

VERİ MADENCİLİĞİ SÜRECİ KULLANILARAK GÜNEŞ IŞINIMI TAHMİNİ

Özlem TERZİ*, Ecir Uğur KÜÇÜKSİLLE, Gülşah ERGİN, Ahmet İLKER

Özet

Yeryüzünde yatay bir düzlem üzerine gelen güneş ışınımının şiddeti, güneşin mevsimlere ve saate göre değişen konumuyla değişiklik göstermektedir. Güneş ışınımının dünya üzerinde, su sıcaklığının yükselmesi ve doğal olaylar için enerji sağlanması gibi etki alanları bulunmaktadır. Güneş enerjisinden mümkün olduğunca fayda sağlayabilmek ve zararlarından sakınabilmek için güneş ışınımı üzerinde yapılan çalışmaların önemi giderek artmaktadır. Bu çalışmada, 2009–2010 yıllarına ait veriler kullanılarak veri madenciliği süreci yardımı ile güneş ışınımı modellenmiştir. Ölçülen değerler ile model sonuçları kıyaslandığında Multilayer Perceptron algoritması ile kurulan üç girdili (P-Ta-Rh) modelin en uygun sonucu verdiği görülmüştür.

Anahtar kelimeler: veri madenciliği süreci, güneş ışınımı, Isparta

ESTIMATION OF SOLAR RADIATION USING DATA MINING PROCESS

Abstract

Intensity of solar radiation on a horizontal plane of the earth varies with the position depending hour and season of the sun. The impact fields of solar radiation over the world are found such as the rise of water temperature and the provision of energy for natural events. It has an increasing importance of works on the solar radiation to benefit and avoid damages from solar energy. In this study, the data mining process is used to estimate solar radiation. Data for 2009-2010 years are modeled with KStar, Linear Regression, RBF Network, Simple Linear Regression, M5Rules, Decision Table, Random Subspace and Multilayer Perceptron algorithms. Comparing model results with measured values, it is shown that the model (P-Ta-Rh) with three input parameters developed using multilayer perceptron algorithm is the most appropriate model.

Key words: data mining process, solar radiation, Isparta.

1. Giriş

Dünyanın en önemli enerji kaynağı güneştir. Güneşin ışınım enerjisi, yer ve atmosfer sistemindeki fiziksel oluşumları etkileyen başlıca enerji kaynağıdır. Dünyadaki madde ve enerji akışları güneş enerjisi sayesinde mümkün olabilmektedir. Rüzgar, deniz dalgası, okyanusta sıcaklık farkı ve biyokütle enerjileri, güneş enerjisinin değişim geçirmiş biçimleridir. Güneş enerjisi, doğadaki su döngüsünün gerçekleşmesinde de rol oynayarak, akarsu gücünü yaratmaktadır. Doğal enerji kaynaklarının pek çoğunun kökeni olan güneş enerjisinden, ısıtma ve elektrik elde etme gibi amaçlarla doğrudan yararlanılmaktadır. Güneş ışınımı diğer enerji türleri vasıtasıyla da doğrudan olmasa bile faydalı enerji haline dönüştürülebilir. Güneş ışınımı enerjisi atmosferdeki hava olaylarının meydana gelmesine

* Süleyman Demirel Üniversitesi, Teknik Eğitim Fakültesi, Yapı Eğitimi Bölümü, 32260 Isparta/ TÜRKİYE, E-posta:ogencer@tef.sdu.edu.tr

sebepler olur. Yeryüzüne ulaşan güneş ışınımının büyük bir kısmı serbest su yüzeyleri (akarsular, göller, deniz ve okyanuslar) tarafından yutularak ısı olarak depolanır [1].

Güneş ışınımı, fotovoltaik, güneş termal sistemleri ve pasif güneş tasarımı gibi güneş enerjisi uygulamalarının temel bir girdisidir. Herhangi bir coğrafi konumda güneş teknolojilerinin tasarımı, optimizasyonu ve performans değerlendirmesi için veriler güncel, güvenilir ve kolay ulaşılabilir olmalıdır. Pahalı ölçüm cihazı ve teknikleri gerektirdiğinden güneş ışınımı ölçümleri gelişmekte olan ülkelerin çoğunda kolaylıkla elde edilemez. Bu nedenle, hazır meteorolojik veriler ile güneş ışınımı tahmin yöntemleri geliştirmek gereklidir [2].

Rivington vd. (2005) güneş ışınımı tahmin metodlarından hava sıcaklığına dayalı iki ve güneşlenme süresi dönüşümüne dayalı bir modelin performans ve karakteristik davranışını İngiltere’de 24 istasyon için değerlendirmişlerdir. Bir yıl için bulanık mantık tabanlı çoklu değerlendirme sistemi ve ortalama hataların zamansal dağılım testlerini kullanarak karşılaştırma yapmışlardır. Güneşlenme süresinden güneş ışınımına dönüşümün en iyi tahminler ürettiğini fakat sistematik mevsimsel hatalar gösterdiğini belirtmişlerdir. Hava sıcaklığı verileri mevcut olduğu durumlarda, hava sıcaklığına dayalı iki yöntemin güvenilir alternatifler olabileceğini belirtmişlerdir [3]. Şen (1998) güneşlenme süresi ölçümlerinden güneş ışınımını tahmin etmek için bulanık mantık yaklaşımını kullanmıştır. Bulanık mantık yaklaşımını Türkiye’nin batısında üç yere uygulamıştır. Bu uygulamanın dünyanın her yerinde herhangi bir ışıma ya da güneşlenme süresi ölçümleri için basit olarak uygulanabileceğini ifade etmiştir [4]. Şen ve Tan (2001) Türkiye’nin kuzeybatı kısmı için yaptıkları kapsamlı saatlik, günlük ve aylık ölçülen güneş ışınımı istatistiksel analizinin, istatistiksel modellere öncülük ettiğini belirtmişlerdir. Bu modellerin aynı parametrelerinin yeniden değerlendirilmesi ile dünyanın her yerinde geçerli olduğunu ifade etmişlerdir. Modellerin üç parametre ile parabolik denklemler şeklinde görüldüğünü ve parametreler arasında önemli bir ilişki olduğunu belirtmişlerdir. Bu denklemler kullanılarak aylık ortalama saatlik küresel veya yaygın ışınım miktarının tahmin edilmesinin mümkün olduğunu ifade etmişlerdir [5]. Wong ve Chow (2001) Hong Kong için günlük ortalama küresel ışınımı tahmin etmişlerdir. Aylık ortalama saatlik küresel ışınım tahminlerini tartıştıklarını belirtmişlerdir. Hong Kong için iki parametrik modeli incelemişler, model tahminleri ile ölçülen veriler arasında karşılaştırma yapmışlardır [6]. Güneş ışınımının belirlenmesi için benzer metodlar çeşitli araştırmacılar tarafından kullanılmıştır [7-9].

Bu çalışmanın amacı, hava sıcaklığı (T_a), nispi nem (R_h), rüzgâr hızı (W_s) ve basınç (P) parametreleri kullanılarak veri madenciliği süreci ile güneş ışınımı tahmin modelleri geliştirmek ve modellerin geçerliliğini değerlendirmektir.

2. Veri Madenciliği Süreci

Veri madenciliği büyük ölçekli veriler arasından bilgiye ulaşma, bilgiyi madenleme işidir. Ya da bir anlamda büyük veri yığınları içerisinde gelecekle ilgili tahminde bulunabilmemizi sağlayabilecek bağıntıların bilgisayar programı kullanılarak aranmasıdır. Veri madenciliği, eldeki verilerden üstü kapalı, çok net olmayan, önceden bilinmeyen ancak potansiyel olarak kullanışlı bilginin çıkarılmasıdır. Bu da; kümeleme, veri özetleme, değişikliklerin analizi, sapmaların tespiti gibi belirli sayıda teknik yaklaşımları içerir. Başka bir deyişle, veri madenciliği, verilerin içerisindeki desenlerin, ilişkilerin, değişimlerin, düzensizliklerin, kuralların ve istatistiksel olarak önemli olan yapıların yarı otomatik olarak keşfedilmesidir.

Temel olarak veri madenciliği, veri setleri arasındaki desenlerin ya da düzenin, verinin analizi ve yazılım tekniklerinin kullanılmasıyla ilgilidir. Veriler arasındaki ilişkiyi, kuralları ve

özellikleri belirlemekten bilgisayar sorumludur. Amaç, daha önceden fark edilmemiş veri desenlerini tespit edebilmektir.

Veri madenciliğini istatistiksel bir yöntemler serisi olarak görmek mümkün olabilir. Ancak veri madenciliği, geleneksel istatistikten birkaç yönde farklılık gösterir. Veri madenciliğinde amaç, kolaylıkla mantıksal kurallara ya da görsel sunumlara çevrilebilecek nitel modellerin çıkarılmasıdır. Bu bağlamda, veri madenciliği insan merkezlidir ve bazen insan – bilgisayar ara yüzü birleştirilir. Veri madenciliği sahası, istatistik, makine bilgisi, veritabanları ve yüksek performanslı işlem gibi temelleri de içerir [10].

Veri Madenciliği bir yöntem değildir bir süreçtir. Bu süreçte ana unsur süreci gerçekleştiren uygulamacıdır. Süreçte bulunan adımlar doğru olarak yerine getirilmediği sürece istenilen sonuca ulaşılması mümkün değildir. Aşağıda veri madenciliği süreci adımları açıklanmıştır.

2.1. Sorunun Tanımlanması

Veri madenciliği çalışmalarında başarılı olmanın ilk şartı, uygulamanın hangi amaçla yapılacağına açık bir şekilde tanımlanmasıdır. Amaç, sorun üzerine odaklanmış ve açık bir dille ifade edilmiş olmalı, elde edilecek sonuçların başarı düzeylerinin nasıl ölçüleceği tanımlanmalıdır. Sorun ile tam örtüşmeyen bir veri madenciliği çalışması, sorunu çözmeye yetmeyeceği gibi sonuçta başka problemlerin de ortaya çıkmasına neden olabilecektir. Ayrıca yanlış kararlarda katlanılacak olan zararlar ve doğru kararlarda kazanılacak faydalara ilişkin öngörülere de bu aşamada yer verilmelidir.

2.2. Verilerin Hazırlanması

Modelin kurulması aşamasında ortaya çıkacak sorunlar, bu aşamaya sık sık geri dönülmesine ve verilerin yeniden düzenlenmesine neden olacaktır. Bu durum verilerin hazırlanması ve modelin kurulması aşamaları için, bir karar vericinin veri keşfi sürecinin toplamı içerisindeki enerji ve zamanının % 50 - % 85' ini harcamasına neden olmaktadır.

2.3. Toplama ve Uyumlaştırma

Tanımlanan sorun için gerekli olduğu düşünülen verilerin ve bu verilerin toplanacağı veri kaynaklarının belirlenmesi adımdır. Hangi veri kaynaklarından yararlanılacağı önemli bir karardır. Çünkü gereğinden az veri kaynağı veri madenciliği çalışmasını eksik bırakacağı gibi, gereğinden fazla veri kaynağı sürecin uzamasına neden olabilecek veri kirliliğine yol açabilecektir.

Veri madenciliğinde kullanılacak verilerin farklı kaynaklardan toplanması, doğal olarak veri uyumsuzluklarına neden olacaktır. Bu uyumsuzluklardan bazıları; verilerin farklı zamanlara ait olması, güncelleme hatası, veri formatlarının farklı olması, kodlama farklılığı (örneğin bir veri tabanında cinsiyet özelliğinin e/k, diğer bir veri tabanında 0/1 olarak kodlanması), farklı ölçü birimleri ve varsayım farklılıklarıdır. Ayrıca verilerin nasıl, nerede ve hangi koşullar altında toplandığı da önem taşımaktadır. Güvenilir olmayan veri kaynaklarının kullanımı tüm veri madenciliği sürecinin de güvenilirliğini etkileyecektir.

Bu nedenlerle, iyi sonuç alınacak veri madenciliği çalışmaları ancak iyi verilerin üzerine kurulabileceği için, toplanan verilerin ne ölçüde uyumlu oldukları bu adımda incelenerek değerlendirilmelidir.

2.4. Birleştirme ve Temizleme

Bu adımda farklı kaynaklardan toplanan verilerde bulunan ve bir önceki adımda belirlenen sorun ve uyumsuzluklar mümkün olduğu ölçüde giderilerek, veriler tek bir veri tabanında toplanır. Ancak basit yöntemlerle ve baştan savma olarak yapılacak sorun giderme işlemlerinin, sonraki aşamalarda daha büyük sorunların kaynağı olacağı unutulmamalıdır.

2.5. Seçim

Bu adımda kurulacak modele bağlı olarak veri seçimi yapılır. Örneğin tahmin edici bir model için bu adım, bağımlı ve bağımsız değişkenlerin ve modelde kullanılacak veri kümesinin seçilmesi anlamını taşımaktadır.

Sıra numarası gibi anlamlı olmayan değişkenlerin modele girmemesi gerekmektedir. Çünkü bu tip değişkenler, diğer değişkenlerin modeldeki ağırlığının azalmasına ve veriye ulaşma zamanlarının uzamasına neden olabilmektedir. Bazı veri madenciliği algoritmaları konu ile ilgisi olmayan bu tip değişkenleri otomatik olarak elese de, pratikte bu işlemin kullanılan yazılıma bırakılmaması daha akılcı olacaktır.

Verilerin görselleştirilmesine olanak sağlayan grafik araçlar ve bunların sunduğu ilişkiler, bağımsız değişkenlerin seçilmesinde önemli yararlar sağlayabilir. Genellikle yanlış veri girişinden veya bir kereye özgü bir olayın gerçekleşmesinden kaynaklanan verilerin, veri kümesinden atılması tercih edilir.

Veri madenciliği çalışmasında geliştirilen modelde kullanılan veri tabanının çok büyük olması durumunda, rastgeleliği bozmayacak şekilde örnekleme yapılması uygun olabilir. Ayrıca burada seçilen örneklem kümesinin tüm popülasyonu temsil edip etmediği de kontrol edilmelidir. Halen kullanılan işletim sistemleri ve paket programlar ne kadar gelişmiş olursa olsun, çok büyük veri tabanları üzerinde çok sayıda modelin denenmesi zaman azlığı nedeni ile mümkün olamamaktadır. Bu nedenle tüm veri tabanını kullanarak bir kaç model denemek yerine, rastgele örneklenmiş bir veri tabanı parçası üzerinde birçok modelin denenmesi ve bunlar arasından en güvenilir ve güçlü modelin seçilmesi daha uygun olacaktır. Diğer bir deyişle modellerin performansları uygun bir karar yöntemi ile sınanmalıdır.

2.6. Modelin Kurulması ve Değerlendirilmesi

Tanımlanan problem için en uygun modelin bulunabilmesi, olabildiğince çok sayıda modelin kurularak denenmesi ile mümkündür. Bu nedenle veri hazırlama ve model kurma aşamaları, en iyi olduğu düşünülen modele varılıncaya kadar yinelenen bir süreçtir.

Model kuruluş süreci, denetimli ve denetimsiz öğrenmenin kullanıldığı modellere göre farklılık göstermektedir.

Örnekten öğrenme olarak da isimlendirilen denetimli öğrenmede, bir denetçi tarafından ilgili sınıflar önceden belirlenen bir kritere göre ayrılarak, her sınıf için çeşitli örnekler verilir. Sistemin amacı verilen örneklerden hareket ederek her bir sınıfa ilişkin özelliklerin bulunması ve bu özelliklerin kural cümleleri ile ifade edilmesidir.

Öğrenme süreci tamamlandığında, tanımlanan kural cümleleri verilen yeni örneklere uygulanır ve yeni örneklerin hangi sınıfa ait olduğu kurulan model tarafından belirlenir.

Denetimsiz öğrenmede, kümeleme analizinde olduğu gibi ilgili örneklerin gözlenmesi ve bu örneklerin özellikleri arasındaki benzerliklerden hareket ederek sınıfların tanımlanması amaçlanmaktadır.

Denetimli öğrenmede seçilen algoritmaya uygun olarak ilgili veriler hazırlandıktan sonra, ilk aşamada verinin bir kısmı modelin öğrenilmesi, diğer kısmı ise modelin geçerliliğinin test edilmesi için ayrılır. Modelin öğrenilmesi, öğrenim kümesi kullanılarak gerçekleştirildikten sonra, test kümesi ile modelin doğruluk derecesi belirlenir.

Bir modelin doğruluğunun test edilmesinde kullanılan en basit algoritma basit geçerlilik testidir. Bu yöntemde tipik olarak verilerin % 5 ile % 33 arasındaki bir kısmı test verileri olarak ayrılır ve kalan kısım üzerinde modelin öğrenimi gerçekleştirildikten sonra, bu veriler üzerinde test işlemi yapılır. Bir sınıflama modelinde yanlış olarak sınıflanan olay sayısının, tüm olay sayısına bölünmesi ile hata oranı, doğru olarak sınıflanan olay sayısının tüm olay sayısına bölünmesi ile ise doğruluk oranı hesaplanır. ($Doğruluk\ Oranı = 1 - Hata\ Oranı$)

Sınırlı miktarda veriye sahip olunması durumunda, kullanılacak diğer bir algoritma, çapraz geçerlilik testidir. Bu yöntemde veri kümesi rastgele iki eşit parçaya ayrılır. İlk aşamada bir parça üzerinde model eğitimi ve diğer parça üzerinde test işlemi; ikinci aşamada ise ikinci parça üzerinde model eğitimi ve birinci parça üzerinde test işlemi yapılarak elde edilen hata oranlarının ortalaması kullanılır.

Bir kaç bin veya daha az satırdan meydana gelen küçük veri tabanlarında, verilerin n gruba ayrıldığı n katlı çapraz geçerlilik testi tercih edilebilir. Verilerin örneğin 10 gruba ayrıldığı bu yöntemde, ilk aşamada birinci grup test, diğer gruplar öğrenim için kullanılır. Bu süreç her defasında bir grubun test, diğer grupların öğrenim amaçlı kullanılması ile sürdürülür. Sonuçta elde edilen hata oranının ortalaması, kurulan modelin tahmini hata oranı olacaktır.

Bootstrapping küçük veri kümeleri için modelin hata düzeyinin tahmininde kullanılan bir başka tekniktir. Çapraz geçerlilikte olduğu gibi model bütün veri kümesi üzerine kurulur. Daha sonra en az 200, bazen binin üzerinde olmak üzere çok fazla sayıda öğrenim kümesi tekrarlı örneklemelerle veri kümesinden oluşturularak hata oranı hesaplanır.

Model kuruluşu çalışmalarının sonucuna bağlı olarak, aynı teknikle farklı parametrelerin kullanıldığı veya başka algoritma ve araçların denendiği değişik modeller kurulabilir. Model kuruluş çalışmalarına başlamadan önce, hangi tekniğin en uygun olduğuna karar verebilmek güçtür. Bu nedenle farklı modeller kurularak, doğruluk derecelerine göre en uygun modeli bulmak üzere sayısız deneme yapılmasında yarar bulunmaktadır.

Kurulan modelin doğruluk derecesi ne denli yüksek olursa olsun, gerçek dünyayı tam anlamı ile modellediğini garanti edebilmek mümkün değildir. Yapılan testler sonucunda geçerli bir modelin doğru olmamasındaki başlıca nedenler, model kuruluşunda kabul edilen varsayımlar ve modelde kullanılan verilerin doğru olmamasıdır.

2.7. Modelin Kullanılması

Kurulan ve geçerliliği kabul edilen model doğrudan bir uygulama olabileceği gibi, bir başka uygulamanın alt parçası olarak kullanılabilir.

2.8. Modelin İzlenmesi

Zaman içerisinde bütün sistemlerin özelliklerinde ve dolayısıyla ürettikleri verilerde ortaya çıkan değişiklikler, kurulan modellerin sürekli olarak izlenmesini ve gerekiyorsa yeniden düzenlenmesini gerektirecektir. Tahmin edilen ve gözlenen değişkenler arasındaki farklılığı gösteren grafikler model sonuçlarının izlenmesinde kullanılan yararlı bir yöntemdir [11].

3. Uygulama

3.1. Sorunun Tanımlanması

Çalışmanın amacı, veri madenciliği süreci kullanarak güneş ışınımı tahmin modelleri geliştirmek ve modellerin geçerliliğinin araştırılmasıdır. Kurulan modellerinin yeterliliği determinasyon katsayısı (R^2) ve karekök ortalama hata (KOH) değerlerine göre belirlenecektir.

3.2. Verilerin Hazırlanması

Bu çalışmada güneş ışınımı tahmini için veri madenciliği süreci kullanılarak geliştirilen modellerde kullanılan meteorolojik veriler 23.05.2009–18.05.2010 tarihleri arasında yarım saatte bir ölçülmüştür. Verilerin günlük ortalamaları alınarak 355 veri elde edilmiştir ve %80'i eğitim %20'si ise test seti olarak ayrılmıştır.

3.3. Toplama ve Uyumlaştırma

Çalışmada kullanılan 355 verinin tamamı tek bir kaynaktan (Süleyman Demirel Üniversitesi'nde kurulan otomatik meteoroloji istasyonundan) elde edildiği için herhangi bir sorun ya da uyumsuz bir durumla karşılaşılmadığı için bu adımda hiçbir işlem yapılmamıştır. Bu istasyon aracılığı ile hava sıcaklığı, basınç, rüzgâr hızı, nisbi nem, güneş ışınımı gibi meteorolojik parametreler yarım saatlik aralıklarla ölçülmektedir.

3.4. Modelin Kurulması ve Değerlendirilmesi

Veri madenciliği süreci ile model geliştirebilmek için günümüzde çeşitli bilgisayar yazılımları kullanılmaktadır. Çalışmada bu yazılımlardan biri olan WEKA programının 3.6 sürümü kullanılmıştır.

Güneş ışınımı tahmini için veri madenciliği süreci kullanılarak Kstar, Linear regression, RBF network, Simple linear regression, M5'Rules, Decision Table, Random Subspace ve Multilayer Perceptron algoritmalarıyla bir, iki, üç ve dört girdili modeller geliştirilmiştir. Geliştirilen modellerde güneş ışınımı tahmini için kullanılan girdi parametreleri güneş ışınımı üzerindeki etki sırasına göre basınç (P), hava sıcaklığı (Ta), rüzgâr hızı (Ws) ve nispi nem (Rh) şeklinde belirlenmiştir. Modellerin yeterliliği, determinasyon katsayısı ve karekök ortalama hata değerlerine göre belirlenmiştir.

Determinasyon katsayısı (R^2) ve karekök ortalama hata (KOH) hesaplanırken sırasıyla aşağıda verilen denklemler kullanılmıştır.

$$R^2 = 1 - \frac{\sum_{i=1}^n (R_{i(\text{gerçek})} - R_{i(\text{model})})^2}{\sum_{i=1}^n (R_{i(\text{gerçek})} - R_{\text{ort}})^2} \quad [1]$$

$$KOH = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_{i(\text{gerçek})} - R_{i(\text{model})})^2} \quad [2]$$

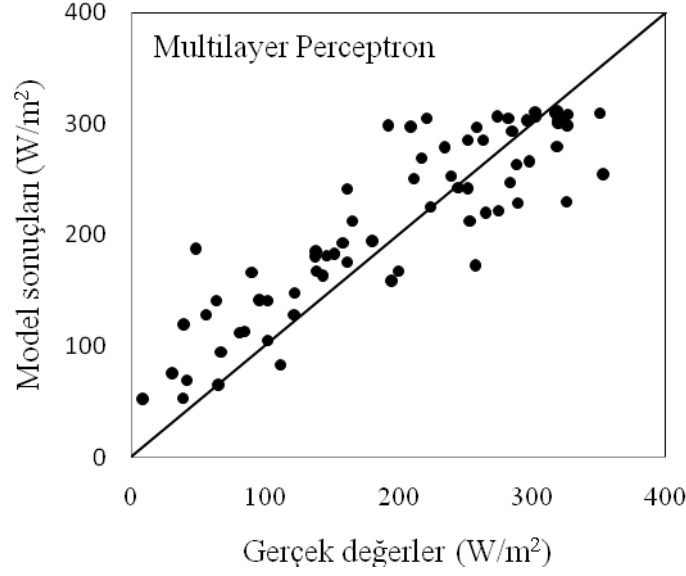
Burada, n toplam veri sayısı, $R_{i(\text{gerçek})}$ ve $R_{i(\text{model})}$ sırasıyla ölçülen güneş ışınımı değeri ve geliştirilen modelin sonuçlarıdır. R_{ort} ise ortalama güneş ışınımı değeridir.

Geliştirilen modellerin test sonuçlarına ait determinasyon katsayısı (R^2) ve karekök ortalama hata (KOH) değerleri Çizelge 1’de verilmiştir.

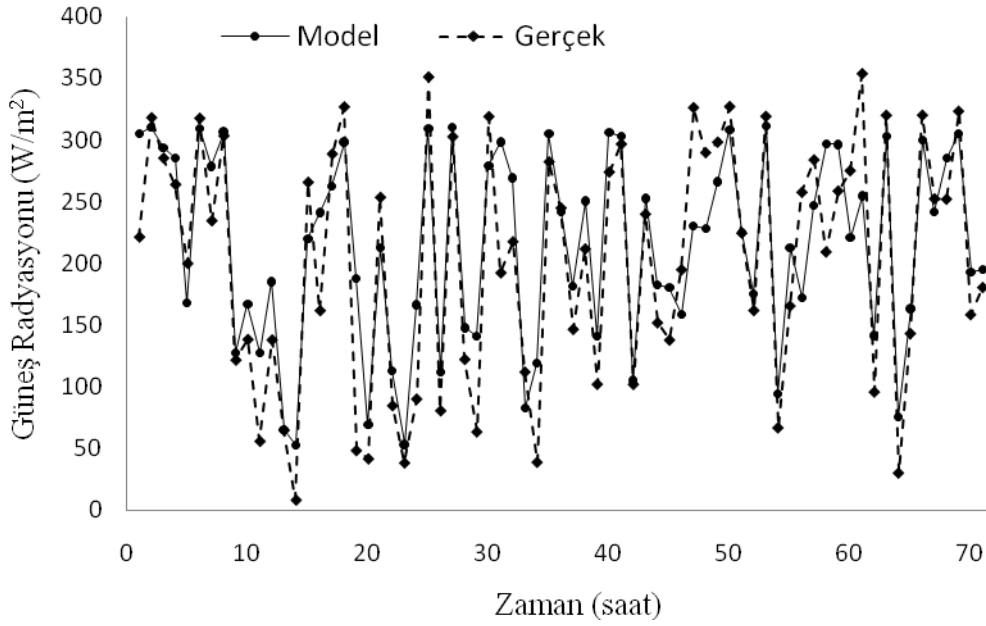
Çizelge 1. Geliştirilen modellerin test setine ait R^2 ve KOH değerleri

Algoritma	KStar	Linear Regression	RBF Network	Simple Linear Regression	M5’Rules	Decision Table	Random Subspace	Multilayer Perceptron
Model	R^2 -KOH	R^2 -KOH	R^2 -KOH	R^2 -KOH	R^2 -KOH	R^2 -KOH	R^2 -KOH	R^2 -KOH
P	0,637-58,35	0,572-63,36	0,567-63,75	0,572-63,36	0,572-63,36	0,597-61,49	0,604-60,98	0,561-64,18
Ta	0,614-60,18	0,612-60,34	0,542-65,57	0,612-60,34	0,612-60,34	0,607-60,74	0,581-62,76	0,509-67,91
Ws	0,410-74,46	0,242-84,34	0,108- 91,53	0,242-84,34	0,408-74,55	0,392- 75,54	0,381- 76,25	0,429- 73,21
Rh	0,111-91,34	0,121- 90,83	0,129- 90,45	0,121-90,83	0,077- 93,10	0,070- 97,31	-0,001-96,93	-0,011- 97,43
P-Ta	0,728- 50,58	0,695- 53,52	0,563- 64,05	0,612- 60,34	0,691- 53,90	0,686-54,27	0,666-55,97	0,679-54,94
Ta-Rh	0,727-50,63	0,648-57,53	0,503- 68,33	0,612-60,34	0,648-57,53	0,607-60,74	0,574-63,21	0,470-70,56
P-Rh	0,698- 53,24	0,604-60,95	0,566-63,86	0,572-63,36	0,693-53,66	0,603-61,07	0,516- 67,41	0,612-60,32
P-Ws	0,687- 54,25	0,621- 59,64	0,577- 63,01	0,572-63,36	0,644-57,80	0,607-60,76	0,616- 60,07	0,581-62,75
Ta-Ws	0,567-63,79	0,658-56,65	0,569-63,59	0,612-60,34	0,658- 56,65	0,607- 60,74	0,567-63,79	0,418-73,91
Ws-Rh	0,420-73,80	0,301-81,02	0,133- 90,23	0,242-84,34	0,409- 74,49	0,346-78,37	0,225-85,30	0,410-74,45
P-Ta-Rh	0,739-49,46	0,732-50,18	0,563-64,04	0,612-60,34	0,742- 49,18	0,686-54,27	0,741-49,32	0,764-47,08
P-Ws-Rh	0,707-52,46	0,636-58,50	0,578-62,93	0,572-63,36	0,736-49,76	0,603-61,07	0,668-55,87	0,565-63,92
Ta-Ws-Rh	0,736-49,76	0,682-54,65	0,533-66,24	0,612-60,34	0,736-49,76	0,607-60,74	0,706-52,52	0,588-62,18
P-Ta-Ws	0,704-52,75	0,721-51,14	0,567-63,80	0,612-60,34	0,697-53,30	0,686-54,27	0,705-52,63	0,606-60,84
P-Ta-Ws-Rh	0,744-49,04	0,748-48,65	0,567-63,77	0,612-60,34	0,748-48,65	0,686-54,27	0,750-48,49	0,584-62,48

Bu modeller arasında en iyi R^2 değerini veren Multilayer Perceptron algoritmasıyla geliştirilen 3 girdili (P-Ta-Rh) model olmuştur. Multilayer Perceptron algoritmasıyla geliştirilen modelin performansını değerlendirmek için model sonuçları ile ölçülen güneş ışınımı değerleri arasındaki ilişki Şekil 1’de verilen saçılma diyagramı ve Şekil 2’de verilen zaman serisi ile gösterilmiştir. Geliştirilen modelin güneş ışınımı tahmini için kullanılabileceği sonucuna varılmıştır.



Şekil 1. Model sonucu ile ölçülen güneş ışıını değerlerine ait saçılma diyagramı



Şekil 2. Model sonucu ile ölçülen günlük güneş ışıını için zaman serisi

4. Sonuçlar

Çalışmada, güneş ışıını tahmini için veri madenciliği süreci kullanılarak farklı girdi kombinasyonlarıyla modeller geliştirilmiştir. Geliştirilen modellere ait sonuçlar ile ölçülen gerçek değerler karşılaştırıldığında, Multilayer Perceptron algoritmasıyla geliştirilen 3 girdili (P-Ta-Rh) modelin diğer modellere kıyasla ölçülmüş güneş ışıını değerleri ile daha uyumlu olduğu ve Random Subspace algoritmasıyla geliştirilen 4 girdili modelin de Multilayer Perceptron algoritmasıyla geliştirilen 3 girdili (P-Ta-Rh) modele yakın bir sonuç verdiği tespit edilmiştir. Sonuç olarak geliştirilen modellerin veri kaybı ya da güneş ışıını ile ilgili öngöründe bulunma gibi durumlarda kullanılabilir olduğu görülmüştür. Zaman içerisinde

ölçülen yeni verilerin güneş ışınımı ile ilişkisi kurulan modellerle yeniden hesaplanıp nasıl sonuçlar vereceği gözlemlenebilir.

Kaynaklar

1. Varınca, K.B., Varank, G., 2005. Güneş kaynaklı farklı enerji üretim sistemlerinde çevresel etkilerin kıyaslanması ve çözüm önerileri. Güneş Enerjisi Sistemleri Sempozyumu ve Sergisi, 148-160, 24-25 Haziran 2005, İçel.
2. Bulut, H., Büyükalaca, O., 2007. Simple model for the generation of daily global solar-radiation data in Turkey. Applied Energy, 84, 477 – 491.
3. Rivington, M., Bellocchi, G., Matthews, K.B., Buchan, K., 2005. Evaluation of Three Model Estimations of Solar Radiation at 24 UK Stations. Agricultural and Forest Meteorology, 132, 228–243.
4. Şen, Z., 1998. Fuzzy Algorithm for Estimation of Solar Irradiation. Solar Energy, 63 (1), 39 – 49.
5. Şen, Z., Tan, E., 2001. Simple models of solar radiation data for northwestern part of Turkey. Energy Conversion and Management, 42, 587 – 598.
6. Wong, L.T., Chow, W.K., 2001. Solar radiation model. Applied Energy, 69, 191 – 224.
7. Dorvlo, A.S.S., Jervase, J.A., Al-Lawati, A., 2002. Solar Radiation Estimation Using Artificial Neural Networks. Applied Energy, 71, 307-319.
8. Sfetsos, A., Coonick, A.H. 2000. Univariate and Multivariate Forecasting of Hourly Solar Radiation With Artificial Intelligence Techniques. Solar Energy, 68(2), 169-178.
9. Sözen, A., Arcaklıoğlu, E., Özalp, M., 2004. Estimation of Solar Potential in Turkey by Artificial Neural Networks Using Meteorological and Geographical Data. Energy Conversion and Management, 45, 3033-3052.
10. http://tr.wikipedia.org/wiki/Veri_madencili%C4%9Fi. (Erişim tarihi: 29.11.2010).
11. Alpaydın, E.2000. Zeki veri madenciliği: ham veriden altın bilgiye ulaşma yöntemleri.Bilişim2000eğitimsemineri.www.deu.edu.tr/userweb/k.yaralioglu/dosyalar/ver_mad.doc (Erişim tarihi: 29.11.2010).