

Acil Servislerde Mortalite Tahmini İçin Veri Madenciliği Tekniklerinin Kullanılması

Deniz ÖZEL^a, Uğur BİLGE^a, Rezan TAHTACI^b

^a Biyoistatistik ve Tıp Bilişimi AD, Akdeniz Üniversitesi, Antalya

^b Numune Eğitim Ve Araştırma Hastanesi Acil Tıp Kliniği, Ankara

Özet: Acil serviste ilk etapta hastalığın ciddiyetinin belirlenmesi, hastanın uygun bölüme alınıp uygun tanı ve tedavi yöntemlerinin kullanılması açısından önemlidir. Acil servislerde hastalığın ciddiyetini değerlendirmek için kullanılan bazı objektif skorlama sistemleri olmasına rağmen, bunların acil servislerdeki kullanımı henüz kesin olarak onaylanmamıştır. Hastaların durumlarının ciddiyetini belirlemede istatistiksel yöntemlerden faydalanılabilir. Bu çalışmanın amacı acil servise başvuran hastaların mortalitesine etki eden demografik ve klinik faktörleri ortaya koymaktır. Bu amaçla, acil servise başvuran hastaların demografik bilgilerinden, öykülerinden, yaşamsal bulgularından ve sıklıkla istenen tetkiklerden elde edilen verilerle hesaplanabilen skorlama sistemleri arasında en önemli dört skorlama sistemini kullanarak tüm bu değişkenlerin mortaliteyi tahmin etmedeki etkinliklerini değerlendirmede veri madenciliği tekniklerinden faydalanılmıştır. Hasta verileri, 1 Ekim 2010-31 Ekim 2010 tarihleri arasında Ankara Numune Eğitim ve Araştırma Hastanesi Acil Tıp Kliniği'ne travma dışı nedenlerle başvuran hastaların dosya ve otomasyon sistemi kayıtlarından toplanmıştır. Toplam 269 hastaya ait veriler Lojistik Regresyon (LR) Analizi ve Genetik Algoritma (GA) ile değerlendirilmiştir. LR analizine katılacak değişkenlerin seçiminde ki-kare testleri yapılmıştır. Ki-kare testlerinde mortalite ile ilişkili bulunan ($p<0,05$) değişkenler LR modeline alınmıştır. LR analizi ile %26,2 duyarlılık, %98,2 seçicilik ile yaş ($p=0,041$) ve ilk 24 saat içindeki klinik durumunu ifade eden APACHE II skoru (Acute Physiology And Chronic Health Evaluation; $p=0,017$) mortaliteyi belirleyen risk faktörleri olarak bulunmuştur. GA'da ise üç farklı kural ile mortaliteyi öngören değişkenler bulunmuştur. 1. kuralda yaş (kesim noktası 45 yaş) ve APACHE II (kesim noktası 0,08) skoru %95,1 duyarlılık, %42,1 seçicilik ile mortalitede etkili bulunurken, 2. kuralda APACHE II (kesim noktası 0,131) ile beraber hastanın yaşamsal bulgularının değerlendirildiği MEW skoru (kesim noktası 1) ve SO_2 (Oksijen Saturasyonu; kesim noktası 94) değişkenleri %65,8 duyarlılık ve %85,1 seçicilikle anlamlı bulunmuştur. Son kural ise duyarlılık (%75,6) ve seçiciliğin (%75,4) birbirine çok yakın olduğu, MEW skoru (Modified Early Warning Score; kesim noktası 1) ve SO_2 (kesim noktası 94) değişkenlerini içermektedir. GA yöntemi ile yüksek duyarlılık veya yüksek seçicilik değerlerine göre, belirli kesim noktalarıyla mortalitede etkili değişkenler seçilebilir. Bu özelliği nedeniyle GA yönteminin, LR yönteminden daha esnek ve klinisyen kararlarında daha kullanışlı olduğu söylenebilir. Sonuç olarak veri madenciliği tekniklerinden faydalanarak acil servise başvuran kritik hastaların önceden tespitiyle mortalite oranları düşürülebilir, hasta triaji ve yönlendirmeleri daha doğru şekilde yapılabilir.

Anahtar Kelimeler: Lojistik Regresyon, Genetik Algoritma, Acil Servis, Mortalite.

Using Data Mining Techniques For The Prediction Of Mortality At Emergency Services

Abstract: Determination of severity of patient's illness at emergency service at the first point is important at the aspect of taking the patient to the appropriate room for the using of appropriate diagnostic and treatment facilities. Although there are some objective scoring systems for identification of severity of sickness at emergency service, there is no approval of these scoring systems. Statistical analyses could be used for the determination of severity of patient's status. The aim of this study is to determine the demographic and clinical factors which are affecting the mortality of patients at emergency services. For this issue; data-mining techniques has been used to determine the efficiency of these factors at the prediction of mortality at emergency services. All data is collected between 1 -31 October 2010 from the patients' files and the computer system

who were admitted to Ankara Numune Eğitim ve Araştırma Hastanesi Acil Tıp Kliniği for all the reasons except from trauma. Information collected from the sources are evaluated with Logistic Regression analysis(LR) and genetic algorithm (GA). Chi-square test has been done for selection of variables which will be assessed with LR analyses. Variables which are related to mortality according to chi-square test ($p < 0.05$) also evaluated with LR model. Age ($p = 0.041$) and the APACHE II score (Acute Physiology And Chronic Health Evaluation, which is explaining the clinical status of patient at first 24 hours)($p = 0.017$) are found with LR analysis as the risk factor which are specifying the mortality of patients (26.2% sensitivity, 98.2% specificity). At the 1st rule age (cut point 45 years) and APACHE II (cut point 0.08) are found as effective at mortality (42.1% sensitivity, 95.1% specificity); at 2nd rule APACHE II (cut point 0,131) and MEW score (cut point 1) which is analyzing the vital findings of patient, and SO_2 (oxygen saturation, cut point 94) are found to be meaningful (65.8% sensitivity, 85.1% specificity). The last rule includes MEW score and SO_2 and for these variables sensitivity (75.6%) and specificity (75.4%) values were so similar. Also with GA method with certain cut points some variables can be selected which are affecting the mortality with high sensitivity or high specificity. For this reason we can say that GA method is more flexible and easy to use for the clinicians than LR method. As a result, by the data-mining techniques using mortality rates of critical patients' can be reduced and triage can be done more accurately.

Sorumlu Yazarın Adresi

Deniz ÖZEL. Akdeniz Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi AD, 07059, Arapsuyu, Antalya. denizozel@akdeniz.edu.tr.