

## **RESUMEN**

La Dirección del Seguro de Salud del ISSFA gestiona los procesos para la prestación del Servicios de Salud, uno de ellos corresponde a la Facturación médica, mediante el proceso de “Pertinencia y Liquidación de Servicios de Salud” se realizan los pagos a los prestadores de salud. En este contexto, se almacenan grandes volúmenes información que requieren análisis a fin de ofrecer valor agregado al ISSFA. Esta investigación presenta el desarrollo de un modelo predictivo para determinar el impacto económico en el pago del seguro de salud del ISSFA en enfermedades músculo - esqueléticas y conocer que factores son influyentes para que un afiliado sea propenso a padecerlas. Con la metodología de minería de datos CRISP, se identifican: los objetivos del negocio y de minería de datos, evaluación inicial de herramientas y técnicas; en la segunda fase se realiza la exploración de los datos y verificación de calidad; en la fase de preparación de los datos se selecciona el conjunto inicial de datos, criterios de exclusión; en la fase de construcción se realiza la integración de datos para definir el conjunto final. Con la herramienta RapidMiner se construyen los modelos: regresión lineal, máquina de vector de soporte y redes neuronales para la predicción del valor a pagar, mientras que para determinar la combinación de factores de una enfermedad se utiliza al árbol de decisión; se evalúan los modelos en función de los la tasa de error, porcentaje de precisión entre otros y finalmente, realiza el análisis de resultados contrastándolo con el presupuesto asignado por servicio de salud, resultando ser la técnica de Regresión lineal la más óptima.

### **Palabras clave:**

- **MINERÍA DE DATOS**
- **MODELO PREDICTIVO**
- **REGRESIÓN LINEAL**
- **ÁRBOLES DE DECISIÓN**

## **ABSTRACT**

The “Dirección del Seguro de Salud del ISSFA” manage the process of Health Services, one of them is Medical Billing. Through the process of “Relevance and liquidation of health services” the payments are made to Hospitals. In this context, there is a large volume of information stored that needs analytics for the purpose of give value added to ISSFA. This research shows the development of one predictive model to determine the financial impact in health services’s payments in health insurance of ISSFA in skeletal muscle diseases and identify which are the influential factors for an affiliate can get. With the methodology of data mining CRISP can identify: business and data mining objectives, initial evaluation of tools and data mining techniques; in second phase it shows the data exploration and quality verification; in data preparation data it selects the initial data set, exclusion criteria; in construction phase the data integration is made to define the final dataset. With the RapidMiner Tool the models: linear regression, support vector machine and neural net are built, for the prediction of value of payment, while to find the factors combination of one disease the decision tree is built; these models are evaluated and taking the error rate, accuracy percentage and others. Finally, the analysis of results is made contrasting with the allocated budget for healthy service, resulting the linear regresión as the more optimal.

### **Keywords:**

- **DATA MINING**
- **PREDICTIVE MODEL**
- **LINEAR REGRESSION**
- **DECISSION TREE**