Journal of the Faculty of Engineering and Architecture of Gazi University 32:3 (2017) 901-907

Mühendislik Mimarlık Fakültesi Dergisi
Journal of The Faculty of Engineering and Architecture of Gazi University

Elektronik / Online ISSN : 1304 - 4915
Basılı / Printed ISSN : 1300 - 1884

Destek vektör regresyonu yöntemiyle karbon nanotüp benzetim süresinin kısaltılması

Mehmet Acı[1*], Mutlu Avcı[2], Çiğdem Acı[1]

[1]Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 33343, Mersin, Türkiye
[2]Çukurova Üniversitesi, Mühendislik-Mimarlık Fakültesi, Biyomedikal Mühendisliği Bölümü, 01330, Adana, Türkiye

Ö N E Ç I K A N L A R

- Destek vektör regresyonu kullanılan tahmin modeli
- Yoğunluk fonksiyonel teorisi hesaplama süresini azaltmak
- Atomik koordinatların yüksek doğrulukla tahmini

Makale Bilgileri
Geliş: 02.06.2016
Kabul: 26.10.2016

DOI:
10.17341/gazimmfd.337642

Anahtar Kelimeler:
Atomik koordinat,
karbon nanotüp,
tahmin,
destek vektör regresyonu

ÖZET

Yoğunluk Fonksiyonel Teorisi'nin (YFT) en önemli uygulamalarından bir tanesi Karbon Nanotüplerdir (KNT). Karbonun kimyasal ve fiziksel karakteristiği sebebiyle KNT benzetimleri nanoteknoloji alanında büyük rol oynamaktadır. YFT hesaplarının saatler hatta günler sürmesi KNT benzetimlerinin en zor yanını oluşturmaktadır. Bu çalışmada KNT'lerin atomik koordinatlarını hesaplayan benzetimlerin süresini kısaltacak bir Destek Vektör Regresyonu (DVR) modeli önerilmiştir. KNT benzetimlerinden elde edilen u, v ve w koordinatları DVR yöntemi kullanılarak yüksek doğrulukta ve dakikalar içinde tahmin edilmiştir. Bu tahmin modeli için CASTEP yazılımı kullanılarak 10721 örnek içeren bir veri kümesi oluşturulmuştur. Veri kümesi atomik koordinatlar ve kiral vektörlerden oluşmaktadır. Önerilen modelin doğruluğunu değerlendirmek için Ortalama Karesel Hata (OKH), Ortalama Mutlak Hata (OMH), Standart Tahmin Hatası (STH) ve Korelasyon Katsayısı (R) ölçütleri kullanılmıştır. Veri kümesi, 10-katlı çapraz doğrulama kullanılarak ve kullanılmadan ayrı ayrı çalışılmıştır. Bu çalışmadan elde edilen sonuçlar iki şekilde kullanılabilir: 1) Atomik koordinatlar hiçbir benzetim programı kullanmadan düşük bir hata oranı ile tahmin edilebilir, 2) Tahmin edilen koordinatlar benzetim programlarına başlangıç değeri olarak verilip benzetimin süresi ciddi oranda kısaltılabilir..

Reducing simulation duration of carbon nanotube using support vector regression method

H I G H L I G H T S

- Prediction model using support vector regression
- Reducing calculation time of density functional theory
- High accuracy estimation of atomic coordinates

Article Info
Received: 02.06.2016
Accepted: 26.10.2016

DOI:
10.17341/gazimmfd.337642

Keywords:
Atomic coordinate,
carbon nanotube,
prediction,
support vector regression

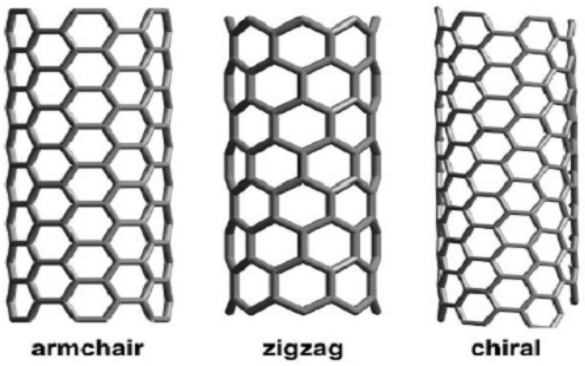
ABSTRACT

Density Functional Theory (DFT) is one of the most important application of Carbon Nanotubes (CNTs). Because of the chemical and physical characteristics of carbon, CNTs play an important role in the field of nanotechnology. The most difficult part of CNT simulations is DFT calculations that take hours or even days. In this study, a Support Vector Regression (SVR) model that reduces the atomic coordinate calculation of CNT simulation duration has been proposed. u, v, and w coordinates which obtained from CNT simulations are predicted with high accuracy using the SVR method within minutes. A dataset containing 10721 samples was created using CASTEP software for the prediction model. The dataset consists of the atomic coordinates and chiral vectors. To evaluate the accuracy of the proposed model, Mean Square Error (MSE), Mean Absolute Error (MAE), Standard Error of the Estimate (SEE) and Correlation Coefficient (R) metrics were used. The dataset is studied separately with and without using 10-fold cross-validation. The results obtained from this study can be used in two ways: 1) The atomic coordinates can be predicted with a low-error without using a simulation program, 2) The estimated results can be used as an initial value of simulation software for reducing duration of the atomic coordinate simulation seriously.

*Sorumlu Yazar/Corresponding Author: maci@mersin.edu.tr / Tel: +90 324 361 00 01 / 17045

1. GİRİŞ (INTRODUCTION)

Yoğunluk Fonksiyonel Teorisi'nin (YFT) temeli, 1927 yılında Thomas ve Fermi tarafından yapılan çalışmaları temel alan Hohenberg ve Kohn [1] teoremleri ve onun devamı olan Kohn ve Sham [2] teoremlerine dayanmaktadır. YFT'de esas olarak etkileşen çok elektronlar sistemlerinin taban durum özelliklerini belirlemek için elektron yoğunluğu temel değişken olarak kabul edilir. YFT, hesaplamalara dayalı yoğun madde fiziği ve malzeme biliminde oldukça yaygın, güncel ve deneylerle uyumlu sonuçlar veren bir yöntemdir ve metaller, yarı iletkenler ve yalıtkanların temel durum özelliklerini belirlemek için oldukça başarılı bir yaklaşımdır. Ayrıca bu yöntemin başarısı, sadece yük hacimli malzemelerle sınırlı olmasından değil aynı zamanda protein ve Karbon Nanotüpler (KNT) gibi karmaşık materyallere de uygulanabilir olmasından kaynaklanmaktadır [3]. Karbonun allotroplarından biri olan grafenin silindir şeklinde sarılması ile KNT'ler elde edilir. KNT'lerin çok küçük boyutta olması ve eşsiz elektronik özelliklerinden dolayı karbon, yarı iletken fiziği üzerinde büyük bir etkiye sahiptir. Kiralite olarak bilinen olası çok çeşitli sarmal geometri nedeniyle KNT'lerin çok değişik çap ve kiralitede elde edilmesi mümkündür. Bir KNT yaklaşık olarak 0,4 nm çapında ve 100 nm uzunluğundadır. Bir tek grafin yüzeyini kıvırarak tek katmanlı KNT elde edilir. Tek katmanlı KNT'lerin çapı genellikle 1 nm olup iki ucu da kapalıdır. Şekil 1'de armchair, zigzag ve chiral tipindeki KNT'ler görülmektedir [4].



Şekil 1. KNT çeşitleri (CNT types)

Bir KNT tasarlamak için öncelikle bu materyalin bilgisayar ortamında benzetiminin yapılması gerekmektedir. Bu benzetimleri yapmak ve materyallerin özelliklerini hesaplamak amacıyla kullanılan Materials Design [5], SCIGRESS [6] ve JCrystalSoft [7] gibi birçok yazılım mevcuttur. CASTEP [8] de bunlardan biridir. KNT benzetimleri, YFT metodu kullanılarak karbon atomlarının 3-boyutlu düzlemdeki koordinatlarının istenen kiraliteye göre belirlenmesi şeklinde yapılır. YFT hesaplamalarının çok uzun sürmesi sebebiyle bir benzetimin sonuncunun alınması saatler hatta günler sürmektedir. Bu süreyi kısaltmanın temelde iki yolu vardır: 1) Hesaplamaları yüksek başarımlı, paralel mimariye sahip bilgisayarlar üzerinde çalıştırmak [9]. 2) Hesaplama sonuçlarını gerçeğe en yakın şekilde tahmin etmek. Birinci yöntem için maliyetleri binlerce doları bulan donanımları edinmek gerekmektedir. İkinci yöntemde ise atomik koordinat değerleri, Makine Öğrenmesi yöntemleriyle eğitilen tahmin modellerinde gerçeğe en yakın şekilde sıradan bilgisayarlar kullanılarak tahmin edilebilmektedir. Bu bağlamda literatürde yapılan çalışmalara bakıldığında, Cheng ve ark. [10] MOSFET hesaplamalarında Yapay Sinir Ağları'nı (YSA) kullanmış ve hesaplama süresini ciddi oranda düşürmüşlerdir. Benzetim programlarının yerine Makine Öğrenmesi yöntemlerinin kullanılması literatürde çalışılan konulardandır: Zayid ve Akay [11] yaptıkları çalışmada çoklu-işlemcili mesaj geçişi mimarisinin başarım ölçütlerini OPNET Modeler [12] yazılımı yerine YSA ve SVR yöntemlerini kullanarak düşük hata oranları ve yüksek korelasyon katsayısı ile tahmin etmişlerdir. Sharma ve ark. [13], biodizel ile çalışan bir motoru YSA kullanarak modellemiş olup motor başarımını ve egzoz emisyonunu yüksek doğrulukta tahmin etmiştir. Kayabaşı ve Akdağlı [14], oyuk yüklü kompakt mikroşerit antenlerin rezonans frekansının hesaplanmasında YSA ve bulanık mantık sitemine dayalı uyarlanır ağ yöntemlerini kullanmış ve diğer metotlara göre daha hızlı ve etkili olduğu sonucunu ortaya çıkarmışlardır. Özocak ve Yurtçu [15], bulanık mantık ve YSA yöntemlerini ince daneli zeminlerin sıkışma indisinin indeks özellikleri vasıtası ile tahmin edilmesinde kullanmışlardır. Fakhim ve ark. [16], çok katmanlı KNT'lerin çimento hidratasyonu üzerindeki etkisini YSA kullanarak incelemişlerdir. Makine Öğrenmesi yöntemlerinin nanofizik hesaplamalarındaki süreye olan etkisi daha önce çalışılmamış bir konudur. Bu çalışmanın ana fikri, KNT'lerin atomik koordinatlarını Destek Vektör Regresyonu (DVR) yöntemi ile gerçeğe en yakın şekilde tahmin etmektir. Elde edilen tahmin sonuçları KNT modellemesinde iki şekilde kullanılabilmektedir: 1) Atomik koordinatlar hiçbir benzetim programı kullanmadan düşük bir hata oranı ile elde edilebilir ve bu haliyle kullanılabilir, 2) Elde edilen koordinatlar benzetim programlarına başlangıç değeri olarak verilip benzetimin süresi ciddi oranda kısaltılabilir. Böylece, Makine Öğrenmesi yöntemleri materyal tasarımında başarı ile kullanılabilir.

2. MATERYAL VE METOT (MATERIAL AND METHOD)

Bu bölümde kullanılan veri kümesi ve kullanılan Makine Öğrenmesi yöntemi ele alınacaktır.

2.1. Veri Kümesi (Dataset)

Bu çalışmada, CASTEP [8] yazılımı kullanılarak modellenen KNT'lerin başlangıç atomik koordinat bilgileri (u, v, w) ve kiral vektörünün (n, m) çeşitli değerleri kullanılarak bir girdi veri kümesi oluşturulmuş ve CASTEP [8] ile hesaplanan gerçek atomik koordinat değerleri (u', v', w') ise çıktı veri kümesi olarak kullanılmıştır. CASTEP benzetiminin ayrıntıları Tablo 1'de verilmiştir. Tablo 2, veri kümesinin istatistiki ayrıntılarını göstermektedir. Girdi kümesi ile çıktı kümesi arasındaki regresyonu

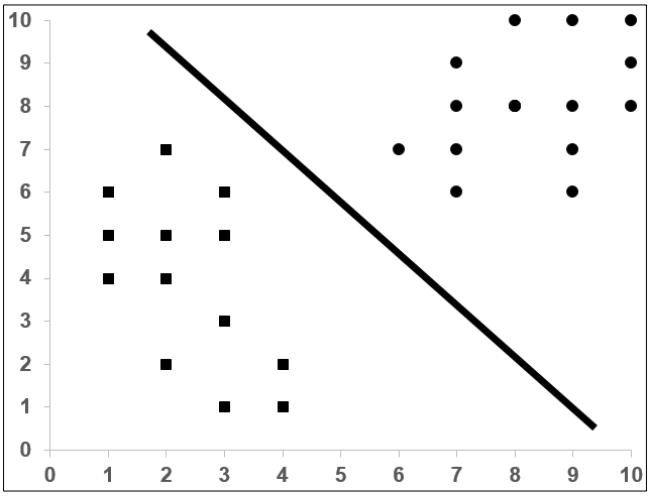
gerçekleştirmek için DVR yöntemi kullanılmıştır. DVR yöntemi, MATLAB [17] yazılımındaki "fitrsvm" fonksiyonu kullanılarak kodlanmış olup çekirdek fonksiyon olarak Radyal Tabanlı Fonksiyon seçilmiştir.

Tablo 1. CASTEP benzetim detayları
(Details of CASTEP simulation)

Özellik	Değer
KNT tipi	Tek katmanlı
Bağ uzunluğu	1,42 Å
Çap	2,07 Å
Koordinatların başlangıç üretimi	Rastgele
Elektriksel Enerji Tolerans Değeri	$1x10^{-5}$ eV

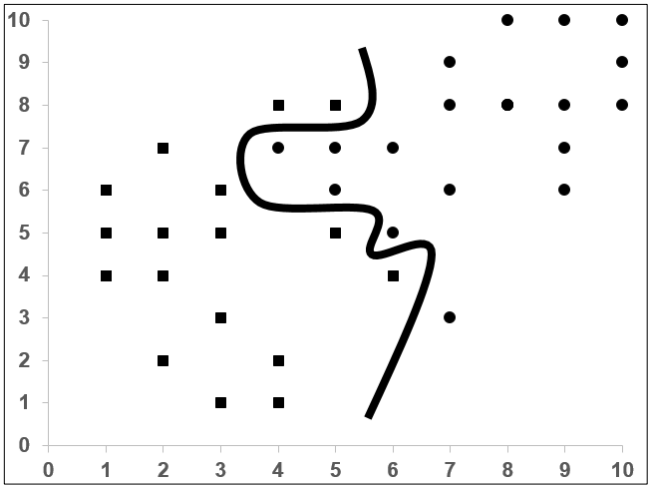
2.2. Destek Vektör Regresyonu (DVR)
(Support Vector Regression) (SVR)

Destek Vektör Makineleri (DVM), Vapnik tarafından geliştirilen eğiticili bir öğrenme yöntemi olup sınıflandırma ve regresyon için kullanılmaktadır [18]. Bu yöntem diğer geleneksel öğrenme yöntemleriyle karşılaştırıldığında doğrusal olmayan problemleri çözmedeki performansı ve yeteneği çok daha iyidir. Yaygın olarak sınıflandırma problemleri için kullanılan DVM'nin regresyon için uyarlaması Smola ve arkadaşları tarafından ileri sürülmüştür [19]. Bu yöntem ise Destek Vektör Regresyonu (DVR) olarak adlandırılmıştır [20]. DVM'lerde karşılaşılabilecek iki durum, verilerin lineer olarak ayrılabilecekleri bir yapıda olması veya lineer olarak ayrılamayan yapıda olmasıdır. Lineer Ayrılabilme Durumu: Eğitim için kullanılacak N elemandan oluşan verinin $\theta=\{x_i, y_i\}$, i = 1, 2,…… N olduğu varsayalım. Burada $y_i \in \{-1, 1\}$ etiket değerleri ve $x_i \in R^d$ özellikler vektörüdür. Lineer olarak ayrılabilme durumunda, bu iki değerli veriler direkt olarak bir aşırı düzlem ile ayrılabilecektir. Bu aşırı düzleme ayırıcı aşırı düzlem adı verilir (Şekil 2). DVM'lerin amacı bu aşırı düzlemin iki ayrı sınıfta bulunan örnek grubuna eş uzaklıkta olmasını sağlamaktır [21].



Şekil 2. Doğrusal olarak ayrılabilen veriler
(Linearly separable data)

Lineer Ayrılamama Durumu: Lineer ayrılabilme durumunda veriler iki sınıfa doğrusal bir düzlem ile ayrılabiliyordu. Uygulamada bu durum her zaman geçerli olmayabilir. Yani doğrusal bir düzlem ile veriler birbirinden ayrılamayabilir (Şekil 3) [21].



Şekil 3. Doğrusal olarak ayrılamayan veriler
(Non-linearly separable data)

Verilerin doğrusal olarak ayrılamadığı durumlarda doğrusal sınıflandırıcı yerine doğrusal olmayan sınıflandırıcılar kullanılabilir. Bu bağlamda Doğrusal Olmayan Özellik Uzayı; $x_i \in R^n$ gözlem vektörünü daha yüksek dereceden bir uzayda z vektörüne dönüştürerek, bu yeni uzayda doğrusal sınıflandırıcıları elde etmek söz konusu olabilir. Bu z vektörünün yer aldığı özellik uzayı F ile gösterilsin. Bu durumda Ø ifadesi $R^n \rightarrow R^F$ eşlemesini yapmak üzere $z = \emptyset(x)$ biçiminde (Eş. 1) ifade edilebilir [22].

$$x \in R^n \rightarrow z(x) = [a_1, \emptyset_1(x), \ldots \ldots a_n, \emptyset_n(x)]^T \in R^F \quad (1)$$

Doğrusal olmayan ayrılabilirlik durumu düşünüldüğü zaman, eğitim örnekleri orijinal giriş uzayında doğrusal olarak ayrılamazlar. Bu tip durumlarda DVM, doğrusal olmayan haritalama fonksiyonu yardımıyla orijinal giriş uzayından doğrusal olarak kolayca sınıflandırma yapabileceği yüksek boyutlu nitelik uzayına dönüşüm yapar. Böylece çekirdek fonksiyonları kullanılarak bütün değerlerin tekrar tekrar çarpım değerlerinin hesaplanarak bulunması yerine doğrudan çekirdek fonksiyonunda değerin yerine koyularak nitelik uzayındaki değerinin bulunması sağlanır. Bu sayede, son derece yüksek boyutlu bir nitelik uzayı ile uğraşma olasılığı kalmaz. Çekirdek fonksiyonlarının diğer bir avantajı da, eğitme aşamasında bir eğitim örneği için fonksiyon kurulup değerler bulunduktan sonra diğer örnekler için artık kalıp değerleri eğitim örneği dışında tamamen hazır olduğu için çok daha kolay hesaplanmasıdır [23]. DVM'de yaygın olarak kullanılan dört çekirdek fonksiyonu vardır. Bunlar; doğrusal, polinomiyal, sigmoid ve radyal tabanlı fonksiyonlardır. DVM'ler ile ilgili daha ayrıntılı bilgiye [18]'dan ulaşılabilir.

3. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

10721 adet örnek içeren veri kümesi, 10-katlı çapraz doğrulama kullanılarak (%90 eğitim, %10 test) ve kullanılmadan (%70 eğitim, %15 doğrulama, %15 test) ayrı ayrı çalışılmıştır. Önerilen modelin doğruluğunu değerlendirmek için test verileri, DVR tahmin modeline uygulanmış ve elde edilen sonuçlar Ortalama Karesel Hata (OKH) (Eş. 2), Ortalama Mutlak Hata (OMH) (Eş. 3), Standart Tahmin Hatası (STH) (Eş. 4) ve Korelasyon Katsayısı (R) (Eş. 5) ölçütlerine göre değerlendirilmiştir.

$$OKH = \frac{1}{n}[\sum_{i=1}^{n}(O_i - P_i)^2] \quad (2)$$

$$OMH = \frac{1}{n}\sum_{i=1}^{n}|O_i - P_i| \quad (3)$$

$$STH = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n}} \quad (4)$$

$$R = \sqrt{1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - O_m)^2}} \quad (5)$$

Burada, n test için kullanılan veri sayısını, P_i tahmin edilen değeri, O_i gözlenen değeri ve O_m ise gözlenen değerlerin ortalamasını göstermektedir. OKH, tahmin edilen değerlerin doğruluk derecesi hakkında en doğru yaklaşımı yapmaktadır. Bu hesaplamada hataların kareleri alındığı için ölçüm içindeki büyük hataların ortalama üzerindeki etkisi daha büyük olmakta ve bu sayede büyük hataların tüm ölçüm üzerindeki etkisi belirlenebilmektedir. OMH, ölçüm değerleri ile model tahminleri arasındaki mutlak hatayı sorgular. STH, tahmin edilen değerlerin regresyon eğrisinden ne kadar uzaklaştığının ortalama değerini verir. OKH, OMH ve STH değerleri sıfıra ne kadar yaklaşırsa, modelinin tahmin yeteneğinin o kadar iyi olduğu sonucu çıkar [24]. R değeri ise tahmin edilen değer ile gerçek değer arasındaki doğrusal ilişkiyi sorgular. R'nin alabileceği en büyük değer +1 olup, R'nin +1'e eşit olması, değişkenler arasında pozitif ve tam doğrusal bir ilişkinin varlığını ortaya

Tablo 2. Veri kümesi istatistikleri (Dataset statistics)

	Girdi Bilgileri					Çıktı Bilgileri		
	u	v	w	m	n	u'	v'	w'
Minimum	0,04	0,04	0	2	1	0,04	0,03	0
Maksimum	0,95	0,95	0,99	12	6	0,96	0,96	0,99
Ortalama	0,50	0,50	0,50	8,22	0,33	0,50	0,49	0,49
Standart Sapma	0,29	0,29	0,29	2,14	1,68	0,29	0,29	0,28

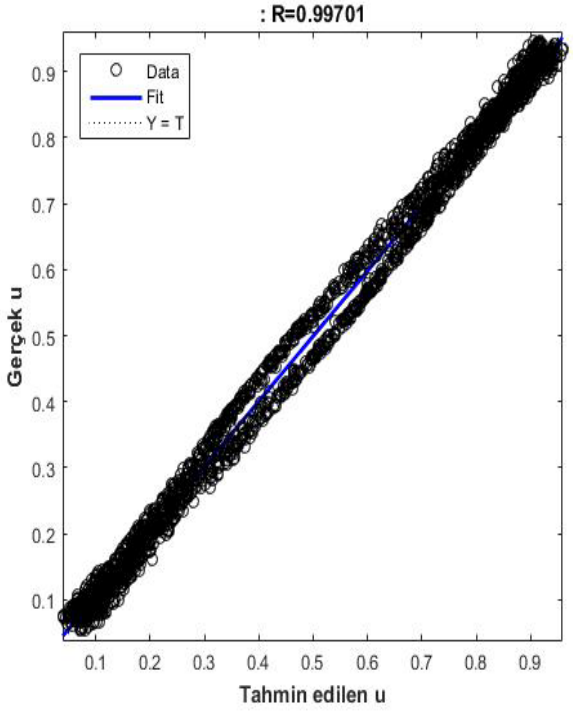
Tablo 3. 10-katlı çapraz doğrulama sonucunda elde edilen u' koordinatı tahmin başarım değerleri
(The performance results of u' coordinate prediction with 10-fold cross validation)

Kat	OKH	OMH	STH	R
1	5,333006e-04	1,990678e-02	9,997333e-01	9,968502e-01
2	5,055757e-04	1,941142e-02	9,997472e-01	9,970751e-01
3	4,890228e-04	1,896345e-02	9,997555e-01	9,971422e-01
4	5,320312e-04	1,996333e-02	9,997339e-01	9,967898e-01
5	5,178528e-04	1,971203e-02	9,997410e-01	9,969124e-01
6	5,065364e-04	1,949053e-02	9,997467e-01	9,970563e-01
7	5,206560e-04	1,963457e-02	9,997396e-01	9,968964e-01
8	4,874628e-04	1,889917e-02	9,997562e-01	9,971268e-01
9	5,105723e-04	1,959983e-02	9,997447e-01	9,969995e-01
10	5,332660e-04	1,985854e-02	9,997333e-01	9,968100e-01
Ortalama	5,136277e-04	1,954396e-02	9,997432e-01	9,969659e-01

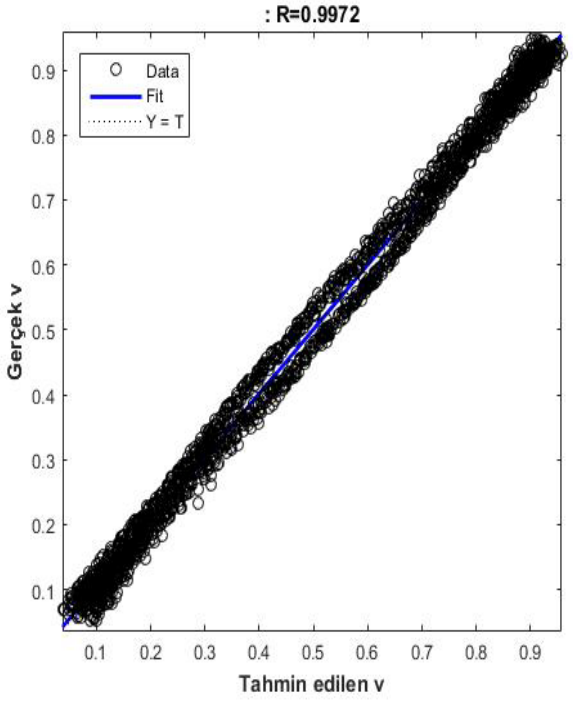
Tablo 4. 10-katlı çapraz doğrulama sonucunda elde edilen v' koordinatı tahmin başarım değerleri
(The performance results of v' coordinate prediction with 10-fold cross validation)

Kat	OKH	OMH	STH	R
1	5,100359e-04	1,959719e-02	9,997449e-01	9,969557e-01
2	5,034257e-04	1,947339e-02	9,997483e-01	9,970456e-01
3	4,876991e-04	1,899584e-02	9,997561e-01	9,971489e-01
4	4,996794e-04	1,932971e-02	9,997501e-01	9,971001e-01
5	4,962250e-04	1,924563e-02	9,997519e-01	9,970568e-01
6	4,948734e-04	1,926974e-02	9,997525e-01	9,970021e-01
7	5,016110e-04	1,932407e-02	9,997492e-01	9,970472e-01
8	4,725569e-04	1,873097e-02	9,997637e-01	9,972077e-01
9	4,821086e-04	1,896942e-02	9,997589e-01	9,971962e-01
10	4,931843e-04	1,926880e-02	9,997534e-01	9,971478e-01
Ortalama	4,941399e-04	1,922048e-02	9,997529e-01	9,970908e-01

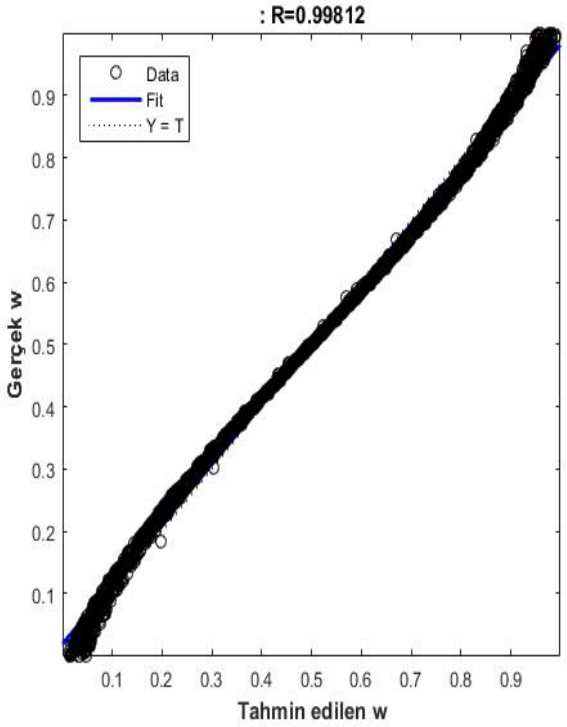
koyar [25]. 10-katlı çapraz doğrulama yapılarak elde edilen atomik koordinat tahmini sonuçları Tablo 3, Tablo 4 ve Tablo 5'te gösterilmiştir. 10-katlı çapraz doğrulama yapılmadan atomik koordinat tahmini yapıldığında Tablo 6'da görülen sonuçlar elde edilmiştir. Şekil 4, Şekil 5 ve Şekil 6, 10-katlı çapraz doğrulama yapılmadan elde edilen sonuçlar ile gerçek sonuçlar arasındaki regresyonu göstermektedir.



Şekil 4. u' koordinatı tahminleri için R değerleri
(R values for u' coordinate predictions)



Şekil 5. v' koordinatı tahminleri için R değerleri
(R values for v' coordinate predictions)



Şekil 6. w' koordinatı tahminleri için R değerleri
(R values for w' coordinate predictions)

Tablo 3, Tablo 4 ve Tablo 5'teki sonuçlar yorumlandığında aşağıdaki sonuçlara ulaşılabilir.

- u', v' ve w' atomik koordinatları bütün katlarda çok düşük tahmin hataları ve oldukça yüksek R değerleri ile tahmin edilmiştir.
- OKH ve OMH değerlerine bakıldığında en düşük hata değerinin 4,343866e-04 ve 1,829499e-02 ile w' koordinatı tahmininde elde edildiği; en düşük STH değerinin 9,997333e-01 ile u' koordinatının tahmininde elde edildiği görülmüştür.
- En yüksek R değeri 9,982015e-01 ile w' koordinatı tahmininde elde edilmiştir.
- Ortalama durum değerlendirildiğinde en iyi OKH, OMH ve R değerlerinin w' koordinatı tahmininde, en iyi STH değerinin ise u' koordinatı tahmininde elde edildiği görülmüştür.

Tablo 6, Şekil 4, Şekil 5 ve Şekil 6 yorumlandığında aşağıdaki sonuçlara ulaşılabilir

- SVR tahmin modeli, u', v' ve w' atomik koordinatlarını çok düşük hata oranları ve yüksek R değeri (1'e çok yakın) ile tahmin edebilmiştir.
- Tüm OKH, OMH, STH ve R ölçütlerinde en iyi başarım w' koordinatı tahmininde elde edilmiştir.
- u' koordinatı ile v' koordinatının tahmin başarısı birbirine çok yakındır.

Bu çalışmada, KNT'lerin atomik koordinatlarını hesaplamak yerine yüksek doğrulukta tahmin ederek uzun benzetim sürelerinin düşürülmesi amaçlanmıştır. Tahmin modeli için

Tablo 5. 10-katlı çapraz doğrulama sonucunda elde edilen w' koordinatı tahmin başarım değerleri
(The performance results of w' coordinate prediction with 10-fold cross validation)

Kat	OKH	OMH	STH	R
1	4,449206e-04	1,846219e-02	9,997775e-01	9,980535e-01
2	4,590819e-04	1,880112e-02	9,997704e-01	9,980209e-01
3	4,404515e-04	1,841490e-02	9,997797e-01	9,981562e-01
4	4,447722e-04	1,848818e-02	9,997776e-01	9,980136e-01
5	4,435198e-04	1,852855e-02	9,997782e-01	9,981186e-01
6	4,623616e-04	1,901068e-02	9,997688e-01	9,980300e-01
7	4,520805e-04	1,863283e-02	9,997739e-01	9,981003e-01
8	4,343866e-04	1,837684e-02	9,997828e-01	9,982015e-01
9	4,410566e-04	1,842806e-02	9,997794e-01	9,980958e-01
10	4,366166e-04	1,829499e-02	9,997817e-01	9,981504e-01
Ortalama	4,459248e-04	1,854384e-02	9,997770e-01	9,980941e-01

SVR yöntemi kullanılmış olup, yöntemde kullanılan veri kümesi CASTEP yazılımı ile hazırlanmıştır. SVR tahmin modeli 10-katlı çapraz doğrulama kullanılarak ve kullanılmadan ayrı ayrı test edilmiştir. OKH, OMH, STH ve R değerleri tahmin sonuçlarının doğruluğunu belirlemek için hesaplanmış olup sıfıra yakın OKH, OMH ve STH değerleri ile bire yakın R değerlerine ulaşılmıştır. Yüksek doğruluk oranları ile dakikalar içinde elde edilen sonuçlar, gerçek koordinat değerleri olarak kullanılabilir. Bunun yanı sıra, elde edilen veriler KNT benzetim yazılımları için başlangıç değerleri olarak kullanılabilir. Başlangıç değerleri kullanılmadığında günler süren benzetimler, bu değerler kullanıldığında saatler içinde bitirilebilir.

Tablo 6. 10-katlı çapraz doğrulama yapılmadan elde edilen u', v' ve w' koordinatları başarım değerleri
(The performance results of u', v' and w' coordinates prediction without 10-fold cross validation)

	u'	v'	w'
OKH	5,152795e-04	4,836680e-04	4,336160e-04
OMH	1,956568e-02	1,890637e-02	1,821846e-02
STH	9,997423e-01	9,997581e-01	9,997832e-01
R	9,970081e-01	9,972008e-01	9,981183e-01

TEŞEKKÜR (ACKNOWLEDGEMENT)

Desteklerinden dolayı Çukurova Üniversitesi Bilimsel Araştırma Projeleri Başkanlığı'na teşekkür ederiz. (Proje kodu: FDK-2015-3170).

KAYNAKLAR (REFERENCES)

1. Hohenberg P., Kohn W., Inhomogeneous electron gas, Physical Review, 136, B864-B871, 1964.
2. Kohn W., Sham L.J., Self-consistent equations including exchange and correlation effects, Physical Review, 140 (4A), A1133-A1138, 1965.
3. Akgenç B., Elektronik yapılarda DFT tabanlı hesaplamalar ve nano-düzeydeki molekül bağıntılarının iletim katsayısı, Master Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Sakarya, 2010.
4. Yetim A., Karbon nano tüpler, Master Tezi, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Adana, 2011.
5. Materials Design. MedeA. http://www. materialsdesign. com/medea. Erişim tarihi Ağustos 16, 2016.
6. FQS Poland. SCIGRESS - Molecular modeling software. http://www.fqs.pl/Chemistry_Materials_Life_Science/products/scigress. Erişim tarihi Ağustos 16, 2016.
7. JCrystalSoft. http://www.jcrystal.com. Erişim tarihi Ağustos 16, 2016.
8. BIOVIA Materials Studio. CASTEP. http://accelrys.com/products/collaborative_science/biovia-materials-studio/quantom-catalysis-software.html. Erişim tarihi Ağustos 16, 2016.
9. Kendall R.A., Aprà E., Bernholdt D.E., Bylaska E.J., Dupuis M., Fann G.I., Harrison R.J., Ju J., Nichols J.A., Nieplocha J., High performance computational chemistry: An overview of NWChem a distributed parallel application, Comput. Phys. Commun., 128 (1), 260-283, 2000.
10. Cheng G., Wu H., Qiang X., Ji Q., Zhao Q., Graphene field-effect transistor modeling based on artificial neural network, International Conference on Mechatronics, Electronic, Industrial and Control Engineering, Shenyang-Çin, 1479-1483, 1-3 Nisan, 2015.
11. Zayid E.I.M., Akay M.F., Predicting the performance measures of a message-passing multiprocessor architecture using artificial neural networks, Neural Computing and Applications, 23 (7-8), 2481-2491, 2013.
12. Riverbed. OPNET is now part of Riverbed SteelCentral. http://www.riverbed.com/gb/products/steelcentral/opnet.html?redirect=opnet. Erişim tarihi Ağustos 18, 2016.
13. Sharma A., Sahoo P.K., Tripathi R., Meher L.C., Artificial neural network-based prediction of performance and emission characteristics of CI engine using polanga as a biodiesel, Int. J. Ambient Energy, 1, 1-12, 2015.
14. Kayabaşı A., Akdağlı A., Usage of ANN and ANFIS methods for computing resonant frequency of slot-loaded compact microstrip antennas, Journal of the Faculty of Engineering and Architecture of Gazi University, 31 (1), 105-117, 2016.
15. Yurtcu Ş., Özocak A., Prediction of compression index of fine-grained soils using statistical and artificial intelligence methods, Journal of the Faculty of

Engineering and Architecture of Gazi University, 31 (3), 597-608, 2016.

16. Fakhim B., Hassani A., Rashidi A., Ghodousi P., Predicting the impact of multiwalled carbon nanotubes on the cement hydration products and durability of cementitious matrix using artificial neural network modeling technique, The Scientific World Journal, 4 (1), 1-8, 2013.

17. MathWorks. MATLAB. http://www.mathworks.com/products/matlab. Erişim tarihi Ağustos 18, 2016.

18. Vapnik V., The nature of statistical learning theory, Springer-Verlag, New York, A.B.D., 2000.

19. Smola A.J., Schölkopf B., A tutorial on support vector regression, Statistics and Computing, 14 (3), 199-222, 2004.

20. Acikkar M., Akay M.F., Akturk E., Gulec M. Intelligent regression techniques for non-exercise prediction of VO(2)max, Signal Processing and Communications Applications Conference, Girne-K.K.T.C., 1-4 ,24-26 Nisan, 2013.

21. Yakut E., Elmas B., Yavuz S., Yapay sinir ağları ve destek vektör makineleri yöntemleriyle borsa endeksi tahmini, Suleyman Demirel University Journal of Faculty of Economics & Administrative Sciences, 19 (1), 139-157, 2014.

22. Yakut E., Veri madenciliği tekniklerinden C5.0 algoritması, destek vektör makineleri ile yapay sinir ağlarının sınıflandırma başarılarının karşılaştırılması: İmalat sektöründe bir uygulama, Doktora Tezi, Atatürk Üniversitesi, Sosyal Bilimleri Enstitütüsü, 2012.

23. Kecman V., Learning and soft computing: Support vector machines, neural networks, and fuzzy logic models, MIT Press, Cambridge, İngiltere, 2001.

24. Aydogmus H.Y., Ekinci A., Erdal H., Optimizing the monthly crude oil price forecasting accuracy via bagging ensemble models, Journal of Economics and International Finance, 7 (5), 127-136, 2015.

25. Yapraklı T.Ş., Erdal H., Firma başarısızlığı tahminlemesi: Makine öğrenmesine dayalı bir uygulama, International Journal of Informatics Technologies, 9 (1), 21-31, 2016.