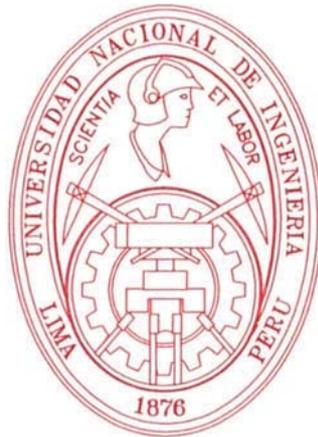


UNIVERSIDAD NACIONAL DE INGENIERÍA
Facultad de Ingeniería Industrial y de Sistemas



**IMPLEMENTACIÓN DEL SISTEMA DE SELECCIÓN DE
CANAL DE CONTROL UTILIZANDO MINERÍA DE
DATOS EN UNA EMPRESA DE ADUANAS**

INFORME DE SUFICIENCIA

PARA OPTAR EL TÍTULO PROFESIONAL DE:

INGENIERO DE SISTEMAS

ARIAS TORRES, JULIO CÉSAR

LIMA – PERU

2012

DEDICATORIA

A mis padres que me dieron la vida, el amor y el firme propósito de alcanzar una carrera digna.

INDICE

DESCRIPTORES TEMÁTICOS	5
LISTADO DE FIGURAS	6
RESUMEN EJECUTIVO	8
INTRODUCCIÓN	10
CAPITULO I	12
ANTECEDENTES	12
1.1. DIAGNÓSTICO FUNCIONAL	12
1.1.1. LA EMPRESA	12
1.1.2. ORGANIGRAMA DE LA EMPRESA.....	13
1.1.3. SERVICIOS.....	15
1.1.4. FUNCIONES	17
1.1.5. CLIENTES.....	21
1.1.6. PROVEEDORES.....	21
1.1.7. PROCESOS.....	22
1.2. DIAGNOSTICO ESTRATÉGICO	26
1.2.1. MISIÓN.....	26
1.2.2. VISIÓN	26
1.2.3. VALORES.....	26
1.2.4. ANALISIS FODA	27
1.2.4.1. ANALISIS INTERNO.....	27
1.2.4.2. ANALISIS EXTERNO.....	28
1.2.5. OBJETIVOS ESTRATÉGICOS IDENTIFICADOS.....	29
CAPITULO II	31
MARCO TEORICO.....	31
2.1. MINERIA DE DATOS.....	31
2.2. ANÁLISIS DISCRIMINANTE	35

2.3. INDICADORES DOING BUSINESS.....	42
2.4. IBM SPSS (Statistical Package for the Social Sciences).....	43
2.5. SAS Enterprise Miner (Herramienta para minería de datos).....	45
2.6. MINERÍA DE DATOS EN EL CONTROL DEL FRAUDE ADUANERO.	47
CAPITULO III	49
PROCESO DE TOMA DE DECISIONES	49
3.1 PLANTEAMIENTO DEL PROBLEMA	49
3.2 ALTERNATIVAS DE SOLUCIÓN	52
3.3 SELECCIÓN DE UNA ALTERNATIVA DE SOLUCIÓN.....	55
CAPITULO IV	60
IMPLEMENTACIÓN DEL SISTEMA	60
4.1. PLAN DE ACCION PARA EL DESARROLLO DE LA SOLUCION PLANTEADA	60
4.2. ANALISIS Y DISEÑO DE LOS MÓDULOS.....	63
4.3. MANUAL DE USUARIO DEL SISTEMA.....	70
CAPITULO V	78
EVALUACION DE RESULTADOS.....	78
CONCLUSIONES Y RECOMENDACIONES.....	90
CONCLUSIONES:.....	90
RECOMENDACIONES:	91
GLOSARIO Y TÉRMINOS	92
BIBLIOGRAFIA	93

DESCRIPTORES TEMÁTICOS

- ↪ Minería de Datos
- ↪ SAS
- ↪ Aduanas
- ↪ Canal de control
- ↪ Fiscalización Aduanera

LISTADO DE FIGURAS

- Grafica 01: Organigrama SUNAT (Fuente SUNAT 2010).
- Grafica 02: Macro procesos del Negocio (Fuente SUNAT 2010).
- Grafica 03: EDT Primer Entregable (Fuente SUNAT 2010).
- Grafica 04: EDT Segundo Entregable (Fuente SUNAT 2010).
- Grafica 05: Diagrama de Paquetes del Sistema de Selección (Fuente SUNAT 2010).
- Grafica 06: Diagrama de Casos de Uso del Sistema (Fuente SUNAT).
- Grafica 07: Diagrama de Casos de Uso SAM - SAC (Fuente SUNAT 2010)
- Grafica 08: Modelo conceptual de datos (Fuente SUNAT 2010)
- Grafica 09: Diagrama de Componentes del Sistema de Selección (Fuente SUNAT 2010)
- Grafica 10: Diagrama de Despliegue del Sistema de Selección (Fuente SUNAT 2010)

- GRAFICO 11: Diagrama de Flujo del proceso de Simulación (Fuente SUNAT 2010)
- Grafica 12: Interface para simular la selección de canal (Fuente SUNAT 2010)
- Grafica 13: Ejemplo de ingreso de datos Caso 1 (Fuente SUNAT 2010)
- Grafica 14: Ejemplo de ingreso de datos Caso 2 (Fuente SUNAT 2010)
- Grafica 15: Ejemplo de ingreso de datos Caso 3 (Fuente SUNAT 2010)
- Grafica 16: Ejemplo de ingreso de datos Caso 4 (Fuente SUNAT 2010)
- Grafica 17: Interface que muestra el resultado de la simulación (Fuente SUNAT 2010)
- Grafica 18: Histograma del Indicador 35 (Elaboración propia en base a los datos del indicador SUNAT).
- Grafica 19: Histograma del Indicador 58 (Elaboración propia en base a los datos del indicador SUNAT).
- Grafica 20: Histograma del Indicador 59 (Elaboración propia en base a los datos del indicador SUNAT).

RESUMEN EJECUTIVO

Las importaciones en el Perú han ido en constante crecimiento en los últimos años, provenientes de distintas partes del mundo. Las importaciones deben pasar por un proceso de control, para verificar que no contienen mercancías fraudulentas, y adicionalmente verificar que están cumpliendo con el pago de todos los tributos asociados a dicha importación.

En la situación planteada, para la SUNAT, existe un principal problema al momento de realizar la fiscalización de las importaciones, ya que se trata de una gran cantidad de importaciones, por lo que es inmanejable poder fiscalizar todas las importaciones que ingresan al País, ya que no se cuenta con la capacidad operativa, ni con el presupuesto para poder realizar una tarea como esta.

Adicionalmente la poca selectividad que existe en el sistema actual de control de riesgo aduanero, no permite a la SUNAT poder encontrar, en mayor porcentaje, aquellas mercancías fraudulentas, lo que genera una pérdida de tiempo y de recursos para la SUNAT.

Debido a todos estos problemas, la SUNAT optó por utilizar herramientas informáticas para mejorar la selectividad del actual sistema de control, tomando como referencia la experiencia que se tiene en otros países, como España y Chile. Para esto se presentaron diversas alternativas tecnológicas con el objetivo de implementar un nuevo sistema de Selección.

En el presente informe se analizarán 2 posibles alternativas que plantean dar solución a los problemas antes mencionados, una utilizando herramientas de Minería de Datos, y la otra utilizando herramientas estadísticas de predicción.

INTRODUCCIÓN

Las importaciones en el Perú han ido en constante crecimiento, debido al crecimiento económico que tiene el país. Esto genera que la carga de trabajo de las diferentes canales de entrada del país (Aduanas Aéreas y Marítimas), haya crecido de manera radical, debido a la gran cantidad de mercancías procedentes de diferentes países del mundo.

En este escenario es que nace la necesidad de poder contar con mecanismos de control para las importaciones, con el objetivo de poder realizar la tarea de fiscalización aduanera de una manera eficiente y eficaz, reduciendo los tiempos de espera en puerto, pero garantizando una correcta detección de los llamados fraudes aduaneros, así como también poder gestionar el riesgo del ingreso de insumos ilegales, o no autorizados.

Las herramientas para poder realizar el control de las importaciones actualmente no cumplen con los objetivos trazados para las aduanas y es por esto que es necesario analizar posibles soluciones para poder mejorar la selectividad de las herramientas actuales, las cuales determinan el canal de control de las mercancías, estos canales se tipifican de la siguiente manera:

Canal Verde: Las Mercancías seleccionadas a canal verde, descargan al Punto de Llegada y son de libre disponibilidad una vez cancelado los derechos arancelarios y demás tributos de importación.

Canal Naranja: Las declaraciones que por efectos del sistema hayan sido seleccionadas a este canal, estarán sometidas a revisión documentaria y son presentadas por el despachador de aduana dentro de las ocho (8) horas siguientes al ingreso de la mercancía al Punto de Llegada.

Canal Rojo: Las declaraciones seleccionadas a este canal estarán sujetas a reconocimiento físico y documental de las mercancías.

La capacidad operativa de las aduanas no es suficiente para poder realizar una revisión física (canal rojo) a todas las mercancías que ingresan al país, por lo que es necesario seleccionar aquellas mercancías que cumplen con ciertos criterios que presentan más riesgo de contener algún tipo de fraude aduanero.

Los Modelos de selección basados en herramientas de minería de datos y herramientas probabilísticas han demostrado tener un buen rendimiento en lo que respecta al control de las importaciones, ya que analizan una fuente de datos lo suficientemente significativa para pronosticar o predecir que mercancías, según ciertos criterios, pueden estar sujetas a fraude.

CAPITULO I

ANTECEDENTES

1.1. DIAGNÓSTICO FUNCIONAL

1.1.1. LA EMPRESA

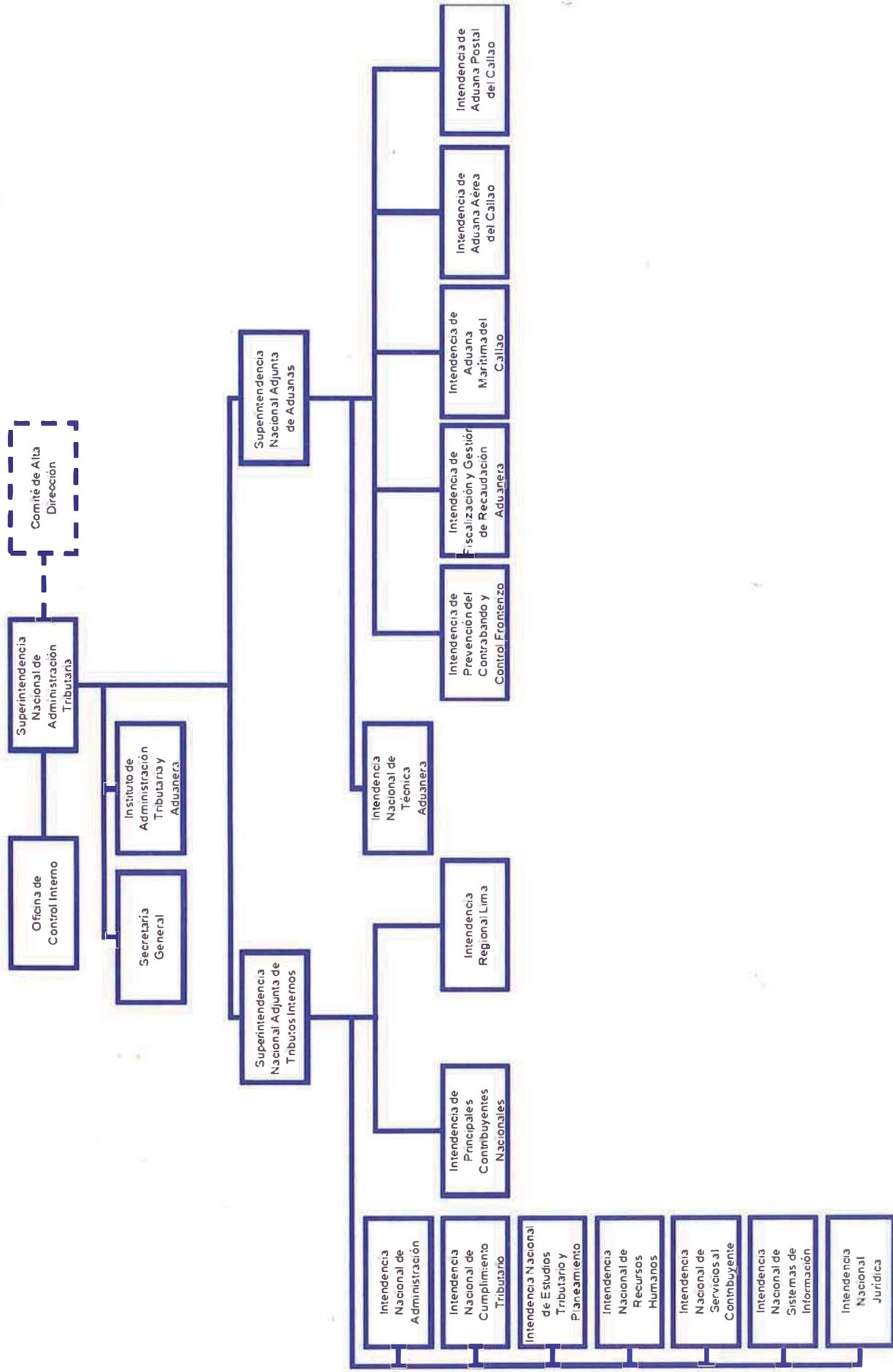
La Superintendencia Nacional de Administración Tributaria es, de acuerdo a su Ley de creación, Ley N° 24829 y a su Ley General aprobada por Decreto Legislativo N° 501, una Institución Pública descentralizada del Sector Economía y Finanzas, dotada de personería jurídica de Derecho Público, patrimonio propio y autonomía económica, administrativa, funcional, técnica y financiera que, en virtud a lo dispuesto por el Decreto Supremo N° 061-2002-PCM, expedido al amparo de lo establecido en el numeral 13.1 del artículo 13° de la Ley N° 27658, ha absorbido a la Superintendencia Nacional de Aduanas, asumiendo las funciones, facultades y atribuciones que por ley, correspondían a esta entidad.

De acuerdo a la Ley General de la Superintendencia Nacional de Administración Tributaria, ésta tiene domicilio legal y sede principal en la ciudad de Lima y puede establecer dependencias en cualquier lugar del territorio nacional.

(Artículo 3° del Reglamento de Organización y Funciones de la SUNAT, aprobado mediante D.S. 115-2002-PCM publicado el 28-10-02).

1.1.2. ORGANIGRAMA DE LA EMPRESA

En el siguiente gráfico se puede observar el organigrama de la SUNAT, en ella se aprecia la jerarquía vertical existente bajo la cual funciona la organización.



Grafica 01: Organigrama SUNAT (Fuente SUNAT 2010)

1.1.3. SERVICIOS

Con el fin de lograr un sistema tributario eficiente, permanente y simple se dictó la Ley Marco del Sistema Tributario Nacional (Decreto Legislativo N° 771), vigente a partir del 1 de enero de 1994.

La ley señala los tributos vigentes e indica quiénes son los acreedores tributarios: el Gobierno Central, los Gobiernos Locales y algunas entidades con fines específicos.

En aplicación del Decreto Supremo 061-2002-PCM, publicado el 12 de julio del 2002, se dispone la fusión por absorción de la Superintendencia Nacional de Aduanas (SUNAD) por la Superintendencia Nacional de Administración Tributaria (SUNAT), pasando la SUNAT a ser el ente administrador de tributos internos y derechos arancelarios del Gobierno Central.

Los principales tributos que administra la SUNAT son los siguientes:

1. Impuesto General a las Ventas: Es el impuesto que se aplica en las operaciones de venta e importación de bienes, así como en la prestación de distintos servicios comerciales, en los contratos de construcción o en la primera venta de inmuebles.
2. Impuesto a la Renta: Es aquél que se aplica a las rentas que provienen del capital, del trabajo o de la aplicación conjunta de ambos.
3. Régimen Especial del Impuesto a la Renta: Es un régimen tributario dirigido a personas naturales y jurídicas, sucesiones indivisas y sociedades conyugales domiciliadas en el país que obtengan rentas de tercera categoría provenientes de las actividades de comercio y/o industria; y actividades de servicios.

4. Nuevo Régimen Único Simplificado: Es un regimen simple que establece un pago único por el Impuesto a la Renta y el Impuesto General a las Ventas (incluyendo al Impuesto de Promoción Municipal). A él pueden acogerse únicamente las personas naturales o sucesiones indivisas, siempre que desarrollen actividades generadoras de rentas de tercera categoría (bodegas, ferreterías, bazares, puestos de mercado, etc.) y cumplan los requisitos y condiciones establecidas.
5. Impuesto Selectivo al Consumo: Es el impuesto que se aplica sólo a la producción o importación de determinados productos como cigarrillos, licores, cervezas, gaseosas, combustibles, etc.
6. Impuesto Extraordinario para la Promoción y Desarrollo Turístico Nacional: Impuesto destinado a financiar las actividades y proyectos destinados a la promoción y desarrollo del turismo nacional.
7. Impuesto Temporal a los Activos Netos: Impuesto aplicable a los generadores de renta de tercera categoría sujetos al régimen general del Impuesto a la Renta, sobre los Activos Netos al 31 de diciembre del año anterior. La obligación surge al 1 de enero de cada ejercicio y se paga desde el mes de abril de cada año.
8. Impuesto a las Transacciones Financieras: El Impuesto grava algunas de las operaciones que se realizan a través de las empresas del Sistema Financiero. Creado por el D.Legislativo N° 939 y modificado por la Ley N° 28194. Vigente desde el 1° de marzo del 2004.
9. Casinos y Tragamonedas: Impuestos que gravan la explotación de casinos y máquinas tragamonedas.

10. Derechos Arancelarios o Ad Valorem, son los derechos aplicados al valor de las mercancías que ingresan al país, contenidas en el arancel de aduanas.
11. Derechos Específicos, son los derechos fijos aplicados a las mercancías de acuerdo a cantidades específicas dispuestas por el Gobierno.
12. Aportaciones al ESSALUD y a la ONP: Mediante la Ley N° 27334 se encarga a la SUNAT la administración de las citadas aportaciones, manteniéndose como acreedor tributario de las mismas el Seguro Social de Salud (ESSALUD) y la Oficina de Normalización Previsional (ONP).
13. Regalías Mineras: Se trata de un concepto no tributario que grava las ventas de minerales metálicos y no metálicos. El artículo 7° de la Ley 28258 - Ley de Regalías Mineras, autoriza a la SUNAT para que realice, todas las funciones asociadas al pago de la regalía minera.

1.1.4. FUNCIONES

Las funciones y atribuciones de la Superintendencia Nacional De Administración Tributaria son las siguientes:

1. Administrar, recaudar y fiscalizar los tributos internos del Gobierno Nacional, con excepción de los municipales, así como las aportaciones al Seguro Social de Salud (ESSALUD) y a la Oficina de Normalización Previsional (ONP), y otros cuya recaudación se le encargue de acuerdo a ley.
2. Proponer al Ministerio de Economía y Finanzas la reglamentación de las normas tributarias y aduaneras.

3. Expedir, dentro del ámbito de su competencia, disposiciones en materia tributaria y aduanera, estableciendo obligaciones de los contribuyentes, responsables y/o usuarios del servicio aduanero, disponer medidas que conduzcan a la simplificación de los regímenes y trámites aduaneros, así como normar los procedimientos que se deriven de éstos.
4. Sistematizar y ordenar la legislación e información estadística de comercio exterior, así como la vinculada con los tributos internos y aduaneros que administra.
5. Proponer al Poder Ejecutivo los lineamientos tributarios para la celebración de acuerdos y convenios internacionales, así como emitir opinión cuando ésta le sea requerida.
6. Celebrar acuerdos y convenios de cooperación técnica y administrativa en materia de su competencia.
7. Promover, coordinar y ejecutar actividades de cooperación técnica, de investigación, de capacitación y perfeccionamiento en materia tributaria y aduanera, en el país o en el extranjero.
8. Otorgar el aplazamiento y/o fraccionamiento para el pago de la deuda tributaria o aduanera, de acuerdo con la Ley.
9. Solicitar, y de ser el caso ejecutar, medidas destinadas a cautelar la percepción de los tributos que administra y disponer la suspensión de las mismas cuando corresponda.
10. Controlar y fiscalizar el tráfico de mercancías, cualquiera sea su origen y naturaleza a nivel nacional.

11. Inspeccionar, fiscalizar y controlar las agencias de aduanas, despachadores oficiales, depósitos autorizados, almacenes fiscales, terminales de almacenamiento, consignatarios y medios de transporte utilizados en el tráfico internacional de personas, mercancías u otros.
12. Prevenir, perseguir y denunciar al contrabando, la defraudación de rentas de aduanas, la defraudación tributaria, el tráfico ilícito de mercancías, así como aplicar medidas en resguardo del interés fiscal.
13. Desarrollar y aplicar sistemas de verificación y control de calidad, cantidad, especie, clase y valor de las mercancías, excepto las que estén en tránsito y transbordo, a efectos de determinar su clasificación en la nomenclatura arancelaria y los derechos que le son aplicables.
14. Desarrollar y administrar los sistemas de análisis y fiscalización de los valores declarados por los usuarios del servicio aduanero.
15. Resolver asuntos contenciosos y no contenciosos y, en este sentido, resolver en vía administrativa los recursos interpuestos por los contribuyentes o responsables; conceder los recursos de apelación y dar cumplimiento a las Resoluciones del Tribunal Fiscal, y en su caso a las del Poder Judicial.
16. Sancionar a quienes contravengan las disposiciones legales y administrativas de carácter tributario y aduanero, con arreglo a Ley.
17. Ejercer los actos y medidas de coerción necesarios para el cobro de deudas por los conceptos indicados en el inciso precedente.
18. Mantener en custodia los bienes incautados, embargados o comisados, efectuando el remate de los mismos cuando ello proceda en el ejercicio de sus funciones.

19. Adjudicar directamente, como modalidad excepcional de disposición de mercancías, aquellas que se encuentren en abandono legal y en comiso administrativo. La adjudicación se hará a las entidades estatales y a aquellas a las que oficialmente se les reconozca fines asistenciales o educacionales, sin fines de lucro.
20. Desarrollar programas de información, divulgación y capacitación en materia tributaria y aduanera.
21. Editar, reproducir y publicar el Arancel Nacional de Aduanas actualizado, los tratados y convenios de carácter aduanero, así como las normas y procedimientos aduaneros.
22. Determinar la correcta aplicación y recaudación de los tributos aduaneros y de otros cuya recaudación se le encargue de acuerdo a ley, así como de los derechos que cobre por los servicios que presta.
23. Participar en la celebración de Convenios y Tratados Internacionales que afecten a la actividad aduanera nacional y colaborar con los Organismos Internacionales de carácter aduanero.
24. Ejercer las demás funciones que sean compatibles con la finalidad de la Superintendencia Nacional de Administración Tributaria.

La Superintendencia Nacional de Administración Tributaria (SUNAT) ejercerá las funciones antes señaladas respecto de las aportaciones al Seguro Social de Salud (ESSALUD) y a la Oficina de Normalización Previsional (ONP), a las que hace referencia la Norma II del Título Preliminar del Texto Único Ordenado del Código Tributario, aprobado por el Decreto Supremo N° 135-99-EF.

La SUNAT también podrá ejercer facultades de administración respecto de otras obligaciones no tributarias de ESSALUD y de la ONP, de acuerdo a lo que se establezca en los convenios interinstitucionales correspondientes. (Artículo 15° del Reglamento de Organización y Funciones de la SUNAT, aprobado mediante D.S. 115-2002-PCM publicado el 28-10-02).

1.1.5. CLIENTES

Los clientes de la Superintendencia Nacional de Administración Tributaria, se pueden agrupar en 4 grupos:

- Operador Tributario

- Operador y Usuario de Comercio Exterior, como son los agentes de aduanas a nivel nacional.

- Entidades del estado, como son los Ministerios, Essalud, etc.

- Ciudadanía, los ciudadanos en general que cumplen con sus obligaciones tributarias.

1.1.6. PROVEEDORES

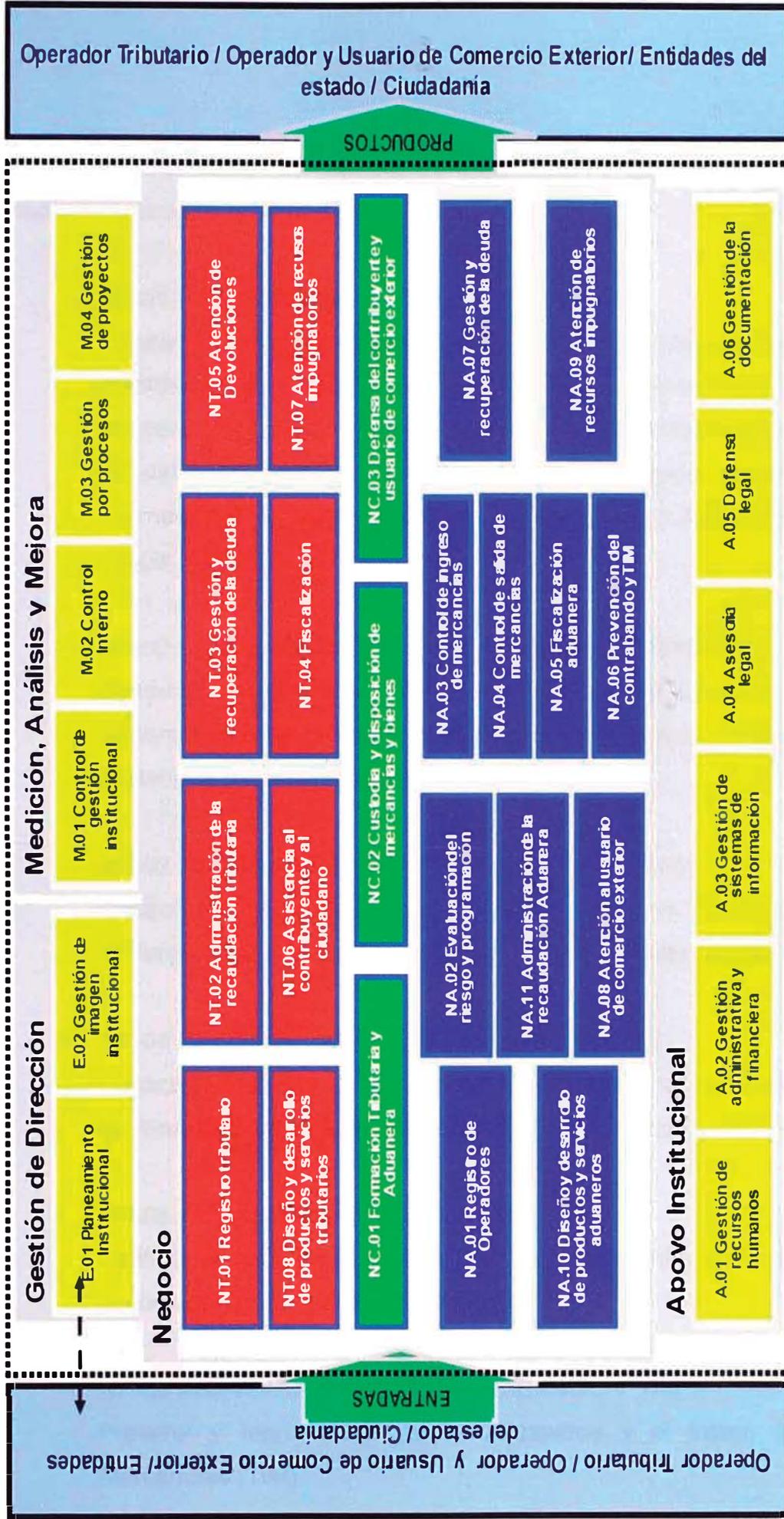
Al tratarse de un proceso de recaudación anual, y repetitivo en el tiempo, los proveedores de la Superintendencia Nacional de Administración Tributaria, serían igualmente los clientes,

Operador Tributario / Operador y Usuario de Comercio Exterior/ Entidades del estado / Ciudadanía.

1.1.7. PROCESOS

A continuación listaremos todos los Macro procesos y subprocesos con los cuales trabaja la **SUNAT**.

SUNAT - MAPA DE MACROPROCESOS INSTITUCIONAL



Grafica 02: Macro proceso del Negocio (Fuente SUNAT 201)

Una breve explicación del objetivo de cada Macro proceso relacionados al negocio aduanero, sería el siguiente:

→ *NA-01 REGISTRO DE OPERADORES*

Brindar a los diferentes operadores de comercio exterior la autorización para ejercer sus actividades, así como la identificación del personal correspondiente. Registrar las autorizaciones recibidas de otros sectores (en el caso de las empresas de transporte internacional terrestre, solo se registran las autorizaciones que otorga el MTC).

→ *NA-02 EVALUACIÓN DEL RIESGO Y PROGRAMACIÓN*

Determinar el tipo de control y alcance a aplicar a los usuarios de comercio exterior, respecto a: Ingreso/ Salida de mercancías, control posterior y prevención de contrabando.

→ *NA-03 CONTROL DE INGRESO DE MERCANCÍAS*

Control de ingreso de mercancías destinadas o no. Cumpliendo con las formalidades aduaneras o su disposición cuando corresponda.

→ *NA-04 CONTROL DE SALIDA DE MERCANCÍAS*

Control de la salida de mercancías destinadas o no, cumpliendo con las formalidades aduaneras o su disposición cuando corresponda.

→ *NA-05 FISCALIZACIÓN ADUANERA*

Verificar el cumplimiento de las obligaciones tributarias y aduaneras de los operadores de comercio exterior

→ *NA-06 PREVENCIÓN DEL CONTRABANDO Y TIM*

Prevenir y reprimir los delitos aduaneros y el tráfico ilícito de mercancías (TIM)

- *NA-07 GESTIÓN Y RECUPERACIÓN DE LA DEUDA*
Asegurar el cumplimiento de pago de la deuda aduanera y/o su extinción

- *NA-08 ATENCIÓN AL USUARIO DE COMERCIO EXTERIOR*
Suministrar información y orientación al usuario y al público en general, respecto a temas aduaneros

- *NA-09 ATENCIÓN DE RECURSOS IMPUGNATORIOS*
Atender los recursos impugnatorios (Reclamaciones, Reconsideraciones, Apelaciones y queja) presentados por los Operadores / usuarios de comercio exterior.

- *NA-10 DISEÑO Y DESARROLLO DE PRODUCTOS Y SERVICIOS ADUANEROS*
Diseñar, desarrollar y normar los productos y servicios relacionados a las actividades de despacho de mercancías, gestión de riesgo, acciones de control, recaudación y prevención del contrabando.

- *NA-11 ADMINISTRACIÓN DE LA RECAUDACIÓN ADUANERA*
Controlar la recaudación, con la finalidad de informar diariamente a los entes beneficiarios de los fondos abonados en sus respectivas cuentas.

1.2. DIAGNOSTICO ESTRATÉGICO

1.2.1. MISIÓN

“Gestionar integralmente el cumplimiento de las obligaciones tributarias y aduaneras, así como la facilitación del comercio exterior, de forma eficiente, transparente, legal y respetando al contribuyente o usuario.”.

1.2.2. VISIÓN

“Constituirse en una institución moderna e innovadora que facilita el comercio exterior y el efectivo cumplimiento tributario y aduanero brindando servicios de excelencia”

1.2.3. VALORES

→ Integridad.

Entendida como la capacidad para ser coherente con los principios y acciones de la Institución. En nuestro trabajo un comportamiento íntegro supone mantener normas y principios éticos, mediante un comportamiento honesto de vocación de servicio. Preservar la buena imagen de servidor público, cuidando que en todo momento nuestra vida privada y comportamiento sea coherente con ella.

→ Excelencia.

Disposición a brindar nuestro mayor esfuerzo para obtener los mejores resultados de manera sostenida en el tiempo, logrando la satisfacción de los contribuyentes, usuarios del comercio exterior y trabajadores, utilizando eficientemente los recursos y optimizando la calidad de nuestros servicios.

→ Compromiso Institucional y Personal.

El Compromiso institucional implica cumplir con nuestras obligaciones laborales haciendo más de lo esperado, con lograr los objetivos. El compromiso personal supone encaminar nuestros actos para fortalecernos como personas en el plano profesional y moral.

1.2.4. ANALISIS FODA

El Análisis FODA de la organización podemos dividirlo en el Análisis Interno y Externo, que mostramos a continuación:

1.2.4.1. ANALISIS INTERNO

→ Fortalezas 2007 – 2011

ITEM	FORTALEZAS 2007 - 2011
F1	Personal con conocimiento técnico de los procesos tributario y aduanero
F2	Acceso a tecnología de última generación
F3	Alto grado de compromiso del estado en los diversos proyectos de la institución (SUNAT).

→ Debilidades 2007 – 2011

ITEM	DEBILIDADES 2007 - 2011
D1	Grado de calidad de datos insuficiente
D2	Insatisfacción de los usuarios
D3	Faltan métricas de calidad
D4	No se cuenta con un diseño de procesos integrado y como consecuencia no se cuenta con un diseño de sistemas de información integrado

D5	Proceso de desarrollo de Sistemas no estandarizado.
D6	Proceso de control de calidad no estandarizado.
D7	Falta de cultura de trabajo en equipo
D8	Los ambientes de desarrollo y pruebas no son los adecuados
D09	Falta potencias y/o actualizar las competencias del personal
D10	Diferentes plataformas tecnológicas (TI y TA)
D11	Bajo nivel de automatización de procesos.

1.2.4.2. ANALISIS EXTERNO

→ Oportunidades 2007 – 2011

ITEM	OPORTUNIDADES 2007 - 2011
O1	Avance en el desarrollo del gobierno electrónico
O2	Apoyo de la Alta Dirección
O3	Se cuenta con presupuesto
O4	Mejoramiento de la gestión de la cartera de proyectos institucionales
O5	Maduración de tecnologías en el mundo
O6	Tercerización

→ Amenazas 2007 - 2011

ITEM	AMENAZAS 2007 - 2011
A1	Falta de un enfoque de sistemas y de Gestión por procesos institucional
A2	No está bien definido el proceso de captura de requerimientos.
A3	El día a día de las áreas operativas impide el involucramiento suficiente en los proyectos informáticos.

A4	La gestión de la cartera de proyectos Institucional de forma parcial
A5	Demora en los procesos de adquisiciones y contrataciones
A6	Plazos políticos, afectan la operatividad de atención de los requerimientos
A7	Falta de un modelo de procesos de negocio integrado.
A8	Demoras en los procesos de adquisición y contratación.
A9	Cambios constantes en las prioridades de los requerimientos.
A10	Definiciones de Requerimientos de baja Calidad.
A11	Falta Institucionalizar el proceso de Gestión de Requerimientos.
A12	Políticas de Personal insuficientes.
A13	Aprobación de una Cartera de Proyectos Institucionales desbalanceada o sub óptima.

1.2.5. OBJETIVOS ESTRATÉGICOS IDENTIFICADOS

Luego de un análisis por parte de la alta gerencia del corporativo y en base a los resultados obtenidos del análisis interno y externo de la empresa, los objetivos estratégicos identificados son:

OE-1: Promover el Cumplimiento Tributario Voluntario.

Mediante la facilitación del cumplimiento tributario, virtualizando los servicios y reduciendo los costos de cumplimiento de las obligaciones tributarias. Asimismo la generación de conciencia tributaria en los distintos segmentos de la sociedad.

OE-2: Facilitar el Comercio Exterior.

Contribuir a la competitividad del país mediante la facilitación del comercio exterior, reduciendo el tiempo de despacho así como, combatiendo la competencia desleal y el contrabando.

OE-3: Reducir el Incumplimiento Tributario.

Generar una efectiva sensación de riesgo, mediante la mejora de los mecanismos de detección de los ilícitos tributarios y aduaneros.

OE-4: Fortalecer la Institucionalidad de la SUNAT.

Implementar y dotar a la organización de las herramientas necesarias para el mejor cumplimiento de su rol en el estado y la sociedad.

CAPITULO II

MARCO TEORICO

2.1. MINERIA DE DATOS

La minería de datos puede definirse inicialmente como un proceso de descubrimiento de nuevas y significativas relaciones, patrones y tendencias al examinar grandes cantidades de datos.

La disponibilidad de grandes volúmenes de información y el uso generalizado de herramientas informáticas ha transformado el análisis de datos orientándolo hacia determinadas técnicas especializadas englobadas bajo el nombre de minería de datos.

Las técnicas de minería de datos persiguen el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en grandes bases de datos. Estas técnicas tienen como objetivo descubrir patrones, perfiles y tendencias a través de análisis de los datos utilizando tecnologías de reconocimiento de patrones, redes neuronales, lógica difusa, algoritmos genéticos y otras técnicas avanzadas de análisis de datos.

No obstante, la minería de datos es ya un concepto muy evolucionado que necesita ser aproximado conceptualmente por etapas. Inicialmente la finalidad de los sistemas de información era recopilar información sobre una

parcela determinada para ayudar en la toma de decisiones. Con la informatización de las organizaciones y la aparición de aplicaciones de software operacionales, la finalidad principal de los sistemas de información era dar soporte a los procesos básicos de la organización (ventas, producción, personal, etc.). Una vez satisfecha la necesidad de tener un soporte informático para los procesos básicos de la organización (sistemas de información para la gestión), las organizaciones exigen nuevas prestaciones de los sistemas de información (sistemas de información para la toma de decisiones).

De esta forma han aparecido diferentes herramientas de negocio para la toma de decisiones que coexisten: EIS, OLAP, consultas e informes, y las propias herramientas de minería de datos.

Un EIS (*Executive Information System*) es un sistema de información y un conjunto de herramientas asociadas que proporcionan a los directivos acceso a la información de estado y sus actividades de gestión. Está especializado en analizar el estado diario de la organización (mediante indicadores clave) para informar rápidamente sobre cambios a los directivos. La información solicitada suele ser, en gran medida numérica (ventas semanales, nivel de stocks, balances parciales, etc.) y representada de forma gráfica al estilo de las hojas de cálculo.

Las herramientas OLAP (*On-Line Analytical Processing*), son más genéricas, funcionan sobre un sistema de información (transaccional o almacén de datos) y permite realizar agregaciones y combinaciones de los datos de manera mucho más complejas y ambiciosas, con objetivos de análisis más estratégico. Las herramientas OLAP están basadas, generalmente, en sistemas o interfaces multidimensionales, que presentan la información de una manera matricial. Las herramientas OLAP proporcionan facilidades para “manejar” y “transformar” los datos, producen otros datos (más agregados y

combinados) y son una gran ayuda para analizar los datos porque producen diferentes vistas de los mismos.

Los sistemas de informes o consultas avanzadas están basados, generalmente, en sistemas relacionales u objeto-relacionales y el resultado se presenta de forma tabular. Generalmente están implementados en bases de datos relacionales.

Las herramientas de minería de datos permiten extraer patrones, tendencias y regularidades para describir y comprender mejor los datos y para predecir comportamientos futuros. La minería de datos analiza los datos y el resto de herramientas citadas anteriormente facilitan el acceso a la información para que el análisis sea más efectivo, es decir, son instrumentos de apoyo a la minería de datos.

No obstante las herramientas anteriormente citadas suelen necesitar de la existencia previa de un almacén de datos. El almacén de datos es el sistema de información central en todo este proceso. Un almacén de datos es una colección de datos orientada a un dominio, integridad, no volátil y variante en el tiempo para ayudar en la toma de decisiones. Un almacén de datos es un conjunto de datos históricos, internos o externos y descriptivos de un contexto o área de estudio, que están integrados y organizados de tal forma que permiten aplicar eficientemente herramientas para resumir, describir y analizar los datos con el fin de ayudar en la toma de decisiones estratégicas.

TECNICAS DE MINERIA DE DATOS:

La clasificación inicial de las técnicas de minería de datos distingue entre técnicas predictivas, en las que las variables pueden clasificarse inicialmente en dependientes e independientes (similares a las técnicas del análisis de la dependencia o métodos explicativos del análisis multivariante), técnicas descriptivas, en las que todas las variables tienen inicialmente el mismo

estatus (similares a las técnicas del análisis de la interdependencia o métodos descriptivos del análisis multivariante) y técnicas auxiliares.

Las *técnicas predictivas* especifican el modelo para los datos en base a un conocimiento teórico previo. El modelo supuesto para los datos debe contrastarse después del proceso de minería de datos antes de aceptarlo como válido. Formalmente, la aplicación de todo modelo debe superar las fases de identificación objetiva (a partir de los datos se aplican reglas que permiten identificar el mejor modelo posible que se ajuste a los datos), estimación (proceso de cálculo de los parámetros del modelo elegido para los datos en la fase de identificación), diagnóstico (proceso de contraste de la validez del modelo estimado) y predicción (proceso de utilización del modelo identificado, estimado y validado para predecir valores futuros de las variables dependientes). En algunos casos el modelo se obtiene como mezcla del conocimiento obtenido antes y después de la Minería de Datos y también debe contrastarse antes de aceptarse como válido. Por ejemplo las *redes neuronales* permiten descubrir modelos complejos y afinarlos a medida que progresa la exploración de datos. Gracias a su capacidad de aprendizaje, permiten descubrir relaciones complejas entre variables sin ninguna intervención externa. Podemos incluir entre estas técnicas todos los tipos de regresión, series temporales, análisis de la varianza y covarianza, árboles de decisión, redes neuronales, algoritmos genéticos y técnicas bayesianas. Tanto los árboles de decisión, como las redes neuronales son a su vez técnicas de clasificación que pueden extraer perfiles de comportamiento o clases, siendo el objetivo construir un modelo que permita clasificar cualquier nuevo dato.

Las *técnicas descriptivas* no asignan ningún papel predeterminado a las variables. No se supone la existencia de variables dependientes ni independientes y tampoco se supone la existencia de un modelo previo para los datos. Los modelos se crean automáticamente partiendo del reconocimiento de patrones. En este grupo se incluyen las técnicas de

agrupamiento y segmentación (que también son técnicas de clasificación en cierto modo), las técnicas de asociación y dependencia, las técnicas de análisis exploratorio de datos y las técnicas de reducción de la dimensión (factorial, componentes principales, correspondencias, etc.) y de escalamiento multidimensional.

Tanto las técnicas predictivas como las técnicas descriptivas están enfocadas al descubrimiento del conocimiento embebido en los datos.

Las técnicas auxiliares son herramientas de apoyo más superficiales y limitadas. Se trata de nuevos métodos basados en técnicas estadísticas descriptivas, consultadas e informes y enfocados en general hacia la *verificación*.¹

2.2. ANÁLISIS DISCRIMINANTE

El análisis discriminante como la regresión logística son técnicas multivariantes de dependencia como la regresión lineal múltiple. La diferencia inicial consiste en que la variable dependiente o criterio es una variable no métrica (categórica).

En primer lugar se especifican los objetivos que se procuran cuando se usa la técnica, después se presentan los supuestos que deben ser considerados para proceder en el uso de la técnica y por último se comenta, de manera sucinta, el proceso usado.

¹ César Perez Lopes – Minería de datos Técnicas y Herramientas
Editorial: Thomson International (Junio 2007)

La Selección de los Casos para el Estudio

Para discriminar entre grupos es necesario contar con un conjunto de casos que hayan sido claramente identificados como componentes indudables de uno de los grupos.

Por lo que resulta de suma importancia seleccionar de manera adecuada los casos que servirán como criterios de selección. De modo que el objetivo de esta fase es identificar en la muestra los casos que habiendo sido ubicados de manera apropiada y precisa como pertenecientes a uno de los grupos de clasificación, dispongan de información completa respecto a las variables que serán usadas para definir los grupos (variables predictoras).

La Verificación de la Diferencia entre los Grupos

Resulta necesario verificar que en efecto los grupos son diferentes, es decir que son excluyentes entre sí. Esto se puede observar por medio de:

- La diferencia entre las medias de los grupos. Se entiende que si los grupos son excluyentes, sus medias deben ser significativamente diferentes.
- El estadístico Lambda de Wilks que resulta ser la razón entre la suma de los cuadrados dentro de los grupos y la suma total de los cuadrados. Un valor Lambda de uno ocurre cuando la media de todos los grupos observados es igual. Lo que significa que a medida que el valor Lambda se acerca a cero, la mayor parte de la variabilidad es atribuible a la diferencia entre las medias de los grupos.

Otro aspecto muy importante a considerar es el grado de interdependencia existente entre las variables dado que los análisis multivariados son muy afectados por la interdependencia.

Esto se puede observar en el SPSS mediante una matriz de correlación en la cual es importante identificar las variables cuyas correlaciones son

grandes, casos que deben constituir una notable minoría; de otra manera es muy probable que no se cuente con material apropiado para realizar el trabajo propuesto.

La Estimación de los Coeficientes b

Una vez determinada la legitimidad de las variables, los casos y los grupos en estudio, se debe proceder a analizar las variables en conjunto.

El análisis discriminante, como otros procedimientos estadísticos multivariados, procura resumir en un índice, toda la información contenida en diversas variables independientes. Para lograr dicha finalidad se buscará estimar el peso de las variables, que permita lograr la mejor separación posible entre grupos.

La ecuación lineal discriminante es similar a la regresión múltiple:

$$D = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_pX_p$$

Siendo X los valores de las variables predictoras y b los coeficientes estimados a partir de los datos. En el análisis discriminante se busca separar lo más posible (diferenciar) a los promedios de cada grupo, pero respecto a la dispersión de los datos. Entre mayor dispersión de datos exista, será mayor el empalme de los datos, por lo que en al valorar la función, es necesario tomar en cuenta la dispersión o variación de los datos, además de la distancia entre los promedios.

A fin que la función lineal discriminante pueda separar o diferenciar entre los grupos, estos deben diferir en sus valores D . Por lo tanto los coeficientes (b) se escogen de manera que los valores de la función discriminante difieran tanto como sea posible entre los grupos. $\{Máxima S (m_1 - m_2)^2 / s^2\}$.

El Cálculo de los Valores Discriminantes

Para cada uno de los casos estudiados se puede calcular un puntaje (valor) discriminante D haciendo uso de los coeficientes de la función. Se los multiplica por el correspondiente valor de la variable en cada caso y luego se suman los productos.

Una vez determinado el puntaje (valor) discriminante D , resulta posible obtener una regla para clasificar cada caso en uno de los grupos. Con ese fin se usa la regla de Bayes, por medio de la cual se estima la probabilidad de que un caso en particular con un puntaje discriminante D pertenezca a un grupo en particular. En este procedimiento se distinguen tres tipos de probabilidades:

- La probabilidad previa o $P(G_i)$ Calculada en una muestra representativa de la población como la proporción observada de los casos correspondientes a cada grupo.
- La probabilidad condicional o $P(D/G_i)$ La cual calcula si los valores D están distribuidos de manera normal para cada grupo y los parámetros de las distribuciones pueden ser estimados. Esta representa la probabilidad de obtener un valor D particular si el caso observado es miembro de un grupo determinado.
- La probabilidad posterior o $P(G_i/D)$ La probabilidad condicional de D dado el grupo, provee una idea de los valores correspondientes a los miembros de un grupo en particular. Pero cuando no se conoce a qué grupo pertenece, se requiere estimar cuánto se asemeja a la membresía en los diferentes grupos a partir de la información disponible. A esto es lo que se denomina probabilidad posterior $P(G_i/D)$ y se puede estimar a partir de $P(D/G_i)$ y $P(G_i)$ usando la regla de Bayes. El caso en observación se clasificará en base a su puntaje discriminante D , en el grupo para el cual la probabilidad posterior se más grande.

Coefficientes de Funciones Discriminantes

Existen diferentes conjuntos de coeficientes según el tipo de función lineal, de los cuales mencionaremos dos:

- Coeficientes de la función lineal discriminante también conocidos como coeficientes de la función canónica discriminante siendo que son idénticos a los que se obtienen por medio de un análisis canónico de correlación. Se trata de pesos discriminantes son determinados por la estructura de varianza de las variables originales a través de los grupos del a variable dependiente (criterio). Las variables independientes con un poder discriminante grande por lo general presentan pesos grandes y las que tienen poco poder discriminante presentan pesos pequeños. Aunque la existencia de multicolinealidad entre las variables independientes puede conducir a una excepción de la regla.

- Coeficientes de la función lineal discriminante de Fisher o coeficientes de clasificación que pueden ser usados directamente para clasificar. Se obtiene un conjunto de coeficientes para cada uno de los grupos y se asigna cada caso al grupo para el cual tiene el puntaje discriminante D más grande. Se trata de un método por el cual se determina una función lineal para cada grupo en la variable dependiente. La clasificación logra calculando un puntaje para cada observación en la función de clasificación de cada grupo y posteriormente asignado cada caso al grupo en el que tiene el puntaje más alto. Esto difiere de lo que se conoce como puntaje Z de discriminación el cual se obtiene para cada función discriminante.

Los resultados de la clasificación son iguales para ambos métodos si se usan todas las funciones canónicas discriminantes (Kshirsager & Arseven, 1975; Green, 1979).

La prueba de hipótesis para el modelo discriminante: si la diferencia en los centroides (los promedios de cada grupo de las variables independientes) o las distancias que presentan los promedios de las variables involucradas en el modelo son significativas. Si en efecto las diferencias son significativas, esto quiere decir que el modelo es bueno para discriminar. Se puede usar la prueba Mahalanobis-D2 en la que se obtiene un valor F calculado para comparar con un F de la tabla la hipótesis nula dice: el modelo discriminante no es bueno para discriminar o clasificar observaciones.

Si la función discriminante es estadísticamente significativa y su capacidad para clasificar es aceptable, el investigador puede dedicarse a realizar interpretaciones sustantivas de sus resultados. Este proceso involucra el examen de la función discriminante para determinar la importancia relativa de cada una de las variables independientes en su capacidad para discriminar entre los grupos. Los tres métodos para determinar la importancia relativa son:

1. Los pesos discriminantes estandarizados. La forma tradicional de interpretar la función discriminante consiste en observar el signo y la magnitud de los pesos discriminantes estandarizados asignados a cada una de las variables en la función. Cuando se ignora el signo, cada peso representa la contribución relativa de su variable a la función. Las variables independientes con mayor peso contribuyen más al poder discriminante de la función. La forma de interpretar esto es análoga a los pesos beta del análisis de regresión y por consiguiente está sujeto a las mismas críticas.
2. La carga discriminante (estructura de correlaciones) miden la correlación lineal simple entre cada variable independiente y la función discriminante. La carga discriminante refleja la varianza que la variable independiente comparte con la función discriminante y puede ser

interpretada como carga factorial al evaluar la contribución relativa de cada variable independiente respecto a la función discriminante.

3. Los valores F parciales. Cuando se utiliza este método se obtiene una forma adicional para interpretar el poder discriminante relativo de las variables independientes por medio del uso de la F parcial. Esto se logra examinando el tamaño absoluto de los valores F significativos y ordenándolos en de mayor a menor. Grandes valores F indican mayor poder discriminatorio. Esto ofrece la oportunidad de observar el nivel de significatividad de cada una de las variables.²

² Lic. Tevni Grajales Guerra – El Análisis Discriminante
Disponibile en: <http://tgrajales.net/estdiscriminante.pdf>

2.3. INDICADORES DOING BUSINESS

El Proyecto *Doing Business* proporciona una medición objetiva de las normas que regulan la actividad empresarial y su puesta en práctica en 183 economías y ciudades seleccionadas en el ámbito subnacional y regional.

El Proyecto *Doing Business* comenzó en 2002 y analiza y compara las normas que regulan las actividades de las pequeñas y medianas empresas locales en 183 economías. El modelo de *Doing Business* y el modelo de costo estándar son las únicas herramientas de carácter homogéneo que estudian un amplio número de jurisdicciones para cuantificar el impacto de la legislación local en la actividad empresarial.

Al recopilar y analizar detalladamente datos cuantitativos para comparar en el tiempo los marcos reguladores de distintas jurisdicciones, *Doing Business* estimula cierto tipo de competencia entre las economías analizadas. También ofrece índices ponderables para reformar y constituye un recurso útil para miembros de la academia, periodistas, investigadores del sector privado y otras personas interesadas en el clima empresarial de cada país.

Además, *Doing Business* ofrece detallados reportes del ámbito subnacional, que cubren de manera exhaustiva la regulación y las reformas de las empresas en diferentes ciudades y regiones de una economía o país. Estos reportes proporcionan datos sobre la facilidad para hacer negocios, clasifican cada localidad y recomiendan reformas para mejorar los resultados en cada una de las áreas de los indicadores. Algunas ciudades pueden comparar sus regulaciones empresariales con otras ciudades vecinas y con las 183 economías que ha clasificado *Doing Business*.

El primer informe *Doing Business*, publicado en 2003, cubría cinco grupos de indicadores en 133 economías. El informe 2011 cubre once grupos de

indicadores en 183 economías. El proyecto se ha beneficiado de la información aportada por gobiernos, la academia, profesionales en ejercicio entre otros. Sin importar los años que han pasado y las distintas adiciones, es importante resaltar que el objetivo del proyecto sigue siendo el mismo: proporcionar una base de datos objetiva para analizar, comprender y mejorar las normas que regulan la actividad empresarial.

2.4. IBM SPSS (Statistical Package for the Social Sciences)

IBM SPSS ofrece los procedimientos estadísticos principales que los gestores y los analistas necesitan para tratar problemas empresariales y de investigación básicos. Este software proporciona herramientas que permiten a los usuarios consultar datos y formular hipótesis de forma rápida, así como ejecutar procedimientos para ayudar a entender las relaciones entre variables, identificar tendencias y realizar predicciones.

IBM SPSS incluye las siguientes funciones clave:

1. Modelos lineales

- Incluye modelos mixtos lineales generalizados para utilizar con datos jerárquicos.
- Este aplicativo dispone de procedimientos de modelos lineales generales (GLM) y de modelos mixtos.
- Incluye modelos lineales generalizados (GENLIN), incluidos los modelos estadísticos más utilizados como la regresión lineal para respuestas distribuidas con normalidad, modelos logísticos para datos binarios y modelos lineales para datos de recuento. GENLIN también ofrece varios modelos estadísticos útiles a través de su propia formulación general de modelos.

- Los procedimientos de las ecuaciones de estimación generalizadas (GEE) amplían los modelos lineales generalizados para acomodar datos longitudinales correlacionados.

2. Modelos no lineales

- La regresión logística multinomial (MLR) prevé resultados categóricos con más de dos categorías.
- La regresión logística binaria clasifica los datos en dos grupos.
- La regresión no lineal (NLR) y la regresión no lineal restringida (CNLR) estima parámetros de modelos no lineales.

3. Tablas personalizadas

- Se comparan medias o proporciones para grupos demográficos, segmentos de clientes, periodos de tiempo u otras variables categóricas al incluir estadísticas inferenciales.
- Incluye tres pruebas de significancia: prueba de chi cuadrado de independencia, comparación de medias de columnas (prueba t) o comparación de proporciones de columna (prueba z).
- Un programa de creación de tablas interactivo proporciona funciones de arrastrar y soltar para la creación de tablas dinámicas.
- Excluye categorías determinadas, muestra celdas de valor perdido y añade subtotales a las tablas.
- Las tablas pueden previsualizarse en tiempo real y modificarse a medida que se crean.
- Las tablas se pueden exportar a Microsoft Word, Excel, PowerPoint o HTML para utilizar en informes.³

³ Descripción del producto - SPSS Statistics Standard (IBM Software).
Disponibile en: <http://www-142.ibm.com/software/products/es/es/spss-stats-standard>

2.5. SAS Enterprise Miner (Herramienta para minería de datos).

SAS Enterprise Miner agiliza el proceso de minería de datos para crear modelos predictivos y descriptivos de alta precisión sobre la base de grandes volúmenes de datos de toda la empresa. Ofrece un conjunto de capacidades fáciles de usar. Con visión de futuro las organizaciones de hoy están utilizando el aplicativo SAS Enterprise Miner para detectar el fraude, minimizar los riesgos, anticiparse a las demandas de recursos, aumentar las tasas de respuesta y frenar la pérdida de clientes.

SAS Enterprise Miner ofrece los siguientes beneficios:

- **Apoyar el procesamiento de datos con un amplio conjunto de capacidades.** Independientemente de su nivel de habilidad, SAS proporciona un software flexible que aborda problemas complejos. Pasando de los datos en bruto a información precisa, orientada a los negocios.
- **Construir mejores modelos con un banco de trabajo de minería de datos versátil.** SAS Enterprise Miner incluye un sistema interactivo de auto documentación que reduce drásticamente el tiempo de desarrollo de un modelo de minería de datos.
- **Permitir a los analistas de negocios de forma rápida y fácil obtener ideas de una manera autosuficiente y automatizada.** El Complemento SAS para Microsoft Office (Excel solamente) permite a los usuarios de negocio y expertos en la materia, con escasos conocimientos estadísticos, generar automáticamente los modelos de predicción de escenarios de negocios comunes y actuar en forma rápida y eficaz.

- **Mejora la precisión de las predicciones y permite compartir fácilmente información fiable para mejorar la calidad de las decisiones.** Mejor rendimiento con los modelos de algoritmos modernos, innovadores y métodos específicos de la industria, esto mejora la estabilidad y la precisión de las predicciones, lo que puede ser verificado fácilmente por el modelo de evaluación visual.

- **Facilitar la implementación del modelo y proceso de calificación de puntuación.** El proceso de aplicación de un modelo con nuevos datos es el resultado final de mucho esfuerzo. SAS Enterprise Miner automatiza el tedioso proceso para todas las etapas del desarrollo del modelo. El código puede ser desplegado en tiempo real o por lotes dentro de los entornos SAS, en la web, directamente en bases de datos relacionales, o integrado a los procesos de negocio.⁴

⁴ Descripción del producto – SAS Enterprise Miner.
Disponible en: <http://www.sas.com/technologies/analytics/datamining/miner/index.html>

2.6. MINERÍA DE DATOS EN EL CONTROL DEL FRAUDE ADUANERO.

El principal recurso con el que cuenta cualquier organización para generar inteligencia y realizar una buena toma de decisiones es la información.

Esto es especialmente cierto en el complejo mundo de la Aduana en donde día a día se registran y almacenan grandes cantidades de datos relacionados con una infinidad de trámites aduaneros que requieren ser fiscalizados a fin de reconocer las operaciones apegadas a derecho de las fraudulentas.

Sin embargo esta gran cantidad de información necesita de herramientas estadísticas que permitan encontrar un orden y sentido en este mar de datos. Parte muy importante de las herramientas que hacen esto posible son las conocidas como “Minería de Datos” en alusión al trabajo de los mineros que mueven montañas de tierra para extraer pequeñas cantidades de metales preciosos, de igual manera, es como la minería de datos trabaja con gran cantidad de datos con el fin de obtener información útil para un fin determinado.

La minería de datos intenta ayudar a comprender el contenido de una base de datos a través de algoritmos matemáticos que buscan relaciones, patrones, comportamientos, agrupaciones, secuencias, tendencias o asociaciones, muchas veces poco evidentes, entre las diferentes variables que componen las bases de datos.

Esto se logra integrando las ventajas de varias disciplinas del llamado “proceso de extracción de conocimiento en bases de datos” tales como la Inteligencia Artificial, la Computación Gráfica, las Bases de Datos y el Procesamiento Masivo.

De forma general, los datos son la materia prima en bruto que necesita ser trabajada para poder obtener información de utilidad. En el momento que el usuario le atribuye algún significado especial a estos datos, los mismos pasan a convertirse en información y cuando esta información logra ser explicada por medio de un modelo matemático entonces la base de datos se ha convertido en conocimiento útil que puede utilizarse para extraer información relevante.

De esta forma, las técnicas de minería de datos nos permiten obtener información útil a partir de bases de datos gigantescas siendo aplicable a campos tan diversos como lo son el análisis bursátil, diagnóstico del cáncer, detección de fraude bancario y aduanero, detección de terroristas, entre otros.

Existen muchas aplicaciones basadas en gran medida en sistemas que incorporan algún proceso de minería de datos, un ejemplo claro que usamos diariamente, son los buscadores de Internet (Google, Yahoo, etc.) que antes de ofrecer los resultados de cualquier consulta deben haber procesado previamente la información de millones de páginas web para agruparlas por categorías y temáticas, esta recopilación y organización se realiza obviamente de forma automática, haciendo uso de herramientas de minería de datos.

De igual forma cuando realizamos una compra utilizando tarjetas de crédito, nuestra transacción es analizada en fracciones de segundo por sistemas informáticos con algoritmos de minería de datos que comparan la transacción en proceso con nuestro patrón usual de compra para detectar alguna variación que pueda hacer sospechar al sistema de que la persona que utiliza la tarjeta no es su dueño verdadero.⁵

⁵ Lic. Alberto Enrique Villalobos Chaves – Aplicación de la minería de datos en el control de fraude Aduanero.

Disponible en: <http://www.scribd.com/doc/11352087/Mineria-de-Datos-en-Aduanas>

CAPITULO III

PROCESO DE TOMA DE DECISIONES

3.1 PLANTEAMIENTO DEL PROBLEMA

Desde hace varios años se observa un crecimiento constante en las importaciones peruanas cuantificado tanto por el número de declaraciones así como por el monto total de importaciones. Este fuerte incremento en la cantidad de declaraciones de importación, hace necesaria la búsqueda de mayor selectividad en los diferentes modelos que se utilizan para la selección de canal de control.

Veamos los aspectos más importantes:

- La demora en la gestión de las importaciones afecta directamente a los indicadores *Doing business* del país, los cuales reflejan que tan rápido se hacen negocios en el país, uno de estos indicadores es el de “Comercio entre Fronteras”, que es el indicador que indica el tiempo que toma realizar el levante de la mercancía en aduanas.⁶
- Existen quejas de los usuarios (empresa Importadoras), que sienten que el tiempo de espera para poder realizar el desaduanaje de su mercancía

⁶ Indicadores Doing Business.
Disponible en: <http://www.doingbusiness.org/rankings>

a veces es mucho, aun cuando ellos cuentan con toda la documentación en regla.

- Los gastos operativos en la fiscalización y control de mercancías que no tienen incidencia relevante para la aduanas, generan costos innecesarios a la SUNAT.
- La poca selectividad del actual sistema, hace que se realice tareas de control innecesarias lo cual hace que constantemente se pierda tiempo, ya que existe una capacidad operativa fija, la cual no es realmente suficiente para la cantidad de importaciones que recibe el país.
- Finalmente un problema latente en el proceso, es el riesgo de manipulación de la información auditada, ya que al conocerse los criterios de selección, existen problemas de seguridad de la información muy graves.

La problemática asociada, se vincula con la aplicación de 2 objetivos estratégicos: **“OE-2: Facilitar el Comercio Exterior”** y **“OE-3: Reducir el Incumplimiento Tributario”**

En base a esto se identifican los siguientes problemas que deben ser atacados y atenuados lo más pronto posible.

P1. Ineficiencia en el actual sistema de selección, por los bajos niveles de selectividad con el que cuenta.

P2. Incremento de gastos operativos, debido al trabajo realizado en fiscalizar mercancías que no tienen incidencia relevante.

De estos problemas podemos integrarlos y plantear el problema de manera más general y concisa.

Problema:

Ineficiencia del sistema de selección de Riesgo Actual, el cual debido a su baja selectividad genera mayores gastos operativos y demora en el manejo de las importaciones que ingresan al País, esto perjudica directamente a clientes y proveedores.

3.2 ALTERNATIVAS DE SOLUCIÓN

Para el problema presentado, la solución planteada fue la siguiente: "Implementación de un nuevo sistema de selección de canal de control de Riesgo para las importaciones" debido a la experiencia de otros países como España y Chile en la implementación de modelos de minería de datos y métodos estadísticos.

Según esta necesidad, la gerencia de Sistemas Aduaneros en conjunto con la Gerencia de Sistemas Tributarios y de Gestión planteó 2 alternativas de solución para este requerimiento, las cuales fueron evaluadas para poder determinar cuál es la mejor opción a implementar.

- A. Implementar un Sistema de selección de canal de control utilizando minería de Datos.
- B. Implementar un Sistema de Selección de canal de control utilizando herramientas Estadísticas.

A continuación detallaremos cada una de estas alternativas identificadas:

A. Implementar un Sistema de selección de canal de control utilizando minería de Datos.

Se decidió esta alternativa debido a un estudio de expertos que según las experiencias observadas en otros países de la región, se planteó poder desarrollar y entrenar diversos modelos de minería de datos, utilizando la herramienta para el modelamiento y desarrollo con Minería de Datos, "SAS Enterprise Miner".

El modelo de minería podía ser entrenado con datos históricos para que así pueda inferir en base a esta información el canal de control adecuado para las DUAs.

Ventajas:

- Alta confiabilidad, los resultados son poco predecibles, ya que las redes neuronales son del tipo caja negra, solo sabemos los datos de entrada y de salida.
- Es dinámico, ya que los modelos de minería son reentrenados constantemente con información histórica de 24 meses hacia atrás.
- La fiscalización utilizando modelos de minería de datos ha sido efectiva en países como Chile y España.

Desventajas:

- No es transparente, ya que al ser una red neuronal, no sabemos al 100% el porqué de algún índice de riesgo obtenido.
- Poco conocimiento de las personas que conocen el negocio de la herramienta SAS, y de la generación de modelos de minería.

B. Implementar un Sistema de Selección de canal de control utilizando herramientas Estadísticas.

Esta segunda alternativa, también debido a un grupo de expertos que conoce temas estadísticos, consideró que un modelo basado en herramientas estadísticas podría predecir de mejor manera el

comportamiento de las importaciones, para esta solución se utilizaría la herramienta de análisis estadístico "SPSS".

Básicamente se estaría manejando un modelo basado en el análisis de discriminante, que según un número de variables puede inferir la selección del canal de control.

Ventajas:

- Las personas que conocen el negocio aduanero, tienen cierto grado de conocimiento en la herramienta SPSS, y el manejo de modelos estadísticos.
- Es abierto, ya que conjuga y pondera las variables predictoras determinando el nivel de riesgo de la Declaración y por consiguiente el canal.
- Transparencia, ya que elimina la discrecionalidad en la elaboración de Listas. Es Auditable.
- Es dinámico, ya que actualización y retroalimentación permanente de los coeficientes y rankings de riesgo sobre la base de información histórica.

Desventajas:

- Al ser un modelo basado en constantes definidas, es posible predecir en cierta forma el resultado, ya que tenemos acceso a la función discriminante, lo cual puede generar problemas en seguridad de la información.

3.3 SELECCIÓN DE UNA ALTERNATIVA DE SOLUCIÓN

Para proceder a realizar una selección adecuada de cual alternativa es la mejor opción a tomar, se procedió a utilizar la técnica de ponderación mediante Juicio de Expertos, para esto se conto con el equipo de la Intendencia de Fiscalización y Gestión del Riesgo Aduanero (IFGRA), este equipo cuenta con una vasta experiencia es fiscalización y el manejo de herramientas de software.

A continuación mostraremos los criterios planteados para la elección de la mejor alternativa:

CRITERIO 1: SEGURIDAD Y CONFIABILIDAD DE LA HERRAMIENTA

Este criterio es uno de los más importantes, el equipo de trabajo consideró que la seguridad de la información es un punto crucial, ya que debe ser muy necesario que el modelo seleccionado no deba ser predecible, ya que las diferentes empresas importadores, podrían aprovecharse de esta falencia.

CRITERIO 2: COSTOS DE DESARROLLO DEL PROYECTO

Este criterio también es importante, ya que el costo de la implantación del sistema no puede llegar a ser tan alto, cada solución debe presentar sus costos ajustados razonablemente.

CRITERIO 3: TIEMPO DE DESARROLLO

El tiempo de desarrollo y de puesta en producción de cada solución debe ser lo más corto posible, ya que este sistema debe resolver el problema de las importaciones lo más pronto posible.

CRITERIO 4: HERRAMIENTA TECNOLÓGICA

El equipo de evaluación considero que se debe tener en cuenta, que la empresa debe contar con herramientas tecnológicas modernas e

innovadoras, para mantenerse actualizada referente a otras empresa del mismo rubro.

CRITERIO 5: EXPERIENCIA CON LA HERRAMIENTA EN OTROS PROYECTOS

Las personas del equipo consideraron que debe tratarse de una herramienta con la cual se cuente con experiencia de aplicación exitosa en otras empresas del mismo rubro.

CRITERIO 6: GRADO DE CONOCIMIENTO ACTUAL EN LA HERRAMIENTA

Se considero relevante considerar el grado de conocimiento con el que cuenta el equipo asociado a la herramienta tecnológica a utilizar, ya que sería mejor que se implantara alguna herramienta ya conocida, y que no requiera una curva de aprendizaje muy alta.

Se estableció un rango del 1 al 5, para poder asignar puntajes a cada criterio y así poder tomar una decisión, siendo 1 el puntaje más bajo y de significado muy malo y 5 el puntaje más alto y de significado muy bueno. También se realizó una ponderación de cada criterio, según cuál sea consideramos de mayor a menor importancia.

PUNTAJE	SIGNIFICADO
1	Muy Bajo
2	Bajo
3	Intermedio
4	Bueno
5	Muy Bueno

Cuadro 01: Puntajes Asociados

SELECCIÓN DE LA ALTERNATIVA

A continuación mostraremos el cuadro comparativo de cada alternativa de solución así como cada ponderación de los criterios, también se mostrará la relación de pesos que considero en total las personas de equipo designado para la selección (IFGRA). Finalmente se escogerá aquella solución que obtenga el mayor puntaje.

Nro	CRITERIO	PESO	PUNT A	PUNT. B	PUNT. POND. A	PUNT. POND. B
1	Seguridad y confiabilidad de la herramienta	30%	5	3	1.5	0.9
2	Costos de desarrollo del proyecto	10%	2	3	0.2	0.3
3	Tiempo de desarrollo	15%	3	3	0.45	0.45
4	Herramienta tecnológica	15%	5	2	0.75	0.3
5	Experiencia con la herramienta en otros proyectos.	20%	5	2	1	0.4
6	Grado de conocimiento actual en la herramienta	10%	1	3	0.1	0.3
	TOTAL	100%			4.0	2.65

Cuadro 02: Resultado de la selección de alternativas.

Como se puede apreciar en el Cuadro 02, la alternativa elegida es la implementación del sistema de selección utilizando los modelos de minería, a continuación explicaremos, el porqué de los puntajes mencionados en el cuadro:

1. El grupo de expertos considero que la solución de minería era mucho más segura, ya que es casi imposible de predecir el resultado de la misma, en cambio en el modelo estadístico de discriminante si se llega a conocer los coeficientes de la función, puede determinarse de cierta forma el resultado de los índices.

2. El costo del desarrollo para ambos proyectos resulto casi el mismo, ya que se contaría con un presupuesto BID, y con consultores externos que serán contratados para cualquiera de los 2 proyectos, pero se consideró que para los modelos de minería se incurriría en un costo de capacitación de las diferentes áreas involucradas, por esta razón se le dio un puntaje mayor a la solución con el modelo estadístico, que es una herramienta ya conocida por muchos de los integrantes del equipo.
3. El tiempo de desarrollo de ambas soluciones, fue considerado el mismo, ya que finalmente a pesar de ser herramientas diferentes ambas necesitan de un análisis masivo de información y de un entrenamiento para obtener los coeficientes para el caso del discriminante, o el entrenamiento de la red neuronal para el caso del modelo de minería de datos.
4. Este punto es importante y fue evaluado debido a que la empresa, quiere utilizar tecnología de punta para todos sus proyectos, el esquema de manejo de modelos de minería refleja una tecnología relativamente moderna a diferencia del la implementación de modelos estadísticos que es algo ya muy usado estos últimos años.
5. Es de conocimiento público que muchos otros países como España y Chile está apostando por la implementación de modelos de minería de datos para la fiscalización aduanera de sus respectivos países, es por esto que se otorga mayor puntaje a la solución con minería de datos, ya que se tienen experiencias de otras Aduanas en el mundo.
6. Se observo mediante una encuesta interna, que muchas personas no está familiarizadas con temas como minería de datos o redes neuronales en la empresa, por lo que se encontró total desconocimiento, pero si mucho interés por aprender. En cambio haciendo la misma encuesta respecto a los modelos estadísticos y a la herramienta SPSS, las

personas respondieron satisfactoriamente, lo cual demostraba que se contaba con un conocimiento y adquirido en lo que respecta a los modelos estadísticos. Es por esto que colocamos un mayor puntaje al proyecto que utilizaría herramientas estadísticas ya conocidas.

CAPITULO IV

IMPLEMENTACIÓN DEL SISTEMA

4.1. PLAN DE ACCION PARA EL DESARROLLO DE LA SOLUCION PLANTEADA

Para el desarrollo del Sistema de Selección de Canal de Control del NPDA, se contrato un consultor externo BID, con el cual se firmaron 2 contratos para la implementación del sistema:

- En la primera fase se realizaron las tareas de migración y adecuaciones del sistema para la nueva plataforma tecnológica del proyecto NPDA, además de diseñarse todo el proceso del Sistema Administrador de Modelos de Minería.

- En la segunda fase se realizaron las estabilización de la primera fase, la implementación de los modelos de minería, y el modulo SAC (Sistema Administrador de catálogos), encargado de la administración de parámetros de los modelos.

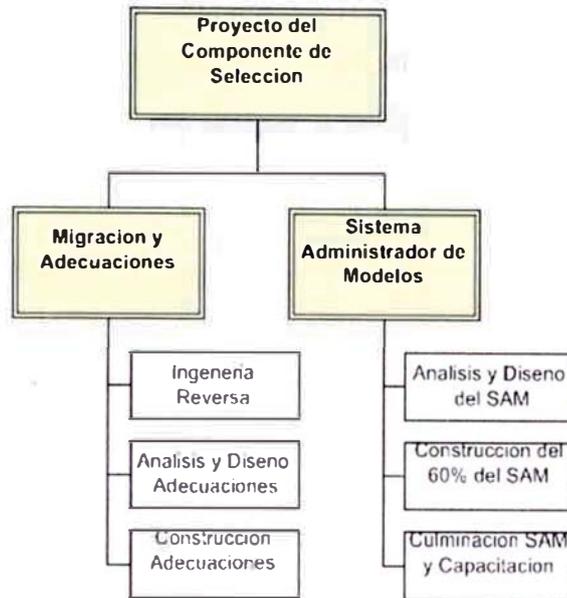


Grafico 03: EDT del Proyecto (Primer Contrato) - Fuente SUNAT

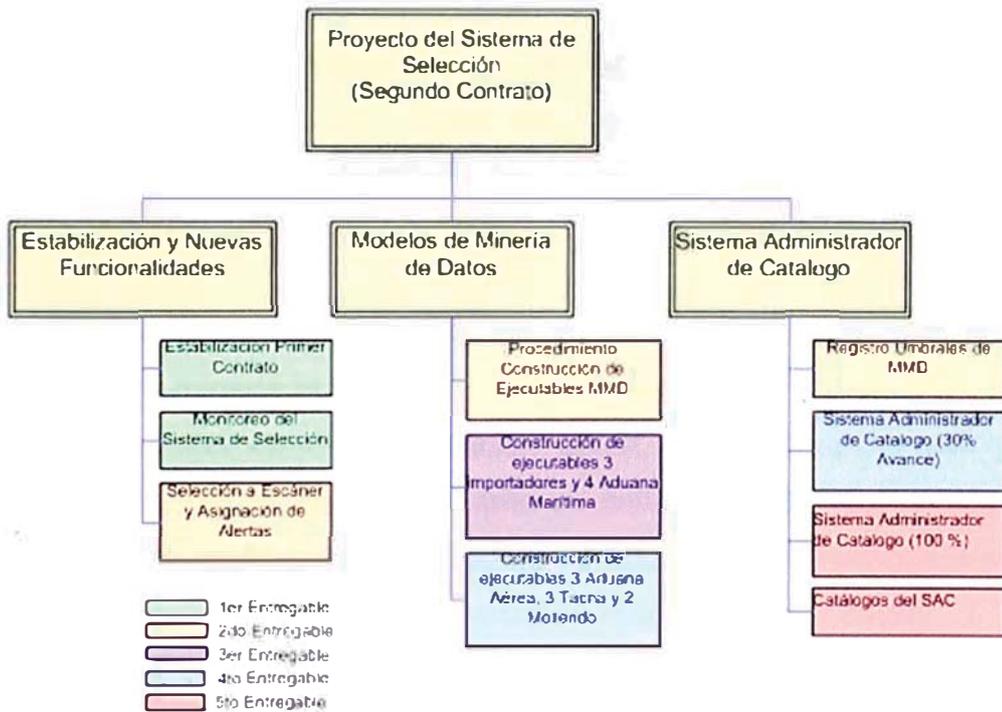


Grafico 04: EDT del Proyecto (Segundo Contrato) - Fuente SUNAT

Para el éxito del proyecto se contaron con consultores especializados en herramientas java y herramientas SAS (Minería de Datos). A continuación una breve explicación de las tareas que formaron parte del proyecto de implementación del sistema.

- **Migración y Adecuaciones:** Esta Actividad abarca principalmente las adecuaciones previas necesarias para la implementación del sistema de selección, ya que necesita ser invocado como un servicio, para que pueda ser independiente y así puedan invocarlo los diferentes actores o sistemas que forman parte del rubro principal del negocio aduanero.
- **Sistema Administrador de Modelos:** La implementación de este módulo abarca toda la infraestructura necesaria para la invocación de los modelos de minería definidos por IFGRA, cada modelo debe poder ser invocado independientemente, según la evaluación inicial de la DUA.
- **Estabilización y nuevas funcionalidades:** Luego de la primera fase del proyecto, será necesario contar con un periodo de estabilización y mejoras al sistema planteado, para poder proceder a realizar las correcciones necesarias para errores que puedan surgir en las pruebas.
- **Modelos de Minería:** Esta etapa corresponde a la implantación de los modelos de minería que creará la IFGRA, para esto es necesario realizar una integración entre la aplicación JAVA, y los modelos que serán instalados como funciones mediante un data Stage de Informix, de esta manera será posible la invocación de cada uno de los modelos, y lo que es más importante, se podrán invocar de manera independiente.
- **Sistema Administrador de Catálogos:** Finalmente será necesario realizar la creación de un modulo para el mantenimiento de los modelos, y tablas paramétricas del sistema, así también este módulo contara con una herramienta de monitoreo, necesaria para poder revisar el

comportamiento del sistema así como la selectividad con la que cuenta cada modelo implementado.

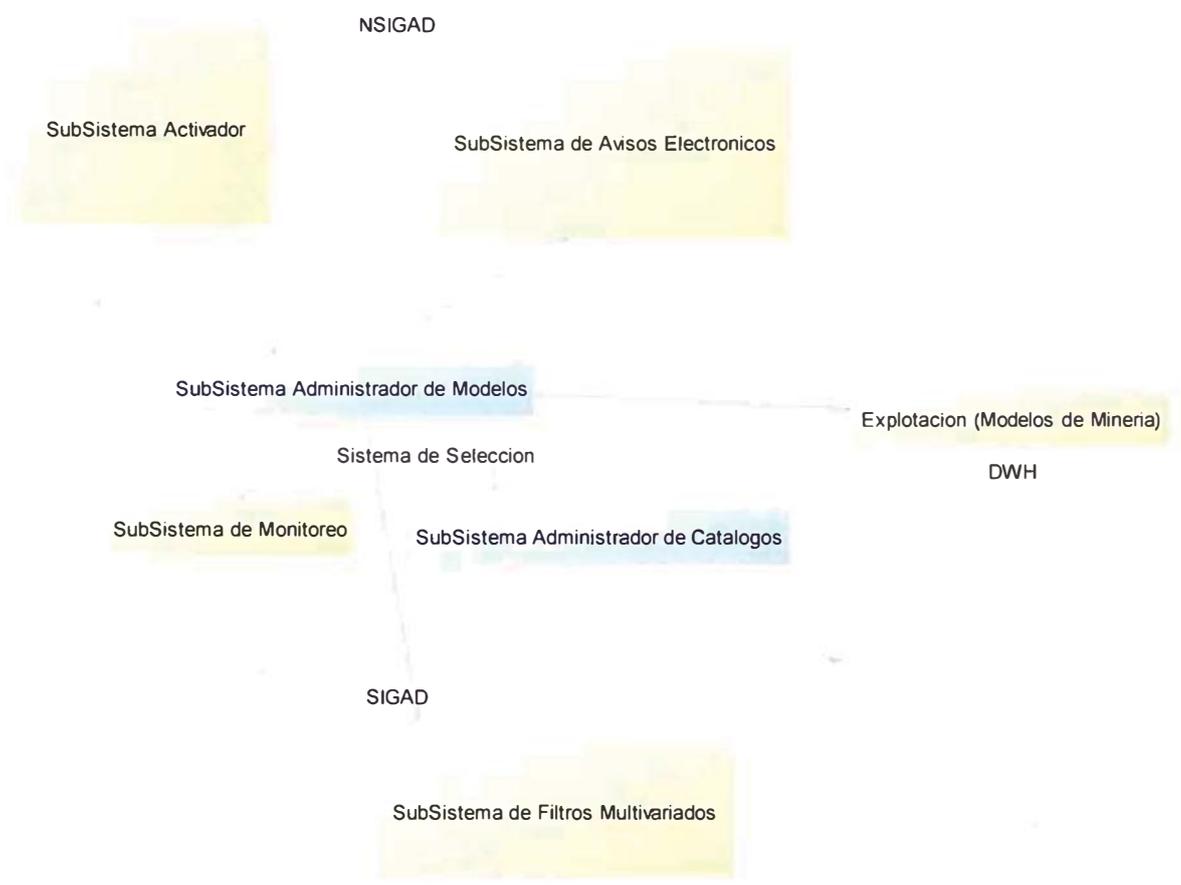
4.2. ANALISIS Y DISEÑO DE LOS MÓDULOS

El sistema desarrollado cuenta básicamente con 2 módulos implementados:

- El Sistema Administrador de Modelos (SAM), es una aplicación que permite la ejecución de las diferentes herramientas de selección de canales, aplica reglas de jerarquía y asigna el canal final, entre esas reglas de selección están ubicados los modelos de Minería de Datos.

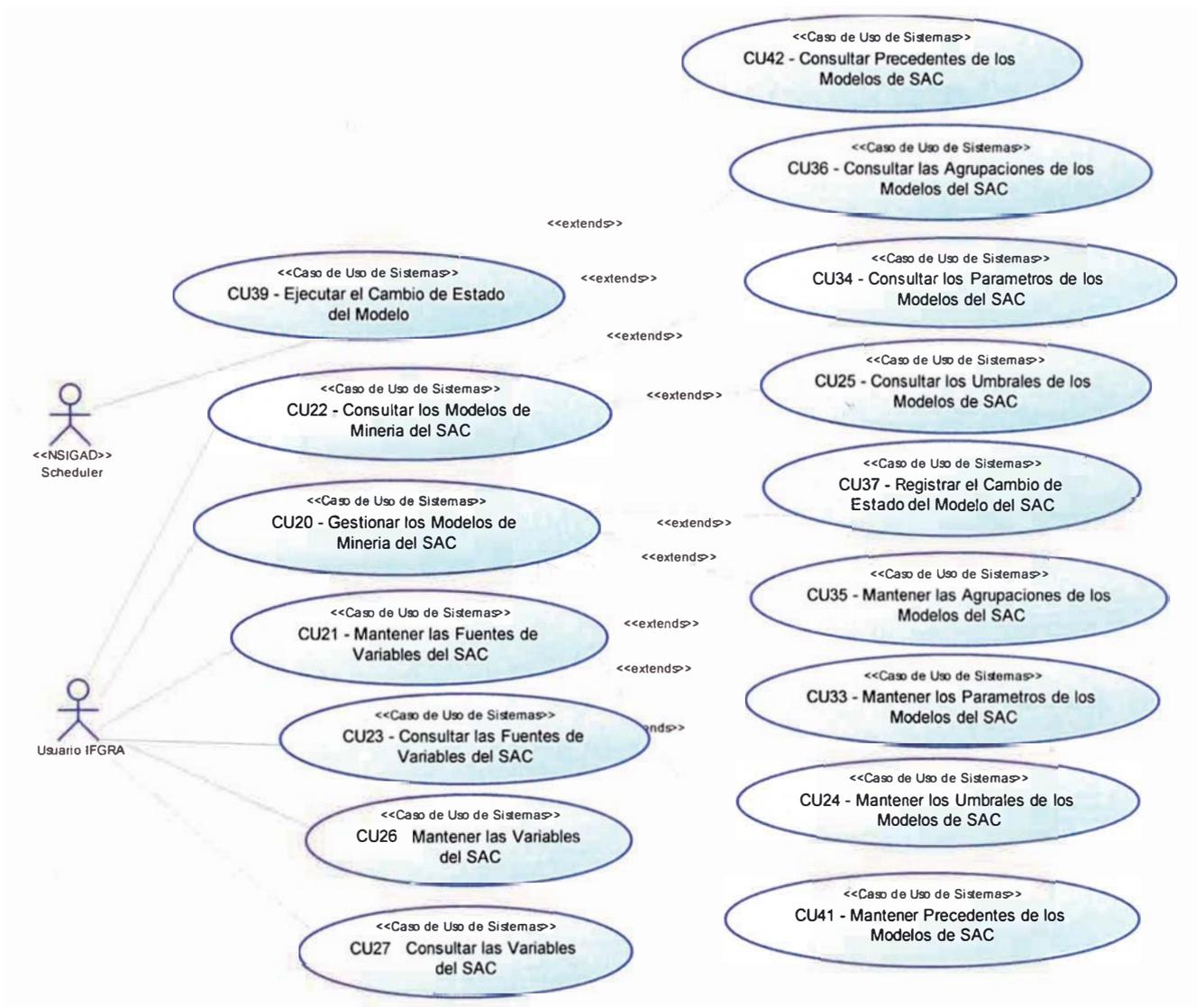
- El Sistema Administrador de Catálogos (SAC), corresponde al mantenimiento y consulta de las tablas de parámetros que permiten soportar la ejecución del SAM.

La arquitectura del sistema la podemos apreciar más claramente en el diagrama de Paquetes indicado en el Grafico 05.

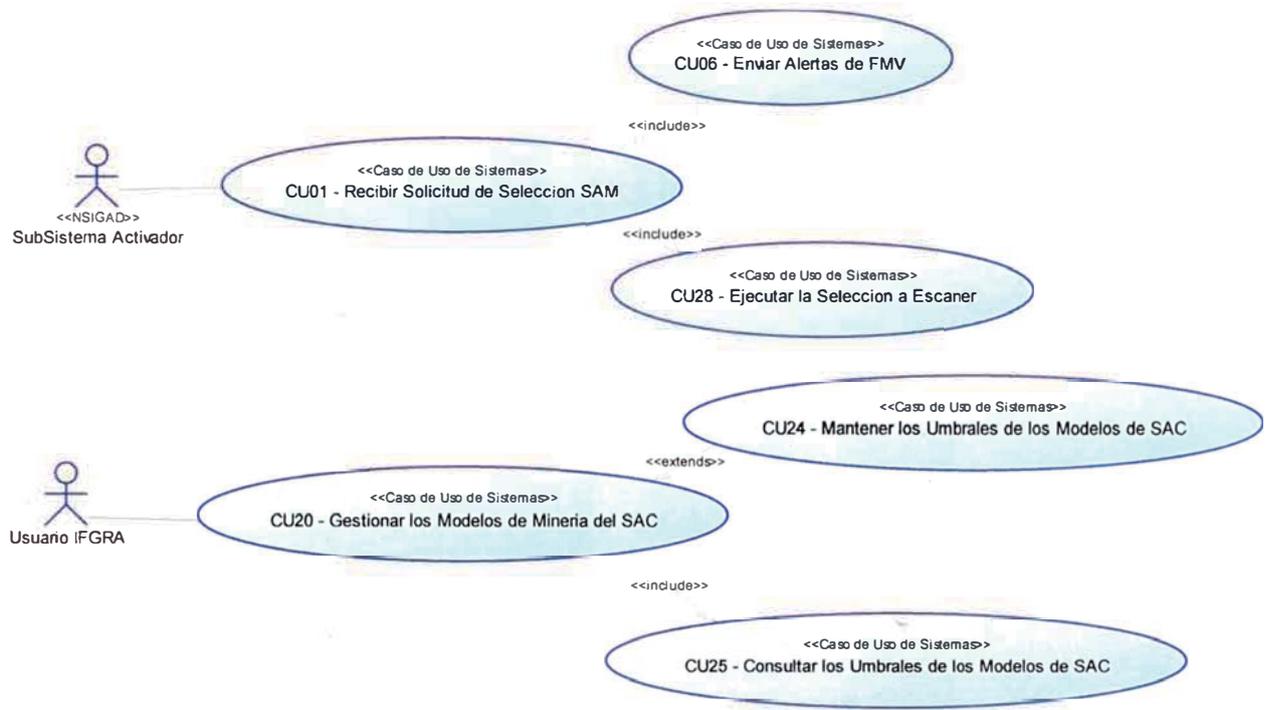


Grafica 05: Diagrama de Paquetes del Sistema de Selección (Fuente SUNAT 2010)

Las funcionalidades principales de ambos módulos se muestran a continuación en el presente diagrama de Casos de Uso, en este diagrama puede apreciarse el CU22 que implica la consulta a los Modelos de Minería de datos, ver Grafico 06 y Grafico 07.



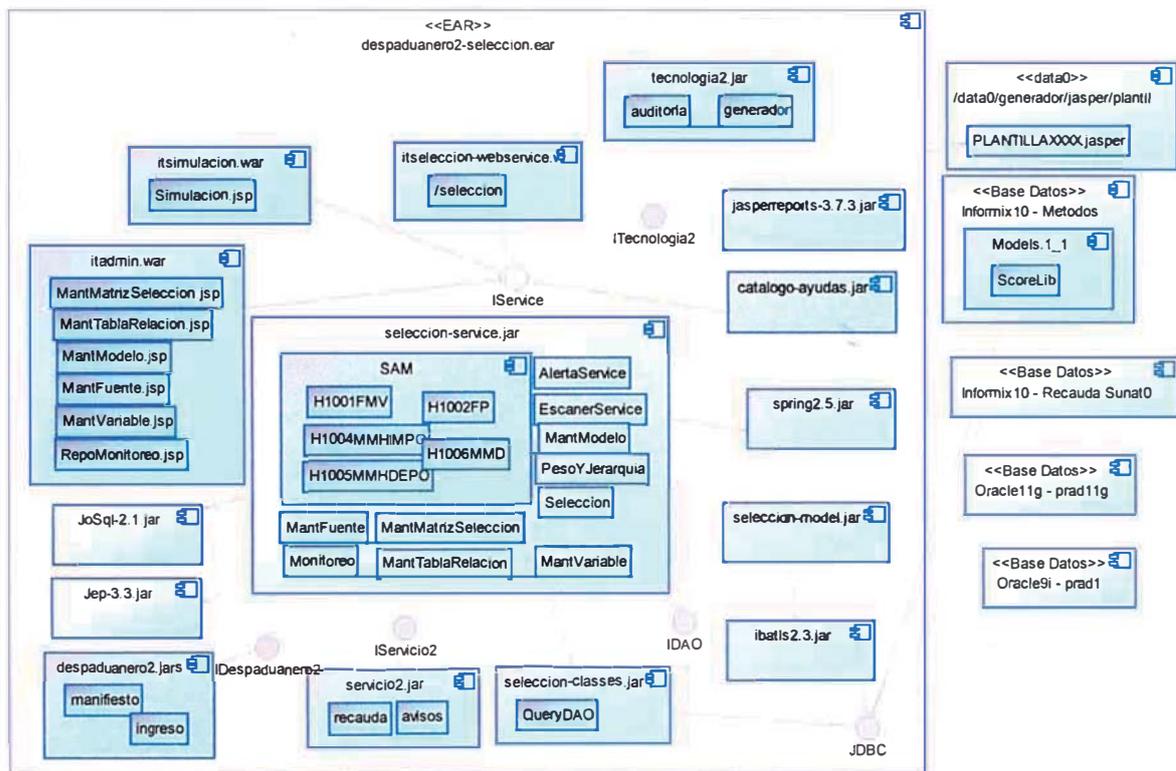
Grafica 06: Diagrama de Casos de Uso SAC (Fuente SUNAT 2010)



Grafica 07: Diagrama de Casos de Uso SAM - SAC (Fuente SUNAT 2010)

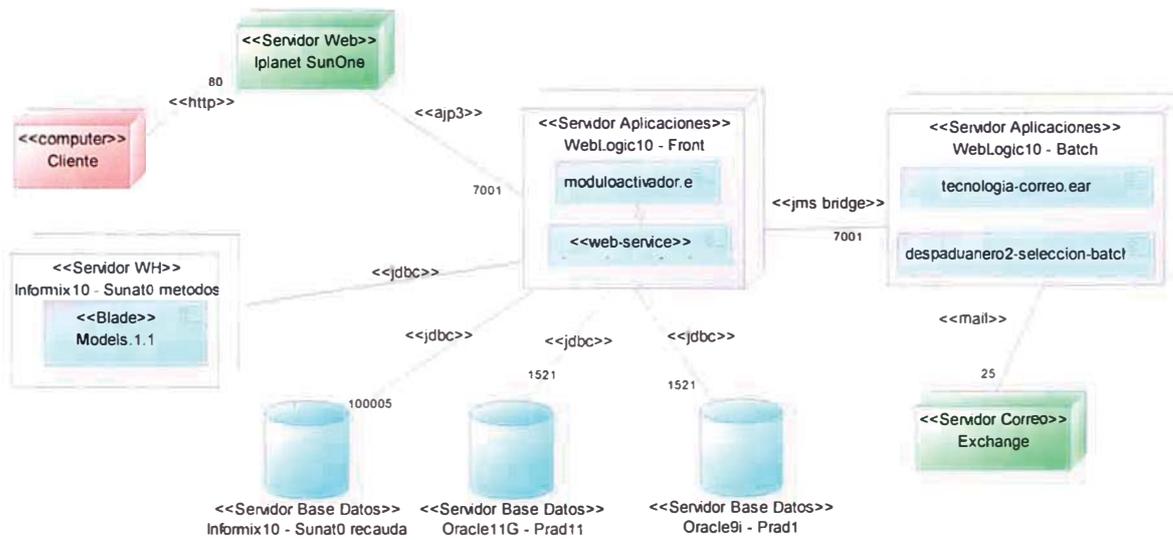
A continuación se presenta el diagrama de entidades, donde podemos identificar la Entidad Modelo, donde se almacena la identificación de los 15 modelos de minería creados por la IFGRA, ver Grafica 08.

La aplicación fue instalada en un servidor de aplicaciones Web, mediante un archivo .EAR. A continuación se muestra el diagrama de componentes del sistema, donde podemos apreciar los diferentes módulos web y librerías utilizadas en todo el proyecto. También podemos observar la conexión con los repositorios externos en Informix y Oracle. Ver Grafico 09.



Grafica 09: Diagrama de Componentes del Sistema de Selección (Fuente SUNAT 2010).

A continuación se muestra el diagrama de despliegue donde podemos identificar los servidores web, servidores de aplicaciones y servidores de base de datos utilizados en la implantación del sistema.



Grafica 10: Diagrama de Despliegue del Sistema de Selección (Fuente SUNAT 2010)

4.3. MANUAL DE USUARIO DEL SISTEMA

El Sistema Administrador de Modelos (SAM) es una herramienta de usuario que permite simular, mantener y consultar las herramientas de selección para la asignación de canal y otras acciones de control de la Aduana Peruana.

El SAM, permite ejecutar las diferentes herramientas por momentos, los cuales son configurados en la matriz de selección.

SIMULACION DE SELECCION DE DECLARACION

1. Flujo de Configuración de Simulación de Declaración

Para iniciar el proceso de simulación se debe revisar el siguiente diagrama de flujo (ver Grafico 11), que indica la verificación de la configuración de los perfiles del Sistema de Filtros, los modelos de minería en el Sistema Administrador de Catálogos.

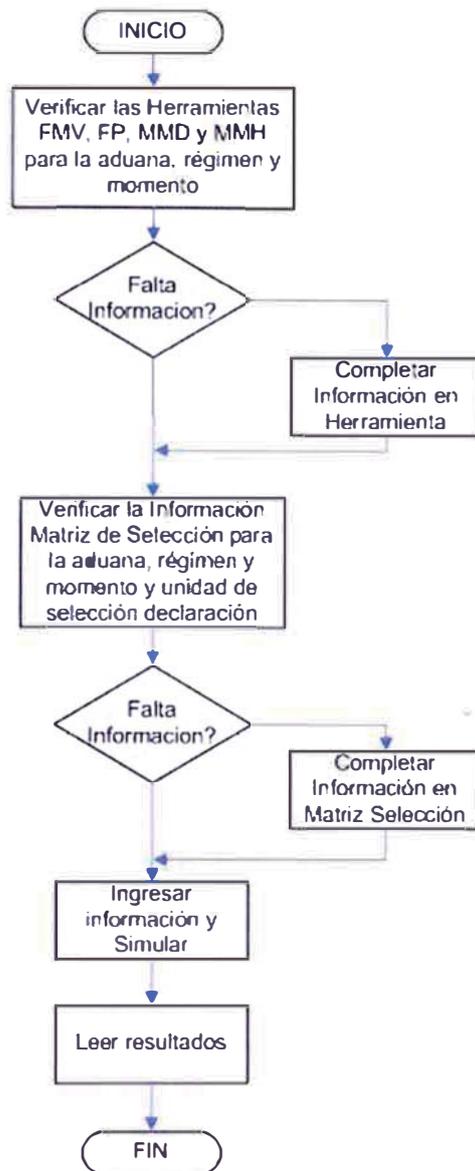
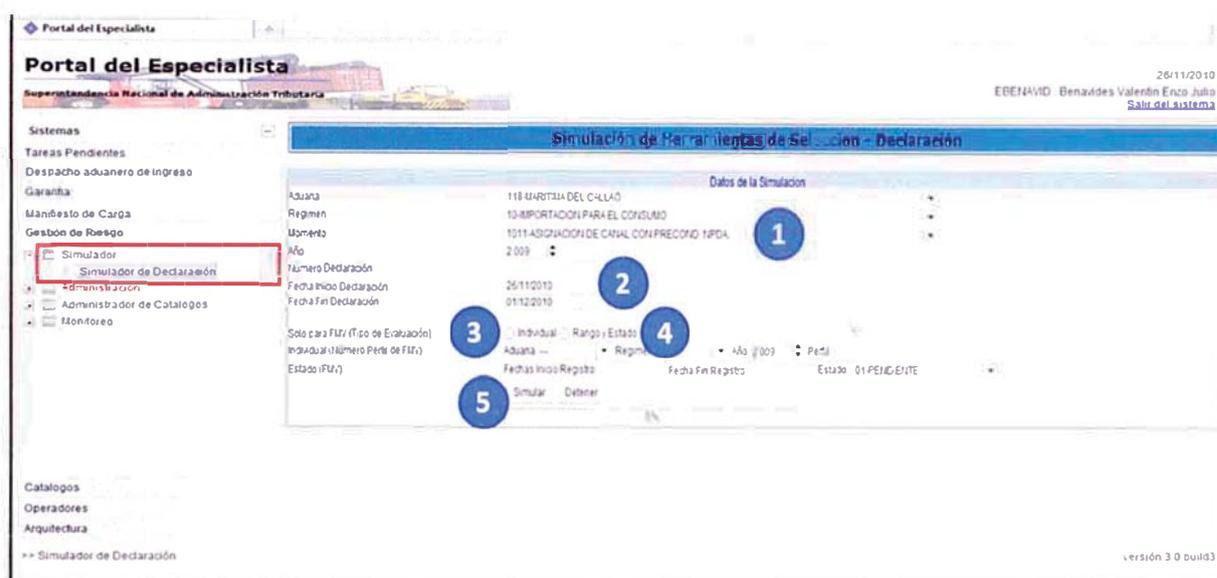


GRAFICO 11: Diagrama de Flujo del proceso de Simulación (Fuente SUNAT 2010)

2. Ingresando a las opciones del Simulador

Las opciones del Simulador están disponibles ingresando a través del Portal del Especialista \ Gestión de Riesgo \ Simulador \ Simulador de Declaración, el perfil necesario para acceder a estas opciones es “ADUANA-ESPECIALISTA-FISCA”



Grafica 12: Interface para simular la selección de canal (Fuente SUNAT 2010)

Los puntos numerados en la interface de “Simulación” se indican a continuación:

1. Opciones de la Declaración: Representan los datos de la Aduana de la Declaración, Régimen de la Declaración, Momento de Evaluación configurado en la herramienta, Año de la Declaración y Numero de la Declaración.
2. Opción Rango de Fechas: Fechas de Numeración de las Declaraciones.
3. Opción Individual: Indica que solo se evaluara un FMV, ingresado en la opción Numero de Perfil o FMV, por defecto el sistema evalúa todos los filtros.

4. Opción Rango y Estado: Indica que se evaluará por fechas de registro de los filtros (FECH_INGSI) y su estado, por defecto el sistema evalúa todos los filtros.
5. Opción Simular: Inicia el proceso de simulación.

3. Casos de Simulación

Al ingresar a la opción se pueden presentar los siguientes casos, que están asociados a las opciones de punto 2.2 del documento:

Caso 1:

Cuando se indica solo los valores de la opción 1, el sistema evalúa todas las herramientas configuradas en la matriz de selección para ese momento y esa unidad de selección, en el caso de los FMV evalúa todos los perfiles sin importar el estado, pero solo para la Declaración indicada.

Portal del Especialista

Portal del Especialista
Superintendencia Nacional de Administración Tributaria

29/11/2010
EBEIVAVID Benaides Valentin Enzo Julio
[Salir del sistema](#)

Simulación de Herramientas de Selección - Declaración

Datos de la Simulación

Aduana: 118-114R/TI114 DEL CALLAO
 Regimen: 10-IMPORTACION PARA EL COMERCIO
 Momento: 1011-ASIGNACION DE CANTIDAD CON PRECIO IDENTIFICADO
 Año: 2009
 Número Declaración: 2345

Solo para FMV (Tipo de Evaluación)
 Individual (Número Perfil de FMV)
 Estado (FLMV)

Individual Rango y Estado
 Aduana: --- Regimen: --- Año: 2009 Perfil: ---
 Fechas Inicio Registro: Fecha Fin Registro: Estado: 01-PENDIENTE

Simular Detener

0%

versión 3.0 build3

Grafica 13: Ejemplo de ingreso de datos Caso 1 (Fuente SUNAT 2010)

Caso 2:

Cuando se indica la opción 1 sin numero de Declaración y con la opción 2, el sistema evalúa todas las herramientas configuradas en la matriz de selección para ese momento y esa unidad de selección, en el caso de los FMV evalúa todos los perfiles sin importar el estado, pero para las declaraciones que se encuentran en el rango de fechas de numeración indicado.

Portal del Especialista

Superintendencia Nacional de Administración Tributaria

29/11/2010
EBERAVID Benavides Valentin Enzo Julio
[Salir del sistema](#)

Simulación de Herramientas de Selección - Declaración

Datos de la Simulación

Aduana	118-MARITIMA DEL CALLAO
Regimen	10-IMPORTACION PARA EL CONSUMO
Momento	1011-ASIGNACION DE CAPITAL CON PRECEDENCIA
Año	2009
Número Declaración	03/11/2010
Fecha Inicio Declaración	05/11/2010

Solo para FMV (Tipo de Evaluación)
Individual (Número Perfil de FMV)
Estado (Filtro)

Individual Rango y Estado

Aduana --- Regimen --- Año 2009 Perfil

Fecha Inicio Registro Fecha Fin Registro Estado 01.PENDIENTE

Simular Detener

versión 3.0 build3

Gráfica 14: Ejemplo de ingreso de datos Caso 2 (Fuente SUNAT 2010)

Caso 3:

Indiferente de las opciones 1 y 2, las opciones 3 y 4 están enfocadas solo a los FMV si se escoge la opción 3 el sistema le pedirá que ingrese un FMV para evaluar.

Portal del Especialista

Superintendencia Nacional de Administración Tributaria

29/11/2010
EBENAVID Benavides Valentin Enzo Julio
[Salir del sistema](#)

Simulación de Herramientas de Selección - Declaración

Datos de la Simulación

Aduana	118-UI-RITMIA DEL CALLAO
Regimen	10-IMPORTACION PARA EL CONSUMO
Momento	1011-ASIGNACION DE CANAL COMPRECONO NPOA
Año	2009
Número Declaración	
Fecha Inicio Declaración	03/11/2010
Fecha Fin Declaración	05/11/2010
Solo para FMV (Tipo de Evaluación)	Individual Rango y Estado
Individual (Número Perfil de FMV)	Aduana 118 Regimen 10 Año 2009 Perfil 000002
Estado (FMV)	Fecha Inicio Registro Fecha Fin Registro Estado 01-PENDIENTE

Simular Detener

0%

versión 3.0 build3

Gráfica 15: Ejemplo de ingreso de datos Caso 3 (Fuente SUNAT 2010)

Caso 4:

Indiferente de las opciones 1 y 2, las opciones 3 y 4 están enfocadas solo a los FMV si se escoge la opción 4 el sistema le pedirá que ingrese un rango de FMV por fecha de registro (FECH_INGSI) y el estado del FMV.

Portal del Especialista

Superintendencia Nacional de Administración Tributaria

29/11/2010
EBE114VD | Benajides Valentin Enzo Julio
[Salir del sistema](#)

Simulación de Herramientas de Selección - Declaración

Datos de la Simulación

Aduana	118-HERRITUA DEL CALLAO
Regimen	10-IMPORTACION PARA EL CONSUMO
Momento	1011-ASIGNACION DE CAPITAL CON PRECOND. NPDA
Año	2009
Número Declaración	
Fecha Inicio Declaración	03/11/2010
Fecha Fin Declaración	05/11/2010

Solo para FMV (Tipo de Evaluación)

Individual (Número Perfil de FMV)

Estado (FMV)

Individual Rango y Estado

Aduana — Regimen — Año 2009 Perfil

Fechas Inicio Registro: 10/11/2010 Fecha Fin Registro: 19/11/2010 Estado: 03-PRODUCCION

Simular Detener

0%

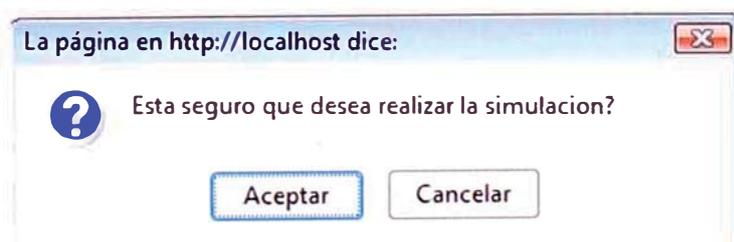
versión 3.0 build3

Grafica 16: Ejemplo de ingreso de datos Caso 4 (Fuente SUNAT 2010)

El sistema siempre evalúa todas las herramientas configuradas en la Matriz de Selección a excepción de la configuración de FMV indicada por el simulador.

4. Simulando la Selección

Una vez seleccionada la opción simular el sistema muestra lo siguiente:



CAPITULO V

EVALUACION DE RESULTADOS

La empresa tiene indicadores ya definidos para poder medir resultados del nuevo sistema a lo largo del año, en general se nota una mejora considerable en los indicadores del año 2010.

5.1. FICHA INDICADOR N° 35: NIVEL DE INCIDENCIA SUSTANCIAL EN CANAL ROJO, EN IMPORTACIONES

CONCEPTOS	DEFINICIÓN
Producto N° 20	Nivel de riesgo asignado
Indicador N° 35	Nivel de incidencia sustancial en canal rojo, en importaciones
Fundamento	Mide la eficacia de la selección de declaraciones sujetas a reconocimiento físico por técnicas de gestión de riesgo
Forma de cálculo	<u>Cantidad de DUAs con Incidencia en reconocimiento físico (2)</u> Cantidad de DUAs seleccionadas por riesgo a reconocimiento (1) Consideraciones Generales: - Aplicable en aduanas principales: Marítima, Aérea, Tacna y Mollendo.

- Aplicable a DUAs excepcionales
- Se tienen en cuenta las declaraciones seleccionadas a canal rojo y las naranja que han sido sujetas a reconocimiento físico.
- Se entiende por una DUA con incidencia a la que tiene registrado una L/C asociado a ajustes de valor, dumping, mala declaración del origen y cambio de clasificación arancelaria con incidencia tributaria.
- Asimismo se consideran las declaraciones con orden de depósito, garantía global o específica como consecuencia de la duda razonable y con levante, no desvirtuadas totalmente.

No se consideran:

- Las DUAs seleccionadas por criterios normativos (*)
- Las DUAs seleccionadas por criterios aleatorios
- Las DUAs presentadas por los importadores frecuentes
- Las liquidaciones de cobranza anuladas.
- Las L/C emitidas producto de acciones de fiscalización posterior (registradas en la tabla L/C programa).
- Las DUAs anuladas sin cancelar, ni las declaraciones canceladas legajadas.
- Las DUAs que no cumplan los criterios de riesgo establecidos
- En el caso de la Aduana de Tacna no se consideran las DUAs que amparan vehículos.

(*) Incluye las declaraciones seleccionadas para verificar el kilometraje declarado en los vehículos usados, de acuerdo a lo prescrito en el Decreto Legislativo 843, modificado con Decreto Supremo 042-2006-MTC.

Consideraciones Específicas:

(1) Determinación de las DUAs seleccionadas por criterios de Riesgo o no normativos.

Para determinar las DUAS seleccionadas por criterios no normativos, se deberá tomar en cuenta los campos TIPO_AFORO (= F) y REGI_PROCE de la Tabla POLIZAI y cruzarla con la tabla CANALES a través del campo CODI_MOTI (REGI_PROCE=CODI_MOTI) cuando el CODI_REGI (=10), CODI_CANAL (= F) en la tabla CANALES, a fin de identificar el valor del campo GRUPO que determinará las siguientes modalidades de selección:

- a) Minería de Datos (Red Neuronal).- Grupo 7
- b) Modelo de MMH.- Grupo 2
- c) Sensible al fraude - Grupo 5
- d) Validación Automatizada de Precios – Grupo 8.
- e) Filtros de Múltiples Variables.- Grupo 6, en este caso además se deberá determinar el FMV que motivo la selección a través de la tabla DUA_TIPO_AFORO, con el número de DUA y cuando el campo CODI_REGI='10', TIPO_AFORO='F' y MOTI_AFORO='80' en la tabla precitada. El número de perfil se obtiene de los campos CODI_ADUAP, ANO_PRESEP y NUME_PERFI. Una vez obtenido el número de perfil, sólo se consideran las DUAs, si es que el perfil tiene los siguientes atributos en la tabla CAB_PERFIL_BASE_CONO: Campo TIPO_PERFI = 0 y el campo CODI_PERFI >= A000 y <= A260 o A520 o A630. Adicionalmente para el caso de la Aduana de Tacna (172) considerar los números de perfil 100004,

100006, 100009 al 100013.

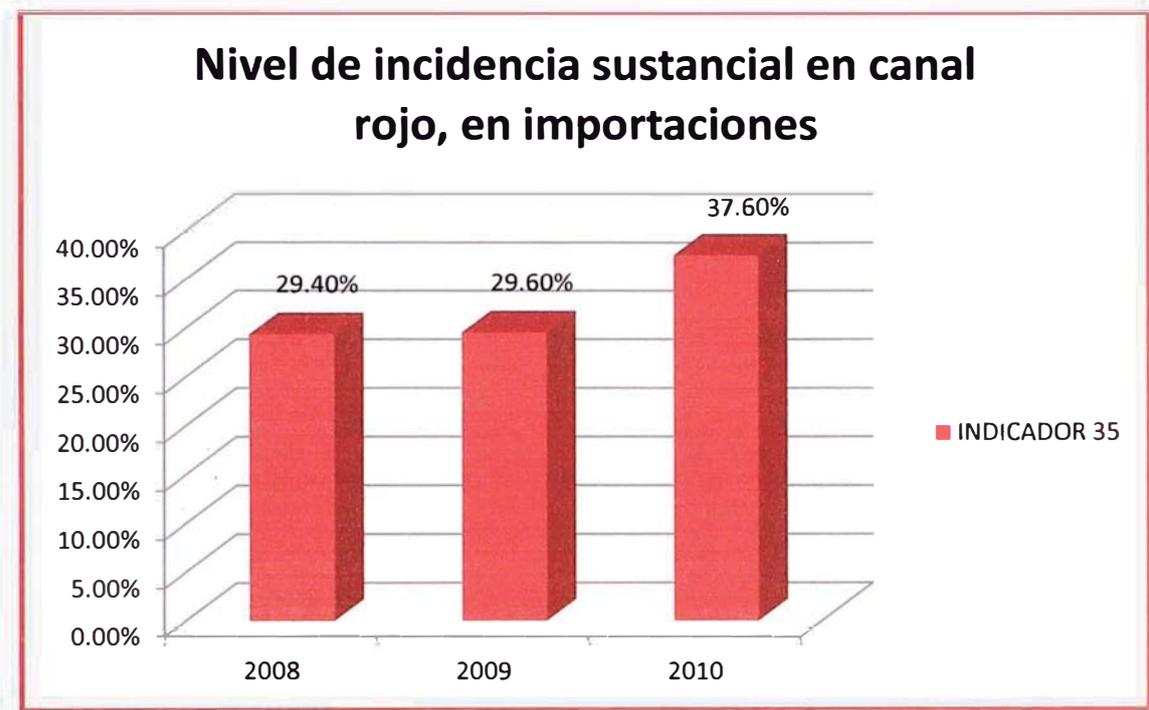
(2) Determinación de las DUAs con incidencia

Se considera DUAS con incidencia a aquellas declaraciones no seleccionadas por criterios normativos y aleatorios, cuando producto del despacho se produzca o genere una L/C relacionada con el incremento en la recaudación tributaria aduanera en los temas de Valor, Dumping, Clasificación Arancelaria y Origen:

- a) Valor.- Se consideran las DUAS con L/C tipo 0029, 0030, 0032 y 0033 (Duda razonable). Adicionalmente se debe considerar las DUAS con L/C tipo 0026, cuando en la tabla TIPOINCI_DECL registre el campo CODI_REGI='10' y en el campo TIPO_INCID los códigos: 02, 05, 16, 19, 42 y 43.
- b) Dumping.-Se consideran las DUAS con L/C Tipo 0026 con código del tributo = '0050' (campo RDCODRUB de la tabla DETALIQ).
- c) Clasificación Arancelaria.- Se han considerado las DUAS con L/C tipo 0003 y 0027 con el campo RLCODCO1 de la tabla LIQUIDA (o los campos RLCODCO11, RLCODCO12, RLCODCO13, RLCODCO21 de la tabla LIQINCMULTA) = 125 o 194 (ley anterior) y D7 o D8 (Ley actual), debiendo considerarse las liquidaciones emitidas en dólares, y en soles sólo cuando la declaración tiene asociado otra L/C tipo 0026 en soles, en donde se acota el ISC. .
- d) Declaración de Origen: Se consideran las DUAS con L/C tipo 0003, 0027, con el campo RLCODCO1 de la tabla LIQUIDA (o los campos RLCODCO11, RLCODCO12, RLCODCO13, RLCODCO21 de la tabla LIQINCMULTA)

	<p>= C01, B31, C15 (Ley actual) ó D3, D5, E1, 188 (Ley anterior); así como las DUAS con L/C tipo 0010, 0013, 0026 con rldetal, rldetal2, rldetal3, rldetal4 y rldetal5 contenga ORIGEN.</p> <p>Finalmente se determinará el cociente entre las DUAs (2) con incidencia seleccionadas por riesgo y el total de DUAs (1) seleccionadas por dichos criterios.</p>			
Unidad de medida	Porcentaje			
Fuente de Información	SIGAD Archivo Plantilla			
Datos históricos	2008: 29.4% 2009: 29.6% 2010: 37.6% (A octubre del 2010)			
Meta para el año 2011	Al 1er. Trimestre	Al 2do Trimestre	Al 3er. Trimestre	Al 4to. Trimestre
	31.5%	31.8%	32.4%	33.7%
Frecuencia de Reporte	Trimestral			
Disponibilidad de la información	A los 10 días hábiles de concluido el trimestre			
Responsable	Intendencia de Fiscalización y Gestión de Recaudación Aduanera			

Cuadro 03: Indicador 35 (Fuente SUNAT)



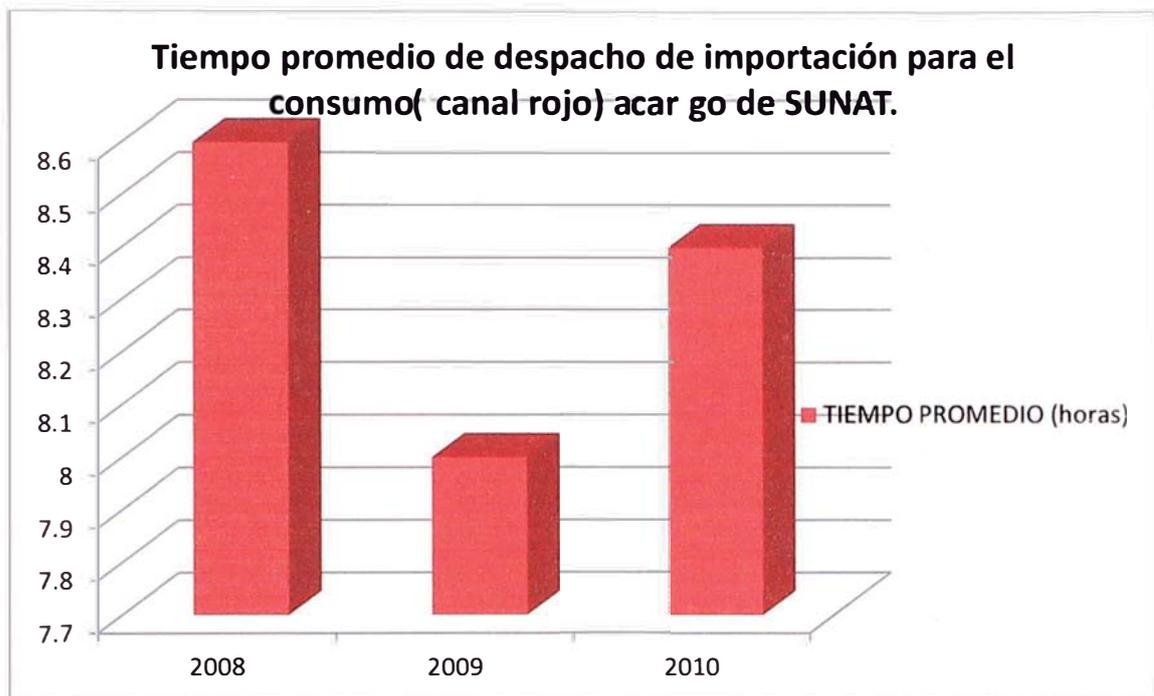
Grafica 18: Histograma del Indicador 35 (Elaboración propia en base a los datos del indicador SUNAT).

5.2. FICHA INDICADOR N° 58: TIEMPO PROMEDIO DE DESPACHO DE IMPORTACIÓN PARA EL CONSUMO (CANAL ROJO) A CARGO DE SUNAT.

CONCEPTOS	DEFINICIÓN
Producto N° 34	Declaración de mercancía con levante
Indicador N° 58	Tiempo promedio de despacho de importación para el consumo (canal rojo) a cargo de SUNAT.
Fundamento	Se busca medir el tiempo de despacho para el régimen de importación para el consumo (despacho excepcional) a cargo de la SUNAT, desde el momento en que se registra la recepción de la Solicitud Electrónica de Reconocimiento Físico (SERF) o GED, una vez recibido los documentos sustentatorios, según corresponda, hasta que se registra la diligencia autorizando su levante.
Forma de cálculo	<p>Tiempo Promedio de Despacho a cargo de la SUNAT = (A – B)</p> <p>Donde:</p> <p><u>Tratándose de la IAMC</u></p> <p>A = Fecha y hora de registro de la diligencia en las declaraciones sujetas a reconocimiento físico (canal rojo) sin incidencia.</p> <p>B = 8:30 am del día siguiente de transmitida la SERF (confirmación electrónica) o emitida la GED, en las declaraciones sujetas a reconocimiento físico (canal rojo).</p> <p><u>Tratándose del resto de aduanas</u></p> <p>A tiene el siguiente valor:</p>

	<p>Fecha de registro de la diligencia en las DUAS sujetas a canal rojo sin incidencia.</p> <p>B tiene el siguiente valor:</p> <p>Fecha de emisión de la SERF (confirmación electrónica o GED) en las DUAS sujetas a canal rojo.</p> <p>Solo se considera las DUAS sin incidencia.</p> <p>Se excluye las DUAS anuladas, legajadas, urgentes, anticipadas, notificadas y aquellas que se encuentren fuera del horario de 8:00 a 17:59 horas de cada día.</p> <p>El tiempo será medido a través del estadístico de la mediana, debido a la distribución sesgada de los tiempos.</p>			
Unidad de medida	Horas			
Fuente de Información	SIGED			
Datos históricos	2008: 8.6 2009: 8.0 2010: 8.4(*)			
Meta para el año 2011	Al 1er. Trimestre	Al 2do Trimestre	Al 3er. Trimestre	Al 4to. Trimestre
	9	9	9	9
Frecuencia de Reporte	Trimestral			
Disponibilidad de la información	10 días hábiles de culminado el trimestre de evaluación (remisión automática)			
Responsable	Intendencia Nacional de Técnica Aduanera			

Cuadro 04: Indicador 58 (Fuente SUNAT)



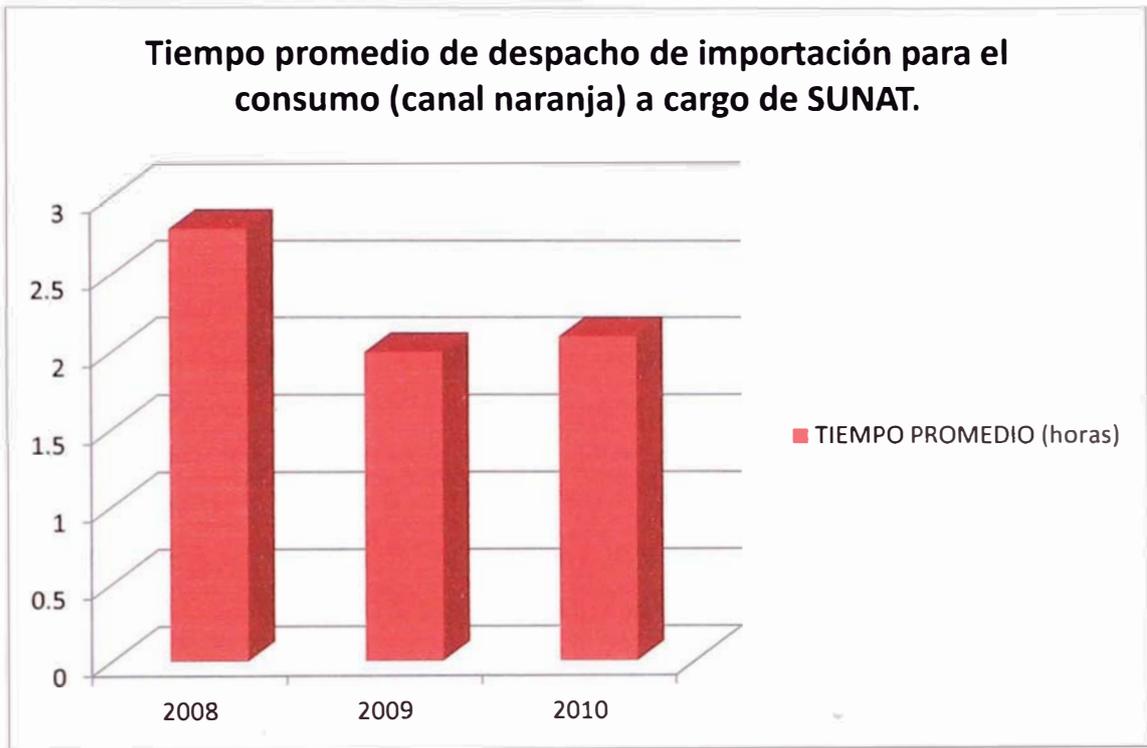
Grafica 19: Histograma del Indicador 58 (Elaboración propia en base a los datos del indicador SUNAT).

5.3. FICHA INDICADOR N° 59: TIEMPO PROMEDIO DE DESPACHO DE IMPORTACIÓN PARA EL CONSUMO (CANAL NARANJA) A CARGO DE SUNAT.

CONCEPTOS	DEFINICIÓN
Producto N° 34	Declaración de mercancía con levante
Indicador N° 59	Tiempo promedio de despacho de importación para el consumo (canal naranja) a cargo de SUNAT.
Fundamento	Se busca medir el tiempo de despacho para el régimen de Importación para el Consumo (despacho excepcional) a cargo de la SUNAT, desde el momento en que se registra la recepción de la Guía de Entrega de Documentos (GED), una vez recibidos los documentos sustentatorios, según corresponda , hasta que se registra la diligencia en el SIGAD.
Forma de cálculo	<p style="text-align: center;">$\text{Tiempo Promedio de Despacho} = (A - B)$</p> <p>Donde:</p> <p>A tiene el siguiente valor: Fecha y hora de registro de la diligencia en el SIGAD de las declaraciones sujetas a revisión documentaria (canal naranja) sin incidencia.</p> <p>B tiene el siguiente valor: Fecha y hora de registro de recepción de la Guía de Entrega de Documentos (GED), según corresponda, en las declaraciones sujetas a revisión documentaria (canal naranja).</p> <ul style="list-style-type: none"> - Solo se considera las Declaraciones sin incidencia. - Se excluye las Declaraciones anuladas, legajadas,

	<p>anticipadas, urgentes, notificadas y aquellas que se encuentren fuera del horario de 8:00 a 17:59 horas de cada día.</p> <p>- El tiempo será medido a través del estadístico de la mediana, debido a la distribución sesgada de los tiempos.</p>			
Unidad de medida	Horas			
Fuente de Información	SIGED			
Datos históricos	2008:2.8 2009: 2 2010:2.1(*)			
Meta para el año 2011	Al 1er. Trimestre	Al 2do Trimestre	Al 3er. Trimestre	Al 4to. Trimestre
	2.5	2.5	2.5	2.5
Frecuencia de Reporte	Trimestral			
Disponibilidad de la información	10 días hábiles de culminado el trimestre de evaluación (remisión automática)			
Responsable	Intendencia Nacional de Técnica Aduanera			

Cuadro 05: Indicador 59 (Fuente SUNAT)



Grafica 20: Histograma del Indicador 58 (Elaboración propia en base a los datos del indicador SUNAT).

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES:

- La implementación de los modelos de Minería, resulto en una mejora directa en la selectividad del sistema de control, ya que al poder evaluar información histórica mediante la minería de datos, se puede predecir o encontrar patrones de comportamiento de las empresas que intentan realizar fraudes en sus importaciones, eso se observa en la mejora de la selectividad que aumento en casi un 8% según los indicadores de la SUNAT.
- La mayor selectividad, tiene como resultado directo una reducción de las tareas operativas a realizar en las diferentes aduanas de ingreso de mercancías, lo cual es realmente una mejora substancial para la empresa, esto se observa en la reducción de casi 1 hora en el tiempo promedio de despacho de importaciones.
- El tiempo de atención de las importaciones se redujo en casi 1 hora, lo cual mejora los indicadores doing business del país.

RECOMENDACIONES:

- Es necesario contar con un soporte directo para la herramienta SAS Enterprise Miner, este soporte se observo necesario, ya que los usuarios luego de la implementación se encontraban con muchas dudas acerca de la herramienta, ya que a pesar de conocer a grandes rasgos la herramienta y de poder definir un modelo relativamente bien, el usuario aun queda con dudas de cómo se debe realizar el proceso de creación de un modelo de la mejor manera.

- Es necesario crear una metodología para la implementación de nuevos modelos de minería, por ser la primera experiencia se realizaron los modelos muy rápidamente y realmente no fueron lo suficientemente documentados.

- Es necesario que el módulo de Selección de Riesgo, cuente con un modulo de seguridad y auditoria interno, para que se pueda gestionar de mejor manera los accesos y quien realiza cambios al sistema o a las tablas de administración, ya que ahora no se está realizando estas tareas de la mejor manera.

GLOSARIO Y TÉRMINOS

- **SI:** Sistema de Información

- **TI:** Tecnología de Información.

- **DUA:** Declaración Única de Aduanas, es un documento administrativo necesario para el cumplimiento de las formalidades aduaneras requeridas para las operaciones de intercambio de mercancías.

- **NPDA:** Nuevo Proceso de Despacho Aduanero.

- **SIGAD:** Sistema de gestiona Aduanera

- **NSIGAD:** Nuevo Sistema de Gestión Aduanera.

BIBLIOGRAFIA

- MINERIA DE DATOS Técnicas y Herramientas
Editorial Thomson International (Junio 2007)
Autor: César Pérez López

- Superintendencia Nacional de Administración Tributaria (SUNAT)
Disponible en: <http://www.sunat.gob.pe>
Visión, Misión, Valores, Servicios (Información Pública).
Autor: Red SUNAT

- Aplicación de la minería de datos en el control de fraude Aduanero.
Disponible en: <http://www.scribd.com/doc/11352087/Mineria-de-Datos-en-Aduanas>
Autor: Lic. Alberto Enrique Villalobos Chaves

- El Análisis Discriminante
Disponible en: <http://tgrajales.net/estdiscriminante.pdf>
Autor: Lic. Tevni Grajales Guerra