

Yazılım Efor Tahmininde Farklı Bir Yaklaşım: Sınır Değerlerine Göre Tahmin

Ömer Faruk Saraç¹, Nevcihan Duru²

¹ Bank Asya Katılım Bankası, İstanbul

² Kocaeli Üniversitesi Bilgisayar Mühendisliği, Kocaeli

Özet. Yazılım efor tahmini, yazılım proje yönetiminde çok önemli bir aşamadır. Tahmin değerinin doğruluğu proje başarıya ya da başarısızlığa doğrudan etki eder. Yöneticiler uygun kaynakları tahmin etmeye çalışırlar ve bu yönetim için zorlayıcı bir durumdur. Araç ve tekniklerin yardımıyla tahmin süreci daha iyi gerçekleştirilebilir. COCOMO en çok kullanılan, parametrik modellerden biri olarak ifade edilen yöntemlerdendir. Aynı zamanda, COCOMO tabanlı Yapay sinir ağlarının (YSA) kullanıldığı çalışmalar da mevcuttur. Bu çalışmalarda toplam performansın arttığı gösterilmiştir. Fakat efor tahmini süreci, genelde, tek çıktı üretir: tahmin edilen efor değeri. Bir proje yöneticisinin tahminlerinde bazı üst ve alt limitleri, sınırları dikkate alması gerektiği bilinen bir gerçektir. Bu çalışmada yeni bir metod önerilmiş, COCOMO tabanlı YSA ile K-Means kümelerine girdi olarak sunulmuş, olası küme belirlenmiş ve buna göre olası üst ve alt limit efor tahmin değerleri üretilmiştir. Deneysel çalışma sonuçları sayısal olarak kabul edilebilir düzeyde çıkmış ve kullanılabilir birden çok tahmin değeri üretilerek daha pratik bir efor tahmin sürecinin olabileceğini göstermiştir.

Anahtar Kelimeler. yazılım efor tahmini, COCOMO, yapay sinir ağları, k-means

1 Giriş

Yazılım efor tahmini, yazılım proje yönetiminde çok önemli bir aşamadır. İhtiyaç duyulan kaynaklar, efor ve zaman değerleri bu aşamada tahmin edilir[1]. Bu nedenle tahmin değerlerinin başarısı çok önemlidir. Tahminlerin fazla ya da az olarak gerçekleşmesi bazı problemlere yol açabilir ve hatta projeyi başarısızlığa götürebilir. Dolayısıyla, tutarlı bir tahmin gerçekleştirme, proje yöneticisinin en önemli işlerindedir.

Yazılım efor tahmininin yapılması zor bir işdir. Tahminler daha sonra, planlamaya, takvim düzenlemelerine, bütçeye ve diğer bazı anlaşmalara temel oluşturur. Planlama ve bütçe değerleri doğru olmadığında projeden beklenen kazanç elde edilemeyebilir. Sürecin zorluğu ve bu şekilde diğer aşamalarla olan doğrudan ilişkisi dolayısıyla proje

yöneticileri karar verme süreçlerini destekleyecek araç ve yöntemlere ihtiyaç duymaktadır.

Efor tahmini yaklaşım itibariyle farklı kategorilere ayrılabilir: uzman değerlendirmesi, parametrik metotlar ve makine öğrenmesi tabanlı metotlar gibi[2]. Her bir kategori altında da belli yöntem ve yaklaşımlar yer alır. En bilinen yöntemlerden biri, B. Boehm tarafından ortaya konulan ve parametrik model olarak ifade edilen COCOMO' dur[3]. COCOMO, Constructive Cost Model ifadesinin kısaltılmışıdır ve basitçe, parametre ve sabit değerleri olan bir formüldür. Özellikle kolay hesaplama imkânı olmasından tercih edilen metotlardandır.

Bir diğer bilinen yöntem, veri madenciliği ve makine öğrenmesi tabanlı, yapay sinir ağlarıdır(YSA). YSA ile karmaşık lineer olmayan problemler modellenilebilir[5]. Ek olarak YSA tahmin değerleri üretebilir. Bu nedenlerden dolayı efor tahmini alanında da kullanılmaktadır. Genelde YSA, COCOMO ile beraber kullanılmaktadır. Farklı mimari yapılarının ve algoritmalarının olmasıyla beraber, en çok tercih edilenlerden birinin ileri beslemeli geri yayımlı YSA olduğu söylenebilir[4,5]. COCOMO maliyet faktörleri YSA için girdi katmanını oluşturur ve sonuç olarak efor tahmini değeri üretilir.

COCOMO ve YSA' nın beraber kullanıldığı ve önemli performans artırımlarının gösterildiği çalışmalar mevcuttur[4,5,6,7]. Bununla beraber, efor tahmini halen çalışma ve iyileştirmelere açık bir alandır. Özellikle performans ölçütlerinin sistem genelinde olması, yüzde olarak hata oranlarının düşmesine rağmen daha iyi sonuçlara ulaşma isteği gibi nedenlerle bu alan üzerinde çalışmalar devam etmektedir. Bilindiği gibi, çoğu model, efor tahmini sonucunda tek bir çıktı üretir. Bu değer belli bir ölçüme göre verilen (adam/ay gibi) efor tahmin değeridir. Ancak, PMI tarafından da ifade edildiği şekliyle, proje planlamalarında istenmeyen durumların da göz önünde bulundurulması ve buna göre planlamada esnekliğin de olması bilinen bir gerçektir[8]. Efor tahmini sürecini göz önünde bulundurduğumuzda, tahmin değerinin yanında olası üst ve alt limitlerin de tahmin edilmesi, karar aşamasında yöneticilere yardımcı olabilir.

Menzies, yazılımın kendisinin de bir veri kaynağı olduğunu ifade etmiştir[9]. Yazılım içerisinde yer alan kaynak kodları, hata veri tabanları, yazışmalar, raporlar ve yazılım geliştirme sürecinde ortaya çıkan diğer bilgiler bu veri kaynaklarından sayılabilir. Her bir kaynağın, süreci geliştirmede kullanılabileceğini ve bu şekilde sürecin kendisinin yine kendi ürettiği bilgiden de beslenebileceğini ifade etmiştir. Bu çalışmada da benzer şekilde, geçmiş veriler kullanılarak, sadece tahmin değeri üretmek yerine olası esnek değerlerin de sunulması hedeflenmiştir.

Bu çalışmada, COCOMO tabanlı YSA modeli kullanılmış, ek olarak, esnek bir yapı sunmak için k-means yöntemi de modele dahil edilerek, efor tahmin değeri ile beraber olası üst ve alt limit değerleri de sunulmuştur. Bu yaklaşımla efor tahmininde kullanılabilecek olası birden çok değer üretilmiştir. Her ne kadar karar destek sistemlerinin basitliği önemli olsa da, tahmin sürecini daha verimli yapabilmek için bu şekilde detaylı bilgilerin sunulması faydalı olacaktır.

COCOMO modelinin seçilmiş olmasının en önemli nedeni, uygulamada görece kolaylık sağlamasıdır. Neticede matematiksel formüle dayandığından, irdeleme ve kontrol olanakları fazladır. Ek olarak, diğer olası modellere referans verilmemesinin

nedeni, yaygın olarak kullanılan ve akademik olarak üzerinde çeşitli çalışma ve iyileştirmelerinin yapıldığı bu model ile, çalışmada ortaya atılan modelin daha rahat sunulabiliyor olmasıdır. Çalışmada amaç COCOMO modelini geliştirmekten ziyade, bilindik bir model üzerine bina edilecek şekilde, efor tahmin yöntemlerini esnetme, olası sınır değerlerini üretme işlemini sürece dahil etme denemesidir.

Çalışma içerisinde, 2. bölümde önceki araştırmalardan örnek verilmiş ve sonuçları kısaca belirtilmiştir. Takip eden bölümde ise önerilen model ortaya konulmuştur. Değerlendirme kriterleri ve test sonuçları çalışmaya dâhil edilmiş ve çıkarımlar belirtilmiştir. Son olarak bulgular neticesinde ortaya çıkan sonuçlar belirtilmiş ve olası gelecek çalışma alanları ifade edilmiştir.

2 İlgili Çalışmalar

Geçen zamanda efor tahmini ile ilgili birçok çalışma yapılmıştır. Jorgensen, yayınladığı makalede bu çalışmaları belli bir düzende incelemiş ve detaylı bilgi vermiştir[11]. Bu yayın ile efor tahminindeki gelişmeler zaman ölçeğinde irdelenebilir.

Giriş bölümünde de ifade edildiği şekliyle, birçok araştırmada efor tahmini gerçekleştiriminde COCOMO tabanlı YSA kullanılmıştır[4,5,6,7]. Bu çalışmalarda kendi içinde mimari seçimler farklılık gösterse de genelde tercih edilenin ileri beslemeli ağ olduğu ifade edilebilir. Bununla beraber, YSA yapısının bir tür kapalı kutu olarak değerlendirilmesi[5,7], model üzerinde bazı ayarlamalar yapmayı da beraberinde getirmiştir[4,5]. Değişiklikler performansı artırmak amacıyla. Aynı zamanda yapılan düzenlemelerle YSA modelinde, ilgili yapıdaki elemanların hangi işlevi gördüklerinin daha iyi ifade edilebilmesi sağlanabilir.

Çalışmalarda genelde, açık olarak ulaşılabilen COCOMO tabanlı veri kümeleri kullanılmaktadır. B.Boehm, COCOMO tanımını da verdiği yayınında bir veri kümesi de sunmuştur ve birçok yerde bu veri kümesi kullanılmıştır. Çalışmalarda bu veri kümesinin tercih edilmesi, modelleri birbiri ile kıyaslama imkânını da sunmuştur.

Kaushik COCOMO tabanlı ileri beslemeli bir YSA modeli önermiştir[4]. Modelde, parametreler logaritmik bir fonksiyondan geçirilmiş ve buna göre tahmin değeri üretilmiştir. Modelin çıktılarının gerçekleşen efor değerlerine yakın olduğu ifade edilmiştir. Çalışmada COCOMO 81 olarak adlandırılan veri kümesi kullanılmıştır. Değerlendirme kriteri olarak MRE (Mean Relative Error) seçilmiştir[12], yüzde olarak beklenen değer ve tahmin edilen değer arasındaki farkın oranı verilmiştir.

Reddy de benzer bir yaklaşıma sahiptir. Çalışmasında yine COCOMO tabanlı YSA kullanılmıştır[5]. Modelde farklı olarak, orijinal COCOMO modelini logaritmik şekle dönüştürülmüş ve bu şekilde 2 temel parçanın olduğu ifade edilmiştir. Bu parçalardan biri, maliyet faktörlerini ifade eder, diğeri ise formül çarpanlarını oluşturur. Buna göre gizli katmanda 2 nöron kullanılmış, ilgili girdi nöronları sadece formülde bağlantısı olduğu nörona sunulmuştur. Çalışma sonuçları, COCOMO tahminleriyle karşılaştırılmış ve yine MRE değerlendirme kriteri olarak verilmiştir. Modelin daha başarılı ve dengeli olduğu çıkarımı yapılmıştır.

Idri ise çalışmasında YSA' nın efor tahmini sürecine uydurulup uydurulamayacağı sorusunu sormuş ve buna göre değerlendirme yapmıştır[6]. Bu çalışmada ileri bes-

lemeli, 3 katmanlı YSA kullanılmış, test için de yine COCOMO 81 veri kümesi tercih edilmiştir. Modelde bulanık mantık kuralları ile YSA modeli eşleştirilmiştir. COCOMO parametreleri incelenmiş, her bir parametrenin limitleri ölçüsünde kural kümeleri oluşturulmuştur. Sonuçların kabul edilebilir bir seviyede olduğu ifade edilmiştir.

Kultur, daha farklı bir yaklaşıma sahiptir. Tek bir YSA yerine, bütünleşik, 20 adet YSA kullanan bir model önermiştir[7]. Modelde, veri kümesinden küçük alt kümeler oluşturulmuş, her bir alt küme, YSA' lara girdi olarak sunulmuştur. Sonuçlar gruplanmış, en büyük gruptaki sonuçların ortalaması değerlendirilmiştir. Ek olarak, geçmiş bilgilere göre bias değeri üretilmiş, tahmin edilen değere dâhil edilmiştir. Bias hesaplamaları k en yakın komşu algoritmasına göre gerçekleştirilmiştir. COCOMO tabanlı veri kümeleri test için kullanılmış ve her bir küme için hem YSA, hem de önerilen modelin sonuçları verilmiştir. Önerilen metodun diğerlerinden başarılı ve daha dengeli olduğu ifade edilmiştir.

Bahsi geçen çalışmalar, temelde COCOMO modelini baz almıştır. Aynı zamanda bu çalışmaya temel teşkil eden YSA modeli üzerine kurgulanmıştır. Efor tahmini süreci, doğal olarak, bir sayısal değer üretme amacı güttüğünden, sonuç listeleri birtakım sayısal değerler ve bu değerlerin belli ölçeklere göre karşılaştırmaları şeklinde olmaktadır. Çalışmalar genelinde kabul görüm MRE değerini kullanmaktadır. Bu nedenle odak olarak MRE değerleri aktarılmıştır.

3 Önerilen Model

İlgili çalışmalarda da ifade edildiği gibi, araştırmalar, YSA' nın efor tahmininde kullanılabileceğini göstermiştir. YSA efor tahmin başarısını genelde artırmıştır. Çalışmalarda model yapısı ve değerlendirme metotları değişse de YSA ortak olarak kalmıştır.

Bu çalışmada da YSA tabanlı efor tahmin modeli sunulmuştur. Efor tahmin sürecinin tek çıktı üreten yapısı esnetilerek, olası üst ve alt limit değerleri de tahmin edilmiştir. YSA, diğer çalışmalardaki gibi COCOMO tabanlıdır. Buna göre COCOMO parametreleri YSA için girdi oluşturmaktadır. YSA için ileri beslemeli geri yayımlı yaklaşım tercih edilmiştir. Orta katmanda 5 nöron bulunmaktadır ve çıktı katmanında 1 nöron vardır. YSA içerisinde herhangi bir bias nöron mevcut değildir. Çalışmada asıl amaç, tek başına efor tahmin değeri üretmek yerine, olası farklı değerlerin üretilmesinin yorumlama ve karar vermede daha faydalı olduğunu göstermektir.

Geçmiş proje bilgileri ki bu çalışma kapsamında COCOMO 81 veri kümesidir, eğitim ve test sürecinde kullanılmaktadır. YSA, bu bilgilere göre efor tahmininden sorumludur. K-means ise, üretilen tahmin değerini destekleyecek üst ve alt limit değerlerini üretmektedir. Model ile ilgili detaylar ve veri kümesi hakkında bilgi aşağıdaki bölümlerde verilmiştir.

3.1 Veri Kümesi

PROMISE sitesi üzerinde, açık olarak ulaşılabilecek efor tahmininde kullanılacak veri kümeleri mevcuttur[10]. Önerilen model için bu kümelere COCOMO 81 veri kümesi kullanılmıştır. Küme içerisinde 63 satır bulunmaktadır. Çalışmada daha sağlıklı sonuç elde edebilmek için tüm değerler normalize edilmiştir. Normalizasyondan sonra sınır değerlere sahip olan minimum ve maksimum satırları çıkarılmıştır.

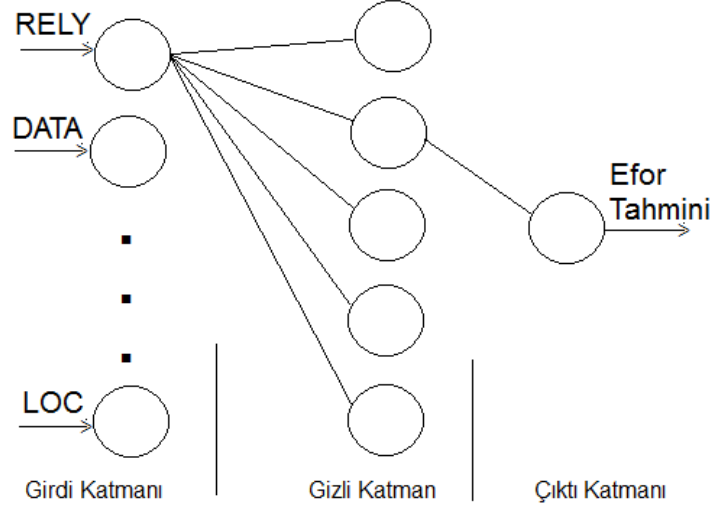
Küme, çalışmalarda 2 bölüme ayrılmış, %75' lik bölümü eğitim için, kalan ise test için kullanılmıştır. Kümelerin oluşturulması için bir random fonksiyon kullanılmıştır. Bu fonksiyona göre, tüm küme içerisinde belli sayıda satırlar 2 yeni alt kümeye oluşturulmuştur. Özellikle eğitim aşamasında veri kümesinin içeriği ile beraber, verilerin sırası da sonuçları etkileyebileceğinden, daha modüler bir yapıya ulaşmak için rastgele elemanların seçilmesi önemlidir. Bilhassa k-means algoritmasında bağlantı değerleri ve kümeler, sonuçta oluşan alt kümeleri çok fazla etkilemektedir. Bu nedenle eğitim ve test veri kümelerinin oluşturulmasında sıra gözlemlenmeden, rastgelelik sağlayan metod tercih edilmiştir.

3.2 YSA Modeli

İleri beslemeli ve geri yayımlı, 3 katmanlı YSA tercih edilmiştir. Girdi katmanında, veri kümesindeki gibi, COCOMO parametrelerini karşılayacak şekilde, girdi nöronları bulunmaktadır. Ek olarak, satır sayısı ölçümünü ifade eden, LOC (Lines Of Code) değeri de girdilere eklenmiştir. Orjinal COCOMO hesaplamasında yer alan bu değerin, YSA içerisine de eklenerek daha sağlıklı sonuç alınması hedeflenmiştir.

Şekil 1' de gösterildiği gibi, modelde 1 gizli katman mevcuttur ve bu katmanda 5 nöron vardır. Çıktı katmanında ise tek nöron, efor tahmin nöronu vardır. Herhangi bir dışsal nöron ya da bias değeri mevcut değildir. Aktivasyon fonksiyonu olarak Sigmoid kullanılmaktadır. Ağırlıklar, -1 ile 1 arasında değişen, rastgele değerler olarak atanmıştır.

Şekil üzerinde görülen, modelin basitleştirilmiş bir gösterimidir. İleri beslemeli bir ağ olduğu düşünülürse, akışın temelde girdi katmanından çıktı katmanına yöneleceği anlaşılabilir. Gösterim kolaylığı olması açısından sadece örnek bağlantılar verilmiştir. Her bir nöron, bir sonraki katmanda ilgili nöronlara bağlanmaktadır. Girdi katmanındaki etiketler ise, model içerisindeki nitelik isimlendirmelerinden alınmış, referans olması açısından orjinal isimlendirmelere sadık kalınmıştır.



Şekil 1. YSA Modelinin Gösterimi

3.3 K-Means Kümeleri

Modele esneklik sağlayan, k-means kullanımıdır. YSA hesaplamasında ortaya konulan tahmin değeri, k-means için kontrol değeri olmaktadır. Sistem tutarlılığını sağlamak için, YSA'ya girdi olarak sunulan veri kümesi, aynı şekilde k-means için de sunulmuştur. Modelin geliştirilmesi sırasında yapılan denemeler sonucu k sayısı, oluşturulacak küme sayısı 4 olarak belirlenmiştir. Küme eleman sayısının azlığı ve içerik, diğer değerlerde istenilen sonucu üretmeyi engellemiştir. Bununla beraber 4 değeri tamamen testler neticesinde ortaya konulmuştur.

Küme oluştururken ölçüm için Öklid uzaklığı kullanılmış, satırların küme merkezlerine olan yakınlığı baz alınmıştır. YSA'daki gibi, başlangıç değerleri -1 ve 1 aralığındaki rastgele değerlerden oluşturulmuştur.

COCOMO parametreleri, karşılığında sabit sayısal değerlerin olduğu kategorik sınıflandırma belirteçleridir. Bu değerler genelde 0,7 ile 1,5 arasında değişmektedir. Her bir parametre için ortalama 6 değer olduğu düşünülürse, sayısal değerlerin birbirine çok yakın ve aynı değerler olduğu anlaşılır. K-means algoritmasıyla kümeleme yapıldığında, kümeler oluşmaya başladığında, küme merkezleri birbirine çok yaklaşmakta ve her bir döngüde veri kümesindeki satırlar aynı kümelerde yoğunlaşmaktadır. Bu sorunu önleyebilmek için doluluk oranı adında yeni bir ölçüm dâhil edilmiştir. Buna göre sistem içerisinde bir eşik değeri belirlenmiş, bu değere ulaşan küme olduğunda, kümeye eleman ekleme aşamasında, bir sonraki en olası küme tercih edilmiştir. Örnek olarak, bir kümede 40 satır varsa ve doluluk oranı %40 olarak belirlenmişse bir kümenin alabileceği maksimum eleman sayısı 10 olacaktır. Doluluk oranı basitçe eleman sayısının küme sayısına oranının belli bir katsayısıdır. Buradaki katsayı ayarlanarak, kümelerin dengesi sağlanabilir.

K-means algoritmasıyla kümeleme işlemi tamamlandığında, her bir küme için belli istatistikler kullanılabilir. Bunlar, küme minimum, maksimum ve ortalama efor değeri gibi verilerdir. Bu değerlerle hem kümelerin homojenliği irdelenebilir hem de küme içerisindeki dağılıma göre tahmin edilen değer hakkında yorum yapılabilir.

3.4 YSA ile K-Means Kümelerinin Birleştirimi

Sistemde, YSA sonucu olarak üretilen değer, temel efor tahmin değeridir. Bu değer, daha sonra k-means ile oluşturulan kümelere sunulur ve olası en yakın üste ve alt limitler hesaplanır. Önerilen model esnekliği bu ilişki ile sağlanmaktadır.

Test kümesindeki her bir eleman için YSA efor tahmin değeri üretilir. Benzer şekilde, bir destek değeri olması için k-means efor değeri hesaplanır. Bu işlem için test satırının hangi kümeye ait olabileceği bulunur ve bu küme içerisindeki en yakın değer çıktı olarak verilir. Daha sonra YSA değeri, bulunan olası küme içerisinde değerlendirilir ve bu değere en yakın üst ve alt satırlar, olası üst ve alt limitler olarak verilir.

Küme içerisindeki en yakın değerleri bulabilmek için, kümedeki elemanlar ve YSA sonuç değeri bir liste olarak tutulur ve sıralanır. YSA değerine en yakın üst ve alt komşusu, sırası ile üst ve alt limit değerlerini oluşturur. Eğer YSA değeri listede ilk ya da son eleman olursa, olası değerler için ancak tek sonuç bulunacaktır. Bu durumda YSA, k-means içerisinde beklenenden daha büyük/küçük değer üretmiş denilebilir. Bununla beraber, test sonuçlarından da görüleceği gibi, modelden beklenen bu duruma düşen satır sayısının çok az olmasıdır.

YSA ile k-means kümelerinin birleşiminden model çıktısı olarak 4 değer üretilir. Bunlar, COCOMO hesabı ile ortaya çıkan hata oranı, YSA hata oranı, k-means hata oranı ve üst-alt limit değerleri için ayrı olarak verilen hata oranlarıdır. Model önerisinde de belirtildiği gibi, sonuç listesinde birden çok değer bulunması, karar destek aşamasında proje yöneticisinin daha geniş bir bakış açısına sahip olmasını sağlayacaktır.

3.5 Kullanıcı Arayüzü ve Model Aracı

Önerilen modelin gerçekleştirimi, Visual Studio üzerinde, C# 2.0 dili ile bir Windows uygulaması olarak yapılmıştır. Kullanıcıya belli değerlerin girilebileceği ve sonuçların tablo olarak gösterildiği bir arayüz hazırlanmıştır. Bu arayüz ile, test ve eğitim kümelerindeki eleman sayısı, k-means küme sayısı, YSA öğrenme oranı gibi temel parametreler verilebilir. Ayrıca, sistem virgülle ayrılmış dosya(csv) girdileri almaktadır. Bu dosyalarda COCOMO modeline uygun veriler vardır ve program girdi dosyasını okuyup ilgili oranda eğitim ve test kümelerini oluşturur. Şekil 2, program arayüzünün örnek bir görüntüsünü sunmaktadır.

Önerilen model, sadece çalışmayı destekleyecek çapta deneme programına dahil edilmiştir. Modelin daha güçlü bir şekilde incelenebilmesi için gerçek bir firma/kurum içerisinde belli bir standarda göre uyarlanması gerekmektedir. Mevcut çalışma için böyle bir imkan oluşmamıştır ancak, model geçerlemesi, tüm parçaları ilgili dil ve platform içerisinde gerçekleştirilmiş, gerçek bir program ile sağlanmıştır.

Bu şekilde model için istenilen seviyede sektör denemesi yapılamamış olsa da, program üzerinde gerçek deneyim ve testlerin yapılması sağlanmıştır. Neticede veri kümeleri açık olarak sunulan, daha önceki gerçek proje niteliklerini içermektedir.

Şekil 2. Kullanıcı Arayüzünden Bir Görünüm

4 Test Sonuçları

Bu bölümde, önerilen modelin, COCOMO 81 veri kümesi ile testinin sonucunda ortaya çıkan sonuçlar belirtilmiş, değerlendirme kriteri olarak kullanılan yöntem ifade edilmiştir. Test sonuçları üzerinden, model hakkında genel değerlendirmeler de sunulmuştur.

Tahmin gerçekleşen sistemlerde temel ölçüm, beklenen değer ile tahmin edilen değer arasındaki fark, ya da uyumdur. Diğer çalışmalarda olduğu gibi, bu çalışmada da değerlendirme ölçütü olarak MRE kullanılmıştır. MRE,

$$(|\text{Beklenen Değer} - \text{Tahmin Edilen Değer}| / \text{Beklenen Değer}) \times 100 \quad (1)$$

bağıntısı ile verilmektedir. MRE değerinin küçük olması, tahmin değerinin başarılı olduğunu ifade etmektedir. Bu nedenle başarı ölçütünde küçük olan MRE değerleri önemlidir.

Tablo 1, COCOMO 81 veri kümesi için ortaya çıkan sonuçları göstermektedir. Tablodan da görülebileceği gibi, k-means genelde daha başarılıdır. Bununla beraber, üst ve alt limit değerleri genelde YSA çıktı değerinden daha başarılıdır. Bunun anlamı, k-means ile bir aralık bulunmasına rağmen, destek değerlerinin YSA değerinden daha iyi olduğudur. Efor tahmini sürecinde düşünüldüğünde, tek başına YSA kullanımından çok daha iyi bir başarı elde edileceği çıkarımı yapılabilir. Üst ve alt değerler, aynı zamanda, proje için olası esneme miktarını verdiklerinden, YSA için mevcut olabilecek kayma miktarı da azalacaktır.

Tablo 1. Test Sonuçları

MRE%	Değerlendirme Türü				
	<i>COCOMO</i>	<i>YSA</i>	<i>K-Means</i>	<i>Alt Limit</i>	<i>Üst Limit</i>
1	18.2	1.76	14.71	14.71	14.71
2	28.36	31.51	0	66.67	5.26
3	34.6	34.44	0	0	0
4	13.01	13.48	18.27	22.12	18.27
5	29.4	53.89	0	14.12	15

MRE%	Değerlendirme Türü				
	<i>COCOMO</i>	<i>YSA</i>	<i>K-Means</i>	<i>Alt Limit</i>	<i>Üst Limit</i>
6	31.29	59.76	1.41	4.23	7.79
7	45.12	34.84	0	4.17	17.24
8	69.35	65.49	8.82	35.29	12.82
9	47.3	27.63	0.83	0.83	0.83
10	22.25	56.50	0	14.12	15
11	43.83	36.90	21.43	48.21	21.43
12	21.22	20.91	0	0	0
13	61.9	68.52	22.41	31.03	20.55
14	62.7	54.48	4.44	4.65	6.52
15	34.57	29.20	1.23	1.23	1.23

Tablodaki bazı satırlarda sonuçların 0 olduğu görülmektedir. Bu satırların sebebi, veri kümesinde aynı efor değerine sahip farklı proje satırları olmasıdır. Yani, farklı 2 değer için aynı sonucun çıkması durumudur. Bazı satırlar için de alt ve üst limit değerlerinin aynı olduğu gözlemlenmektedir. Bunun sebebi ise, model tanımında da belirtildiği gibi, tahmin edilen değer, ilgili k-means küme içerisindeki değerlerin en üst ya da en altında çıkmasıdır. Bu durumlarda tek limit değeri olacaktır. COCOMO hata oranlarının yüksek çıkmış olması ilginç bir durum olmakla beraber, efor tahmininde model geliştirme üzerinde çalışılmasının gerekçesini de ortaya koymaktadır.

Test sonuçları incelendiğinde, beklentinin genel olarak gerçekleştiği söylenebilir. 8 numaralı satır göz önünde bulundurulduğunda, COCOMO için 69 çıkan hata oranı, YSA için de yüksek sayılabilecek seviyede çıkmış, 68 olarak gerçekleşmiştir. Bununla beraber, alt ve üst limit değerleri ise, 35 ve 12 olarak ölçülmüştür. Bu değerler ışığında yorum yapılacak olursa, YSA kullanımında olası yüksek hata oranı mevcutken, bu değere ek olarak limit değerleri kullanıldığında hem hata oranı azalmış hem de proje karakteristiğine göre yöneticinin seçeceği yöne bağlı olarak olası esneme miktarı da tahmin edilebilecektir. Sonuç olarak tablodaki değerlere göre modelin dengeli olduğu çıkarımı yapılabilir.

5 Sonuç ve Gelecek Çalışmalar

Efor tahmini proje yönetiminde en önemli konulardan biridir. Bu süreci destekleyecek araç ve yöntemlerin ortaya konması, proje yöneticilerinin daha başarılı tahmin yapmalarını ve dolayısıyla daha iyi proje yönetimine sahip olmalarını sağlayacaktır.

Bu çalışmada, efor tahmininde farklı bir yaklaşım olarak, üst ve alt limitlere göre efor tahmini yapılmış ve önerilen model, test verileri ile denenmiştir. Sonuçlardan yapılabilecek çıkarımlara göre önerilen modelin dengeli olduğu, amaca hizmet edebileceği, karar destek aşamalarında birden çok çıktı ile yorum yapmanın daha sağlıklı

olabileceği gösterilmiştir. Model, COCOMO üzerine bina edilmiş ve YSA ile k-means yöntemleri kullanılmıştır. Model sonuçlarının tek başına hem COCOMO hem de YSA' dan daha başarılı olduğu gözlemlenmiştir.

Çalışma, COCOMO 81 adlı veri kümesi ile yapılmıştır ve bu kümede 63 satır bulunmaktadır. Modelin veri kümesinin büyüklüğünden etkilenip etkilenmediği bilinmemektedir. Aynı zamanda, limit değerleri basitçe bir sıralama fonksiyonu sonrası, YSA tahmin değerine en yakın satırların alınması şeklinde hesaplanmıştır. Proje karakteristiğine göre bu hesaplamaların daha esnek bir yapıya kavuşturulması, sonuç değerlerini daha anlamlı hale getirebilir. Ek olarak çalışmada standart YSA kullanılmış ve herhangi bir bias nöron dâhil edilmemiştir. Bias değerinin sisteme eklenmesi ve limit değerlerinin aralığının daha modüler şekilde hesaplanması olası gelecek çalışmaları arasında sayılabilir.

Kaynaklar

1. Huges, B. ve Cotterell, M.: Software Project Management, 4th Ed., McGraw Hill, ABD. (2004)
2. Stellman, A. ve Greene, J.: Applied Software Project Management, O'Reilly, ABD. (2009)
3. Boehm, B., Software Engineering Economics, Prentice Hall, ABD. (1994)
4. Kaushik, A., Chauhan, A., Mittal, D. ve Gupta, S.: COCOMO Estimates Using Neural Networks, International Journal of Intelligent Systems and Applications, Cilt 4, No 9, pp. 18-22 (2012)
5. Reddy, Ch.S. ve Raju, KVSVN.: A Concise Neural Network Model for Estimating Software Effort, International Journal of Recent Trends in Engineering, Cilt 1, No 1, pp. 188-193 (2009)
6. Idri A, Khoshgoftaar T, Abran A.: Can neural networks be easily interpreted in software cost estimation?, Proceedings of the IEEE International Conference on Fuzzy Systems, FUZZ-IEEE'02, Vol.: 2, pp. 1162-1167 (2002)
7. Kultur, Y., Turhan, B. ve Bener, A.: Ensemble of neural networks with associative memory (ENNA) for estimating software development costs, Knowledge-Based Systems, Cilt 22, No 6, pp. 395-402 (2009)
8. PMBOK, Project Management Body Of Knowledge, 4th Ed., Project Management Institute, ABD. (2008)
9. Marcus, A. ve Menzies, T.: Software Is Data Too , International Symposium on Foundations of Software Engineering, Santa Fe, NM, ADB., 229-232 (2010)
10. PROMISE Veri Kümeleri, <http://promise.site.uottawa.ca/SERepository/datasets-page.html>
11. Jørgensen, M. ve Shepperd, M.: A Systematic Review of Software Development Cost Estimation Studies, IEEE transactions on software engineering, Cilt 33, No 1, 33-53 (2007)
12. Dave, VS. ve Dutta, K.: Application of Feed-Forward Neural Network in Estimation of Software Effort, Proceedings on International Symposium on Devices MEMS, Intelligent Systems & Communication, No 5, 5-9 (2011)