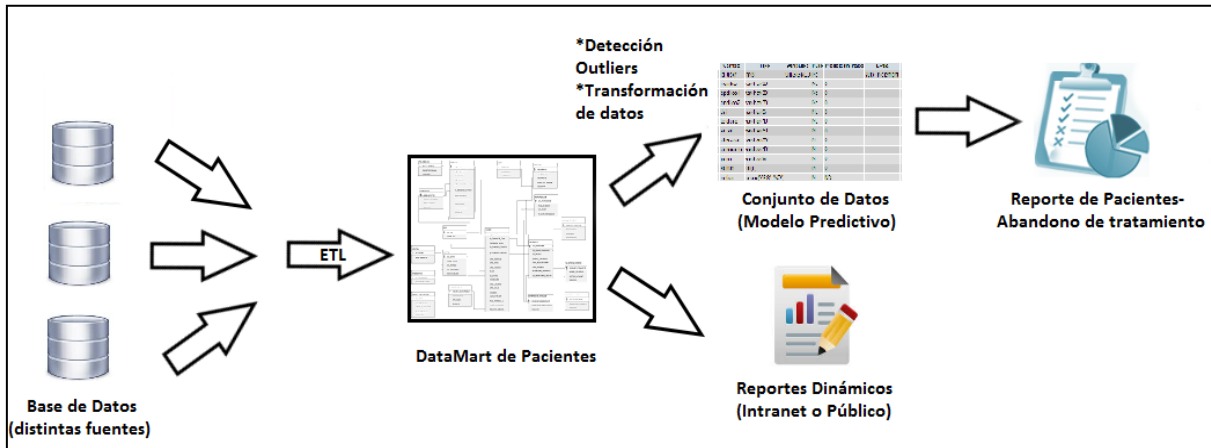


9 ANEXOS

9.1 Proceso General del proyecto para el INEN y Católica



9.2 Tabla: Diccionario de Datos del modelo del DataMart

| Tabla: DTM_Paciente_Mov | | | |
|-------------------------|--------|--------------|--|
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_paciente | 11 | Entero | Identificador de cada paciente (se debe mantener la protección de la identidad del paciente, por tal motivo, corresponde a un identificador diferente al DNI del paciente) |
| cod_ubigeo_nacimiento | 11 | Entero | Llave foránea de la tabla DTM_Ubigeo_Mae |
| cod_ubigeo_residencia | 11 | Entero | Llave foránea de la tabla DTM_Ubigeo_Mae |
| cod_cond_economica | 11 | Entero | Llave foránea de la tabla DTM_Cond_Socioeconomica_Mae |
| cod_estado_civil | 11 | Entero | Llave foránea de la tabla DTM_Tabla_Parametrica_Mae. Situación de las personas determinada por su relación proveniente del matrimonio |
| cod_grado_instruccion | 11 | Entero | Llave foránea de la tabla DTM_Tabla_Parametrica_Mae. Indica el nivel académico máximo obtenido por el paciente |

| | | | |
|-----------------|----|----------|--|
| cod_religion | 11 | Entero | Llave foránea de la tabla DTM_Tabla_Parametrica_Mae. Indica la creencia religiosa del paciente |
| des_sexo | 10 | Caracter | Llave foránea de la tabla DTM_Tabla_Parametrica_Mae. Indica el sexo del paciente. |
| fec_nacimiento | 7 | Fecha | Fecha de nacimiento del paciente |
| fec_admision | 7 | Fecha | Fecha en que el paciente se registro en el INEN, apertura de la historia clínica |
| estado_paciente | 25 | Caracter | Estado actual del paciente |

Tabla: DTM_Tratamiento_Eschema_Mov

| Atributo | Tamaño | Tipo de Dato | Descripción |
|---------------------------|--------|--------------|---|
| cod_tratamiento_esquema | 11 | Entero | Identificador del esquema de tratamiento que viene recibiendo el paciente. Se asigna un tipo diferente de esquema de acuerdo al tipo y estadio del cáncer |
| cod_departamento_atencion | 11 | Entero | Llave foránea de la tabla DTM_Departamento_Atencion_Mae |
| cod_tipo_tratamiento | 11 | Entero | Llave foránea de la tabla DTM_Tipo_Tratamiento_Mae |
| cod_tipo_esquema | 11 | Entero | Llave foránea de la tabla DTM_Tipo_Eschema_Mae |
| cod_cond_economica | 11 | Entero | Llave foránea de la tabla DTM_Cond_Socioeconomica_Mae. Indica la condición socioeconómica del paciente al momento de recibir el tratamiento |
| cod_diagnostico_paciente | | | Llave foránea de la tabla DTM_Diagnostico_Paciente_Mov |
| fec_inicio_tratamiento | 7 | Fecha | Fecha en que el paciente inicio el tratamiento |
| num_sesiones_programadas | 11 | Entero | Número de sesiones programadas para dicho esquema de tratamiento |

Tabla: DTM_Sesion_Tratamiento_Mov

| Atributo | Tamaño | Tipo de Dato | Descripción |
|-------------------------|--------|--------------|--|
| cod_sesion_tratamiento | 11 | Entero | Identificador de cada sesión asistida por el paciente para el tratamiento respectivo |
| cod_tratamiento_esquema | 11 | Entero | Llave foránea de la tabla DTM_Tratamiento_Eschema_Mov |
| fec_sesion | 7 | Fecha | Fecha en que se realizó la atención para la sesión del tratamiento |

Tabla: DTM_Tipo_Tratamiento_Mae

| Atributo | Tamaño | Tipo de Dato | Descripción |
|----------|--------|--------------|-------------|
|----------|--------|--------------|-------------|

| cod_tipo_tratamiento | 11 | Entero | Identificador del tipo de tratamiento. El modelo contemplará solo los tratamientos de quimioterapia; sin embargo, se puede adaptar para abarcar otros tipos de tratamiento en un futuro |
|---|--------|--------------|---|
| des_tipo_tratamiento | 50 | Caracter | Nombre del tipo de tratamiento |
| Tabla: DTM_Tipo_Eschema_Mae | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_tipo_esquema | 11 | Entero | Identificador del tipo de esquema. Dependiendo del tipo de cáncer y del estadio del paciente, el tratamiento tendrá un tipo de esquema el cual dictará la forma en que se llevará el tratamiento. |
| des_tipo_esquema | 50 | Caracter | Nombre del tipo de esquema según el INEN |
| Tabla: DTM_Ubigeo_Mae | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_ubigeo | 11 | Entero | Identificador de la ubicación geográfica (incluye países, departamentos, provincias y distritos) |
| cod_ubigeo_superior | 11 | Entero | Llave foránea de la tabla DTM_Ubigeo_Mae. Representa el nivel superior del lugar de procedencia. |
| des_ubigeo | 50 | Caracter | Nombre o denominación de la ubicación geográfica. |
| des_tipo_ubigeo | 25 | Caracter | Tipo de ubigeo. Puede ser "Pais", "Departamento", "Provincia" y "Distrito" |
| distancia_de_lima | 32 | Real | Distancia desde la ubicación geográfica a la ciudad Lima, para asegurar una mayor precisión en el modelo. |
| Tabla: DTM_Cond_Socioeconomica_Mae | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_cond_socioeconomica | 11 | Entero | Identificador de la condición socioeconómica. El INEN maneja esta variable como el nombre del seguro con el cual acude el paciente como una aproximación de su condición económica |
| des_seguro | 25 | Caracter | Nombre del seguro del paciente |
| des_tipo_seguro | 10 | Caracter | Tipo de seguro del paciente. Puede ser "Privado" y "Público" |
| valor_normalizacion | 11 | Entero | Valor que se utilizará para representar al tipo de seguro |

| Tabla: DTM Hospitalizacion_Mov | | | |
|--------------------------------------|--------|--------------|--|
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_hospitalizacion | 11 | Entero | Identificador de una hospitalización del paciente. Servicio orientado a brindar cuidados básicos en los ambientes del hospital durante un periodo de tiempo. |
| cod_paciente | 11 | Entero | Llave foránea de la tabla DTM_Paciente_Mov |
| fec_inicio_hospitalizacion | 7 | Fecha | Fecha en que el paciente se hospitalizó |
| fec_fin_hospitalizacion | 7 | Fecha | Fecha en que el paciente se retiró de su ambiente de hospitalización |
| Tabla: DTM Operacion_Mov | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_operacion | 11 | Entero | Identificador de una operación del paciente. |
| cod_paciente | 11 | Entero | Llave foránea de la tabla DTM_Paciente_Mov |
| fec_operacion | 7 | Fecha | Fecha en que se realizó la operación |
| Tabla: DTM Consulta_Ext_Mov | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_consulta_ext | 11 | Entero | Identificador de una consulta de paciente por atención externa. No incluye las consultas por emergencia |
| cod_paciente | 11 | Entero | Llave foránea de la tabla DTM_Paciente_Mov |
| cod_departamento_atencion | 11 | Entero | Llave foránea de la tabla DTM_Departamento_Atencion_Mae |
| fec_consulta | 7 | Fecha | Fecha en que se realizó la consulta |
| Tabla: DTM Departamento_Atencion_Mae | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_departamento_atencion | 11 | Entero | Identificador del departamento de atención. Se encuentra dividido de acuerdo al tipo de cáncer que se trata |
| des_departamento_atencion | 50 | Caracter | Nombre del departamento de atención |
| Tabla: DTM Diagnostico_Mae | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_diagnostico | 11 | Entero | Identificador del diagnóstico o del cáncer. Será de acuerdo a la clasificación del CIE-10 |

| cod_departamento_atencion | 11 | Entero | Llave foránea de la tabla DTM_Departamento_Atencion_Mae |
|--|--------|--------------|--|
| des_diagnostico | 100 | Caracter | Nombre del diagnóstico o cáncer |
| des_tipo_diagnostico | 25 | Caracter | Tipo de diagnóstico al cual pertenece el cáncer. Puede ser "Morfológico" y "Topográfico" |
| des_version_clasificacion | 25 | Caracter | Clasificación de los tipos de cáncer de acuerdo a la versión de la CIE |
| des_grupo_cancer | 50 | Caracter | Grupo de cáncer al cual pertenece el diagnóstico. Se encuentra organizado por el órgano donde se encuentra al cáncer |
| Tabla: DTM_Diagnostico_Paciente_Mov | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_diagnostico_paciente | 11 | Entero | Identificador de cada diagnóstico que se le brinda al paciente. |
| cod_diagnostico_topo | 11 | Entero | Llave foránea de la tabla DTM_Diagnostico_Mae Indica el diagnóstico topografico del cáncer. |
| cod_diagnostico_morfo | 11 | Entero | Llave foránea de la tabla DTM_Diagnostico_Mae Indica el diagnóstico morfologico del cáncer. |
| cod_paciente | 11 | Entero | Llave foránea de la tabla DTM_Paciente_Mov |
| cod_estadio_cancer_actual | 11 | Entero | Llave foránea de la tabla DTM_Estadio_Cancer_Mae. |
| fec_diagnostico | 7 | Fecha | Fecha en que se diagnosticó el cáncer al paciente |
| Tabla: DTM_Estadio_Cancer_Mae | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_estadio_cancer | 11 | Entero | Identificador del estadio del cáncer. Indica el grado o nivel en el que se encuentra el cáncer (un valor otorgado por un doctor) |
| des_estadio_cancer | 50 | Caracter | Nombre o descripción del estadio del cáncer |
| valor_normalizacion | 11 | Entero | Valor que representa al estadio del cáncer durante la normalización |
| Tabla: DTM_Tabla_Parametrica_Mae | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_parametro | 11 | Entero | Identificador del parametro. |
| des_nombre_tabla | 25 | Carácter | Nombre de la tabla a la cual pertenece el parámetro |
| valor | 50 | Carácter | Valor del parámetro en caracteres. |
| cod_real | 11 | Entero | Código original de la tabla que |

| | | | proviene |
|--|--------|--------------|--|
| Tabla: DTM_Historial_Actualizacion_Mov | | | |
| Atributo | Tamaño | Tipo de Dato | Descripción |
| cod_historial | 11 | Entero | Identificador de la actualización del modelo de DataMart |
| fec_actualizacion | 7 | Fecha | Fecha en que se realizó la actualización |
| estado_actualizacion | 25 | Carácter | Estado de la actualización del DataMart |
| num_pacientes_nuevos | 11 | Entero | Cantidad de pacientes insertados o nuevos |
| observaciones | 100 | Carácter | Observaciones del proceso de actualización |



9.3 Descripción de los principales procesos ETL para la carga del DataMart

| Nombre del Proceso | Nombre del archivo | Definición |
|-------------------------------------|--------------------------------------|--|
| Obtener Variables Iniciales | getVariablesDataMart.ktr | Permite obtener los parámetros iniciales para el proceso ETL como fecha de actualización, fecha de última actualización, constantes, entre otros |
| DTM_TablasMaestras | carga_DataMart_MAE.kjb | Permite realizar la carga de todas las tablas maestras |
| Nombre del Sub proceso-Maestras | Nombre del archivo | Definición |
| 1) Actualizar Tabla Parametrica | carga_DM_parametricaMAE.ktr | Permite realizar la actualización de la tabla maestra "dtm_tabla_parametrica_mae" |
| 2) Actualizar Estadio Cancer | carga_DM_estadioMAE.ktr | Permite realizar la actualización de la tabla maestra "dtm_estadio_cancer_mae" |
| 3) Actualizar Condición Economica | carga_DM_condicionMAE.ktr | Permite realizar la actualización de la tabla maestra "dtm_cond_socioeconomica_mae" |
| 4) Actualizar Departamento Atencion | carga_DM_departamentoatencionMAE.ktr | Permite realizar la actualización de la tabla maestra "dtm_departamento_atencion_mae" |
| 5) Actualizar Diagnostico | carga_DM_diagnosticoMAE.ktr | Permite realizar la actualización de la tabla maestra "dtm_diagnostico_mae" |
| 6) Actualizar Tipo Tratamiento | carga_DM_tipotratamientoMAE.ktr | Permite realizar la actualización de la tabla maestra "dtm_tipo_tratamiento_mae" |
| 7) Actualizar Ubigeo | carga_DM_ubigeoMAE.ktr | Permite realizar la actualización de la tabla maestra "dtm_ubigeo_mae" |
| DTM_TablasTransaccionales | carga_DATAMart_MOV.kjb | Permite realizar la carga de todas las tablas transaccionales |

| Nombre del Sub proceso-Transaccionales | Nombre del archivo | Definición |
|--|-------------------------------------|---|
| 1) Actualizar Pacientes | carga_DM_pacienteMOV.ktr | Permite realizar la actualización de la tabla maestra "dtm_paciente_mov" |
| 2) Actualizar Operaciones | carga_DM_operacionMOV.ktr | Permite realizar la actualización de la tabla maestra "dtm_operacion_mov" |
| 3) Actualizar Hospitalizaciones | carga_DM_hospitalizacionMOV.ktr | Permite realizar la actualización de la tabla maestra "dtm_hospitalizacion_mov" |
| 4) Actualizar Consultas Externas | carga_DM_consultaExtMOV.ktr | Permite realizar la actualización de la tabla maestra "dtm_consulta_ext_mov" |
| 5) Actualizar Diagnostico Paciente | carga_DM_diagnosticoPacienteMOV.ktr | Permite realizar la actualización de la tabla maestra "dtm_diagnostico_paciente_mov" |
| 6) Actualizar Tratamiento | carga_DM_tratamientoMOV.ktr | Permite realizar la actualización de la tabla maestra "dtm_tratamiento_esquema_mov" y la tabla "dtm_sesion_tratamiento_mov" |

9.4 Métodos de normalización para cada variable

| Variables | Métodos de normalización utilizados |
|--------------------------|--|
| Condición Económica | 1) Transformación cadena de caracteres-Variable Ordinal 2) Normalización Mínimo-Máximo 3) Suavización por redondeo |
| Estado civil | |
| Lugar nacimiento | 1) Normalización Mínimo-Máximo 2) Suavización por redondeo |
| Lugar residencia | 1) Normalización Mínimo-Máximo 2) Suavización por redondeo |
| Sexo | |
| Edad | 1) Normalización Mínimo-Máximo 2) Suavización por redondeo |
| Primera Consulta | 1) Normalización Mínimo-Máximo 2) Suavización por redondeo |
| Diagnostico topográfico | 1) Suavización por agrupación |
| Estadio | 1) Transformación cadena de caracteres-Variable Ordinal 2) Normalización Mínimo-Máximo 3) Suavización por redondeo |
| Numero operaciones | 1) Normalización Mínimo-Máximo 2) Suavización por redondeo |
| Numero hospitalizaciones | 1) Normalización Mínimo-Máximo 2) Suavización por redondeo |
| Brecha | 1) Normalización Mínimo-Máximo 2) Suavización por redondeo |
| Tratamiento Abandona | |

9.5 Descripción de los principales procesos ETL para el proceso automatizado de Minería de Datos

| Nombre del Proceso | Nombre del archivo | Definición |
|--------------------|--------------------------|---|
| Parámetros | Parametros_Mining.ktr | Permite obtener los parámetros iniciales para la ejecución del proceso de Minería de Datos automatizado, tales como: tamaño del modelo, periodo de ejecución, parámetros para los algoritmos, tipos de reportes, entre otros. |
| Today Folder | | Permite crear el nuevo directorio para los archivos que se crearan en la ejecución actual. |
| Extracción | Extraccion_Mining.ktr | Permite extraer la información para el modelo y de las nuevas instancias desde el DataMart. Además, permite crear la variable objetivo (Abandona o Continua) para las muestras del modelo. |
| Valores Nulos | Nulos_Mining.ktr | Permite filtrar las muestras del modelo que tengan exceso de valores nulos o inconsistentes. Finalmente, se asignará un valor apropiado a los campos que aún permanezcan como nulos. |
| Normalización I | Normalizacion_Mining.ktr | Permite realizar la normalización de todas las instancias por el método de transformación de caracteres. |

| | | |
|-------------------|-------------------------------|---|
| Min-Max Variables | MinMax_Variables.ktr | Permite obtener los parámetros de mínimos y máximos para los campos que seguirán el método de normalización Min-Max. |
| Normalización II | MinMax_Suavizacion_Mining.ktr | Permite realizar los métodos de normalización Min-Max y de Suavización. Devuelve los archivos de extensión arff que conforman el modelo y las nuevas instancias a probar. |
| Valores Anómalos | Outlier_Mining.kjb | Permite ejecutar un archivo .jar con el algoritmo de detección y eliminación de valores anómalos. Finalmente, se tendrá el nuevo modelo de datos más refinado. |
| J48 DataMining | J48DataMining.kjb | Permite ejecutar un archivo .jar con el algoritmo de árboles de decisión de la librería Weka, con el fin de clasificar cada nueva instancia en el modelo de datos. Finalmente, se generarán reportes por departamento de atención y reportes generales. |