

UNIVERSIDAD NACIONAL DEL SANTA
ESCUELA DE POSGRADO
Programa de Doctorado en Estadística Matemática



UNS
ESCUELA DE
POSGRADO

**Un modelo de Análisis multivariante para
la medición de la pobreza en el Perú**

**Tesis para optar el grado académico de
Doctor en Estadística Matemática**

Autor:

Mg. Caycho Chumpitaz, Carlos Teodoro

Asesora:

Dra. González Castro, Jeanette Baldramina
Código ORCID: 0000-0003-3731-9847
DNI. N° 41493243

Línea de Investigación
Educación Estadística

Nuevo Chimbote - PERÚ
2023



UNS
ESCUELA DE
POSGRADO

CONSTANCIA DE ASESORAMIENTO DE TESIS

Yo, Dra. Jeanette Baldramina González Castro, mediante la presente certifico mi asesoramiento de la Tesis Doctoral titulada:

“UN MODELO DE ANALISIS MULTIVARIANTE PARA LA MEDICION DE LA POBREZA EN EL PERU”, elaborada por el magister Carlos Teodoro Caycho Chumpitaz para obtener el Grado Académico de Doctor en Estadística Matemática en la Escuela de Posgrado de la Universidad Nacional del Santa.

Nuevo Chimbote, Agosto del 2023

.....
GONZÁLEZ CASTRO, JEANETTE BALDRAMINA
ASESORA
CODIGO ORCID: 0000-0003-4661-7447
DNI N°17907323



UNS
ESCUELA DE
POSGRADO

CONFORMIDAD DEL JURADO EVALUADOR

**“UN MODELO DE ANALISIS MULTIVARIANTE PARA LA
MEDICION DE LA POBREZA EN EL PERU”**

TESIS PARA OPTAR EL GRADO DE DOCTOR EN ESTADÍSTICA MATEMÁTICA

Revisado y Aprobado por el Jurado Evaluador:

MOORE FLORES, TEODORO
PRESIDENTE
CODIGO ORCID 0000-0002-1755-3459
DNI N° 32763522

CALDERON YARLEQUÉ, LUIS ALBERTO
SECRETARIO
CODIGO ORCID 0000-0002-6870-5096
DNI N° 40097132

GONZÁLEZ CASTRO, JEANETTE BALDRAMINA
VOCAL
CODIGO ORCID: 0000-0003-4661-7447
DNI N°17907323



UNS
ESCUELA DE
POSGRADO

ACTA DE EVALUACIÓN DE SUSTENTACIÓN DE TESIS

A los dieciocho días del mes de agosto del año 2023, siendo las 11:00 horas, en el aula multimedia N° 01 de la Escuela de Posgrado de la Universidad Nacional del Santa, se reunieron los miembros del Jurado Evaluador conformado por los docentes: Dr. Teodoro Moore Flores (Presidente), Dr. Luis Alberto Calderón Yarlequé (Secretario) y la Dra. Jeanette Baldramina González Castro (Vocal), designados mediante Resolución Directoral N° 169-2023-EPGUNS de fecha 17 de julio 2023, con la finalidad de evaluar la tesis titulada: **UN MODELO DE ANÁLISIS MULTIVARIANTE PARA LA MEDICIÓN DE LA POBREZA EN EL PERÚ**; presentado por el tesista **Carlos Teodoro Caycho Chumpitaz**, egresado del programa de **Doctorado en Estadística Matemática**.

Sustentación autorizada mediante Resolución Directoral N° 186-2023-EPG-UNS de fecha 14 de agosto de 2023.

El presidente del jurado autorizó el inicio del acto académico; producido y concluido el acto de sustentación de tesis, los miembros del jurado procedieron a la evaluación respectiva, haciendo una serie de preguntas y recomendaciones al tesista, quien dio respuestas a las interrogantes y observaciones.

El jurado después de deliberar sobre aspectos relacionados con el trabajo, contenido y sustentación del mismo y con las sugerencias pertinentes, declara la sustentación como: APROBADO asignándole la calificación de: Dieciocho (18).

Siendo las 12:00 horas del mismo día se da por finalizado el acto académico, firmando la presente acta en señal de conformidad.


Dr. Teodoro Moore Flores
Presidente


Dr. Luis Alberto Calderón Yarlequé
Secretario


Dra. Jeanette Baldramina González Castro
Vocal



Recibo digital

Este recibo confirma que su trabajo ha sido recibido por Turnitin. A continuación podrá ver la información del recibo con respecto a su entrega.

La primera página de tus entregas se muestra abajo.

Autor de la entrega:	Carlos Caycho
Título del ejercicio:	PIC_maestría
Título de la entrega:	Un modelo de análisis multivariante para la medición de la ...
Nombre del archivo:	ccaycho.pdf
Tamaño del archivo:	2.87M
Total páginas:	82
Total de palabras:	18,800
Total de caracteres:	97,946
Fecha de entrega:	25-ago.-2023 11:50p. m. (UTC-0500)
Identificador de la entrega...	2151529857

I INTRODUCCIÓN

En el Perú el Instituto Nacional de Estadística e Informática (INEI), es el ente oficial que mide la pobreza en el Perú y lo hace usando un enfoque monetario y de necesidades básicas insatisfechas (NBI) de manera separada para entender un fenómeno social, complejo y multidimensional; sin embargo, existen otros enfoques de medición de pobreza como el enfoque de medición multidimensional que permite una mejor comprensión, porque brinda una mejor visión y más detallada de este fenómeno, dado que se puede identificar a la pobreza con aspectos relacionados con la educación, la salud, la vivienda, entre otros.

No obstante, a la fecha el INEI aún no cuenta con indicadores de pobreza bajo el enfoque de medición multidimensional en su base de estadísticas a diferencia de varios países de América Latina, a pesar de que este enfoque de medición no es algo nuevo. Sin embargo, algunas entidades públicas y privadas han presentado sus propios indicadores de pobreza haciendo uso del enfoque de medición multidimensional como el Banco Central de Reserva del Perú (BCRP) o la Sociedad de Comercio Exterior del Perú (COMEX). Como se observa en la Figura 1.1, existe una gran diferencia entre los indicadores de pobreza monetaria y NBI frente al indicador de pobreza multidimensional. Solo en el 2021, la diferencia entre los indicadores de pobreza monetaria y multidimensional es de un 14%; es decir, alrededor de 4 millones de personas no estarían siendo consideradas pobres en términos económicos, pero sí tendrían carencias y/o brechas en educación, salud y vivienda.

Un modelo de análisis multivariante para la medición de la pobreza en el Perú

INFORME DE ORIGINALIDAD

24%

INDICE DE SIMILITUD

24%

FUENTES DE INTERNET

8%

PUBLICACIONES

11%

TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	bookdown.org Fuente de Internet	5%
2	hdl.handle.net Fuente de Internet	2%
3	diario.latercera.com Fuente de Internet	1%
4	www.inei.gob.pe Fuente de Internet	1%
5	americasolidaria.org Fuente de Internet	1%
6	docplayer.es Fuente de Internet	1%
7	repositorio.minedu.gob.pe Fuente de Internet	1%
8	tesis.unap.edu.pe Fuente de Internet	1%
9	1library.co Fuente de Internet	

DEDICATORIA

**A mis padres:
Laura y Carlos**

AGRADECIMIENTOS

**A mi hermana Laura y
a mi sobrino Carlos
por su gran apoyo con la
revisión y comentarios**

ÍNDICE

RESUMEN	XIII
ABSTRACT	XIV
I INTRODUCCIÓN	15
1.1. Planteamiento del problema de investigación	18
1.2. Formulación del problema de investigación	21
1.3. Objetivos de la investigación.....	21
1.3.1 Objetivo General.....	21
1.3.2 Objetivos específicos.....	21
1.4. Formulación de la hipótesis de la investigación	22
1.5. Justificación e importancia de la investigación	22
1.6. Delimitación del estudio	23
II MARCO TEÓRICO	25
2.1. Antecedentes de la investigación.....	25
2.2. Fundamentos teóricos de la investigación	35
2.3. Marco conceptual.....	41
2.3.1 Medición de la pobreza.....	41
2.3.2 Árboles de Decisión.....	52
III MATERIALES Y MÉTODOS.....	62
3.1. Variables e indicadores de la investigación.....	62
3.1.1 Variable dependiente.....	62
3.1.2 Variables independientes	62
3.2. Métodos de la investigación	68
3.3. Diseño o esquema de la investigación	69
3.4. Población y muestra.....	70
3.4.1 Población objetivo	70

3.4.2	Muestra.....	70
3.5.	Actividades del proceso investigativo.....	72
3.6.	Técnicas e instrumentos de la investigación	73
3.7.	Procedimiento para la recolección de datos.....	73
3.8.	Técnicas de procesamiento y análisis de los datos	74
3.8.1	Procesamiento de datos	74
3.8.2	Archivo de datos.....	74
IV	RESULTADOS Y DISCUSIÓN.....	76
4.1.	Aplicación del análisis de árboles de decisión	76
4.2.	Detalles del análisis CHAID	77
4.3.	Aplicación del análisis CHAID	78
4.4.	Resultados del análisis de árboles de decisión	85
4.4.1	Análisis del riesgo del árbol de decisión	85
4.4.2	Análisis de la clasificación de la muestra con el árbol de decisión	86
V	CONCLUSIONES Y RECOMENDACIONES	88
5.1.	Conclusiones.....	88
5.2.	Recomendaciones.....	89
VI	REFERENCIAS BIBLIOGRÁFICAS Y VIRTUALES.....	90
VII	ANEXOS	97

ÍNDICE DE TABLAS

Tabla 3.1 – Listado de variables independientes del modelo.	63
Tabla 3.2 – Matriz operacional	64
Tabla 3.3 – Hogares entrevistados, por departamento; según área geográfica.....	71
Tabla 3.4 – Hogares entrevistados, por dominio geográfico, según área geográfica;.....	72
Tabla 3.5 – Variables del archivo de datos, nombres y etiquetas.....	74
Tabla 4.1 – Resumen del modelo estadístico.....	78
Tabla 4.2 – Resumen del árbol de decisión	79
Tabla 4.3 – Ganancias por nodos del modelo	84
Tabla 4.4 – Riesgo del modelo	85
Tabla 4.5 – Clasificación del modelo	86

ÍNDICE DE FIGURAS

Figura 1.1 – Evolución de la pobreza en el Perú en el periodo 2011-2021...	16
Figura 1.2 – Objetivos de Desarrollo Sostenible (ODS)	17
Figura 2.1 – Dimensiones de la Pobreza Multidimensional según indicadores y ODS de impacto	15
Figura 2.2 – Enfoques para la asignación de peso en indicadores.....	17
Figura 2.3 – Teorías sobre pobreza: Conductual, Estructural y Política.....	26
Figura 2.4 – Clasificación de las técnicas de la minería de datos.....	41
Figura 4.1 – Árbol de decisión.....	67

ÍNDICE DE ANEXOS

Anexo 1 – Sintaxis para la ejecución de árbol de decisión.	85
Anexo 2 – Reglas obtenidas en el proceso de generación del árbol de decisión	87

RESUMEN

Esta investigación tuvo como propósito construir un modelo de análisis multivariante para determinar si un hogar es pobre o no pobre, a partir de 21 variables obtenidas de la base de datos Sumaria de la Encuesta Nacional de Hogares del 2018 (ENAH0-2018) levantada por el Instituto Nacional de Estadística e Informática (INEI) del Perú. Para la construcción del modelo se usó el enfoque de medición multidimensional de la pobreza y la técnica de análisis estadístico multivariante árboles de decisión. Las 21 variables independientes seleccionadas pertenecen a las dimensiones: características de los miembros del hogar, características de los servicios de la vivienda, escolaridad de los miembros del hogar, combustible para cocinar, activos del hogar, e ingresos y gastos del hogar; y la condición de pobreza como variable dependiente pobreza_2 que tiene las categorías de los hogares: 1 = "pobre" y 2 = "no pobre". En este sentido, se estaría contribuyendo con uno de los Objetivos de Desarrollo Sostenible (ODS), el ODS 1: Fin de la Pobreza. Se encontró que el modelo de análisis multivariante consideró a las variables: gasto monetario per cápita mensual, número de miembros del hogar y si la vivienda del hogar tiene piso de tierra, como las variables que determinan si un hogar es pobre o no pobre.

Palabras clave: Pobreza Multidimensional, análisis multivariante, ODS 1, Perú

ABSTRACT

The purpose of this research was to build a multivariate analysis model to determine if a household is poor or not poor, based on 21 variables obtained from the Summary database of the 2018 National Household Survey (ENAH0-2018) collected by the Instituto Nacional de Estadística e Informática (INEI) of Peru. For the construction of the model, the multidimensional poverty measurement approach and the multivariate decision trees statistical analysis technique were used. The 21 selected independent variables belong to the dimensions: characteristics of household members, characteristics of housing services, schooling of household members, cooking fuel, household assets, and household income and expenses; and the poverty condition as a dependent variable `pobreza_2` that has the household categories: 1 = "poor" and 2 = "not poor". In this sense, it would be contributing to one of the Sustainable Development Goals (SDG), SDG 1: End of Poverty. It was found that the multivariate analysis model considered the variables: monthly per capita monetary expenditure, number of household members, and whether the household's dwelling has a dirt floor, as the variables that determine whether a household is poor or not poor.

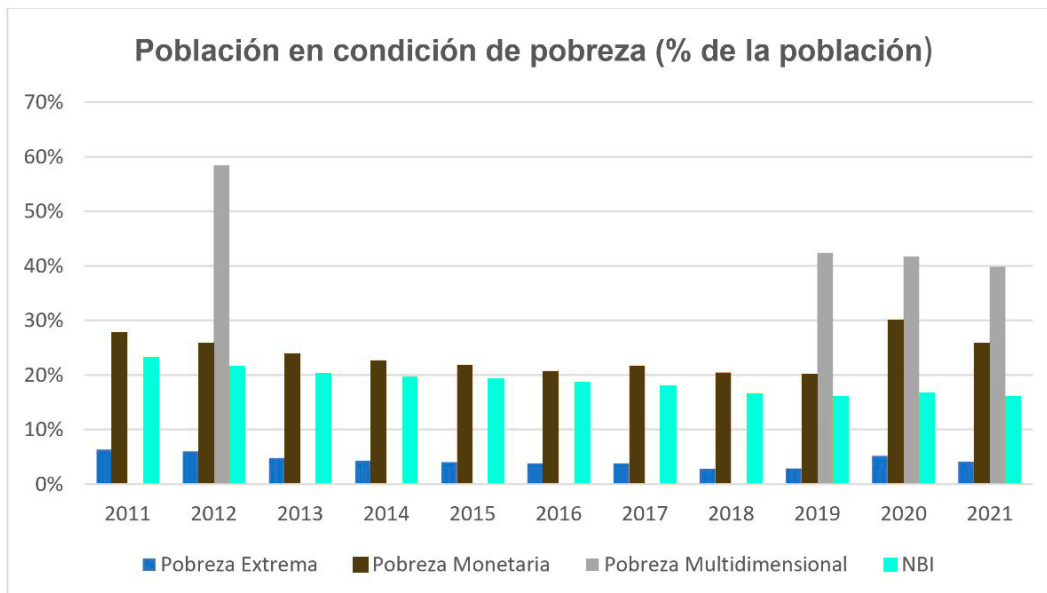
Keywords: Multidimensional Poverty, multivariate analysis, SDG 1, Peru

I INTRODUCCIÓN

En el Perú el Instituto Nacional de Estadística e Informática (INEI), es el ente oficial que mide la pobreza en el Perú y lo hace usando un enfoque monetario y de necesidades básicas insatisfechas (NBI) de manera separada para entender un fenómeno social, complejo y multidimensional; sin embargo, existen otros enfoques de medición de pobreza como el enfoque de medición multidimensional que permite una mejor comprensión, porque brinda una mejor visión y más detallada de este fenómeno, dado que se puede identificar a la pobreza con aspectos relacionados con la educación, la salud, la vivienda, entre otros.

No obstante, a la fecha el INEI aún no cuenta con indicadores de pobreza bajo el enfoque de medición multidimensional en su base de estadísticas a diferencia de varios países de América Latina, a pesar de que este enfoque de medición no es algo nuevo. Sin embargo, algunas entidades públicas y privadas han presentado sus propios indicadores de pobreza haciendo uso del enfoque de medición multidimensional como el Banco Central de Reserva del Perú (BCRP) o la Sociedad de Comercio Exterior del Perú (COMEX). Como se observa en la Figura 1.1, existe una gran diferencia entre los indicadores de pobreza monetaria y NBI frente al indicador de pobreza multidimensional. Solo en el 2021, la diferencia entre los indicadores de pobreza monetaria y multidimensional es de un 14%; es decir, alrededor de 4 millones de personas no estarían siendo consideradas pobres en términos económicos, pero si tendrían carencias y/o brechas en educación, salud y vivienda.

Figura 1.1
Evolución de la pobreza en el Perú en el periodo 2011-2021



Nota. Los indicadores de Pobreza Extrema, Pobreza Monetaria y NBI fueron extraídos de las estadísticas sociales del INEI, mientras que el índice de Pobreza Multidimensional fue extraído de la Sociedad de Comercio Exterior del Perú (COMEX).

Como se indicó líneas arriba existen diversos enfoques metodológicos para la medición de la pobreza que permiten un mejor entendimiento sobre la pobreza. De acuerdo con Pérez-Campos y Rodríguez Saldarriaga (2015) se tiene: enfoque monetario, enfoque multidimensional, enfoque de necesidades básicas insatisfechas, enfoque subjetivo, enfoque relativo y enfoque mixto integrado.

La propuesta de este trabajo de investigación fue obtener un modelo para la medición de la pobreza en el Perú, utilizando el enfoque de medición multidimensional de la pobreza, metodología propuesta por Alkire y Foster (2008), y la técnica de análisis multivariante arboles de decisión, considerando 21 variables independientes y la condición de pobreza como variable dependiente que tiene las categorías de los hogares: 1 = “pobre” y 2 = “no pobre” de la base de datos Sumaria de la ENAHO 2018.

No obstante, al igual que otras investigaciones para medir la pobreza bajo el enfoque multidimensional, existen limitaciones en los datos dada la complejidad de conseguir información sobre educación, salud, seguridad y nutrición de la misma población. En nuestro caso, la base de datos Sumaria de la ENAHO 2018 no cubre aspectos sobre salud, seguridad y nutrición.

Figura 1.2
Objetivos de Desarrollo Sostenible (ODS)



Nota. Información disponible en la página de las Naciones Unidas.
<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>

Por otro lado, dada la creciente preocupación por parte de los gobiernos, empresas, la academia y miembros de la sociedad civil por la sostenibilidad, es necesario incluir a las universidades en la promoción de la investigación y educación en las diferentes ramas de estudio que dispongan alineadas a este nuevo paradigma que busca un balance entre lo económico, lo social y lo ambiental. La sostenibilidad como tal puede ser estudiada por medio de los Objetivos de Desarrollo Sostenible (ODS) de las Naciones Unidas como se muestran en la Figura 1.2. Además, la literatura indica que el enfoque de medición multidimensional de la pobreza se ha ido consolidando en los últimos

años en parte por los ODS, y porque algunos países usan estos mismos objetivos para desarrollar sus propios indicadores.

De manera desagregada específicamente el incluir el enfoque de la pobreza multidimensional a nivel nacional estaría contribuyendo con el Objetivo 1: Fin de la Pobreza, y sobre todo con el Objetivo 1.2. el cual busca erradicar al menos a la mitad de la proporción de la población que vive en situaciones de pobreza en todas sus dimensiones (UN,2015). De igual manera dicho enfoque también contribuye a otros ODS en base a la dimensión en que se encuentran como salud (ODS 3), educación (ODS 4) y vivienda (ODS 11).

Por ello, además de estudiar la pobreza bajo un enfoque de medición multidimensional haciendo uso del análisis multivariante, se está contribuyendo a la investigación científica al abordar una nueva metodología de como medir la pobreza en el país, también se permitiría analizar cuáles son las principales variables que se necesitan abordar a manera de política de Estado para reducir las tasas de pobreza de una manera sostenible y multidimensional. Además de servir como un referente para futuras investigaciones que busquen medir la pobreza multidimensional con análisis estadísticos diferentes como el análisis multivariante, que es un conjunto de técnicas estadísticas que permiten analizar simultáneamente una serie de datos y/o múltiples variables de cada individuo o unidad de investigación. En este trabajo usamos la técnica multivariante, árboles de decisión, que son algoritmos que se aplican sobre un conjunto de datos a fin de obtener un modelo (Berlanga et al., 2013; Pérez López y Santin Gónzales, 2007).

1.1. Planteamiento del problema de investigación

En la región de América Latina, la pobreza no es que no se haya considerado como un fenómeno multidimensional, sino todo lo contrario,

pues según Clausen Lizárraga (2019) en el Perú como en el resto de los países de Latinoamérica se ha ido incluyendo indicadores de NBI con la finalidad de hacer un mejor seguimiento sobre las condiciones de educación, salud y vivienda de la población; además del enfoque monetarista en la medición de la pobreza. En los últimos años, el INEI ha estado emitiendo informes como el “Perú: Perfil de la Pobreza por Dominios Geográficos” donde expone diferentes clases de estadísticas que buscan medir la incidencia de pobreza en el país no solo desde el enfoque monetario ni de las NBI ya que incluye información complementaria como empleo, vivienda, gasto e ingreso per cápita, entre otros.

Asimismo, a pesar de que el enfoque multidimensional de la pobreza no es algo nuevo en el mundo académico, el gobierno peruano ha ido sumando esfuerzos desde hace años para incluir esta metodología en la forma en cómo se mide la pobreza en el Perú. Por ejemplo, en el 2016 se celebró el “Taller de Medición de la Pobreza Multidimensional”, participando los miembros del INEI, la Comisión Consultiva para la Medición de la Pobreza, el Banco Mundial y *Oxford Poverty and Human Development Initiative* (OPHI) de la Universidad de Oxford. Dicho taller tuvo como objetivo la presentación de la metodología Alkire-Foster y del Índice de Pobreza Multidimensional (IPM) de la OPHI y su relación con los ODS, donde el INEI y la OPHI terminarían suscribiendo un convenio de cooperación (INEI, 2016). No obstante, el INEI hasta la fecha no ha publicado ningún tipo de indicador de pobreza multidimensional.

Adicionalmente, por medio del Ministerio de Desarrollo e Inclusión Social (MIDIS) en los últimos años ha ido promoviendo la inclusión del enfoque multidimensional a la mesa. En 2021, el MIDIS compartió al INEI el informe "Transversalización del Enfoque de Pobreza Multidimensional en el Proceso de Actualización de la Política Nacional de Desarrollo e Inclusión Social", con el objetivo de brindar las métricas y herramientas

necesarios para iniciar la medición bajo este enfoque. Un año más tarde, estaría anunciando un estudio con sus contrapartes de los países de la Alianza del Pacífico sobre la “Pobreza multidimensional y los efectos de la pandemia por la COVID-19 en el crecimiento económico y las brechas de género” (MIDIS,2021; MIDIS,2022). Por lo que si bien, se puede observar una mayor acción estatal por promover la metodología multidimensional tanto a nivel nacional como internacional, los esfuerzos pueden verse afectados por la inestabilidad política que han aquejado al país en los últimos dos años.

Desde el BCRP, Huaranca y Castillo (2021) proponen un indicador de pobreza multidimensional empleando la metodología Alkire-Foster por medio de 6 dimensiones en base al IPM Global y al IPM implementado por otros países de Latinoamérica como Chile, Colombia, Ecuador y Panamá. Mientras que, desde el sector privado COMEX (2021) también emplea la misma metodología, aunque solo usa 3 dimensiones o pilares de la ENAHO. No obstante, al medir la pobreza a través de índices hace que ambos indicadores se encuentran sujetos a las limitaciones intrínsecas de este enfoque de medición de la pobreza multidimensional; es decir, reducir la pobreza a solo un índice en vez de un tablero de indicadores independientes o la selección de dimensiones e indicadores.

En ese contexto, la presente investigación tiene el propósito de buscar un modelo que determine qué variables y qué valores de estas variables pueden calificar a un hogar como pobre o no pobre. Para ello se usó la metodología o enfoque de medición de pobreza multidimensional. El enfoque multidimensional incluye varias dimensiones como son: las características de los miembros del hogar, las características de la vivienda, la escolaridad de los miembros del hogar, el combustible para cocinar, los activos del hogar, e ingresos y gastos del hogar. Este enfoque da una visión más integral del fenómeno de la pobreza pues brinda un

panorama más realista en la medición de la pobreza (Ximenes et al., 2016).

Como herramienta para la obtención del modelo se utiliza la técnica del análisis multivariante árboles de decisión. Entre las facilidades de esta técnica es que el análisis es rápido, porque cuantifica el costo y las probabilidades de los diferentes resultados obtenidos, hecho que ayuda en la toma de decisiones (Zuniga y Abgar, 2011). En ese sentido, nuestro fundamento teórico para la medición de la pobreza está basado en el enfoque de medición de la pobreza multidimensional y como herramienta para determinar el modelo de medición, la técnica de árboles de decisión.

1.2. Formulación del problema de investigación

¿Será que el modelo de análisis estadístico multivariado nos permitirá que en los hogares peruanos estimar la condición en “pobres” y “no pobres”?

1.3. Objetivos de la investigación

1.3.1 Objetivo General

Determinar un modelo de análisis estadístico multivariado que permita estimar en los hogares peruanos la condición de “pobre” o “no pobre”.

1.3.2 Objetivos específicos

1. Determinar las variables predictoras más importantes para el modelo; esto es, las variables que permitan clasificar adecuadamente si un hogar es “pobre” o “no pobre” en el Perú.
2. Determinar la proporción de error del modelo construido.
3. Determinar la probabilidad de acierto del modelo construido.

1.4. Formulación de la hipótesis de la investigación

El modelo de análisis estadístico multivariado nos permite estimar si los hogares peruanos tienen la condición de pobre o no pobre.

1.5. Justificación e importancia de la investigación

Esta investigación usa el enfoque de medición multidimensional de la pobreza y la técnica multivariante árboles de decisión, vínculo que no ha sido tratado, hasta la fecha, en el Perú.

Al utilizar el análisis de datos multivariantes, específicamente la técnica de árboles de decisión nos permitió extraer información “oculta” de cada una de las 21 variables predictoras alineadas a la metodología Alkire-Foster (AF) y contribuyen a determinar las variables con mayor importancia para un grupo de estudio en específico. Esta ventaja vendría a ser de interés para las áreas de debate que giran en torno a la medición de la pobreza multidimensional, pues de manera estadísticas se pueden elegir las variables más importantes a diferencia de los índices actuales de pobreza multidimensional que asignan pesos a cada variable.

En ese sentido, al determinar con mayor exactitud la medición de la pobreza en el Perú, ayudará a los hacedores de políticas para que redireccionen los programas sociales y que éstos lleguen a los lugares donde los hogares realmente lo necesiten.

El modelo de análisis multivariante que se ha obtenido para la medición de la pobreza permitirá difundir las técnicas multivariantes y la aplicación de las mismas es una medición objetiva de la pobreza. Es decir, la importancia está en difundir las técnicas multivariantes que dan una lectura menos economista de la medición de la pobreza y mucho más sistémica y objetiva, lo cual podría tener un impacto en cómo se deben desarrollar las políticas públicas que buscan reducir los niveles de incidencia de pobreza multidimensional.

Asimismo, existe a la fecha limitada literatura sobre la pobreza multidimensional en el Perú y mucho menos que empleen el análisis estadístico multivariante, por lo que la presente investigación busca ser un referente para futuras investigaciones en el tema; así como proponer una metodología alternativa. En ese sentido, de manera más amplia y como se ha mencionado anteriormente, se estaría también contribuyendo con los ODS, y en especial al ODS 1: Fin de la Pobreza, el enfoque multidimensional también abarca dimensiones educativas, de salud y calidad de vida que incluirían a los ODS 4: Educación de Calidad, ODS 8: Crecimiento Económico y Empleo Decente, ODS 10: Reducción de las Desigualdades.

1.6. Delimitación del estudio

En este trabajo de investigación se desarrolló haciendo uso del enfoque de medición multidimensional de la pobreza; y al igual que otras

investigaciones para medir la pobreza bajo este enfoque, existen limitaciones en los datos dada la complejidad de conseguir información sobre educación, salud, seguridad y nutrición de la misma población. En nuestro caso, la base de datos Sumaria de la ENAHO 2018 no cubre aspectos sobre salud, seguridad y nutrición. Por ello, se recomienda que en futuras investigaciones que empleen la misma metodología para estudiar la pobreza en el Perú deberán tomar en consideración las limitaciones de la data de la ENAHO.

II MARCO TEÓRICO

2.1. Antecedentes de la investigación

En la revisión de literatura de este trabajo de investigación, se llegó a la conclusión de que la metodología utilizada para la medición de la pobreza tiene diversos enfoques. De acuerdo con Pérez-Campos y Rodríguez Saldarriaga (2015) se tiene: enfoque monetario, enfoque multidimensional, enfoque de necesidades básicas insatisfechas, enfoque subjetivo, enfoque relativo y enfoque mixto integrado.

Cuando hablamos del Índice de Pobreza Multidimensional, el PNUD (2011) explica que es un indicador que tiene como objetivo medir los graves déficits en las dimensiones de la salud, educación y estándares de vida, así como tener en cuenta tanto a la población desfavorecida como la intensidad de las privaciones de sus necesidades básicas. Según Clausen Lizárraga (2019) desde fines del siglo XX a nivel global el interés por el concepto de pobreza multidimensional, así como los esfuerzos para comprenderla, medirla y eliminarla, han ido cobrando fuerza. Esto se debe en parte al trabajo articulado para medir la pobreza en una escala mundial usando indicadores como el índice de Pobreza Humana (IPH) y posteriormente el establecimiento de lineamientos de las Naciones Unidas, como los Objetivos del Milenio (ODM) en el 2000 y los Objetivos del Desarrollo Sostenible (ODS) en el 2015, los cuales han permitido que se consolide el interés durante las últimas décadas. Diferentes países de ingresos bajos y medios como Armenia, Chile, Colombia, Ecuador, Filipinas, México, Mozambique, Pakistán y Vietnam han desarrollado e

implementado medidas multidimensionales de la pobreza en el marco de los ODS (Zavaleta et al., 2018)

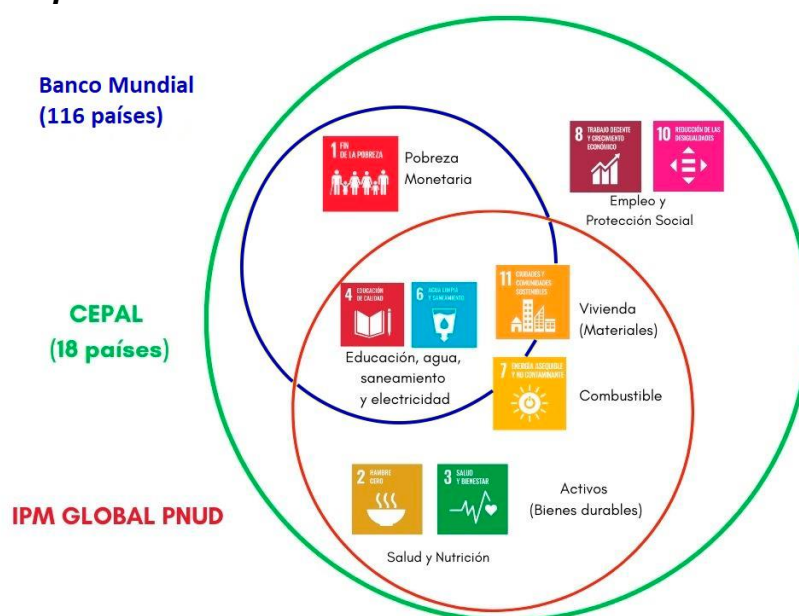
Del mismo modo el Programa de las Naciones Unidas para el Desarrollo (PNUD) ha promovido el diseño de nuevos indicadores que permitan analizar y estudiar de manera multidimensional el desarrollo humano como el Índice de Desarrollo Humano (IDH), por ello en el 2010 desarrolló juntamente con el OPHI el Índice de Pobreza Multidimensional (IPM-G). Este indicador tiene como objetivo consolidarse como una medida de pobreza aguda no monetaria que permita hacer un análisis comparativo entre economías de ingresos bajos y medios (Alkire y Santos, 2014). No obstante, la revisión de literatura indica que fue a inicios del siglo XXI donde las investigaciones como las de Bourguignon y Chakravarty (2003) y Atkison (2003) sobre la medición multidimensional de la pobreza sentaron una base para medir las interrelaciones entre las diferentes dimensiones de la pobreza, usando como referencia al índice de Pobreza Foster Greer Throbecke (FGT) y el enfoque de conteo de pobreza que se había implementado en Europa.

Asimismo, otras divisiones de las Naciones Unidas también tuvieron otras propuestas en cuanto a la metodología de la medición multidimensional de la pobreza. Por ejemplo, la Comisión Económica para América Latina y el Caribe (2014) propuso un índice de pobreza multidimensional específicamente para la realidad latinoamericana, usando información de las Necesidades Básicas Insatisfechas de 18 países de la región.

De igual forma, el Fondo de las Naciones Unidas para la Infancia, más conocido como UNICEF (2015) tuvo la iniciativa de medir la pobreza multidimensional en las poblaciones más jóvenes a nivel internacional de un grupo de más de 40 países basado en el *Multiple Overlapping Deprivation Analsys* (MODA).

Por otro lado, entidades como el Banco Mundial (2018) también tuvo la iniciativa de medir el nivel de pobreza multidimensional en 116 países con un enfoque distinto al IPM-G del PNUD y el OPHI, para complementar indicadores de pobreza monetarias basados en la línea de pobreza. También incluye un piloto donde se pueden encontrar mayores dimensiones e indicadores para 6 países.

Figura 2.1
Dimensiones de la Pobreza Multidimensional según indicadores y ODS de impacto



Nota. En base al aporte de Clausen Lizárraga (2019). Se añadieron los ODS en cada dimensión según las sinergias que existen con cada uno de los objetivos y metas de la Agenda 2030.
Elaboración propia

Como se observa en la Figura 2.1, el indicador de pobreza de la CEPAL incluye un mayor número de dimensiones, seguida del IPM del PNUD. Al hablar de agua, saneamiento, electricidad, combustible y vivienda se propone que se tomen en cuenta en conjunto los ODS 6: Agua y Saneamiento, ODS 7: Energía Asequible y ODS 11: Ciudades y Comunidades Sostenibles. En el marco de la Agenda 2030, algunos de

los indicadores del IPM-G han ido adaptándose para estar alineados con diferentes objetivos y metas de los Objetivos de Desarrollo Sostenible (Alkire y Jahan, 2018).

Si bien los nuevos indicadores para medir la incidencia de pobreza buscan suplir las carencias del enfoque monetario al incluir diferentes dimensiones al estudio, la metodología y el enfoque empleado no es ajeno al debate académico. De acuerdo con Clausen Lizárraga (2019) y el INEI (2017) existen 6 áreas de discusión en torno a la medición de la pobreza multidimensional: i) sobre la pertinencia de construir un índice de pobreza en vez de un tablero de indicadores, ii) sobre la elección de dimensiones e indicadores, iii) sobre las potencialidades y limitaciones del método Alkire-Foster, iv) sobre la elección de la unidad de identificación y sus implicancias, v) sobre la definición de ponderadores dimensionales, vi) sobre la definición de umbrales de pobreza multidimensional.

Bajo esa línea por ejemplo autores como Ravallion (2011) señalan que para medir la pobreza multidimensional se deberían emplear tableros de control que permita visualizar de manera independiente cada una de las dimensiones de la pobreza, a diferencia de la metodología Alkire-Foster (AF) al medir la pobreza mediante un índice al verse limitada por depender de los datos de las diferentes dimensiones de la pobreza del mismo hogar muestreado. Ante la limitación de contar a los hogares como unidades de análisis, se suman críticas en torno a que un solo indicador no puede identificar las privaciones de cada miembro de la unidad de análisis. La *Australian National University* (2013) tuvo la iniciativa de desarrollar el *Individual Deprivation Measure* (IDM) como un enfoque alternativo que incluye más de 10 dimensiones multidimensionales y ha sido empleada para estudiar a más de 5 países de ingresos bajos y medios de Asia y África.

En cuanto a la selección de dimensiones e indicadores, al igual que sucede en este trabajo de investigación existe información limitada con la que se puede trabajar, más aún si sumas los limitantes mencionados en el párrafo anterior en torno a las unidades de estudio. Un paso más adelante ya habiendo seleccionado las dimensiones e indicadores según disponibilidad, la siguiente crítica vendría a ser el razonamiento detrás de la asignación de ponderaciones específicas para cada dimensión y variables, algo que también vendría a limitar a la metodología AF. El uso de ponderaciones en índices para medir el bienestar social tiene dos posturas opuestas pues existen quienes afirman que las ponderaciones no existen o no reflejan correctamente la realidad de pobreza multidimensional y aquellos quienes opinan lo contrario (Lustig, 2011). De cualquier forma, es innegable que el uso de ponderaciones para los indicadores de pobreza multidimensional podría considerarse ya como una costumbre arraigada en cuánto a la forma de medición de la pobreza al emplear la metodología AF.

Figura 2.2
Enfoques para la asignación de peso en indicadores



Nota. Información extraída del trabajo de Decanq y Lugo (2013).
Elaboración propia.

De manera más detallada, Decanq y Lugo (2013) explican que la asignación de ponderaciones se puede basar en tres enfoques que son *data driven*, los normativos y los híbridos como se puede observar en la Figura 2.2. Los autores indican que dentro de los tres enfoques el menos usado es el híbrido dada su complejidad de estimación, mientras que dentro de los *data driven* los pesos más favorables suponen un reto a la asignación de esquemas de peso diferentes entre distintos individuos de una misma muestra. En contraste el enfoque normativo parece guardar una mayor relación con la metodología AF y el IPM de la OPHI-PNUD dado que toman en cuenta la opinión de expertos, y dependiendo del investigador se asignan pesos iguales o arbitrarios según el número de dimensiones e indicadores que pueden tener sus investigaciones.

Es necesario también mencionar investigaciones que miden la pobreza multidimensional en el Perú como en otros países usando la metodología AF aunque no empleen el análisis multivariante al igual que este trabajo de investigación, con el objetivo de que el presente trabajo sirva como referente para futuras investigaciones que busquen medir la pobreza multidimensional. Como se ha mencionado anteriormente, a nivel nacional la entidad encargada de estimar los niveles de pobreza es el INEI, el cual cuenta con la ayuda de la Comisión Consultiva para la Estimación de la Pobreza desde el 2010. Dicha comisión tiene como objetivo asesorar, evaluar y validar las estimaciones de pobreza y también presentar nuevas metodologías de medición de pobreza, y reúne a académicos, representantes de organismos internacionales, funcionarios públicos, expertos de entidades nacionales, entre otros (Decreto Supremo N° 097-2010-PMC, 2010). Adicionalmente, las metodologías empleadas en el Perú por el INEI para medir la incidencia de pobreza además de emplear el método de las NBI y el enfoque monetario por medio de la Línea de Pobreza (LP), plantea un Método Integrado (MI) que resulta del cruce entre ambos enfoques con el objetivo de suplir las carencias entre ambos indicadores; sin embargo, otros factores multidimensionales de la pobreza

como la educación, salud y vivienda no llegan todavía a ser considerados en un indicador de pobreza de manera integrada.

La relación entre la pobreza monetaria y pobreza multidimensional en niños y jóvenes de Etiopía, India, Perú y Vietnam fue estudiada por Kim (2018) usando la información del *Young Lives Survey* del 2006 y 2009 junto a otros indicadores para armar las dimensiones de pobreza multidimensional. Dentro de los principales resultados el autor resalta que los niños y jóvenes que se mantienen en la pobreza monetaria son menos probables a salir de la pobreza multidimensional; sin embargo, no siempre que se escapa de la pobreza monetaria se logra salir de la pobreza multidimensional. En el caso de Perú esta dinámica entre la pobreza monetaria y multidimensional muestra relación moderada en grupos mayores.

Con un enfoque más sectorial, Clausen y Trivelli (2019) miden la pobreza multidimensional en el país de manera comprensiva y sensible al contexto rural peruano por medio del método de identificación y agregación de IPM-RP. Se emplearon 22 indicadores distribuidos en 9 dimensiones usando como insumos información disponible en la ENAHO del 2018. Los autores indican que los niveles de pobreza multidimensional en zonas rurales con baja densidad poblacional oscilan entre el 35.5% y el 75.7%, mientras en las zonas con mayor densidad poblacional el rango es de 25.5% y el 60.8%.

El trabajo de Clausen y Barrantes (2021) busca conocer como las personas con discapacidad se pueden ver afectadas por la pobreza desde un enfoque multidimensional en el Perú por medio de la metodología AF, donde se tomaron en cuenta 8 dimensiones de las cuales 4 estaban orientadas a indicadores de discapacidad: asistencia médica; educación, trabajo y jubilación; actividades básicas diarias; y conexión social. Los

resultados muestran que el 41.1% de la población con discapacidad en el país sufre necesidades en por lo menos 3 de 8 dimensiones, registrando los mayores niveles de pobreza en personas con discapacidad son las mujeres, también pobladores de zonas rurales y de etnia indígena.

Desde el sector público, Huaranca y Castillo (2021) miden la pobreza multidimensional en base a 6 dimensiones de pobreza como lo son la salud, el acceso a los servicios básicos, el entorno físico, la participación social y económica, y la educación; asignándole pesos diferentes a cada dimensión y dándole mayor importancia a las dimensiones de salud y educación. Además de incluir un análisis según zonas geográficas y determinar qué dimensiones contribuyen más en la pobreza multidimensional. Dentro de los resultados los autores indican que en el país tanto la selva como las zonas rurales son en donde existe mayor incidencia de pobreza multidimensional, mientras que las dimensiones que más contribuyen a este problema son la participación económica y el entorno físico, donde los principales cambios que contribuirían en la erradicación de la pobreza serían mejorar las condiciones de empleo para evitar la precariedad laboral así como mejorar las condiciones de los hogares peruanos como el uso de energía para la cocina y los materiales de construcción de las viviendas.

En la revisión de literatura se puede observar diversos trabajos que han empleado la metodología del análisis multivariante para medir los niveles de pobreza:

El análisis multivariante es una técnica que se está usando mucho de un tiempo a esta parte. De acuerdo con la investigación de Källestål et al. (2019), la cual emplea la metodología de árboles de decisión aplicada a las bases de datos de vigilancia demográfica y sanitaria de Cuatro Santos, al norte de Nicaragua, para identificar las intervenciones necesarias para

la reducción continua de la pobreza, medida por el índice de Necesidades Básicas Insatisfechas (NBI). Los árboles de decisión fueron validados mediante la aplicación de análisis de bosques aleatorios condicionales para evaluar y clasificar la importancia de los predictores sobre su capacidad para explicar la variación del resultado de la "pobreza". La mayoría de los hogares de Cuatro Santos proporcionaron información y las variables incluidas midieron las condiciones de la vivienda, los activos y las experiencias demográficas desde la última actualización (5 años), la participación anterior en las intervenciones y la seguridad alimentaria durante las últimas 4 semanas. Lo que les permitió concluir que los predictores de pobreza son útiles para dirigir las intervenciones y en el área de Cuatro Santos y que es más importante priorizar la educación; así como el desarrollo en entornos rurales pobres similares en todo el mundo.

Por otro lado, el enfoque de medición de la pobreza multidimensional es el que se está usando en estos últimos tiempos. Por ejemplo, según Ximenes et al. (2016) para analizar la pobreza multidimensional y sus aspectos subjetivos en contextos rurales y urbanos en un estado del Brasil, se consideró una muestra de 417 participantes. Lo que le permitió crear índice de medición de pobreza multidimensional por medio de las privaciones en las dimensiones de la vivienda, la educación, el trabajo/ingreso, la salud y los aspectos subjetivos de la pobreza.

En ese mismo sentido, se tiene que de acuerdo con Ordóñez et al. (2018) para identificar las causas fundamentales de la pobreza multidimensional en Ecuador y sus interrelaciones incorporando variables que no son exclusivamente económicas, sino que también factores socioeconómicos, se utiliza el método de ecuaciones estructurales y también se utilizan las técnicas multivariantes. Este trabajo, sostiene que las características sociodemográficas que han sido establecidas y la complejidad de los factores que han sido analizados para medir la pobreza multidimensional son corroborados por la experimentación por el análisis realizado,

mediante las ecuaciones estructurales. Es decir, este trabajo conjuga la medición de la pobreza multidimensional y el análisis multivariante.

La perspectiva para analizar y medir la pobreza, en estos últimos tiempos, vincula el uso del enfoque multidimensional de la pobreza con el análisis estadístico multivariante. Al respecto, de acuerdo con Aguilar Estrada et al. (2018a;2018b) para distinguir las intensidades de la pobreza multidimensional en los distintos municipios más pobres de México se realizó una investigación con métodos estadísticos multivariantes como el análisis de conglomerados; encontrándose tres conglomerados que se utilizaron para poder establecer características específicas de la distribución geoespacial para focalizar los programas de combate contra la pobreza. La metodología para medir la pobreza multidimensional en México muestra que la incidencia es muy elevada pero también heterogénea, a partir de los resultados se establecieron los factores, incidentes principales por la cual la pobreza en México es bastante elevada entre los municipios estudiados.

Otro estudio desde el enfoque multidimensional es el propuesto por Waisgrais et al. (2017) en el que se presentan y discuten los resultados de la pobreza en la niñez en Argentina desde la perspectiva multidimensional. Este estudio cuenta con 10 dimensiones y 28 indicadores para el tratamiento de la información, también se realiza una encuesta de análisis múltiples por conglomerados. De los resultados se han obtenido hallazgos que son robustos a los cambios de umbral dimensional de la pobreza. Por lo tanto, de las 10 dimensiones consideradas cuatro dimensiones explican más del 63%, en su caso serían las privaciones de información, la violencia, la salud y el ocio en Argentina.

Otro trabajo que aborda el tema de pobreza multidimensional y análisis multivariante es el propuesto por Colafranceschi et al. (2011) quienes encuentran que un grupo poblacional se identifica como pobre multidimensional a través de distintas dimensiones. Para esto también se realiza un análisis de conglomerados con el objetivo de realizar una comparación entre los resultados obtenidos del estudio y los que surgen a partir de la línea de pobreza. En dicho estudio, se aplicó la técnicas multivariantes para el análisis de la pobreza multidimensional como el análisis de correspondencias múltiples y el análisis de conglomerados y luego de presentar los resultados se planteó que esta investigación se compare con otras a futuro para poder cotejar los resultados con futuras investigaciones y también informar que no fueron realizadas en el trabajo las líneas de investigación alternativas y considerar que se debería realizar a futuro un análisis de grupos difusos.

En contraste, con el objetivo de construir un indicador de pobreza multidimensional para detectar los principales focos de riesgo de pobreza en España, en el trabajo de Poza Lara y Fernández Cornejo (2010) se usó análisis factorial de segundo orden. Por lo que el estudio ha usado la acepción multidimensional relativo y acorde con los países desarrollados y con un indicador indirecto construido mediante análisis factorial para su posterior aplicación al caso español.

2.2. Fundamentos teóricos de la investigación

Definir la pobreza es bastante complejo, ya que ésta en sí misma es un problema con una importante dimensión política, por los intereses de los diferentes grupos que tienen fuerte influencia en los modelos de distribución y en la existencia de la pobreza (Wilson, 1996). Por lo que ahondar, en las raíces de la pobreza supone plantear cuestiones difíciles

y conflictivas, esto explica las reticencias y los rechazos que acompañan el proceso del conocimiento de ésta (Pérez Muñoz, 2012, pág. 20).

Desde un punto de vista filosófico, la pobreza ha sido una cuestión vinculada a la justicia desde Platón a Thomas Pogge. Además, Aristóteles establece que la pobreza es un mal en sí misma, fuente de conflictos sociales que una sociedad justa debería evitar. Por una parte, el lado histórico de la teoría aristotélica de la pobreza plantea que la democracia paradigmática promueve alguna clase de bienestar. Por otra parte, la faceta utópica expone los pensamientos de Aristóteles acerca de una sociedad ideal donde el crecimiento de la clase media reduce la lucha entre ciudadanos ricos y pobres (Méndez Aguirre, 2011). Aristóteles define a los pobres (*hoi pénētes*) como aquellos que carecen de lo necesario. Los ricos, al contrario de los pobres, sí tienen lo necesario para vivir bien e incluso cuentan con lo superfluo. La dicotomía “ricos-pobres”, como bien observa Sancho Rocher, es una “simplificación”; pero es ampliamente utilizada en la literatura griega arcaica y clásica (Álvarez Yágüez, 2009; Sancho Rocher, 2011), y que además este enfoque monetarista de la pobreza sigue vigente hasta la fecha.

Por otro lado, Dubois (2005) bajo el enfoque utilitarista, señala la concepción dominante en el siglo pasado se basó en un concepto absoluto de pobreza definido a partir de lo que se denomina el umbral de la pobreza. Dicho umbral se determina en función del ingreso o renta necesario para que una persona pueda sobrevivir, y una vez fijado se convierte en la referencia para determinar quienes son pobres; luego que identifican a los pobres, se procede a su cuantificación. Los términos de pobreza y pobre están asociados a un estado de necesidad y carencia; dicha carencia se va relacionar con lo necesario para el sustento de la vida. Lo que significa que la pobreza es una inevitable situación de comparación entre lo observado y una condición normativa (Boltvinik y Hernández Laos, 2000, pág. 30).

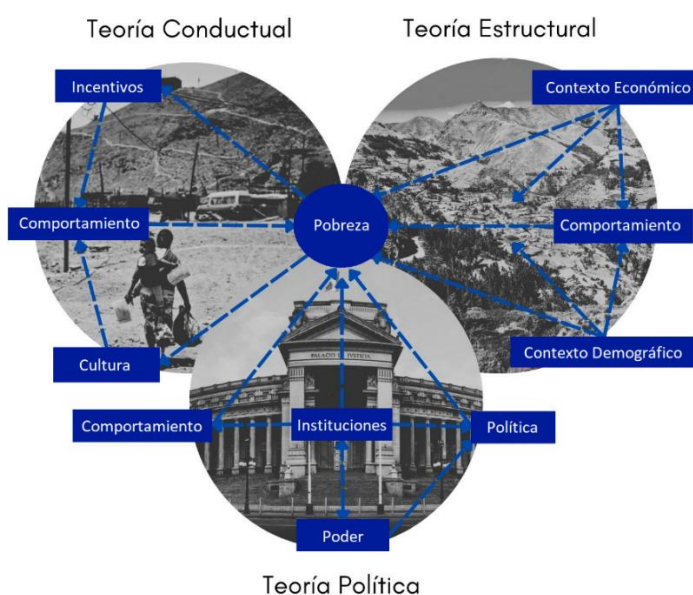
El economista Amartya Sen (1992) señala que antes de dar un concepto de la pobreza hay que definir quién debe de estar en el centro de nuestro interés; en un sentido lógico, la pobreza es una característica de los pobres, como un grupo que son parte de la sociedad, y es en ellos que debemos de centrar nuestro análisis para conocer sus características. Además el autor señala el concepto de pobreza a partir de las capacidades, es decir lo que la gente puede hacer, definiendo a la pobreza como ausencia de capacidades básicas que permiten a cualquier individuo insertarse en la sociedad, a través del ejercicio de su voluntad; entonces la pobreza es la incapacidad de conseguir bienestar precisamente debido a la ausencia de medios. Años atrás Sen (1982) propuso dos métodos para medir la pobreza. En primer lugar, el método directo mide la pobreza según el nivel de satisfacción de un conjunto de necesidades básicas, derechos o en línea con el enfoque de capacidades. En segundo lugar, el método indirecto mide el nivel de pobreza según el criterio de la línea de pobreza; es decir, el nivel de ingresos que una persona debe tener como mínimo para satisfacer sus necesidades básicas. Se ha implementado en medidas oficiales el método indirecto en la mayoría de los países del mundo mientras que el método directo en regiones como Europa y Latinoamérica (Alkire y Santos,2011).

Las investigaciones de Amartya Sen fomentaron que el enfoque monetarista de la pobreza fuera adquiriendo un enfoque más interdisciplinario, lo que sumado a otros factores ya antes mencionados como la Agenda 2030 dieron paso para que el enfoque de pobreza multidimensional se consolide. Por ello la definición oficial del INEI (2000, pág. 1) explica que:

“La pobreza es una condición en la cual una o más personas tienen un nivel de bienestar inferior al mínimo socialmente aceptado. En una primera aproximación, la pobreza se asocia con la incapacidad de las personas para satisfacer sus necesidades

básicas de alimentación. Luego, se considera un concepto más amplio que incluye la salud, las condiciones de vivienda, educación, empleo, ingresos, gastos y aspectos más extensos como la identidad, los derechos humanos, la participación popular, entre otros.”

Figura 2.3
Teorías sobre pobreza: Conductual, Estructural y Política



Nota. Traducido del trabajo de Brady (2019).
Elaboración propia.

En ese contexto, si bien se han hecho avances en cuánto a la medición de la pobreza, Sawhill (1988) señala el estancamiento de la investigación en torno a la pobreza al limitar a esta problemática social al nivel de ingreso de las personas, además del número y composición de aquellos que se encuentran en las colas inferiores de ingresos, pero no tanto a los procesos que generan dichos datos. Según Brady (2019) son tres las principales corrientes teóricas sobre la pobreza, la teoría conductual, la teoría estructural y la teórica política, como se muestra en la Figura 2.3. El autor señala que las futuras investigaciones en pobreza deberían abarcar un enfoque interdisciplinario. De acuerdo con las teorías de la pobreza puede ser explicada ya sea por el comportamiento de la población que se encuentra bajo la influencia de la cultura y los incentivos,

el contexto demográfico y económico, la política y las instituciones; por lo que las mediciones de pobreza deberían considerar estas diferentes causalidades y no limitarse solo al nivel de ingreso.

En primer lugar, tenemos a la teoría conductual que señala tanto que los incentivos y la cultura causan el comportamiento el cual influye en la pobreza, que al mismo tiempo tiene una causalidad en los incentivos y la cultura. Es decir, el comportamiento es el principal factor que causaría la pobreza, donde las personas son pobres debido a que toman decisiones contraproducentes que aumentan el riesgo de incrementar la pobreza como el ser madre soltera o el desempleo (Durlauf, 2006). Según Brady (2019) algunos de los desafíos que tiene este enfoque vendrían a ser que no se toma en cuenta otros factores que determinan la pobreza que no sean los incentivos, la cultura y el comportamiento. Asimismo, a diferencia de otras teorías de pobreza la teoría del comportamiento a veces se encuentra entre la línea de causalidad de pobreza y de tener un rol en hacer posible o probable la pobreza debido a la falta de diseño de investigaciones comparativas que sustenten mejor la relación de causalidad de pobreza (Small et al. 2010; Streib et al., 2016).

En segundo lugar, la teoría estructural tanto en el contexto económico como en el contexto demográfico indirectamente causan la pobreza del comportamiento, que a su vez actúan como mediador entre la relación causal comportamiento-pobreza. Cuando se habla del contexto económico incluye al crecimiento económico y el desarrollo industrial, entre otros; mientras que el contexto demográfico alude a las disparidades entre hogares, composición de sexo y edad, segregación por lugar de residencia y urbanización (Brady, 2019). Al igual que la anterior teoría el enfoque estructural no es ajeno a críticas, ya que como señala Contreras (2012) se puede considerar al contexto estructural como una causa subyacente de la pobreza, también se puede considerar que la cultura es un mecanismo entre el contexto estructural y el comportamiento.

Por último, la teoría política explica que tanto el poder como las instituciones causan la política que a su vez causa pobreza y modera la relación entre comportamiento-pobreza. Dada la importancia de las instituciones en este enfoque, la calidad de la institucionalidad de un país es un factor determinante para controlar la pobreza. Huber y Stephen (2001) indican que una débil institucionalidad guía el cuándo y el cómo las políticas públicas pueden dar forma a la pobreza; en contraste al hecho de contar con una institucionalidad fuerte, sería sinónimo de que las instituciones públicas consolidadas están por encima de las políticas de turno para reducir los niveles de pobreza en forma más asertiva.

Estas tres teorías que explican las causas de la pobreza reafirman la necesidad de desentenderse de un enfoque monetarista y reduccionista sobre la pobreza y como medirla, además de hacer eco dentro de las áreas de debate de la medición de la pobreza multidimensional.

Por otro lado, desde la ciencia estadística se deben buscar nuevas y mejores formas que nos permitan entender mejor la pobreza multidimensional, donde los principales indicadores no son ajenos a críticas válidas por parte de la comunidad académica. Abordar el problema multidimensional de la pobreza con un enfoque filosófico de la estadística puede contribuir a que las investigaciones sobre pobreza dejen de estancarse. La filosofía y la estadística tienen en común el problema de la inducción que consiste en sustentar las inferencias que se derivan de los datos como la extrapolación de predicciones y clasificaciones, dado que la estadística como ciencia fáctica es la herramienta del método científico (Bunge, 2005).

Bajo esa línea la filosofía de la estadística consiste en los fundamentos y la interpretación apropiada de los métodos estadísticos, los datos de entrada que requiere y sus resultados. La estadística utiliza hechos empíricos que son codificados y estructurados en datos que luego son utilizados para probar diferentes hipótesis de la realidad. La estadística es utilizada para comunicar resultados que se encuentran en las

investigaciones que se realizan en las diferentes áreas de la ciencia (Stanford Encyclopedia of Philosophy, 2014). De acuerdo con Bandyopahdyay y Forster (2011) las diferentes vertientes estadísticas como la estadística del error, el criterio de información Akaike o la estadística bayesiana han contribuido a entender mejor el problema de la inducción y los medios por los cuales se puede solucionar, aunque cada uno de estos enfoques tiene características diferentes y aportan diferentes visiones de los temas subyacentes que presentan. En ese contexto, no está demás medir la pobreza multidimensional empleando diferentes enfoques a los tradicionales, en este caso por medio del análisis multivariante, para descubrir nuevas características acerca de las dimensiones e indicadores que juegan un rol clave en la incidencia de la pobreza usando los árboles de decisiones CHAID.

Con este trabajo de tesis se busca contribuir a la generación de nuevo conocimiento o mejorar la existente convirtiéndose en una rama de la epistemología (Shapiro, 1983). Esto es, el modelo del análisis multivariante que se ha obtenido para la medición de la pobreza de los hogares del Perú permitirá difundir las técnicas multivariantes y la aplicación de éstas en una medición más objetiva de la pobreza para obtener una lectura menos monetarista de la medición de la pobreza y mucho más sistémica y objetiva.

2.3. Marco conceptual

2.3.1 Medición de la pobreza

i) Terminología de la pobreza

Antes de comenzar con los métodos de medición de la pobreza, es necesario mencionar las principales definiciones sobre la pobreza

que se emplean en el Perú, donde de acuerdo con el INEI (2000, pág. 2) los conceptos más importantes que giran en torno a la pobreza son:

Pobreza Total, cuando el consumo per cápita es inferior al costo de una canasta total de bienes y servicios mínimos esenciales.

Pobreza Extrema, cuando los hogares de las personas tienen ingresos o consumos per cápita inferiores al valor de la canasta mínima de alimentos.

Línea de pobreza total, como el costo de una canasta mínima de bienes, esto es, alimentos y servicios; es decir, se toma en cuenta el gasto total per cápita mensual del valor de la canasta de consumo alimenticio y de servicios.

Línea de pobreza extrema, como el costo de una canasta mínima de alimentos; es decir, el valor mensual del consumo de una canasta de bienes alimenticios.

Brecha de la pobreza, como la diferencia entre los ingresos de los pobres y el valor de las líneas de pobreza, sean estas de la pobreza extrema o de la pobreza total.

Severidad de la pobreza, como un indicador de desigualdad entre los pobres, que se mide como el valor promedio de los cuadrados de las diferencias entre los ingresos de los pobres y la línea de la pobreza, que estaría representando la variabilidad promedio de los ingresos y la línea de pobreza.

Población con necesidades básicas insatisfechas, cuando ésta tiene por lo menos una necesidad básica insatisfecha.

Bajo este concepto, se puede definir como *pobreza absoluta o extrema* cuando se carece de todas las necesidades básicas para sobrevivir o la falta de alimento, vivienda inadecuada, vestimenta o medicina insuficiente, de manera que se encuentran luchando por sobrevivir; y *pobreza relativa* cuando los recursos con los que cuentan son insuficientes para poder alcanzar un estándar aceptable en la sociedad en la que se encuentran (Pérez-Campos y Rodríguez Saldarriaga, 2015; EAPN, 2009; Rowntree, 1901).

Además, Poza Lara (2007) comenta que una persona es pobre de acuerdo a su nivel de vida relacionándola con su entorno y si este se encuentra debajo de un determinado umbral que sea considerado mínimo. De acuerdo con Atkinson (1993;1987) y Martínez Álvarez, (2003) se puede distinguir dos definiciones fundamentales, tales como el concepto de *pobreza objetiva* que es aquella situación medida a través de circunstancias objetivas de los hogares o de las personas, y sus mediciones tienen como base el nivel de vida mínimo. Aquí se consideran el nivel de ingresos, de gastos y el equipamiento que tenga la vivienda. Mientras que la *pobreza subjetiva*, considerada pobreza como insatisfacción, en este caso se define como pobre a quien no se encuentra satisfecho con su situación, ya que se considera excluido de lo que se considera el modo normal de vida, independientemente de sus posibilidades económicas.

Por su parte, el INEI (2016) indica que para medir la incidencia de la pobreza se considera pobre a todas las personas residentes en

viviendas particulares,¹ cuyo gasto per cápita valorizado monetariamente, no supera el umbral de la línea de pobreza. Se incluyen otros modos de adquisición como son: el autosuministro y autoconsumo, el pago en especie, y las donaciones públicas y privadas. Se considera que es *pobreza objetiva* porque la medición no resulta de la percepción subjetiva, sino de un conjunto de procedimientos, que determinan si un individuo se encuentra o no por debajo del umbral que lo clasifica como pobre o no. Se considera que es *pobreza absoluta* porque la medición es respecto a un valor de la línea que no depende de la distribución relativa del bienestar de los hogares. Para ello se usan dos tipos de líneas: *de pobreza extrema* (cuando los hogares tienen ingresos o consumos per cápita inferiores al valor de la canasta mínima de alimentos) y *de pobreza* (cuando el consumo per cápita es inferior al costo de una canasta total de bienes y servicios mínimos esenciales).

ii) Organismos Internacionales que miden la pobreza

Según Jimeno Ocares (2012) los tres organismos internacionales que hacen mediciones de pobreza en el mundo son:

La *Comisión Económica para América Latina y el Caribe (CEPAL)* emplea el enfoque de línea de pobreza absoluta, basándose en los datos de encuestas de hogares. Además, define la canasta básica de alimentos de aproximadamente 2300 calorías diarias y también considera los hábitos de consumo, la disponibilidad de alimentos y las diferencias de precios tanto en las áreas urbanas como en las áreas rurales.

¹ Son aquellas edificaciones destinadas para ser habitada por una o más personas, con vínculo consanguíneo o sin él, que viven bajo normas de convivencia familiar. Véase en: <http://censos.inei.gob.pe/cpv2007/tabulados/Docs/Glosario.pdf>

El *Banco Mundial (BM)*, también considera el enfoque de pobreza absoluta, pero con otro umbral, en lugar de una línea de pobreza, emplea un rango de US\$ 1,25 a US\$ 2 diarios por paridad de poder adquisitivo. De tal manera que con un ingreso hasta US\$ 2 diarios, es pobre; con US\$ 1,25, es indigente.

La *Organización para la Cooperación y el Desarrollo Económico (OCDE)*, emplea el enfoque de pobreza relativa, pero también medidas multidimensionales como pobreza infantil, situación habitacional y calidad de empleo. También realiza contrastes en los ingresos personales con el ingreso promedio del país. Y para definir si un individuo es pobre, mide la pobreza con el 60% de la mediana de ingresos disponibles o el 50% de la media de ingresos.

iii) Enfoques de medición de pobreza

Los enfoques nos ayudan a identificar a la pobreza; y ésta puede ser considerada un fenómeno absoluto o relativo; además se puede determinar a través de la medición del consumo realizado o de la capacidad para realizarlo. El *enfoque relativo* se refiere a la percepción de las personas de su propio bienestar en función del bienestar de los demás. Es decir, que la pobreza de una persona va depender de lo que tenga su grupo social de referencia, y al no tener tanto como él va implicar una condición de privación relativa. Adoptar esta perspectiva implica analizar la pobreza como si fuera un subconjunto de un tema mayor, que es la desigualdad en la distribución de los ingresos. El *enfoque absoluto*, si bien la sociedad determina ciertas necesidades, no se puede negar la existencia de un núcleo irreducible de pobreza absoluta, que es independiente del nivel de ingresos de un grupo referencial. Si nos

referimos a este núcleo, el que está conformado por necesidades cuya satisfacción, es una situación de privación, como la inanición (CIEPYC-UNLP,2010).

Por otro lado, Pérez-Campos y Rodríguez Saldarriaga (2015) proponen que se tiene los siguientes enfoques de medición:

1. El *Enfoque de Pobreza Monetaria*, es el mismo que emplea el INEI y que se detalla en su informe técnico, citado en (INEI, 2017); esto es:

“Se considera como pobres monetarios a las personas que residen en hogares cuyo gasto per cápita es insuficiente para adquirir una canasta básica de alimentos y no alimentos... . Son pobres extremos aquellas personas que integran hogares cuyos gastos per cápita están por debajo del costo de la canasta básica de alimentos.”

2. El *Enfoque de Pobreza Multidimensional*, que desarrolla una metodología para medir la pobreza a través del Índice de Pobreza Multidimensional. Esta definición ha tomado vigencia durante los últimos años y contempla una medición más integral sobre la pobreza a través de la incidencia de esta, estableciendo un hogar con privaciones, al que se mide tres dimensiones: educativa, salud y condiciones de vida. Así se considera a una persona pobre multidimensional si en al menos una de estas dimensiones existe carencia o privación (Alkire y Foster, 2011; Alkire y Santos, 2011).

3. El *Enfoque de Necesidades Básicas Insatisfechas*, que considera a un hogar como pobre (INEI,2016; Koltko-

Rivera,2006), si presenta alguna de las siguientes características:

- Si residen en viviendas con características físicas inadecuadas, es decir con paredes exteriores de estera o quincha, piedra con barro o madera y piso de tierra;
- Si reside en viviendas en hacinamiento, más de 3 a 4 personas por habitación, que no cuentan con baño, cocina, pasadizo y garaje;
- Si reside en viviendas sin ningún tipo de servicio higiénico; si en los hogares tienen niños de 6 a 12 años que no asisten al colegio; si la condición de ocupación del jefe de familia es de desocupado, con nivel de educación de primaria incompleta.

4. El *Enfoque de Pobreza Subjetiva*. se considera a los hogares con una percepción de bienestar precario. (Pérez-Campos y Rodríguez Saldarriaga, 2015, pág. 213)

5. El *Enfoque de Pobreza Relativa*, se establece una valla de pobreza alternativa basándose en algún indicador monetario (ingreso, gasto o consumo) (Pérez-Campos y Rodríguez Saldarriaga, 2015, pág. 213).

6. El *Enfoque de Pobreza Mixta Integrada*, se basa en la intersección de dos características simultáneas. Es decir, considera la intersección de la canasta de alimentos de un hogar, con la condición de vida de dicho hogar (Busso, 2005).

iv) Métodos de medición.

De acuerdo con CIEPYC-UNLP (2010, págs. 31-32) para medir la pobreza se puede utilizar un método directo o uno indirecto. A estos métodos se ha agregado los métodos de la línea de pobreza, el de las necesidades básicas insatisfechas y el integrado, que combina las dos anteriores.

1. El *Método directo*, en este método se considera que una persona es pobre cuando no puede satisfacer una o varias necesidades básicas, como una vivienda decente, acceder a una educación básica, o una alimentación adecuada. Además en este método se relaciona el bienestar con el consumo efectivamente realizado para satisfacer las necesidades básicas.
2. El *Método indirecto*, en este método se clasifica como pobres a todas las personas que no cuenten con los recursos suficientes para satisfacer sus necesidades básicas. Y se evalúa a través de la capacidad para realizarlo. Además en este método una persona que no haya satisfecho varias necesidades básicas podría no ser considerado pobre.
3. El *Método de las necesidades insatisfechas*, en este método se toma en consideración un conjunto de indicadores relacionados con necesidades básicas estructurales, como son: educación, vivienda, salud, infraestructura pública, entre otras; con las que se va evaluar el bienestar individual. Se considera una población pobre, a la que tiene al menos una necesidad básica insatisfecha y como pobres extremos a los que tienen dos o más indicadores insatisfechos.

4. El *Método de medición integrado*, en este método va integrar los métodos de línea de pobreza y necesidades básicas insatisfechas; y las va clasificar en cuatro grupos:
- Pobres crónicos, es el grupo más vulnerables porque tienen al menos una necesidad básica insatisfecha y sus ingresos o gastos están por debajo de la línea de pobreza.
 - Pobres recientes, es el grupo que tienen todas las necesidades básicas satisfechas y sus ingresos o gastos están por debajo de la línea de pobreza.
 - Pobres inerciales, es el grupo que tiene al menos una necesidad básica insatisfecha y sus ingresos o gastos están por encima de la línea de pobreza.
 - Pobre integrados socialmente, es el grupo que no tienen necesidades básicas insatisfechas y sus gastos están por arriba de la línea de pobreza.

v) Medición actual de la pobreza entre países

En cuanto a lo que hacen algunos países, Jimeno Ocares (2012) señala en su investigación que podemos decir que en:

Estados Unidos de América, emplea una línea de pobreza absoluta, según una canasta básica que fue fijada en 1955, pero ajustada de acuerdo a la inflación.

Canadá, no tiene medición oficial, pero la oficina de estadísticas pública tiene tres tipos de cálculos con el porcentaje de población con bajos ingresos: *LICO (Low Income Cutoffs)*, la más antigua línea de pobreza relativa que representa el monto según el cual la mayoría de los *canadienses* gasta en *promedio* un 20% más que el

promedio en alimentación, vivienda y vestuario; *LIM (Low Income Measures)*, línea de pobreza relativa calculada como el 50% del ingreso medio, usa una tabla de equivalencia basada en tamaño y composición del hogar; y *MBM (Market Basket Measure)*, línea de pobreza absoluta, según el costo de una canasta básica (alimentos, vestuario, vivienda, transporte y otros gastos esenciales) (Jimeno Ocares, 2012).

Reino Unido, desde 1997 usa un enfoque multidimensional y tienen más de 50 indicadores para grupos de la población (Jimeno Ocares, 2012).

Irlanda, usa tres enfoques; el de línea de pobreza tienen varias líneas relativas, de acuerdo a los porcentajes del ingreso promedio (40%, 50% y 60%) y de ingreso mediano (50%, 60% y 70%); y el enfoque multidimensional contempla 23 indicadores de carencia agrupados en cuatro categorías (carencia de estilo de vida básico y secundario, de vivienda y otros) que luego suma. El combinado, la unión de los dos anteriores enfoques, usa distintas líneas de pobreza y ocho indicadores de carencia de estilo de vida básico para medir lo que definen como pobreza consistente (Jimeno Ocares, 2012).

Australia, no tiene medición oficial, su principal referente es una línea de pobreza absoluta fijada en los años 70. También publican líneas relativas de pobreza y el gobierno está implementando mediciones de inclusión social (Jimeno Ocares, 2012).

Nueva Zelanda, no tiene medición oficial, el ministerio de desarrollo social publica datos con enfoque de línea de pobreza y de enfoque multidimensional. En el primero, usa líneas relativas calculadas

como 50% y 60% del ingreso mediano. En el segundo, emplea un indicador de mediciones subjetivas que evalúa 14 ítems de tenencia restringida, 7 de participación social, 16 comportamientos económicos y 2 autoevaluaciones (Jimeno Ocares, 2012).

Para el caso de los países latinoamericanos, según Barneche et al. (2010) se tiene que en:

Chile, desde 1987, las mediciones de pobreza son realizadas por el Ministerio de Planificación (MIDEPLAN) en conjunto con la Comisión Económica para América Latina y el Caribe (CEPAL), emplean el método de línea de pobreza, esto es ingreso per cápita del hogar. La información es provista por la Encuesta de Caracterización Socioeconómica (CASEN), se busca evaluar la situación socioeconómica y el impacto de los programas sociales en las condiciones de vida de las personas de bajos ingresos (Barneche et al., 2010).

Argentina, usa el enfoque de línea de pobreza absoluta. Para esto calculan la cantidad de hogares y personas bajo línea de pobreza, según datos de ingresos reportados en la Encuesta Permanente de Hogares (Barneche et al., 2010).

Brasil, el Instituto Brasileño de Geografía y Estadística (IBGE) es el encargado de medir la pobreza sobre la base de dos líneas: de indigencia como ingreso mínimo para cubrir necesidades básicas de alimentación (de 2313 calorías en zonas urbanas y 2400 en rurales), y de pobreza, como ingreso mínimo mensual que satisface una canasta alimentaria y no alimentaria. Tiene periodicidad variable (Barneche et al., 2010).

México, a fines del año 2000 crea un comité técnico para la medición de la pobreza, que definió 3 líneas absolutas. En el año 2004, se creó el Consejo Nacional para la Evaluación de la Política Social (CONEVAL), el cual emplean el enfoque de línea de pobreza multidimensional. En el año 2009, se usó el enfoque combinado como el oficial. Actualmente bajo el enfoque de línea de pobreza utilizan dos líneas absolutas (alimenticia y no alimenticia), en lo multidimensional usan un índice de carencia en seis dimensiones: rezago educativo, servicios de salud, seguridad social, calidad y espacio de vivienda, servicios básicos de vivienda y acceso a la alimentación. En el combinado calculan la intersección entre la población bajo pobreza y con carencia (Consejo Nacional de Evaluación de la Política de Desarrollo Social, 2015; Barneche et al., 2010).

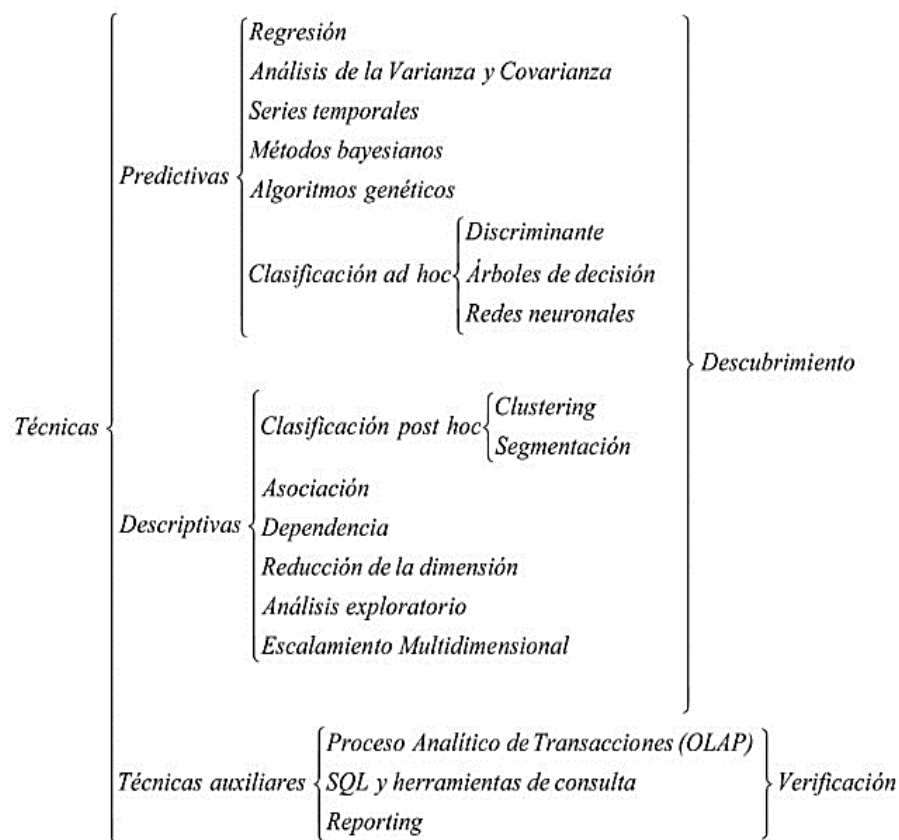
Perú, de acuerdo con (INEI, 2016, pág. 43), se utiliza el enfoque de pobreza monetaria.

2.3.2 Árboles de Decisión

En este trabajo de investigación se aplica el análisis multivariante, y la razón básica, radica en tratar de dar una mejor explicación del fenómeno objeto de estudio, información que los métodos estadísticos univariados y bivariados son incapaces de conseguir. Entre las técnicas del análisis multivariante, se tienen a las técnicas de la minería de datos que provienen de la inteligencia artificial y de la estadística como son los árboles de decisión que son algoritmos que se aplican sobre un conjunto de datos a fin de obtener un modelo (Berlanga et al., 2013; Pérez López y Santin Gónzales, 2007).

Por ello Berlanga et al. (2013) explica que entre las técnicas de la minería se tienen a las redes neuronales, regresión lineal, árboles de decisión, modelos estadísticos, agrupamiento o *clustering* y reglas de asociación. La técnica de los árboles de decisión prepara, sondea y explora los datos para sacar la información oculta en ellos. Se aborda la solución a problemas de predicción, clasificación y segmentación.

Figura 2.4
Clasificación de las técnicas de la minería de datos



Nota. Tomado de Pérez López y Santin Gónzales (2007).

Las técnicas de la minería de datos de acuerdo con Pérez y Santín (2007) se clasifican entre técnicas predictivas, técnicas descriptivas y técnicas auxiliares tal como se muestra en la Figura 2.4:

Otra de las razones para usar los árboles de decisión es porque por su estructura son fáciles de comprender y analizar. Entre dichas facilidades, permite un análisis rápido que cuantifica el costo y las probabilidades de

los diferentes resultados brindando así ayuda en la toma de decisiones (Zuniga y Abgar, 2011).

Se puede indicar que los árboles de decisión, de acuerdo con Hair et al. (2004) es una técnica del análisis multivariante, que están basadas en reglas que consisten en nodos (puntos de decisión) y de ramas (conexiones entre nodos) que van a alcanzar múltiples resultados al estar de paso entre dos o más nodos. Además, tienen un aspecto familiar, por ejemplo, con los dendrogramas del análisis de clúster jerárquico, pero que se construye e interpreta en forma distinta. En esta técnica se hacen particiones secuenciales en el conjunto o matriz de datos con el objeto de maximizar las diferencias en la variable dependiente.

En ese sentido, de acuerdo con Hair et al. (2004) se podría afirmar que esta técnica permite representar en forma gráfica y analítica los eventos que pueden surgir a partir de una decisión asumida en cierto momento; y tomar las decisiones más acertadas, con base a las probabilidades, dentro de un abanico de posibles decisiones.

i) Antecedentes de los árboles de decisión

El algoritmo genérico se denomina "*Greedy algorithm*" que va desde la raíz hacia abajo buscando de manera recursiva los atributos que generan el mejor árbol hasta encontrar el óptimo global con una estructura de árbol lo más simple posible. La diferencia entre estos algoritmos es la posibilidad de obtener valores reales. Estos algoritmos se caracterizan por su capacidad de procesar datos de manera eficiente y clasifican los distintos atributos del árbol y la eficiencia de estos, con el objetivo de obtener

el mejor árbol y que éste sea lo más simple posible (Dupouy Berrios, 2014).

Los árboles de decisión permiten: plantear claramente el problema de tal manera que todas las opciones sean analizadas; desarrollar un análisis rápido de todas las consecuencias de las posibles decisiones; y utilizar un esquema que cuantifica el costo de los resultados y las probabilidades de que los diferentes resultados aparezcan. Hecho que es de mucha ayuda en la toma de decisiones adecuadas (Zuniga y Abgar, 2011).

ii) Definición y conceptos

Los árboles de decisión se basan en algoritmos para clasificar elementos u observaciones, utilizando particiones sucesivas; es una técnica que permite entender el proceso del modelo e interpretar las decisiones tomadas por el modelo, revelando las formas complejas en la estructura de datos. Esta es una técnica explicativa y de descomposición que utiliza un proceso de división secuencial, iterativa y descendente; y considerando una variable dependiente forma grupos homogéneos definidos específicamente mediante combinaciones de variables independientes en las que se incluyen la totalidad de los casos recogidos en la muestra (Parra, 2019).

En esta técnica se hace una división jerárquica y secuencial del problema, y en cada una de estas divisiones o nodos se describen gráficamente las decisiones posibles y los resultados de las distintas combinaciones de decisiones y eventos. A cada evento se le asignan probabilidades y a cada una de las ramas se le determina un resultado. Los árboles se representan en nodos en

los que se pueden entender el conocimiento que genera el árbol (Dupouy Berrios, 2014).

En los diferentes algoritmos de los árboles de clasificación se genera nodos terminales y una escala de probabilidades con los posibles valores para el resultado final y la probabilidad empleada tanto para validar el modelo, como para el cálculo de la pérdida esperada (Cardona H., 2004).

iii) Características de los algoritmos de clasificación

La característica más importante es que se asume que los grupos deban ser disjuntos. Dado que los árboles de decisión conducirán un objeto asignado a una sola hoja; por tanto, un único grupo a un objeto. Para ello es necesario que las particiones existentes deban ser disjuntas.

En Parra (2019) se indica que los elementos y herramientas que determinan la construcción de un árbol son:

- El criterio para determinar la partición de cada nodo;
- La regla que declara un nodo terminal;
- La asignación de una clase a cada nodo terminal, lo que determina la regla de clasificación;
- La fusión entre la relación con la variable dependiente, las categorías de las variables predictoras no significativas se agrupan para formar categorías combinadas que sean significativas;
- La partición, es la selección del punto de división y la variable utilizada para dividir el conjunto de todos los datos que se elige por comparación con todas las demás;

- La poda, es el proceso que permite eliminar las ramas que añaden poco valor de predicción al árbol;
- La evaluación de la bondad del clasificador obtenido; y
- La estimación de la validación del árbol y el cálculo del riesgo.

iv) Tipos de árboles de decisión

Se pueden distinguir dos tipos de árboles:

- De clasificación, CHAID (*Chisquared Automatic Interaction Detection*), en los que se emplea variables categóricas, tanto nominales como ordinales; y
- De regresión, CART (*Classification and Regression Trees*), que se aplica a variables continuas.

v) Los algoritmos más usados

1. Árboles CHAID,

Su antecesor es el algoritmo AID (*Automatic Interaction Detection*), las variables explicativas deben ser de escala nominal u ordinal y la variable a explicar debe de ser de escala razón o de intervalo o de escala dicotómica (0 y 1). El análisis constituye en un análisis de la varianza entre las categorías de la variable independiente, que maximiza la varianza secuencial que se realiza mediante divisiones dicotómicas de la variable dependiente que busca en cada etapa la partición inter-grupos y minimiza la varianza intragrupo. La agrupación de categorías se efectúa probando todas las combinaciones binarias posibles de las variables. El algoritmo presenta dos limitaciones: por una parte, requiere un gran número de elementos muestrales para efectuar el análisis; y por otra, no brinda un modelo explícito que

explique o determine la relación existente entre la variable dependiente y las variables explicativas (Parra, 2019).

Las limitaciones del AID son corregidas por CHAID (*Chisquared Automatic Interaction Detection*), porque en vez de usar el análisis de la varianza, el árbol se desarrolla sobre las tablas de contingencia y el estadístico χ^2 . En este caso, se usa variables categóricas nominales, categóricas ordinales y variables continuas, permitiendo generar tanto árboles de decisión para resolver problemas de clasificación como árboles de regresión. En este algoritmo los nodos se pueden dividir en más de dos ramas; y la construcción del árbol se basa en el cálculo de la significación de un contraste estadístico como criterio para definir la jerarquía de las variables predictoras o de salida, al igual que para establecer las agrupaciones de valores similares respecto a las variables de salida, a la vez que conserva inalterables todos los valores distintos. Todos los valores estadísticamente homogéneos son clasificados en una misma categoría y asignados a una única rama. Como medida estadística, si la prueba es continua, se utiliza la prueba F, mientras que si la variable predicha es categórica se utiliza la prueba Chi-cuadrado (Parra, 2019).

2. Árboles CART

El algoritmo CART (*Classification and Regression Trees*) genera árboles de decisión binarios lo que significa que, cada nodo se divide en exactamente dos ramas. El modelo permite usar variables de entrada y de salida nominales, ordinales y continuas y puede resolver problemas de clasificación como de regresión. Para calcular la medida de impureza, el algoritmo utiliza el índice de Gini, el mismo que toma el valor cero cuando

un grupo es completamente homogéneo y un mayor valor cuando el grupo es heterogéneo.

3. Árboles *QUEST*

El algoritmo *QUEST* (*Quick, Unbiased, Efficient Statistical Tree*) corrige y restringe la búsqueda exhaustiva de las particiones significativas que se generan en los algoritmos AID, CHAID y CART. Para ello, primero elige la mejor variable predictora cuyo objetivo es que el número de categorías que poseen las variables no afecte la elección de la mejor variable: Luego para realizar la mejor segmentación de la variable que se ha seleccionado. Entonces, el modelo sólo puede ser utilizado si la variable de salida es categórica nominal. En el desarrollo se obtiene la mejor división de los valores de la variable elegida. Además, otros de los cambios propuestos en este algoritmo es la eliminación de la poda, la transformación de las variables cualitativas en cuantitativas a través del procedimiento *CRIMCOORD* y un cambio en los valores perdidos de los clasificadores en los distintos nodos (Muñoz Texzocotla, 2011). Cabe mencionar que Loh & Shih (1997) indican que la transformación *CRIMCOORD*, de datos categóricos a numéricos, consta de dos pasos: primero, los elementos del conjunto de datos que toman valores de una variable categórica se mapean a vectores ficticios 0-1; y segundo, los vectores ficticios son proyectados a su coordenada discriminante más grande.

Adicionalmente, el algoritmo contiene la posibilidad de construir particiones no binarias como en el método CHAID y, a semejanza del método CART, el rechazo a la validación cruzada previa propuesta por Breiman et al. (1984). Respecto a estos algoritmos, la diferencia está en la forma de particionar los

nodos. Asimismo, los autores proponen una clasificación basada en el análisis discriminante, a la que llamaron *FACT* (*Fast Algorithm for Classification Trees*). Así, una vez que se ha seleccionado la variable, se procede a ver cuál es la mejor partición binaria del nodo, donde nos podemos encontrar en alguno de los casos siguientes:

Por un lado, para reducir a dos categorías a la variable dependiente se usa el procedimiento de conglomerados k-means, escogiendo las medias muestrales los valores extremos y para cada media adicional se calcula la distancia cuadrática a los centros anteriormente elegidos y se agrupan al más cercano. De ese modo, se vuelven a recalcular los centros y se vuelve a asignar un grupo dependiendo de la proximidad a los nuevos centros recalculados.

Mientras, que si la variable elegida es nominal hay que convertirla en un vector de variables ficticias empleando el análisis discriminante, que convierte cada valor discreto en otro continuo con valores entre -1 y +1.

4. Árboles C5.0

El algoritmo C5 es uno de los algoritmos más utilizados en el ámbito de los árboles de clasificación. La forma de inferir árboles de decisión a través de este algoritmo es el resultado de la evolución del algoritmo C4.5 y su versión ID3 (Quinlan, 1993).

Según la investigación de Parra (2019), el algoritmo genera modelos de árboles de clasificación, usando sólo variables de

salida categórica. En tanto que las variables de entrada pueden ser continua o categórica. El algoritmo básico ID3 construye los árboles de decisión de manera descendente y se empieza indagando qué atributo es el que debería ser colocado en la raíz del árbol. Para resolver esta cuestión cada atributo es evaluado a través de una prueba estadística que determina cómo clasificar. Cuando se selecciona el mejor atributo éste es colocado en la raíz del árbol. Entonces una rama y su nodo se crea para cada valor posible del atributo en cuestión. Los ejemplos de entrenamiento son repartidos en los nodos descendentes de acuerdo con el valor que tengan para el atributo de la raíz. El algoritmo se detiene cuando en el entrenamiento se comparte el mismo valor para el atributo que está siendo probado. Sin embargo, es posible utilizar otros criterios para finalizar la búsqueda. Por un lado, la cobertura mínima de tal forma que el número de ejemplos por cada nodo está por debajo de cierto umbral; y, por otro lado, las pruebas estadísticas para probar si las distribuciones de las clases en los subárboles difieren significativamente.

Por tanto, este algoritmo está basado en el concepto de “ganancia de información” que es una de las maneras de cuantificar la bondad de un atributo que consiste en considerar la cantidad e información que proveerá ese atributo tal y como está definido en la teoría de la información.

III MATERIALES Y MÉTODOS

3.1. Variables e indicadores de la investigación

3.1.1 Variable dependiente

La base de datos Sumaria de la ENAHO 2018 tiene la variable original “**pobreza**” que es resultado del procesamiento INEI-ENAHO 2018 que tiene tres categorías (un hogar es 1 = “pobre extremo”, 2 = “pobre no extremo”, 3 = “no pobre”). Por simplicidad en la aplicación de la técnica de árboles de decisión en el presente trabajo de investigación, la variable dependiente es **pobreza_2**, obtenida de la variable original “**pobreza**” mediante una recodificación a dos grupos; esto es, un hogar es: “1 = *pobre*”, “2 = *no pobre*”.

3.1.2 Variables independientes

Las variables independientes fueron obtenidas de la base de datos Sumaria de la ENAHO 2018, y en total son 21 variables, cualitativas nominales, cuantitativas discretas y cuantitativas continuas, y son las siguientes:

Tabla 3.1
Listado de variables independientes del modelo

Núm.	Denominación	Variables
1	<i>percepho</i>	número total de perceptores de ingresos en el hogar
2	<i>mieperho</i>	número total de miembros en el hogar
3	<i>estrsocial</i>	estrato socioeconómico
4	<i>luz1</i>	en el hogar no se tiene electricidad
5	<i>telef2</i>	en el hogar se tiene teléfono fijo
6	<i>potable2</i>	en el hogar no se tiene acceso al agua potable
7	<i>combusti</i>	en el hogar se cocina con bosta, estiércol, leña o carbón
8	<i>radio</i>	en el hogar se tiene radio
9	<i>tvcolor</i>	en el hogar se tiene TV a color
10	<i>bici</i>	en el hogar se tiene bicicleta
11	<i>refri</i>	en el hogar se tiene refrigeradora/congeladora
12	<i>auto</i>	en el hogar se tiene auto/camioneta
13	<i>moto</i>	en el hogar se tiene moto
14	<i>camion</i>	en el hogar se tiene camión
15	<i>miem_esco</i>	número de miembros del hogar en edad escolar
16	<i>toestud1</i>	número de miembros del hogar que no asisten al colegio
17	<i>anoestu</i>	número total de años de escolaridad de los miembros del hogar
18	<i>pisost</i>	la vivienda del hogar tiene piso de tierra o arena
19	<i>ingmonepercap</i>	ingreso monetario per cápita del hogar
20	<i>gasmonepercap</i>	gasto monetario per cápita del hogar
21	<i>Area</i>	Área Geográfica

A continuación, se presenta en la Tabla 3.2, la matriz operacionalización de variables de este trabajo de tesis:

Tabla 3.2.
Matriz de operacionalización

Variable	Definición conceptual	Definición operacional	Dimensión	Indicador
Variable independiente 1 percepto Perceptores de ingresos en el hogar	Los perceptores son los miembros del hogar que perciben ingresos como remuneración por su trabajo.	Personas en el hogar	Característica de los miembros del hogar	Número total de perceptores
Variable independiente 2 mieperho Miembros del hogar	Los miembros del hogar son los integrantes tales como: padres, hijos, etc..	Personas en el hogar		Número total de miembros en el hogar
Variable independiente 3 mieperho Estrato socioeconómico	De acuerdo con los niveles de ingresos estimados de los hogares se han clasificado en cinco estratos: Alto, Medio Alto, Medio, Medio Bajo y Bajo.	1 = Alto, 2 = Medio Alto, 3 = Medio, 4 = Medio Bajo, 5 = Bajo		Estrato socioeconómico al que pertenece el hogar
Variable independiente 4 luz1 Electricidad en el hogar	Indica si un hogar tiene acceso a la conexión del fluido eléctrico.	1 = No tiene, 0 = Si tiene	Características de la vivienda	En el hogar no se tiene electricidad
Variable independiente 5 telef2 Telefonía fija en el hogar	Indica si un hogar tiene acceso a la conexión de la telefonía fija.	1 = Si tiene, 0 = No tiene		En el hogar se tiene teléfono fijo

.../

/...

Tabla 3.2.
Matriz de operacionalización

Variable	Definición conceptual	Definición operacional	Dimensión	Indicador
Variable independiente 6 potable2 Acceso al agua potable	Indica si un hogar tiene acceso a la conexión a red pública de agua potable.	1 = No tiene, 0 = Si tiene	Características de la vivienda	En el hogar no se tiene acceso al agua potable
Variable independiente 18 pisost Piso de la vivienda del hogar	Determina el tipo de piso que tiene la vivienda de un hogar	1 = Tiene piso de tierra o arena, 0 = Tiene otro tipo de piso		La vivienda del hogar tiene piso de tierra o arena
Variable independiente 21 Area Area geográfica	Indica que área geográfica se ubica la vivienda del hogar	1 = Área urbana 2 = rural		Área geográfica donde se ubica la vivienda del hogar
Variable independiente 7 combusti Combustible para cocinar en el hogar	Indica el tipo de combustible que un hogar usa para cocinar	1 = Usa bosta, estiércol, leña o carbón, 0 = Usa otro tipo de combustible	Combustible para cocinar	En el hogar se cocina con bosta, estiércol, leña o carbón
Variable independiente 8 radio El hogar posee radio	Indica si un hogar posee un aparato de radio AM o FM	1 = Si tiene, 0 = No tiene	Activos del hogar	En el hogar se tiene radio
Variable independiente 9 tvcolor El hogar posee TV a color	Indica si un hogar posee un aparato TV a color	1 = Si tiene, 0 = No tiene		En el hogar se tiene TV a color

...//

//...

Tabla 3.2.
Matriz de operacionalización

Variable	Definición conceptual	Definición operacional	Dimensión	Indicador
Variable independiente 10 bici El hogar posee bicicleta	Indica si un hogar posee bicicleta	1 = Si tiene, 0 = No tiene		En el hogar se tiene bicicleta
Variable independiente 11 refri El hogar posee refrigeradora/congeladora	Indica si un hogar posee un aparato refrigerador / congelador	1 = Si tiene, 0 = No tiene	Activos del hogar	En el hogar se tiene refrigeradora/congeladora
Variable independiente 12 auto El hogar posee auto/camioneta	Indica si un hogar posee un auto/camioneta	1 = Si tiene, 0 = No tiene		En el hogar se tiene auto/camioneta
Variable independiente 13 moto El hogar posee moto	Indica si un hogar posee una moto	1 = Si tiene, 0 = No tiene		En el hogar se tiene moto
Variable independiente 14 camion El hogar posee camión	Indica si un hogar posee un camión	1 = Si tiene, 0 = No tiene		En el hogar se tiene camión
Variable independiente 15 mim_besco Miembros del hogar en edad escolar	Determina cuantos miembros en edad escolar existen en el hogar	Personas en el hogar	Escolaridad	Número de miembros del hogar en edad escolar

...///

///...

Tabla 3.2.
Matriz de operacionalización

Variable	Definición conceptual	Definición operacional	Dimensión	Indicador
Variable independiente 16 foestud1 Asistencia al colegio de los miembros del hogar	Determina cuantos miembros asisten al colegio en el hogar	Personas en el hogar		Número de miembros del hogar que no asisten al colegio
Variable independiente 17 anoestu Escolaridad de los miembros del hogar (en años)	Determina la escolaridad de los miembros hogar	Años de escolaridad	Escolaridad	Número total de años de escolaridad de los miembros del hogar
Variable independiente 19 ingmonepercap Ingreso monetario per cápita del hogar	Determina el monto de ingreso per cápita del hogar	Monto mensual en soles	Ingresos y gastos del hogar	Monto del ingreso monetario per cápita del hogar
Variable independiente 20 gasmonepercap Gasto monetario per cápita del hogar	Determina el monto de gasto per cápita del hogar	Monto mensual en soles		Monto del gasto monetario per cápita del hogar
Variable dependiente pobreza_2 Clasificación del hogar como "pobre" o "no pobre"	Indica si el hogar ha sido considerado como "pobre" o "no pobre"	1 = Pobre 2 = No pobre		Clasificación del hogar: "1 = pobre" o "2 = no pobre"

3.2. Métodos de la investigación

El presente trabajo usa un método cuantitativo deductivo; es cuantitativo porque se obtuvo un modelo de análisis multivariante mediante los datos de la ENAHO 2018; y es deductivo porque se basó en la metodología de medición multidimensional de la pobreza para la construcción del modelo.

El enfoque de pobreza multidimensional, que se usó fue propuesto por Alkire y Foster (2011) y Alkire y Santos (2011) que ha tomado vigencia durante los últimos años y contempla una medición más integral sobre la pobreza a través de la incidencia de esta, estableciendo un hogar con privaciones, al que se mide tres dimensiones: educativa, salud y condiciones de vida. Así se considera a un hogar pobre multidimensional si en al menos una de estas dimensiones existe carencia o privación.

En este caso, no se va a considerar el *Índice Global de Pobreza Multidimensional* planteado en la metodología AF; sin embargo, se va a considerar algunas de las variables que se utilizan en la construcción del índice antes mencionado.

Mediante los árboles de decisión CHAID se obtuvo el modelo multivariante que permite determinar si un hogar es pobre o no pobre en el Perú, técnica que construye un árbol basado en un método estadístico que examina las relaciones entre la variable dependiente y las 21 variables independientes. Esto es, crea un diagrama de árbol que identifica las variables independientes que predicen en forma significativa a la condición de “pobre” y “no pobre” de un hogar.

3.3. Diseño o esquema de la investigación

Esta investigación es descriptiva porque selecciona las características fundamentales del objeto de estudio y su descripción detallada de las partes, categorías o clases de ese objeto (Bernal Torres, 2000). Es una investigación descriptiva porque: proponemos un modelo estadístico y porque hemos hecho la observación y la revisión documental al usar la base de datos Sumaria de ENAHO 2018; y no se limita a la recolección de datos, sino a la predicción e identificación de las relaciones que existen entre dos o más variables a fin de extraer generalizaciones significativas que contribuyan al conocimiento.

Esta investigación es transversal, porque se observa la base de datos Sumaria de ENAHO al final de 2018, es decir no se compara los datos de los hogares en el transcurso del año 2018. Además, porque es el más adecuado para estudiar el desarrollo humano.

Esta investigación es descriptiva correlacional, porque indaga sobre la relación entre variables y cómo interactúan entre ellas. La investigación de correlación es popular y útil en muchos estudios científicos, particularmente en las ciencias sociales. Esto es, se miden dos o más variables que se pretende ver si están o no relacionadas en los mismos sujetos y después se analiza la correlación. La utilidad y el propósito principal de los estudios correlacionales es saber cómo se puede comportar un concepto o variable conociendo el comportamiento de otra u otras variables relacionadas (Sampieri et al., 2006).

3.4. Población y muestra

3.4.1 Población objetivo

La población bajo estudio está constituida por todos los hogares que residen en las viviendas particulares del área urbana y rural correspondientes a todos y cada uno de los departamentos del país y de la provincia constitucional del Callao. La unidad de análisis está constituida por: el hogar y las personas; siendo que la unidad de investigación es el hogar. De acuerdo con los datos levantados por los Censos Nacionales 2017: XII de Población, VII de Vivienda y III de Comunidades Indígenas, el tamaño de la población es de 8 millones 252 mil 284 hogares residentes en viviendas particulares con ocupantes presentes.²

3.4.2 Muestra

Según (INEI, 2016, pág. 147), la muestra es del tipo probabilística, de áreas, estratificada, multietápica e independiente en cada departamento de estudio. La muestra contempla la medición de los cambios en el comportamiento de algunas características de la población para investigar la dinámica de las causas y sus relaciones. El tamaño de muestra de la ENAHO 2018 es de 37 mil 642 hogares en vivienda particulares con ocupantes entrevistados.

² Véase Nota de Prensa N° 155 – 07 Setiembre 2018 del INEI.
https://m.inei.gob.pe/media/MenuRecursivo/noticias/nota-de-prensa-n-155-2018-inei_1.pdf

Tabla 3.3
Hogares entrevistados, por departamento; según área geográfica

		Área Geográfica		
		Área Urbana	Área Rural	Total
Departamento	Amazonas	555	670	1 225
	Áncash	1 133	801	1 934
	Apurímac	383	567	950
	Arequipa	1 807	620	2 427
	Ayacucho	623	569	1 192
	Cajamarca	555	912	1 467
	Callao	1 011	0	1 011
	Cusco	608	685	1 293
	Huancavelica	351	688	1 039
	Huánuco	553	735	1 288
	Ica	1 297	257	1 554
	Junín	899	668	1 567
	La Libertad	1 151	415	1 566
	Lambayeque	1 084	353	1 437
	Lima	4 163	356	4 519
	Loreto	924	564	1 488
	Madre de Dios	674	311	985
	Moquegua	837	160	997
	Pasco	669	610	1 279
	Piura	1 188	473	1 661
	Puno	753	1 085	1 838
	San Martín	869	473	1 342
	Tacna	1 178	193	1 371
	Tumbes	700	164	864
	Ucayali	813	355	1 168
	Total	24 778	12 684	37 462

Nota. Elaborado en base a la información de INEI-ENAH0 2018

El marco muestral para la selección de la muestra de la Encuesta Nacional de Hogares (ENAH0) se basa en la información estadística proveniente de los Censos Nacionales de Población y Vivienda, así como del material cartográfico respectivo.

Tabla 3.4**Hogares entrevistados, por área geográfica; según dominio geográfico**

		Área Geográfica		
		Área Urbana	Área Rural	Total
Dominio Geográfico	Costa Norte	3 867	914	4 781
	Costa Centro	2 841	543	3 384
	Costa Sur	2 115	327	2 442
	Sierra Norte	966	1 420	2 386
	Sierra Centro	3 335	3 324	6 659
	Sierra Sur	3 379	2 817	6 196
	Selva	4 212	3 339	7 551
	Lima Metropolitana	4 063	0	4 063
	Total	24 778	12 684	37 462

Nota. Elaborado en base a la información de INEI-ENAH0 2018

3.5. Actividades del proceso investigativo

En este trabajo de investigación se han desarrollado las siguientes actividades:

- i)** Revisión de literatura actualizada sobre pobreza multidimensional y metodología estadística multivariada para su medición.
- ii)** Formulación del problema de investigación y de la hipótesis
- iii)** Redacción del estado del arte sobre enfoques, organismos que miden y métodos de medición de la pobreza,
- iv)** Redacción de la metodología a seguir para el logro de los objetivos
- v)** Recolección de datos de la base de datos del INEI, vía web <https://proyectos.inei.gob.pe/microdatos/>
- vi)** Procesamiento de los datos, usando el IBM SPSS Statistics, versión 26 y el software estadístico r y rStudio.
- vii)** Análisis e interpretación de los resultados y comparación de estos con los resultados de otros investigadores

3.6. Técnicas e instrumentos de la investigación

En este trabajo se usó la metodología de medición multidimensional de la pobreza; mientras que para la obtención del modelo se usó la técnica de análisis multivariante árboles de decisión y en específico el análisis CHAID, que construye un árbol basado de un método estadístico que examina las relaciones entre la variable dependiente y las 21 variables independientes de la base de datos Sumaria de la ENAHO de 2018 producida por el INEI, disponible en la web de dicha institución.

3.7. Procedimiento para la recolección de datos

En este trabajo la unidad de investigación es el hogar, que de acuerdo con el INEI (2018):

“... es la agrupación primaria que permite atender las necesidades de supervivencia de sus miembros y, a su vez, es el elemento primario de la organización social. En general, los hogares adoptan diferentes tipos de arreglos, que tienen como núcleo fundamental a la pareja con o sin hijos, al que se agregan otras personas con las que pueden o no tener relaciones de parentesco.”

Por lo tanto, para el año 2018, el tamaño de la muestra programada a nivel nacional fue de 37 mil 462 hogares, el 66,14% correspondieron al área urbana (24 mil 778 hogares) y el 33,86% al área rural (12 mil 684) hogares. En tanto que, la mayor parte, el 12,06% corresponden al departamento de Lima (4 mil 519 hogares), la menor parte, el 2,31% corresponde al departamento de Ucayali (864 hogares).³

³ Véase Tabla 3.3 y Tabla 3.4.

3.8. Técnicas de procesamiento y análisis de los datos

3.8.1 Procesamiento de datos

En este trabajo de investigación se hizo uso de los softwares estadísticos siguientes: Software estadístico r⁴ y el IBM® SPSS® Statistics, versión 26. Por un lado, con el Software estadístico r se investigó las diferentes formas de aplicación de la técnica de los árboles de decisión, encontrando una gama enorme de versiones y modos de aplicación, luego de elegido el análisis CHAID y se aplicó el IBM® SPSS® Statistics porque la base de datos Sumaria de la ENAHO-2018 estaba en el formato “SAV” del SPSS y dicho software cuenta con la técnica del análisis CHAID.

3.8.2 Archivo de datos

La matriz de datos se ha obtenido del archivo sumaria de la ENAHO 2018 con algunas modificaciones que nos ha permitido generar un archivo en el software estadístico IBM® SPSS® Statistics, versión 26, que tiene la siguiente estructura:

Tabla 3.5
Variables del archivo de datos, nombres y etiquetas

Variable	Etiqueta
año	año de la encuesta
mes	mes de ejecución de la encuesta
nconglome	número de conglomerado (proveniente del marco)
conglome	número de conglomerado
vivienda	número de selección de vivienda
hogar	número secuencial del hogar
ubigeo	ubicación geográfica

⁴ R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Variable	Etiqueta
dominio	dominio geográfico
estrato	estrato geográfico
percepho	número total de perceptores de ingresos en el hogar
mieperho	número total de miembros de ingresos en el hogar
totmieho	número total de personas en el hogar, al momento de la entrevista
Ld	deflactor espacial
estrsocial	estrato socioeconómico
linpe	línea de pobreza alimentaria
linea	línea de pobreza total
pobreza	código de pobreza INEI
factor07	factor de expansión anual proyecciones cpv-2007
luz1	en el hogar no se tiene electricidad
telef2	en el hogar se tiene teléfono fijo
potable2	en el hogar no se tiene acceso a agua potable
combusti	en el hogar se cocina con bosta, estiércol, leña o carbón
radio	en el hogar se tiene radio
tvcolor	en el hogar se tiene TV a color
bici	en el hogar se tiene bicicleta
refri	en el hogar se tiene refrigeradora/congeladora
auto	en el hogar se tiene auto/camioneta
moto	en el hogar se tiene moto
camion	en el hogar se tiene moto; es una variable cualitativa nomina
miem_esco	número de miembros del hogar en edad escolar
toestud1	número de miembros del hogar que no asisten al colegio
anoestu	número total de años de escolaridad de los miembros del hogar
pisost	la vivienda del hogar tiene piso de tierra o arena
ingmonepercap	ingreso monetario per cápita del hogar; es una variable cuantitativa continua
gasmonepercap	gasto monetario per cápita del hogar
ingmonepercap_ord	ingreso monetario per cápita (Agrupada)
gasmonepercap_ord	gasto monetario per cápita (Agrupada)
pobreza_2	Dos grupos
Depto	Departamento
Area	Área Geográfica

Nota. Elaborado en base a la información de INEI-ENAH0 2018

IV RESULTADOS Y DISCUSIÓN

Mediante los árboles de decisión se ha construido el modelo multivariante que permite determinar si un hogar es pobre o no pobre en el Perú. Es decir, un modelo que clasifica a los hogares como pobres y no pobres; para ello se ha utilizado como variable dependiente (*pobreza_2*) basado en los valores de las 21 variables independientes la base de datos Sumaria de la ENAHO-2018.

4.1. Aplicación del análisis de árboles de decisión

En nuestro caso, la variable dependiente "*pobreza_2*" que es aquella que registra la condición de "*1 = pobre*" y "*2 = no pobre*" que tiene un hogar, frente a las 21 variables independientes que se describen en la Tabla 3.1

Para ello hemos usado el análisis CHAID, que construye un árbol basado en un método estadístico que examina las relaciones entre las 21 variables independientes y la variable dependiente. Este análisis, crea un diagrama de árbol que identifica las variables independientes que predicen en forma significativa a la condición de "pobre" de un hogar. En nuestro caso, contamos con la base de datos "Sumaria" de la ENAHO-2018, que es un resumen a nivel de hogar de esta encuesta anual de hogares que han sido clasificados como "pobres" o "no pobres", la técnica va a generar una serie de tablas de contingencia con las que puede relacionar a la variable dependiente con cada una de las 21 variables independientes.

4.2. Detalles del análisis CHAID

El análisis CHAID es el más usado de los 4 métodos que proporciona el IBM® SPSS® Statistics, versión 26; esto es, CHAID, CHAID Exhaustivo, CRT y QUEST. No es un análisis binario, puede producir más de dos grupos en cualquier nivel de crecimiento del árbol. Por lo cual, crea un árbol más amplio que los métodos crecientes binarios. Se puede sustentar que es aplicable para todo tipo de variable.

Este análisis inicia aplicando la prueba de independencia chi-cuadrado entre cada una de las variables independientes y la variable dependiente. Así selecciona a la variable independiente que tiene una asociación más significativa con la variable dependiente. En el caso, de que una variable independiente tenga más de 2 categorías, las compara y une las categorías que no muestran ninguna diferencia y muestra resultados en forma conjunta. Este proceso de unir categorías finaliza cuando todas las categorías restantes difieren a un nivel de significación específico.

De esta manera, y en forma sucesiva, crea sus grupos y los divide por la variable independiente más relacionada con la variable dependiente; proceso que finaliza cuando no encuentra variables independientes que estén relacionadas con la variable dependiente en los grupos en formados o bien por alguna regla específica de finalización. Luego, denomina nodos terminales a los grupos que quedaron formados.

4.3. Aplicación del análisis CHAID

Al aplicar el análisis CHAID, se ha obtenido la Tabla 4.1, que muestra una síntesis de las opciones utilizadas en la creación del árbol de decisión. Allí se muestra la lista de las variables utilizadas, dependiente e independientes. Es decir, variable dependiente “*pobreza_2*” y las 21 variables independientes. Se muestra también la profundidad alcanzada: los casos mínimos en el nodo parental (padre) y nodos filiales (hijos), la cantidad de nodos hijos o terminales en la construcción del árbol y las variables independientes que se incluyeron en el modelo.

Tabla 4.1
Resumen del modelo estadístico

Especificaciones	Método de crecimiento	CHAID
	Variable dependiente	pobreza_2
	Variables independientes	mieperho, totmieho, luz1, telef2, potable2, combusti, radio, tvcolor, bici, refri, auto, moto, camion, miem_esco, toestud1, anoestu, pisost, ingmonepercap, gasmonepercap, Area
	Validación	Ninguna
	Máxima profundidad del árbol	3
	Casos mínimos en nodo padre	1000
	Casos mínimos en nodo hijo	500
Resultados	Variables independientes incluidas	gasmonepercap, mieperho, pisost, Area, radio
	Número de nodos	31
	Número de nodos terminales	21
	Profundidad	3

Figura 4.2

Árbol de decisión

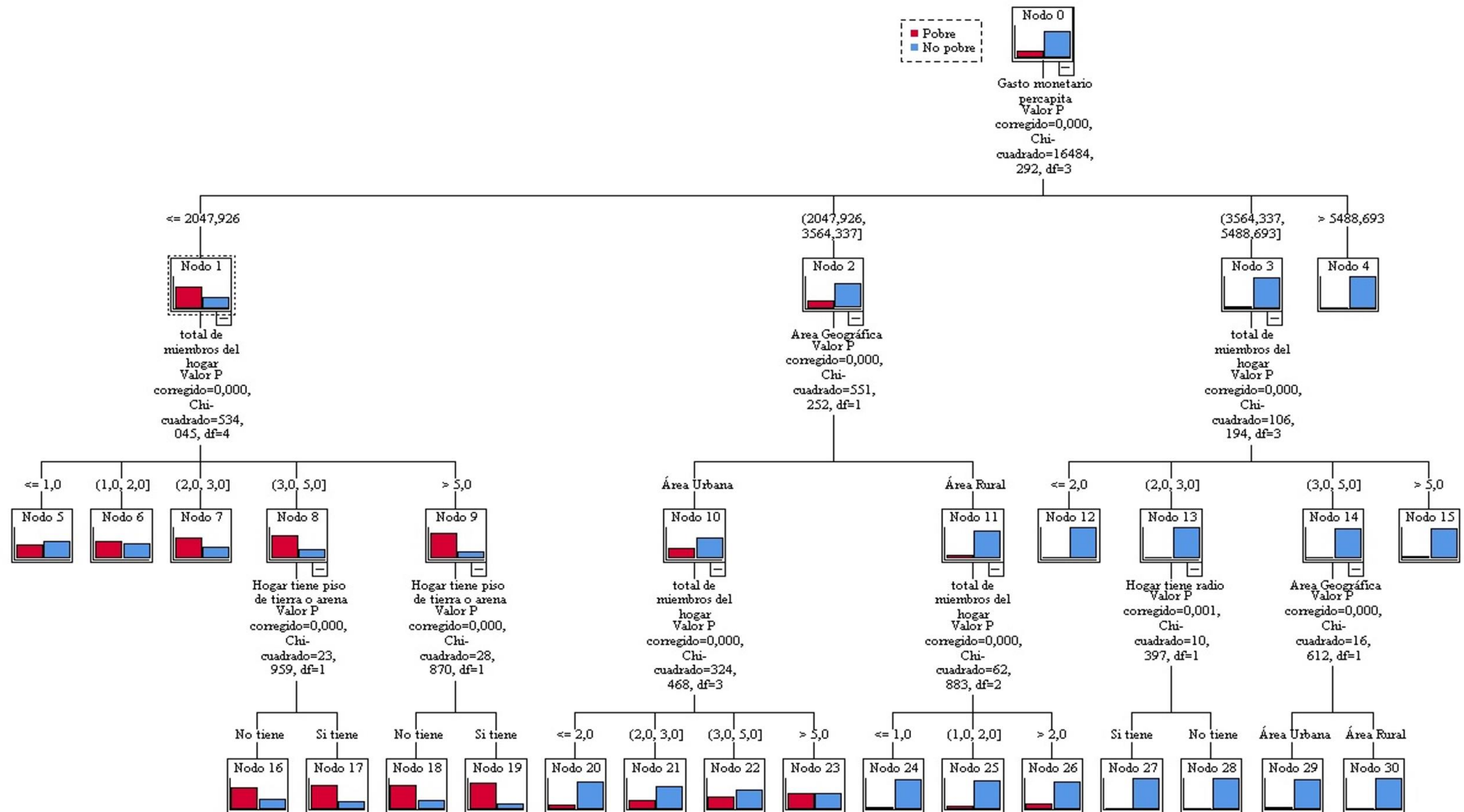


Tabla 4.2

Resumen del árbol de decisión

Nodo	1		2		Total		Categoría pronosticada	Nodo padre	Variable independiente primaria	Sig. ^a	Chi-cuadrado	gl	Valores de división
	N	Porcentaje	N	Porcentaje	N	Porcentaje							
0	6749	18,0%	30713	82,0%	37462	100,0%	2						
1	4962	66,2%	2530	33,8%	7492	20,0%	1	0	gasmonepercap	0,000	16484,292	3	<= 2047,926
2	1646	22,0%	5847	78,0%	7493	20,0%	2	0	gasmonepercap	0,000	16484,292	3	(2047,926, 3564,337]
3	141	1,9%	7351	98,1%	7492	20,0%	2	0	gasmonepercap	0,000	16484,292	3	(3564,337, 5488,693]
4	0	0,0%	14985	100,0%	14985	40,0%	2	0	gasmonepercap	0,000	16484,292	3	> 5488,693
5	461	44,2%	581	55,8%	1042	2,8%	2	1	mieperho	0,000	534,045	4	<= 1,0
6	849	54,7%	703	45,3%	1552	4,1%	1	1	mieperho	0,000	534,045	4	(1,0, 2,0]
7	701	66,0%	361	34,0%	1062	2,8%	1	1	mieperho	0,000	534,045	4	(2,0, 3,0]
8	1711	74,1%	597	25,9%	2308	6,2%	1	1	mieperho	0,000	534,045	4	(3,0, 5,0]
9	1240	81,2%	288	18,8%	1528	4,1%	1	1	mieperho	0,000	534,045	4	> 5,0
10	1273	32,8%	2608	67,2%	3881	10,4%	2	2	Area	0,000	551,252	1	Área Urbana
11	373	10,3%	3239	89,7%	3612	9,6%	2	2	Area	0,000	551,252	1	Área Rural
12	3	0,1%	2041	99,9%	2044	5,5%	2	3	mieperho	0,000	106,194	3	<= 2,0
13	14	1,0%	1455	99,0%	1469	3,9%	2	3	mieperho	0,000	106,194	3	(2,0, 3,0]
14	63	2,3%	2690	97,7%	2753	7,3%	2	3	mieperho	0,000	106,194	3	(3,0, 5,0]
15	61	5,0%	1165	95,0%	1226	3,3%	2	3	mieperho	0,000	106,194	3	> 5,0
16	523	67,8%	248	32,2%	771	2,1%	1	8	pisost	0,000	23,959	1	No tiene
17	1188	77,3%	349	22,7%	1537	4,1%	1	8	pisost	0,000	23,959	1	Si tiene
18	442	74,4%	152	25,6%	594	1,6%	1	9	pisost	0,000	28,870	1	No tiene
19	798	85,4%	136	14,6%	934	2,5%	1	9	pisost	0,000	28,870	1	Si tiene
20	134	13,1%	890	86,9%	1024	2,7%	2	10	mieperho	0,000	324,468	3	<= 2,0
21	179	28,3%	454	71,7%	633	1,7%	2	10	mieperho	0,000	324,468	3	(2,0, 3,0]
22	564	39,2%	876	60,8%	1440	3,8%	2	10	mieperho	0,000	324,468	3	(3,0, 5,0]
23	396	50,5%	388	49,5%	784	2,1%	1	10	mieperho	0,000	324,468	3	> 5,0
24	22	3,7%	578	96,3%	600	1,6%	2	11	mieperho	0,000	62,883	2	<= 1,0
25	53	6,7%	738	93,3%	791	2,1%	2	11	mieperho	0,000	62,883	2	(1,0, 2,0]
26	298	13,4%	1923	86,6%	2221	5,9%	2	11	mieperho	0,000	62,883	2	> 2,0
27	2	0,2%	832	99,8%	834	2,2%	2	13	radio	0,001	10,397	1	Si tiene
28	12	1,9%	623	98,1%	635	1,7%	2	13	radio	0,001	10,397	1	No tiene
29	63	2,9%	2126	97,1%	2189	5,8%	2	14	Area	0,000	16,612	1	Área Urbana
30	0	0,0%	564	100,0%	564	1,5%	2	14	Area	0,000	16,612	1	Área Rural

Método de crecimiento: CHAID
 Variable dependiente: pobreza_2
 a. Bonferroni ajustado

En la Figura 4.1 y en la Tabla 4.2 se muestra tanto el árbol de decisión como la tabla del resumen del árbol de decisión; cada caja del diagrama representa un grupo o subgrupo de casos que se denominan “*nodo*”. El primer nodo, el Nodo 0, representa la muestra entera de los casos; esto es, los 6749 hogares “pobres” (*pobreza_2 = 1*) que representan el 18,0% del total de hogares del archivo de datos; y los 30713 hogares “no pobres” (*pobreza_2 = 2*) que corresponden al 82,0% del total de hogares del archivo de datos; a este nodo raíz se le denomina nodo parental o padre.

De las 21 variables independientes que se utilizaron, la variable “gasto monetario per cápita” (*gasmonepercap*) es la que tiene la más alta asociación con la variable dependiente y se puede considerar como la mejor predictora para considerar a un hogar como “pobre”, y por esta variable la muestra se ha clasificado en cuatro grupos de hogares:

- *gasmonepercap* $\leq 2047,926$, 20% de los hogares (Nodo 1);
- $2047,926 < \textit{gasmonepercap} \leq 3564,337$, 20% de los hogares (Nodo 2);
- $3564,337 < \textit{gasmonepercap} \leq 5488,693$, 20% de los hogares (Nodo 3); y
- *gasmonepercap* $> 5488,693$, 40% de los hogares (Nodo 4).

A estos nodos, Nodo 1, Nodo 2, Nodo 3 y Nodo 4, se les denomina nodos filiales.

Luego se continúa con la división de la muestra en cada uno de estos nodos filiales. Para el caso del Nodo 1, cuando *gasmonepercap* $\leq 2047,926$ se consideró a la variable “total de miembros del hogar” (*mieperho*) por ser la siguiente variable con la más alta asociación con la variable dependiente.

En este caso, a los 7492 hogares (20%) se clasificaron en cinco grupos:

- $mieperho \leq 1,0$, 1042 (2,8%) hogares (Nodo 5);
- $1,0 < mieperho \leq 2,0$, 1552 (4,1%) hogares (Nodo 6);
- $2,0 < mieperho \leq 3,0$, 1062 (2,8%) hogares (Nodo 7);
- $3,0 < mieperho \leq 5,0$, 2308 (6,2%) hogares (Nodo 8); y
- $mieperho > 5,0$, 1528 (4,1%) hogares (Nodo 9).

Si se observa el Nodo 9, se consideró a la variable “hogar tiene piso de tierra o arena” (*pisost*) por ser la siguiente variable con la más alta asociación con la variable dependiente. En este caso, a los 1528 hogares (4,1%) se clasificaron en dos grupos:

- hogar no tiene piso de tierra o arena ($pisost = 2$), 594 (1,6%) hogares (Nodo 18); y
- hogar si tiene piso de tierra o arena ($pisost = 1$), 934 (2,5%) hogares (Nodo 19).

Estos últimos nodos, Nodo 18 y Nodo 19 se les considera como nodos terminales.

Si se analiza con más énfasis la Figura 4.1 y la Tabla 4.2, se puede observar que los subgrupos formados por $gasmonepercap \leq 2047,926$, la proporción de hogares “pobres” es del 66,2%. Dentro de este subgrupo, se puede observar que cuando se pone atención a los subgrupos formados por $mieperho > 5,0$, la proporción de hogares “pobres” es del 81,2%. Si se observa dentro de este penúltimo subgrupo, analizando los subgrupos formados por hogar si tiene piso de tierra o arena ($pisost = 1$), la proporción de hogares “pobres” es del 85,4%. Estas proporciones de hogares pobres son mucho más altas, que el 18% obtenido en la muestra global.

Como se puede apreciar, con este análisis CHAID al desarrollarse las particiones sucesivas de la muestra se puede determinar las variables predictoras más importantes de la variable dependiente. En este caso, el árbol de decisión tiene una profundidad de 3 niveles, porque se han considerado tres particiones por las variables:

- Gasto monetario per cápita: $gasmonepercap \leq 2047,926$,
- Total de miembros del hogar: $mieperho > 5,0$
- Hogar tiene piso de tierra o arena: $pisost = 1$.

En la Tabla 4.3, se muestra las ganancias en cada uno de los 21 nodos terminales que tiene el diagrama de árbol que hemos construido; esta tabla presenta el resumen de cada uno de estos 21 nodos terminales los cuales se pueden identificar con el número del nodo en la primera columna de la izquierda de la tabla. Los nodos terminales están ordenados en forma descendente por la columna respuesta.

Al respecto, vamos a detallar cada una de las columnas de la Tabla 4.3.

- Columna *Nodo*, es el número del nodo y se puede utilizar para ubicar el nodo dentro del diagrama del árbol.
- Columna *Nodo N*, es la cantidad de casos del nodo de cada fila, en particular.
- Columna *Nodo Porcentaje*, es el porcentaje de casos del nodo con respecto a la muestra total.
- Columna *Ganancia N*, indica la cantidad de casos en ese modo que corresponden a la categoría de interés; esto es, la cantidad de hogares “pobres”.
- Columna *Ganancia Porcentaje*, es el porcentaje de casos del nodo que corresponden a la categoría de interés, respecto a la cantidad de casos que corresponden a la categoría de interés en la muestra total.
- Columna *Respuesta*, es el porcentaje de casos de la categoría de interés, respecto a la cantidad de casos en el nodo.

- Columna *Índice*, es la razón del porcentaje de casos para la categoría interés en el nodo entre el porcentaje de casos de la categoría interés en la muestra total.

Tabla 4.3
Ganancias por nodos del modelo

Nodo	Nodo		Ganancia		Respuesta	Índice
	N	Porcentaje	N	Porcentaje		
19	934	2,5%	798	11,8%	85,4%	474,3%
17	1537	4,1%	1188	17,6%	77,3%	429,0%
18	594	1,6%	442	6,5%	74,4%	413,0%
16	771	2,1%	523	7,7%	67,8%	376,5%
7	1062	2,8%	701	10,4%	66,0%	366,4%
6	1552	4,1%	849	12,6%	54,7%	303,6%
23	784	2,1%	396	5,9%	50,5%	280,4%
5	1042	2,8%	461	6,8%	44,2%	245,6%
22	1440	3,8%	564	8,4%	39,2%	217,4%
21	633	1,7%	179	2,7%	28,3%	157,0%
26	2221	5,9%	298	4,4%	13,4%	74,5%
20	1024	2,7%	134	2,0%	13,1%	72,6%
25	791	2,1%	53	0,8%	6,7%	37,2%
15	1226	3,3%	61	0,9%	5,0%	27,6%
24	600	1,6%	22	0,3%	3,7%	20,4%
29	2189	5,8%	63	0,9%	2,9%	16,0%
28	635	1,7%	12	0,2%	1,9%	10,5%
27	834	2,2%	2	0,0%	0,2%	1,3%
12	2044	5,5%	3	0,0%	0,1%	0,8%
4	14985	40,0%	0	0,0%	0,0%	0,0%
30	564	1,5%	0	0,0%	0,0%	0,0%

Método de crecimiento: CHAID
Variable dependiente: pobreza_2

Examinando la Tabla 4.3 de ganancia, se puede ver como el nodo 19 es el nodo que mejor desempeño tiene y representa aproximadamente el 85,4% de los 934 hogares –que representan el 2,5% de los 37462 hogares observados–, hay 798 hogares que son “pobres”, esto es, $(798/934)*100\% = 85,4\%$. Esto significa que hay una mayor proporción de hogares “pobres” que la muestra total; esto es, $(6749/ 37462)*100\%$

= 18.0%. El índice para este nodo indica $(85,4\%/18,0\%) = 474,3\%$ lo cual nos permite entender que esperamos tener casi 5 veces más hogares pobres que si hubiéramos utilizado la selección de casos aleatoria en la muestra completa.

4.4. Resultados del análisis de árboles de decisión

4.4.1 Análisis del riesgo del árbol de decisión

En la Tabla 4.4, se muestra la proporción de error para nuestro modelo. Con esta estimación se puede conocer la exactitud predictiva del árbol. Acompañando a esta medida se muestra el error típico de esta proporción de error.

Tabla 4.4
Riesgo del modelo

Estimación	Desv. Error
0,112	0,002

Método de crecimiento: CHAID
Variable dependiente: pobreza_2

En nuestro caso, se han clasificado incorrectamente a 1852 hogares “pobres” y a 2337 hogares “no pobres”; es decir, a $1852 + 2337 = 4189$ hogares, lo que indica que hay una proporción de error del $(4189/37462) * 100\% = 11,2\%$ de los hogares observados en la muestra. El error típico de esta proporción de error se calcula dividiendo la proporción de error entre el tamaño de la muestra y luego se toma la raíz cuadrada del valor;

$$\text{esto es: } \hat{\sigma}_p = \sqrt{\frac{0,112}{37462}} = 0,002.$$

Con este error típico podríamos estimar la proporción de error “ π ” en toda la población, aplicando una confianza del 95%. Es decir, obtener el intervalo de confianza del 95% ($Z_{0,975} = 1,96$) para la proporción de error en la población mediante la siguiente expresión: $IC(\pi) = (\hat{p} \pm Z_{1-\alpha/2}\hat{\sigma}_{\hat{p}}) = (0,10808; 0,11592)$. Esto significa, que nuestra proporción de error estaría entre 10,8% y 11,6% con un error de estimación del 0,392%.

4.4.2 Análisis de la clasificación de la muestra con el árbol de decisión

En la Tabla 4.5, se muestra una matriz que indica el número de clasificaciones correctas e incorrectas que se ha generado con este modelo. Para cada nodo se registra cuantos hogares “pobres” y hogares “no pobres” fueron correctamente clasificados y cuando no se asignaron correctamente.

Tabla 4.5
Clasificación del modelo

Observado	Pronosticado		
	1	2	Porcentaje correcto
1	4897	1852	72,6%
2	2337	28376	92,4%
Porcentaje global	19,3%	80,7%	88,8%

Método de crecimiento: CHAID
Variable dependiente: pobreza_2

En nuestro caso, se puede observar que de los 6749 hogares “pobres” el modelo clasifico correctamente a 4897; es decir la proporción de acierto fue del $(4897/6749)*100\% = 72,6\%$. En tanto, que de los 30713 hogares “no pobres” el modelo clasifico correctamente a 28376; siendo la proporción de acierto del $(28376/30713)*100\% = 92,4\%$. A nivel global, la proporción de acierto del modelo es del $((4897 + 28376) / (6749$

+ 30713) *100% = 88,8%. Lo que significa que el modelo es muy bueno porque tiene casi 90% de acierto.

V CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

De acuerdo con la Figura N° 4.1, se puede identificar que el nodo 19 como el que tiene mejor desempeño. Esto significa que hay una mayor proporción de hogares “pobres” que en la muestra total; esto es, $(6749/37462)*100\% = 18.0\%$. El índice para este nodo indica $(85,4\%/18,0\%) = 474,3\%$ lo cual nos permite entender que esperamos tener casi 5 veces más hogares pobres que si hubiéramos utilizado la selección de casos aleatoria en la muestra completa.

Por tanto, se considera a los hogares como pobres a aquellos que tienen un gasto per cápita de a lo más 2047,926 soles, que tienen más de 5 miembros del hogar y que la vivienda del hogar tenga un piso de tierra.

En ese sentido, se puede indicar que las variables predictoras más importantes para el modelo; esto es, las variables que permitan clasificar adecuadamente si un hogar es “pobre” o “no pobre” en el Perú son:

- El gasto per cápita mensual, para que un hogar se considere como pobre debe ser de a lo más 2047,926 soles;
- El número de miembros de hogar, para que un hogar se considere como pobre debe ser de más de 5 miembros del hogar; y
- El piso de la vivienda del hogar, para que un hogar se considere como pobre su vivienda debe tener un piso de tierra.

El modelo tiene una proporción de error que estaría entre 10,8% y 11,6% con un error de estimación del 0,392% y un nivel de confianza del 95%.

A nivel global, la proporción de acierto del modelo es del 88,8% = $((4897 + 28376) / (6749 + 30713))$. Lo que significa que el modelo es muy bueno porque tiene casi 90% de acierto.

5.2. Recomendaciones

En la medición de la pobreza se debería tener en cuenta el enfoque multidimensional y debería usarse el análisis multivariante “árboles de decisión”. Por lo que se sugiere al INEI cambio del enfoque de medición de la pobreza monetaria al enfoque de medición de la pobreza multidimensional.

En primera instancia obteniendo el índice de pobreza monetaria, luego hacer uso de las técnicas multivariantes como el análisis de árboles de decisión.

En una segunda instancia se podría usar un sistema geo referencial para lograr la focalización de los hogares pobres y así obtener los mapas de pobreza a nivel de región, provincia, distrito, centros poblados y asentamientos humanos, tanto a nivel urbano como rural.

VI REFERENCIAS BIBLIOGRÁFICAS Y VIRTUALES

- Aguilar Estrada, A., Caamal Cauich, I., & Portillo Vázquez, M. (2018). Intensidades de pobreza multidimensional en México a nivel municipal. *Revista mexicana de ciencias agrícolas*, 9(1), 251-258.
- Aguilar-Estrada, A., Caamal-Cahuich, I., & Ortiz-Rosales, M. (2018). Pobreza multidimensional en Chiapas: generalizada pero heterogénea. *Revista LiminaR. Estudios Sociales y Humanísticos*, 2, 105-117.
- Alkire, S., & Foster, J. (2008). Counting and Multidimensional Poverty Measurement. *Oxford Poverty and Human Development Initiative (OPHI)*, 33 p.
- Alkire, S., & Jahan, S. (2018). *The new global MPI 2018: Aligning with the sustainable development goals*. Oxford Poverty & Human Development Initiative (OPHI).
- Alkire, S., & Santos, M. (2011). Acute Multidimensional Poverty: A New Index for Developing Countries. *Proceedings of the German Development Economics Conference*. Berlin.
- Alkire, S., & Santos, M. (2014). Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index. *World Development*(59), 251–274. Obtenido de <https://doi.org/10.1016/J.WORLDDEV.2014.01.026>
- Álvarez Yágüez, J. (Julio - Diciembre de 2009). Aristóteles: Perí demokratías. La cuestión de la democracia. *ISEGORÍA, Revista de Filosofía Moral y Política*, 49, 60-101. Madrid, España.
- Atkinson, A. (1983). *The Economics of Inequality* (Second Edition ed.). Oxford: Clarendon Press.
- Atkinson, A. (1987). On the Measurement of Poverty. *Econometrica*(Nº 4), 749-764.

- Banco Mundial. (2005).
Generación de Ingreso y Protección Social para los Pobres.
 Washington, D.C.: Banco Internacional de Reconstrucción y Fomento - BM.
- Bank, T. W. (2005). *The World Bank Annual Report 2005.* Washington DC .
- Barneche, P., Bugallo, A., Ferrea, H., Ilarregui, M., Monterde, C., Pérez, M., & Angeletti, K. (2010). Métodos de Medición de la Pobreza. Conceptos y aplicaciones en América Latina. *Entrelíneas de la Política Económica*, 26(4), 31-41.
- Berlanga, V., Rubio Hurtado, M., & Vilà Baños, R. (2013). Cómo aplicar árboles de decisión en SPSS. *REIRE. Revista d'Innovació i Recerca en Educació*, 2013, 6(1), 65-79.
- Bernal Torres, C. (2000). *Metodología de la investigación para administración y economía.* Pearson.
- Boltvinik, J., & Hernández Laos, E. (2000). *Pobreza y Distribución del Ingreso en México* (2da ed.). México D.F.: Siglo XXI Editores.
- Borga, L., & D'ambrosio, C. (2021). Social protection and multidimensional poverty: Lessons from Ethiopia, India and Peru. . *World Development*(147), 105634.
- Bourguignon, F., & Chakravarty, S. (2003). The Measurement of Multidimensional Poverty. *The Journal of Economic Inequality*, 1, 25–49.
 Obtenido de <https://doi.org/10.1023/A:1023913831342>
- Brady, D. (2019). Theories of the Causes of Poverty. *Annual Review of Sociology*(45), 155–175. Obtenido de <https://doi.org/10.1146/ANNUREV-SOC-073018-022550>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification And Regression Trees*,. Belmont, CA.: Wadsworth.
- Bunge, M. (2005). *La ciencia. Su método y su filosofía.* Debolsillo.
- Busso, G. (2005). Pobreza, exclusión y vulnerabilidad social. Usos, limitaciones y potencialidades para el diseño de políticas de desarrollo y de población. *VII Jornadas Argentinas de Estudios de Población.*
- Cardona H., P. (Diciembre de 2004). Aplicación de árboles de decisión en modelos de riesgo crediticio. *Revista Colombiana de Estadística*, 27(2), 139-151.

- Cepal, N. U. (2019). *Panorama Social de América Latina 2018*. Obtenido de <https://www.cepal.org/es/publicaciones/37626-panorama-social-america-latina-2014>
- CIEPYC-UNLP. (Agosto de 2010). Métodos de Medición de la Pobreza. Conceptos y Aplicaciones en América Latina. *Entrelíneas de la Política Económica, Año 4*. (CIEPYC-UNLP, Ed.) La Plata, Buenos Aires, Argentina. Obtenido de www.ciepyc.unlp.edu.ar
- Clausen Lizarrága, J. (2019). *Posibilidades y Desafíos de la Medición de la Pobreza Multidimensional en el Perú: Elementos para potenciar el debate*. Instituto Nacional de Estadísticas e Informática (INEI). Obtenido de https://www.inei.gob.pe/media/MenuRecursivo/investigaciones/pobreza_multidimensional.pdf
- Clausen, J., & Barrantes, N. (2020). Implementing a Group-Specific Multidimensional Poverty Measure: The Case of Persons with Disabilities in Peru. . *Journal of Human Development and Capabilities, 21(4), 8., 21(4), 355–38*. Obtenido de <https://doi.org/10.1080/19452829.2020.182831>
- Colafranceschi, M., Peyrou, M., & Sanguinetti, M. (2011). Pobreza multidimensional en Uruguay: una aplicación de técnicas multivariantes. *Quantum: Revista de Administración, Contabilidad y Economía, 6(1), 28-55*.
- Consejo Nacional de Evaluación de la Política de Desarrollo Social, C. (2015). *Medición de la Pobreza en México y en las Entidades Federativas 2014*. INEGI.
- Contreras, R. (2013). *The stickup kids: Race, drugs, violence, and the American dream*. Univ of California Press.
- Decancq, K., & Lugo, M. (2013). Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews, 32(1), 7-34*. Obtenido de <https://doi.org/10.1080/07474938.2012.690641>
- Dubois, A. (2005). *Diccionario de Acción Humanitaria y Cooperación al Desarrollo*. Recuperado el 12 de julio de 2017, de <http://www.dicc.hegoa.ehu.es/listar/mostrar/172>

- Dupouy Berrios, C. (Julio de 2014). *Aplicación de árboles de decisión para la estimación del escenario económico y la estimación de movimiento la tasa de interés en Chile*. Santiago: Postgrado Economía y Negocios, Universidad de Chile. Recuperado el 21 de Diciembre de 2019
- Durlauf, S. N. (2006). Groups, social influences, and inequality. *Poverty traps*, 141-175. Obtenido de <https://doi.org/10.1515/9781400841295.141>
- EAPN. (2009). *Poverty and Inequality in the UE*. European Anti-Poverty Network.
- Hair, J. F., Anderson, R., Tatham, R., & Black, W. (2004). *Análisis Multivariante* (5ta ed.). Madrid: Pearson Prentice Hall Editores.
- Huaranca, M., & Castillo, L. (2021). *Índice de Pobreza Multidimensional*. BCRP. Banco Central de Reserva del Perú -BCRP. Obtenido de <https://www.bcrp.gob.pe/docs/Proyeccion-Institucional/Encuentro-de-Economistas/2021/ee-2021-s6-castillo-huaranca.pdf>
- Huber, E., & Stephens, J. (2001). *Development and Crisis of the Welfare State: Parties and policies in global markets*. Chicago: Univ. Chicago Press. Obtenido de <https://press.uchicago.edu/ucp/books/book/chicago/D/bo3633601.html>
- INEI. (Enero de 2000). Metodología para la Medición de la Pobreza en el Perú. *Metodologías Estadísticas*(Año 1), 8.
- INEI. (2016). *Evolución de la Pobreza Monetaria 2009 – 2015*. Lima.
- INEI. (17 de Enero de 2017). Obtenido de Sistema de documentacion virtual de las investigaciones estadísticas: http://webinei.inei.gob.pe/anda_inei/index.php/catalog/195
- INEI. (2018). *Perú: Perfil Sociodemográfico, Informe Nacional, Censos Nacionales 2017: XII de Población, VII de Vivienda y III de Comunidades Indígena*. Lima: INEI. Obtenido de https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1539/
- Jimeno Ocares, P. (03 de Noviembre de 2012). *Las distintas formas de medir la pobreza que se usan en el mundo*. Recuperado el 08 de Julio de 2017, de <http://diario.latercera.com/>: <http://diario.latercera.com/edicionimpresa/las-distintas-formas-de-medir-la-pobreza-que-se-usan-en-el-mundo/>

- Källestål, C., Zelaya, E., Peña, R., Pérez, W., Contreras, M., Persson, L., . . . Selling, K. (2019). Predicting poverty. Data mining approaches to the health and demographic surveillance system in Cuatro Santos, Nicaragua. *International journal for equity in health*, 18(1), 165.
- Koltko-Rivera, M. (2006). Rediscovering the Later Version of Maslow's Hierarchy of Needs: Self-Transcendence and Opportunities for Theory, Research, and Unification. *Review of General Psychology*, 10(4), 302–317.
- Loh, W., & Shih, Y. (1997). Split selection methods for classification trees. *Statistica sinica*, 815-840.
- Martínez Álvarez, J. (2003). *Economía de la pobreza*. Madrid: UNED Ediciones, Colección Variada.
- Méndez Aguirre, V. (2011). El Problema de la Pobreza en la Utopía Aristotélica. *Actas del VI Coloquio Internacional ATΩN Competencia y Cooperación de la antigua Grecia a la actualidad*. México: Centro de Estudios Clásicos Universidad Autónoma de México.
- MIDIS. (30 de agosto de 2022). *Midis realizará estudio sobre pobreza multidimensional en el marco de la Alianza del Pacífico*. Obtenido de <https://www.gob.pe/institucion/midis/noticias/645494-midis-realizara-estudio-sobre-pobreza-multidimensional-en-el-marco-de-la-alianza-del-pacifico>
- Muñoz Texzocotetla, O. (abril de 2011). Aprendizaje Maquinal Multivalores. *Tesis para obtener grado de Maestro en Ciencias*. México D.F., México: Universidad Autónoma Metropolitana.
- Ordóñez, M., Cajilima, J., y Cordero, E. (2018). Determinantes de la pobreza multidimensional en el Ecuador y sus interrelaciones. Una aproximación a través de modelos de ecuaciones estructurales para los años 2009 y 2016. (U. d. Cuenca, Ed.)
- Parra, F. (21 de enero de 2019). *Estadística y Machine Learning con R*. Obtenido de R Pubs Blog: <https://bookdown.org/content/2274/metodos-de-clasificacion.html#arboles-de-clasificacion>
- Pérez López, C., y Santin Gónzales, D. (2007). *Técnicas y herramientas: técnicas y herramientas*. Madrid, España: Editorial Paraninfo.

- Pérez Muñoz, A. (2012). *Un modelo explicativo procesual de la pobreza desde la psicología social comunitaria y el enfoque de los DDHH*. Madrid: Universidad Complutense de Madrid.
- Pérez, C., y Santín, D. (2007). *Minería de datos. Técnicas y herramientas*,. Madrid: Thomson.
- Peréz-Campos, P., y Rodríguez Saldarriaga, A. (2015). El ejercicio de medir la pobreza en el Perú. *Evidencia para una política de inversión en el talento, Documento de discusión(27)*, 201-224.
- Poza Lara, C. (2007). *Pobreza multidimensional: El caso específico español a través del panel de hogares de la unión europea*. Madrid: (Tesis doctotal) Universidad Complutense de Madrid.
- Poza Lara, C., y Fernández Cornejo, J. (2010). Una aproximación a la construcción de un indicador de pobreza multidimensional: ¿Cuáles son los focos de riesgo en España? *Revista de Métodos Cuantitativos para la Economía y la Empresa*, 10, 43-72.
- Quinlan, J. (1993). *C4.5: programs for machine learning*. San Francisco,CA.: Morgan Kaufmann Publishers Inc.
- Rowntree, B. (1901). *Poverty: a study of town life*. London: McMillan.
- Sampieri, R., Collado, C., y Lucio, P. (2006). *Metodología de la investigación*. MacGraw-Hill, Chile Smith, M.
- Sancho Rocher, L. (2011). Riqueza, impiedad y ηψβρις en el Contra Midias de Demóstenes. *EMERITA, Revista de Lingüística y Filología Clásica(LXXIX 1)*, 31-54.
- Sawhill, I. V. (1988). Poverty in the US: Why is it so Persistent? *Journal of Economic Literature*, 26(3), 1073-1119. Obtenido de <https://www.jstor.org/stable/2726525>
- Sen, A. (1992). Sobre conceptos y medidas de pobreza. *Comercio Exterior*, 42(4), 310-324. Obtenido de <http://tinyurl.com/ycrlaoh>
- Sen, A. K. (1982). *Choice, Welfare and Measurement*. Oxford: Basil Blackwell.
- Shapiro, S. (1983). Mathematics and Reality. *Philosophy of Science*, 523-548.
- Small, M., Harding, D., & Lamont, M. (2010). Reconsidering culture and poverty. *The annals of the American academy of political and social science*, 629(1), 6-27. Obtenido de

https://doi.org/10.1177/0002716210362077/SUPPL_FILE/INTERVIEW-WITH-MARIO-SMALL.MP3

- Stanford Encyclopedia of Philosophy. (19 de August de 2014). *Philosophy of Statistics*. Obtenido de <http://plato.stanford.edu/entries/statistics/>
- Streib, J., Verma, S., Welsh, W., & Burton, L. (2016). Life, Death, and Resurrections. *The Culture of Poverty Perspective*(1). Obtenido de <https://doi.org/10.1093/OXFORDHB/9780199914050.013.12>
- UNICEF. (2015). *Multidimensional Overlapping Deprivation Analysis (MODA)*. Obtenido de <https://www.unicef-irc.org/research/multidimensional-overlapping-deprivation-analysis-moda-201-2015/>
- Waisgrais, S., Paz, y Curcio, J. (2017). Medición y análisis multidimensionales del bienestar y la pobreza en niñas, niños y adolescentes en Argentina. (I. d. Económico, Ed.) (17).
- Wilson, F. (1996). *Drawing Together Some Regional Perspectives on Poverty*. Londres: Harvester Wheatsheaf.
- Ximenes, V., Júnior, J., Cruz, J., Da Silva, L., & Sarriera, J. (2016). Pobreza multidimensional y sus aspectos subjetivos en contextos rurales y urbanos en las zonas del noreste. *Estudios de Psicología*, 21(2), 146-156.
- Zavaleta, D., Moreno, C., y Santos, M. (2018). La medición de la pobreza multidimensional en América Latina. En S. Deneulin, J. Clausen, & A. Valencia, *Introducción al enfoque de las capacidades: aportes para el desarrollo humano en América Latina* (págs. 253-274). Buenos Aires: Ediciones Manantial & Pontificia Universidad Católica del Perú.
- Zuniga, C., y Abgar, N. (2011). *Breve aproximación a la técnica de árbol de decisiones*. Obtenido de niefcz.files.wordpress.com: <https://niefcz.files.wordpress.com/2011/07/breve-aproximacion-a-la-tecnica-de-arbol-de-decisiones.pd>.

VII ANEXOS

Anexo 1. Sintaxis para la ejecución de árbol de decisión

```
GET
  FILE='D:\TESIS DOCTORAL\Sumaria 2018 con otras
    variables.sav'.
DATASET NAME ConjuntoDatos1 WINDOW=FRONT.
TREE pobreza_2 [n] BY mieperho [s] totmieho [s] luz1 [n]
  telef2 [n] potable2 [n] combusti [n] radio [n]
  tvcolor [n] bici [n] refri [n] auto [n] moto [n]
  camion [n] miem_esco [s] toestud1 [s] anoestu [s]
  pisost [n] ingmonepercap [s] gasmonepercap [s]
  Area [n]
/TREE DISPLAY=TOPDOWN NODES=CHART BRANCHSTATISTICS=YES
  NODEDEFS=YES SCALE=AUTO
/DEPCATEGORIES USEVALUES=[1.00 2.00] TARGET=[1.00]
/PRINT MODELSUMMARY CLASSIFICATION RISK TREETABLE
/GAIN CATEGORYTABLE=YES TYPE=[NODE] SORT=DESCENDING
  CUMULATIVE=NO
/PLOT GAIN INDEX RESPONSE INCREMENT=10
/RULES NODES=TERMINAL SYNTAX=INTERNAL TYPE=SCORING
/METHOD TYPE=CHAID
```

```
/GROWTHLIMIT MAXDEPTH=AUTO MINPARENTSIZE=1000
      MINCHILDSIZE=500
/VALIDATION TYPE=NONE OUTPUT=BOTHSAMPLES
/CHAID ALPHASPLIT=0.05 ALPHAMERGE=0.05 SPLITMERGED=NO
      CHISQUARE=PEARSON CONVERGE=0.001
      MAXITERATIONS=100 ADJUST=BONFERRONI INTERVALS=5
/COSTS EQUAL
/MISSING NOMINALMISSING=MISSING.
```

Anexo 2. Reglas obtenidas en el proceso de generación del árbol de decisión

```
/* Node 5 */.
```

```
DO IF (VALUE(gasmonepercap) LE 2047.92578125) AND  
(VALUE(mieperho) LE 1).
```

```
COMPUTE nod_001 = 5.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 = 0.557582.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 6 */.
```

```
DO IF (VALUE(gasmonepercap) LE 2047.92578125) AND  
(VALUE(mieperho) GT 1 AND VALUE(mieperho) LE 2).
```

```
COMPUTE nod_001 = 6.
```

```
COMPUTE pre_001 = 1.
```

```
COMPUTE prb_001 = 0.547036.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 7 */.
```

```
DO IF (VALUE(gasmonepercap) LE 2047.92578125) AND  
(VALUE(mieperho) GT 2 AND VALUE(mieperho) LE 3).
```

```
COMPUTE nod_001 = 7.
```

```
COMPUTE pre_001 = 1.
```

```
COMPUTE prb_001 = 0.660075.
```

```
END IF.
```

EXECUTE.

/* Node 16 */.

DO IF (VALUE(gasmonepercap) LE 2047.92578125) AND
(SYSMIS(mieperho) OR (VALUE(mieperho) GT 3 AND
VALUE(mieperho) LE 5)) AND (VALUE(pisost) EQ 0).

COMPUTE nod_001 = 16.

COMPUTE pre_001 = 1.

COMPUTE prb_001 = 0.678340.

END IF.

EXECUTE.

/* Node 17 */.

DO IF (VALUE(gasmonepercap) LE 2047.92578125) AND
(SYSMIS(mieperho) OR (VALUE(mieperho) GT 3 AND
VALUE(mieperho) LE 5)) AND (SYSMIS(pisost) OR
VALUE(pisost) NE 0).

COMPUTE nod_001 = 17.

COMPUTE pre_001 = 1.

COMPUTE prb_001 = 0.772934.

END

IF.

EXECUTE.

/* Node 18 */.

DO IF (VALUE(gasmonepercap) LE 2047.92578125) AND
(VALUE(mieperho) GT 5) AND (VALUE(pisost) EQ 0).

COMPUTE nod_001 = 18.

COMPUTE pre_001 = 1.

COMPUTE prb_001 = 0.744108.

END IF.

EXECUTE.

/* Node 19 */.

DO IF (VALUE(gasmonepercap) LE 2047.92578125) AND
(VALUE(mieperho) GT 5) AND (SYSMIS(pisost) OR
VALUE(pisost) NE 0).

COMPUTE nod_001 = 19.

COMPUTE pre_001 = 1.

COMPUTE prb_001 = 0.854390.

END IF.

EXECUTE.

/* Node 20 */.

DO IF (VALUE(gasmonepercap) GT 2047.92578125 AND
VALUE(gasmonepercap) LE 3564.336669921875) AND
(SYSMIS(Area) OR VALUE(Area) NE 2) AND (VALUE(mieperho)
LE 2).

COMPUTE nod_001 = 20.

COMPUTE pre_001 = 2.

COMPUTE prb_001 = 0.869141.

END IF.

EXECUTE.

/* Node 21 */.

DO IF (VALUE(gasmonepercap) GT 2047.92578125 AND
VALUE(gasmonepercap) LE 3564.336669921875) AND
(SYSMIS(Area) OR VALUE(Area) NE 2) AND (VALUE(mieperho)
GT 2 AND VALUE(mieperho) LE 3).

COMPUTE nod_001 = 21.

COMPUTE pre_001 = 2.

COMPUTE prb_001 =

0.717220.

```
END IF.  
EXECUTE.
```

```
/* Node 22 */.
```

```
DO IF (VALUE(gasmonepercap) GT 2047.92578125 AND  
VALUE(gasmonepercap) LE 3564.336669921875) AND  
(SYSMIS(Area) OR VALUE(Area) NE 2) AND (SYSMIS(mieperho)  
OR (VALUE(mieperho) GT 3 AND VALUE(mieperho) LE 5)).
```

```
COMPUTE nod_001 = 22.
```

```
COMPUTE pre_001 =
```

```
2.
```

```
COMPUTE prb_001 = 0.608333.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 23 */.
```

```
DO IF (VALUE(gasmonepercap) GT 2047.92578125 AND  
VALUE(gasmonepercap) LE 3564.336669921875) AND  
(SYSMIS(Area) OR VALUE(Area) NE 2) AND (VALUE(mieperho)  
GT 5).
```

```
COMPUTE nod_001 = 23.
```

```
COMPUTE pre_001 = 1.
```

```
COMPUTE prb_001 = 0.505102.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 24 */.
```

```
DO IF (VALUE(gasmonepercap) GT 2047.92578125 AND  
VALUE(gasmonepercap) LE 3564.336669921875) AND  
(VALUE(Area) EQ 2) AND (VALUE(mieperho) LE 1).
```

```
COMPUTE nod_001 = 24.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 = 0.963333.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 25 */.
```

```
DO IF (VALUE(gasmonepercap) GT 2047.92578125 AND  
VALUE(gasmonepercap) LE 3564.336669921875) AND  
(VALUE(Area) EQ 2) AND (VALUE(mieperho) GT 1 AND  
VALUE(mieperho) LE 2).
```

```
COMPUTE nod_001 = 25.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 = 0.932996.
```

```
END
```

```
IF.
```

```
EXECUTE.
```

```
/* Node 26 */.
```

```
DO IF (VALUE(gasmonepercap) GT 2047.92578125 AND  
VALUE(gasmonepercap) LE 3564.336669921875) AND  
(VALUE(Area) EQ 2) AND (SYSMIS(mieperho) OR  
(VALUE(mieperho) GT 2)).
```

```
COMPUTE nod_001 = 26.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 = 0.865826.
```

```
END
```

```
IF.
```

```
EXECUTE.
```

```
/* Node 12 */.
```



```
DO IF (VALUE(gasmonepercap) GT 3564.336669921875 AND
VALUE(gasmonepercap) LE 5488.692708333333) AND
(VALUE(mieperho) LE 2).
```

```
COMPUTE nod_001 = 12.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 = 0.998532.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 27 */.
```

```
DO IF (VALUE(gasmonepercap) GT 3564.336669921875 AND
VALUE(gasmonepercap) LE 5488.692708333333) AND
(VALUE(mieperho) GT 2 AND VALUE(mieperho) LE 3) AND
(SYSMIS(radio) OR VALUE(radio) NE 0).
```

```
COMPUTE nod_001 = 27.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE
```

```
    prb_001 = 0.997602.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 28 */.
```

```
DO IF (VALUE(gasmonepercap) GT 3564.336669921875 AND
VALUE(gasmonepercap) LE 5488.692708333333) AND
(VALUE(mieperho) GT 2 AND VALUE(mieperho) LE 3) AND
(VALUE(radio) EQ 0).
```

```
COMPUTE nod_001 = 28.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 =
```

```
    0.981102.
```

```
END IF.
```

EXECUTE.

/* Node 29 */.

DO IF (VALUE(gasmonepercap) GT 3564.336669921875 AND
VALUE(gasmonepercap) LE 5488.692708333333) AND
(SYSMIS(mieperho) OR (VALUE(mieperho) GT 3 AND
VALUE(mieperho) LE 5)) AND (SYSMIS(Area) OR VALUE(Area)
NE 2).

COMPUTE nod_001 = 29.

COMPUTE

pre_001 = 2.

COMPUTE prb_001 = 0.971220.

END IF.

EXECUTE.

/* Node 30 */.

DO IF (VALUE(gasmonepercap) GT 3564.336669921875 AND
VALUE(gasmonepercap) LE 5488.692708333333) AND
(SYSMIS(mieperho) OR (VALUE(mieperho) GT 3 AND
VALUE(mieperho) LE 5)) AND (VALUE(Area) EQ 2).

COMPUTE nod_001 = 30.

COMPUTE pre_001 = 2.

COMPUTE

prb_001 = 1.000000.

END IF.

EXECUTE.

/* Node 15 */.

DO IF (VALUE(gasmonepercap) GT 3564.336669921875 AND
VALUE(gasmonepercap) LE 5488.692708333333) AND
(VALUE(mieperho) GT 5).

COMPUTE nod_001 = 15.

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 = 0.950245.
```

```
END IF.
```

```
EXECUTE.
```

```
/* Node 4 */.
```

```
DO IF (SYSMIS(gasmonepercap) OR (VALUE(gasmonepercap) GT  
5488.692708333333)).
```

```
COMPUTE nod_001 = 4.
```

```
COMPUTE pre_001 = 2.
```

```
COMPUTE prb_001 = 1.000000.
```

```
END IF.
```

```
EXECUTE.
```

Un modelo de análisis multivariante para la medición de la pobreza en el Perú

por Carlos Caycho

Fecha de entrega: 25-ago-2023 11:50p.m. (UTC-0500)

Identificador de la entrega: 2151529857

Nombre del archivo: ccaycho.pdf (2.87M)

Total de palabras: 18800

Total de caracteres: 97946

1 %

10

Submitted to Universiteit Hasselt

Trabajo del estudiante

1 %

11

prezi.com

Fuente de Internet

1 %

12

repositorio.unp.edu.pe

Fuente de Internet

<1 %

13

Submitted to Universidad Internacional de la Rioja

Trabajo del estudiante

<1 %

14

repositorio.ug.edu.ec

Fuente de Internet

<1 %

15

Submitted to Pontificia Universidad Catolica del Peru

Trabajo del estudiante

<1 %

16

Submitted to Universidad Cesar Vallejo

Trabajo del estudiante

<1 %

17

es.scribd.com

Fuente de Internet

<1 %

18

Submitted to CONACYT

Trabajo del estudiante

<1 %

19

www.redalyc.org

Fuente de Internet

<1 %

20	www.scielo.br Fuente de Internet	<1 %
21	www.icane.es Fuente de Internet	<1 %
22	inba.info Fuente de Internet	<1 %
23	dspace.ucuenca.edu.ec Fuente de Internet	<1 %
24	www.condesan.org Fuente de Internet	<1 %
25	www.euskara.euskadi.eus Fuente de Internet	<1 %
26	documentop.com Fuente de Internet	<1 %
27	repositorio.unheval.edu.pe Fuente de Internet	<1 %
28	Submitted to Universidad Católica de Santa María Trabajo del estudiante	<1 %
29	Submitted to Universidad de Lima Trabajo del estudiante	<1 %
30	pirhua.udep.edu.pe Fuente de Internet	<1 %
31	Submitted to Universidad Nacional del Santa	

<1 %

32

www.adra.org.pe

Fuente de Internet

<1 %

33

tesis.ucsm.edu.pe

Fuente de Internet

<1 %

34

www.eumed.net

Fuente de Internet

<1 %

35

cies.org.pe

Fuente de Internet

<1 %

36

www.carm.es

Fuente de Internet

<1 %

37

www.defensoria.gob.pe

Fuente de Internet

<1 %

38

CESEL S A. "MEIA del Proyecto Ampliación de la Central Hidroeléctrica Santa Teresa-IGA0001581", R.D. N° 310-2016-MEM/DGAAE, 2020

Publicación

<1 %

39

www.rlc.fao.org

Fuente de Internet

<1 %

40

eprints.ucm.es

Fuente de Internet

<1 %

41

repositorio.unc.edu.pe

Fuente de Internet

<1 %

42	revistas.pucp.edu.pe Fuente de Internet	<1 %
43	ideas.repec.org Fuente de Internet	<1 %
44	www.coursehero.com Fuente de Internet	<1 %
45	cathi.uacj.mx Fuente de Internet	<1 %
46	repositorio.cepal.org Fuente de Internet	<1 %
47	Alipio Orco Díaz. "Gasto público en inversiones y reducción de la pobreza regional en el Perú, período 2009-2018", Quipukamayoc, 2020 Publicación	<1 %
48	Submitted to Universidad Andina del Cusco Trabajo del estudiante	<1 %
49	repositorio.espe.edu.ec Fuente de Internet	<1 %
50	repositorio.uc.cl Fuente de Internet	<1 %
51	repositorio.ucv.edu.pe Fuente de Internet	<1 %
52	biblioteca.uteg.edu.ec:8080 Fuente de Internet	<1 %

53	issuu.com Fuente de Internet	<1 %
54	repositorio.unfv.edu.pe Fuente de Internet	<1 %
55	www.economicas.unsa.edu.ar Fuente de Internet	<1 %
56	Submitted to Universidad Carlos III de Madrid Trabajo del estudiante	<1 %
57	biblioteca2.ucab.edu.ve Fuente de Internet	<1 %
58	humanright2water.org Fuente de Internet	<1 %
59	repositorio.unac.edu.pe Fuente de Internet	<1 %
60	tesis.pucp.edu.pe Fuente de Internet	<1 %
61	www.ipe.org.pe Fuente de Internet	<1 %
62	www.mef.gob.pe Fuente de Internet	<1 %
63	repositorio.undac.edu.pe Fuente de Internet	<1 %
64	studyres.es Fuente de Internet	<1 %

Excluir citas Activo

Excluir coincidencias < 15 words

Excluir bibliografía Activo