Makine Öğrenmesi Yöntemleri İle Firma Başarısızlığı Tahminlemesi

Çalışmanın Amacı:

Bu çalışmada, Erzurum ilinde 38 yıldır faaliyet gösteren bir inşaat malzemeleri toptancısı firmanın müşterisi 157 firmanın verileri kullanılarak, YSA, DVM, TBA-YSA ve TBA-DVM yöntemleri ile borçlarını ödeme ve ödememe olasılıkları firma başarısızlığı kapsamında ele alınmış ve hangi müşterilere vadeli satış yapılmasına karar verilebilmesi için firma başarısızlığı tahminlemesi gerçekleştirilmiştir.

Yöntem:

Bu bölümde, firma başarısızlığı tahminlemesinde kullanılacak TBA, YSA ve DVM’ler hakkında kısa açıklayıcı bilgiler sunulacaktır.

Temel Bileşenler Analizi (Principle Component Analysis):

TBA çok sayıda birbiri ile ilişkili değişkenler içeren veri setinin boyutlarını veri içerisinde varolan değişimlerin mümkün olduğunca korunarak daha az boyuta indirgenmesini sağlar . Bu yönüyle TBA verideki gerekli öznitelikleri ortaya çıkarmada oldukça etkili bir yöntemdir. Yüksek boyutlu verilerdeki genel özellikleri bularak boyut sayısının azaltılmasını ve verinin sıkıştırılmasını sağlar. Boyut azalmasıyla bazı özellikler kaybedilebilir ama amaçlanan, bu kaybolan özelliklerin veri seti hakkında çok az bilgi içeriyor olmasıdır

Yapay Sinir Ağları (Artificial Neural Networks):

YSA insan beyninin çalışma şeklinin simüle edilmesiyle bilgi işleme fonksiyonundan esinlenerek geliştirilmiştir. İlk YSA modeli 1943 yılında W. McCulloch ve W. Pitts tarafından sunulmuştur . 1948 yılında Wiener “Cybernetics” isimli eserinde, sinirlerin işlevi ve çalışma prensiplerini, 1949’da ise Hebb, “Organization of Behavior” isimli eserinde öğrenme ile ilgili temel teoriyi ele almıştır. Hebb eserinde, öğrenebilen ve uyum sağlayabilen sinir ağları modeli için temel oluşturacak “Hebb Kuralı”nı ortaya koymuştur . 1957 yılında ise F. Rosenblatt’ın, insan beyninin işlevlerini modelleyebilmek amacıyla geliştirdiği tek katmanlı, eğitilebilen ve tek çıkışlı bir ağ modeli olan “Perceptron”dan sonra, YSA ile yapılan çalışmalar artış göstermiştir .

Destek Vektör Makineleri (Support Vector Machines):

1960’larda Vapnik tarafından teorik temelleri atılan DVM’ler, ilk defa 1995 yılında bir sınıflandırma probleminin çözümünde kullanılmıştır. DVM’ler en basit ifadesiyle sınıflandırma (classification) ve regresyon (regression) problemlerinin analizinde kullanılan bir öğrenen makinedir.

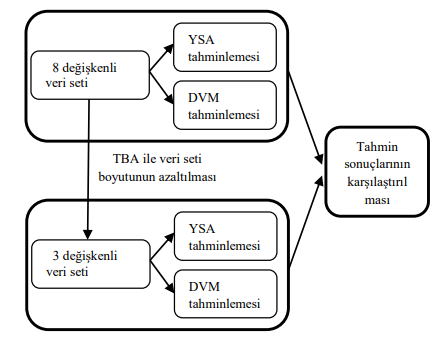
DVM’ler esas olarak iki grup örneklem arasındaki mesafeyi maksimize eden optimum ayıraç hiperdüzlem (hyperplane) ile ikili (binary) sınıflandırma yaparlar. Hiperdüzlemin sınırlarına (margins) en yakın grup elemanlarına destek vektörleri (support vectors) denir.

DVM’ler finans ve ekonomiden mühendislik uygulamalarına kadar pek çok farklı alanda etkin olarak kullanılmaktadır.

Veri Seti Ve Analiz:

Bu çalışmada kullanılan veri seti, Erzurum ilinde 38 senedir faaliyet yürüten bir inşaat malzemeleri toptancısı firmanın 157 müşterisinin 2006-2013 yılları arasındaki gerçek verileri ile oluşturulmuştur.

Bu bölümde, öncelikle firma başarısızlığı tahminlemesi için elde edilen veri seti tanıtılacak, daha sonra uygulanan prosedürler açıklanarak analizler gerçekleştirilecek ve elde edilen sonuçlar yorumlanacaktır. Bu çalışmada firma başarısının tahminlenmesi için önerilen metodolojinin adımları Aşağıda şematik olarak sunulmuştur.

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

Çalışmanın bundan sonraki aşamasında uygulaması gerçekleştirilecek TBA-YSA ve TBA-DVM hibrit yöntemlerinde “Firma Yaşı”, “Vadesi Geçen Toplam Borç Aralığı” ve “Ortalama Sipariş Tutarı” değişkenleri esas alınarak tahminleme gerçekleştirilecektir.

Bulgular:

Bu çalışmada firma başarısızlığı tahminlemesi için uygun bir makine öğrenmesi yöntemi araştırılmıştır. Bu kapsamda öncelikle Model-1 ve Model-2 için yalın YSA ve DVM yöntemlerinin tahmin başarısı değerlendirilmiştir. Her ne kadar yalın modeller ile tahmin başarıları için tatmin edici sonuçlar elde edilmiş olsa da daha iyi sonuçlar elde edilememesinin veri boyutundan kaynaklandığı değerlendirilmiş, veri seti boyutunun azaltılması ve tahmin performansının artırılması için ise TBA uygulanmıştır. TBA kullanılarak oluşturulan hibrit modellerin tahmin başarısını artırdığı gözlenmiştir. Benzer şekilde Model-1 ve Model-2 için hibrit modellerin performansları da karşılaştırılmıştır.

tablo içeren bir resim

Açıklama otomatik olarak oluşturuldu

Sonuç:

Bu çalışmada Erzurum ilinde 38 yıldır faaliyet gösteren inşaat malzemeleri toptancısı bir firmanın müşterilerinin vadeli borçlarını ödeme/ödememe riskleri firma başarısızlığı kapsamında ele alınmış ve firma başarısızlığının tahmininde uygun bir makine öğrenmesi yöntemi araştırılmıştır.

Yalın YSA ve DVM yöntemleri ile gerçekleştirilen tahminlemelerde firma başarısızlığına etki eden 8 farklı değişken ve verinin eğitim ile test verisi olarak ayrılması için 2 farklı model kullanılmıştır. Yalın YSA ve DVM’ler, model performanslarına göre karşılaştırıldığında hem Model-1, hem de Model-2 için DVM’lerden daha tatmin edici tahmin sonuçları elde edilmiştir.

Hibrit yöntemler içinse firma başarısızlığına etki eden değişkenlerin belirlenmesi ve veri setinin boyutunun ve gürültünün azaltılması için TBA ile değişken sayısı 3’e indirgenmiştir. Benzer şekilde karşılaştırma yapabilmek amacıyla bu değişkenler ile de verinin eğitim ve test verisi olarak ayrılması için aynı 2 farklı model kullanılmıştır. Model-1 ve Model-2 sonuçlarına göre hibrit modeller ile yalın modellere göre daha iyi sonuçlar elde edilmiştir. Bu sayede hibrit yöntemlerin tahmin başarısını arttırdığı tespit edilmiştir. Hibrit yöntemlerin model performanslarına göre karşılaştırıldığında ise yine hem Model-1 hem de Model-2 için TBA-DVM hibrit yöntemi ile daha tatmin edici tahmin sonuçları elde edilmiştir.