

Sağlıkta Veri Madenciliğiyle Anti-Fraud Uygulamalar: Veri Sorunlarının Çözümü İçin Veri Madenciliğine Dayalı Profillendirme Yaklaşımı

Ali Serhan KOYUNCUGİL^a, Nermin ÖZGÜLBAŞ^b

^aSermaye Piyasası Kurulu Araştırma Dairesi, Ankara

^bBaşkent Üniversitesi Sağlık Kurumları İşletmeciliği Bölümü, Ankara

Anti-Fraud Applications via Data Mining in Healthcare: Profiling Approach for Solution of Data Problems Based on Data Mining

Özet

Günümüzde, bilişim teknolojilerine dayalı gözetim ve denetim sistemlerinde en yaygın ve etkin olarak kullanılan analitik yaklaşım veri madenciliğidir. Veri madenciliği, yaygın ve en basit tanımıyla, büyük hacimde veri içerisinde gizli kalmış, faydalı bilgi, ilişki ve örüntülerin otomatik olarak açığa çıkarılıp, karar destek amaçlı olarak kullanılması olarak ifade edilmektedir. Veri madenciliğinin iş hayatına girmesiyle beraber, devasa veri kümeleri içerisinde tespit edilebilen davranışsal örüntüler, fraud (hilekarlık) tespitinde yaygın olarak kullanım imkanı bulmasına yol açmıştır. Dolayısıyla, günümüzde etkin olarak kullanılan anti-fraud (hilekarlık önleme) sistemlerin pek çoğu veri madenciliğini temel almaktadır. Hastalar, sağlık personeli, hastane yönetimi veya tedarikçi kuruluşlar tarafından oluşturulacak hilekarlıkların tespiti için anti-fraud sistemlerin geliştirilmesi ekonomik olduğu kadar toplumsal anlamda da önem taşımakta, konuyla ilgili soruşturma sonuçlarının cezai yaptırımlara yol açabileceği düşünüldüğünde de fraud tespitinin çok hassas bir dengede yürütülmesi gerekliliği ortaya çıkmaktadır.

Sağlıklı işleyen bir anti-fraud sistemin ortaya konulması, ancak, veri madenciliğinin temel süreci veya bir başka deyişle özü diyebileceğimiz modelleme aşamasında doğru bir yapılandırmaya gidilmesiyle mümkün olmaktadır. Doğru bir modelleme yapmak aşağıdaki gibi özetlenebilecek bir dizi ardışık ve karmaşık alt süreci beraberinde getirmektedir; (1) Problemi iyi ve doğru tanımlama, (2) Doğru amaç tespit etme, (3) Amaca uygun doğru veriyi tespit etme, (4) Veriye ve amaca uygun doğru yöntemi tespit etme, (5) Çıktıları doğru anlayabilme ve uygulayabilme. Buna karşın, modelleme sürecinin en önemli zaafı temiz veriye olan ihtiyaç olarak karşımıza çıkmaktadır. Şüphesiz ki, modellemeye geçmezden önce verinin içerdiği sorunların belirlenip, çözülmesi gerekmektedir.

Bu çalışmada amaç, sağlık sektöründeki hasta, sağlık personeli, hastane yöneticileri veya tedarikçilerine yönelik veri madenciliğine dayalı anti-fraud modellemelerinde sıklıkla sorun haline gelen ve uygun bir model üretmek için en önemli aşamalardan biri olan veri sorunun çözümü için bir yaklaşım önermektir. Veri sorunları denildiğinde, temel anlamda, kayıp veri, tutarsız veri, aykırı değer, uç değer biçiminde özetlenebilir. Bu çalışmada, parametrik olmayan, dinamik bir yaklaşım izleyip verinin kendi içerisindeki anomalilerin tespitinden hareket ederek veri sorunlarını gidermeyi amaçlayan veri madenciliğine dayalı profillendirme yaklaşımı sunulmaktadır. Bu çalışma kapsamında önerilecek olan yaklaşım ile sağlık sektöründe anti-fraud uygulamalar gerçekleştirecek araştırmacılar, bir şablon model kullanmak yerine, veri madenciliği yöntemlerinin kullanımıyla, her farklı örüntü içerisinde yer alan veri sorununu, örüntünün özellikleri dikkate alınarak giderebilme imkanına kavuşacaktır. Böylelikle, sağlık sektörü anti-fraud modelleri az yarılgıya izin verecek bir çözüme ulaşmış olacaktır.

Anahtar kelimeler

Veri madenciliği; Anti-fraud sistemi; Profillendirme; Gözetim ve denetim sistemleri; Veri sorunları; Kayıp veri; Tutarsız veri; Aykırı değer; Uç değer; Sağlıkta veri madenciliği