MODELO PARA DETECTAR LA EVASIÓN DEL IMPUESTO A LA RENTA EN EL SECTOR DE LOS PRINCIPALES CONTRIBUYENTES PERUANOS.

by

Frank Jonislla Pillaca

Una disertación presentada en cumplimiento parcial

de los requisitos para el grado

Magister en Ingeniería de Sistemas e Informática

mención Ingeniería de software.

UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Diciembre 2018

|  |
| --- |
| NOTE: the copyright page is optional, but you will need to include it in the ProQuest version when it is submitted.  The copyright is no longer than a one-page document. In the event that the research study uses a published survey, a questionnaire, figures, or charts from published materials, the copyright notices of approval must accompany these materials. |

© 2018 by Frank, fjp., Jonislla  
ALL RIGHTS RESERVED

APPROVAL PAGE

MODELO PARA DETECTAR LA EVASIÓN DEL IMPUESTO A LA RENTA EN EL SECTOR DE LOS PRINCIPALES CONTRIBUYENTES PERUANOS.

by

Frank Jonislla Pillaca

Month2018

Approved:

Full name, Credential, Role

Full name, Credential, Role

Full name, Credential, Role

Accepted and Signed: Month, Day, Year

Typed Full Name of Mentor

Accepted and Signed: Month, Day, Year

Typed Full Name of Committee Member

Accepted and Signed: Month, Day, Year

Typed Full Name of Committee Member

Month, Day, Year

Name of Dissertation Approval Authority

Title of Dissertation Approval Authority  
Universidad Nacional Mayor de San Marcos.

Abstract

|  |
| --- |
| To ensure the ABSTRACT title does NOT appear in the Table of Contents, use the “Normal Indent” formatting (APA Formatting toolbar or heading 5a on the headings toolbar) – note: it is already preformatted.  The abstract is to be completed after chapters 4 and 5 are complete. The abstract provides a clear summary of the paper, indicating both content and tone of the paper. Abstracts include the statement of the problem, the research methods used to analyze the problem, a brief description of the research design, a listing of the key results, a brief description of the significance of the results, selected key conclusions, and selected key recommendations. First-person narrative should not be used in the abstract. If learners want to publish the abstract of their research project to *Dissertation Abstracts International* (DAI), a clearinghouse of abstracts, the abstract should be no longer than 350 words. APA-publishable abstracts, however, must be no longer than 120 words. The abstract paragraph should not be indented. To begin typing the abstract, click once in the gray field. |

DEDICATION

|  |
| --- |
| To ensure the DEDICATION title does NOT appear in the Table of Contents, use the “Normal Indent” formatting (APA Formatting toolbar or heading 5a on the headings toolbar) – note: it is already preformatted.  This **optional** section is no longer than one page, double-spaced, and follows the title page. You may delete the dedication page by highlighting everything on the page and hitting the delete key. Be sure the ACKNOWLEDGEMENTS page moves to the top of the page.  The word *dedication* appears in uppercase letters centered at the top of the page. To begin typing the dedication, click once in the gray field. |

|  |
| --- |
| ! IMPORTANT!  If you DO NOT have a COPYRIGHT PAGE, you must adjust the page number of the DEDICATION page. To adjust the page number:  Click on the Insert menu  Select Page Numbers…  Click on Format…  Change start at “iv” to start at “iii” by clicking the down arrow |

Acknowledgments

|  |
| --- |
| To ensure the ACKNOWLEDGMENTS title does NOT appear in the Table of Contents, use the “Normal Indent” formatting (APA Formatting toolbar or heading 5a on the headings toolbar) – note: it is already preformatted.  This **optional** section is no longer than one page, double-spaced, and follows the title page, or the dedication page if one is included. Many writers choose to recognize the help of friends, colleagues, mentors, assistants, editors, and family members on this page. Research funding, grants, and permission to reprint copyrighted materials also should be acknowledged on this page.  The word *acknowledgments*, appears in uppercase letters centered at the top of the page. Note the word is spelled without the letter *e* after the *g*. The page number of the Acknowledgments page will vary depending upon how many of the optional pages are included before it.  To add a DEDICATION or other optional page, simply copy and paste the title and gray field from this page, place your cursor at the BOTTOM OF THE PAGE BEFORE THE NEW PAGE, hold the control key down and hit the enter key to create a new page, place your cursor at the top of the new page, and paste (control+V).  To begin typing the acknowledgments, click once in the gray field. |

|  |
| --- |
| ! IMPORTANT!  If you DO NOT have a COPYRIGHT PAGE and you DELETE the DEDICATION PAGE, you must adjust the page number of the ACKNOWLEDGMENTS page. To adjust the page number:  Click on the Insert menu  Select Page Numbers…  Click on Format…  Change start at “iv” to start at “iii” by clicking the down arrow |

TABLE OF CONTENTS

|  |
| --- |
| To ensure the TABLE OF CONTENTS title does NOT appear in the Table of Contents, use the “Normal Indent” formatting (APA Formatting toolbar or heading 5a on the headings toolbar) – note: it is already preformatted.  To update the Table of Contents, RIGHT CLICK anywhere in the Table of Contents, select UPDATE FIELD, then select UPDATE ENTIRE TABLE and click OK.  If an additional letter appears after the last appendix, place your cursor AFTER THE PAGE NUMBER OF THE PREVIOUS LINE AND HIT DELETE.  You can use the table of contents to HYPERLINK to any point in the dissertation – simply double click on the heading… |

[LIST OF TABLES viii](#_Toc532899742)

[LIST OF FIGURES ix](#_Toc532899743)

[CAPITULO 1: INTRODUCcIóN 1](#_Toc532899744)

[**Antecedentes** 2](#_Toc532899745)

[**Problema** 4](#_Toc532899746)

[**Importancia del problema** 4](#_Toc532899747)

[**Motivación** 7](#_Toc532899748)

[**Objetivos** 7](#_Toc532899749)

[**Organización** 8](#_Toc532899750)

[CHAPTER 2: LITERATURE REVIEW 9](#_Toc532899751)

[**Metodología De Investigación** 9](#_Toc532899752)

[Planificación de la revisión 9](#_Toc532899753)

[Desarrollo de la revisión. 10](#_Toc532899754)

[Resultado de la revisión. 12](#_Toc532899755)

[Analisis 15](#_Toc532899756)

[¿Qué modelos existen para predecir la evasión de impuestos? 15](#_Toc532899757)

[¿Qué algoritmos de machine learning existen para predecir la evasión de impuesto? 16](#_Toc532899758)

[Second Subtopic - heading level 3 18](#_Toc532899759)

[Summary 19](#_Toc532899760)

[Conclusion 19](#_Toc532899761)

[CHAPTER 3: METHODOLOGY 20](#_Toc532899762)

[Research Design 20](#_Toc532899763)

[Appropriateness of Design 20](#_Toc532899764)

[Research Questions 21](#_Toc532899765)

[Population 21](#_Toc532899766)

[Informed Consent 21](#_Toc532899767)

[Sampling Frame 21](#_Toc532899768)

[Confidentiality 22](#_Toc532899769)

[Geographic Location 22](#_Toc532899770)

[Instrumentation 22](#_Toc532899771)

[Data Collection 23](#_Toc532899772)

[Data Analysis 23](#_Toc532899773)

[Validity and Reliability 23](#_Toc532899774)

[Summary 24](#_Toc532899775)

[CHAPTER 4: PRESENTATION AND ANALYSIS OF DATA 25](#_Toc532899776)

[Results 25](#_Toc532899777)

[Summary 29](#_Toc532899778)

[CHAPTER 5: SUMMARY AND RECOMMENDATIONS 30](#_Toc532899779)

[Conclusion 30](#_Toc532899780)

[Implications 31](#_Toc532899781)

[Recommendations 31](#_Toc532899782)

[References 32](#_Toc532899783)

[Appendixes 33](#_Toc532899784)

1. [Title that identifies Appendix to the Reader 33](#_Toc532899785)
2. [Informed Consent Form 35](#_Toc532899786)
3. [Transcripts of Interviews 37](#_Toc532899787)

##### LIST OF TABLES

[Tabla 1 *Criterios de inclusión de los artículos seleccionados* 10](#_Toc532899788)

[Tabla 2 *Lista de criterios de exclusión de artículos.* 10](#_Toc532899789)

[Tabla 3 *Modelos de evasión de impuestos.* 15](#_Toc532899790)

[Tabla 4 *Algoritmos usados para detectar la evasión de impuestos.* 16](#_Toc532899791)

[Table 1 *Title of table formatting for manuscripts requires the title typed flush left at the top of the table, capitalization of the initial letters of the principal words (see APA section 3.1.3), italicizing the title, and double space if the title takes up more than one line, and beginning subsequent lines under the first line* 26](#_Toc532899792)

[Table 2 *Title* 27](#_Toc532899793)

|  |
| --- |
| To ensure the LIST OF TABLES and LIST OF FIGURES appear in the table of contents, use HEADING 5 formatting (Headings toolbar). (Note: it IS already pre-formatted; no need to change it.)  To update the LIST OF TABLES and LIST OF FIGURES, RIGHT CLICK on either Table 1 or Table 2 (or Figure 1), select UPDATE FIELD, then select UPDATE ENTIRE TABLE and click OK.  You can use the Index of Tables to HYPERLINK to any table or figure in the dissertation – simply double click on the heading…  For information on how to properly format tables and figures, see the instructions accompanying the sample tables. |

##### LIST OF FIGURES

[Figura 1. *Estimación de la evasión tributaria de los clientes informales del sistema financiero Fuente(http://www.bcrp.gob.pe/docs/Publicaciones/Documentos-de-Trabajo/2016/documento-de-trabajo-11-2016.pdf)* 5](#_Toc532899802)

[Figura 2. *Proceso de selección de artículos.* 12](#_Toc532899803)

[Figura 3. *Tendencia de los artículos seleccionados por años.* 13](#_Toc532899804)

[Figura 4. *Tendencia de los artículos seleccionados por años.* 13](#_Toc532899805)

[Figura 5. *Distribución de artículos por países.* 14](#_Toc532899806)

[Figura 6. *Distribución de artículos por área de conocimiento.* 14](#_Toc532899807)

[Figure 1. Use sentence case and double space if the caption takes up more than one line; keep second and subsequent lines flush left (note: use the FIGURE CAPTION button on the APA TABLES toolbar). 29](#_Toc532899808)

##### CAPITULO 1: INTRODUCcIóN

|  |
| --- |
| Use “Heading 5” formatting (APA Headings toolbar) for chapter names and numbers (chapter 1, chapter 2, etc.), so they appear correctly in the Table of Contents. All of the chapter headings and titles are preformatted… there is no need to change them. However, if you do accidentally change or delete them,  Remember that CHAPTER should be in ALL CAPS and *must have Heading 5 in the style window*.  NOTE: The chapter title is a required field. If you delete it, the chapter will not appear in the table of contents. |

Chapter one addresses clearly and succinctly describe the problem.

Independent and dependent variables to be investigated are: list and briefly describe key variables. The research methodology will be      .

|  |
| --- |
| The chapter begins with an introductory section, it is approximately one page, and introduces the reader to the topic by citing recent studies. **APA style says NOT to put the title “Introduction”.** The purpose of this chapter is to frame the entire study and capture the attention of the reader. It is key to put the research study into perspective/context and establish, through a succinct problem statement, the need for the research. It is particularly important because if all of its components fit together, the rest of the dissertation will flow smoothly.  To use the Citations CUSTOM TOOLBARS, simply place your cursor in the text where you want to insert a citation. Select the appropriate style from the toolbar and click – it’s that easy! For more information on the citation style, the toolbars have the APA Manual’s page numbers for easy reference.  To begin typing this section, click once in the gray field. |

|  |
| --- |
| USING THE CUSTOM TOOLBARS IN MASTERPLATE\_FIVE: If you have disabled the macros function, the toolbars will not work.  Enable the macros, then close the file and re-open it.  When you open the template for the first time, you may see a pop up security window: CHECK THE <Always trust macros from this source> box.  IF YOU DID NOT GET THE POP UP WINDOW, Go to the TOOLS menu, click on MACRO, then SECURITY. Set the security settings to HIGH (Word 2000 or lower, LOW Word XP), then close the file and re-open it. |

# **Antecedentes**

**¿Qué son los impuestos?**

“Existen necesidades comunes a mucha gente cuya satisfacción exige esfuerzo colectivo, porque cada persona por separado no tendría capacidad para financiarlas”. Para satisfacer estas necesidades colectivas, existen administraciones públicas, como la SUNAT, que se encargan de planificar los gastos de los servicios públicos que disfrutan los ciudadanos. Pero para poder pagar estos gastos, las administraciones públicas tienen que obtener unos ingresos que proceden de los ciudadanos; estos ingresos son los IMPUESTOS”.

**Características de los impuestos**

Pueden ser: Permanentes o transitorios; Directos o indirectos, proporcionales, progresivos, regresivos. Por ejemplo, el impuesto de renta es considerado un impuesto progresivo en la medida en que los agentes con mayores ingresos deben pagar una tasa más alta. Por el contrario, un impuesto regresivo es aquel en el que se capta un porcentaje menor en la medida en que el ingreso aumenta.

Pueden ser reales o personales; Impuestos reales son aquellos cuya naturaleza se determina con independencia del elemento personal de la relación tributaria, y pueden ser definidos por la norma sin referencia a ningún sujeto determinado y gravan una manifestación de riqueza sin ponerse en relación con una determinada persona.

Si todos como sociedad nos involucramos más activamente en los impuestos, seremos una sociedad más participativa en la generación de riqueza, en la aplicación de los recursos gubernamentales, en la exigencia de la transparencia y en la rendición de cuentas a nuestros servidores públicos.

Ejemplos de impuestos:

• Impuesto General a las Ventas

• Impuesto Extraordinario de Solidaridad

• Impuesto general a las rentas

• Impuesto a la Renta

**Impuesto a la Renta.**

El impuesto a la renta es un tributo que se determina anualmente y su ejercicio inicia el 01 de enero y finaliza el 31 de diciembre.

Este impuesto se aplica a los ingresos que provienen del arrendamiento u otro tipo de cesión de bienes muebles o inmuebles, acciones u otros valores mobiliarios, y/o del trabajo realizado de forma dependiente o independiente.

Para la determinación del Impuesto a la Renta de Personas Naturales que no realizan actividad empresarial, corresponde lo siguiente:

Rentas de Capital - Primera Categoría: generadas por el arrendamiento, subarrendamiento o cualquier tipo de cesión de bienes muebles o inmuebles. El pago debes hacerlo tú mismo y corresponde al 6.25% sobre el monto obtenido por la renta.

Rentas de Capital - Segunda Categoría: corresponde a los intereses por colocación de capitales, regalías, patentes, rentas vitalicias, derechos de llave y otros.

Tercera categoría: las derivadas de actividades comerciales, industriales, servicios o negocios.

Rentas del Trabajo (Cuarta y/o Quinta categoría) y Renta de Fuente Extranjera: Si eres trabajador independiente o dependiente y tu ingreso anual es superior a 7UIT S/. deberás pagar impuesto a la renta.

En el Perú el impuesto a la renta es el que más se evade.

Evasión Fiscal.

Existe evasión fiscal cuando una persona infringiendo la ley, deja de pagar todo o una parte de un impuesto al que está obligada. Al incumplir de manera intencional con el pago de las contribuciones que le corresponden como contribuyente y ciudadano, estará cometiendo un delito. Este incumplimiento causa un deterioro económico en las Finanzas Públicas y el contribuyente obtiene un beneficio que no le correspondería, razón por la cual incurre en el delito de defraudación fiscal; para detener este flagelo es necesario generar herramientas que permitan detectar la evasión tributaria, principalmente la evasión del impuesto a la renta.

# **Problema**

El problema es la detección de empresas que no registran sus ganancias de forma completa, en este estudio nos enfocaremos en la evasión de impuesto a la renta por parte del sector de los principales contribuyentes peruanos.

# **Importancia del problema**

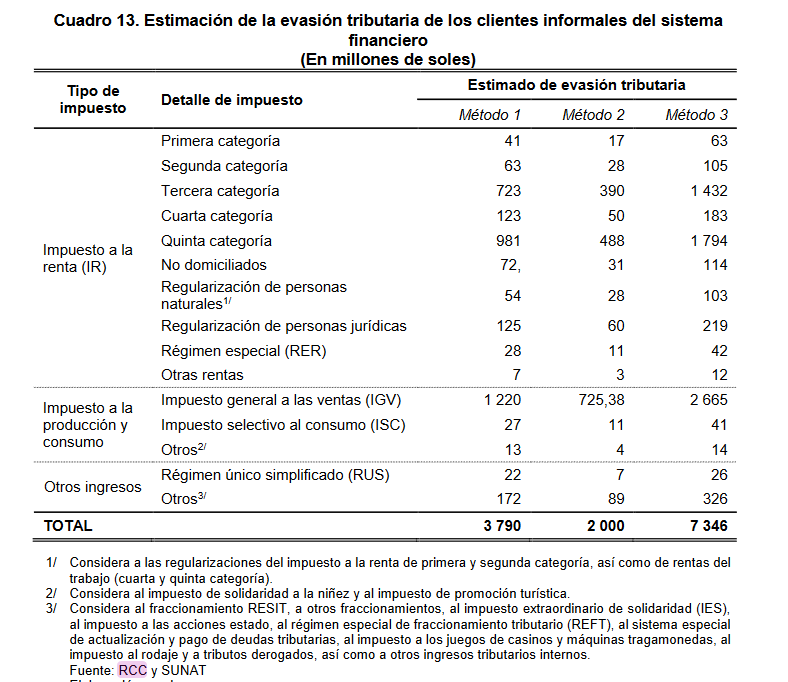
El fraude tributario es un fenómeno mundial que afecta a todas las economías del planeta en mayor o menor medida con pérdidas económicas que van desde el 8 hasta el 63% del PBI dependiendo de cada país analizado (Banco Mundial, 2017). En el Perú contamos con una economía altamente informal donde 8 de cada 10 dólares se mueven fuera del sistema financiero y operan principalmente en mercado con el uso de efectivo que se inserta al sistema y evade impuestos, no es rastreable y no opera de acuerdo con el contexto formal de la economía. Existen cerca de 6.5 millones de empresas y 10 millones de empresarios individuales, pero solo 12 700 empresas representan el 75% del total de la recaudación tributaria nacional.

La evasión tributaria de IGV en el Perú alcanza el 36% del potencial de recaudación y la evasión del impuesto a la renta (IR) es del orden del 57% perdiéndose en general por el total de evasión tributaria cerca de 22 000 millones de dólares anuales que el Estado peruano deja de percibir por este problema (SUNAT, 2018).

El Cuadro 13 muestra los resultados de la estimación de la evasión tributaria de los clientes informales del sistema financiero, a nivel total y por tributo.

Tomando en cuenta la información de los tres métodos empleados, se observa que la evasión total podría ubicarse dentro del rango de S/ 2 000 millones (0,3 por ciento del PBI) y S/ 7 346 millones (1,3 por ciento del PBI), aproximadamente.

Figura 1. *Estimación de la evasión tributaria de los clientes informales del sistema financiero Fuente(*[*http://www.bcrp.gob.pe/docs/Publicaciones/Documentos-de-Trabajo/2016/documento-de-trabajo-11-2016.pdf*](http://www.bcrp.gob.pe/docs/Publicaciones/Documentos-de-Trabajo/2016/documento-de-trabajo-11-2016.pdf)*)*



|  |
| --- |
| Approximately 3/4 of a page.The purpose statement details the reason why the study is being conducted. The purpose statement is distilled to one or two declarative sentences from which the entire study will emanate. Purpose statements can be supplemented with additional information for clarification, but a single, succinct sentence that captures the essence of the study should identify the (a) research method, (b) dependent and (c) independent variables, (d) the audience to which the problem is significant, and (e) the setting.  A sample purpose statement that illustrates the above elements follows: “The purpose of this (a) qualitative, phenomenological research study is to explore the (b) personal value patterns/profiles of (c) first-level supervisors at a (d) manufacturing facility in the Pacific Northwest.”  A sample purpose statement that illustrates the above elements follows: “The purpose of this mixed methodology (quantitative and qualitative) research study is to determine which of three teaching methods are most effective in improving student retention, and to discover through personal interviews why students felt the instruction was (or was not) effective.”  To begin typing this section, click once in the gray field. |

# **Motivación**

Los artículos ofrecen una efectividad medida en los resultados, mas no en el tiempo en que se obtienen estos; por ello se diseñara un modelo que muestre la efectividad medido en resultados correctos y tiempos de ejecución.

|  |
| --- |
| Approximately 1/2 page. This section defines the study’s original contribution to the academic field, to society, to a profession, to the community, and/or to a particular population. The significance section establishes a global reason for doing a worthwhile study. This section explains why the study is unique and who will benefit from its completion. The significance should be appropriate to the field of study. **Discuss two areas of significance: (1) Significance to the field of study (i.e. study contribution) and (2) significance to leadership.**  To begin typing this section, click once in the gray field. |

# **Objetivos**

**Objetivo general**: Diseñar un modelo de detección de evasión de impuestos usando machine learning.

**Objetivo específico**: Diseñar un modelo que sea capaz de seleccionar contribuyentes que no estén cumpliendo sus deberes tributarios

# **Organización**

1. Type research question one here.
2. Type research question two here.

##### CHAPTER 2: LITERATURE REVIEW

# **Metodología De Investigación**

Este estudio a seguido una secuencia de pasos sistemático y simple, es capaz de ser reproducido sin problemas y de manera intuitiva, esta metodología consta de 3 etapas.

Planificación de la revisión: En esta fase, se plantean las preguntas de investigación y se define el protocolo de revisión.

Realización de la revisión: en esta fase, el plan se ejecuta y se seleccionan los artículos primarios, de acuerdo con los criterios de inclusión y exclusión establecidos para el efecto.

Resultado de la revisión: En esta fase, se muestran las estadísticas y el análisis de los documentos encontrados y seleccionados, y que se analizan a continuación. (SYSTEMATIC LITERATURE REVIEW OF CRITICAL SUCCESS FACTORS OF INFORMATION TECHNOLOGY STARTUPS)

### Planificación de la revisión

Para responder a la pregunta de investigación, se hacen las siguientes preguntas sobre las técnicas y herramientas necesarias para detectar la evasión de impuestos.

P1: ¿Qué métodos existen para predecir la evasión de impuestos?

P2: ¿Qué algoritmos de machine learning existen para predecir la evasión de impuesto?

P3: ¿Qué métricas se usan para medir la efectividad de un modelo de evasión de impuestos?

La fuente de búsqueda es proporcionada por el metabuscador: “Web of Science”, el cual es un servicio en línea de información científica. El período de búsqueda comienza en el año 2016, ya que es un periodo razonable para contar con información reciente y actualizada; puesto que la tecnología avanza de forma rápida.

Usamos la siguiente cadena de búsqueda en los títulos, resumen y palabras clave: ts=(((tax AND income) OR (tax AND evasion) OR (tax AND avoidance) OR (tax AND default) OR( tax AND compliance)) AND ((big AND data) OR (machine AND learning) OR (neural AND network) OR (genetic AND algorithm))).

### Desarrollo de la revisión.

Los criterios de inclusión y exclusión se han considerado, como se muestra en las Tablas 1 y 2, respectivamente.

Tabla 1 *Criterios de inclusión de los artículos seleccionados*

|  |  |
| --- | --- |
| **Criterios de inclusión** | |
| **Criterio de inclusión** | **Razón de inclusión** |
| Objetivo de la investigación | Estudios que permiten predecir la evasión de impuestos. |
| Estudios empíricos cuantitativos. | Estos artículos se incluyen porque proporcionan evidencia empírica existente, lo que representa el interés principal de esta revisión. |
| Lenguaje ingles | Sólo artículos en ingles fueron considerados. |

Tabla 2 *Lista de criterios de exclusión de artículos.*

|  |  |
| --- | --- |
| **Criterios de exclusión** | |
| **Criterio de exclusión** | **Razón de exclusión** |
| Tipo de publicación | Se excluyeron libros, capítulos de libros y disertaciones. |
| Unidad de Análisis | Se excluyeron estudios que no consideren soluciones basadas en machine learning. |
| Enfoque de la investigación | Se excluyeron estudios que no muestran metodología de investigación, pruebas numéricas (estadística descriptiva) y análisis o discusión. |

Como primer paso, las palabras clave y sus respectivas descripciones se utilizaron para buscar los artículos principales. La revisión se limitó a artículos en revistas revisadas por pares, omitiendo libros, capítulos de libros y actas de congresos, ya que los artículos de revistas se consideran un conocimiento válido y representan declaraciones autorizadas sobre el tema (Ardito et al., 2015). Específicamente, la búsqueda se limitó a revistas de alta calidad en el campo de "Evasión de impuestos" y "Machine learning". El procedimiento de búsqueda consideró las ediciones disponibles de revistas del período 2005. Se identificaron un total de 59 estudios potenciales, estos fueron sometidos a un proceso de selección de acuerdo con los criterios de inclusión y exclusión establecidos. Así, primero, dividimos los artículos en "artículos excluidos" (19 artículos) y "artículos para el chequeo de texto completo" (40 artículos), fue necesario realizar una revisión previa de los títulos y resúmenes. Sin embargo, muchos resúmenes no proporcionan una comprensión clara del propósito de los artículos, y se seleccionaron 20 referencias. A continuación, un análisis más detallado de los artículos de "verificación de texto completo", se procedió a leer la introducción y conclusiones, obteniendo 16 referencias. Finalmente, procedimos a leer el contenido completo del artículo para determinar su relevancia para el presente estudio y, principalmente, para determinar si estos estudios identifican los factores críticos de éxito, al notar que la cantidad de artículos no era suficiente se tuvo que recurrir a artículos presentados en conferencias que nos permitan tener una cantidad aceptable de documentos a tomar en cuenta para el estudio. Así, la muestra final consta de 16 estudios primarios y 4 de conferencias. Nos encontramos con un gran inconveniente al buscar artículos relacionados a la evasión de impuestos y aprendizaje automático.

Los procesos aplicados, así como los resultados obtenidos en cada paso del proceso, se representan en la Figura 2.

Construir un gráfico de procesos

Figura 2. *Proceso de selección de artículos.*

### Resultado de la revisión.

El resultado del proceso de selección de los artículos nos dio 59, de los cuales 20 fueron seleccionados. Los cuales fueron analizados para responder las preguntas de investigación.

De acuerdo a las tendencias de citación de los artículos nos podemos dar cuenta que el área de evasión de impuestos y machine learning, no ha sido muy estudiado; la tendencia indica que a partir del 2017 los artículos referentes al estudio, vienen siendo referenciados.

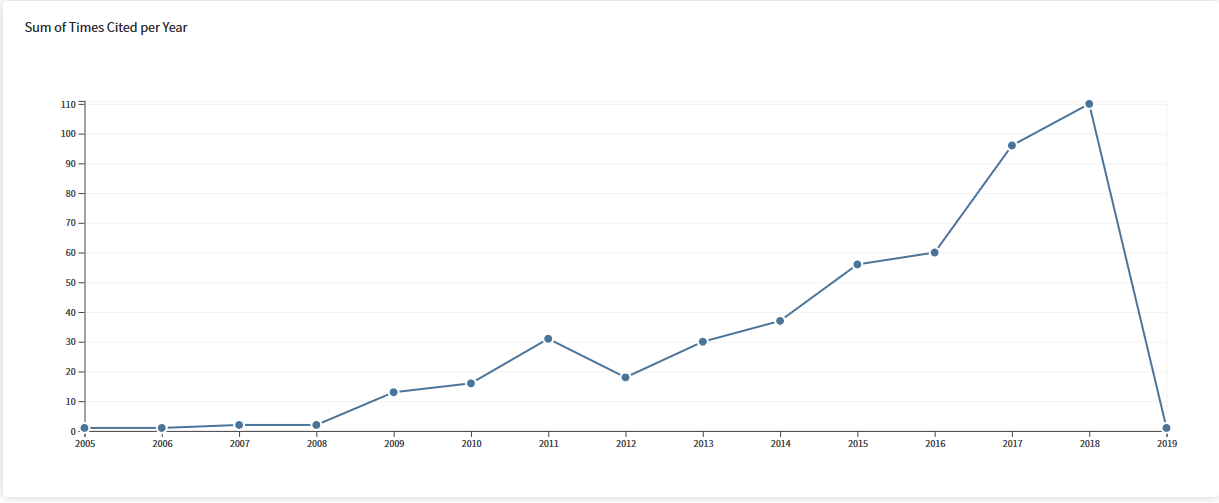


Figura 3. *Tendencia de los artículos seleccionados por años.*

La mayor cantidad de artículos, fueron encontrados en el año 2017, tal como en la figura 4, lo cual demuestra que el área de evasión tributaria y machine learning no muestra mucho interés en la academia.

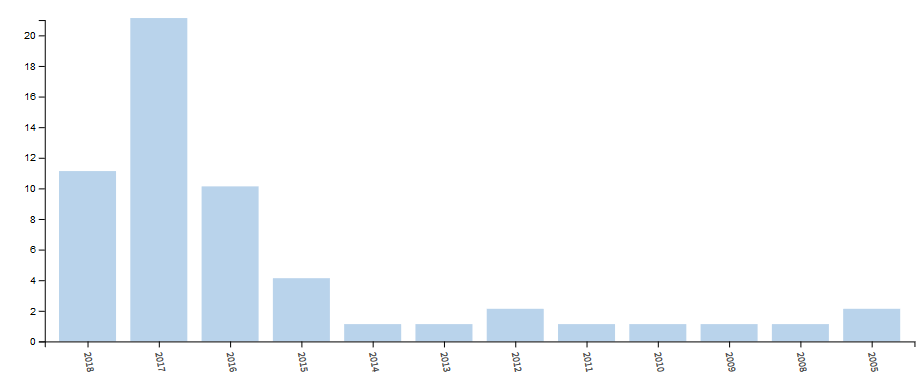


Figura 4. *Tendencia de los artículos seleccionados por años.*

Los países con mayor cantidad de artículos referentes al estudio son estados unidos e Inglaterra, puesto que los países con más ingresos son los que más perdidas tienen debido a la evasión de impuestos.



Figura 5. *Distribución de artículos por países.*

Como era de espera la mayor cantidad de artículos se encontraron en las áreas de economía, ingeniería y ciencias de la computación.

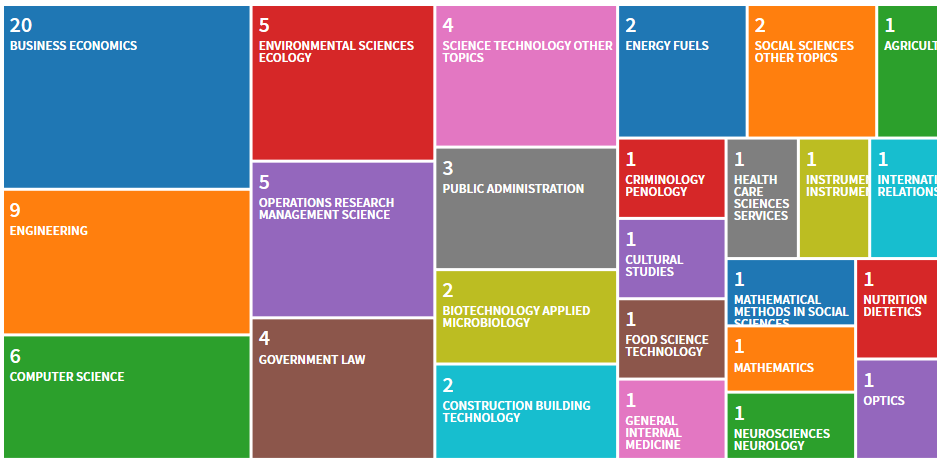


Figura 6. *Distribución de artículos por área de conocimiento.*

# Análisis

Esta sección responde las preguntas de investigación lanzadas en la etapa de planeamiento.

### ¿Qué modelos existen para predecir la evasión de impuestos?

En trabajos relacionados al modelo de detección de evasión de impuestos se pueden agrupar en dos categorías principales: i) analítico (macroeconómico y basado en el principio de agentes), y ii) computacional (basado en agentes, basado en la simulación) (Goumagias, Hristu-Varsakelis, and Assael 2018) en esta investigación nos centraremos en los modelos basados en la simulación.

Tabla 3 *Modelos de evasión de impuestos.*

|  |  |
| --- | --- |
| **Modelo** | **Referencia** |
| modelo de co-evolución de los esquemas de Evasión tributarias con las políticas de auditoría. | (Hemberg et al. 2016) |
| Modelo de evaluación de informes financieros anulas, aplicando la ley de Benford. | (Ausloos, Cerqueti, and Mir 2017) |
| Dinámica de la evasión fiscal a través de un modelo de tipo epidémico; estudia la probabilidad de que un contribuyente cambie de comportamiento de acuerdo al comportamiento de sus relaciones. | (Brum and Crokidakis 2017) |
| Modelo para el estudio de la aversión al riesgo, para comprender la Evasión fiscal de las empresas. | (Goumagias, Hristu-Varsakelis, and Assael 2018) |
| Modelo para determinar un subconjunto óptimo o casi óptimo de variables para un modelo de análisis discriminante lineal. | (Höglund 2017) |
| Modelo para detectar Evasión tributaria usando un sistema inteligente hibrido. | (Rahimikia et al. 2017; Jupri and Sarno 2018; Mabe-Madisa 2018) |
| Modelo para predecir la Evasión de impuestos usando el proceso de regresión Gaussiana. | (Babu and Vasavi 2017) |
| Modelo de red de color para representar el comportamiento económico. | (Tian et al. 2016) |
| Modelo de análisis visual que da soporte a la detección de evasión de impuestos | (Didimo et al. 2018) |
| Modelo de predicción de evasión de impuestos mediante el análisis de redes sociales. | (Lismont et al. 2018) |
| Modelo de detección de fraude tributario usando técnicas de machine learning no supervisado. | (de Roux et al. 2018) |
| Modelo de detección de fraude inter-regiones usando transfer-learning. | (Zhu et al. 2018) |
| Modelo de detección de evasión de impuestos usando patrones de contraste. | (Zheng et al. 2016) |

### ¿Qué algoritmos de machine learning existen para predecir la evasión de impuesto?

El campo del aprendizaje automático se divide en tres subdominios: aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo (Qiu et al. 2016).

El aprendizaje supervisado requiere capacitación con datos etiquetados que tienen entradas y salidas deseadas.

El aprendizaje no supervisado no requiere datos de capacitación etiquetados y el entorno solo proporciona entradas sin los objetivos deseados.

El aprendizaje por refuerzo permite aprender de los comentarios recibidos a través de interacciones con un entorno externo.

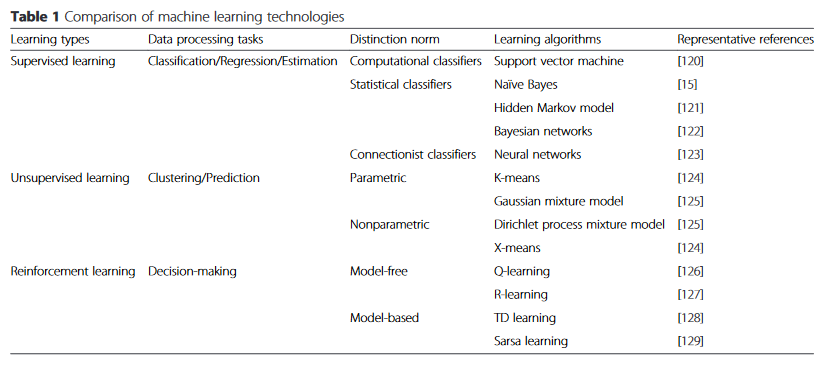


Figura 7. *Clasificación de algoritmos de aprendizaje según el tipo de aprendizaje.*

Tabla 4 *Algoritmos usados para detectar la evasión de impuestos.*

|  |  |
| --- | --- |
| **Algoritmo** | **Referencia** |
| MLP (Perceptrón multicapa) | (Jupri and Sarno 2018; Rahimikia et al. 2017) |
| Naive Bayes | (Jupri and Sarno 2018; Mabe-Madisa 2018) |
| SVM (Máquina de vectores de soporte) | (Jupri and Sarno 2018; Rahimikia et al. 2017) |
| KNN (Vecino más cercano K) | (Jupri and Sarno 2018) |
| C4.5 | (Jupri and Sarno 2018) |
| Transfer Component Analysis (TCA) | (Zhu et al. 2018) |
| LightGBM | (Zhu et al. 2018) |
| Transfer Adaboost | (Zhu et al. 2018) |
| árbol de decisión | (Lismont et al. 2018; Mabe-Madisa 2018) |
| Teoría de grafos | (Tian et al. 2016; Didimo et al. 2018) |
| La ley del primer dígito de Benford | (Ausloos, Cerqueti, and Mir 2017) |
| Redes neuronales profundas | (Goumagias, Hristu-Varsakelis, and Assael 2018) |
| Q-learning | (Goumagias, Hristu-Varsakelis, and Assael 2018) |
| Análisis discriminante lineal | (Höglund 2017) |
| Logistic regression (LR) | (Rahimikia et al. 2017) |
| Harmony search (HS) optimization algorithm | (Rahimikia et al. 2017) |
| Gaussian Regression process with varying hyper parameters | (Babu and Vasavi 2017) |
| Regresión logística | (Lismont et al. 2018) |
| Random forest | (Lismont et al. 2018) |

# Summary

##### References

|  |
| --- |
| Use the menus for basic formats. Consult the pages referenced as needed. Other useful pages in the APA manual are: pp. 217-219, abbreviations and cities that need no state or country; pp. 219-222, order of items in the reference list; pp. 232-236, index to the examples by type of work; pp. 237-239, index to elements in the examples. NOTE: USE HEADING 5 FOR THE REFERENCES HEADER SO IT WILL UPDATE CORRECTLY IN THE TABLE OF CONTENTS.  Remember that REFERENCES must appear in the Table of Contents (the heading is already correctly formatted for the TOC… no need to change it)  NOTE: For online references, substitute 'n.d.' if no posting date is known.  HINTS ON USING THE REFERENCES TOOLBAR:  To insert a reference, click on the reference toolbar and select the appropriate type of reference. Be sure your cursor in placed where you want the reference to appear BEFORE inserting the reference.  To put the references in alphabetical order:  Select the entire reference section by putting your cursor in front of the first letter of the first word, scrolling down to the last letter of the last reference, hold the control key down and hit the letter ‘A’,  On the Table menu, select the <SORT> option.  If the reference *is not inserted with a hanging indent,* simply place your cursor anywhere on the reference and select REFERENCE FORMAT on the APA FORMATTING toolbar. |