

Asemptomatik Karotid Arter Stenozunun Veri Madenciliği Yöntemleriyle İncelenmesi

Selen BOZKURT^a, Ugur BİLGE^a, Sedat DURMAZ^b, Kemal Hakan GULKESEN^a,
Saim YILMAZ^b

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

Application of Data Mining Techniques for Detecting Asymptomatic Carotid Artery Stenosis

Abstract: Asymptomatic carotid artery stenosis, one of the etiological factors for stroke, has several risk factors such as hypertension, smoking, diabetes, and physical inactivity. Understanding and determining factors that predispose to asymptomatic carotid stenosis will help in the design of acute stroke trials and in prevention programs. The goal of this study is to explore rules and relationships that might be used to detect possible asymptomatic carotid stenosis by using data mining techniques. For this purpose, Genetic Algorithms (GA), Logistic Regression (LR), and Chi-square tests have been applied to the patient dataset. Results of these tests have also been compared. We think that our study clearly shows that coronary artery disease (CAD) and hypertension are important indicators of prognosis for Asymptomatic Carotid Artery Stenosis patients. Both methods give better results than single approaches, LR having a higher ROC AUC. GA can be an alternative to LR, because of its simplicity to use in daily practice. The GA technique should be further investigated for the prediction of CAD.

Key Words: Genetic algorithms; logistic regression; data mining; radiology, asymptomatic carotid artery stenosis

Özet: Asemptomatik karotid arter stenozu inme için bir risk faktörü olarak kabul edilmekte olup hipertansiyon, sigara, diyabet ve yetersiz egzersiz gibi çeşitli risk faktörleri bulunmaktadır. Asemptomatik karotid arter stenozu risk faktörlerini anlayıp tanımlayabilmek inmeyle ilgili erken tanı ve komplikasyonların önlenmesiyle ilgili çalışmalara katkı sağlayacaktır. Bu çalışmanın amacı, asemptomatik karotid arter stenozu vakalarının saptanmasında kullanılabilecek kural ve ilişkileri veri madenciliği yöntemleriyle araştırmaktır. Bu amaçla, Genetik Algoritma (GA), Lojistik Regresyon (LR), ve ki-kare testi yöntemleri kullanılmıştır ve bu farklı yöntemlerin sonuçları karşılaştırılmıştır. Çalışmanın bulgularına göre koroner arter hastalığı veya hipertansiyon varlığı asemptomatik karotid arter stenozu hastalarının prognozunda önemli belirteçlerdir. Veri madenciliği yöntemleri karşılaştırıldığında, iki yöntem de birbirine yakın ve ki-kare testlerinden daha iyi sonuçlar vermişlerdir. LR'nun ROC altında kalan alan değerinin daha yüksek olduğu görülmüştür. Yine de GA'nın uygulamadaki kullanım kolaylığı avantajı nedeniyle LR'na iyi bir alternatif olduğu düşünülmektedir.

Anahtar Kelimeler: Genetik algoritma; lojistik regresyon; veri madenciliği; semptomatik karotid arter stenozu

Sorumlu Yazarın Adresi

Yrd. Doç. Dr. Ugur Bilge, Biyoistatistik ve Tıp Bilişimi AD, Akdeniz Üniversitesi, Antalya, ubilge@akdeniz.edu.tr