

2.6. Veri Madenciliği

Veri madenciliği, denetimsiz öğrenim yoluyla keşifsel veri analizine odaklanan ilgili bir çalışma alanıdır. Verilerdeki (geçmiş) bilinmeyen özelliklerin keşfedilmesine odaklanır. Bu veri tabanlarında bilgi keşfi analizinin bir adımıdır.

Veri Madenciliği ve Makine öğrenimi arasındaki fark nedir?

Makine öğrenimi, bilgisayarlara açıkça programlanmadan öğrenme yeteneği veren algoritmaların incelenmesi, tasarımı ve geliştirilmesi ile ilgilidir. Veri madenciliği, yapılandırılmamış verilerin bilgiyi veya bilinmeyen ilginç kalıpları çıkarmaya çalıştığı süreç olarak tanımlanabilir. Veri madenciliği işlemi sırasında, öğrenme algoritmaları kullanılır.

Makine öğrenmesi ve veri madenciliği genellikle aynı yöntemleri kullanır ve önemli ölçüde örtüşür, ancak makine öğrenmesi öğrenme verilerinden öğrenilen öngörüye odaklanırken, veri madenciliği verilerde (önceden) bilinmeyen özelliklerin keşfine odaklanır (bu veritabanlarında bilgi keşfinin analiz basamağı). Veri madenciliği birçok makine öğrenme yöntemi kullanır, ancak farklı hedefleri vardır.

Makine öğrenmesi ile genellikle aynı yöntemleri kullanır ve önemli ölçüde örtüşür, ancak makine öğrenmesi öğrenme verilerinden öğrenilen bilinen özelliklere dayanarak öngörüye odaklanırken, veri madenciliği verilerde (önceden) bilinmeyen özelliklerin keşfine odaklanır. Veritabanlarında bilgi keşfinin ilk basamağıdır.

Veri Madenciliği, büyük miktardaki veri yığını içerisinde desenlerin, ilişkilerin, önemli bilgilerin keşfedilmesi tekniğidir. Kötüye kullanım tespiti ve anormallik tespiti yöntemleri kullanılırken yapay sinir ağları, bayes ağlar ve KNN gibi sınıflandırma yöntemleri; bölünmeli, çizge tabanlı ve hiyerarşik demetleme yöntemleri, karar ağaçları ve genetik algoritmalar da mevcuttur. Ayrıca farklı yöntemlerin birleştirilmesi ile oluşturulan hibrit yöntemler de kullanılmaktadır.

Karar Destek Sistemleri

Karar Destek Sistemleri, değişik kaynaklardan toplanan bilgilerin düzenlenerek, karar modellenerek, bilgiler analiz edilerek ve değerlendirme sonuçlarını karar vericiye sunan bilgisayar tabanlı sistemlerdir. Bir karar verici için verilen kararın doğruluğu, onun yeteneklerine, deneyimine ve bilgi birikimine olduğu kadar sahip olduğu veri kümesinin yeterliliğine de bağlıdır. Kararın başarısında, verilerin doğru depolanması, doğru sınıflanması, doğru ayıklanıp işlenmesi ve doğru yorumlanması çok önemli bir rol oynar. Bu nedenle, veri madenciliği, Karar Destek Sistemleri için etkili araçlardan biridir.