriza por ser transeccional retrospectivo documental de rasgo. En la Figura 34, se muestra un esquema de los diseños de investigación.

Figura 34

Esquema para seleccionar los diseños de investigación



Fuente: (Hurtado de Barrera, 2010)

3.2. Técnicas de Recolección de Datos

3.2.1. Técnicas

La técnica aplicada en la presente investigación es la técnica de la documentación. Supo Condori & Zacarías Ventura (2020) explican que la documentación es una técnica de recolección de datos que consiste en copiar los datos de registros donde previamente han sido almacenados, la documentación no sólo implica revisar documentos físicos, sino también documentos digitales, el manejo de la información de los medios digitales son altamente eficientes, siendo recomendable el uso de las bases de datos en los estudios predictivos y aplicativos. En este caso, los proyectos de ley publicados en el portal del Congreso de la República provienen de una fuente de origen digital.

3.2.2. Instrumentos

Según (Supo Condori & Zacarías Ventura, 2020) mencionan que, los estudios basados en la documentación no cuentan con instrumentos de medición, únicamente cuentan con una ficha de recolección de datos, donde se requiere copiar o trasladar la información previamente registrada o que fue medida con otros fines distintos a la investigación en curso, algunos investigadores suelen confundir una historia clínica como un instrumento de medición o un archivo a partir del cual obtuvieron

la información, en este caso, la historia clínica o el archivo consultado no puede considerarse un instrumento de medición, la finalidad de ambos documentos solamente es registrar la información. Por lo tanto, las proposiciones parlamentarias publicadas en el Portal del Congreso no se pueden considerar un instrumento de medición sino una fuente digital de información similar a una ficha de datos física.

3.3. Técnicas Estadísticas para el Procesamiento de la Información

Hurtado de Barrera (2010) clasifica las técnicas de estadísticas en las técnicas univariadas para el análisis descriptivo/comparativo, correlacionales, análisis de tendencias, análisis multivariable/de interdependencia/dependencia. En la presente investigación se contó primeramente con el uso de las técnicas de análisis descriptivo como son las frecuencias y los porcentajes. En vista que, el presente estudio tiene como soporte a la aplicación de las ANN no se hizo uso de las técnicas estadísticas relacionadas al cálculo predictivo.

3.4. Diseño Muestral

3.4.1. Población

La población está conformada por 377 proposiciones parlamentarias presentadas en el periodo anual de sesiones 2019 perteneciente al periodo parlamentario 2016. Para el presente estudio se consideró los proyectos de ley y las resoluciones legislativas como proposiciones parlamentarias ya que estas

unidades de estudio pueden alcanzar a convertirse en ley y cumplen con el propósito del estudio.

3.4.2. Muestra

Para determinar la muestra se utilizó el muestreo según el criterio del investigador, (Alzamora de los Godos Urcia & Calderón Saldaña, 2010) señala que, la elección de una muestra dependerá de los objetivos del estudio, del esquema de investigación y de la contribución que se desea hacer con ella. Por lo tanto, la muestra también está conformada por 377 proposiciones parlamentarias presentadas en el periodo anual de sesiones 2019 perteneciente al periodo parlamentario 2016.

3.5. Aspectos Éticos

En la presente investigación se garantiza los siguientes principios éticos:

- No se utiliza el trabajo de otros investigadores o autores previa citación con las adecuadas fuentes que se hayan incluido en el estudio.
- Se reconoce apropiadamente las contribuciones de todos los participantes en la investigación.
- Se trata con la debida reserva la información obtenida y no se utiliza para propósitos distintos a los objetivos de ésta investigación.

- El investigador garantiza la fidelidad de la información y los datos generados como producto de su investigación.

CAPÍTULO IV

RESULTADOS

4.1. Presentación de los Resultados Descriptivos

Para la presentación de los resultados obtenidos se hizo uso de los programas SPSS versión 25 y Microsoft Office Excel 365. A continuación, se describe el análisis de los variables que sirvieron como datos de entrada (atributos) para crear el modelo algorítmico.

Tipo de Proponente (Variable 1)

Esta variable está compuesta por tres valores:

0 = Congresista de la República

1 = Poder Ejecutivo

2 = Otra entidad o grupo que tiene iniciativa legislativa

Interpretación y análisis: no se consideró su análisis ya que sólo se tomó las proposiciones elaboradas por los congresistas. Al volverse una constante es una variable no significativa.

Tipo de Proposición Legislativa (Variable 2)

Esta variable está compuesta por tres valores:

1 = Proyecto de ley

2 = Proyecto de resolución legislativa

3 = Proyecto de ley de reforma constitucional

Interpretación y análisis: en la Tabla 13, se observó que, de 377 proposiciones de la muestra, la proposición del tipo proyecto de ley con 364 casos representaron el 96.6%; seguido del tipo proyecto de resolución legislativa con 2 casos y representaron el 0.5%; y finalmente tenemos con 11 casos al tipo proyecto de ley de reforma constitucional con un 2.9%. Esto nos indica que las proposiciones que se elaboraron en el Congreso fueron mayormente del tipo Proyecto de ley. Para su mejor visualización véase la Figura 35.

Tabla 13

Tipo de proposición legislativa

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Proyecto de ley	364	96,6	96,6	96,6
Proyecto de resolución legislativa	2	,5	,5	97,1
Proyecto de ley de reforma constitucional	11	2,9	2,9	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia.

Figura 35

Tipo de proposición legislativa



Fuente: elaboración propia.

Número de Comisiones que fue Enviada la Proposición (Variable 3)

Esta variable está compuesta por tres valores:

0 = A ninguna comisión

1 = A una sola comisión

2 = A dos comisiones (o más comisiones)

Interpretación y análisis: en la Tabla 14, se observó que, de 377 proposiciones de la muestra, el número de proposiciones que no fueron enviadas a ninguna comisión fueron 264 casos representaron el 70%; seguido de las proposiciones enviadas a una comisión con 86 casos que representaron el 22.8%; y finalmente tenemos 27 casos de proposiciones enviadas por lo menos dos veces que representaron el 7.2%. Esto nos indica que la mayoría de las proposiciones que se elaboraron en el Congreso no se enviaron a ninguna comisión. Para su mejor visualización véase la Figura 36.

Tabla 14

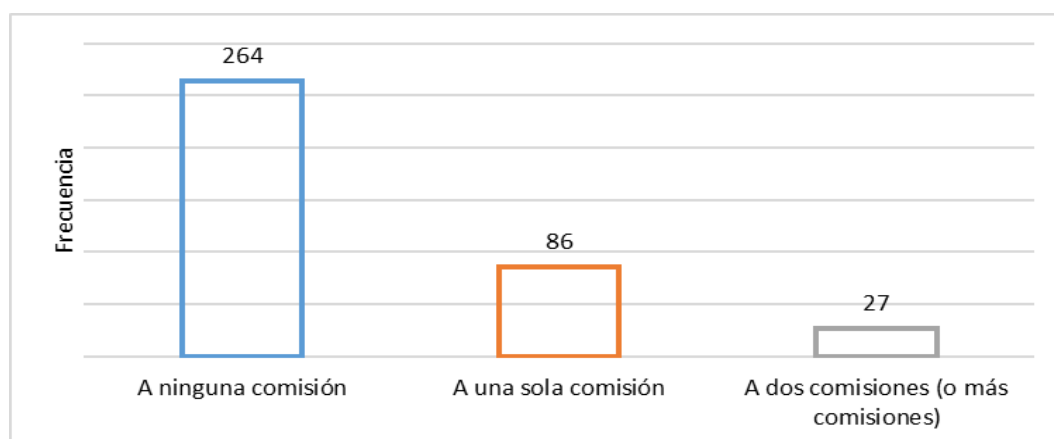
Número de comisiones que fue enviada la proposición

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
A ninguna comisión	264	70,0	70,0	70,0
A una sola comisión	86	22,8	22,8	92,8
A dos comisiones (o más comisiones)	27	7,2	7,2	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia.

Figura 36

Número de comisiones que fue enviada la proposición



Fuente: elaboración propia.

Lista de Comisiones a que fue Enviada la Proposición (Variable 4)

Interpretación y análisis: no se consideró su análisis ya que, para los fines del presente estudio, resulta poco significativo analizar si existe relación entre la viabilidad de las iniciativas y la comisión ordinaria del Congreso a que es derivada cada proposición legislativa en particular.

Comisión de Carácter Económico de la Proposición (Variable 5)

Esta variable está compuesta por dos valores:

0 = No es de carácter económico

1 = Sí es de carácter económico

Interpretación y análisis: en la Tabla 15 se observó que, de 377 proposiciones de la muestra, las proposiciones de carácter no económico con 264 casos representaron el 70%; seguido del tipo de carácter económico con 113 casos que representaron el 30%. Esto nos indica que las proposiciones que se elaboraron en el Congreso fueron en su mayoría de carácter no económico. Para su mejor visualización véase la Figura 37.

Tabla 15

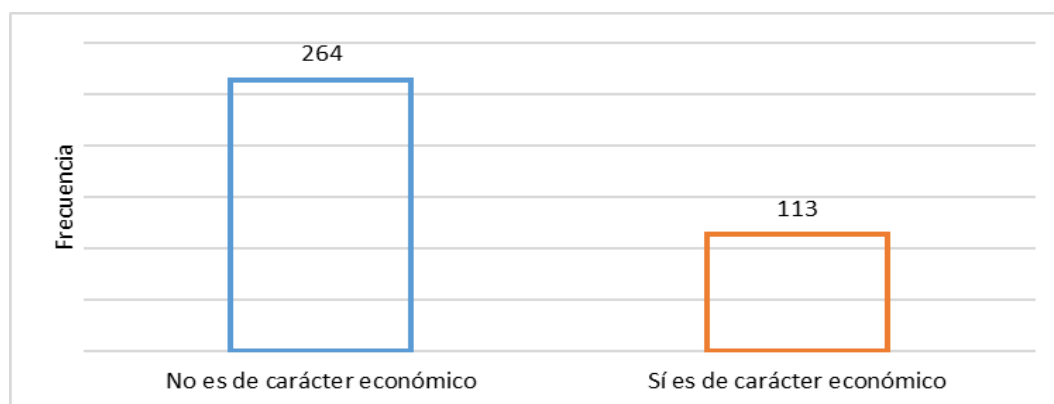
Comisión de carácter económico de la proposición

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
No es de carácter económico	264	70,0	70,0	70,0
Sí es de carácter económico	113	30,0	30,0	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia.

Figura 37

Comisión de carácter económico de la proposición



Fuente: elaboración propia.

La Proposición Tiene Incidencia Económica Directa (Variable 6)

Esta variable está compuesta por dos valores:

0 = No tiene incidencia económica directa

1 = Sí tiene incidencia económica directa

Interpretación y análisis: en la Tabla 16, se observó que, de 377 proposiciones de la muestra, las proposiciones sin incidencia económica directa fueron 244 casos y representaron el 64.7%; seguido del tipo con incidencia económica directa con 133 casos que representaron el 35.3%. Esto nos indica que las proposiciones que se elaboraron en el Congreso fueron en su mayoría sin incidencia económica directa. Para su mejor visualización véase la Figura 38.

Tabla 16

La proposición tiene incidencia económica directa

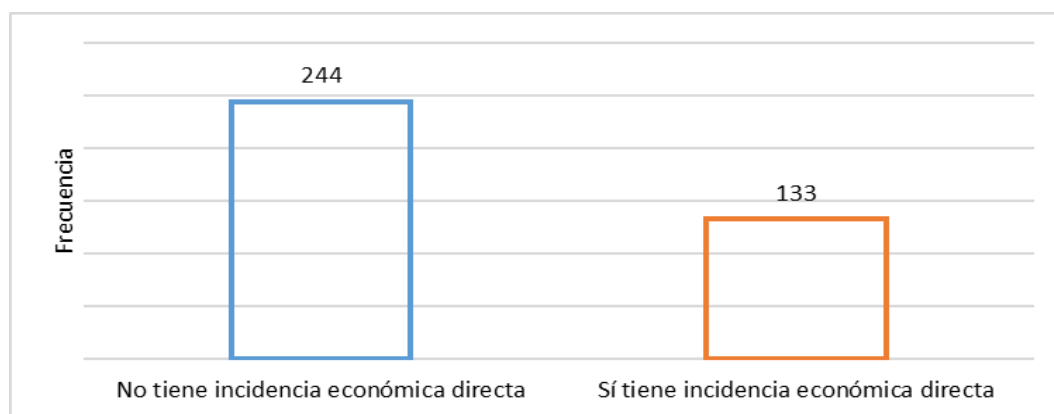
Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
---------	------------	------------	-------------------	----------------------

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
No tiene incidencia económica directa	244	64,7	64,7	64,7
Sí tiene incidencia económica directa	133	35,3	35,3	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia.

Figura 38

La proposición tiene incidencia económica directa



Fuente: elaboración propia.

Nombre del Congresista Proponente (Variable 7)

Interpretación y análisis: a juicio propio no se consideró su análisis ya que en el proceso se realizará predicciones para las proposiciones, no por cada congresista.

Número de Congresistas que Firman la Propuesta (Variable 8)

Esta variable está compuesta por tres valores:

0 = Seis congresistas o menos, cuando el grupo parlamentario está conformado por menos de seis congresistas

1 = Más de seis congresistas

2 = Más de diez congresistas

Interpretación y análisis: en la Tabla 17, se observó que, de 377 proposiciones de la muestra, el número de proposiciones con firmas entre seis o menos participantes fueron 300 casos y representaron el 79.6%; seguido de más de seis firmas de los participantes con 64 casos y representaron el 17%; y con 13 casos, las proposiciones con más de diez firmas de los participantes que representaron el 3.4%. Esto nos indica que la mayoría de las proposiciones que se elaboraron en el Congreso cuentan con seis firmas o menos firmas de participantes. Para su mejor visualización véase la Figura 39.

Tabla 17

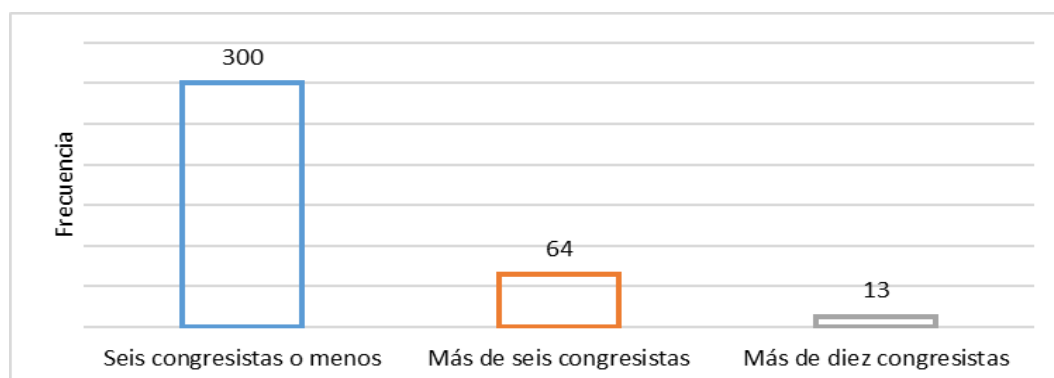
Número de congresistas que firman la propuesta

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Seis congresistas o menos	300	79,6	79,6	79,6
Más de seis congresistas	64	17,0	17,0	96,6
Más de diez congresistas	13	3,4	3,4	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia.

Figura 39

Número de congresistas que firman la propuesta



Fuente: elaboración propia.

Grupo Parlamentario del Proponente Principal (Variable 9)

Interpretación y análisis: a juicio propio no se consideró su análisis ya que en el proceso se realizará predicciones para las proposiciones, no por cada grupo parlamentario.

El Proponente Forma Parte de la Mayoría Parlamentaria (Variable 10)

Esta variable está compuesta por tres valores:

0 = Es de la mayoría parlamentaria

1 = Es de la oposición. Criterio: por lo mínimo mayor a 6 integrantes y menor a la mayoría.

2 = Ni lo uno ni lo otro. Criterio: por lo máximo 6 integrantes.

Interpretación y análisis: en la Tabla 18, se observó que, de 377 proposiciones de la muestra, el número de proponentes de la mayoría parlamentaria fueron 34 casos que representaron el 9%; seguido de la

oposición con 299 casos que representaron el 79.3%; y finalmente tenemos con 44 casos las proponentes que no pertenecen a la mayoría ni a la oposición que representaron el 11.7%. Esto nos indica que la mayoría de las proposiciones que se elaboraron en el Congreso tienen predominio en la oposición parlamentaria. Para su mejor visualización véase la Figura 40.

Tabla 18

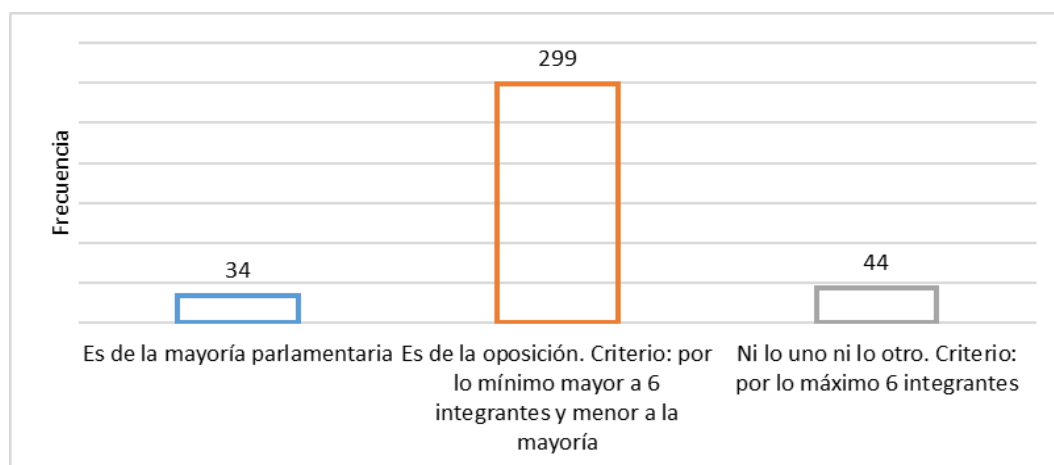
El proponente forma parte de la mayoría parlamentaria

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Es de la mayoría parlamentaria	34	9,0	9,0	9,0
Es de la oposición. Criterio: por lo mínimo mayor a 6 integrantes y menor a la mayoría	299	79,3	79,3	88,3
Ni lo uno ni lo otro. Criterio: por lo máximo 6 integrantes	44	11,7	11,7	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia.

Figura 40

El proponente forma parte de la mayoría parlamentaria



Fuente: elaboración propia.

El Proponente Recibió Capacitación en la Metodología ACB (Variable 11)

Esta variable está compuesta por dos valores:

0 = Es de la mayoría parlamentaria

1 = Ni lo uno ni lo otro

Interpretación y análisis: a juicio propio no se consideró su análisis ya que no se contó con las facilidades de la obtención de los datos en la página web de Infobae.

El Proponente Tiene Experiencia Parlamentaria (Variable 12)

Esta variable está compuesta por dos valores:

0 = No tiene experiencia parlamentaria previa

1 = Sí tiene experiencia parlamentaria previa

Interpretación y análisis: en la Tabla 19, se observó que, de 377 proposiciones de la muestra, los proponentes sin experiencia parlamentaria fueron 355 casos y representó el 94.2%; seguido del tipo con experiencia parlamentaria con 22 casos y representaron el 5.8%. Esto nos indica que las proposiciones que se elaboraron en el Congreso fueron en su mayoría proponentes sin experiencia parlamentaria previa. Para su mejor visualización véase la Figura 41.

Tabla 19

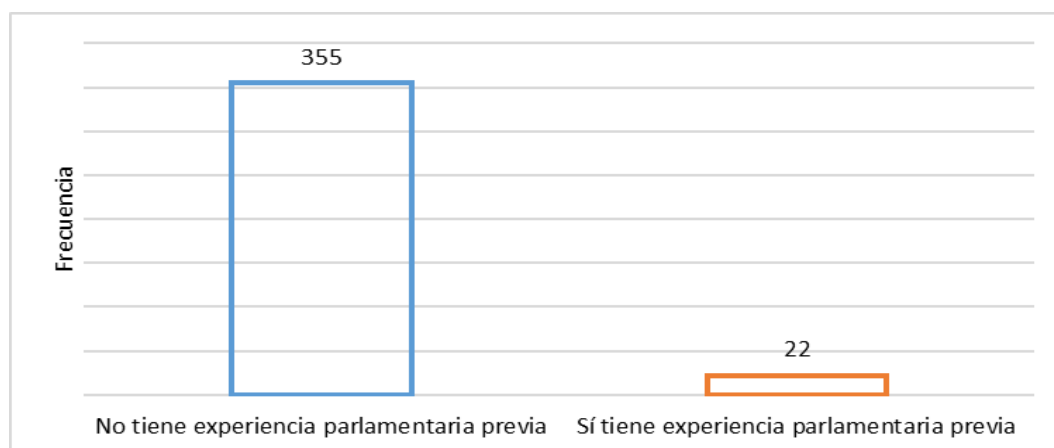
El proponente tiene experiencia parlamentaria

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
No tiene experiencia parlamentaria previa	355	94,2	94,2	94,2
Sí tiene experiencia parlamentaria previa	22	5,8	5,8	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia

Figura 41

El proponente tiene experiencia parlamentaria



Fuente: elaboración propia.

El Proponente Tiene Experiencia Administrativa (Variable 13)

Esta variable está compuesta por dos valores:

0 = No tiene experiencia administrativa

1 = Sí tiene experiencia administrativa

Interpretación y análisis: a juicio propio no se consideró su análisis ya que no se contó con las facilidades de la obtención de los datos en la página web de Infobae.

Sexo del Proponente (Variable 14)

Esta variable está compuesta por dos valores:

0 = Femenino

1 = Masculino

Interpretación y análisis: en la Tabla 20, se observó que, de 377 proposiciones de la muestra, las proposiciones cuyo autor es de sexo femenino fueron 98 casos y representaron el 26%; seguido del sexo masculino con 279 casos que representaron el 74%. Esto nos indica que las proposiciones que se elaboraron en el Congreso fueron en su mayoría proposiciones presentadas por congresistas del sexo masculino. Para su mejor visualización véase la Figura 42.

Tabla 20

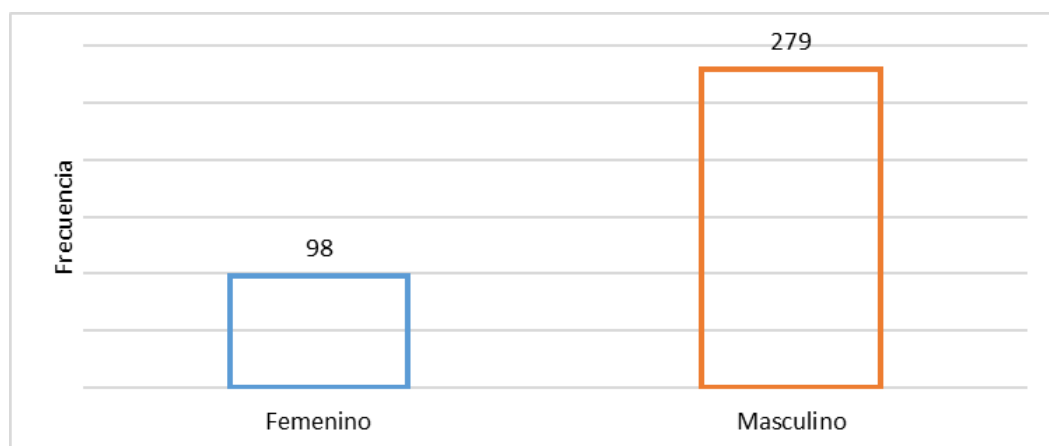
Sexo del proponente

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Femenino	98	26,0	26,0	26,0
Masculino	279	74,0	74,0	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia

Figura 42

Sexo del proponente



Fuente: elaboración propia.

Número de Proposiciones Presentadas por el Congresista (Variable 15)

Esta variable está compuesta por cuatro valores:

0 = Por lo menos mayor a 80

1 = Por lo menos mayor a 20

2 = Por lo menos mayor a 1

3 = Ninguno presentado (en el periodo anual de sesiones elegido por lo menos un congresista presentó 1 proposición)

Interpretación y análisis: en la Tabla 21, se observó que, de 377 proposiciones de la muestra, los congresistas que presentaron al menos más de 80 proposiciones fueron 12 casos y representaron el 3.2%; seguido por 263 casos, los congresistas que presentaron por lo menos más de 20 proposiciones y representaron el 69.8%; y finalmente, tenemos con 102 casos, los que presentaron por lo menos más de una proposición

que representaron el 27.1%. Esto nos indica que la mayoría de las propuestas que se elaboraron en el Congreso tienen predominio en los congresistas que presentaron por lo menos más de 20 y menos de 80 proposiciones. Para su mejor visualización véase la Figura 43.

Tabla 21

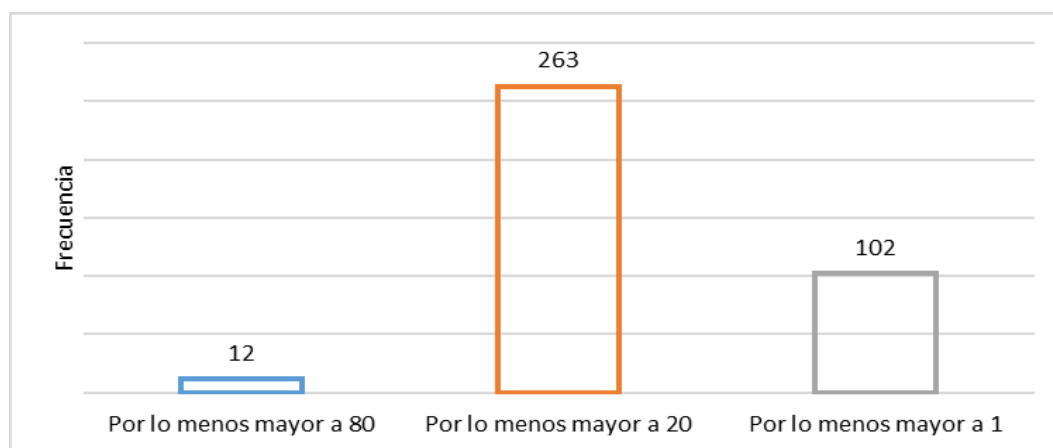
Número de proposiciones presentadas por el congresista

Valores	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Por lo menos mayor a 80	12	3,2	3,2	3,2
Por lo menos mayor a 20	263	69,8	69,8	72,9
Por lo menos mayor a 1	102	27,1	27,1	100,0
Total	377	100,0	100,0	

Fuente: elaboración propia

Figura 43

Número de proposiciones presentadas por el congresista



Fuente: elaboración propia.

4.2. Presentación de los Resultados de la Predicción

Para la presentación de los resultados de la predicción se hizo uso de la herramienta SPSS versión 25, mediante la opción de redes neuronales artificiales multicapa, se eligió este algoritmo ya que este tipo de redes copian el comportamiento de cualquier sistema de datos usando entradas y salidas definidas por el investigador.

En la Tabla 22, se observa la división de la base de datos en el grupo de entrenamiento en la que se crea el modelo, se seleccionaron 274 proposiciones para realizar el entrenamiento de la predicción que representaron el 72.7% de la muestra total. En el grupo de pruebas, con la que se evalúa el modelo, se seleccionaron 103 proposiciones para realizar las pruebas de predicción que representaron el 27.3% de la muestra total.

Tabla 22
Resumen de procesamiento de casos

		N	Porcentaje
Muestra	Entrenamiento	274	72,7%
	Pruebas	103	27,3%
Válido		377	100,0%
Excluido		0	
Total		377	

Fuente: elaboración propia

En la Tabla 23, se describe la información de la red neuronal. A continuación, se describen las capas de la red neuronal generada,

compuesta por tres secciones: la capa de entrada, las capas ocultas y la capa de salida:

- La sección de la capa de entrada de la red neuronal está conformada por nueve factores que son los atributos de entradas o variables independientes (ítems) y 23 unidades del total de las categorías o valores de los atributos de entrada.
- La sección de la capa oculta está conformada por una capa oculta y esta capa oculta única está conformada por ocho unidades.
- La sección de la capa de salida, lo conforma una variable dependiente y dos número de unidades que son sus dos categorías o valores .

Tabla 23

Información de red

Capa de entrada	Factores	1	2. Tipo de proposición legislativa
		2	3. Número de comisiones que fue enviada la proposición
		3	5. Comisión de carácter económico de la proposición
		4	6. La proposición tiene incidencia económica directa
		5	8. Número de congresistas que firman la propuesta
		6	10. El proponente forma parte de la mayoría parlamentaria
		7	12. El proponente tiene experiencia parlamentaria
		8	14. Sexo del proponente
		9	15. Número de proposiciones presentadas por el congresista
	Número de unidades ^a		23
Capas ocultas	Número de capas ocultas		1
	Número de unidades en la capa oculta 1 ^a		8
	Función de activación		Tangente hiperbólica
Capa de salida	Variables dependientes	1	Viabilidad
	Número de unidades		2
	Función de activación		Softmax

a. Se excluye la unidad de sesgo

Fuente: elaboración propia

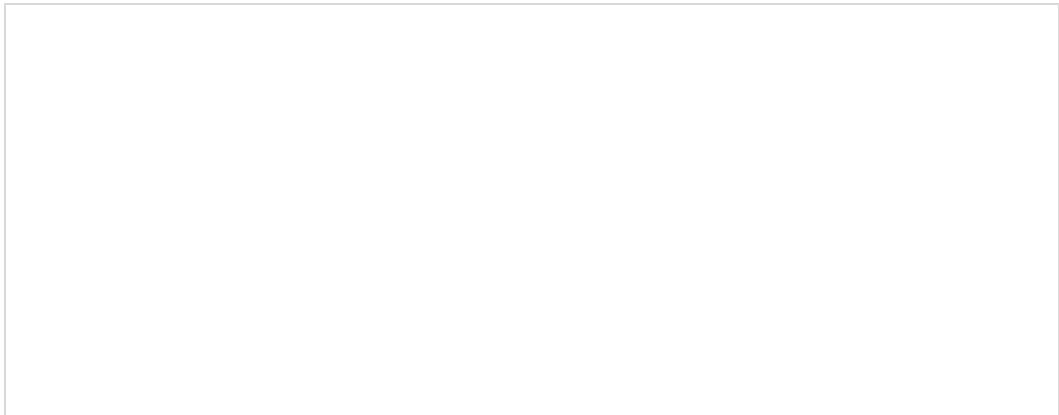
En la Figura 44, se muestra la red neuronal, sus capas y sus conexiones sinápticas. Sus capas contienen tres secciones: la capa de entrada, las capas ocultas y la capa de salida. En la capa de entrada de la red neuronal lo conforman, una unidad de sesgo y 23 unidades. Las 23 unidades en la capa de entrada representan el total de las categorías de las variables independientes, ya sea dicotómica o multicotómica, por ejemplo, el primer factor se divide en tres categorías, 1 (Proyecto de ley), 2 (Proyecto de resolución legislativa) y 3 (Proyecto de ley de reforma constitucional). En la sección de las capas ocultas lo conforman, una unidad de sesgo y ocho unidades ocultas. En la sección de la capa de salida lo conforman una unidad de sesgo y dos unidades. Las dos unidades representan las dos categorías de la variable independiente correspondientes a las categorías de viable y no viable. Finalmente, se muestran las conexiones sinápticas, conformadas por las ponderaciones sinápticas mayores a cero de color gris y las ponderaciones sinápticas menores a cero de color azul

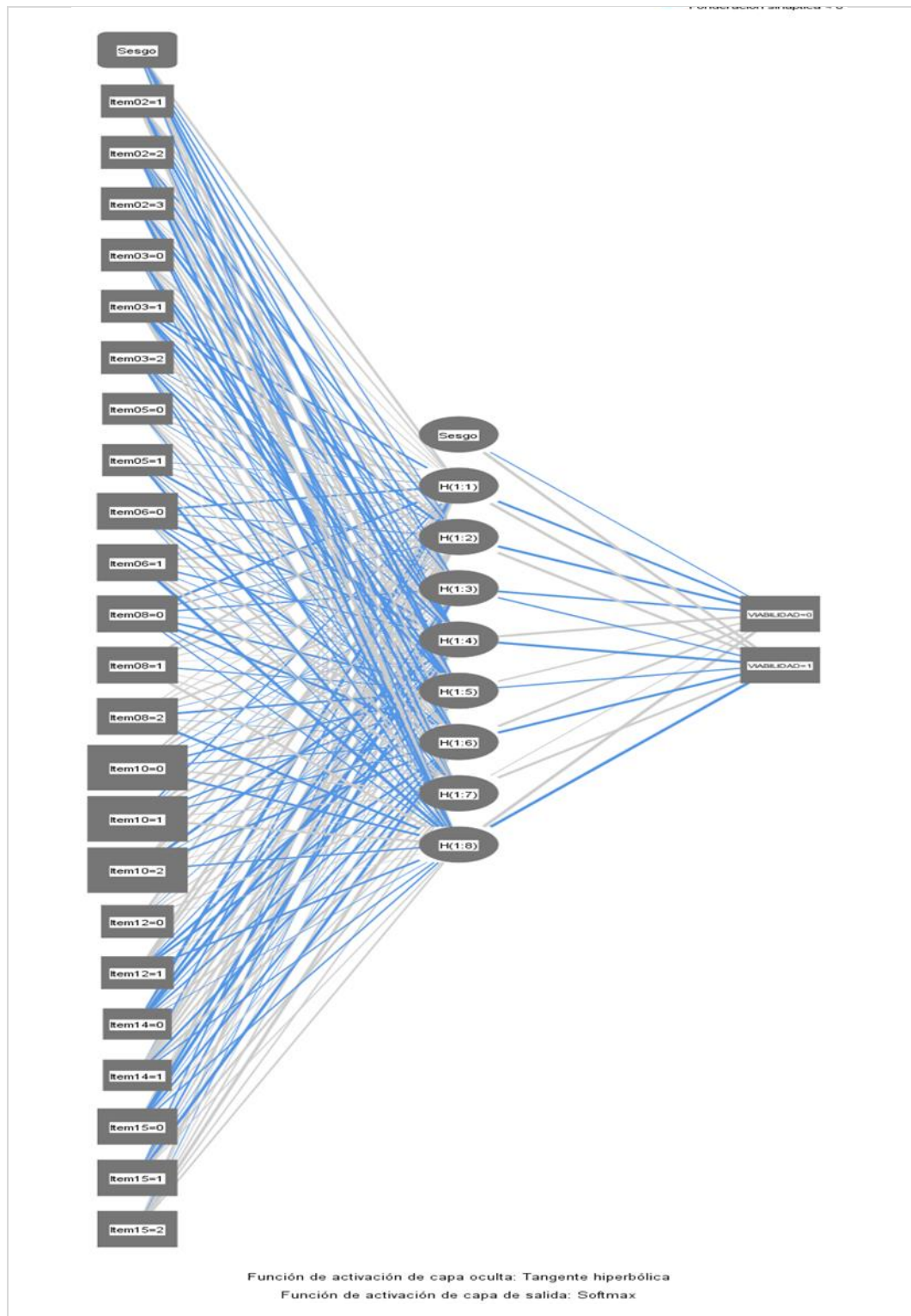
En la Tabla 24, se describen las estimaciones de parámetro de la red neuronal, la capa de entrada genera a la capa oculta y la capa oculta genera la capa de salida por eso se requiere pesos sinápticos. Los pesos sinápticos indican el grado de participación que tiene cada una de las variables al momento de generar la siguiente capa. En este caso sería,

los pesos entre la capa de entrada con la capa oculta y la capa oculta con la capa de salida.

Figura 44

Diagrama de la red neuronal artificial





Fuente: elaboración propia.

Tabla 24
Estimaciones de parámetro

Predictor	Pronosticado								Pronosticado	
	Capa oculta 1								Capa de salida	
	H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	H(1:8)	[VIABILIDAD=0]	[VIABILIDAD=1]
(Sesgo)	0.405	-0.160	-0.258	-0.274	-0.491	0.323	0.266	-0.236		
[ítem02=1]	0.240	0.434	-0.061	-0.243	-0.395	-0.004	-0.324	0.634		
[ítem02=2]	0.043	-0.312	0.441	-0.346	0.294	-0.428	0.383	-0.488		
[ítem02=3]	0.139	0.144	-0.202	-0.473	0.183	0.294	-0.250	-0.401		
[ítem03=0]	0.010	0.679	0.240	0.308	-0.420	-0.418	0.348	-0.107		
[ítem03=1]	-0.409	-0.492	0.183	-0.403	-0.629	-0.061	0.406	0.243		
[ítem03=2]	0.326	-0.156	0.333	-0.224	-0.452	0.312	-0.173	-0.278		
[ítem05=0]	0.470	0.373	-0.141	0.121	0.121	0.511	-0.219	0.030		
[ítem05=1]	-0.085	-0.042	0.328	-0.497	0.227	-0.176	0.278	-0.482		
[ítem06=0]	-0.536	-0.269	0.355	-0.620	-0.155	-0.435	0.458	-0.072		
[ítem06=1]	0.394	0.888	0.340	-0.397	-0.087	0.466	0.495	-1.080		
[ítem08=0]	-0.696	-0.217	-0.387	-0.033	-0.500	-0.844	-0.196	-0.242		
[ítem08=1]	0.624	0.706	0.214	0.031	-0.284	0.092	0.242	0.836		
[ítem08=2]	0.146	0.187	0.491	-0.087	-0.538	-0.009	0.071	-0.723		
[ítem10=0]	0.342	0.455	-0.141	-0.054	-0.430	-0.297	0.411	-0.571		
[ítem10=1]	-0.265	-0.016	0.465	0.620	-0.354	-0.350	0.260	0.912		
[ítem10=2]	0.258	0.500	0.045	-0.644	-0.076	-0.509	0.156	-0.403		
[ítem12=0]	0.394	0.010	0.103	-0.057	0.237	0.234	0.309	0.044		
[ítem12=1]	0.486	0.493	0.211	-0.152	-0.483	-0.480	0.126	-0.451		
[ítem14=0]	-0.651	0.159	-0.442	-1.008	-0.112	-0.328	0.283	-0.067		
[ítem14=1]	-0.272	0.624	0.415	0.406	0.284	-0.411	0.477	-0.250		
[ítem15=0]	0.012	0.726	-0.431	-0.575	-0.003	-0.350	0.188	-0.256		
[ítem15=1]	-0.239	0.022	-0.308	0.250	-0.021	-0.601	0.383	0.176		
[ítem15=2]	0.406	0.063	-0.061	0.051	0.249	0.640	0.239	0.282		
(Sesgo)									-0.201	0.851
H(1:1)									-0.549	0.929
H(1:2)									-0.930	1.298
H(1:3)									-0.491	-0.334
H(1:4)									0.939	-1.277
H(1:5)									0.374	-0.331
H(1:6)									0.557	-0.855
H(1:7)									0.145	0.599
H(1:8)									1.304	-1.302

Fuente: elaboración propia

En la Tabla 25, se describen las variables implicadas con un valor de importancia para la predicción del modelo, la variable con mayor influencia se considera 100% y el resto de las variables se analizan en función de la primera como una proporción de la variable más importante. En este caso, la variable con mayor importancia en la generación del modelo fue la variable 10, esto quiere decir que, la predicción de la viabilidad legislativa ha sido fuertemente influenciada en un 100% por la variable 10, seguido de la variable 6 en un 53.6% y sucesivamente.

Tabla 25

Importancia de las variables independientes

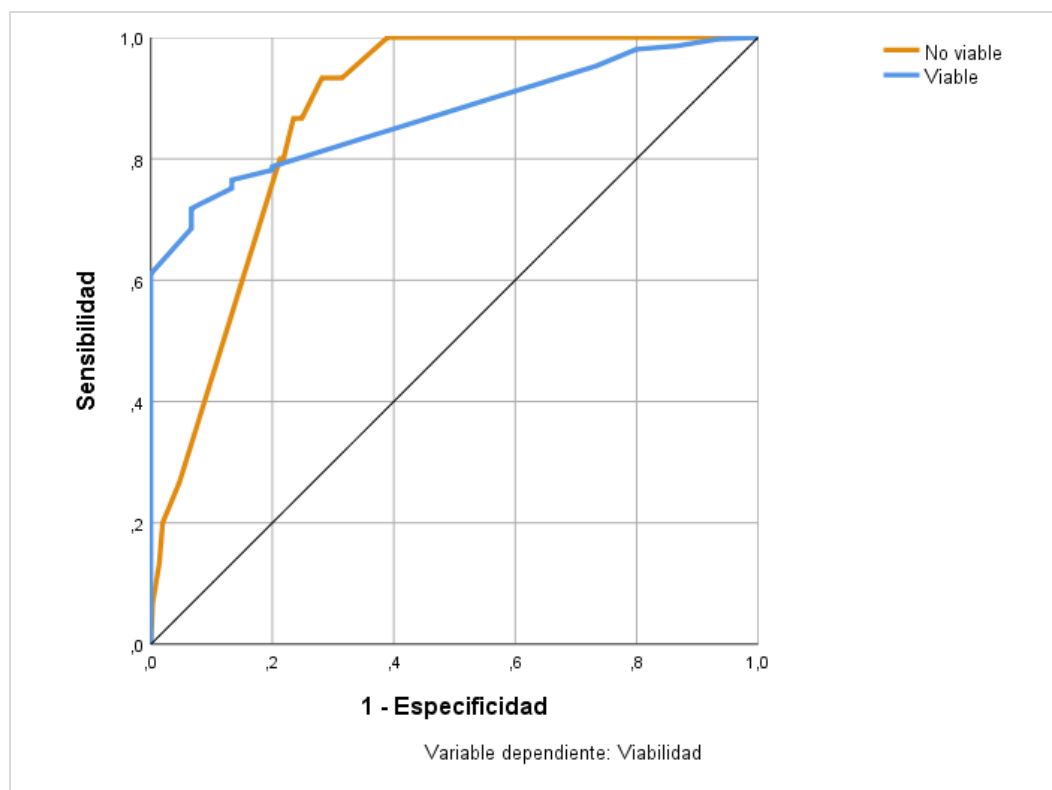
	Importancia	Importancia normalizada
10. El proponente forma parte de la mayoría parlamentaria	0.245	100.0%
6. La proposición tiene incidencia económica directa	0.131	53.6%
8. Número de congresistas que firman la propuesta	0.131	53.5%
15. Número de proposiciones presentadas por el congresista	0.125	50.9%
2. Tipo de proposición legislativa	0.082	33.4%
3. Número de comisiones que fue enviada la proposición	0.080	32.8%
12. El proponente tiene experiencia parlamentaria	0.080	32.6%
5. Comisión de carácter económico de la proposición	0.068	27.7%
14. Sexo del proponente	0.058	23.7%

Fuente: elaboración propia

En la Figura 45, se muestra la curva ROC (Receiver Operating Characteristic), la variable dependiente contiene dos categorías, por lo tanto, existen dos intensiones a predecir, en azul se representa la viabilidad y en rojo la no viabilidad. La curva ROC encierra un área bajo la curva, mientras más grande sea el área, mejor es la capacidad predictiva.

Figura 45

Curva ROC (curva rendimiento diagnóstico)



Fuente: elaboración propia

En la Tabla 26, se describen los valores exactos del área bajo la curva ROC con respecto al modelo generado. Se puede observar un 87.1% del área bajo la curva para ambas categorías, esto quiere decir que, la capacidad predictiva para encontrar la viabilidad legislativa es igual entre las dos categorías.

Tabla 26

Área bajo la curva

		Áreas
Viabilidad	0	,871
	1	,871

Fuente: elaboración propia

En la Tabla 27, se observan los resultados del proceso de cálculo de pronósticos incorrectos en los modelos predictivos divididos en dos secciones, de entrenamiento y de pruebas. En la sección de entrenamiento, el porcentaje de pronósticos incorrectos es del 4% con un error de entropía cruzada de 35.261. En la sección de pruebas, el error de entropía cruzada es de 12.151 y el porcentaje el porcentaje de pronósticos incorrectos equivale al 3.9%, éste es un indicador de la falencia del porcentaje en que el modelo falla. Este valor es el complemento del 96.1% que se muestra en la sección de Pruebas de la Tabla 28.

Tabla 27
Resumen del modelo

Entrenamiento	Error de entropía cruzada	35,161
	Porcentaje de pronósticos incorrectos	4,0%
	Regla de parada utilizada	1 paso(s) consecutivo(s) sin disminución del error ^a
	Tiempo de entrenamiento	0:00:00.16
Pruebas	Error de entropía cruzada	12,151
	Porcentaje de pronósticos incorrectos	3,9%

Variable dependiente: Viabilidad

a. Los cálculos de error se basan en la muestra de comprobación.

Fuente: elaboración propia

En la Tabla 28, se muestran los resultados del proceso de cálculo de precisiones en los modelos predictivos divididos también en dos secciones, la de entrenamiento y la de pruebas. En la sección de entrenamiento, se observa el número de acuerdos, coincidencias o concordancias entre el valor real y el valor pronóstico, el porcentaje global equivale a un 96%. En la sección de pruebas, de comprobación o también

llamado grupo de evaluación, el porcentaje global equivale al 96.1%, que es la cantidad que el modelo acierta.

En resumen, el modelo generado alcanzó la precisión del 96.1% de predicción de la viabilidad legislativa con un 3.9% de errores de predicción. Por lo tanto, se puede considerar como un modelo confiable de alta precisión.

Tabla 28
Clasificación

Muestra	Observado	Pronosticado		
		0	1	Porcentaje correcto
Entrenamiento	0	0	11	0,0%
	1	0	263	100,0%
	Porcentaje global	0,0%	100,0%	96,0%
Pruebas	0	0	4	0,0%
	1	0	99	100,0%
	Porcentaje global	0,0%	100,0%	96,1%

Variable dependiente: Viabilidad

Fuente: elaboración propia

CAPÍTULO V

DISCUSIÓN

En el estudio realizado por (Zevallos Salazar, 2017) obtuvo una precisión del 84% para predecir el rendimiento académico de los estudiantes de tres colegios del nivel primario del distrito de Ventanilla. Comparando los resultados con la presente investigación existe una diferencia de 12.1% a favor de la presente. El estudio de (Zevallos Salazar, 2017) se caracterizó por usar 33 factores sobre una población de 1,073 estudiantes, un número mayor en factores y en población al presente trabajo. La diferencia puede estar relacionada a los factores identificados o a la necesidad de aumentar la población de estudio.

En la investigación presentada por (Aguilar Vilca & Camargo Ramos, 2021) obtuvieron un resultado del 97.38% para predecir la deserción de clientes de microcréditos de bancos. Los resultados obtenidos por (Aguilar Vilca & Camargo Ramos, 2021) dan una diferencia del 1.28% a favor de su estudio. La población de estudio estuvo conformada por 24,420 clientes, que ha comparación con la presente, es una población grande. El trabajo se caracterizó por fusionar tres técnicas de predicción para mejorar su capacidad predictiva. La diferencia puede estar relacionada al tamaño de la población y a la hibridación de técnicas, ya que, en la presente, sólo se aplicó una técnica.

En el trabajo publicado por (Kornilova, Argyle, & Eidelman, 2018) obtuvieron una precisión de acierto al 86.21% entre ocho modelos evaluados basados en el uso de metadatos. En comparación con la presente investigación

hace una diferencia de 10.11% a favor del presente estudio. La precisión del modelo evaluado por (Kornilova, Argyle, & Eidelman, 2018) es notable en el uso de grandes cantidades de datos ingresados no solo referido a la cantidad de proyectos de ley (aproximadamente 20,000 proyectos de ley en cada sesión) sino también de la inclusión de alrededor de 2,683 metadatos.

En la investigación realizada por (Khashman & Khashman, 2016), referente a la predicción de votos en el Congreso Estadounidense obtuvieron una precisión del valor global de 96.33% utilizando el modelo ANN y el 95.29% utilizando el modelo SVM. A comparación de los resultados del presente estudio, los resultados dejan una diferencia de precisión en un 0.23%. Uno de los factores de esta diferencia es que no se recolectaron datos secundarios sino datos primarios, asociados con las opiniones de los congresistas estadounidenses sobre temas nacionales. A pesar de tener una población de estudio de 435, que es relativamente mayor, contó con más datos de entrada que consistieron en 16 atributos frente a nueve atributos del presente estudio.

En su investigación (Zolghadr, Niaki, & Akhavan Niaki, 2016), obtuvieron un 93.14% para la SVM y 92% de precisión en las ANN para la predicción de las elecciones presidenciales. A comparación el presente estudio, en la que se obtuvo un 96.1%, la diferencia de precisión fue de 2.96% a favor de la presente. En el presente estudio se recolectó datos de un periodo anual de sesiones, mientras que en la investigación de (Zolghadr, Niaki, & Akhavan Niaki, 2016) fue de 12 años (datos de las elecciones presidenciales de 2004, 2008 y 2012) con una población mayor equivalente a 791 votos. El presente estudio se basó en la aplicación de las ANN, mientras que, el modelo SVR fue

evaluado como el de mejor rendimiento en el estudio de (Zolghadr, Niaki, & Akhavan Niaki, 2016).

Colliri & Zhao (2019) obtuvieron un 90% de precisión para predecir las condenas por corrupción de los congresistas brasileños aplicando redes temporales, es decir, un 6.1% a favor de la presente investigación. Una ventaja en la investigación realizada por (Colliri & Zhao, 2019), es que propusieron una técnica basada en la red para analizar los datos de votación de proyectos de ley que comprendían votos de los congresistas brasileños, estos conjuntos de datos comprendieron 3,407 sesiones de votación de proyectos de ley en la Cámara de Representantes, lo que implica que es posible revelar el comportamiento ilegal de los políticos a través de sus actividades públicas legales.

En la investigación realizada por (Kaur & Bozié, 2019) referente a la predicción de las sentencias del Tribunal Europeo de Derechos Humanos utilizaron dos algoritmos de predicción, alcanzando una precisión media de 82% para las CNN y un 75% para el modelo de SVM. En el caso de la presente investigación, se alcanzó una precisión de acierto del 96.1%, dando así, una diferencia de 14.1% mayor a la capacidad predictiva del modelo de (Kaur & Bozié, 2019). Entre las causas de esta diferencia se puede mencionar el tamaño de la población, en la presente investigación se trabajó con una población de 377 proposiciones, mientras que en la investigación de (Kaur & Bozié, 2019) se trabajó con una población de 11,532 sentencias como datos de entrada. Puede considerarse también, la diferencia de atributos por cada registro de entrada, ya que la estructura de atributos de una sentencia difiere

de la estructura de un proyecto de ley. Finalmente, se puede mencionar el uso de las CNN y las SVM, utilizado por (Kaur & Bozié, 2019), que son otro tipo de modelo algorítmico con capacidades predictivas.

CONCLUSIONES

- En la presente investigación se confirmó que, es prioritario determinar la estructura de la información en la fase CRISP-DM de comprensión de datos. Mediante el estudio realizado por (Arias-Schreiber Barba, Peña Jumba, & Valdivieso Serrano, 2019) ha permitido conocer y aplicar los atributos (variables exógenas) y las etiquetas (variables endógenas), e identificar los valores categóricos de los datos que han sido la base de la investigación.
- Los resultados de la investigación confirman que es necesario determinar la calidad de los datos mediante la fase CRISP-DM de preparación de los datos. Esto ha permitido asegurar un tratamiento eficiente de 377 proposiciones parlamentarias correspondientes al periodo anual de sesiones 2019 perteneciente al periodo parlamentario 2016, y además, permitió evaluar nueve atributos de los quince atributos seleccionados.
- En la fase CRISP-DM de modelamiento se confirmó que se requiere determinar los procesos de aprendizaje y las técnicas de aprendizaje para su aplicación de acuerdo a la finalidad del estudio, en este caso, se definió el uso de las redes neuronales artificiales de perceptrón multicapa.
- Por último, mediante la fase CRISP-DM de evaluación, se determinó la obtención de un modelo altamente preciso y fiable, comprobando que el modelo arroja una alta probabilidad de predicción (dada su cercanía al 100%) en la viabilidad legislativa, alcanzando una precisión del 96.1% en la que el modelo acierta. Por lo tanto, se concluye que, las redes neuronales

artificiales predicen la viabilidad legislativa, quedando demostrado en el presente estudio realizado.

42

RECOMENDACIONES

- Para un estudio de nivel predictivo es deseable contar con una plataforma big data para facilitar la extracción de una mayor cantidad de datos y sea más asequible mapear la estructura de la información para ordenarla y clasificarla. En caso, no se contara con una plataforma big data, es recomendable optar por programas scraping, siempre en cuando, tratándose de orígenes digitales en portales web. También es recomendable opcionalmente que, los atributos sean sustentados por estudios realizados con anterioridad para dar más consistencia a la investigación.
- Es preciso contar con una información fiable, asegurando la calidad de los datos. Esto permitirá definir la cantidad de datos y atributos válidos para la realización del estudio. Se recomienda usar herramientas de tratamientos de datos.
- Para la elección de una técnica algorítmica del modelado se recomienda contar con conocimientos previos de los procesos de aprendizaje, las técnicas de aprendizaje, la experiencia de otros investigadores de acuerdo a los datos que se están tratando y a la capacidad predictiva, por ejemplo, últimamente se aplican las CNN para el análisis de imágenes. En el presente estudio se optó por las ANN ya que este tipo de redes copian el comportamiento de cualquier sistema usando entradas y salidas definidas por el investigador. Es recomendable mencionar que, para iniciar el uso práctico de las técnicas de aprendizaje bastaría con usar programas que las

incorporan, para que el investigador se sumerja en el análisis de los resultados, luego de comprenderlas, en base a su inquietud de conocimientos, personalizarlas y experimentar.

- Finalmente, para obtener una mayor precisión de aciertos y una mejora en el aprendizaje en los modelos de las ANN, y por ende, una reducción en el porcentaje de error, es recomendable trabajar con grandes cantidades de datos, ya que existe una relación directa entre un mayor acierto de predicción y una mayor cantidad de datos procesados. Referente al estudio realizado es recomendable incrementar la cantidad de proposiciones parlamentarias en estudio y establecer más de nueve atributos por cada proposición. Sería recomendable también, probar la hibridación de múltiples técnicas para reforzar y aumentar la precisión predictiva.

FUENTES DE INFORMACIÓN

Aguilar Vilca, D., & Camargo Ramos, J. C. (2021). *Sistema inteligente basado en redes neuronales, máquina de soporte vectorial y random forest para la predicción de deserción de clientes en microcréditos de bancos.*

Obtenido de Página web de Cybertesis del repositorio de tesis digitales UNMSM: <https://cybertesis.unmsm.edu.pe/handle/20.500.12672/16390>

Alzamora de los Godos Urcia, L. A., & Calderón Saldaña, J. P. (2010).

Metodología de investigación científica en posgrado. Lima: Editorial Lulu.

Analytics Steps. (22 de Agosto de 2019). *7 Types of Activation Functions in*

Neural Network of Dinesh Kumawat. Obtenido de Página web de

Analytics Steps: <https://www.analyticssteps.com/blogs/7-types-activation-functions-neural-network>

Analytics Vidhya. (23 de Enero de 2019). *Brief History of Neural Networks.*

Obtenido de Página web de Analytics Vidhya:

<https://medium.com/analytics-vidhya/brief-history-of-neural-networks-44c2bf72eec>

Arias-Schreiber Barba, F., Peña Jumba, A., & Valdivieso Serrano, L. (2019). *La*

Evaluación de las Leyes en el Perú, El análisis costo-beneficio en el

Congreso de la República. Lima: Fondo Editorial de la Pontificia

Universidad Católica del Perú.

Awad, M., & Khanna, R. (2015). *Efficient Learning Machines: Theories,*

Concepts, and Applications for Engineers and. New York: ApressOpen.

Barrientos Carles, V. (1 de Agosto de 2020). *Conocimiento computacional:*

racionalidad versus empirismo. Obtenido de Página web del portal

Gazeta: <https://gazeta.gt/conocimiento-computacional-racionalidad-versus-empirismo/>

Beltrán Castañón, C. (7 de septiembre de 2020). *La Inteligencia Artificial como*

herramienta de desarrollo en Perú. Obtenido de Página Facebook de

UNESCO Perú:

<https://www.facebook.com/oficinaunescolima/videos/1028715270912806>

- Berzal, F. (2018). *Redes neuronales & deep learning*. Granada: Independently published.
- Bianchi, J. (8 de junio de 2018). *Jurimetría: desde la predictibilidad de las sentencias a la viabilidad de proyectos de ley*. Obtenido de Página web de Idealex: <https://idealex.press/jurimetria-predictibilidad-de-sentencias-a-la-viabilidad-de-proyectos-de-ley/>
- Bunge, M. (2001). *La ciencia, su método y su filosofía*. Buenos Aires: Editorial Sudamericana S.A.
- Chollet, F. (2018). *Deep Learning with Python*. New York: Editorial Manning Publications Co.
- Colliri, T., & Zhao, L. (2019). Analyzing the Bills-Voting Dynamics and Predicting Corruption-Convictions Among Brazilian Congressmen Through Temporal Networks. São Paulo, São Paulo, Brazil.
- Congreso de la República. (2012). *Escala renumerativa*. Obtenido de Portal del Congreso de la República:
https://www.peru.gob.pe/docs/PLANES/16/PLAN_16_Escala_Remunerativa_Empleados__D.L._728_2012.pdf
- Congreso de la República. (2016). *Flujograma legislativo ¿Cómo se crea una Ley?* Obtenido de Portal del Congreso de la República:
<https://www.congreso.gob.pe/Flujogramalegislativo/>
- Congreso de la República. (1 de Diciembre de 2017). *Manual de Aplicación del Análisis Costo Beneficio (ACB) y del Análisis Costo Efectividad (ACE) para la presentación o evaluación de proyectos de ley*. Obtenido de Portal del Centro de Estudios Constitucionales y Parlamentarios del Congreso de la República:

[https://www2.congreso.gob.pe/Sicr/CCEP/CCEPpub.nsf/ID_Revistas/L006/\\$FILE/libro-6.pdf](https://www2.congreso.gob.pe/Sicr/CCEP/CCEPpub.nsf/ID_Revistas/L006/$FILE/libro-6.pdf)

Congreso de la República. (01 de Diciembre de 2020). *Proyectos de ley*.

Obtenido de Portal del Congreso de la República:

<http://www.congreso.gob.pe/pley-2016-2021>

Congreso de la República. (2020b). *Proyectos de ley archivados*. Obtenido de

Portal del Congreso de la República:

<https://www4.congreso.gob.pe/comisiones/1995/trabajo/317.htm>

Congreso de la República. (01 de Febrero de 2021). *Consulta del personal*.

Obtenido de Portal del Congreso de la República:

<http://www.congreso.gob.pe/consulta-personal>

Congreso de la República. (1 de Abril de 2021b). *Manual de la técnica*

legislativa. Obtenido de Página web del Departamento de Comisiones

Congreso de la República:

<https://www.congreso.gob.pe/Docs/dgp/files/manual-tecnica-legislativa-3raedicion.pdf>

Congreso de la República. (1 de Julio de 2021c). *Reglamento del Congreso de*

la República. Obtenido de Portal de la Biblioteca César Vallejo del

Congreso de la República:

<https://www.congreso.gob.pe/Docs/files/reglamento/reglamento-2021-01-07-2021.pdf>

Delgado Guembes, C. (2020). *Introducción al estudio del Parlamento Peruano*.

Lima: Editorial Universidad San Ignacio de Loyola.

Fundación Directorio Legislativo. (Agosto de 2018). *Directorio Legislativo*.

Obtenido de Página web de Directorio Legislativo:

<https://directoriolegislativo.org/>

García Montero, M. (30 de Mayo de 2007). *La actividad legislativa en América*

Latina: sobre el papel reactivo y proactivo de Presidentes y Parlamentos.

Obtenido de Página web del Instituto de Iberoamérica de la Universidad de Salamanca:

http://americo.usal.es/oir/legislatina/papers/Garcia_LA.pdf

Gershenson, C. (24 de Marzo de 2020). *Pensamiento científico*. Obtenido de Pagina web de Coursera: <https://www.coursera.org/learn/ciencia>

Glave Remy, M., Quintanilla Chacón , A. E., Huilca Flores, I. I., Foronda Farro, M. E., Pariona Tarqui, T. E., & Dammert Ego Aguirre, M. E. (16 de Agosto de 2016). *Proyectos de ley*. Obtenido de Portal del Congreso de la República:

<https://www2.congreso.gob.pe/Sicr/TraDocEstProc/CLProLey2016.nsf/sicr/tradocestproc/CLProLey2016.nsf/debusqueda/52143A9B23A2F4330525801100790DD2?opendocument>

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge: Editorial The MIT Press.

Google Developers. (Abril de 2019). *Glosario sobre aprendizaje automático*.

Obtenido de Página web de Google Developers:

<https://developers.google.com/machine-learning/crash-course/glossary?hl=es>

Haro Carranza, J. (2012). Análisis Costo Beneficio. Sincerando las propuestas legislativas en el Perú. *Cuadernos Parlamentarios*, 16-30.

Haykin, S. (2016). *Neural networks and learning machines*. New Jersey: Editorial Pearson Education.

Huertas Hernández, S. A. (1 de Febrero de 2019). *¿Qué factores explican la productividad legislativa? Un análisis a Ecuador y Perú*. Obtenido de Página web de Repositorio Flacsoandes:

<https://repositorio.flacsoandes.edu.ec/xmlui/handle/10469/15479>

Hurtado de Barrera, J. (2010). *Metodología de la investigación: guía de una comprensión holística de la ciencia*. Caracas: Quirón Ediciones.

- Hurtado de Barrera, J. (2014). *El proyecto de la investigación. Comprensión holística de la metodología y la investigación*. Caracas: Ediciones Quirón.
- IBM. (2020). *Introducción al CRISP-DM*. Obtenido de Página web de IBM: <https://www.ibm.com/docs/es/spss-modeler/SaaS?topic=guide-introduction-crisp-dm>
- Iglesias Rodríguez, E., García Zaballos, A., Puig Gabarró, P., & Benzaqué, I. (2020). *Inteligencia artificial. Gran oportunidad del siglo XXI. Documento de reflexión y propuesta de actuación*. Obtenido de Página web del Banco Interamericano de Desarrollo: <https://publications.iadb.org/publications/spanish/document/Inteligencia-artificial-Gran-oportunidad-del-siglo-XXI-Documento-de-reflexion-y-propuesta-de-actuacion.pdf>
- Kaur, A., & Bozié, B. (2019). Convolutional Neural Network-based Automatic Prediction of Judgments of the European Court of Human Rights. Dublin, Dublin, Irlanda.
- Khashman, Z., & Khashman, A. (2016). Anticipation of political party voting using artificial intelligence. 115-120. Lefkosa, Mersin, Turkey.
- Kornilova, A., Argyle, D., & Eidelman, V. (21 de 5 de 2018). Enhancing Legislative Embeddings with Author Attributes for Vote Prediction. Columbia, Washington, Estados Unidos.
- Kumar, R. (2019). *Research Methodology, a step by step guide for beginners*. London: SAGE Publications Ltd.
- Loevinger, L. (1949). Jurimetrics. The Next Step Forward. *Minnesota Law Review*, 455-493.
- López, A. (11 de Junio de 2018). *Jurimetría, la estadística al servicio del Derecho*. Obtenido de Página web de Tiempo Judicial:

<http://tiempojudicial.com/2018/06/11/jurimetria-la-estadistica-al-servicio-del-derecho/>

Management Solutions. (24 de Julio de 2018). *Machine Learning, una pieza clave en la transformación de los modelos de negocio*. Obtenido de Página web de Managementsolutions:
<https://www.managementsolutions.com/sites/default/files/publicaciones/es/machine-learning.pdf>

Nicolini Del Castillo, M. (2013). *Glosario de términos parlamentarios peruanos*. Lima: Fondo Editorial Congreso de la República.

Perez Paredes, Y. J. (1 de Diciembre de 2020). *La evaluación de la ley: ¿etapa final del procedimiento legislativo o parte del contenido de la técnica legislativa?* Obtenido de Página web de la revista Cuadernos Parlamentarios del Congreso de la República :
[https://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/76E4141387F2CEB505257EA800770757/\\$FILE/68_pdfsam_ccep_11.pdf](https://www2.congreso.gob.pe/sicr/cendocbib/con4_uibd.nsf/76E4141387F2CEB505257EA800770757/$FILE/68_pdfsam_ccep_11.pdf)

Shmueli, G., Bruce, P. C., Stephens, M. L., & Patel, N. R. (2017). *Data mining for business analytics*. New Jersey: Jhon Wiley & Sons, Inc.

Sonderegger, D. (24 de Abril de 2018). *Can You Predict What Congress Will Do?* Obtenido de Página web de Above The Law:
<https://abovethelaw.com/2018/04/can-you-predict-what-congress-will-do/>

Stanford University's Computer Science Department. (2020). *Neural networks history*. Obtenido de Página web de la sección de Ciencias de la Computación de la Universidad de Stanford:
<https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/History/history1.html>

Supo Condori, J. A., & Zacarías Ventura, H. R. (2020). *Metodología de la Investigación Científica*. Arequipa: Editorial Bioestadístico EEDU E.I.R.L.

- The Royal Society. (Abril de 2017). *Machine learning: the power and promise of computers that learn by example*. Obtenido de Página web de la Royal Society: <https://royalsociety.org/~media/policy/projects/machine-learning/publications/machine-learning-report.pdf?la=en>
- United States Congress. (20 de Abril de 2021). *Tracking the United States Congress*. Obtenido de Página web del United States Congress: <https://www.govtrack.us/congress/bills/statistics>
- Universitat Oberta de Catalunya. (2020). *Metodologías y estándares por Jordi Gironés Roig*. Obtenido de Portal web de la Universitat Oberta de Catalunya: http://openaccess.uoc.edu/webapps/o2/bitstream/10609/71345/4/Business%20analytics_M%C3%B3dulo%203_Metodolog%C3%ADas%20y%20est%C3%A1ndares.pdf
- Véliz Capuñay, C. (2020). *Aprendizaje automático, Introducción al aprendizaje profundo*. Lima: Fondo Editorial de la Pontificia Universidad Católica del Perú.
- Wolters Kluwer. (15 de 10 de 2020). *Jurimetría*. Obtenido de Página web de Wolters Kluwer: <https://www.wolterskluwer.es/productos/bases-de-datos/jurimetria.html>
- Zevallos Salazar, R. J. (2017). *Predicción del rendimiento académico mediante redes neuronales UNAC*. Obtenido de Página web del Repositorio Institucional Digital: <http://repositorio.unac.edu.pe/handle/UNAC/2728>
- Zolghadr, M., Niaki, S. T., & Akhavan Niaki, S. A. (2016). Modeling and forecasting US presidential election using learning. Morgantown, Virginia Occidental, Estados Unidos.

ANEXOS

Tabla 29

Matriz de consistencia.

TEMA: REDES NEURONALES ARTIFICIALES Y LA PREDICCIÓN DE LA VIABILIDAD LEGISLATIVA EN LAS PROPOSICIONES PARLAMENTARIAS, PERÚ, AÑO 2020						
PROBLEMA	OBJETIVO	HIPÓTESIS	VARIABLES	DIMENSIÓN	INDICADOR	METODOLOGÍA
<u>PROBLEMA GENERAL</u> ¿En qué medida las redes neuronales artificiales predicen la viabilidad legislativa en las proposiciones parlamentarias, Perú, año 2020?	<u>OBJETIVO GENERAL</u> Predecir la viabilidad legislativa aplicando las redes neuronales artificiales en las proposiciones parlamentarias, Perú, año 2020.	<u>HIPÓTESIS GENERAL</u> No aplica	<u>VARIABLE EXÓGENA</u> Redes neuronales artificiales	Fase 2	F2.1, F2.2, F2.3, F2.4	<u>TIPO DE INVESTIGACIÓN</u> Básica
				Fase 3	F3.1, F3.2, F3.3, F3.4, F3.5	<u>NIVEL DE INVESTIGACIÓN</u> Predictivo
<u>PROBLEMAS DERIVADOS</u> ¿Se ha determinado la estructura de la información para predecir la viabilidad legislativa en las proposiciones parlamentarias?	<u>OBJETIVOS ESPECÍFICOS</u> Determinar la estructura de la información mediante la fase CRISP-DM de comprensión de los datos para predecir la viabilidad legislativa en las proposiciones parlamentarias.	<u>HIPÓTESIS DERIVADAS</u> No aplica		Fase 4	F4.1, F4.2, F4.3, F4.4	<u>MÉTODO DE LA INVESTIGACIÓN</u> Método inductivo
¿Se ha determinado la calidad de la información para predecir la viabilidad legislativa en las proposiciones parlamentarias?	Determinar la calidad de la información mediante la fase CRISP-DM de preparación de los datos para predecir la viabilidad legislativa en las proposiciones parlamentarias.			Fase 5	F5.1, F5.2, F5.3	<u>DISEÑO DE LA INVESTIGACIÓN</u> Transeccional retrospectivo documental de rasgo
¿Se ha determinado la técnica algorítmica para su aplicación para predecir la viabilidad legislativa en las proposiciones parlamentarias?	Determinar la técnica algorítmica para su aplicación mediante la fase CRISP-DM de modelamiento para predecir la viabilidad legislativa en las proposiciones parlamentarias.		<u>VARIABLE ENDÓGENA</u> Predicción de la viabilidad legislativa	Predicción de la viabilidad legislativa	Porcentaje de predicción de la viabilidad legislativa	<u>POBLACIÓN DE ESTUDIO</u> 377 proposiciones parlamentarias
¿Se ha determinado los resultados para predecir la viabilidad legislativa en las proposiciones parlamentarias?	Determinar los resultados mediante la fase CRISP-DM de evaluación para predecir la viabilidad legislativa en las proposiciones parlamentarias.					<u>MUESTRA DE ESTUDIO</u> 377 proposiciones parlamentarias
						<u>TÉCNICA DE RECOLECCIÓN DE DATOS</u> Documentación
						<u>INSTRUMENTO</u> No aplica.
						<u>ESTADÍSTICO</u> Análisis descriptivo y ANNs

Fuente: elaboración propia.